PHONETICS LABORATORY
DEPARTMENT OF GENERAL LINGUISTICS
LUND UNIVERSITY

WORKING PAPERS

11 · 1975

MONA LINDAU	Vowel features	1
	A phonetic explanation to reduced vowel harmony systems	43
SIDNEY WOOD	The weakness of the tongue-arching model of vowel articulation	55
	Tense and lax vowels - degree of constriction or pharyngeal volume?	109

Previous numbers and contents:

WP 1 · 1969	K. HADDING M. HIRANO T. SMITH	Electromyographic study of lip activity in Swedish CV:C and CVC: syllables
	KURT JOHANSSON	Försök avseende vokaltransitions riktning och dess betydelse för "plats"-distinktionen bland tonande klusiler
WP 2 · 1970	MONA LINDAU	Prosodic problems in a generative phonology of Swedish
WP 3 · 1970	GÖSTA BRUCE	Diphthongization in the Malmö dialect
	EVA GÅRDING	Word tones and larynx muscles
	KURT JOHANSSON	Perceptual experiments with Swedish disyllabic accent-1 and accent-2 words
WP 4 · 1971	ROBERT BANNERT	Hat das Deutsche zugrundeliegende stimmhafte Spiranten?
	EVA GÅRDING	Laryngeal boundary signals
	KARIN KITZING	Contrastive acoustic analysis of vowel phonemes, pronounced by some North German and South Swedish high school pupils (A summary)
	SIDNEY WOOD	A spectrographic study of allophonic variation and vowel reduction in West Greenlandic Eskimo
WP 5 · 1971	MICHAEL STUDDERT-KENNEDY KERSTIN HADDING	Auditory and linguistic processes in the perception of intonation contours
	ANDERS LÖFQVIST	Some observations on supraglottal air pressure
WP 6 · 1972	ROBERT BANNERT	Zur Stimmhaftigkeit und Quantität in einem bairischen Dialekt
WP 7 · 1973	GÖSTA BRUCE	Tonal accent rules for compound stressed words in the Malmö dialect
	EVA GÅRDING PER LINDBLAD	Constancy and variation in Swedish word accent patterns
	KERSTIN HADDING MICHAEL STUDDERT-KENNEDY	Are you asking me, telling me or talking to yourself?
WP 8 · 1973	EVA GÅRDING	The Scandinavian word accents
WP 9 · 1973	PHONETIC SYMPOSIUM	Postgraduate students Stockholm · Lund
	SIDNEY WOOD	Speech tempo

WP 10 · 1975	ROBERT BANNERT ANNE-CHRISTINE BREDVAD-JENSEN	Temporal organization of Swedish tonal accents: The effect of vowel duration
	GÖSTA BRUCE	Stockholm accents in focus
	EVA GÅRDING OSAMU FUJIMURA HAJIME HIROSE ZYUN'ICI SIMADA	Laryngeal control of Swedish word accents
	KURT JOHANSSON	Perceptual characteristics of vowels
	ANDERS LÖFQVIST	Some phonetic correlates of emphatic stress in Swedish
	BERTIL MALMBERG	Niveaux, choix et systemes approximatifs dans le langage
	KERSTIN NAUCLER	Some thoughts on reading and writing
	THORE PETTERSSON	In favour of the archiphoneme
	EVA WIGFORSS	Foreign accent and bilingualism
	SIDNEY WOOD	What is the difference between English and Swedish dental stops?

Correspondence concerning this series should be addressed to:

Robert Bannert
Phonetics Laboratory
Lund University
Kävlingevägen 20
S-222 40 LUND
Sweden

PHONETICS LABORATORY
DEPARTMENT OF GENERAL LINGUISTICS
LUND UNIVERSITY

WORKING PAPERS

11 · 1975

MONA LINDAU	Vowel features	1
	A phonetic explanation to reduced vowel harmony systems	43
SIDNEY WOOD	The weakness of the tongue-arching model of vowel articulation	55
	Tense and lax vowels - degree of constriction or pharyngeal volume?	109

VOWEL FEATURES

Mona Lindau*

The two most obvious functions of features are their classificatory function and phonetic function. Features classify the distinctive sounds of a language by specifying the contrasts between them. The phonetic quality of a sound is specified by assigning an intrinsic phonetic quality to the features. Another use of features in generative phonology is to make possible the definition of a natural class. A fourth function is the specification of sound patterns and sound changes in such a way that expected natural patterns and changes are formally distinguished from "unnatural" ones. A sound change, or alternation, is natural when some physical reason can be found as its underlying cause, as opposed to such sound changes that happen for reasons that have not apparent relation to the sounds. A phonological process described in terms of features should make its degree of naturalness explicit as a function of the formalism. Ideally, a description of a phonological process in terms of features should permit an explanation as well. The very least we expect from a formalised description is that it provide an accurate statement of the process involved. In order to accomplish this, features must be related to the correct physical parameters that control the speech mechanism. If one significant difference between the American English vowel in bird [bɚd] and the vowel in cut [kət] is in the lowering of the third and fourth formants in bird, as opposed to no such lowering in cut, then the distinctive feature may best be labelled [Lowered frequency of the third and fourth formants] or [Lowered F_3, F_4]. As usual, the first second and third formants are written F1, F2, F3, respectively. The frequency of the first, second, and third

---

* I would like to thank Kay Williamson and Peter Ladefoged for their contributions to earlier versions of this paper.

formants are written F_1, F_2, and F_3, respectively. The feature could of course be labelled [Rhotacized] after the perceptual effect of the lowered third and fourth formants, but then we must also include a convention [Rhotacized] --→ [Lowered F_3, F_4] to apply in all environments for the correct phonetic specification. In any case it would be inappropriate to label the distinctive feature with the articulatory term [Retroflex] as that is not even factually correct. The labelling refers to formal specification. Informally we may prefer to refer to the more familiar labels of "retroflex". I will for example continue to use "height" and "back" in the informal discussion.

The search for "true" correlates of features over the years has demonstrated that it is not possible to relate all features to acoustic parameters, as was attempted by Jakobson, Fant, and Halle (1951), nor to exclusively articulatory parameters as was done by the International Phonetic Association (1949), or by Chomsky and Halle (1968). Some features may best be described as articulatory scales, others as acoustic or perceptual, some perhaps as combinations. Moreover, it is suggested that if variations occur as points along one continuous parameter it is more explanatory to describe that variation as a change of values along a single multivalued feature rather than in terms of switching between binary features (Ladefoged 1971). Our primary goal as phonologists-phoneticians is to come up with an accurate description and an explanation of phonological processes.

This chapter is an attempt to provide a first approximation to a set of features that are required to specify contrasts, and phonological processes that involve vowels. The proposed set of vowel features is exhaustive as far as I know. I have attempted to relate each feature to its physical correlate, and to specify the number of phonological values necessary for each feature. The problem of how to deal with cross-language comparisons of the values of a multivalued feature at the lexical level has not been sorted out at this stage. This problem will for example occur when one wants to compare a language with two values of feature to a language with four values of feature, where the lowest vowel in both languages functions in the same way as in rules. As I have concentrated on classificatory features of vowels, problems with features involved in interactions between consonants and vowels have not been considered here.

The basic vowel parameters

The most basic vowel parameter is vowel height. There is no language that does not contrast vowels along a vertical scale. Another basic contrast occurs along a horizontal scale. There are very few languages that do not contrast front and back vowels. Vowel height and backness then form a basic two dimensional vowel space that is required for almost all languages of the world. Additional contrasts, like lip rounding, pharyngal size, nasality can be considered to be superimposed on this basic vowel space.

"Height"

What is the physical correlate of vowel height? There is abundant evidence against the traditional concept of vowel height as the height of the highest point of the tongue. Using X-ray data from Ngwe vowels and cardinal vowels, Ladefoged (1964, 1975) demonstrated that particularly the tongue height of back vowels bears very little relation to vowel height. Figure 1 is a plot of the highest points of the tongue of the cardinal vowels. The tongue height is approximately the same for [o] and [ɔ]. In addition the distance between the tongue heights of [i] and [a] is considerably smaller than that between [u] and [ɑ], which is contrary to how the vowels are heard (Ladefoged, 1967).

X-ray data from vowel production of one speaker each of Akan, Dho Luo, Ateso, and German were analysed by Lindau

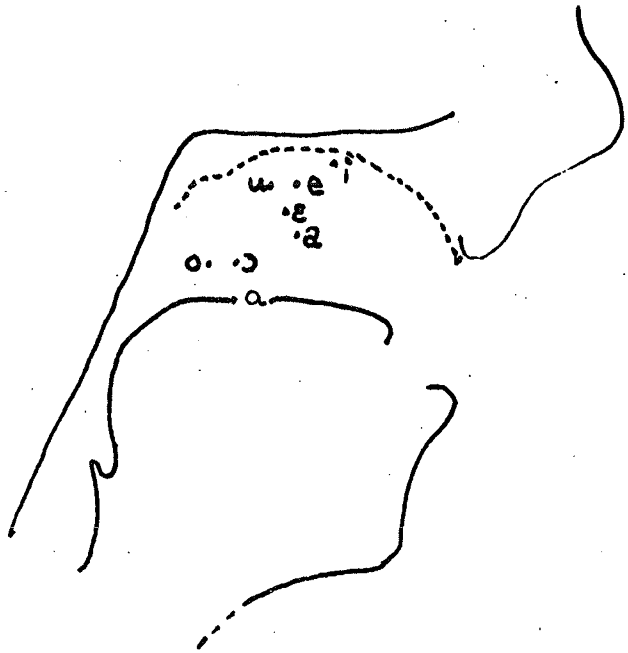


Figure 1. The highest points of the tongue as shown in a published set of x-rays of cardinal vowels. The outline of the upper surface of the vocal tract is not clear on the x-rays, and it is estimated. (From Ladefoged 1975:198.)

et al. (1972). The vowels were [i ɩ e ɛ u ɷ o ɔ]. Figure 2 is a plot of the relative tongue height of these eight vowels in the four languages. The Ateso speaker is the only case where tongue height is related to vowel height; the speakers of the other languages do not use tongue height to produce different vowel heights. Thus tongue height cannot be the primary underlying mechanism of variation in vowel "height".

Lindblom and Sundberg (1969, 1971) proposed relative jaw opening as the main difference between high, mid and low vowels. If this were correct, then the tongue-shapes ought to stay the same within the jaw, and the jaw opening vary with vowel height (provided of course front and back vowels are regarded separately). Lindblom and Sundberg showed that for their single Swedish subject the tongue shapes did remain constant with respect to the jaw.

Ladefoged et al. (1972) studied vowel productions of six American English speakers by use of cineradiography. Figure 3 is from this study. It shows the front lax vowels /ɩ ɛ æ/ - as in bit, bet, bat - superimposed onto a fixed jaw for each of the six subjects so as to show only the movement of the tongue (if any) with respect to the jaw. It is clear that even when we confine the discussion to /ɩ ɛ æ/, we find that only subject 2 behaves as predicted by Lindblom and Sundberg. Subject 1 has similar tongue shapes for /ɛ/ and /æ/ and uses jaw opening to distinguish between two out of three vowels. None of the others have similar tongue shapes in any of the three vowels. They cannot then be using primarily different degrees of jaw opening to control vowel height.

Figure 4 is a plot of relative jaw opening in the eight vowels in Akan, Dho Luo, Ateso, and German. Jaw opening in Dho Luo, at least in this speaker, shows a good ordering relationship to vowel height but the distances between the vowel points do not correspond very well to how they are heard. The vowel points of the other languages show a

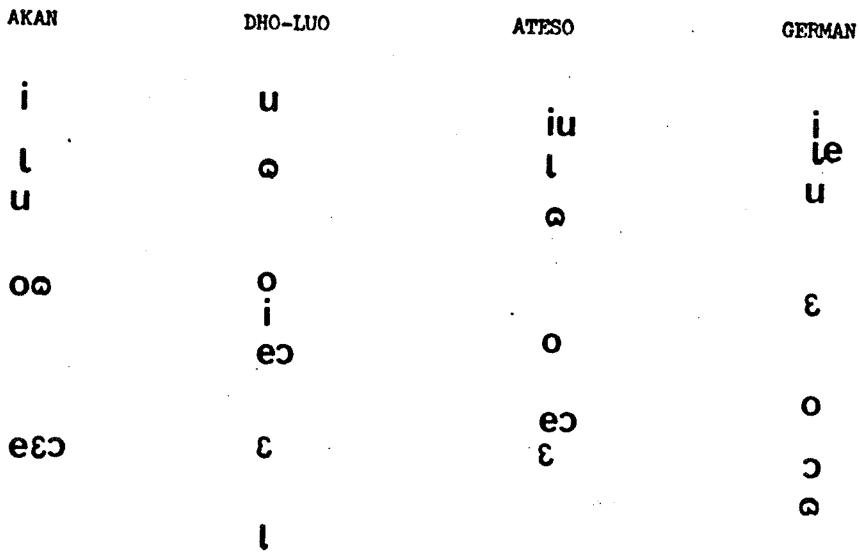


Figure 2. Vertical positions of the highest points of the tongue in eight vowels of the speakers of Akan, Dho-Luo, Ateso, and German.

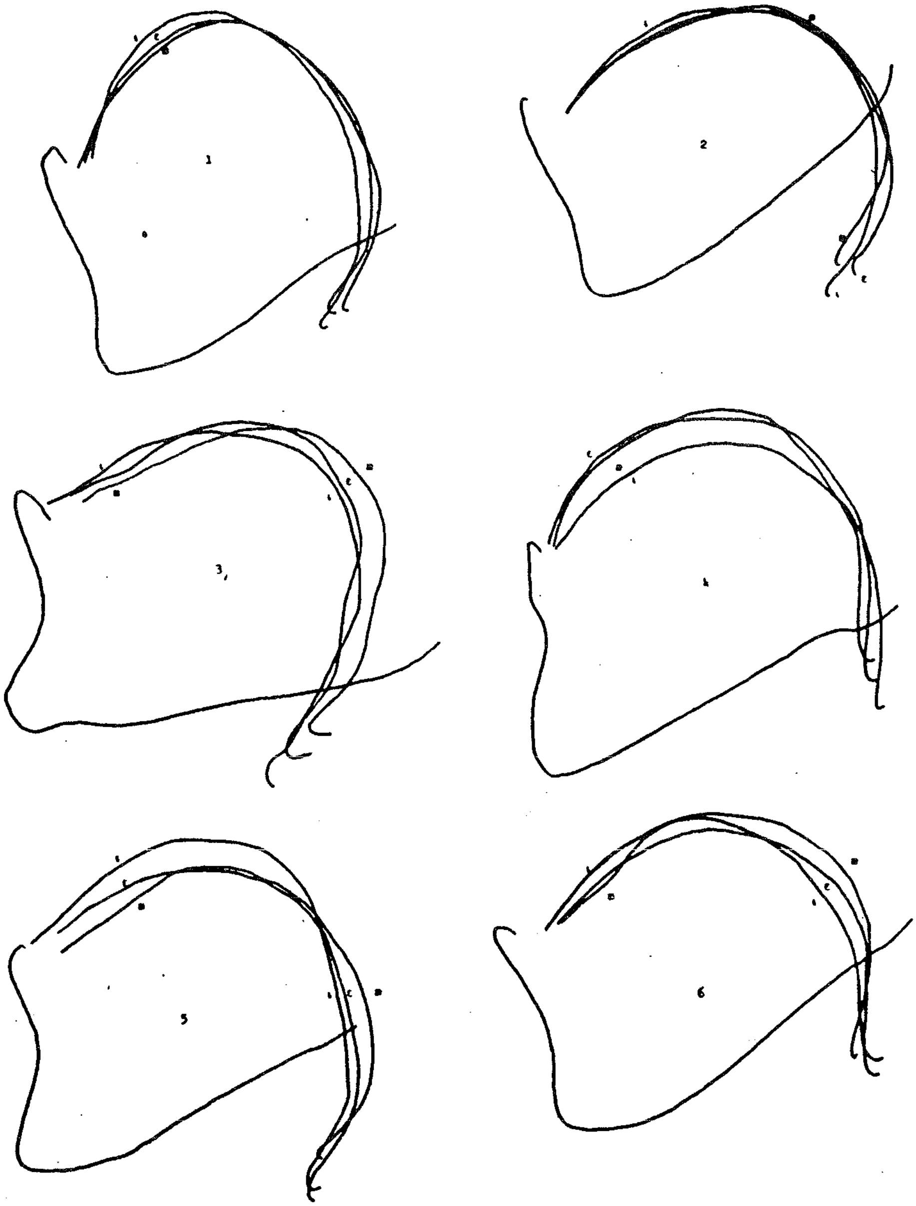


Figure 3. The lax vowels /ɪ ɛ æ / in English superimposed onto a fixed mandible for each of six subjects so as to show only the movement of the tongue (if any) with respect to the mandible. (Ladefoged, DeClerk, Lindau, and Papçun 1972.)

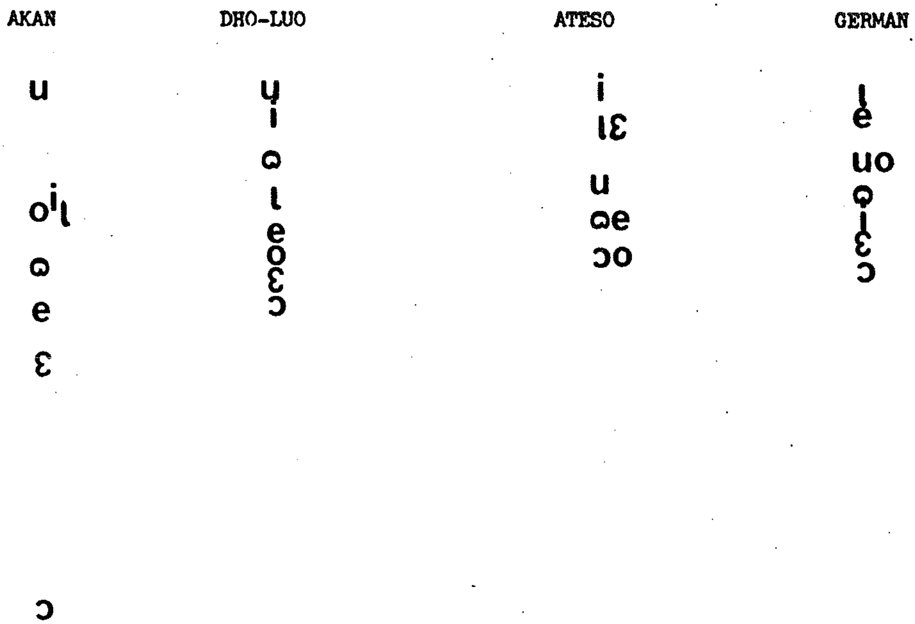


Figure 4. Vertical positions of the jaw opening in eight vowels of the speakers of Akan, Dho-Luo, Ateso, and German.

better relation between jaw opening and vowel height than was the case when tongue height and vowel height were compared, but the relationship is not good enough to support Lindblom and Sundberg's claim about the jaw opening as the universal phonetic correlate of vowel height. The use of jaw opening to distinguish between high, mid, and low vowels by some speakers only shows that this is one possible way of achieving vowel height. It does not justify postulating jaw opening as a necessary correlate of vowel height.

In summary, all available evidence points to the fact that a speaker has several possible gestures available for producing a certain point in the basic vowel space, and that different speakers also do make use of all available mechanisms to achieve the same acoustic result. The invariance in vowel height is not of any articulatory kind but rather acoustic. Formant frequencies plotted on a formant chart usually show a much better relation to how the vowels are perceived (Ladefoged 1964, 1971, 1975). The cardinal vowels as spoken by Daniel Jones were plotted on the formant chart in Figure 5. The formant frequencies were inferred from a formant chart in Lindblom and Sundberg (1969), and plotted on a formant chart with F_1 against the difference between F_2 and F_1. The resulting figure is much closer to the traditional quadrilateral than the figure described by the highest point of the tongue (Figure 1). Vowel height is related in a straight-forward way to the frequency of the first formant (F_1). High vowels have relatively low F_1, and low vowels have relatively high F_1. Articulatorily based features like Tongue Height, Jaw opening, Stricture (Williamson 1974)[1] are less appropriate for vowels. With the correlate of vowel height being F_1 the most appropriate features label of vowel height is of course [F_1].

The feature [F_1] is multivalued because vowels may contrast more than two values along this single scale. Phonological processes involving this feature shifts the vowels up and

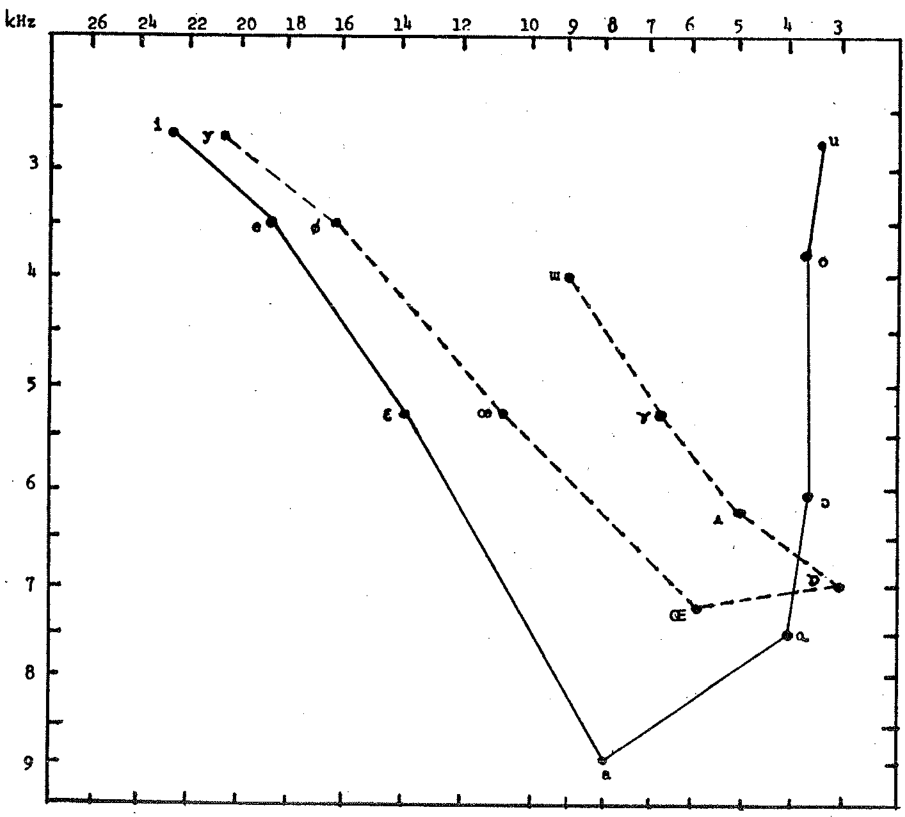


Figure 5. Formant chart with the frequency of the first formant on the vertical axis and the distance between the frequencies of the first and second formants on the horizontal axis for the cardinal vowels. The formant frequencies are from Lindblom and Sundberg (1969).

down one scale. The use of binary features to express movements along one physical scale would make a wrong claim about the relationship between the vowels. There are values of $[F_1]$ that simply cannot be expressed correctly with two binary features. A Swedish dialect, Scanian, (as spoken in Malmö) diphthongizes long vowels as below (Bruce 1970):

/i:/ → [ei]	/y:/ → [øy]	/u:/ → [eu]
/e:/ → [εe]	/ʉ:/ → [øu]	/o:/ → [εo]
/ε:/ → [æε]	/ø:/ → [œø]	/ɑ:/ → [æɑ]

A vowel insertion rule must specify a vowel one step lower than the underlying vowel. As four heights are involved, Chomsky and Halle's [High] and [Low] cannot be used but we can try Wang's [High] and [Mid] with the use of paired variables (Wang 1968).

$$\emptyset \rightarrow \begin{bmatrix} V \\ \alpha \text{high} \\ \beta \text{mid} \end{bmatrix} \Big/ \text{---} \begin{bmatrix} \beta \text{ high} \\ -\alpha \text{mid} \end{bmatrix}$$

(The rounding variation has been ignored, since it is not pertinent to the point). This rule generates the desired output, [ei], [εe], [æε], etc. but because of the switching nature of the rule it also generates a fourth type of diphthong [æ i], [œ y], and [æ u] that is not only not desired but makes the wrong claim that this would be the most likely extension of diphthongization in Scanian. The only way to avoid it is by the use of n-ary values. This also makes the rule formally simpler.

$$\emptyset \rightarrow \begin{bmatrix} V \\ n + 1 \; F_1 \end{bmatrix} \Big/ \text{---} \begin{bmatrix} n \; F_1 \end{bmatrix}$$

How many values are needed for F_1 ? Some languages contrast only two values of vowel height, e.g. Kabardian (Halle 1970) with the vowel system

ɵ
a

or Turkish with a system of eight vowels on two heights:

i y ɯu
e ø ɑo

Sedlak (1969) lists some twenty languages with two vowel heights. The maximum number of values for this vowel feature seems to be four. Ladefoged (1971) reports Danish and English, and Hockett (1955) two Polish dialects with four heights. Dan has a system with at least four central vowels:

i	ɨ	u
e	ɘ	o
	ɵ	
(æ)	a	(ɒ)

(/æ/ and /ɒ/ are included by Bearth and Zemp (1967) but not by Welmers (1973).[2]

Five vowel heights have been reported for Ngemba by Eastlack (1968);

i	ɨ	u
I		
e	ë	o
	ə	
	a	

The vowel /I/ could easily be distinguished from the others by some other feature than height. Moreover, the maximal contrast at any value of Backness is still only four heights. So Ngemba has at least no more than four contrastive heights. Even so, this analysis makes the system look suspiciously inefficient with respect to the use of available acoustic space.

For contrastive purposes we thus need four values of [F_1]. If the glides /j ʋ w ɥ/ are regarded as end points of the vowel height continuum, they can be included at one end of this scale as [0 F_1].

"Back"

The second basic vowel dimension places vowels as points along a horizontal scale, usually called Backness. Backness has traditionally been regarded as an articulatory dimension. While it is true that the tongue is further back in back vowels than in front vowels, there is, however, not a good correspondence between the highest points of the tongue on the horizontal dimension and the way in which corresponding vowels are located on a vowel chart. Compare the positions of [o] and [ɔ] in Figure 1. Again, we look to acoustic dimensions for a better correlate of Backness. The obvious candidate is the frequency of the second formant, F_2. F_2 is relatively low for back vowels, relatively high for front vowels, and in between for central vowels. When F_2 is plotted against F_1 on the ordinary type of formant chart the resulting figure forms a traditional vowel triangle. Acoustically and perceptually, however, back vowels are usually not on a slope like the right hand side of a triangle, but distributed more on a straight vertical line. The acoustic, and probably the perceptual vowel space, is in fact more like the Jonesian quadrilateral than a triangle. If we plot F_1 against the difference between F_2 and F_1, instead of against F_2, a quadrilateral vowel figure is obtained. The slope of the front vowels also improves in relation to the auditory chart. Backness is thus better related to the difference between F_2 and F_1 than simply to F_2, and the feature will be labelled [F_2-F_1].

Some real evidence for [F_2-F_1] comes from studies of acoustic and perceptual vowel spaces using a type of factor analysis, PARAFAC, which "incorporates, within the factor model, certain basic tests for determining the explanatory factors" (Harshman 1971:14). This procedure of factor analysis provides a unique, "true" solution for a set of adequate data. Three factors were extracted from a data set of formant frequencies of Swedish vowels of several speakers. The vowels along the factor corresponding to the "back" dimension were distributed in such a way that they are much better related to $F_2 - F_1$ than to F_2 (Lindau et al. 1971).

There are languages that do not contrast vowels along the horizontal dimension. When there thus is only one value of $[F_2 - F_1]$ that value refers to central vowels. These systems occur in some Caucasian languages, e.g. Kabardian. Hockett (1955) mentions Adyge, possibly Abkhaz, and Udykh with a system of

ɨ

ə

a

Mohrlang (1971) analyses Higi as a system of three central phonemes.[3] The occurrence of such vowel systems constitutes a violation of Sedlak's proposed universal no. 4:

"All languages have a high or lower high front vowel." Of course, both Kabardian and Abkhaz have extremely rich inventories of <u>phonetic</u> vowels that are derived from assimilations to features of surrounding consonants - including [i]s - but I presume Sedlak refers to vowel systems on the phonological level. These facts further imply the nonexistence of any universal to the effect that there is at least one particular vowel that occurs in all languages of the world. There simply is not. Another universal suggests itself, and I propose it here:

"If a language has no horizontal contrast, all the vowels will be central."

I do not know of any language with only back or only front vowels.

The majority of languages contrasts two horizontal values. In the vowel systems I have looked at, these two values equal front and back, i.e. the maximum and minimum values of the feature [F_2 - F_1]. I propose a second universal of this feature to complement the first one.

"If a language has horizontal contrasts, then it has front and back vowels."

The feature [Back] in the SPE system has a maximum of two values. This excludes the possibility of specifying central vowels on the systematic phonemic level. Consequently, in languages with three heights, as in the very common seven vowel system of /i e ɛ a ɔ o u/, the vowel /a/ is forced into a [+Back] classification, and it is distinguished from /ɔ/ by the feature [Round] . This implies a very curious claim that the third vowel height somehow "causes" /a/ to be [+Back] , when really the way in which /a/ functions as front, central, or back in different languages does not have any obvious relation to the number of heights or rounding there are. Moreover, many languages have other central vowels that function as vowels between front and back vowels, and not as unrounded back ones.

There are also languages that contrast three horizontal values with the same value of rounding. Norwegian has four high vowels, out of which three are rounded (Vanvik 1972), namely /i y ʉ u/. The vowels /ʉ/ and /u/ could conceivably be derived from underlying /u/ and /o/ respectively, but I do not consider a neater system and a reduplication of historical process justification enough for this in present day Norwegian, where the alternation patterns do not support this "solution". There is no alternation [ʉ]~[u] nor [u]~[o]. Norwegian contrasts three rounded horizontal values. Another language with three horizontal contrasts is Brôu (Miller 1967). This language has 41 vowels, including short and long vowels, and diphthongs, It seems the system can be reduced to 17 long and short vowels, or to the following ten or eleven basic vowels.

i	ɨ	u
e	ə	(ʌ)o
	a	ɑ ɔ

I have retranscribed Miller's transcription into that of IPA for easier reference. My symbols are chosen from studying Miller's detailed phonetic descriptions and acoustic charts.

On the acoustic charts /a/ is clearly central, right between front and back vowels. The system is symmetric with /ɨ ə a/ as central vowels. From the literature I do not know of any strong evidence that /a/ behaves as a phonological front vowel. There is thus no reason to postulate /a/ as front and low rather than central. Thus Brôu contrasts four low vowels, three of which are unrounded; so also here three values of $[F_2 - F_1]$ are essential.

As three contrasts constitute the maximum number of horizontal contrasts, another universal suggests itself:

"No language contrasts more than three horizontal values."

<u>Features of the lips</u>

The feature $[F_2 - F_1]$ is not quite independent. A constriction at the front of the vocal tract results in a larger distance between F_2 and F_1 than a constriction in the middle (where back vowels are). When we add variation at the ends of the vocal tract this affects F_3 and F_2, and thus also the distance between F_2 and F_1. A decrease of the size of either end of the vocal tract will lower F_3 and F_2. Thus the relatively small difference $F_2 - F_1$ that results from a constriction in the middle of the vocal tract is made even smaller by decreasing the mouth opening. Front vowels will have a larger distance between F_2 and F_1 if pronounced with spread lips (and wide low pharynx). The maximal horizontal distance is obtained by maximising the mouth opening for front vowels and decreasing it for back vowels - which is why front vowels are basically unrounded, and back vowels basically rounded. Variation of the size of the mouth opening may be used to create more vowels. Decreasing the lip opening for front vowels, and increasing it for back will add sets of vowels inside the "basic" maximal vowel space.

The lip opening can be decreased in two days: by protruding the lips or by compressing them by vertical forces so that the lip opening becomes a narrow slit. These two possibilities have been recognized since Sweet (1877). Both mechanisms involve lip action, or labiality, but only the first type is protruded. Labial consonants are produced by lip compression, and protrusion may be superimposed. Protrusion implies labiality, but not vice versa. Many phonological rules also apply to rounded vowels and labial consonants, so a feature is needed to cover both types of lip action - [Labial]. Protrusion is as usual specified with Round . Both lip features have invariant articulatory correlates, and complex acoustic ones.

Round

The feature [Round] may serve to contrast two types of front vowels and two types of back vowels. I have not come across any language with a rounding contrast for central vowels. [Round] is a binary feature. Phonetic degrees of lip protrusion are predictable from the value of F_1 (vowel height).

Systems with a single front rounded vowel are rare. Chacobo, Basque, Mandarin Chinese are reported by Sedlak. Two front rounded vowels occur in a substantial number of languages e.g. German, Icelandic, Faroese, Norwegian, Swedish, French, Albanian, Turkish, Hungarian, Estonian, Tibetan, Akha. No language has more than three contrastive front rounded vowels. Systems with three rounded vowels are not very common. Sedlak lists Icelandic with three front rounded vowels. But most analyses come up with one or two front vowels (Einarsson 1928, Haugen 1958, Benediktsson 1959). They occur in those versions of French that distinguish for example jeûne [ʒø:n] 'fast' and jeune [ʒœn] 'young'.

Systems with one back unrounded vowel occur in Chinese.

iy   ɯu

ə

a

Two back unrounded vowels occur in Turkish for example:

iy   ɯu

eø   ɑo

Akha (Lewis 1968)[4]

i y   ɯ u

e ø   ɤ o

ɛ   a   ɔ

As for front rounded vowels, the maximum number of back unrounded vowels is three, as in Vietnamese:

i   ɯ u

e   ɤ o

æ   ʌ ɔ

a

or in Fe' Fe' (Hyman 1972):

i   ɯ   u

e   ɤ   o

a   ɑ

The above languages also demonstrate that front rounded and back unrounded vowels may co-occur in a system.

Central vowels are mostly unrounded. Rounded ones occur in for example Norwegian (p. 15). There is no language that contrasts rounded and unrounded central vowels at the same height. In languages with a single central unrounded vowel, that vowel is usually /a/. Sedlak lists a number of languages with two central unrounded vowels. Three central unrounded vowels are not very common but occur for example in Brôu (p. 16), Ngwe, and Kashmiri:

Ngwe (Dunstan 1966):

i	ɨ	u
e	ə	o
ɛ		ɔ
	a	

Kashmiri (Kelkar 1964):

i	ɨ	u
e	ə	o (+/:/)
	a	

Four central unrounded vowels occur in Dan (p. 12).

There is a problem with assessing systems with reported central or back unrounded vowels. Linguists do not consistently use the same symbols for these vowel classes. As it turns out it may be a pseudoproblem: these two vowel classes never contrast for non-low vowels. The low /a/ and /ɑ/ may contrast as in Brôu, though this is very rare.

The non-contrastiveness of unrounded high central and unrounded high back vowels seems to have an acoustic reason. Apparently it has to do with non-linear relationships between articulation and acoustic effects. Consider Figure 5. Rounding non-low front vowels lowers F_2 some 200 Hz, while unrounding back vowels has a much larger effect on F_2, which increases by about 700 Hz. This relatively large increase of F_2 will place the "back" unrounded vowels acuostically very close to a central position. Vowels in this acoustic area are notoriously unstable. This is the most difficult area for a speaker in which to produce constant and stable vowel qualities, and for a listener to distinguish between vowel qualities. The instability of the central unrounded and back unrounded vowels is predictable from Fant's Maxima Theory (Fant 1960, Gunnilstam 1973). Vowels are more stable at those areas in the vocal tract where a constriction produces formant curves (as on logograms) where two formant curves have their respective maximum and minimum simultaneously. At these places a small articulatory movement causes no acoustic change. But where formant curves have a steep slope, a small articulatory change will have large acoustic effects. A study of Fant's logogram of the effect on formants as a function of the place of constriction with various degree of lip rounding shows that at 10 cm from the glottis (approximately [u]) unrounding will cause a considerable upward slope of F_2 (Fant 1960, p. 82). Very small articulatory displacements in the back to central area will cause relatively large shifts of F_2 as long as the lips are not rounded.

Labial

Vertical lip compression is a much less usual way of decreasing the lip opening for vowels than lip protrusion. In fact, the only language I am aware where this occurs, is Swedish. Swedish contrasts lip protrusion, [Round], and lip compression, [Labial], for high vowels:

i	y ʉ	u
e	ø	o
		ɑ

/y/ and /ʉ/ are both non-back with decreased lip opening. The vowel /y/ is produced with lip protrusion, /ʉ/ with the gesture for [Labial]. A second reason for classifying /ʉ/ as [Labial] is in the nature of its offglide. In Swedish long high vowels have an approximant offglide at the same place of articulation as the vowel. The offglide after [ʉ:] is a labial [β]. The others are [ij, yɥ, uw].

Urhobo approximants supply another example of a Round-Labial contrast. Urhobo has a round /w/ and a labial /ʋ/. Before high back vowels both are also velar. Before rounded vowels both /w/ and /ʋ/ are influenced by the rounding -

but not in the same way. In producing /w/ in /ũ̃wũro/ 'bend in the knee' and /owɔ/ 'leg' my informant's lips are quite strongly protruded, but while producing /ʋ/ in /ɔ̃dũ:ʋũ/ 'a kind of animal trap' and /ʋurɛ/ 'sever' the lip opening is decreased but not by protrusion (cf. Kelly 1966).

Both in Swedish and Urhobo the vowels and approximants differ by the use of two separate lip gestures, not by different degrees of the same gesture, so they should be characterized by separate features.

Expanded

In many Niger-Congo languages of West Africa and in Nilo-Saharan languages of East Africa vowels may be distinguished by a mechanism involving the size of the pharynx, as controlled by variation in the positions of the root of the tongue and the larynx (Ladefoged 1964; Pike 1967; Stewart 1967; Lindau et al. 1972; Antell et al. 1974; Lindau 1975). This mechanism consistently underlies one phonological process only: vowel harmony. On the basis of evidence from the same speaker Halle and Stevens (1969) and Perkell (1971) suggest that the root of the tongue distinguishes the "tense" and "lax" vowels in English in the same way as harmonizing sets are distinguished in the African languages. But it is quite clear that, when more speakers are considered, not all speakers of English separate "tense" and "lax" vowels using the tongue-root (Ladefoged et al. 1972). In the African languages the size of the pharynx separates two harmonizing sets of vowels. The maximal system is 5 + 5 vowels: five vowels /i e ɜ o u/ with a large pharynx and five vowels /ɩ ɛ a ɔ ɷ/ with a small pharynx. The ten-vowel systems are relatively rare. They have been reported for some Kwa languages, namely Sele (Allen 1974), Abe (Stewart 1971), Igede (Bergman 1971), and Engenni (Thomas 1969), for some Benue-Congo languages, namely Ọgbia (Williamson 1972), Abuan (Wolff 1969), and Kohumono (Cook 1969), and for some Gur languages: Kasem, Sisala, Mianka (Bendor-Samuel 1971). Among Nilo-Saharan languages ten vowel systems are found in Kalenjin, Päkot, Acholi, Lotuko (Antell et al. 1974). Nine vowel systems where /ɜ/ has merged with some other low vowel are fairly common. They occur for example in Akan languages, Delta Ịjọ, and some Central Delta languages. The vowel /a/ tends to be neutral to vowel harmony and the 4 + 4 + /a/ system patterns like below:

1.		2.	
i	u	ɩ	ʊ
e	o	ɛ	ɔ

a

Many languages have reduced the nine vowel system to a partially harmonizing seven vowel system. By the time the system has reduced to a five vowel system the vowel harmony will be lost (Williamson 1974).

Over the years many features have been proposed for African vowel harmony: Tense, Raised Height, Breathy, Covered - just to mention a few. There is now substantial evidence that the main phonetic control of the vowel harmony is the movement of the tongue root (Lindau et al. 1972; Retard 1973; Painter 1973). The tongue root mechanism is mostly - but not always - combined with vertical larynx displacements, and sometimes with movements ofthe back pharyngal wall. It thus seems that what a speaker tries to accomplish is variation of the pharyngal size. As illustrated in Figure 6 the Akan speaker produces the set 1 vowels /i e/ with a relatively large pharynx by advancing the root of the tongue beyond a "normal" position for that vowel, and by lowering the larynx. The relatively small pharynx of the set 2 vowels /ɩ ɛ/ is produced by retracting the root of the tongue beyond its "normal" position, and by a relatively high larynx.

Figure 7 functions as a summary statement of the formant space in Akan. A comparison of Figure 6 and Figure 7 will give some idea of articulatory-acoustic relationships.

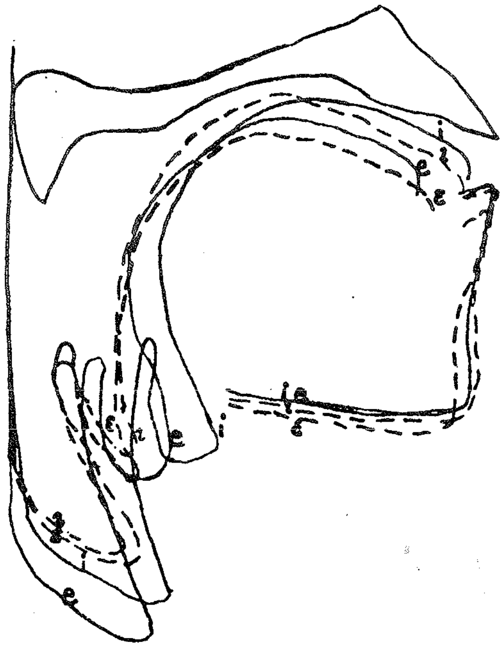


Figure 6. Selected tracings of /i   e   / of one speaker of Akan.

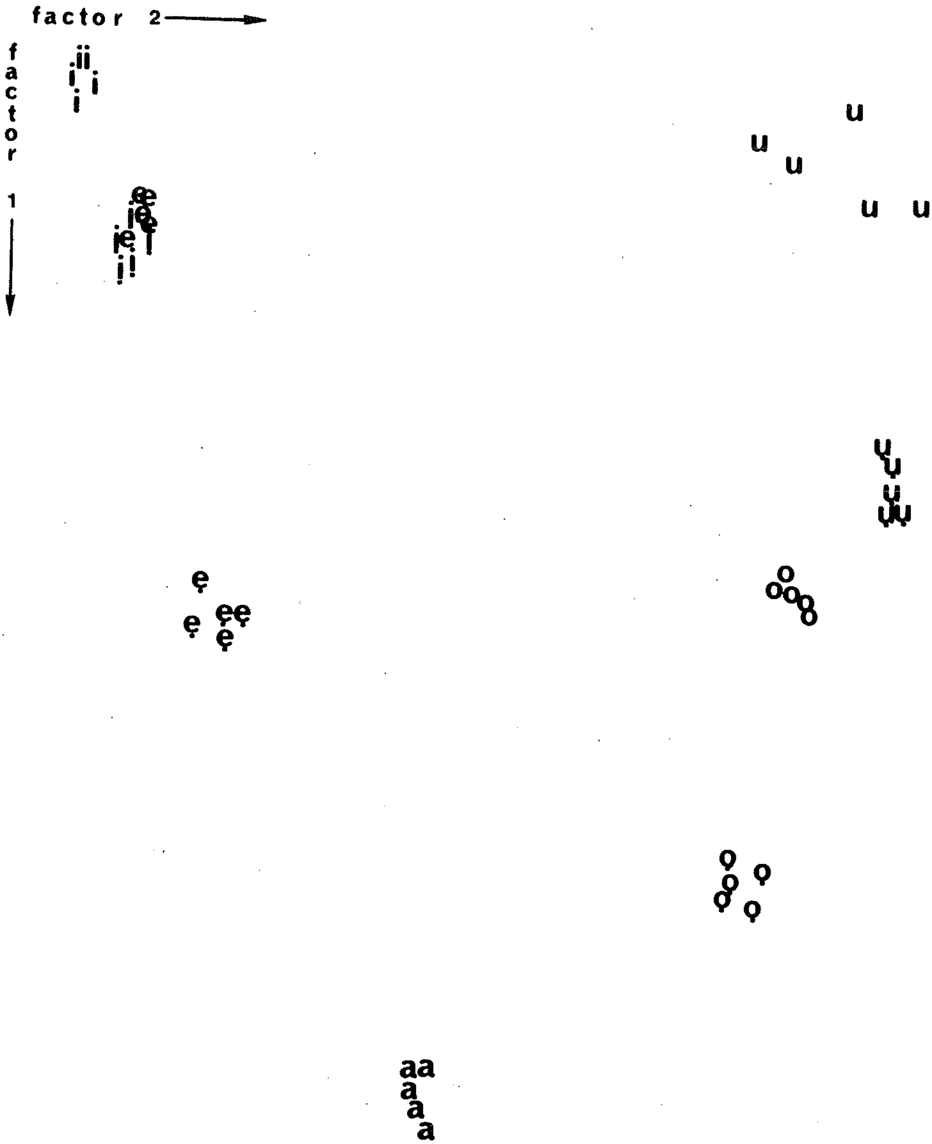


Figure 7. Two factor solutions of factor analysis of two formant frequencies of five tokens each of nine vowels of four speakers, using the Parafac-procedure (Harshman 1970). Language: Akan. The dotted vowels represent set 2 vowels. Factor 1 = F_1, factor 2 = inverse of F_2, correlation = 993, mean square error = 1016.7 (or 31.5 mel). The factor solution is here used as a normalization procedure for formant frequencies.

Varying the size of the pharynx, as between /i/ and /ɩ/, and /e/ and /ɛ/ affects F_1: that is, it has the same general acoustic effect as varying the size of a constriction in the front of the mouth. Decreasing the pharynx size (by retracting the tongue root) increases F_1 in the same way as apening up a constriction of the mouth does (by lowering the body of the tongue). That is, going from [i] to [e] by mainly increasing the size of the mouth constriction, and going from [i] to [ɩ] by decreasing the pharyngal size will have very much the same acoustic effect. For an attempt to explain this, see Lindau (1973). This is clearly shown by the acoustic merging of /ɩ/ and /e/ in Figure 7.

While there is more than one articulatory way of varying vowel height ([i] → [e]), the difference between [i] and [ɩ],and between the other harmonizing pairs has a consistent articulatory correlate. This is not just variation of the tongue root. The larynx and the back pharyngal wall are also involved. What is consistent is the variation of pharyngal size. So the corresponding feature will not be labelled [Advanced Tongue Root] but [Expanded], referring to pharyngal expansion.

When there is no contrast, the tongue root is not especially advanced or retracted. This state is regarded as a zero value of the feature [Expanded]. In the African languages the contrast is achieved by deviating in opposite directions from that zero value. So the feature values are:

[1 Expanded]	=	Wide pharynx
[0 Expanded]	=	Neutral pharynx
[-1 Expanded]	=	Narrow pharynx

It is conceivable that the same mechanism is involved in distinguishing between emphatic and non-emphatic consonants. If that is so, both [0 Expanded] and [-1 Expanded] occur in languages with pharyngalized consonants, like Arabic. It is evident from the cineradiographic data presented by Ali and Daniloff (1970) that vowels in the environment of pharyngealized consonants are all produced with a retracted tongue root, very similar to that in the vowel harmony languages. Some speakers of English seem to produce the so called tense vowels with an advanced tongue root and the lax vowels with a neutral tongue root, so here the difference is between [1 Expanded] and [0 Expanded]. But it is obvious from our data in Ladefoged et al. (1972) that speakers are not consistent in distinguishing tense and lax in this way, so the feature [Expanded] cannot be used to distinguish English vowels.

Vowel systems in many Mon Khmer languages are characterized by so called <u>voice registers</u>, where the vowels fall into two sets called First and Second Register. K. Gregersen (1973) summarized a good number of impressionistic phonetic descriptions of the two registers. On the basis of this he proposes that the Mon Khmer registers are really controlled by the same mechanism as vowel harmony in African languages. There are striking similarities in these impressionistic descriptions to the earlier descriptions of the African vowels. No conclusive evidence in terms of x-ray data occurs as yet that I am aware of, but Gregersen's hypothesis sounds very likely. If he is correct, then some Mon Khmer languages contrast [-1 Expanded] and [0 Expanded], others contrast [0 Expanded] and [1 Expanded]. Gregersen points out that one set is "normal" and the other set may deviate in either direction.

"<u>Retroflex</u>"

So called "retroflex" vowels have been reported for Badaga, a Dravidian language. Emeneau (1939) analyses the Badaga vowel system into 30 contrastive vowels:

i u ỉ ủ i̎ u̎

e o ẻ ỏ e̎ o̎

a ả a̎

V̉ = slightly retroflex vowel.

V̎ = strongly retroflex vowel.

Each vowel occurs long and short as well. The Badaga contrasts call for a ternary feature of retroflexion. It is worthwhile to point out that these threeway contrasts have not been noticed elsewhere.

Emeneau described the retroflex vowels as being produced with the tip of the tongue curled upwards and backwards to a smaller or greater extent. Ladefoged (1975) points to the vowel in American English <u>sir</u>, <u>cur</u>, <u>bird</u> and he notes that although these vowels are strongly r-coloured, they are nevertheless not always retroflex. Some speakers produce the r-colouring with the tip of the tongue down. There is also a constriction in the pharynx below the epiglottis.

The acoustic effect of both gestures for r-colouring is a lowered third and fourth formant. It seems that again we have a feature where the invariant physical reality lies in the acoustic domain rather than in the articulatory domain. The articulatory term "retroflex" is therefore inappropriate as label for the feature. Ladefoged labels the [ɚ] vowels with an auditorily based term "rhotacized". I suggest that as we already have acoustic features [F_1] and [$F_2 - F_1$] and this correlate is also acoustic, the most appropriate label is acoustic, [Lowered F_3, F_4]. The three contrastive values 0, 1, and 2 of [Lowered F_3, F_4] refer to plain vowels, slightly "retroflex" vowels and strongly "retroflex" vowels, respectively.

<u>Nasal</u>

Properties and processes involving nasalization in vowels have been discussed extensively by Ferguson (1963), Ladefoged (1971), Ruhlen (1973) among others. Nasalized vowels occur frequently phonetically in the environment of nasal consonants. But many languages show a true contrast between oral and nasalized vowels, e.g. many Kwa languages in West Africa. The feature is [Nasal] with an obvious articulatory correlate: the state of the velum. The acoustic effects of lowering the velum are very complex. They include an increase of F_1 (House and Stevens 1956, Ohala 1971), as well as increases in the bandwidths of the formants. Nasalized vowels will thus sound "lowered" without changing the rest of the vocal tract which is why nasal vowels tend to lower systematically (Ohala 1971, Hombert 1974). The feature [Nasal] is probably binary, although several degrees of nasality occur phonetically.

Long

Long and short vowels occur in many languages. The durational differences are, however, not always interpretable as contrastive length. The domain of a length feature may be the syllable in which case vowel duration is predictable from the syllable structure. This is the case in for example Icelandic, Norwegian, and Swedish. In Swedish closed long syllables may end in V:C, or VCC. In other languages, where long vowels function alike to diphthongs, long vowels may be derived from VV-sequences, as in Finnish (Lehiste 1970). The interpretation of [V:] as /VV/ is also standard in such tone languages as have tonal glides or double tone over a long vowel, as happens in many Niger-Congo languages.

Vowel length is accompanied by qualitative differences in many languages. Problems arise in the interpretation when trying to decide on which is significant. The vowel quality differences manifest themselves in centralization of short vowels. This is the case in German, Swedish, English, Czech, Serbocroatian, where the two sets of vowel qualities are referred to as "tense" in long vowels and "lax" in short vowels. A listening experiment conducted by Hadding and Abramson (1964) showed that in Swedish the durational differences became less important when a vowel pair differed substantially in quality. It thus seems that when vowels differ in both respects, quality differences are a primary cue provided these differences are large enough.

There are undoubtably also languages like Luganda, Estonian, Mixe, where vowels differ solely as to segmental quality, so a feature [Long] must be included in a universal inventory. Probably only two values are contrastive: short and long. Ladefoged (1974) reports four values in Kamba, but some are grammatically conditioned. The question of two or three contrastive lengths in Estonian has been debated for years (Lehiste 1970). Lehiste demonstrates that Estonian has unquestionable three ranges of durational vowel differences - short, long, overlong - but there are alternative interpretations of the overlong vowel. Hoogshagen (1959) reports three vowels lengths in Mixe (Mexico) V, V˙ and V:, interpreting them as /V/, /V˙/ and /V·h/, respectively. More than two lexically contrastive lengths have not been demonstrated unambiguously - yet. Length is therefore a binary feature. Short vowels are [-Long], long vowels are [+Long].

"Tense"

The tense/lax distinction has been extensively discussed since the time of Melville Bell (1867). A feature like Tense is clearly needed in many phonological rules. Whether this feature is truly also needed for contrastive purposes is not that obvious, and what phonetic mechanism controls the feature seems to be a wide open question, judging from the literature. The range of proposed correlates covers most conceivable parameters from "muscular energy" to perceptual "colour" dimensions. For a discussion of the literature the reader is referred to Miller (1974). What is meant by a tense/lax distinction is usually the kind of vowel quality differences that accompany long and short vowels in European languages like English, German, Swedish, Czech and in some languages spoken in India, e.g. Kannada. The long vowels here are perceptually more peripheral and the corresponding short vowels more centralized towards a schwa. In English tense vowels are also diphthongized.

When tenseness could be predicted from length in these languages the feature [Tense] may not be needed on the systematic phonemic level. But because the vowel quality sometimes is the primary one (p. 28), we might want to keep Tense as a contrastive feature for phonetic reasons.

There are also languages where Tense apparently is independent of length. Hindi-Urdu apparently has tense-lax contrastive differences without length differences (Sedlak). So does Friulian, also according to Sedlak. As [Tense] can be independent from [Long] it must be included as a separate feature.

The qualitative difference between Tense and Lax is described as peripheral vs. central. There is no consistent articulatory mechanism corresponding to this (Ladefoged et al. 1972). Perceptual and acoustic relations correspond quite well. On an acoustic chart the lax vowels are inside the tense vowels, on an axis towards a[ə]. Although the feature is better regarded as acoustic rather than articulatory, there is no obvious single acoustic parameter that exactly corresponds to that axis. For laxing, we could use something like "formant frequencies approaching $F_1 = 500$, $F_2 = 1500$, $F_3 = 2500$ Hz". It is worth stressing again here that also from an acoustic point of view Tense is not the same as the feature Expanded. Tenseness is on a central - peripheral acoustic axis, while Expanded is on a vertical (F_1) axis.

The feature of tenseness will be labelled [Peripheral]. It is a binary feature. So called tense vowels are [+Peripheral]. [-Peripheral] vowels are inside their [+Peripheral] counterparts approaching formant frequencies of 500, 1500 and 2500 Hz.

Welmers (1973) reports a remarkable vowel system for Dinka with three phonetic degrees of centralization. But the three degrees of [Peripheral] are also accompanied by differences in length and phonation types, so it seems unlikely that the peripheral - central differences are contrastive. Besides, as some of the central - peripheral vowels in Dinka are controlled by differences in pharyngal size (L. Jacobson, personal communication) it is apparently not the feature [Peripheral] that is involved but the feature [Expanded].

Phonation types

Differences in phonation types among vowels are usually non-contrastive. Voiceless vowels occur in many languages, but always conditioned by surrounding voiceless consonants. Hindi vowels may be somewhat breathy voiced from preceding breathy voiced consonants. There are a few languages where different states of the glottis are contrastive. Ladefoged (1971) reports Gujerati contrasts between voiced and breathy voiced vowels, at least on the systematic phonetic level. Lango contrasts voiced and laryngealized vowels. Ladefoged's feature is [Glottal Stricture] with nine possible categories. Only two of these may contrast for vowels.

\+ + + + +

It remains to mention two features apart from [Peripheral], that do not seem to function to classify sounds into contrastive categories, but that are needed for correct specification of phonological processes. The feature [Grave] is not contrastive independently of other features. Grave vowels are always back, and grave consonants are all classified after their place of articulation. But labial and velar consonants often function together as a class, and interact with back vowels. The common property of grave sounds is an acoustic one: low spectral energy.

As an example of this feature in phonological rules let us take a comparison between British and American English. Both dialects have a vowel /ju:/ and a vowel /u:/, but the British /ju:/ has become /u:/ in some varieties of American English in stressed syllables in the environment after dental and alveolar consonants, but not after labial and velar consonants. Cf. the American pronunciation of pew, spew, beauty, few, view, mute, cute, gules; but enthuse, tune, stew, dune, lute, nude, rude, sue, presume. This historical sound change is best described in terms of the feature Grave . The nongrave /j/ may disappear after a nongrave segment, but not after a grave segment:

$$\begin{bmatrix} -\text{vocalic} \\ -\text{consonantal} \\ -\text{grave} \end{bmatrix} \rightarrow \emptyset \; / \; [-\text{grave}] \; \text{---}$$

Other examples of the use of this feature can be found in Hyman (1972).

The second "rule" feature occurs in Dinka. Welmers (1973) arranges the Dinka vowels in a system like an eight spoke wheel with the top spoke missing:

```
ī              ū
   ï        ü
      i   u
ē ë e        o ö ō
      ɛ   ɔ
        a
   ɛ̈    ä    ɔ̈
ɛ̄       ā       ɔ̄
```

V̄ = long brassy peripheral
V̈ = medium long, breathy, somewhat centralized
V = very short, very centralized

Morphophonemic alternations take place as follows according to Welmers (1973:29):

> "Alternations between noun singulars and plurals appear to involve most commonly a movement clockwise to the next spoke but in the same position on the spoke; that is if the singular has /ü/, the plural has /ö/; if the singular has /ö/, the plural has /ɔ̈/, and so on until if the singular has /ë/ the plural has /ï/; but if the singular has /ï/ there is no change in the plural (since there is no spoke in the next position clockwise). A less common pattern is precisely the reverse, with the alternation in the plural one spoke counter-clockwise from the vowel of the singular; if the singular is on the /u/ spoke there is no change in the plural. Still other alternations are one step in or out on the same spoke: /ō/ to /ö/, /a/ to /ä/, and the like."

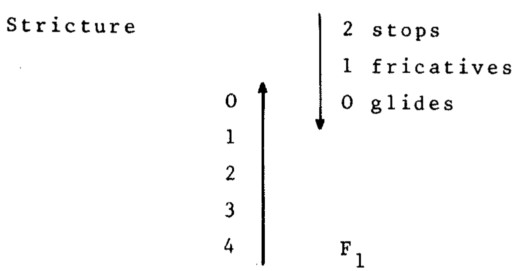
Alternations one step in or out on the same spoke could be accounted for by the n-ary feature [Glottal Stricture]. But there is no feature that could do the "around the clock" patterns. The underlying mechanism must be acoustic, in fact it corresponds very well to the frequency of the second formant, F_2. From any position on the spokes, going clockwise or counterclockwise there is a continuous change of F_2. Thus we need a multivalued feature [F_2]. For Dinka there are seven values. The feature [F_2] is associated with this "around the clock" variation, and it is different from variation in "backness" in our framework.

<u>List of features</u>	<u>Maximum contrasts</u>
F_1	4 (5, if O = glide, fn. 1)
$F_2 - F_1$	3
Round	2
Labial	2
Expanded	2 (3 values)
Lowered F_3, F_4	3
Nasal	2
Long	2
Peripheral	2
Glottal Stricture	2 (7)

<u>"Rule features"</u>	<u>Number of values</u>
Grave	2
F_2	n

Footnotes

1. In K. Williamson's framework [Stricture] is an n-ary feature referring to the size of the passage between two articulators, ranging from complete closure to wide open (for low vowels). Including consonants and vowels in a single feature is probably not correct. At the point where the stricture changes from obstruent to sonorant (i.e. to a glide) the phonetic correlate changes from a basically articulatory to an acoustic mechanism. As both stricture rules and vowel rules may involve glides, what we have is perhaps two features that overlap at the point of glides. If we regard glides as the zero value for each feature, we could also describe how, when a weakening process results in glides, the next step is deletion of the whole segment. Glides are also regarded as end points of the vowel space.



2. The Dan vowels occur long and short. /æ:/ and /ɒ:/ occur only as long vowels in the data from Bearth and Zemp (1967).

dɨ spear	fi: unpleasant odour	bu rotte
dɩ tree	we: salt	bo: beetle
dɵ father	wɛ: sleeping place	dɔ: termite
za judgment	wæ: to collect	bɒ: helper

3. According to Mohrlang (1971) Higi contrasts /i e ə a/ word finally and /ə e a/ word medially. As the phonetic values of /ə e a/ are determined by the surrounding consonants, Mohrlang analyses them all as phonemically central. It must also be pointed out here that the analysis of Higi vowels is by no means clearcut. Wolff (1959) analyses Higi as a six vowel system:

i	ɨ	u
e		o
	a	

4. Lewis' transcription has been converted into that of IPA.

Languages mentioned

Language	Classification
Abe	Kwa/Niger-Congo
Abkhaz	Caucasian
Abuan	Benue-Congo/niger-Congo
Acholi	Eastern Sudanic/Nilo-Saharan
Adyge	Caucasian
Agwagwune	Benue-Congo/Niger-Congo
Akha	Burmese-Lolo/Sino-Tibetan
Albanian	Indo-European
Arabic	Semitic/Afro-Asiatic
Ateso	Eastern Sudanic/Nilo-Saharan
Badaga	Dravidian
Basque	undetermined
Brôu	Mon-Khmer/Austro-Asiatic
Chacobo	Tacana-Pano/Ge-Pano-Carib
Chinese	Han-Chinese/Sino-Tibetan
Czech	Slavic/Indo-European
Dan	Mande/Niger-Congo
Danish	Germanic/Indo-European
Dho Luo	Eastern Sudanic/Nilo-Saharan
Dinka	Eastern Sudanic/Nilo-Saharan
Engenni	Kwa/Niger-Congo
English	Germanic/Indo-European
Estonian	Uralic/Altaic
Faroese	Germanic/Indo-European
Fe'fe'	Benue-Congo/Niger-Congo
Finnish	Uralic/Altaic
French	Italic/Indo-European
Friulian	Italic/Indo-European
German	Germanic/Indo-European
Gujerati	Indic/Indo-European
Higi	Chadic/Afro-Asiatic
Hindi	Indic/Indo-European
Hungarian	Uralic/Altaic
Icelandic	Germanic/Indo-European
Igede	Kwa/Niger-Congo
Ịjọ	Kwa/Niger-Congo
Japanese	Japanese-Ryukyuan/Altaic
Kabardian	Caucasian
Kalenjin	Eastern Sudanic/Nilo-Saharan
Kannada	Dravidian
Kasem	Gur/Niger-Congo
Kashmiri	Indo-Iranian/Indo-European
Kohumono	Benue-Congo/Niger-Congo
Lango	Eastern Sudanic/Nilo-Saharan
Lotuko	Eastern Sudanic/Nilo-Saharan
Luganda	Bantu/Niger-Congo
Mianka	Gur/Niger-Congo
Mixe	no information - spoken in Mexico
Ngemba	Benue-Congo/Niger-Congo
Ngwe	Benue-Congo/Niger-Congo

Norwegian	Germanic/Indo-European
Ọgbia	Benue-Congo/Niger-Congo
Päkot	Eastern Sudanic/Nilo-Saharan
Polish	Slavic/Indo-European
Sele	Kwa/Niger-Congo
Serbocroatian	Slavic/Indo-European
Sisala	Gur/Niger-Congo
Swedish	Germanic/Indo-European
Tibetan	Sino-Tibetan
Turkish	Turkic/Altaic
Twi/Akan	Kwa/Niger-Congo
Udykh	Caucasian
Urhobo	Kwa/Niger-Congo
Vietnamese	Austro-Asiatic

Bibliography

Ali L.H. and R.G. Daniloff. 1972. A cineflurographic investigation of emphatic sound assimilation in Arabic. In A. Rigault (ed.), Proceedings of the 7th International Congress of Phonetic Sciences. Mouton, The Hague, 639-648

Allen C.A, 1974. Studies in the phonology of Sele - the language of Santrokofi. MA thesis. University of Ghana, Legon

Antell S.A., G.K. Cherono, B.L. Hall, R.M.R. Hall, A. Myers, M. Pam. 1974. Nilo-Saharan vowel harmony from the vantage point of Kalenjin. Research Notes 7. University of Ibadan

Bearth T. and H. Zemp. 1967. The phonology of Dan (Santa). Journal of African Languages 6.1:9-29

Bell M. 1867. Visible Speech. London

Bendor-Samuel J.T. 1971. Niger-Congo: Gur. In T. Sebeok (ed.), Current Trends in Linguistics, Vol. VII, Mouton, The Hague

Benediktsson H. 1959. The vowel system of Icelandic: a survey of its history. Word 15:282:312

Bergman R. 1971. Vowel sandhi and word division in Igede. Journal of West African languages 8.1:13-25

Bruce B. 1970. Diphthongization in the Malmö dialect. Working Papers 3. Phonetics Laboratory, Lund University

Chomsky N. and M. Halle. 1968. The Sound Pattern of English. Harper & Row, New York

Cook T.L. 1969. Notes on Kohumono. Research Notes 2.3. University of Ibadan

Dunstan E. 1966. Tone and Concord in Ngwe Nominals. PhD thesis. University of London

Eastlack C.L. 1968. The noun classes of Ngemba. Journal of African Languages 7.1:33-40

Einarsson S. 1928. Beiträge zur Phonetik der isländischen Sprache. Oslo

Emeneau M.B. 1939. The vowels of the Badaga language. Language 15.1:43-47

Fant G. 1960. Acoustic Theory of Speech Production. Mouton, The Hague

Ferguson C.A. 1963. Assumptions about nasals: a sample study in phonological universals. In Greenberg (ed.), Universals of Language. M.I.T. Press, Cambridge, Mass., 53-60

Gregersen K.J. 1973. Tongue-root and register in Mon-Khmer. Forthcoming in Papers from the 1st International Conference on Austroasiatic Linguistics. Oceanic Linguistics. University of Hawaii Press

Gunnilstam O. L973. The theory of local linearity. Reports from the Uppsala University Department of Linguistics 2:38-69

Hadding K. and A.S. Abramson. 1964. Duration versus spectrum in Swedish vowels: some perceptual experiments. Studia Linguistica 18:94-107

Halle M. 1970. Is Kabardian a vowelless language? Foundations of Language 6.7:95-103

Halle M. and K. Stevens. 1969. On the feature 'advanced tongue root'. MIT-QPR 94:209-215

Harshman R. 1970. Parafac. Working Papers in Phonetics No. 16. UCLA

Harshman R. 1971. Parafac: An "Explanatory" Factor Analysis Procedure. Working Papers in Phonetics No. 19:14-16. UCLA

Haugen E. 1958. The phonemics of Modern Icelandic. Language 34:55-88

Hockett C.F. 1955. A Manual of Phonology. Memoir 11 of Indiana University Publications in Anthropology and Linguistics

Hombert J.M. 1974. The role of phonetics in phonology. Seminar paper. University of Ibadan

Hoogshagen S. 1959. Three contrastive vowel lengths in Mixe. Zeitschrift für Phonetik und Allgemeine Sprachwissenschaft 12:111-115

House A.S. and K.N. Stevens. 1956. Analog studies of the nasalization in vowels. Journal of Speech and Hearing Disorders 21:218-232

Hyman L. 1972. A Phonological Study of Fe'fe'. Studies in African Linguistics. Suppl. No. 4. Los Angeles

Jacobson R., G. Fant, and M. Halle. 1951, rev. ed. 1963. Preliminaries to Speech Analysis. M.I.T. Press, Cambridge, Mass.

Kelkar A.R. and P.N. Trisal. 1964. Kashmiri word phonology: a first sketch. Anthropological Linguistics 6.1:13-22

Kelly J. 1966. A note on lip-rounding. Le Maitre Phonétique 125:8-9

Ladefoged P. 1964. A Phonetic Study of West African Languages. Cambridge University Press, Cambridge

Ladefoged P. 1971. Preliminaries to Linguistic Phonetics. University of Chicago Press, Chicago

Ladefoged P. 1975. A Course in Phonetics. Harcourt Brace Jovanovich Inc., New York

Ladefoged P., J. DeClerk, M. Lindau, and G. Papçun. 1972. An auditory theory of vowel production. Working Papers in Phonetics 22. University of California, Los Angeles, 48-75

Lehiste I. 1970. Suprasegmentals. M.I.T. Press, Cambridge, Mass.

Lewis P. 1968. Akha phonology. Anthropological Linguistics 10.2:8-18

Lindau M. 1973. Different vocal tract shapes with similar formant frequencies. Paper presented at the 86th meeting of the Acoustical Society of America

Lindau M. Forthcoming. Phonetic mechanisms of vowel harmony in African languages. Proceedings of the Eighth International Congress of Phonetic Sciences, Leeds 1975

Lindau M., L. Jacobson, and P. Ladefoged. 1972. The feature "Advanced Tongue Root". Working Papers in Phonetics 22. University of California, Los Angeles, 76-92

Lindblom B. and J. Sundberg. 1969. A quantitative model of vowel production and the distinctive features of Swedish vowels. Quarterly Progress and Status Report 1. Speech Transmission Laboratory, Stockholm, 14-32

Lindblom B. and J. Sundberg. 1969. A quantitative theory of cardinal vowels and the teaching of pronunciation. Quarterly Progress and Status Report 2-3, Speech Transmission Laboratory, Stockholm, 19-25

Lindblom B. and J. Sundberg. 1971. Acoustical consequences of lip, tongue, jaw, and larynx movement. Journal of the Acoustical Society of America 50.4:1166-1179

Miller J.D. 1967. An acoustical study of Brôu vowels. Phonetica 17:149-177

Miller P.D. 1974. A critical bibliography on the Tense/Lax distinction in vowels. Ohio State University Working Papers in Linguistics 17:222-231

Mohrlang R. 1971. Vectors, prosodies, and Higi vowels. Journal of African Languages 10.1:75

Ohala J. 1971. The role of physiological and acoustic models in explaining the direction of sound change. Project on Linguistic Analysis Reports, 2nd series 15. University of California, Berkeley, 25-40

Painter C. 1973. Cineradiographic data on the feature 'Covered' in Twi Vowel Harmony. Phonetica 28:97-120

Perkell J.S. 1971. Physiology of Speech Production. M.I.T. Press, Cambridge

Pike K. 1967. Tongue root position in practical phonetics. Phonetica 17:129-140

The Principles of the International Phonetic Association. 1940. University College. London

Retard G.L.A. 1972. L'Agni - Variete Dialectale Sanvi. Annales de l'Universite d'Abidjan Serie - H - V

Ruhlen M. 1973. Nasal vowels. Working Papers on Language Universals 12. Stanford University, 1-36

Sedlak P. 1969. Typological considerations of vowel quality systems. Working Papers on Language Universals 1. Stanford University, 1-40

Stewart J.M. 1967. Tongue root position in Akan vowel harmony. Phonetica 16:185-204

Stewart J.M. 1971. Niger-Congo, Kwa. In T. Sebeok (ed.), Current Trends in Linguistics, Vol. VII. Mouton, The Hague

Sweet H. 1877. Handbook of Phonetics. London

Thomas E. 1969. A Grammatical Description of the Engenni Language. PhD thesis. University of London

Vanvik A. 1972. A phonetic-phonemic analysis of Standard Eastern Norwegian, Part I. Norwegian Journal of Linguistics 26.2:119-164

Wang W. 1968. Vowel features, paired variables and the English vowel shift. Language 44.4:695-708

Welmers W. 1973. African Language Structures. University of California Press, Berkeley and Los Angeles

Williamson K. 1972. Assimilation in Ọgbia. Research Notes 2-3. University of Ibadan

Williamson K. 1974. Some reduced vowel harmony systems. Research Notes 7. University of Ibadan

Williamson K. 1974. Some proposals for multivalued features. Seminar paper. University of Ibadan

Wolff H. 1959. Subsystem typologies and area linguistics. Anthropological Linguistics 1.7 (monograph)

Wolff H. 1969. A Comparative Vocabulary of Abuan Dialects. North Western University Press, Evanston

A PHONETIC EXPLANATION TO REDUCED VOWEL HARMONY SYSTEMS*

Mona Lindau

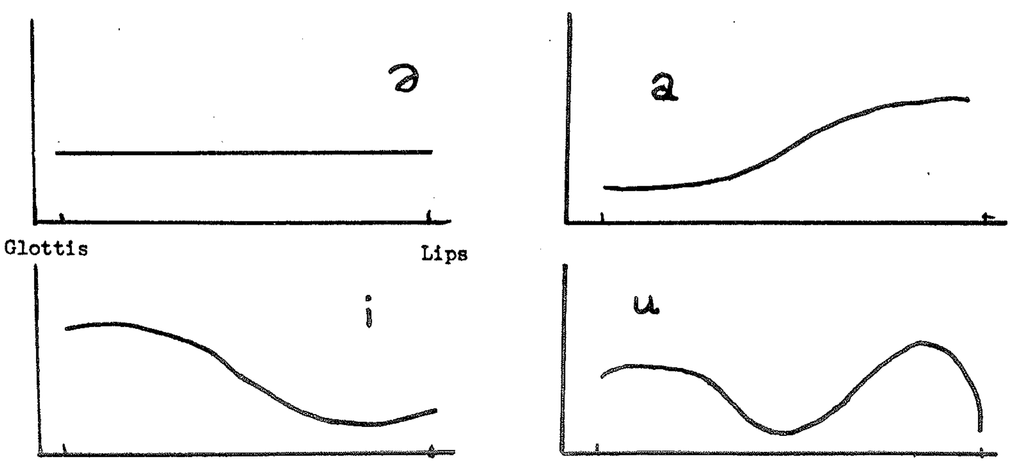
Proto-Kwa and other Proto-Niger-Congo languages are being reconstructed with a vowel system of five tongue-root advanced vowels and five tongue-root-retracted vowels (Stewart, 1971). There are modern Niger-Congo languages that still exhibit a ten vowel system, but most of them are reduced to nine- or seven-vowel systems, and in the case of most Lower Niger languages to eight-vowel systems. In this paper I will consider some common patterns of reduction, and attempt to provide a phonetic explanation for these patterns. The explanations involve predictions from a theory that is developed independent from theories of vowel production and phonological constraints on vowel systems, namely acoustic perturbation theory. Because of the independence of perturbation theory and theories of vowel systems the proposed explanation would be a theoretically strong one, provided it also stands up to closer scrutiny.

Perturbation theory[1] has recently been applied to theories of speech production to answer questions about the relationships between articulatory configurations and corresponding formant frequencies. We can ask questions about these relationships from two angles, either given a certain articulatory configuration and an articulatory change, what is the acoustic effect? or given a certain point in a formant space, and a certain formant change, what articulatory configurations could have accomplished this? At the moment the first question has had the larger amount of research devoted to it, so the discussion here will be restricted to questions of the first type.

---

* I wish to thank Kay Williamson and Peter Ladefoged for discussions and comments on earlier versions of this paper.

In applying perturbation theory to vowel production, the articulatory configurations are described in terms of area functions of the vocal tract. An area function "described the cross-sectional area of the oto-pharyngal cavity as measured perpendicularly to the longitudinal midline of this cavity. This midline runs from the glottis to the labial orofice of the mouth." (Öhman, 1973.) The cross-sectional area at any point along the vocal tract is calculated from the saggital distance at the same point along the vocal tract. Area functions of [ə, i, a, u] are illustrated below (Öhman, 1973, adapted from Fant, 1960).



The abscissa shows the distance from the glottis, the ordinate the cross-sectional area. The curves show approximate cross-sectional areas along the midline from glottis to lips. The vowel [ə] is simplified to a straight line, because the formant frequencies of this simple tube, closed at one end and open at the other, are very regularly distributed. F_1, F_2, and F_3 are 500 Hz, 1500 Hz, and 2500 Hz, respectively, for a vocal tract length of 17 cm. All other vowels are regarded as perturbations in different directions from the "neutral" schwa. An algorithm has been defined by

Öhman for calculating the formant frequencies of the continuous cross-sectional area curves of vowels that deviate from schwa. The formant frequencies corresponding to such deviation curves depend on whether the deviation curve is symmetric or antisymmetric about the midpoint. The symmetry of a curve is defined by taking the midpoint on the x-axis (distance from the glottis) and look at how the two parts of the curve to the left and right of the midpoint relate to each other. In a symmetric deviation curve the right part of the curve is a positive mirror image of the left part. If folded in half along a line midway from the glottis to the lips the right and left parts would cover each other. The deviation curve of [u] has a strong tendency to symmatry. In an antisymmetric curve the right half is negative mirror image of the left half. Both [i] and [a] have antisymmetric deviation curves.

Calculation of formant frequencies of a vocal tract tube with minimal termination impedance (i.e. unrounded lips) demonstrates two interesting facts. Firstly, any perturbation of such a vocal tract where the deviation curve is symmetric about the midpoint will have no acoustic effect on any of the formant frequencies. Secondly, the largest acoustic effects in such a vocal tract tube will be achieved by such perturbation of the vocal tract of schwa as result in deviation curves that are antisymmetric about the midpoint. In other words, moving the front of the tongue and the tongue-root in the same direction in relation to the roof of the mouth and the back pharyngal wall will have no acoustic effect, as long as there is no liprounding. Given the same lip condition, large acoustic differences will be obtained by moving the tongue body and tongue-root in different directions in relation to the roof of the mouth and the back pharyngal wall. Notice at this point that in going from i to [a] (where [a] is taken to be a low vowel midway between cardinal [a] and [ɑ]), the articulatory configurations are mainly antisymmetric. So for conditions with closure at the glottis and no impedance at the lips the most efficient way to vary formant frequencies (i.e. vowel qualities) corresponds precisely to that of the so called front vowels.

Conditions with higher impedance at the lips (i.e. rounded or closed lips) have not been worked out in any detail. But it is known that if a tube is terminated with a constriction providing a considerable impedance, then symmetric perturbations do affect formant frequencies.
It is, however, worth pointing out that back vowels that are basically rounded, also have their main constriction around the midpoint of the vocal tract, so that their vocal tract configurations tend to be symmetric. Apparently, the most efficient way to vary vowel qualities with rounded lips is by means of symmetric configurations[2].

We are now in a position to discuss real vowel systems. The following is an attempt to explain some common patterns of vowel mergings that occur in Niger-Congo languages with tongue root harmony.

One common pattern in these languages is that the vowels /ị/ and /e/, and /ụ/ and /o/ have merged, so that an earlier nine vowel system has become a seven vowel system. It is interesting to note that these sound changes seem to start with the merging of the two front unrounded vowels. We find languages today where these two front vowels have merged, or are merging, but the back rounded ones are not. Akan constitutes an example of a language where /ị/ and /e/ are in the process of merging, but there is no sign of /ụ/ and /o/ merging. Figure 1 is a typical formant frequency chart of eight Akan vowels (/a/ is excluded). The vowels were pronounced in short utterances by one speaker, and each utterance repeated five times. There is complete overlap of the formant frequencies of /ị/ and /e/. The corresponding back vowels /ụ/ and /o/ are kept acoustically separate by the second formant. It is worth noticing here that in Akan /ị/ and /e/ do not seem to contrast in stems, while /ụ/ and /o/ do.

Figure 2 shows typical tongue shapes of /ị/ and /e/ superimposed on each other, and of /ụ/ and /o/ superimposed on each other. Note that the highest point of the tongue in the vowels [ị] and [ụ] which are traditionally called high vowels, is much the same as it is in the vowels [e] and [o], which are traditionally called mid vowels. Although the tongue-shapes of /ị/ and /e/, and of /ụ/ and /o/ differ in very much the same way, the corresponding acoustic effects do not differ in the same way. The lack of a one-to-one correspondence between articulatory configurations and acoustic results is of course well known, and the more interesting question is the specification of what kinds of tongue shapes result in the same formant frequencies, and what kinds result in different formant frequencies. The phonologically mid /e/ and /o/ have an advanced tongue-root, and thus a larger pharyngal cavity than the phonologically high /ị/ and /ụ/. The mid /e/ and /o/ also have a lower front of the tongue, and thus a larger mouth cavity than the high /ị/ and /ụ/. Both halves of the vocal tract tube are larger in the mid vowels than in the high vowels. In other words, both /ị/ and /e/, and /ụ/ and /o/ are produced by perturbing the neutral schwa in such a way that the deviation curves are symmetric about the midpoint. This fact now provides an explanation as to why /ị/ and /e/ should merge, but not /ụ/ and /o/, as perturbation theory predicts that symmetric deviation curves will not differ as to formant frequencies for unrounded lip conditions, but will for rounded vowels. The deviation curves of all four vowels are symmetric, but only /ị/ and /e/ are unrounded. Therefore /ị/ and /e/ predictably have the same formant frequencies, while those of /ụ/ and /o/ differ because of the lip-rounding. Naturally, vocal tract shapes from real speakers will not be perfectly symmetric about the midpoint. In this case the high and mid vowels are approximately symmetric, and the acoustic effect is overlapping of the formant frequencies for the unrounded vowels, and non-overlapping for the rounded ones.

By relating the independently developed perturbation theory to phonetics the merging of /ị/ and /e/ has now been provided with an explanation that does not require the use of adhoc concepts like "marked - unmarked", and the like. I propose that in all those Niger-Congo languages where /ị/ and /e/ have merged, or are in the process of merging, the reasons are not to be found in looking for the tongue shapes becoming the same, but the explanation for the merging has an acoustic basis. Because of their both having symmetric tongue shapes they become acoustically the same, and are therefore starting to be perceived as the "same" vowel.

So why do earlier nine-vowel systems mostly become seven vowel systems by also merging /ụ/ and /o/ later, and not just eight vowel systems? I have no neat explanation for this, but I suggest that once /ị/ and /e/ overlap acoustically, this will create a structural pressure towards making the systems symmetric again by merging the corresponding back vowels. It is a fact that Niger-Congo vowel systems with harmony have a strong tendency towards symmetry. To sum up: a common development from a nine vowel system to a seven vowel system starts by unconditional merging of /ị/ and /e/ for acoustic reasons, then /ụ/ and /o/ merge for reasons of structural pressure towards symmetry.

When the tongue-root mechanism is involved in vowel production there is of course a very good possibility that /ị/ and /e/ will be articulated with symmetric tongue shapes and therefore merge, as has happened in many Niger-Congo languages. But there is another possibility. The vowels /ị/ and /e/ will merge, unless the pressures of communication within the language act to prevent this from happening. If a speaker wants to keep these two vowels distinct, he can easily do so by changing the tongue shape of /e/ to a more assymmetric shape. This can be accomplished by the following strategies: just lowering the front of the tongue, or just retracting the tongue root, or combining the two gestures. In the first and third case the acoustic effect is an increase of F_1. Published sources do not show how a pharyngal decrease by itself would affect the formant frequencies, but it seems likely that it would have the same effect, namely an increase of F_1. So whatever adjustment towards an asymmetric tongue shape a speaker chooses, the effect will be vowel lowering, and /e/ becoming more [ɛ]-like in quality. It might even merge with /ẹ/.

As there is no problem in keeping the rounded vowels with symmetric shapes separate, one would expect /u/, /ụ/, /o/ and /ọ/ to stay intact.

The pattern discussed above, where /ị/ and /e/ are kept separate, but /e/ has merged, or is merging, with /ẹ/, is exhibited by the development from Proto-Lower Niger to modern Lower Niger (Williamson, 1975). Most dialects of Igbo have an eight vowel system that is considered to be curiously skewed:

i   u
   ị   ụ
e[ɛ]   o
   a   ọ

K. Williamson (1975) posits a ten vowel system in Proto-Lower Niger. The Onitsha dialect of Igbo has the same phonological eight vowel system as most dialects of Igbo, but it has phonetically a nine-vowel system:

i   u
   ị   ụ
e   o
   [ɛ]   ọ
      a

e and [ɛ] do not contrast: /e/ is realized as [ɛ] before tongueretracted vowels, as [e] elsewhere. Most dialects of Igbo and the Onitsha dialect can be interpreted in terms of the above discussion. If Igbo /ị/ and /e/ are to be kept separate, the tongue shape of /e/ must become more asymmetric. The distribution of [e] and [ɛ] in Onitsha indicates that these speakers have chosen to retract the tongue root of /e/ as the particular strategy for making /e/ more different from /ị/. In another dialect of Igbo, namely the Umuchu dialect, where radiographic evidence is available (Lindau, forthcoming), it is clear that the speaker has lowered /e/ by lowering the body of the tongue but retaining the tongue root distinction. (See figure 3.) At an earlier stage in the development from Proto-Lower Niger to today's Lower Niger languages the tongue-root advanced /a/ merged with /ạ/ (as in Onitsha) or with /ẹ/ (as in Ịka). As expected from perturbation theory these unconditioned mergers have affected only unrounded vowels, while the rounded vowels remain unaffected.

Looking at the Igbo vowel system in the light of perturbation theory thus explains the apparently "unnatural" skewness as a quite natural system, arising from an original "desire" to keep /ị/ and /e/ distinct, without making /e/ "higher" than /ị/.

Notes

1. The following description of perturbation theory is summarized from Öhman (1973).

2. There is as far as I know no evidence of what happens with rounded lips and antisymmetric tongueshapes. It is worth noting however, that front rounded vowels do differ acoustically, but the acoustic space of the front rounded vowels is considerably smaller than that of the front unrounded.

References

Fant G. 1960. Acoustic Theory of Speech Production. The Hague, Mouton

Lindau M. 1973. "Different vocal tract shapes with similar formant frequencies". Papers given to the 86th meeting of the Acoustical Society of America

Lindau M. Forthcoming. UCLA Ph. D. Dissertation

Öhman S. 1973. Untitled ms. UCLA

Stewart J.M. 1971. "Niger-Congo: Kwa". In Sebeok (ed.), Current Trends in Linguistics, Vol. VII: 179-212

Williamson K. 1975. "Proto-Lower Niger vowels: some first thoughts". Seminar paper, Department of Linguistics and Nigerian Languages, University of Ibadan

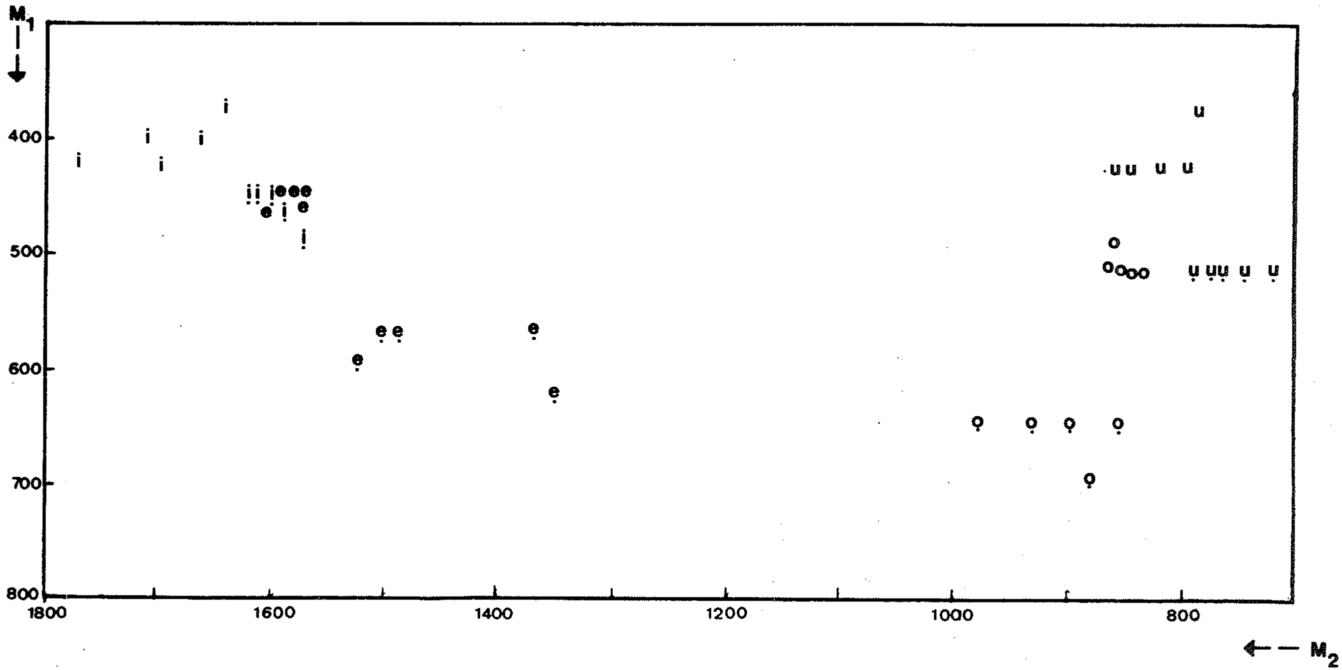


Figure 1. Formant chart with pitches in mel of eight Akan vowels of one speaker.

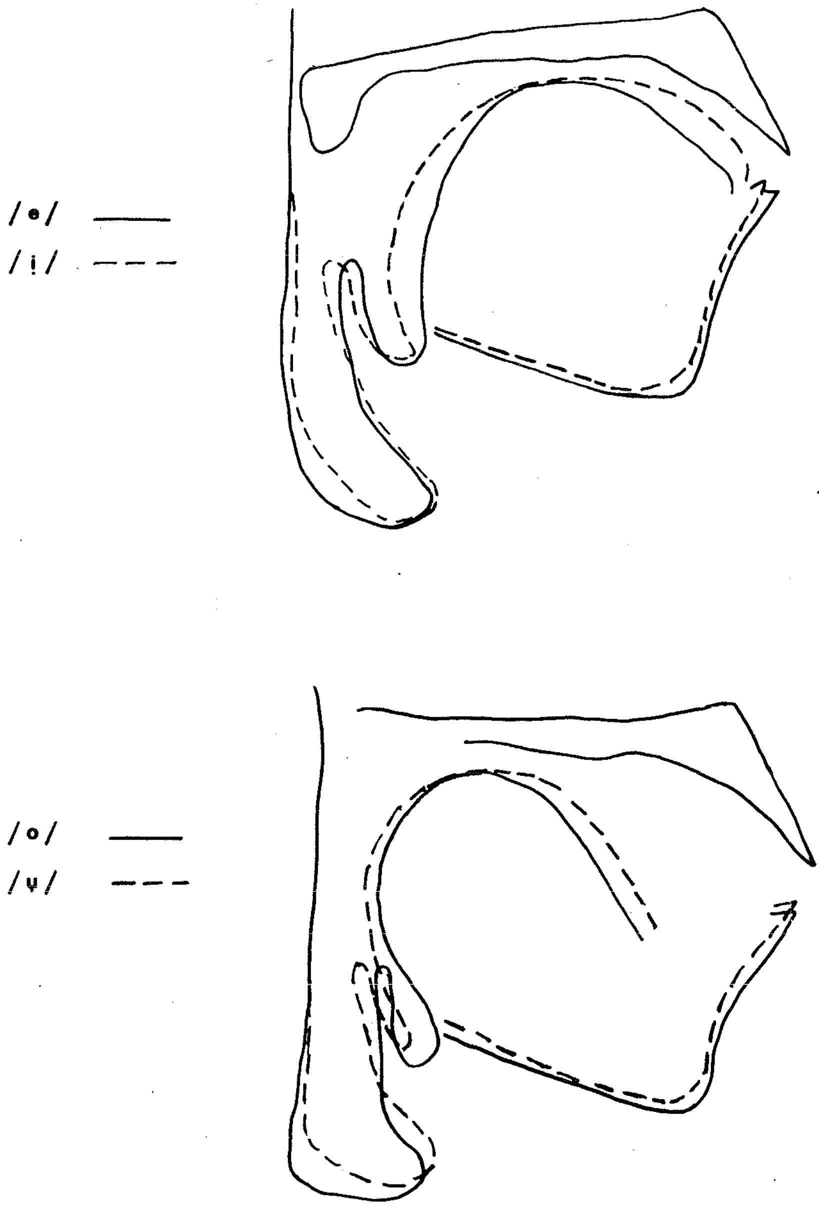


Figure 2. Tongue shapes of /e/ and /i̩/, and /o/ and /u̩/ of a spekaer of Akan.

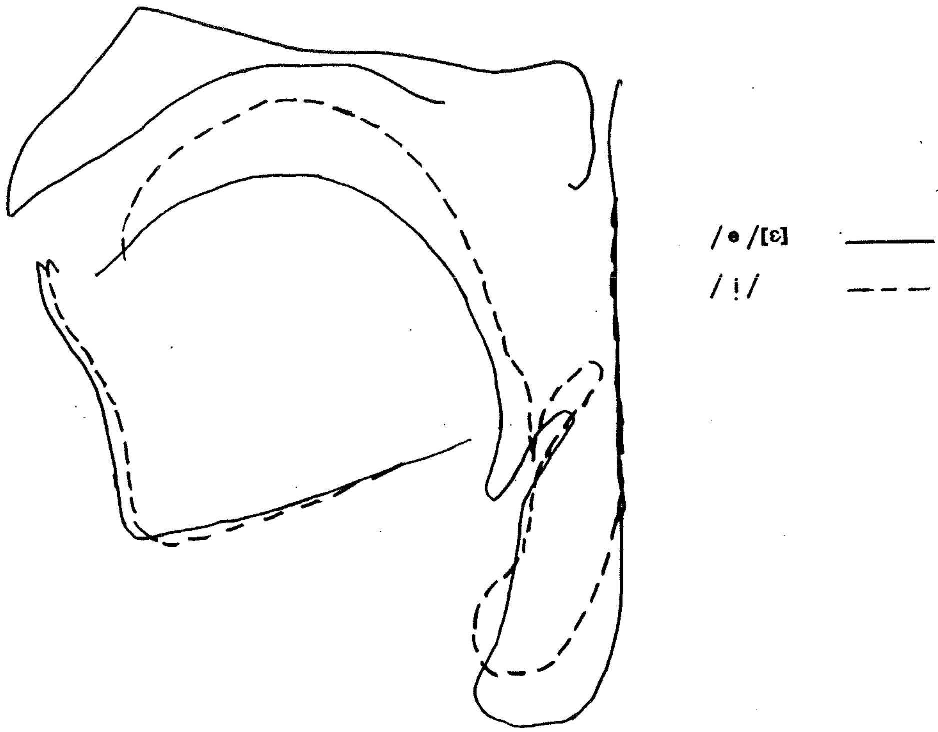


Figure 3. Tongue shapes of /e/ ([ɛ]) and /i̩/ of a speaker of the Umuchu dialect of Igbo.

Working Papers 11·1975 Lund University
Phonetics Laboratory General Linguistics

THE WEAKNESSES OF THE TONGUE-ARCHING MODEL OF VOWEL ARTICULATION

Sidney Wood

SUMMARY

Published X-ray tracings of vowel articulations are examined in the light of criticisms made against the tongue-arching model during the past 70 years. This corroborates the charges made against the model of failing to prescribe tongue position correctly. The implications of this failure are discussed. The constancy of vocal tract configurations, compared with the ambiguity of tongue arch position, points to a more suitable type of model for vowel articulation in which individual gestures combine to shape the vocal cavities to the resonator configurations appropriate to the sound quality.

CONTENTS

Introduction
The historical background
Examination of published X-ray tracings
   Methods and material
   Tongue retraction
   Tongue height
The weaknesses of the tongue-arching model
   Physiological weaknesses
   Predictive capability
   Explanatory power
An alternative articulatory model
Some implications for phonology

INTRODUCTION

The two-dimensional tongue-arching model has provided the predominant theoretical vowel articulation framework for phonetics and phonology during the past 100 years. There was never any real opportunity to test the physiological basis of the model before the introduction of radiology at the turn of the century provided the first means of investigating tongue positions and vocal tract configurations. Since then, a number of unexpected tongue positions have regularly been reported for some vowels, which raises serious doubts as to the predictive capability of the model. The following report contains an examination of 38 sets of X-ray profiles (published during the past half century or so), in order to ascertain how widespread and serious the apparently anomalous tongue arch positions might be. This is followed by a discussion of the implications of the results. Such anomalies need not necessarily be serious for phonetic theory - it might suffice to revise minor details of model design. However, current knowledge of vocal tract acoustics and the neuromotor level of speech production show respectively that the explanatory power and physiological foundation of the model are also very weak. The sum of these weaknesses is that the representation of vowel articulation provided by the model is not only inaccurate but also irrelevant to the processes of speech production. The model has constituted an unnecessarily weak link in current linguistic theory.

The main reason for the survival of the tongue-arching model to the present day has been the absence of a substitute articulatory model. Examination of the published tracings indicates that the vocal tract configuration is more constant than the tongue arch position. This matches the known regularity of spectral character. From this I conclude that the speaker is striving to create a definite target resonator shape appropriate to the intended quality and that it is reasonable to expect similar regularity at the articulatory and neuromotor stages. This provides a framework for a substitute model in which articulatory gestures, with known neuromotor activity, combine to create specific resonator shapes with known resonance properties. Since such a model is a more effective instrument of prediction and explanation, it will yield more realistic phonetic solutions to phonological problems. Phonology has therefore much to gain from adopting such a model in place of the tongue-arching model.

There has always been a school of phonetics during this same period that has expressed scepticism over some or all of the attempts to describe speech in articulatory terms of any sort. Its adherents have instead emphasized that speech communication is possible because definite sound qualities are heard and understood by a listener. They have therefore insisted that speech should be described in acoustic or perceptual terms. Some, especially towards the end of the 19th century when there was bitter rivalry between "acoustic" and "organic" schools, wished to shun articulation altogether. Others have since then continued to disregard articulation on the grounds that a speaker can utter a sound in a variety of ways, this inconstancy providing an apparent proof of the irrelevance of articulation. However, articulation is undeniably a necessary stage in the speech chain that merits description not only for its own sake but also because it is an indispensable link in speech communication between speaker and listener. Phonetics requires a comprehensive account of speech production and not a one-sided description restricted to any single phase of speech communication.

THE HISTORICAL BACKGROUND

The tongue-arching model portrayed vowel articulation in terms of two dimensions, the vertical and horizontal movement of the top of the tongue hump, by which vowels could be located in a Cartesian coordinate system (or, as D. Jones put it [1967: § 151], "by means of a system similar to the latitude and longitude principle used in geography"). Each vowel was said to have its own tongue position coordinates in the high/low and front/back dimensions, and a complete vowel system appeared as a polygon whose shape was characteristic for that language. It seemed perfectly natural to discuss vowel systems in geometrical terms by referring to the spatial relationships between points in the polygon. Other articulatory variables were often disregarded in the simple two-dimensional portrayal since they were said to be correlated with tongue arch coordinates for positions in the vowel polygon - for example, rounded front vowels have been described as "slightly retracted" relative to their unrounded counterparts, lax vowels "centralized" relative to the corresponding tense vowels, and so on.

Prior to the introduction of the tongue-arching model (in fact, ever since antiquity) vowel production had been understood in terms of three distinct tongue gestures (aimed at the pharynx, hard palate or velum), jaw opening and lip position. These gestures could easily be seen but in the absence of adequate acoustic theory their spectral consequences could not be properly understood or even known.[1] The mid-19th century philologists and Christian missionaries, handicapped by their limited knowledge of vocal tract acoustics and by the impossibility of making quantitative investigations of internal articulation, had found it increasingly difficult to account for finer distinctions of vowel quality or to accomodate the unfamiliar vowel qualities that were being discovered in the languages of the world. The new tongue-arching model appeared far more attractive and superseded the ancient model during the second half of the 19th century. For some reason, it also gained the reputation of being more scientific than the ancient model. The new model was almost universally adopted by the new movements that dominated work in phonetics towards the end of the century - the neogrammarians, the language teaching reformers and the IPA - and while some controversy between supporters of the rival schools still lingered on, the ancient model hardly survived into the present century apart from newer editions of earlier works. Helmholtz (1863) had referred to the ancient model, and his book reappeared in a 6th edition in 1913. The same model was preferred by the laryngologist Gutzmann for his speech handbook (1909) and he still retained it in the 2nd edition in 1929. Russel (1928) found that the ancient model gave a better picture of vowel articulation and the shaping of the vocal cavities (although above all he preferred to describe vowels by their acoustic and impressionistic characteristics). But among phoneticians and phonologists, the ancient model was already lost.

There is a fundamental conceptual difference between the two types of model regarding tongue movement between front and back. The ancient model recognized distinct pharyngeal, palatal and velar gestures. In the early years of the 19th century it was common to portray the ancient model in the form of a tree (n.b. not a triangle) with velar and palatal series branching off from the basic pharyngeal configuration, in simplified form thus:

```
      a
   o     e
   u     i
```

In practice, the tree was augmented with additional branches for rounded palatals and plain velars. Contrary to widely held belief, the insertion of these brances between those depicted above never implied intermediate tongue positions[2]. In contrast, the tongue was allowed free movement in any direction in the tongue-arching model and, in particular, the tongue hump was said to occupy any position along the front/back axis. The possibility of intermediate tongue positions between front and back was explicit. Bell (1867) recounted how, after a sleepless night spent puzzling over the articulation of the vowel of sir, he came upon the idea of the tongue not only rising up to the hard and soft palates but also centrally between them. At a stroke of the imagination he created a whole new series of vacant matrix cells for the "difficult" vowels. This invention was revolutionary. The next step - to envisage the front/back axis as a continuum with any number of positions - was easy. An essential component of the tongue-arching model was this division of the horizontal axis into at least three positions. Many phoneticians, believing in a concept of continuous advancement or retraction, claimed that small horizontal adjustments of tongue position yielded modified vowel qualities. They spoke of an "advanced" [i] or a "retracted" [e], for example. It was this feature - alien to the ancient model - that made it so attractive in the 19th century, providing a seemingly simple tool for describing finer or unusual contrasts of vowels.

Attempts were made to relate the tongue arch positions to the vowel spectrum. It was formerly believed that the top of the tongue arch was the limit of a buchal cavity in which a characteristic vowel resonance was formed, and later that the arch constituted a neck between a buchal cavity and a pharyngeal cavity, each with its own resonance. The role of varying tongue height and retraction was said to be to vary the volume, and hence the resonance, of the buchal cavity. Now that the acoustics of the vocal tract are better understood and the source-filter theory generally accepted (Chiba and Kajiyama, 1941; Stevens and House, 1955, 1961; Fant, 1960) we have learnt that this role attributed to the tongue arch was a misconception. The location of the top of the tongue arch below the palate is only indirectly (and not always predictably) related to the configuration of the vocal cavities and the true place of narrowing in the vocal tract (cf. Fant, 1960: §§ 2.32, 2.33). The true place of narrowing can theoretically occur at any point along the vocal tract although in practice it occurs at one of four - along the hard palate, along the soft palate, in the upper pharynx and in the lower pharynx[3].

The tongue-arching representation of vowel articulation was never confirmed - on the contrary, it was discredited in one of the first genuine opportunities for testing its validity (Meyer, 1910). From the 1860s until the introduction of radiology at the turn of the century there had been no means available for observing or measuring the shapes, positions and sizes of the internal articulators and cavities, apart from palatography or the mirror and probe. The articulatory hypothesis underlying the tongue-arching model was refutable in principle but in reality the means for testing it were not available for a further three or four decades. Grandgent (1890) had devised a novel method of fitting different sized discs into various parts of the vocal tract to measure its cross-section and the overall picture he obtained of the cavities was remarkably good. In particular he was one of the first to point out how the back of the tongue falls away sharply in palatal vowels, leaving a far larger pharyngeal cavity than anyone had hitherto reckoned with. For comparison, the speech physiologist Brücke's (1856) profiles, based on anatomical sections, had a distinctly bulging pharyngeal tongue outline for [i]. But the numerous repetitions of a vowel articulation necessary for Grandgent's method meant that his measurements were very coarse and concealed differences of tongue arch position smaller than a millimetre or so. They did not therefore show up the anomalous tongue heights that were later reported from X-ray investigations. Atkinson (1898) had used a similar probing method. Even more ingenious was Meyer's plasto-palatographic method (1910) in which fine strips of metal foil suspended from a false palate were deformed by the tongue so that they retained an imprint of its contour. Meyer found that the tongue was lower for "lax" /ɪ/ than for "tense" /ē/ (German, Dutch and Swedish informants) contrary to expectations and contradicting the predictions of the tongue-arching model. Meyer published these results in the Festschrift honouring Vietor, who (1914) agreed that they showed earlier notions about tongue articulation to have been largely erroneous. Vietor announced his intention of altering his popular textbooks of phonetics but he never did so. Chlumsky (1913) received Meyer's work with caution. In particular he was unable to obtain good results with the plastopalatographic method.

The first X-ray inspection had been performed just before the turn of the century as soon as the new invention had become available (Scheier, 1909) and a little later it had become possible to photograph the image and thereby conserve a more faithful and accurate reproduction (Meyer, 1907). These authors had investigated tense German vowels, and the omission of the lax vowels meant they had no opportunity to observe the unexpected /ē – ɪ/ tongue height "inversion" subsequently discovered by

Meyer. Kruisinger (1925) noted that "high" /ī/ and "mid" /ē/ were equally "high" on Meyer's radiograms. Russel (1928) took his first radiograms in order to demonstrate to his students the tongue-arching model, but failed to obtain a set of tongue positions that were convincing enough for the purpose. After taking several thousand radiograms from over 400 subjects, he concluded instead that the model was fallaceous. In addition to the [ɪ - e] height inversion, Russel observed that [ɔ] was often lower than [a]. He availed himself of every possible opportunity to attack the model, e.g. (1935). On the other hand, Carmody's (1937) faith in the model was not shaken by the irregular tongue arch positions he had discovered in Holbrook's sets of radiograms (for example, that "low back vowels depend mostly on lip position for their distinctive quality and so must be merged into a vague field which bounds their variations", and again, that "English ʌ is too variable to locate without further material since in our two tracings it falls once inside the quadrilateral and once directly behind o"). He found it meaningful to superimpose tongue arch diagrams for different speakers and languages and to describe the differences in terms of advancement-retraction and raising-lowering. He dismissed criticism of the model as coming "unfortunately from teachers acquainted with phonetics only at second hand". I wonder what Russel, whom he had named, said to that. On the other hand, Russel's own references to dogmatic acceptance of "unproved theories founded on fantasy" and to "philologists and others unacquainted with scientific phonetics" doubtless also upset many scholars in the 1920s and 1930s. Nevertheless, lateral profile radiograms of the vocal tract did frequently seem to reveal tongue arch positions that were confusing rather than enlightening with reference to the tongue-arching model. Many investigators must have experienced misgivings if not direct disappointment over puzzling X-ray results after all the trouble, expense and (not least) dangers involved in their work.

Much of the criticism of the tongue-arching model in the 19th century was internal and was concerned with the definition of features and the correct feature specifications of particular vowels or with the design of the model. For example, Bell classified the vowel of English <u>let</u> as "low-front-narrow" while Ellis, Sweet and Storm preferred "mid-front-wide". There was controversy towards the end of the 19th century as to whether "height" referred to the mandible (the traditional view) or the tongue (the new view, referring to internal resonator configuration). Not until very recently (Lindblom and Sundberg, 1971) have the individual contributions of the jaw-opening and tongue elevation been assessed separately.

At the same time, there was external opposition, especially from those who insisted that since speech consisted of sounds it should only be described in acoustic or auditory terms. Lloyd (1890) deplored the hostile rivalry and mutual disregard between the "organic" and "acoustic" schools. He pleaded "it is evident to a dispassionate observer that there is here no true place for partisanship, that neither line of investigation ought rightly to exclude or overlook the other, but that each is necessary to the other's completeness". The supposed physiological foundation of the model was undermined by Meyer's work in the first decade of the present century and finally destroyed by Russel's in the 1920s and 1930s. In the 1940s there came a new attack from a different angle. Joos (1948: §§2.35, 2.36) insisted that those phoneticians who believed they could feel the tongue positions by some kinesthetic sense were the victims of self-deception. They were really judging the vowels by auditory impressions. A similar conviction had already been expressed by Russel, but Joos had spectral evidence to strengthen this view. Judgments of height are usually related to the frequency of the first formant and judgments of advancement-retraction to the frequency of the second formant. Further confirmation has been provided by the experiments of Ladefoged (1967: chapt. 2).

Although the tongue-arching model has been discredited for more than half a century, it has never been completely disavowed. It still occupies a central position in phonetic theory, both for teaching and research as well as for phonology, as a glance through the phonetics and linguistics manuals and journals will show. But Meyer's and Russel's results were embarrasing and the reactions varied. Meyer's own solution to the crisis was a proposal that "tense" and "lax" vowels differ in vocal fold presure and in air flow (1913). Chlumsky (1914) was critical and the idea was hardly taken seriously by other phoneticians. A rare exception was a philologist and master at the Imperial High School of Zaborze, M. Leky, who while on war service completed a treatise on phonetics in which airflow variation is given a central role (1917).

Many, like Kruisinga (1925) or Russel (1928), held that the acoustic school's impressionistic analysis of speech was the better way. It seemed that there was a far greater constancy in the spectral character of speech than in articulation. Many held that articulation, seemingly so variable, was irrelevant in contrast to the spectral constancy. This coincided with the advances in design of spectro-analysers and other acoustic instruments (Joos, 1948; Fant, 1958) and a new and hitherto largely unexplored field was opened up to determine the spectral character of speech segments for many languages and to discover the acoustic contrasts and cues preferred by listeners.

Others, either sceptical and preferring to wait and see, or wanting for something better, retained the tongue-arching model. Jespersen, in later editions of his phonetics handbook, faithfully reported the anomalous tongue heights found by Meyer and observed that vowel theory had been shaken. But hesitated to draw the consequences because of the subjectively felt affinity of [i] to [ɪ] and [e] to [ɛ] and he therefore retained the traditional view: "und wenn ich trotz aller Annerkennung von Meyers vorzüglicher Arbeit auch in dieser Ausgabe im wesentlichen die alte Lehre festgehelten habe, geschiet dies, weil m.E. der übereinstimmenden subjektiven Abschätzung vieler Beobachter auf Grund überaus zahlreicher Warnehmungen ein grosser Wert beizumessen ist". He hoped further investigations would be made and suggested that the behaviour of the dorsum of the tongue would turn out to be more important than the front for vowel articulation. Many phoneticians doubted whether experimental design and methods had been satisfactory. Chlumsky (1913) failed to reproduce Meyer's plastopalatographic results. Others feared that contrast chains and sustained utterances distorted the articulation of X-ray subjects, despite the assurances of practitioners like Russel (1928) or Gutzmann (1930), or public demonstrations by S. Jones (1929) who pronounced the name of the Welsh village Llanfairpwllgwyngyllgogerychwyrndrobwllllantisiliogogogoch with one silver chain along the tongue and another thorugh the nose and down over the velum. Meyer's results were rarely mentioned in other phonetics handbooks.[4] The model continued to enjoy popular acceptance.

Many have continued to rely on the model simply because it has provided a convenient abstract classification system fulfilling a foremost requirement of linguistics during this period however shaky the model of production on which the classifying features have been based. Any other set of features would have served equally well. Classification is an example of the lowest level of measurement, the nominal scale, where one-to-one transformations of the classifying labels are permissible. An abstract classifying system is consequently not affected by any errors of fact regarding speech production providing the categories remain intact. Scholars, whose only requirement has been for a classification system, have been able to continue, deaf to the theoretical crisis surrounding tongue articulation.

One reason for the retention of the tongue-arching model has been the lack of a substitute. Even recently, Ladefoged (1971; chapt. 8), after recognizing that the terms of the tongue-arching model are often not in accord with the physiological facts and that "it is difficult to understand how phoneticians could persist in considering that the traditional articulatory categories provide an adequate specification of vowels", has nevertheless once again retained the tongue-arching model in an elementary text book. He added the reservation that "in descriptions of vowels, although a pseudo-articulatory terminology may provide an adequate set of labels for auditory descriptions, we have seen that we do not have, as yet, a set of articulatory parameters which will specify vowel quality". In the purely acoustic tradition of phonetics, Russel had suggested a set of impressionistic features for describing vowel qualities. Similarly, there are the acoustic features of Jakobson et al. (1952) based on the spectral character of speech segments. The simplest acoustic alternative has been a one-to-one substitution of falling F_1 for "height" judgments and falling F_2 for "retraction" judgments (Joos, 1948; Delattre, 1951) or falling F_2-F_1 difference for "retraction (Ladefoged et al. 1971a). But for articulation, the ancient model displaced by the tongue-arching model belonged irretrievably to the unscientific past. Yet it is interesting to note that three of the acoustically relevant constriction locations in the vocal tract coincide with the three tongue gestures ofthe ancient model (pharyngeal, palatal and velar), showing the latter to have been a sounder view of vowel articulation than its 19th century opponents in the tongue-arching school were prepared to admit. There has been a slender tradition among acoustics theorists from Helmholtz through Paget and Russel to Chiba, Kajiyama, Stevens, House and Fant on which an alternative to the tongue-arching model may be based.

EXAMINATION OF PUBLISHED X-RAY TRACINGS

Methods and material

I have examined 38 sets of X-ray tracings of vowel articulations from 15 different languages (published during the past half century of so) in order to discover how widespread and serious the irregular tongue arch positions might be. If the anomalies are rare, they may be looked upon as accidentally deviant articulations that can be disregarded. If they occur more frequently, it will be necessary to consider just how misleading the tongue-arching model might be and to weigh the implications for phonology.

I have collected the following sets, whose authors covered a wide range of interests such as language teaching, linguistics theory, dialectology, acoustics, speech therapy, laryngology and so on:

Meyer (1907), German; Scheier (1909), German; Polland and Hála (1926), Czeck; Parmenter and Treviño (1932), Spanish; Carmody (1936), Holbrook's German; Carmody (1937), Holbrook's French (3), Spanish, Italian, Portuguese, Am. English (2), S. Br. English, Russian, Polish; Chlumsky et al. (1938), French; Sovijärvi (1938), Finnish; Chiba and Kajiyama (1941), Japanese, German; Mazlovà (1949), Zábřeh dialect of Czeck; Ohnesorg and Švarny (1955), Chinese (3); Skaličková (1955), Korean; Koneczna and Zawadowski (1956), Russian (4); Korlén and Malmberg (1959), Strenger's German; Strenger, Swedish; Hála (1959), S. Br. English; Fant (1960), Russian; Wängler (1961), German; Malmberg (1966), Strenger's Spanish; Perkell (1969), Am. English; Perkell (1971), Am. English; Pétursson (1974), Icelandic.

Each tracing has been photographed and enlarged to natural size. The tracings have been reproduced to a scale that provides overall vocal tract lengths in the range 15-19 cms (depending on the vowel) for male speakers and somewhat shorter for female speakers. Comparison of such features as cervicle segments, incisors and mandible, maxilla and hyoid bones ensured that all articulations in one set were reproduced to the same scale.

Some authors warned that it would be impossible to superimpose their tracings for comparison owing to distortion arising from different points

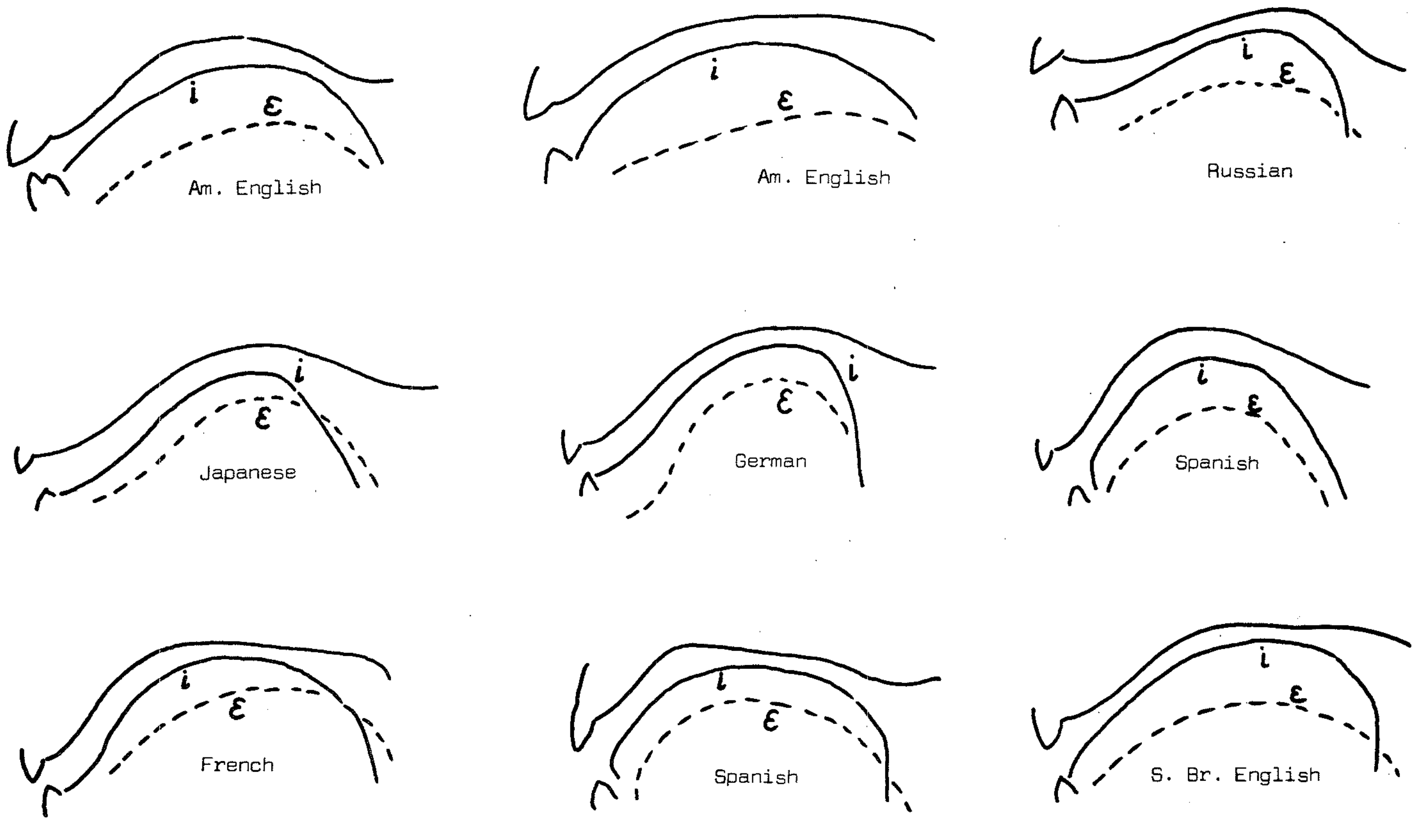
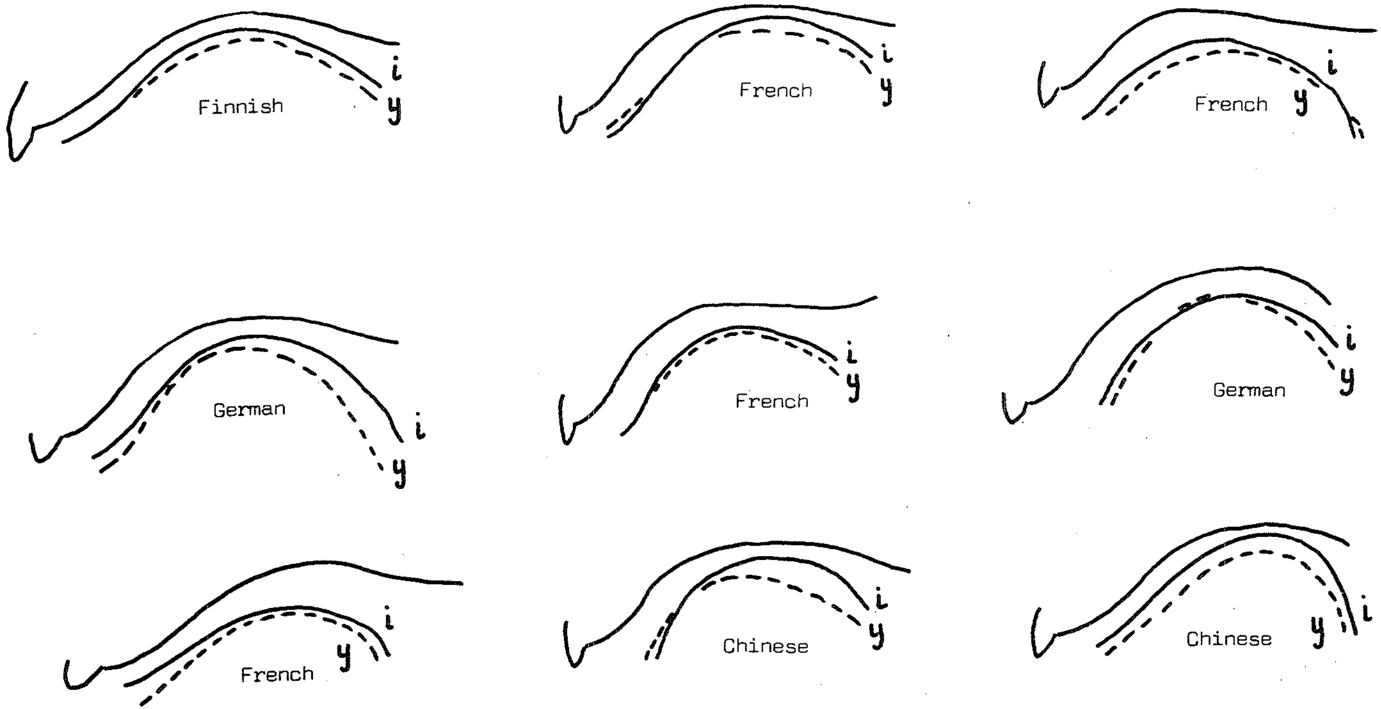


Fig. 1. A selection of [i] and [ε] profiles, illustrating how the tongue closely follows the individual shape of the hard palate (cf. also Fig. 2).



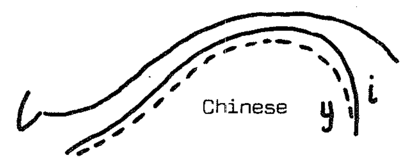


Fig. 2. A selection of [i] and [y] profiles, showing that the tongue is not more retracted for [y] (but it is slightly lower). These profiles also show how the tongue closely follows the shape of the hard palate (cf. Fig. 1).

of aim of the X-ray beam. After normalizing the scale of reproduction, I have hardly found this to be so. Differences between the relative sizes of the hard features on successive exposures are rarely larger than would be expected from simple random tracing errors. Distortion errors would not seem to be a major component of the total experimental error. On the other hand, tracings in some of the sets certainly cannot be superimposed exactly in the form published because their authors had used a different scale of reproduction for each separate picture.

I have used the vocal tract area function as a model for cavity configuration, the volume of a section of the tract being proportional to the cross-section area of that section. Distances across the vocal tract measured on the tracings have been transformed into cross-section areas according to two functions published by Sundberg (1969) for the palatal region and for the upper pharyngeal region. Sundberg's pharyngeal cross-distance/cross-area functions differ from others (Fant, 1960, Ladefoged et al., 1971). He argues that the side walls of the pharynx are drawn inwards when the cross-distance exceeds about 25 mm with the result that further sagittal widening of the pharynx produces a net reduction in the cross-area. The same procedure was followed by Lindblom and Sundberg (1971) except that this effect was not observed in the lower pharynx and they therefore used two functions for the pharyngeal region, one for above the epiglottis and one for the remainder. Following their example, I have also used a third function for the lower pharynx derived from data published by Fant (1960). The areas and lengths of the lip sections have been estimated with the help of the procedures and data given by Lindblom and Sundberg (1971).

Regarding the history of speech radiography and technical procedures, there are two comprehensive surveys, MacMillan and Keleman (1952) and Simon (1961). Standard sources of technical procedures for phoneticians in recent decades have been Subtelny et al. (1957) and Strenger (1968) while current cineradiographic techniques have been described by Moll (1960), Perkell (1969) and Kent (1972).

Tongue retraction

Fig. 1 contains tongue profiles for [i]-like and [ɛ]-like vowels from a selection of sets. The following features should be noted:

(i) There is wide variation of hard palate shapes between speakers from sharply domed to relatively flat.

(ii) The tongue of each speaker, irrespective of language, closely follows the contour of the hard palate for [i], leaving a very narrow passage (cross-section area about 0.5–1.0 cm^2).

(iii) Consequently, the top of the tongue arch may be further forward or further back for different speakers, depending on the shape of the hard palate.

In addition, the tongue profile for [ɛ] is also dominated by the contour of the palate, which still determines the location of the highest part of the hump. Essentially, as has been pointed out by Lindblom and Sundberg, the tongue profiles of [i] and [ɛ] for each speaker are very similar with reference to the mandible. Characteristic for [ɛ] is the wider channel along the palate (cross-section area about 3 cm^2). I conclude therefore that failure to consider the shape of the hard palate is a possible source of error that spuriously indicates "retraction" as a major difference when tracings for different subjects are being compared. See also Fig. 2 which contains a further selection of [i] profiles, this time related to [y].

The belief in several degrees of retraction has been further encouraged by incorrect articulatory interpretation of vowel spectra. It has been recognized for several decades that the traditional subjective judgment "retraction" was really based on auditory sensation and was related to the frequency of the second formant. Unfortunately, it has been too easy to assume the converse, that the frequency of the second formant will therefore reflect horizontal movement of the tongue (see, for example, Delattre 1951). The relationship between tongue movement and vowel spectrum is more complex. Fant pointed out that tongue <u>lowering</u> can cause F_2 to fall. This can be illustrated by an example from the published sets of X-ray tracings, the "tense-lax" quality difference between English or German /ī/ and /ɪ/ where the F_2 difference is some 300 or 400 Hz. The mandibular and lingual articulations of /ī, ɪ/ by Chiba and Kajiyama's German subject are given at Figs. 3 (a, d). These

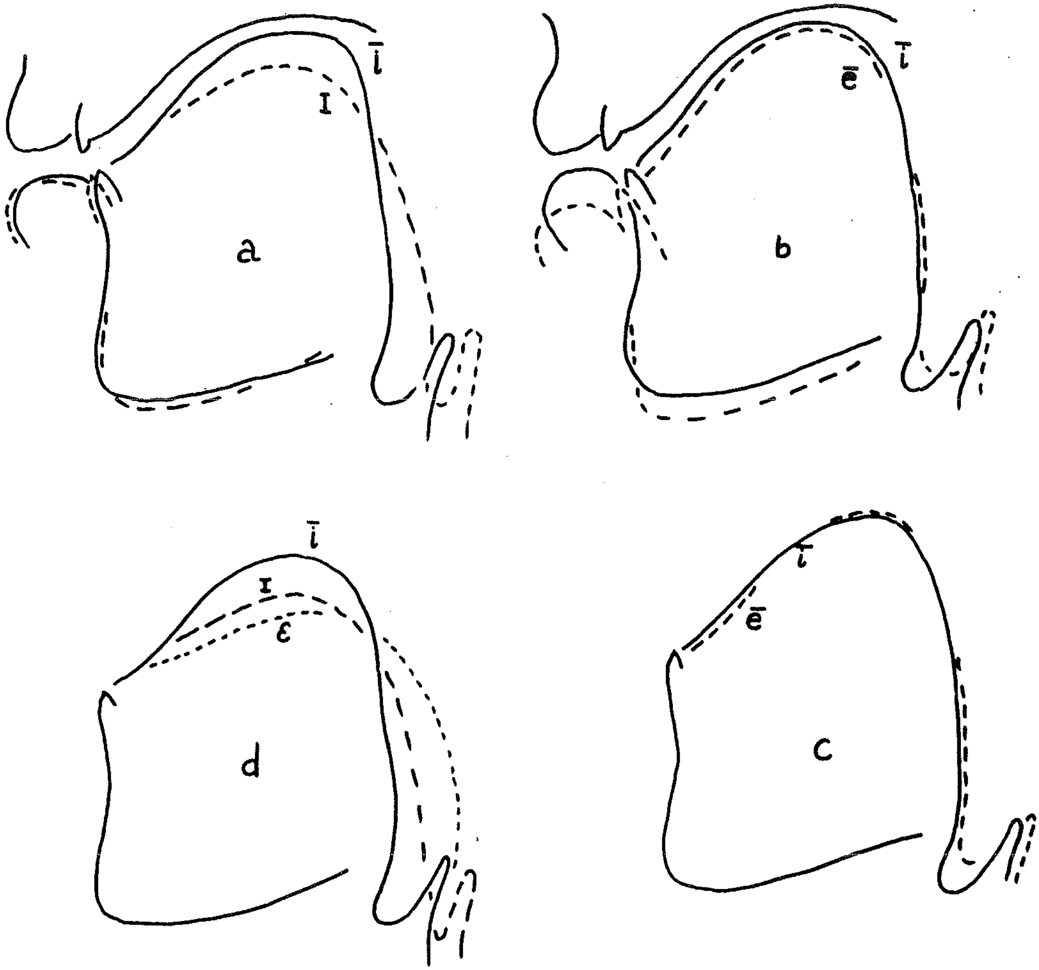


Fig. 3. Tongue height, jaw position and tongue shapes in German /ī, ɪ , ē, ɛ / from Chiba and Kajiyama's (1941) X-ray tracings.

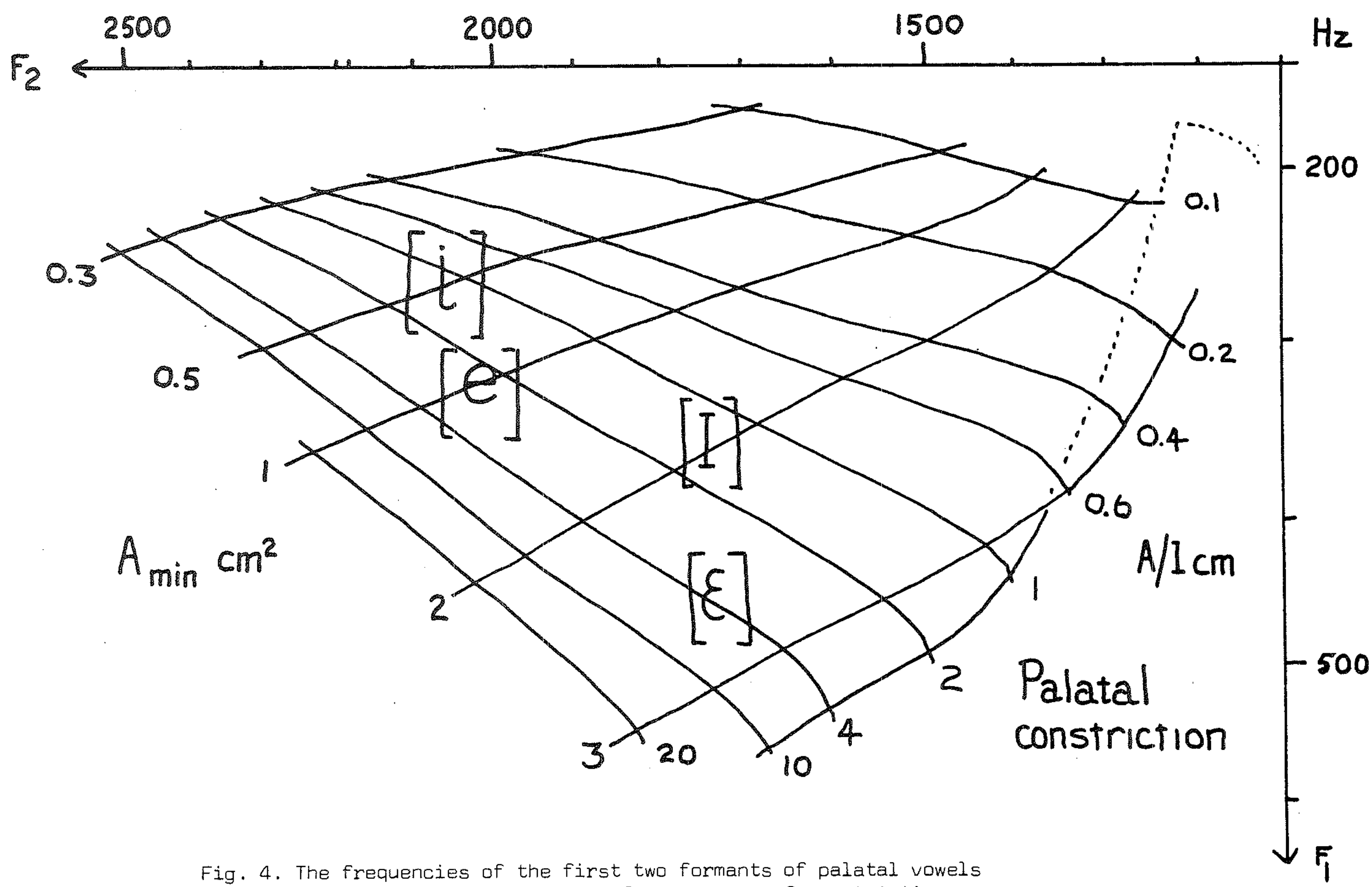


Fig. 4. The frequencies of the first two formants of palatal vowels for different combinations of the degree of constriction (A_{min} cm^2) and opening area/length ratio (A/1 cm). Based on Stevens and House's (1955) nomograms, for d_o = 12 cm.

show that the tongue was lower relative to the mandible for /ɪ/ than for /ī/, but that the jaws and lips were very similar.[3] Consequently, the mouth-opening was much the same for both vowels, but the constricted palatal passage was much wider for /ɪ/ and the pharynx much narrower. Reference to published nomograms such as those given by Stevens and House (1955) or Fant (1960) show that widening the palatal constriction from about 0.5 cm^2 for [i] to about 2.0 cm^2 for [ɪ], but keeping the same degree of mouth-opening, yields precisely the spectral difference between these vowels including the F2 difference of about 300 Hz. The Stevens and House nomograms have been redrawn at Fig. 4 for the palatal vowels[5]. The tracings at Fig. 3 do not indicate any tongue-arch retraction for /ɪ/, only lowering. Any of the [i] configurations at Figs. 1 and 2 can be transformed to an [ɪ] configuration by doubling the cross-section area of the palatal constriction from about 0.5 - 1.0 cm^2 to about 1.5 - 2.0 cm^2 while leaving the mouth-opening (jaw and lips) the same. The speaker does this by lowering the tongue about 3 mm with reference to the mandible. At the same time, lowering the tongue within the mandible causes the root of the tongue to narrow the lower pharynx. Both of these modifications, varying the degree of constriction at the hard palate and the volume of the lower pharynx, are relevant for the resonances of the vocal tract for these two vowels. Had the tongue been retracted instead of lowered, the constriction would have had to be withdrawn by as much as 2 cms to make F2 fall by 300 or 400 Hz, i.e. almost to the palatovelar location of [u]-like vowels. Stevens (1972) has pointed out that the plain palatal vowels are particularly insensitive to small variations of constriction location. It is just not acoustically profitable to make small tongue retractions for the palatal vowels. On the other hand, the nomograms show that very small variations of the degree of constriction yield relatively large spectral differences. Gunnilstam (1974) has underlined the role of varying the degree of constriction for producing large spectral differences.

I have also considered the traditional belief that the tongue arch is retracted slightly for rounded palatal vowels. Fig. 2 shows a selection of [i] and [y] tongue arches, none of which indicates such retraction. On the other hand they all show the tongue to be slightly lower for [y], irrespective of language. This difference can be entirely attributed to the mandible being slightly lower for the [y] renderings, the

[i] and [y] tongue profiles coinciding completely with reference to the mandible. Stevens's argument implies that tongue retraction for [y] would not contribute much to the spectral difference between [i] and [y], especially in comparison with the large spectral difference already obtained by rounding and protruding the lips.

There is one case where it is relevant to refer to gradual advancement or retraction of the tongue. This is for the difference between [ɑ, a, ,æ]-like vowels. The graver the low pharyngeal vowel, the further the tongue root is drawn into the lower pharynx to make the constriction even narrower. This is in fact the same parameter as for [i, ɪ], namely the degree of constriction at the narrowest part of the vocal tract. Owing to the 90° bend in the vocal tract, this parameter is varied by raising or lowering the tongue for the palatal constrictions but by advancing or retracting the tongue for pharyngeal constrictions. Carmody found the tongue positions of Holbook's two examples of American English /ʌ/ very variable, one falling right outside the tongue arch polygon, "behind o". These two cases are illustrated at Fig. 5 (b, c). The "lowest" vowel of all for subject Z was /o/ while the "position" of /ʌ/ was identical with /æ/. For H, the "position" of /ʌ/ was "higher" and "further back" than /ɔ/ (right behind /o/ as Carmody observed). Carmody hoped that this puzzling situation could be resolved by examining more radiographs. However, I shall demonstrate that these cases are only bewildering in relation to the tongue-arching model. The very same pair of X-ray sets can be given a very different interpretation that finds both examples very similar and typical not only for these vowels but for [a]-like vowels generally.

The area functions at Fig. 6 (a, b) show that the resonator configurations in both sets were very similar. All three vowels expressing /ɑ, ʌ, æ/ had the same low pharyngeal place of constriction at about 5 cms above the glottis. The main difference was in the degree of constriction, narrowest for /ɑ/ and widest for /æ/:

	/ɑ/	/ʌ/	/æ/
cross-section area at constriction	0.5 – 1.0 cm^2	1.5 – 2.0 cm^2	2.5 – 3.0 cm^2

This can be compared with the area functions of French /ɑ, a/ renderings at Fig. 6 (d), where /ɑ/ has a constriction of about 0.8 cm^2 and

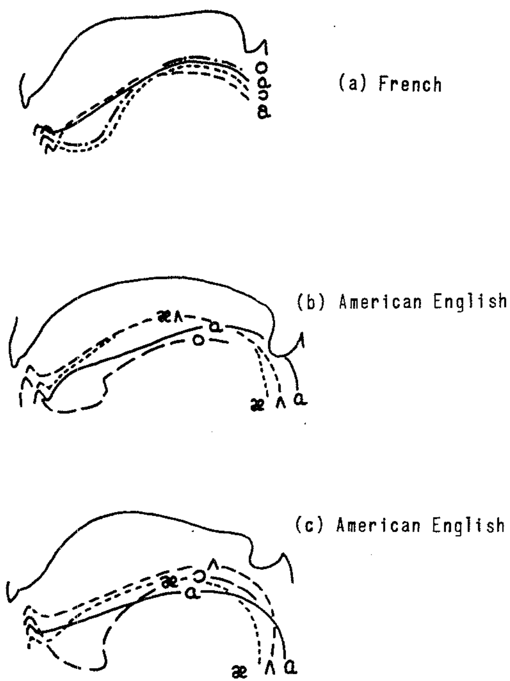


Fig. 5. Some confusing tongue heights: (a) mid and low back French vowels, and (b, c) English /ʌ/ related to some neighbouring vowels.

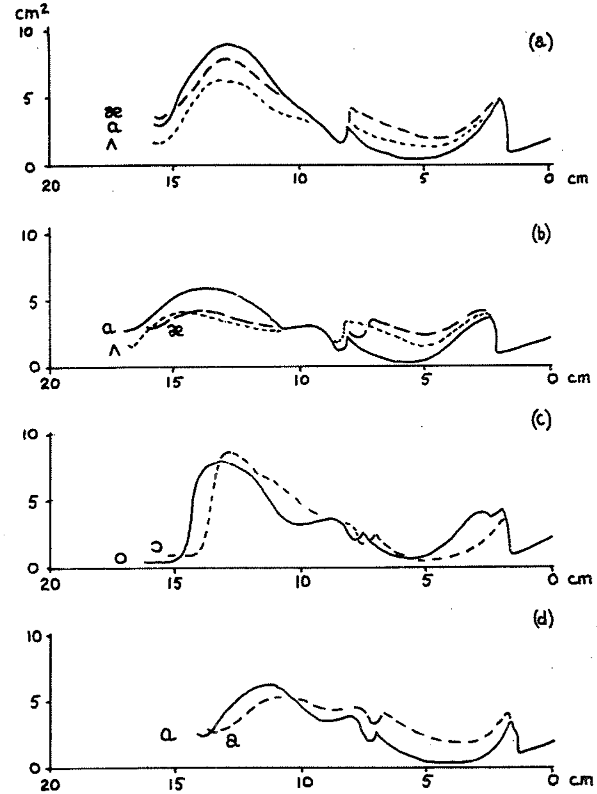


Fig. 6. (a, b) Area functions of the English vowels at Fig. 5 (b, c), and (c, d) area functions of the French vowels at Fig. 5(a).

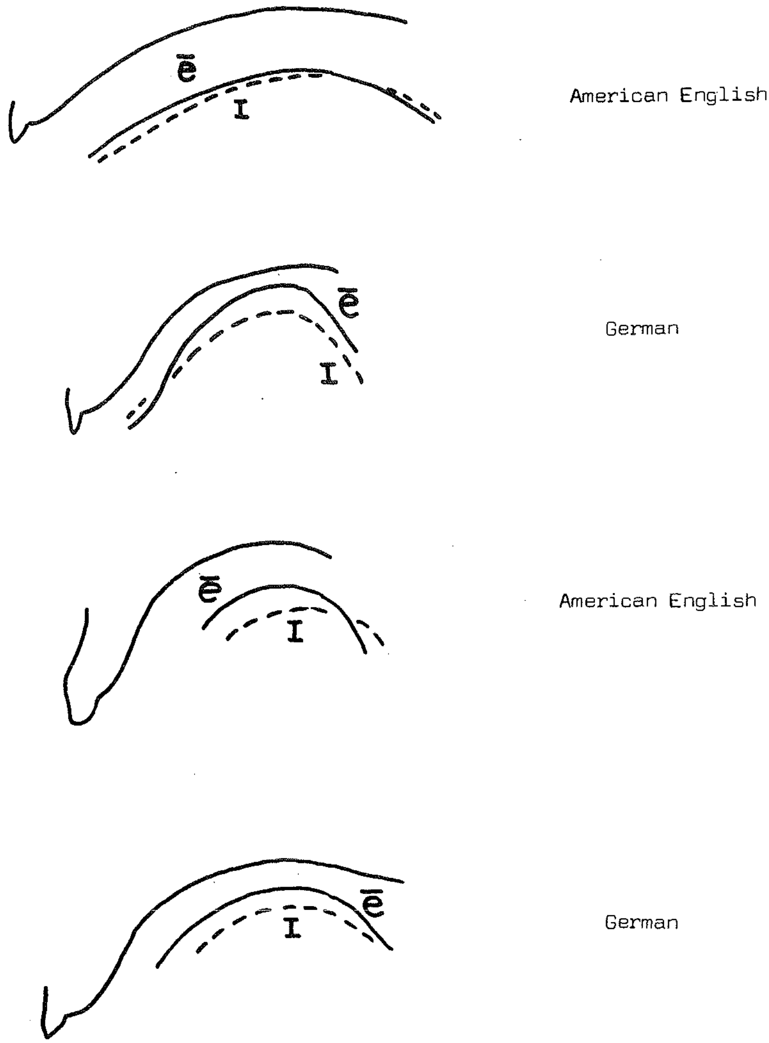


Fig. 7. Four cases of lax [ɪ] with lower tongue height than tense [ē].

/a/ about 2.0 cm^2. Clearly, the graver the [a]-like vowel, the narrower its pharyngeal constriction. The same conclusion for French [a]-like vowels was made by Mettas et al. (1971) after deducing the probable articulations for their observed spectra by referring them to Fant's nomograms. It has frequently been suggested that English /ʌ/ has an [a]-like quality, especially in Southern British English. But also Peterson and Barney's (1952) 76 American informants (men, women and children) all produced /ɑ, ʌ, æ/ with the highest first formant frequencies of all vowels (at least 600 Hz) and differentiated between them with the second formant frequency in three separate ranges between 1000 and 2000 Hz. The average F1 and F2 frequencies for the 33 men in that group were:

	/ɑ/	/ʌ/	/æ/
F1 (Hz)	730	640	660
F2 (Hz)	1030	1190	1720

These Peterson and Barney /ʌ/ spectra are certainly [a]-like.[6]

By comparing the vocal tract configurations of these vowels, and especially the place and degree of constriction, I have shown that Carmody's supposedly variable and inconclusive /ʌ/ renderings were in fact very similar and had the same resonator characteristics. This example, together with the comparison with the French [ɑ - a]-like vowels from one of Holbrook's French sets confirms the relevance of tongue body advancement and retraction for [ɑ, a, æ]-like vowels. But the relevant factor for shaping the resonating cavities is not the tongue arch position but the width of the constricted lower pharynx.

Tongue height

Meyer's and Russel's criticisms had mainly concerned tongue height, especially that the tongue was lower for [ɪ] than for [e] and lower for [ɔ] than for [a]. In addition my collected material also contained examples of [o] lower than [a] and confusion of the heights of [ʊ] and [o].

[ɪ] - [e]

This case arises in languages with quality contrasts /i - ɪ, e - ɛ/. In the X-ray sets I have collected, this applies for American English, and German. Southern British English dialects have quality contrasts for three of the monophthongs /i – ɪ – ɛ/ but a diphthong monophthong contrast /ɛj – ɛ/ is used for raid–red. Swedish has mainly quantity contrasts among the plain palatal vowels /i: – i, e:, ɛ: – ɛ/ (Elert, 1964) while other pairs of vowels may also have quality differences (Hadding-Koch and Abrahamson, 1966). There is general agreement that there is a quantity contrast for Swedish /i: – i/ with little quality difference, although there may be some variation between dialects.

Four examples from American English and German are illustrated at Fig. 7. All showing higher tongue arch for /ē/ than for /ɪ/. In all 7 possible sets for these languages, /ē/ was "higher" than /ɪ/ (other sets from these languages did not contain examples of both vowels). The higher tongue arch for /ē/ than for /ɪ/ was also reported for 5 out of 6 subjects by Ladefoged et al. (1972b) in a cineradiographic study of American English speech.

An early criticism of this observed "height inversion" was that tongue articulation was distorted by the use of chains for emphasizing the tongue outline or by the unnaturally sustained or repeated renderings of vowels necessitated by long exposure times, but the same result is still found when presentday cineradiographic techniques are used. The outlines of soft tissues are nowadays enhanced by applying a bismuth or barium compound to the articulators, and electronic intensification of the image makes possible very brief exposure times (50 to 200 frames/sec with only a few milliseconds radiation per frame).

Strenger's Swedish profiles show a higher tongue for short /i/ than for long /e:/ but this is to be expected if the subject had the non-qualitative /i: – i/ quantity contrast, so that this case is not necessarily an exception to the reported anomaly. On the other hand, Meyer's plastopalatograms from a Swedish subject (1910) had shown /e:/ to be higher.

The failure of the tongue-arching model to get the heights of [ɪ] and [e] right could of course be looked upon as an easily rectifiable mistake. These vowels need only be put in their correct places. But then the affinity of [i] with [ɪ] and of [e] with [ɛ] would be lost (Jespersen's objection). In either case, the model would still fail to capture the true articulatory relationship between "tense" and "lax" palatal vowels. I shall refer once again to the German example at

Fig. 3 to demonstrate how it is possible for "half-close tense" [e] to be "higher" than "close lax" [ɪ].[3]

Fig. 3 (a) shows only very slight mandibular difference between /ī/ and /ɪ/ and virtually the same lip separation, which means that the mouth-opening and hence the radiation were much the same for both renderings. The main difference between them is that the tongue is considerably lower for /ɪ/ than for /ī/, widening the palatal constriction and bulging into the pharyngeal cavity.

Fig. 3 (b) shows that the mandible was lowered much more for /ē/ than for /ī/ while the palatal passage was only slightly widened and there was consequently only a little bulging movement back into the pharynx. In fact, /ī/ and /ē/ have very nearly the same "tongue height", similar to what Kruisinger (1925) had noted on Meyer's radiograms.

Fig. 3 (c) shows that /ī/ and /ē/ had the same tongue shape relative to the mandible. Fig. 3 (d) shows that /ɪ/ and /ɛ/ had the same tongue shape relative to the mandible, both decidedly lower than for /ī, ē/.

The "tense" vowels /ī, ē/ were thus differentiated from the "lax" vowels /ɪ, ɛ/ by the height of the tongue within the mandible, while the "close" vowels /ī, ɪ/ were differentiated from the "open" vowels /ē, ɛ/ by the degree of mandibular depression. The component gestures shaping the vocal tract for these vowels are thus as follows:

PALATAL VOWELS:

		TONGUE IN JAW: HIGHER	TONGUE IN JAW: LOWER
JAW	HIGHER	i	ɪ
	LOWER	e	ɛ

In the terms of the tongue-arching model, the tongue is "more central" for "lax" vowels than for "tense" vowels. Jakobson and Halle (1964) quote several examples expressing this view and conclude "tense phonemes are produced with more deviation from the neutral, central position than the corresponding lax phonemes". For a case similar to that described above, the English /ī – ɪ/ opposition, D. Jones (1967: § 160) wrote that the tongue was "lowered" and "retracted" for /ɪ/ with respect to /i/. Fig. 3 confirms the lower tongue but it does not support the notion of retraction. The relevant difference is in the respective mandibular and lingual components as outlined above, which in turn determine the degree of mouth-opening, the degree of palatal constriction and the size of the pharyngeal cavity.

Fig. 3 shows how the mandibular and lingual movements combine in opposite directions for /ɪ/ and /ē/ (high jaw and low tongue versus low jaw and high tongue respectively). The difference is sufficient for /ē/ to come out "higher" than /ɪ/, a simple explanation for what has hitherto appeared to be a perplexing anomaly in the terms of the tongue-arching model.

When these articulations are referred to the Stevens and House nomograms (Fig. 4), basing the parameter values on the vocal tract area function for each vowel[5], the following approximate frequencies are found for the first two formants:

	F1	F2
/ī/	250 - 300 Hz	2000 - 2100 Hz
/ē/	300 - 350 Hz	1950 - 2050 Hz
/ɪ/	325 - 375 Hz	1700 - 1800 Hz
/ɛ/	425 - 475 Hz	1650 - 1750 Hz

These are not far from what we might expect to find on spectrograms of adult male speech.

The result of the mandibular change from /ī/ to /ē/ (doubling the mouth-opening and widening the constricted palatal passage a little) appears mainly to result in a rise in F1 while F2 falls only slightly. The result of the lingual change from /ī/ to /ɪ/ (leaving the mouth-opening unchanged but widening the constricted palatal passage) is a simultaneous raising of F1 and lowering of F2. In the latter case the spectrum is "centralized" towards 500, 1500, 2500 ... etc. Hz, but without corresponding articulatory "centralization".

The separate mandibular and lingual differences between /ī/ - /ē/ and /ī/ - /ɪ/ respectively are thus compatible with the spectral character of these vowels. I realize that nothing has been proved by describing one example, although it is typical of the whole material. I have demonstrated that a pair of tongue heights that have been puzzling in traditional terms can be given an interpretation that is intimately related to the physiology and acoustics of vowel production, an impossibility within the framework of the tongue-arching model.

Ladefoged et al. (1972a) are sceptical of the type of solution outlined above, on the basis of their factor analysis of forces presumed to be acting on tongue shape (or, more precisely, the displacement of specified points along the dorsum of the tongue). They found individuality between six subjects in the way they utilized and coordinated mandibular and lingual movement. However, they record that three of the six had "a very bunched, tense, shape of the tongue in heed and hayed, and a flatter, lax, shape in hid, head and had", which is the same as that illustrated at Fig. 3. I shall do no more here than underline that five of their six subjects had /ē/ higher than /ɪ/ and that three of the six agreed with the case described above regarding different lingual gestures for "tense" and "lax" vowels, while the remaining three subjects disagreed both with that pattern and with each other[7].

[ɔ] - [ɑ, a] and [o] - [ɑ, a]

The second situation, conflicting tongue heights for [ɔ] and [ɑ, a] is expected to occur where there are two qualitatively different [o]-like sounds, whether the difference is distinctive (as in English) or allophonic (as in Spanish). It is the "lower" vowel [ɔ] that has been reported with tongue arch lower than for [a]. Fig. 8 shows two examples of [ɔ] lower than [a] or [ɑ], 4 examples of [o] lower than [a, ɑ] (unexpected) and 2 examples with both [o - ɔ] lower than [ɑ, a], one of them with [o] lower than [ɔ] (quite unforeseen and in complete contradiction to the tongue-arching model). The relative "heights" of these vowels in all the sets collected from the literature are compared at Fig. 9. Of the 22 sets where this comparison was possible, 8 had [ɔ] higher than [ɑ, a], 8 about the same height and 6 lower. In only one third of the possible sets was [ɔ] definitely higher than [ɑ, a] in accordance with the model. In addition, 6 had [o] lower than [ɑ, a] and 6 almost the same as [ɑ, a]. In two thirds of the possible sets, [o] was higher than [ɑ, a] in accordance with the model.

Notwithstanding the random character of the tongue "heights" of [o]-like and [a]-like vowels, it is interesting to discover that when the area functions for these vowels are compared, the rounded [o - ɔ]-like

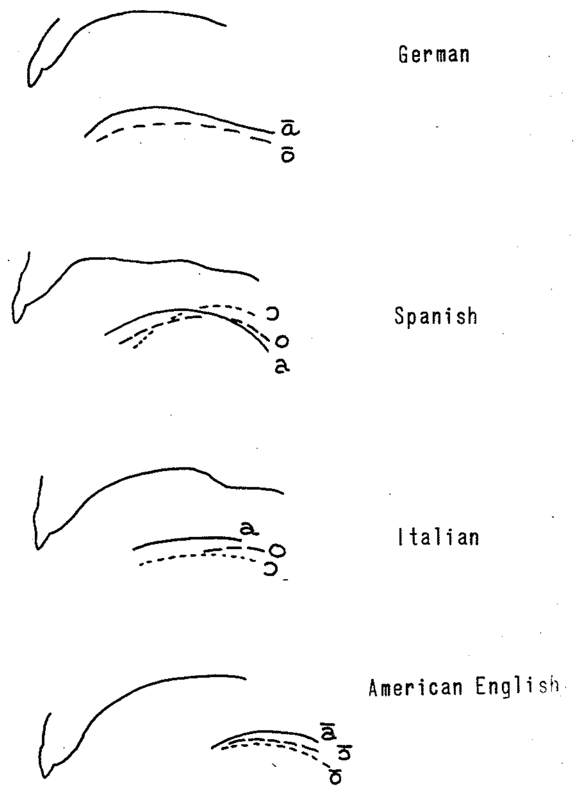


Fig. 8. Four cases of confused tongue heights of mid and low back vowels.

Relative tongue heights, [o,ɔ]-like and [ɑ, a]-like vowels

	[ɔ] higher than [ɑ, a]	[ɔ] same as [ɑ, a]	[ɔ] lower than [ɑ, a]	[o] same as [ɑ, a]	[o] lower than [ɑ, a]
Czeck	1			1	
English (Am)[8]	1	1	1		1
English (S Br)[8]	1	1			
French		3	1		
German	2		1		1
Icelandic	1				
Italian			1		1
Japanese					1
Korean	1			1	
Polish			1		
Portuguese		1			
Russian			1	2	1
Spanish		2		1	1
Swedish	1			1	
	8	8	6	6	6
Total compared		22		34	

Fig. 9. Comparison of tongue heights of [o, ɔ] -like vowels from a large number of sets of published X-ray tracings.

vowels always have a place of constriction a little farther from larynx than the [ɑ - a]-like vowels. I shall illustrate one case, Holbrook's French subject C (female). Fig. 5 (a) shows that the tongue "heights" for her "back" vowels were ranked [o, ɑ, ɔ, a] from "higher" to "lower". Fig. 6 (c, d) shows that the /o/ and /ɔ/ renderings constricted the pharynx at a distance between 5 and 6 cms above the glottis, while the /ɑ/ and /a/ renderings constricted the pharynx between 4 and 5 cms above the glottis. Whatever the orders of tongue heights in sets of vowels, examination of all the sets indicates that this relation of constriction locations for the two types of vowel is a very strong constant of speech[3].

[ʊ] - [o]

I also found several instances of [ʊ] lower than [o]. The only languages where this might be expected are those with clear quality contrasts for /ū - ʊ/, that is, English and German in this material. There were 7 sets containing the [ʊ - o] pair among the 38. Of these, 3 had [ʊ] higher than [o] (1 German, 1 American English, 1 British English), 1 the same (1 German) and 3 with [ʊ] lower than [o] (1 German, 1 American English, 1 British English). This suggests the distribution for this pair is random rather than language specific.

THE WEAKNESSES OF THE TONGUE-ARCHING MODEL

The comparison of published X-ray tracings in the preceding section confirms the anomalous tongue positions that were said to contradict the tongue-arching model. "Close" [ɪ] is more "open" than "half-close" [e]. The "heights" of "half-open" [ɔ] and "open" [a] are random. In only two-thirds of the cases was the tongue "higher" for "half-close" [o] than for "open" [a]. The concept of gradual retraction is without foundation. Tongue arch position in terms of "height" and "retraction" is ambiguous with regard to resonator shaping and consequently to the spectrum of the vowel generated in the vocal tract. The vocal tract configuration is dependent on a number of other factors, information on which is not generally available in vowel descriptions based on the tongue-arching model since they are external to it and customarily disregarded. The advantage of the tongue-arching model was that it had seemed to offer 19th century phoneticians better possibilities for describing finer shades of vowel quality than could be generated by the ancient model it displaced. Its weaknesses are related to the fact that it was a product of the imagination that was never confirmed in serious tests. It is physiologically unsound since it was based on a misconceived notion of tongue articulation for vowels. Consequently it fails to predict the values of its parameters correctly for many vowels. The ambiguity of the relationship between the values of its parameters, physiological activity, resonator configuration and spectral output means that it is powerless to explain central areas of speech production.

Physiological weaknesses

There are two serious physiological weaknesses.

Firstly, the model neglects the pharynx completely. The earliest radiograms had shown the low pharyngeal constriction for [a]-like vowels, and its significance was underlined by Russel. Carmody (1941) made a detailed analysis of the pharyngeal cavity from Holbrook's radiograms, but he did not try to relate his findings to the tongue-arching model, in which he remained a firm believer. More recently, the pharyngeal cavity has been explored by tomography (Fant, 1960, 1964) and ultrasound (Minifie et al. 1970). Delattre (1971) has studied pharyngeal articulations with cineradiography. Ladefoged et al. (1971) have made casts of the living pharynx and Lindqvist and Sundberg (1971) have inspected the pharynx with a fibrescope. The shaping and acoustical significance of the pharyngeal cavity are outside the domain of the tongue-arching model. although the tongue root position proposal (Stewart 1967, Halle and Stevens 1969, Perkell 1971, Lindau et al. 1972) is an attempt to relate the difference between tense and lax vowels to the volume of the lower pharynx. The meaning of such supplementary concepts as "uvularization" and "pharyngealization" is not clear. The extrinsic muscles of the tongue (which are generally held to be mainly responsible for tongue shape and position in vowels) all contract in the pharyngeal region and whatever task these muscles may otherwise be performing they always immediately and directly alter the pharyngeal cavity. Three pairs of muscles contract in the lower or mid pharyngeal region - the hyoglossi, the posterior genioglossi and the glossopharyngei. The fourth pair, the styloglossi, contract across the upper pharynx. Nothing of this is captured by the tongue-arching model, despite its supposedly physiological basis.

The second major physiological weakness is that the location of the tongue arch cannot readily be related to knowledge of the state of the tongue muscles. The ancients had only been hampered by their insufficient knowledge. The celebrated Persian physician and philospher Ibn Sīnā (1000), better known to mediaevel Europeans as Avicenna, had made a detailed description of the muscular structure of the tongue, but had to admit failure in his attempt to relate it to tongue movement during vowels. This was probably due to the fact that either he, or Galen whom he may have been quoting, had dissected the tongue of the ape and not that of man (Singer, 1957: p. 53)[9]. But since at least the treatise of Hellwag (1781)[10], there has been virtual agreement about the role of the extrinsic muscles of the tongue for directing lingual gestures to form constrictions in the vocal tract. The presentday view is given by, for example, MacNeilage and Sholes (1964), Zemlin (1968: p. 281), Harris (1971), Raphael (1971a, 1971b), Smith (1971). The hyoglossi draw the tongue bodily downwards to narrow the lower pharynx. The posterior genioglossi pull the tongue root forward to widen the pharynx and assist in raising the body of the tongue towards the palate. The glossopharyngei (fibres of the pharyngeal constrictors that insert into the sides of the tongue) draw the tongue back into the mid-pharynx. The styloglossi draw the tongue upwards and rearwards towards the soft palate. The effect of contracting these muscles, alone or in combination, is to narrow different regions of the vocal tract, controlling the location of the constriction and the volumes of the cavities. The amount of contraction, together with the movement of the mandible, controls the degree of constriction. A type of model based on constriction location is compatible with observable motor activity. But specific muscular activity is not unambiguously and exclusively related to the raising or lowering, advancement or retraction of the tongue-arch, so that the tongue-arching model constitutes a very weak link between neuromotor activity and articulation. This means it provides a bewildering articulatory framework, not least for electromyographic investigations of tongue movement.

Predictive capability

One aspect of the weak predictive capability of the tongue-arching model has already been described. The tongue positions that can be observed in speech are not always the same as those prescribed by the model. The analysis of Carmody's difficulties with "low back" vowels and the English /ʌ/ profiles showed that while the coordinates defining "tongue position" are largely irrelevant for the articulation of vowels, a type of model based on the vocal tract configuration did not find these profiles in any way enigmatic. It would not therefore be sufficient simply to rectify the location of the errant vowels in the polygon by assigning the "correct" coordinates. It would still be impossible to predict the resonator configuration satisfactorily, and hence the spectrum, from the coordinates. It would similarly remain impossible to predict the underlying motor activity.

Explanatory power

In view of its unsound physiological foundation and ambiguous relation to vocal tract shaping and resonance properties, and its consequently unsatisfactory predictive capability, the tongue-arching model failed to provide a smooth and direct link between articulation and acoustics. It could not therefore explain the relationships between the successive links of the speech chain, the systematic preferences for the structure of vowel systems, the phonetic processes involved in sound changes and so on. It is not surprising that the esteem of articulation fell when compared with the progress made in speech acoustics. Advances in the analysis of the acoustic structure of speech and in psycholinguistics have made it possible to elucidate much of the role played by acoustic cues in perception. Acoustic contrast has been accorded a firm position in speech theory[11]. But the bewildering relationship between the parameter values of the tongue-arching model and the spectral character of vowels has prevented the construction of a comprehensive view embracing and integrating all phases of vowel production.

AN ALTERNATIVE ARTICULATORY MODEL

Spectrographic analysis over the past few decades has demonstrated that the spectral character of vowels is relatively constant, especially for the same speaker (for example, Joos, 1948; Potter and Steinberg, 1950; Peterson and Barney, 1952) confirming the isolated examples of spectral analysis published in previous years (Malmberg, 1952: pp 89-97). Differences of formant frequency range between speakers due to differences of vocal tract scale are regular and predictable. Spectral variations within the same spaker's speech are regular and can be related to such factors as consonant environment, degree of stress, style or temporal constraints (Tiffany, 1959; Stevens and House, 1963; Lindblom, 1963). The relative spectral contrasts utilized for phonemic distinctions are universal (Jakobson et al. 1952). Contrasting this spectral constancy with the apparent variability and confusion of tongue articulation and knowing that there is theoretically an infinite number of possible resonator shapes for a given spectrum, it was natural that many phoneticians preferred to believe that there was no constancy at all in articulation and that the speaker's only concern was to produce the correct spectrum. For example, Malmberg (1952:99) has written "on peut changer un [e] en un [ø] en arrondissant les lèvres. Mais on peut produire à peu près le même effet en retirant un peu la langue. Les deux procédés amènent un abaissement du formant haut de la voyelle ... C'est par cette différence dans la structure acoustique, et non pas par la position des organes, que le [ø] se distingue du [e]".

However, the examples discussed indicate that the speaker is nevertheless striving to create a constant vocal tract configuration for a given vowel, thereby confirming constancy in two adjacent links of the speech chain - resonator shape as well as spectrum. Irrespective of language, the [o - ɔ]-like vowels always have a tongue constriction a little higher in the pharynx than the [ɑ - a]-like vowels. Similarly, all [ɑ - a - æ]-like vowels are produced by constricting the lower pharynx, the degree of gravity being related to the degree of constriction.[3] If the speaker is, as it seems, always striving to produce one constant configuration for a vowel, then it is also reasonable to look for constancy in the manner of forming these configurations. Is there, for example, a simple set of underlying gestures that are combined in various ways to achieve the desired configurations? The preceding discussion concerning the "tense-lax" palatal vowels [i, ɪ, e, ɛ] indicated that mandibular and lingual gestures are combined in different ways for these vowels. Further, the similarity of the [i, ɛ, y] profiles compared at Figs. 1 and 2 adds further strength to the notion that these speakers have produced comparable cavity configurations by using the same means.

There has been a growing tendency in recent years to look in this direction for a substitute articulatory model. Stevens and House (1955) found that "X-ray studies indicate that during the articulation of vowels the dimensions along the vocal tract are controlled primarily by the position of the tongue constriction and by the degree of constriction". Kaneko (1957) has compared the American English and Japanese vowel systems with reference to the vocal tract configurations. Lindblom and Sundberg have described the Swedish vowel system (1969a) and the cardinal vowels (1969b) using the place and degree of constriction to define the place of articulation. I have myself described the West Greenlandic vowel system (1971) using the place of constriction as a phonological feature that can be used in generative rules[3]. Pétursson (1974) has described the Icelandic vowels with reference to the constriction location. Lindblom and Sundberg (1971), simulating physiological factors that determine the vocal tract area function, have explored and described the spectral consequences of individual mandible, tongue and lip movements. Stevens and House (1955) noted the controversy regarding articulation and vowel diagrams, but did not wish at that moment to suggest "that the present data validate any theory of static positions for vowel production". However, the preceding discussion concerning palatal and pharyngeal vowels indicated that the place and degree of constriction describe more relevant and constant differences between vowel articulations than did the parameters of the tongue-arching model. The features of cavity configuration, and their manner of formation can provide the basis for an alternative description of vowel articulation, as a substitute for the tongue-arching model.

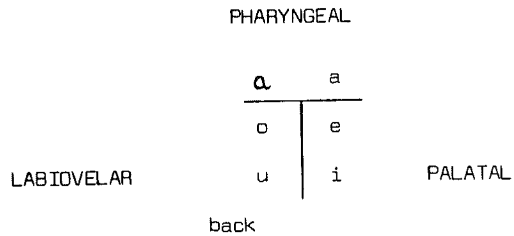
SOME IMPLICATIONS FOR PHONOLOGY

The difficulties arising from the irrelevancies of the tongue-arching model for vocal tract shaping are acutely felt today when so many phoneticians wish to model the vocal tract and simultaneously discuss the articulation, acoustics and phonological relationships of vowels, or ponder the suitability and phonetic meaning of phonological features (for example Ladefoged 1964; Lindblom and Sundberg 1969a, 1969b, 1971; Perkel 1971; Lieberman and Crelin 1971; Ladefoged et al. 1971a, 1972b; Lindau et al. 1972; Lindblom 1972; Stevens 1972).

Phoneticians of the acoustical school, recognizing the fallacies of the tongue-arching model, had aimed instead to use the spectral characteristics of speech segments as descriptors and features when analysing phonemic systems or when dealing with problems of phonology. The best known scheme of this kind was that of Jakobson et al. This was repeated by Halle (1964) and Jakobson and Halle (1968), and was widely used for nearly two decades. But the acoustically oriented basis has given way to articulation again, while the function of features has undergone revision. In particular, McCawley (1967) pointed out certain inadequacies arising from the different roles played by feature systems in Preliminaries to Speech Analysis and in Halle's Sound Pattern of Russian of 1959. Jakobson had amphasized the contrastive function of the features for denoting phonemic distinctions, whereas Halle was using the features for the complete systematic specification of segments necessary for the ordered rules of a generative phonology. McCawley found that Jakobson's desire for the set to be minimal, achieved for example by subsuming the spectral characteristics of both lip-rounding and pharyngealization under the one feature flat, meant that the set was too small for a generative phonologist and made it impossible for him to distinguish the very different processes involved in for example labial and pharyngeal assimilations. The enormous expansion of the set of features, the shift to articulation and the new role of features can be seen in Chomsky and Halle (1968: chapt. 7) where it is explained that "the totality of phonetic features can be said to represent the speech-producing capabilities of the human vocal apparatus".

Following the renewed focus on articulation, there is a grave risk that the tongue-arching model will become even more firmly entrenched. Chomsky and Halle's three features - high, low, back - denote six positions of the tongue body (no longer the tongue arch) and indicate that the tongue is raised, lowered or retracted from an arbitrary origin, the position for [ɛ][12]. This small set of features avoids the erroneous gradual retraction concept of the tongue-arching model proper. It is possible to translate the feature specifications of the small primary set of vowels generated by this framework into vocal cavity configurations, although the procedure is complex. This can be done because the underlying arrangement of the vowel scheme bears a closer affinity to the ancient pharynx-velum-palate-apperture typer of model (which reflects cavity configuration) than to the tongue-arching model. Phoneticians in the 19th century occasionally split the pharyngeal

[a]-like vowels and divided them between the palatal and labiovelar series, thus:



From this follows the fact that the only unique relation between the parameters of the two arrangements is that low vowels have a low pharyngeal constriction and non-low vowels do not.

The six vowels are as follows (including the redundant labialization of non-low back vowels [u, o]):

	ɑ	a	i	ɛ	u	o	
high	–	–	+	–	+	–	
back	+	–	–	–	+	+	(I)
low	+	+	–	–	–	–	
labial	–	–	–	–	+	+	

The discarded information about vocal tract configurations can be filled in from general phonetic knowledge of the articulation of this small set of very frequent vowels. One possible arrangement is as follows:

(i) Constriction location: [i, ɛ] have palatal constrictions, [u] a palatovelar constriction, [o] a pharyngovelar constriction and [ɑ, a] a low pharyngeal constriction.

(ii) Degree of narrowing: [ɛ, a] have wider constrictions, [i, u, o, ɑ] have narrower constrictions.

(iii) Mouth-opening: [ɛ, o] have wider mouth-openings relative to [i, u] respectively, and [ɑ, a] have wide openings.

(iv) Lip-rounding: [u, o] are labial

From this information, a new matrix can be constructed (where the features palatal, velar, pharyngeal define constriction locations):

	ɑ	a	i	ɛ	u	o
palatal	−	−	+	+	+	−
velar	−	−	−	−	+	+
pharyngeal	+	+	−	−	−	+
narrow	+	−	+	−	+	+
labial	−	−	−	−	+	+
open	+	+	−	+	−	+

(II)

There is no reciprocal one-to-one relation between the features and values of (I) and those of (II), except that low vowels have their constriction low in the pharynx. High vowels [i, u] are non-pharyngeal, but it does not follow that non-high vowels are pharyngeal ([ɛ] is palatal). Not all back vowels are non-palatal ([u] has a palatovelar constriction), nor are all non-back vowels palatal ([a] has a pharyngeal constriction).

The features and values of matrix (I) can be transformed to those of matrix (II) by a conditional statement such as (III):

$$
[\beta\,\text{high}] \longrightarrow
\left\langle
\begin{array}{ll}
\beta & \left\langle \begin{bmatrix} +\text{palatal} \\ \alpha\,\text{velar} \\ -\text{pharyngeal} \\ -\text{open} \\ +\text{narrow} \end{bmatrix} \right\rangle \;/\; \begin{bmatrix} \overline{} \\ \alpha\,\text{back} \end{bmatrix} \quad \text{(IIIa)} \\[2ex]
-\beta & \left\langle \begin{array}{ll}
-\gamma & \left\langle \begin{bmatrix} -\alpha\,\text{palatal} \\ \alpha\,\text{velar} \\ \alpha\,\text{pharyngeal} \\ +\text{open} \\ \alpha\,\text{narrow} \end{bmatrix} \right\rangle \quad \text{(IIIb)} \\[2ex]
\gamma & \left\langle \begin{bmatrix} -\text{palatal} \\ -\text{velar} \\ +\text{pharyngeal} \\ +\text{open} \\ \alpha\,\text{narrow} \end{bmatrix} \right\rangle \quad \text{(IIIc)}
\end{array} \right. \;/\; \begin{bmatrix} \overline{} \\ \alpha\,\text{back} \\ \gamma\,\text{low} \end{bmatrix}
\end{array}
\right\rangle
$$

Chomsky and Halle's three tongue body features (I) can only provide 2 x 3 positions. If a need is felt for more position categories, extra features have to be provided – for example front and mid which will permit 3 x 4 tongue positions. An even more formidable statement than

(III) will be needed to derive vocal tract configurations, if it were at all possible still to do so. Worst of all, the addition of front would reintroduce the error of the central tongue position. Further, apart from the confusion and ombiguity of tongue height, there has never been agreement about the features needed to generate four degrees of opening for vowels[13]. The Chomsky-Halle arrangement does not therefore avoid the weaknesses of the tongue-arching model when treating vowel systems requiring more tongue positions than the basic six. They share the situation of the early 19th century users of the ancient type of model - the number of parameters available is insufficient to generate the number of vowel categories observed in more complex systems. Moreover, if every available possible feature combination is utilized to provide pigeon holes for difficult vowels, disregarding physiological and acoustic data, unrealistic solutions will result. This is the course resorted to by Chomsky and Halle when they pair off English /a/ and /ʌ/ with /ɔ/ and /o/ respectively as non-labial members of low and intermediate tongue height categories. It was demonstrated above that with regard to vocal tract shaping, English /ʌ , a/ share the lower pharyngeal constriction location, and /o, ɔ/ the higher pharyngovelar constriction location.

The translation of (I) into (II) by (III) seems to be a very clumsy necessity to have to go thorugh before the phonological component can yield its output in a form that is related to sound quality via spectral character and resonator configuration. Yet it is the ability of a model to relate phenomena at the separate links of the speech chain that sharpens the predictive capability and increases the explanatory power of a theory for phonology. Matrix (I) will generate 6 letters of a phonetic alphabet. If we stop there, the phonology output will be a phonetic transcription where each letter stands for a set of feature specifications. This was Chomsky and Halle's goal[14]. But the goal can be constrained even further, to make the phonological component deliver its output in a form compatible with current speech production theory. Fant (1969) has written: "Before we can accomplish the happy marriage between phonology and phonetics we have to work out the rules for predicting the speech event given the output of the phonological component of the grammar. To me this is the central, though much neglected, problem of phonetics."

It is well known that opinions differ between linguists as to whether or not generative rules have psychological reality. At least phoneticians can demand that when rules are written in a set of features that have an articulatory basis, the standard of physiological accuracy should be set high. Critics such as Fant (1969) and Vennemann and Ladefoged (1971) accept the linguist's need for abstract classificatory or "cover" features alongside strictly phonetic features. One can always imagine a feature interpreting component that will clothe the features of the phonology output with the appropriate phonetic character. Statement (III) above would be part of such a component. The problem is whether a feature interpreting component is always necessary, and if it is then what form it should have and where and how smoothly it should operate. With respect to the present specific issue, the movement of the tongue and the shaping of the vocal tract for vowels, I believe it would be more suitable to write the phonology straight away in features similar to those of (II) rather than switch terminology and conceptual framework halfway by translating the present features (I) into features shaping the vocal tract (II) with some heavy interpretive device such as (III).

Ladefoged (1970) in debate with Fromkin (1970) expressed the view that it would be unwise to try to claim for any current feature system "any more than that it is a summary of the data we know we now have available, and there are several limitations on our present data." Chomsky and Halle (p. 298) had themselves recognized that the many gaps in their knowledge made the success of their ambition to "cover every inherent phonetic feature" somewhat problematical. The set of features in which the phonological component is written is constantly under review and I believe that a revision as suggested above would be a valuable modification, bringing the phonetic apparatus of phonology more closely into line with current speech production theory and thereby increasing its explanatory power.

CONCLUSIONS

Cases had been reported of [ɪ] having a lower tongue height than [e], and [ɔ] lower than [ɑ], contradictory to the tongue heights prescribed by the tongue-arching model. My examination of all possible cases in

38 sets of X-ray tracings published during the past 70 years confirms these reports. In every example, [ɪ] was lower than [e]. The tongue height relation between [ɔ] and [ɑ - a] was random. Unexpectedly, some cases of [o] lower than [ɔ] were also found and even [o] lower than [ɑ]. In addition, a few cases suggested the relation between [ʊ] and [o] was also random. Further, the concept of a continuous scale of tongue retraction was found to be false. The tongue positions prescribed by the model do not agree with those observed in actual speech. This model therefore gives an inaccurate representation of vowel articulation.

The concept of tongue height is ambiguous with regard to vocal tract shaping. Its effect varies according to the location of the tongue constriction in the vocal tract. Its two components (mandibular and lingual) have different acoustic consequences, the former altering the radiation characteristics of the mouth-opening while the latter does not. This ambiguity means that tongue height is useless as an articulatory parameter of vocal tract shaping. Close comparison of one set of German /ī, ɪ , ē, ɛ/ profiles and area functions indicated that the mandibular and lingual components must be treated separately as independent gestures, the difference between [i, ɪ]-like and [e, ɛ]-like vowels being mandibular, and between "tense" and "lax" vowels lingual.

Vocal cavity configurations are constant for different renderings of the same vowels, although the X-rayed tongue arch positions may be as random as those observed for [ɔ]-like and [a]-like vowels. The same was true for a particularly puzzling pair of English /ʌ/ tongue arch positions that Carmody was unwilling to accept, but whose area functions were nevertheless remarkably similar. The variability of tongue arch position and the constancy of vocal tract configuration indicates that it would be more fruitful to describe how articulation strives to achieve the constant cavity configurations.

The physiological basis, the predictive capability and the explanatory power of the tongue-arching model are all very weak. It neglects the pharynx. The coordinates of tongue arch position (height and fronting) are not related to the observed activity of the extrinsic muscles of the tongue. The tongue arch positions it aims to describe are found to be very variable in actual speech and are largely irrelevant to the detailed shaping of the vocal cavities. Its parameters are ambiguous with regard to the acoustic output of the vocal tract. Consequently, the tongue-arching model is a poor articulatory medium for relating neuromotor activity, movements of the articulators, vocal tract configurations and the acoustic character of speech. It obstructs the building of a comprehensive description of speech production in which each of the successive stages (neuromotor, articulation, cavity shaping, acoustic output) are unambiguously related to each other. It therefore constitutes an unnecessarily weak link in phonetic and linguistics theory.

For several decades, there has seemed to be greater regularity in the spectral character of vowels than in their articulation in terms of the tongue-arching model, and many phoneticians have therefore preferred an acoustic or auditory rather than articulatory description of speech. However, it was seen above that this spectral constancy is matched by similar constancy in the vocal tract configurations. I suggest that articulation is not in itself inconstant, but that it has instead been described in terms of an unsatisfactory model whose parameter values have provided a bewildering and variable picture of actual speech. Given that the spectral character of speech and the cavity configurations are constant, there is probably similar constancy in the coordination of the gestures that create the resonating cavities and in the packages of motor commands necessary for these gestures.

A more suitable definition of the place of artculation of vowels would be in terms of the place and degree of tongue constriction. Now that the acoustics of the vocal tract are more thoroughly understood, there is a growing tendency to look in this direction for an alternative to the tongue-arching model. The explanatory power of the phonological component of grammar would be greatly enchanced if the features of tongue movement were based on this type of model instead of on the tongue-arching model.

NOTES

1. Except in general terms. For example the 6th or 5th century BC authors of the Sanskrit Atharva-Veda Prātiśākhya recitation manual (Whitney, 1862: § I.i.36) taught that the short Sanskrit [a] for /a/ was "obscured" by narrowing the mouth-opening relative to the long [a:] for /a+a/. Another example can be found in a treatise of the 2nd century AD Roman grammarian Terentianus Maurus (Keil, vol. 6) who described how the "tragic tone of the mouth cavern" of [o] and the "graver quality" of [u] are produced by rounding and protruding the lips.

2. These additional branches were known as mixed because they combined the tongue of one basic series (palatal or velar) with the lips of the other (plain or round). The same term mixed later came to denote the "central" vowels of the tongue-arching model, whence the subsequent confusion in interpreting the older trees.

3. There are several references in this paper to conclusions based on the collection of published sets of X-ray tracings. More detailed accounts will be given in a forthcoming thesis. The four constriction locations were found in every published set examined. The jaw and tongue positions described for "tense" and "lax" palatal vowels are typical for the whole collection of X-ray tracings and for my own X-ray film of Southern British speech. The description of the W. Greenlandic vowel system has been fully revised and fresh cineradiographic material added. See also note 13.

4. Exceptions, where they were given prominence, were Kruisinga's textbook of English pronunciation, Jespersen's handbook and Russel's polemic treatise on vowel theory. Vietor had agreed with Meyer. D. Jones (1967: § 129) mentioned that "the late Dr. E. Meyer of Stockholm obtained excellent diagrams of the tongue positions of vowels by means of a row of fine leaden threads attached to an artificial palate" but did not report that these same excellent diagrams contradicted part of what he himself was teaching in the book. Malmberg has frequently pointed out how Meyer's findings, both in this and other fields, have been confirmed by later investigators. For example, in 1952 and in his obituary tribute to Meyer (1953) where he wrote of

"...seine plastographische Methode...wodurch er die der älteren Palatographie ersetzte und dank welcher er dann auch die althergebrachte Vorstellung von einer festen Beziehung zwischen Zungenstellung und Lautklang als principiell falsch ablehnen konnte", and "jedenfalls hat die moderne Phonetik durch eine Kombination von Röntgenographie und akustischer Lautspektrographie die Richtigkeit der Meyerschen Ergebnisse in erstaunender Weise bestätigt". It should be noted that my thesis is that Meyer discredited one particular tongue articulation model for vowels. I still maintain that there is a firm relationship between tongue gestures and vowel quality.

5. The Stevens and House nomograms give the frequencies of the first two formants generated by a three parameter vocal tract model for different constriction locations, for different mouth-openings (represented by the values of the lip-opening area/length ratio A/l cms) and for different cross-section areas at the constriction. When these numbers are inserted in their equation, a close approximation to a natural vocal tract area function is obtained. Fig. 4 gives the case where the constriction is located 12 cms from the source, a suitable value to represent the palatal vowels.

6. This comment applies only to the examples quoted here. There are differences between dialects. The quality of the vowel segment denoted /ʌ/ today has changed from an [u]-like quality over the past few hundred years, the older quality still being preserved in spellings. There are dialects, especially in northern and central England, where this change has only been partial, the corresponding vowel having an [ʊ] or [ɤ]-like quality. While handbooks of American English frequently quote an [ə]-like quality for /ʌ/, the Peterson and Barney spectra have the typical high F_1 of an [a]-like quality.

7. This disagreement prompted Ladefoged and his colleagues to conclude that each indvidual speaker evolves articulatory behaviour that is peculiar to himself.

8. For the English dialects, [ɔ] represents the vowel of caught in American English and cot in southern British English, while [o] represents the vowel of coat in American English and caught in southern British English.

9. Galen's own account of the structure, movements and innervation of the tongue is contained in books IX (cranium, brain and cranial nerves) and XI (face and jaws) of his "On the uses of the parts of the body of man" (Darembourg, 1854). It is hardly surprising that reference to apes was not helpful for this point. A necessary element of the ontogenesis of the speech organs (compared with the oral anatomy of the non-human primates) is the 90° bend in the vocal tract, resulting from the erect posture of man, that permits a wide range of variation of the vocal cavities by means of tongue movement. This has been particularly stressed by Lieberman (1972).

10. One of the principal anatomical references of Hellwag is to Albinus's edition (1744) of the plates of Eustachius that had remained unpublished from 1522 until 1714. They are considered to be more accurate than those of Eustachius's contemporary Vesalius (Singer, 1957, p. 135). For the tongue in particular, Hellwag referred to a work of Heuermann, De lingua humana, 1749.

11. But the doctrine of maximum auditory contrast will need to be modified. It is not true that acoustic contrasts are always as large as possible or that when the number of phonemes in a vowel system is increased the new contrasts are necessarily the largest available. Consider for example the simple observation that numerous qualitative distinctions may be evolved among the palatal vowels (as many as /i, ɪ, e, ɛ, y, ʏ, ø/ with consequently very small spectral contrasts) while the whole spectral range of [a]-like vowels frequently remains phonemically undivided. One rarely finds as many as three qualitatively different [a]-like phonemes (as in English, /æ, ʌ, ɑ/). This cannot be explained in terms of seeking out maximum contrasts. I shall argue in a forthcoming paper that there are good articulatory reasons for vowel systems to develop in this way. I can mention in particular the excellent tactile feedback afforded by the tongue in contact with the upper teeth and the opportunity for individual exploitation of mandibular and lingual gestures, permitting precise control of palatal vowels. For [a]-like vowels the mouth-opening must be wide (a close opening would endanger contrasts with [ɛ], [ɜ] and [ɔ]) and labial contrasts are not practical (lip-rounding might lead to confusion with [œ] and [ɔ]). This means that the only articulatory variable that can

be exploited for useful spectral contrasts among [a]-like vowels is the degree of tongue constriction in the lower pharynx. Here is the key to why the predictions of the magnetic repulsion analogy of Liljencrantz and Lindblom (1972) became weaker for systems of seven vowels or more. Instead of exploiting more constrasts among the palatal vowels, it preferred to arrange vowels equidistantly around the boundary of the spectral space.

12. Ladefoged has pointed out that their belief in this as a neutral tongue position is not well founded. They refer to what they call a preparatory position seen at the beginning of X-ray motion films. They may have been misled by a superficial similarity between the normal breathing configuration and the palatal [ɛ] configuration.

13. Suppose it is necessary to transform the three "heights" [i, ɛ, a] (+high -low, -high -low, -high +low) into four for [i, e, ɛ, a]. The introduction of a feature mid necessitates a redefinition of the other features to enable [ɛ] to become +low (+high -mid, +high +mid, +low +mid, +low -mid). Wang's suggestion (1968) for +high -mid, +high +mid, -high +mid and -high -mid also necessitates amendments to the initial set of features prior to the introduction of the fourth "height". I propose instead a set of articulatory features based on gestures that shape the vocal cavities. This will generate any relevant number of vowels along the [i - ɑ] scale as follows:

	ɑ	a	æ	i	ɪ	e	ɛ
pharyngeal	+	+	+	–	–	–	–
palatal	–	–	+	+	+	+	+
"tense"	+	–	(–)	+	–	+	–
open	+	+	+	–	–	+	+

Part of the foundation for this approach has already been described elsewhere in this paper. The complete scheme, with definitions and phonetic evidence, will be described in a forthcoming thesis[3] where it will be applied to the solution of a number of phonological problems.

14. "The phonological component accepts as input a structurally analysed string. As output it provides the 'phonetic representation' of this string. The phonetic representation consists of a sequence of 'phonetic segments' each of which is nothing more than a set of 'phonetic feature specifications'..." (p. 164).

BIBLIOGRAPHICAL REFERENCES

Atkinson H.W. 1898. Methods of mouth-mapping. Die Neueren Sprachen 6:494-503

Bell A.M. 1867. Visible speech: universal alphabetics or self-interpreting physiological letters for the writing of all languages in one alphabet. London

Bravmann M. 1934. Materialen und Untersuchungen zu den Phonetischen Lehren der Araber, Göttingen. (Inaugural dissertation). With a German translation of Ibn Sīnā's treatise

Brücke E. 1856. Grundzüger der Physiologie und Systematik der Sprachlaute für Linguisten und Taubstummenlehrer. Vienna

Carmody F.J. 1936. Radiographs of 13 German vowels. Archives Néerlandaises de Phonétique Expérimentale 12:27-33

Carmody F.J. 1937. X-ray studies of speech articulations. Notes and X-ray films of the late Richard T. Holbrook. University of California Publications in Modern Philology, vol. 20, no. 4

Carmody F.J. 1941. An X-ray study of pharyngeal articulation. University of California Publications in Modern Philology, vol. 21, no. 5

Chiba T. and Kajiyama M. 1941. The vowel, its nature and structure. Tokyo

Chlumsky J. 1913. Méthodes pour obtenir le profil de la langue pendant l'articulation. Revue de Phonétique 3:167-173

Chlumsky J. 1914. A review of E. Meyer's article: Das Problem der Vokalspannung. Revue de Phonétique 4:90-95

Chlumsky J., Pauphilet A. and Polland B. 1938. Radiografie francouzských samohlásek a polosamohlásek. Rozpravy České Akademie Věd a Umění, Trida 3, Cislo 75. Prague

Chomsky N. and Halle M. 1968. The sound pattern of English. Englewood Cliffs, NJ

Darembourg C. 1854. Oeuvres anatomiques, physiologiques et médicales de Galien. Paris, 2 vols

Delattre P. 1951. The physiological interpretation of sound spectrograms. Publications of the Modern Languages Association of America 66:864. Reprinted in Studies in French and comparative phonetics, the Hague, 1966

Delattre P. 1971. Pharyngeal features in the consonants of Arabic, German, Spanish, French and American English. Phonetica 23: 129-155

Elert C.-C. 1964. Phonologic studies of quantity in Swedish based on material from Stockholm speakers. Uppsala. Monografier utgivna av Stockholms kommunalförvaltning, no. 27

Fant C.G.M. 1958. Modern instruments and methods for acoustic studies of speech. Acta Polytechnica Scandinavica. Also in Proceedings of the 8th international congress of linguists, Oslo, pp. 282-362

Fant C.G.M. 1960. The acoustic theory of speech production. The Hague

Fant C.G.M. 1964. Formants and cavities. Proceedings of the 5th international congress of phonetic sciences, Münster, pp. 120-141

Fant C.G.M. 1969. Distinctive features and phonetic dimensions. STL-QPSR 2-3/1969: 1-18, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm

Fromkin V. 1970. A reply to P. Ladefoged's "The phonetic framework of generative phonology". UCLA Working Papers in Phonetics no. 14: 34-39

Galen. On the uses of the parts of the human body. French translation by Darembourg (1854).

Grandgent C.H. 1890. Vowel measurements. Publications of the Modern Languages Association of America 5: 148-174

Gunnilstam O. 1974. The theory of local linearity. Journal of Phonetics 2:91–108

Gutzmann H. 1909. Physiologie der Stimme und Sprache. Braunschweig

Gutzmann H. 1930. Röntgenaufnahmen von Zunge und Gaumensegel bei Vokalen und Dauerkonsonanten. Fortschritte auf dem Gebiete der Röntgenstrahlen 41:392–404

Hadding-Koch K. and Abrahamson A.S. 1966. Duration versus spectrum in Swedish vowels: some perceptual experiments. Studia Linguistica 18:94–107

Hála B. 1959. English vowels in phonetic pictures. Rozpravy Československé Akademie Věd, Sešit 5, Ročnik 69

Halle M. 1959. The sound pattern of Russian. The Hague

Halle M. 1964. On the bases of phonology. In Fodor and Katz (eds), The structure of language, Englewood Cliffs, NJ, pp. 324–333

Halle M. and Stevens K.N. 1969. On the feature Advanced Tongue Root. QPR 94:209–215. MIT Research Laboratory of Electronics

Harris K. 1971. Action of the extrinsic tongue musculature in the control of tongue position. Preliminary report. SR 25/26, pp. 87–96, Haskins Laboratory

Hellwag C.F. 1781. Dissertatio inauguralis physiologico-medica de formatione loquelae. Tübingen. Reprinted by Vietor, 1886, Heilbronn

Helmholtz H. von. 1863. Die Lehre von den Tonempfindungen. Braunschweig

Ibn Sīnā. Ca 1000. Risalah fi asbāb ḥuduth al-ḥurūf (a treatise on the points of articulation of the speech sounds). See Bravmann (1934) and Semaan (1963) for translations.

Jakobson R., Fant C.G.M. and Halle M. 1952. Preliminaries to speech analysis. Cambridge, Mass.

Jakobson R. and Halle M. 1964. Tenseness and laxness. In Abercrombie et al. (eds), In honour of Daniel Jones, London. Also in reprints of Jakobson et al. (1952)

Jakobson R. and Halle M. 1968. Phonology in relation to phonetics. In Malmberg (ed), Manual of phonetics, the Hague. Chapt. 13

Jones D. 1967. An outline of English phonetics. 9th edn. Cambridge. First published, 1918

Jones S. 1929. Radiography and pronunciation. The British Journal of Radiography 2: 149-150

Joos M. 1948. Acoustic phonetics. Language Monograph no. 23. Supplement to Language, vol. 24

Kaneko N. 1957. Specific X-ray observations on the movement of the lower jaw, lips and tongue in pronunciation. The Study of Sounds 8: 1-18 and 450-451. Tokyo

Keil F. 1855-1880. Grammatici latini. 7 vols. Leipzig

Kent R.D. 1972. Some considerations in the fluorographic analysis of tongue movements during speech. Phonetica 26: 16-32

Koneczna H. and Zawadowski W. 1956. Obrazy rentgenografizne glosek rosijskich. Warsaw

Korlén G. and Malmberg B. 1959. Tysk ljudlära. Lund

Kruisinga E. 1925. A handbook of presentday English. Part I: English sounds. Utrecht

Ladefoged P. 1964. Some possibilities in speech synthesis. Language and speech 7:205-214

Ladefoged P. 1967. Three areas of experimental phonetics. Oxford. Chapt. 2

Ladefoged P. 1970. The phonetic framework of generative phonology. UCLA Working Papers in Phonetics no. 14:25-32

Ladefoged P. 1971. Preliminaries to linguistic phonetics. Chicago

Ladefoged P., Anthony J.F.K. and Riley C. 1971. Direct measurement of the vocal tract. UCLA Working Papers in Phonetics 19:4-13

Ladefoged P., DeClerk J.L. and Harshman R. 1972. Control of the tongue in vowels. Proceedings of the 7th international congress of phonetic sciences, the Hague, pp. 349-354

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

Leky M. 1917. Grundlagen einer allgemeinen Phonetik als Vorstufe zur Sprachwissenschaft. Cologne

Lieberman P. 1972. The speech of primates. The Hague

Lieberman P. and Crelin E.S. 1971. On the speech of Neanderthal man. Linguistic Inquiry 2:203–222. Also reprinted in Lieberman 1972.

Liljencrants J. and Lindblom B. 1972. Numerical simulation of vowel quality systems. The role of perceptual contrast. Language 48: 839–862

Lindau M., Jacobson L. and Ladefoged P. 1972. The feature advanced tongue root. UCLA Working Papers in Phonetics 22:76–94

Lindblom B. 1963. Spectrographic study of vowel reduction. Journal of the Acoustical Society of America 35:1773–1781

Lindblom B. 1972. Phonetics and the description of language. Proceedings of the 7th international congress of phonetic sciences, the Hague, pp. 63–97

Lindblom B. and Sundberg J. 1969a. A quantitative model of vowel production and the distinctive features of Swedish vowels. STL-QPSR 1/1969, pp. 14–32. Speech Transmission Laboratory, Royal Institute of Technology, Stockholm

Lindblom B. and Sundberg J. 1969b. A quantitative theory of cardinal vowels and the teaching of pronunciation. STL-QPSR 2–3/1969, pp. 19–25. Speech Transmission Laboratory, Royal Institute of Technology, Stockholm

Lindblom B. and Sundberg J. 1971. Acoustical consequences of lip, tongue, jaw and larynx movements. Journal of the Acoustical Society of America 50:1166–1179

Lindqvist J. and Sundberg J. 1971. Pharyngeal constrictions. STL-QPSR 4/1971, pp. 26–31. Speech Transmission Laboratory, Royal Institute of Technology, Stockholm

Lloyd R.J. 1890–1892. Speech sounds, their nature and causation. Phonetische Studien, vols 3, 4 and 5

McCawley J.D. 1967. Le rôle d'un système de traits phonologiques dans une théorie du langage. Langages, no. 8, pp. 112–123

MacMillan A.S. and Keleman G. 1952. Radiography of the supraglottic speech organs. Archives of Otolaryngology 55:671-688

MacNeilage P. and Sholes G.N. 1964. An electromyographic study of the tongue during vowel production. Journal of Speech and Hearing Research 17:211-232

Malmberg B. 1952. Le problème du classement des sons du langage et quelques questions connexes. Studia Linguistica 6:1-56

Malmberg B. 1953. Ernst A. Meyer. Studia Linguistica 7:119-122

Malmberg B. 1966. Spansk fonetik

Mazlovà V. 1949. Výslovnost na Zábřežsku, fonetická studie z moravské dialektologie. Prague

Mettas O., Fant C.G.M. and Stålhammar U. 1971. The A-vowels of Parisian French. STL-QPSR 4/1971, pp. 1-8. Speech Transmission Laboratory, Royal Institute of Technology, Stockholm

Meyer E. 1907. Röntgenographische Lautbilder. Medizinisch-Pädagogische Monatschrift für die gesamte Sprachheilkunde. Vol 17

Meyer E. 1910. Untersuchungen über Lautbildung. Festschrift Wilhelm Vietor, pp. 166-248. (Special number of Die Neueren Sprachen)

Meyer E. 1913. Das Problem der Vokalspannung. Die Neueren Sprachen 21:65-86 and 146-171

Minifie F.D., Hixon T.J., Kelsey C.A. and Woodhouse R.J. 1970. Lateral pharyngeal wall movement during speech production. Journal of Speech and Hearing Research 13:584-594

Moll K.L. 1960. Cinefluorographic techniques in speech research. Journal of Speech and Hearing Research 3:227-241

Ohnesorg K. and Švarny O. 1955. Etudes expérimentales des articulations chinoises. Rospravý Československé Akademie Věd, Ročník 65, Sešit 5

Permenter C.E. and Treviño S.N. 1932. An X-ray study of Spanish vowels. Hispania, 15:483-496. Stanford University

Perkell J.S. 1969. Physiology of speech production: results and implications of a quantitative cineradiographic study. Cambridge, Mass.

Perkell J.S. 1971. Physiology of speech production: a preliminary study of two suggested revisions of the features specifying vowels. QPR 102: 123–139. MIT Research Laboratory of Electronics

Peterson G. and Barney H. 1952. Control methods used in a study of the vowels. Journal of the Acoustical Society of America 24: 175–184

Pétursson M. 1974a. Peut-on interpréter les données de la radiociné-matographie en fonction du tube acoustique à section uniforme? Phonetica 29;22–79

Pétursson M. 1974b. Les articulations de l'islandais à la lumière de! la radiocinématographie. Société de Linguistique de Paris, collection linguistique no. 68

Polland B. and Hála B. 1926. Artikulace Českych zvukû v roentgenových obrazech. Ná kladem České Akademie Věd a Umění. Prague

Potter R.K., Kopp A.G. and Green H.C. 1947. Visible speech. New York

Potter R.K. and Steinberg J.C. 1950. Toward the specification of speech. Journal of the Acoustical Society of America 22:807–820

Raphael L.J. 1971a. An electromyographic investigation of the tense-lax feature in some English vowels. SR 25/26, pp. 131–140, Haskins Laboratories

Raphael L.J. 1971b. An electromyographic investigation of the feature of tension in some American English vowels. SR 28, pp. 179–191, Haskins Laboratories

Russel O. 1928. The vowel. Columbus, Ohio

Russel O. 1935. Synchronised X-ray, oscillograph, sound and movie experiments, showing the fallacy of vowel triangle and open-close theories. Proceedings of the second international congress of phonetic sciences, London, pp. 198–205

Scheier M. 1909. Die Bedeutung des Röntgenverfahrens für die Physiologie der Sprache und Stimme. Archiv für Laryngologie und Rhinologie, 22: 175–208

Semaan K.I. 1963. Arabic phonetics. Ibn Sīnā's Risālah on the points of articulation of the speech sounds, translated from the mediaevel Arabic. Arthur Jeffery Memorial Monographs no. 2, Lahore

Simon P. 1961. Films radiologiques des articulations et les aspects génétiques des sons du langage. Orbis 10:47-68

Singer C. 1957. A short history of anatomy and physiology from the Greeks to Harvey. New York

Skaličková A. 1955. The Korean vowels. Archiv Orientální 23:29-51

Smith T. 1971. A phonetic study of the phonetic function of the extrinsic tongue muscles. UCLA Working Papers in Phonetics, no 18

Sovijärvi A. 1938. Die gehalteten, geflüsterten und gesungenen Vokale der Finnischen Sprache. Helsinki

Stevens K.N. 1972. The quantal nature of speech, evidence from articulatory-acoustic data. In David E.E. and Denes P.B. (eds.): Human communication, a unified view, pp. 51-56

Stevens K.N. and House A.S. 1955. Development of a quantitative description of vowel articulation. Journal of the Acoustical Society of America 27:484-495

Stevens K.N. and House A.S. 1961. An acoustical theory of vowel production and some of its implications. Journal of Speech and Hearing Research 4:303-320

Stevens K.N. and House A.S. 1963. Perturbation of vowel articulations by consonantal context. Journal of Speech and Hearing Research 6:111-128

Stewart J.M. 1967. Tongue root position in Akan vowel harmony. Phonetica 16:185-204

Strenger F. 1968. Radiographic, palatographic and labiographic methods in phonetics. In Malmberg B. (ed.): Manual of phonetics, Amsterdam, pp. 334-364

Subtelny J.D., Pruzansky S. and Subtelny J. 1957. The application of Roentgenography in the study of speech. In Kaiser L. (ed.): Manual of phonetics, Amsterdam, pp. 166-179

Sundberg J. 1969. Articulatory differences between spoken and sung vowels in singers. STL-QPSR 1/1969:33-45, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm

Tiffany W.R. 1959. Non-random sources of variation in vowel quality. Journal of Speech and Hearing Research 2:305-317

Vennemann T. and Ladefoged P. 1971. Phonetic features and phonological features. UCLA Working Papers in Phonetics 21:13:24

Vietor W. 1914. Zur Systematik der Vokalartikulation. Miscellanea Phonetica 1:1-5

Wang W.S.-Y. 1968. Vowel Features, paired variables and the English vowel shift. Language 44:695-708

Wängler H.H. 1961. Atlas deutscher Sprachlaute. Berlin

Whitney W.D. 1962. The Atharva-Veda Prātiçākhya, or Çaunakīyā Çaturādhyāyikā: text, translation and notes. Journal of the American Oriental Society 7:333-615

Wood S. 1971. A spectrographic study of allophonic variation and vowel reduction in West Greenlandic Eskimo. Working Papers no. 4, pp. 58-94, Phonetics Laboratory, Department of General Linguistics, Lund University. See note 3

Zemlin W.R. 1968. Speech and hearing science - anatomy and physiology. Englewood Cliffs, N.J.

Working Papers 11·1975 Lund University
Phonetics Laboratory General Linguistics

TENSE AND LAX VOWELS - DEGREE OF CONSTRICTION OR PHARYNGEAL VOLUME?

Sidney Wood
September 1975

Introduction

There are differences of both the degree of tongue constriction and the volume of the lower pharynx between tense and lax vowels. These factors are modifications of the configuration of the vocal tract and will consequently alter its resonances. For a complete account of the production of different vowel categories, it is necessary to know the magnitude of acoustical difference that can be referred to any particular articulatory variable. The nomograms published by Stevens and House (1955) and Fant (1960) based on the three-parameter model have been very helpful in describing the acoustical properties of the vocal tract but their usefulness is strictly limited by the difficulty of relating the model parameters to specific articulatory manoeuvres in a number of situations. The exploration of the acoustical consequences of lip, tongue, jaw and larynx movement by Lindblom and Sundberg (1971) has shown the way to the solution of this type of problem. A midsaggital profile of the vocal tract is deliberately altered and the resonances of each configuration are measured or calculated. This can be done either by computer or with the aid of an electrical analogue. The experiments to be described below were designed to assess how much of the acoustical difference within pairs of tense and lax vowels can be attributed to the degree of constriction and how much to the pharyngeal volume. Midsaggital profiles of the vocal tract were systematically modified, the corresponding area functions set on an electrical vocal tract analogue (LEA[1]) and the resonance frequencies found and measured.

Tense and lax vowels

The terms tense and lax are notoriously ambiguous in both phonetics and phonology. There are two types of ambiguity I particularly wish to underline. The one concerns the physiological and acoustical character of the contrasts. This ambiguity is not so serious since it reflects our limited knowledge of the production processes involved. As our knowledge improves, this amgiguity will be resolved. Far more serious is the confusion of tenseness and laxness with vowel length or quantity.

I shall restrict the terms tense and lax exclusively to the timbre differences in such pairs as [i-ɪ, e-ɛ, u-ʊ, o-ɔ, ɑ-a] (and the rounded palatals [y-ʏ, ø-œ] which for the remainder of this report will be subsumed with the spread-lip palatals). This usage is not inconsistant with the traditional definition in terms of muscular tension of the tongue which implies differences of lingual articulation and consequently of vocal tract configuration and resonance. There is necessarily an acoustical difference between tense and lax vowels.

There is a well known tendency for tense vowels to be longer than lax vowels. This is usually said to be due to the tense gestures taking more time to execute. It is an undeniable fact that in many languages tense vowels are long and lax vowels short. But other relationships are also found such as timbre contrasts between vowels of the same length or quantity contrasts between vowels of the same timbre. The relationship between tenseness and quantity can vary synchronically from language to language and diachronically from period to period in one and the same language. The relationship between tense vowels, long vowels and diphthongs is complex and does not become simpler if tenseness and quantity are treated as equivalent. The examples at Table 1 follow from distinguishing between tense-lax timbre contrasts and long-short quantity contrasts.

	TENSE					LAX				
LONG	i:	e:	u:	o:	ɑ:		ɛ:		ɔ:	a:
SHORT	i	e	u		ɑ	ɪ	ɛ	ʊ	ɔ	a

Table 1. Tense-lax and long-short pairs of vowels. The contrast /i:-i/ is long versus short (tense). A contrast /i:-ɪ/ is long and tense versus short and lax. A contrast /i-ɪ/ is tense versus lax.

Vocal tract differences

Tracings of X-rayed vowel articulations reveal consistent differences of both degree of constriction and of pharyngeal volume between tense and lax vowels. In addition, there are also differences of lip position (less rounded, sometimes less spread, for lax vowels) and larynx position (deeper for tense vowels, especially for rounded vowels). The articulatory gestures involved appear to be much the same irrespective of language, which points to a universal physiological and biological basis for the acoustical contrasts founded on this difference. I have drawn this conclusion from analysis of the same collection of published sets of X-ray tracings as was used for my criticism of the tongue-arching model (1975). As a control on these conclusions, I have also analysed five X-ray motion films (English, Egyptian, Southern Swedish and West Greenlandic Eskimo) that have been made in Lund[2]. The following is a summary of the findings that are relevant to the present problem[3].

The degree of constriction is quantified as the cross-section area of the vocal tract at the tongue constriction.

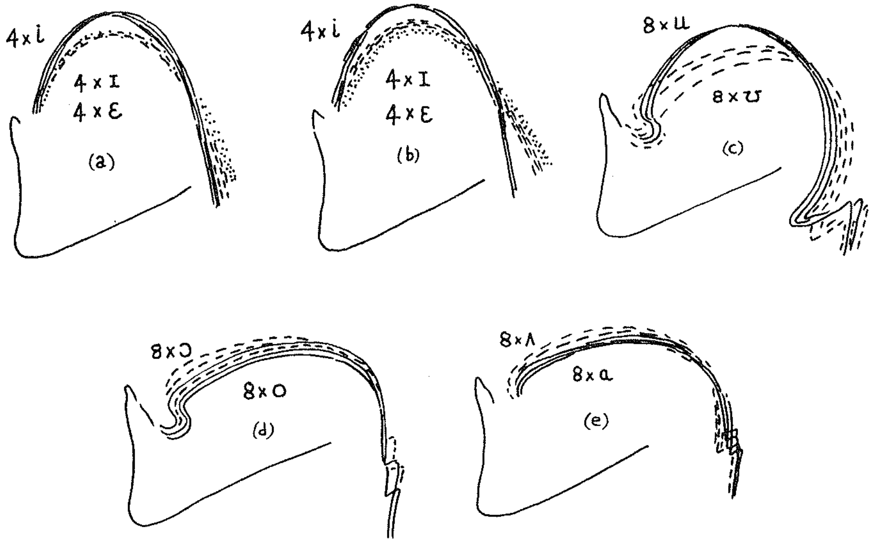


Fig. 1. Sets of tongue profiles for tense-lax pairs by a Southern British English subject. There are 8 examples of each vowel, 4 uttered a little slower than average everyday speech (4.5 syllables/sec) and 4 a little faster (6.5 syllables/sec). The main articulatory consequence of the rate difference was a narrower jaw-opening for open vowels [ɛ , o, ɔ, ɑ, ʌ]. There was hardly any influence on the tongue profile, except for the palatal [ɛ] where the tongue was lower relative to the mandible in the faster set (b) to compensate for the higher position.

There is considerable similarity of constriction size for similar vowel qualities irrespective of language. Typical ranges are given in Table 2.

CONSTRICTION	HARD PALATE		SOFT PALATE	UPPER PHARYNX	LOWER PHARYNX	
VOWEL PAIR	i/ɪ	e/ɛ	u/ʊ	o/ɔ	ɑ/a	
TENSE VOWEL	0.5-1.0	1.0-1.7	0.5-1.0	0.6-1.0	0.5-1.0	cm^2
LAX VOWEL	1.6-2.2	2.5-3.0	1.5-	0.4-0.7	1.3-1.7	

Table 2. Cross-section area of the vocal tract at the tongue constriction, representing the degree of constriction. The tense vowel has the narrower constriction, except for the [o-ɔ] pair.

Each pair is characterized by a widening of the constricted passage by 3-4 mm for the lax vowel. The exception is the [o-ɔ] pair where the lax vowel just has the narrower con-

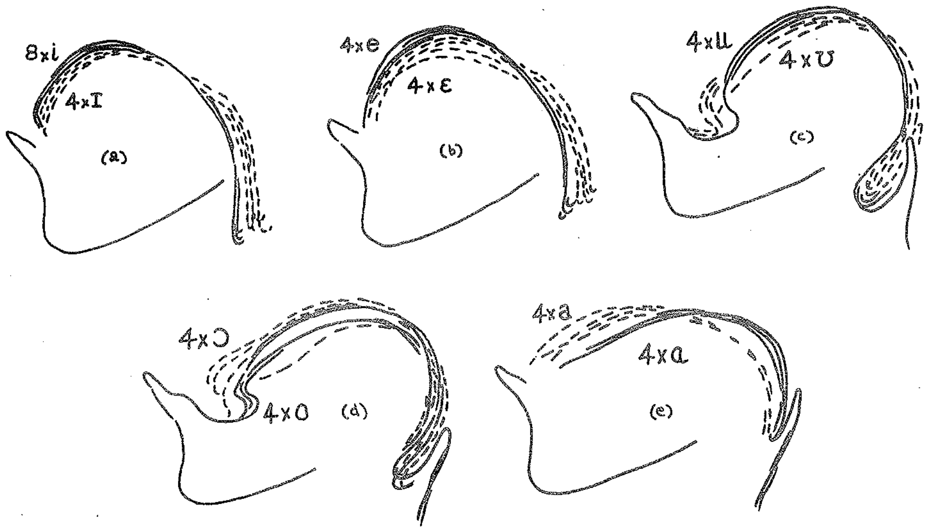


Fig. 2. Sets of tongue profiles for tense-lax pairs by an Egyptian subject. There are four examples of each vowel, except for [i]. The [a] quality represents /a/ in a "non-emphatic" environment while [ɑ] represents 2 x /a/ and 2 x /a+a/ in an "emphatic" environment.

striction although both ranges virtually overlap. In the case of [ʊ], when the velar passage is widened beyond 2.0 cm^2 the back of the tongue begins to constrict the upper pharynx instead. The quoted ranges are characteristic for each vowel quality.

For all these pairs (except [ɑ-a]), there are corresponding differences in the lower pharynx (Table 3). In the case of the [ɑ-a]-like vowels, the lower pharynx is constricted by the tongue so that variation of low pharyngeal width therefore modifies the constriction itself. Moreover, the tense vowel [ɑ] has the narrower pharynx.

Physiologically, these differences of degree of constriction and low pharyngeal volume are created by the movement of the tongue. This movement must be broken into its lingual and mandibular components (Lindblom and Sundberg, 1971). The tongue constriction is formed by directing the tongue

CONSTRICTION	HARD PALATE		SOFT PALATE	UPPER PHARYNX	
VOWEL PAIR	i/ɪ	e/ɛ	u/ʊ	o/ɔ	
TENSE VOWEL	25-30	19-23	25-30	15-22	mm
LAX VOWEL	19-23	16-20	19-23	11-19	

Table 3. Typical ranges of low pharyngeal width from the tongue to the rear pharyngeal wall at the epiglottis. The absolute measure depends on the size of the subject's valleculae and is highly variable between individuals. The tense vowel always has the wider lower pharynx.

itself towards (i) the hard palate (for palatal [i-ɪ, e-ɛ]-like vowels), (ii) the soft palate (for palatovelar [u-ʊ]-like vowels), (iii) the upper pharynx (for pharyngovelar [o-ɔ]-like vowels) and (iv) the lower pharynx (for low pharyngeal [ɑ-a]-like vowels) as can be seen at Figs. 1 and 2. At the same time the tongue is raised or lowered bodily by the jaw. This contributes to the constrictions made against the roof of the mouth, i.e. for the palatal and palatovelar vowels. Constrictions in the pharynx are hardly affected by mandibular movement. The jaw occupies two relevant positions during vowels - a closer opening of 5-10 mm for [i, ɪ , u, ʊ]-like vowels and a wider opening of 11-16 mm or more for [e, ɛ , o, ɔ, ɑ, a]-like vowels. The variation depends on such factors as articulation rate and speaking effort. The tongue compensates for the freedom of jaw movement in order to maintain a suitable palatal or palatovelar constriction size (mandibular movement is in the direction of the constriction in these cases). Such lingual compensation is not necessary for the pharyngeal constrictions (but the lips compensate for variation of jaw position in all rounded vowels).

It has been reported that the tongue root is further forward for tense than for lax vowels. The proposed feature advanced tongue root was based on this observation (Halle and Stevens 1969, Perkell 1971). One consequence of advancing the tongue root is to widen the lower pharynx and thus increase its volume. A second consequence is to raise the tongue body, which is in the direction of the constriction in the case of the palatal and palatovelar vowels. The muscles that would pull the tongue root forward are the posterior fibres of the genioglossi. These fibres are also said to assist in raising the tongue. This manoeuvre is necessary for all vowels with a constriction against the roof of the mouth ([i, ɪ, e, ɛ, u, ʊ]). Figs. 1 and 2 show how the tongue root is drawn forward for all these vowels and also how differences of tongue root position between tense and lax vowels in this group are correlated with the height of the tongue relative to the mandible. For the vowels with constricted pharynx ([o, ɔ, ɑ, a]) contraction of the posterior fibres of the genioglossi would be contrary to the rearward constriction-forming gestures. In the case of the pharyngovelar [o, ɔ]-like vowels, it is nevertheless theoretically possible to vary the tongue root position below the upper pharyngeal constriction. Figs. 1 and 2 suggest there was little difference of tongue root position between [o] and [ɔ] for these two subjects, but the tendency was for the tongue root to be more advanced for [o]. In the case of the low pharyngeal [ɑ, a]-like vowels, advancing the tongue root would immediately widen the constriction towards the lax vowel and cannot therefore be utilized for the tense vowel. Figs. 1 and 2 show that for this pair the tongue root is advanced to widen the low pharyngeal constriction for the lax vowel.

The role of the degree of constriction

Sweet (1906) noted that the passage above the tongue appeared to be narrower for tense vowels, the tongue being more "convex". This represents a modification of tongue height (i.e. the sum of the vertical lingual and mandibular

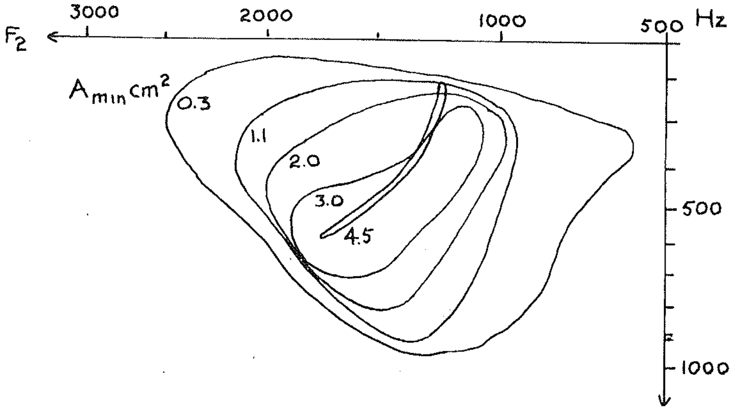


Fig. 3. The maximum possible spectral ranges for F1 and F2 at different degrees of constriction (A_{min} cm^2). This is based on the Stevens and House (1955) three-parameter model nomograms. Each ring encloses the spectra generated by all combinations of constriction location and mouth-opening size for the stated constriction size.

gestures). Tongue height modifies the tongue constriction only in the case of the [i, ɪ, e, ɛ] -like vowels (constricted hard palate) and the [u, ʊ] -like vowels (constricted soft palate). For the vowels with constricted pharynx, the degree of constriction is hardly related to tongue height. In the case of the vowels with constricted lower pharynx, the constriction is indeed narrower for [ɑ] and wider for [a]. This is not exactly what Sweet had had in mind, however, although it is a natural extension of his original idea. He admitted that his distinction between narrow and wide vowels was "not clear in the back vowels where the convexity of the tongue seems to be accompanied by tension of the uvula and soft palate". Sweet was on the track of the truth, that the degree of constriction is a relevant resonator variable in the vocal tract and that differences in the degree of constriction are associated with tenseness and laxness. But his preference for the tongue-arching model, coupled with the impossibility of observing internal articulations and configurations before the discovery of X-rays, effectively concealed the solution from him.

What is the effect of varying the degree of constriction?

The vocal tract is divided into two cavities, one above and one below the tongue constriction. The degree of constriction determines the amount of coupling between the two cavities - that is, the extent to which they resonate together or indepedently of each other. At the one extreme, the constriction is so narrow that the two cavities influence each other relatively little. At the other extreme, the constriction is so wide that the tract becomes a single uninterrupted pipe. Some idea of the consequence of varying the degree of constriction between these extremes is illustrated by Fig. 3 which is based on the Stevens and House nomograms. The degree of constriction is represented by the cross-section area at the constriction, A_{min} cm^2. Each ring encloses an area representing the frequencies of the first and second formants generated by all combinations of constriction location and mouth-opening size for the stated degree of constriction. A constriction of 0.3 cm^2 is about the narrowest possible for pure vowel sounds, further narrowing leading to the production of turbulence in the constriction. At a constriction of 4.5 cm^2, the vocal tract approaches the uniform tube configuration so that the constriction location no longer exerts any influence. Fig. 3 suggests that the possible spectral range is dependent on the degree of constriction. For the maximum possible spectral range, the very small constriction size would be necessary. As the constricted opening widens, the possible spectral range would be reduced. This would mean that the vocal tract resonances are very sensitive to the degree of constriction, as has also been suggested by Gunnilstam (1974). A few millimetres of tongue movement at the constriction would cause a considerable spectral difference. Unfortunately, we cannot be certain that this is due to the degree of constriction alone, since modification of the degree of constriction in the three-parameter model simultaneously involves a change of the low pharyngeal volume.

The role of the lower pharynx

In Sweet's day, the existence of more than one vocal tract resonance was a highly controversial subject among most phoneticians and interest was limited to the bucchal cavity and the crown of the tongue arch. Once the resonance dispute had been settled, the arch was said to divide the tract into two cavities each with its characteristic resonance - the mouth formant and the throat formant. We know today that the tongue arch does not form the dividing constriction and also that the formants have complex cavity affiliations. Nevertheless it is true that modification of the volume of the pharynx will affect the resonances of the vocal tract and that any articulatory modification of the pharynx is therefore acoustically relevant.

Attention was drawn by Stewart (1967) to the role played by the width of the lower pharynx in vowel harmony in the West African language Akan. This harmony difference is very similar to the tense-lax difference, although there are differing opinions as to whether they are both examples of the same phenomenon from the production point of view (Lindau et al. 1972, Lindau 1975). The advanced tongue root proposal claimed to cover both cases. The different tongue root positions for my English and Egyptian subjects have already been seen at Figs. 1 and 2. As already explained, the rule cannot hold for the low pharyngeal [ɑ, a] pair since the lower pharynx is now the location of the constriction.

What is the effect of varying the volume of the lower pharynx? Enlargement due to tongue root advancement occurs in the region of the epiglottis, that is, at about 2 to 4 cm above the glottis. Halle and Stevens recall Chiba and Kajiyama's observation that expansion of an acoustical tube in the vicinity of a sound pressure maximum in the standing wave for a particular natural frequency tends to lower that natural frequency. There is always a maximum in sound pressure distribution close to the glottis for all natural frequencies and in the case of F1 this maximum extends over the first 4 cm of the vocal tract. Hence expansion in this region always causes lowering of F_1. Halle and Stevens also point out that F2 has a pressure minimum at about 2 to 6 cm above the glottis for front vowels and a pressure maximum at about 4 cm above the glottis for back vowels. Expansion in this region will thus cause an upward shift of F_2 for front vowels and a downward shift of F_2 for back vowels. They note that these spectral differences are in the direction observed in acoustic data for tense-lax pairs.

<u>The problem</u>

In both natural speech and in the three-parameter model, the degree of constriction and the lower pharyngeal volume are largely inseparable. It is not therefore immediately apparent which, if either, of these two variables provides the greater contribution to the spectral differences between tense and lax vowels.

It is generally accepted that advancing the tongue root tends to bunch the tongue body towards the roof of the mouth. This manoeuvre thus simultaneously widens the lower pharynx and narrows the palatal or palatovelar constrictions. For the [o , ɔ]-like vowels with constricted upper pharynx, advancing the tongue root in the lower pharynx below the constriction is partially antagonistic to the narrowing of the upper pharynx by the contracting pharyngeal constrictor muscles. As recorded in Tables 2 and 3, I have found a difference of low pharyngeal width in this class but little difference in the degree of constriction (unlike other tense-lax pairs). For all the pharyngeal vowels, any tongue raising associated with tongue root advancement will diminish the volume of the bucchal cavity but at the same time such diminution is countered by any downward movement of the jaw.

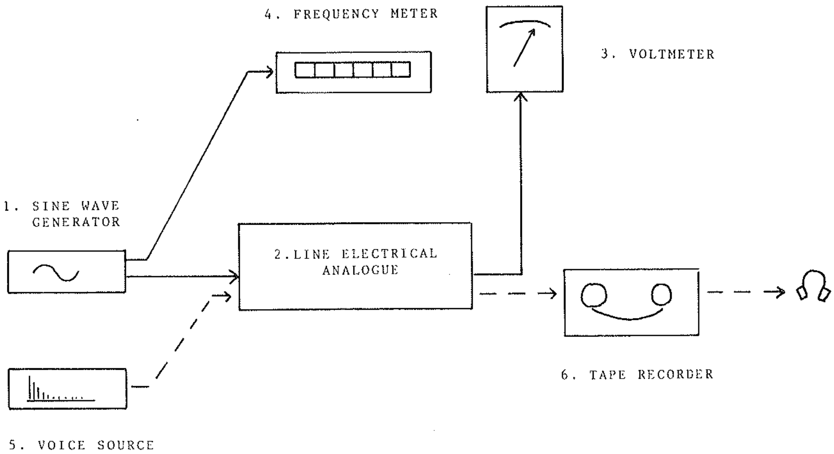


Fig. 4. For sweeping and measuring resonances, a sinus wave from the generator (1) is passed through the analogue LEA (2) to a voltmeter (3). Voltage maxima occur at resonance frequencies which can be read off from the frequency meter (4). For monitoring and recording synthetic vowel qualities, a voice spectrum from a voice source (5) passes through the analogue to a tape recorder (6).

CONSTRICTION	PALATAL				PALATO-VELAR		PHARYNGO-VELAR		LOW PHARYNGEAL	
VOWEL	i	ɪ	e	ɛ	u	ʊ	o	ɔ	ɑ	a
PALATAL	+	+	+	+	+	+	–	–	–	–
VELAR	–	–	–	–	+	+	+	+	–	–
PHARYNGEAL	–	–	–	–	–	–	+	+	+	+
OPEN	–	–	+	+	–	–	+	+	+	+
ROUND	–	–	–	–	+	+	+	+	(–)	–
TENSE	+	–	+	–	+	–	+	–	+	–

Table 4. A matrix showing how the different articulatory components are combined. Each component is defined in the text and by the values given in Tables 2 and 3.

In the three-parameter model, the equation that models the area function relates the opening of the passage above the tongue to the volume of the lower pharynx in a similar fashion to natural speech. Consequently, the different degrees of constriction at Fig. 3 are linked to corresponding pharyngeal differences. It is impossible to say whether the spectral reduction illustrated by this figure is the result of widening the constriction, narrowing the lower pharynx or both (if so in what proportions?). However, we have in the electrical vocal tract analogue a tool that permits us to alter the values of these variables at will. The underlying principle of the experiments reported below is to alter the vocal tract area function in steps from one configuration to another and to note the spectral difference arising from each step.

Method

By careful examination and analysis of motion X-ray films as outlined above, I have isolated the component gestures used by the human speaker to shape the vocal tract. Realistic modifications can be made to a vocal tract replica (a mid-saggital outline of a vocal tract) by reproducing the gestures of natural speech. This has resulted in a building kit that consists of a vocal tract (maxilla and pharynx), a mandible, a tongue for palatal constrictions, a tongue for palatovelar constrictions, a tongue for pharyngovelar constrictions, a tongue for low pharyngeal constrictions, a larynx that can be lowered 5 or 10 mm, sets of lips (spread, plain, slightly rounded, well rounded) and a tongue blade that can be depressed. These components are put together according to the matrix at Table 4.

Open is defined as a jaw-opening larger than 10 mm. For the experiments a jaw-opening of 14 mm was used. Non-open is a jaw-opening smaller than 10 mm. An opening of 8 mm was used.

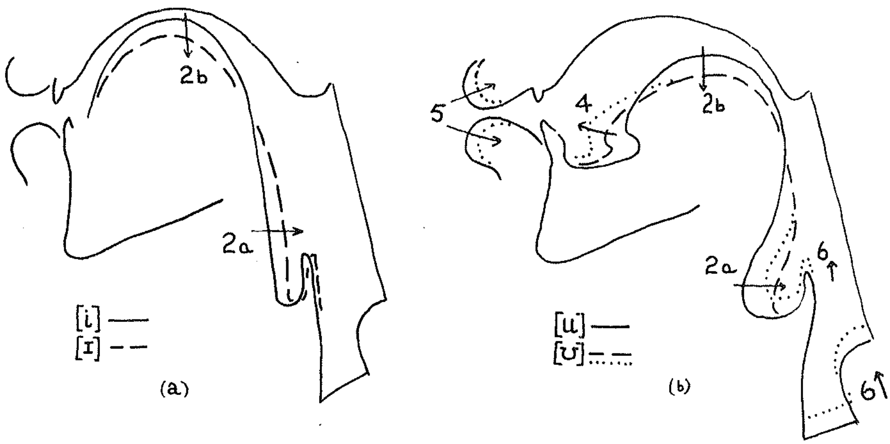


Fig. 5. Modifications made to the model profile for (a) palatal vowels and (b) palatovelar vowels.

For all tense vowels except low pharyngeal, the tongue root was advanced and the tongue body raised. This narrows the constriction of palatals and palatovelars. The constriction of the pharyngovelars is not altered. For the low pharyngeals, the tongue was drawn further into the pharynx to narrow the constriction. For rounded vowels, the larynx was lowered, 10 mm for tense and 5 mm for lax. The lips were more rounded for tense, less rounded for lax. The tongue blade was depressed more for tense rounded vowels, less for lax.

For each configuration, the cross-distances along the tract were transformed into cross-section areas using conversion data published by Sundberg (1969) for the palatal and upper pharyngeal region and by Fant (1960) for the lower pharyngeal region. The area functions thus obtained were then set on the electrical analogue and the resonances measured (Fig. 4).

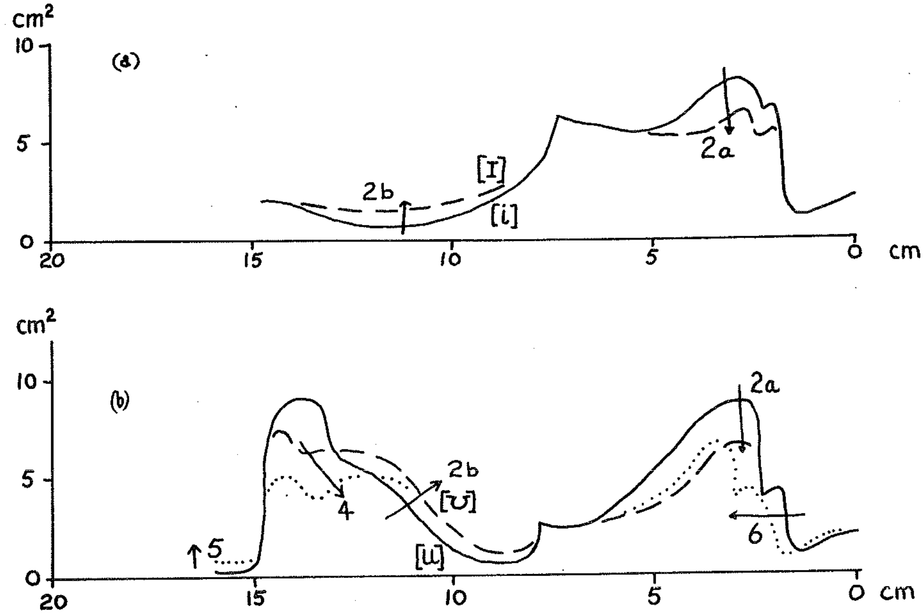


Fig. 6. Area functions for the configurations at Fig. 5(a, b).

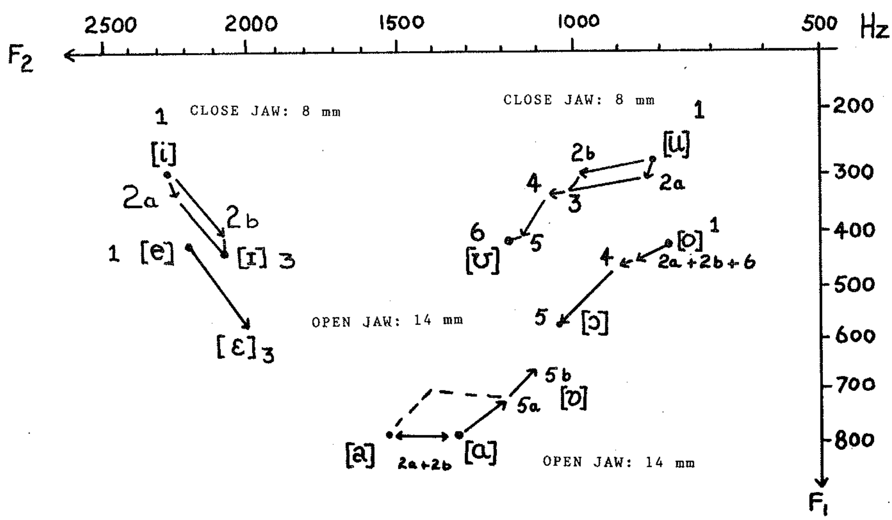


Fig. 7.
The spectral consequences of making the articulatory modifications illustrated at Figs. 5, 6 9 and 10. (1) is the initial tense configuration, (2a) retracted tongue root, (2b) lowered tongue arch, (3) the sum of 2a + 2b, (4) less depressed tongue blade, (5) less rounded lips, (6) less depressed larynx.

The same notation has been used for modifications for all vowels: (1) the initial tense vowel contour, (2a) retracted tongue root, (2b) widened constriction, (3) the sum of 2a and 2b, (4) less depressed tongue blade, (5) less rounded lips, (6) larynx less depressed by 5 mm.

Palatal constrictions

A tense [i] configuration was altered to a lax [ɪ] configuration (Figs. 5a and 6a) by lowering the tongue relative to the mandible. To avoid the necessity for compensatory movements, the same jaw-opening (8 mm) was used for both. The results were as follows (see also Fig. 7):

Retracted tongue root	F_1	+20 Hz	F_2	-20 Hz
Widened constriction	F_1	+90 Hz	F_2	-200 Hz

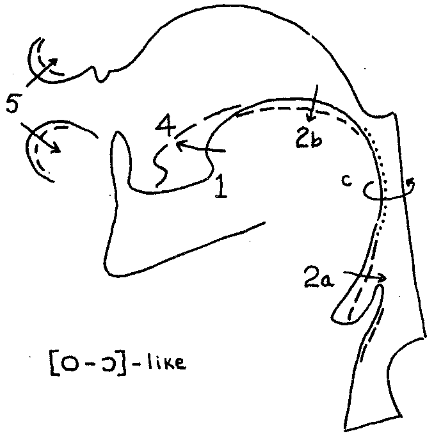
Both are in the right direction, but the contribution of the narrowed pharynx was small compared with that of the widened constriction.

The experiment was repeated for [e-ɛ], using the same tongue profiles relative to the mandible but with a jaw-opening of 14 mm. A similar result was obtained.

Palatovelar constrictions

In addition to the different constriction sizes and tongue root positions between [u] and [ʊ] there are also differences of laminal depression, laryngeal depression and degree of rounding. The jaw-opening was 8 mm for both vowels. The modifications are illustrated at Figs. 5b and 6b and the results at Fig. 7.

Retracted tongue root	F_1	+25 Hz	F_2	+15 Hz
Widened constriction	F_1	+15 Hz	F_2	+185 Hz



	2a narrower lower pharynx	2b lower tongue arch	c narrower constriction	4 less laminal depression	5 less lip rounding
F_1 Hz	+20	+7	−10~15	+10~20	+100~130
F_2 Hz	+15	+30	−70	+40~80	+100~200

Fig. 8. Modifications made to the model profile for [o,ɔ]-like vowels and the spectral consequence of each with reference to the initial configuration.

Both are in the right direction. Here too, by far the largest contribution came from the widened constriction.

Fig. 7 also shows that the sum of these two modifications (point 3) is barely half the maximum possible spectral difference. Raising the tongue blade (4) and raising the larynx 5 mm (6) made moderate contributions to F_2 (+45 Hz and +35 Hz respectively) whereas relaxing the lips slightly (5) added as much as 80 or 90 Hz to both formants.

Pharyngovelar constrictions

Fig. 8 illustrates similar modifications for [o, ɔ] -like

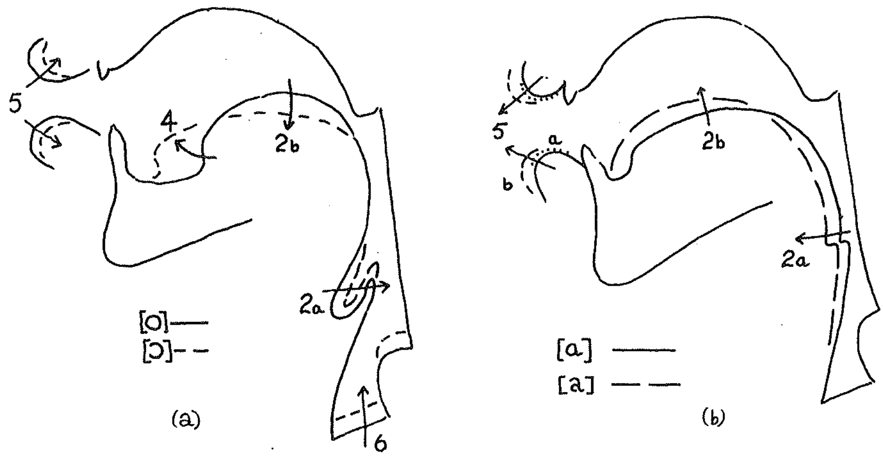


Fig. 9. Modifications made to the model profile for (a) pharyngovelar vowels and (b) low pharyngeal vowels.

vowels and gives the results. The jaw-opening was 14 mm. Modifications were made one at a time, always with reference to the same initial configuration. Both the retracted tongue root (2a) and lower tongue arch (2b) yielded small contributions. Narrowing the constriction from 1.0 cm^2 to 0.65 cm^2 lowered F_1 and F_2. Any tendency for [o] to have a wider constriction (cf. Table 2) is therefore spectrally disadvantageous to the contrast and constitutes a penalty that must be made up by some other factor (e.g. 2a+2b, 4). Less lip-rounding (5) produced a considerable spectral difference.

Figs. 9a and 10a illustrate stepwise modifications from [o] to [ɔ] with the same 0.65 cm^2 constriction for both (i.e. no penalty this time). The jaw-opening was 14 mm. The results are given at Fig. 7. Factor (5) (less lip-rounding) yielded as large a spectral difference as all the other factors (2a+2b+6+4) together.

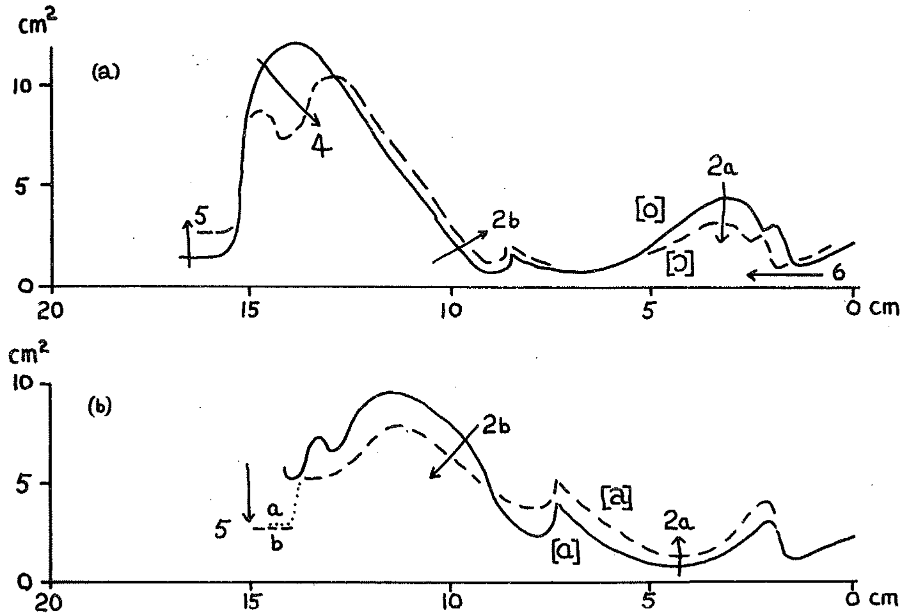


Fig. 10. Area functions for the configurations at Fig. 9(a, b).

Low pharyngeals

Figs. 9b and 10b illustrate modifications from an [ɑ] configuration to an [a] configuration. The jaw-opening was 14 mm. Spread lips (basic configuration) and neutral (5a) were applied to both since examples of both vowels with either lip position occur in natural speech. In addition, slightly rounded lips (5b) were applied to modify [ɑ] to [ɒ] . This is a grave variant of [ɑ] that occurs for Swedish /a:/ in some dialects. The results are given at Fig. 7.

The consequence of widening the constriction from 0.65 cm^2 for [ɑ] to 1.3 cm^2 for [a] was to raise F_2 by at least 200 Hz, with either spread or neutral lips. The difference between spread and neutral lips was about 80 Hz for F1 and 130 Hz for F2. Other experiments indicated that each 2 mm increment to jaw-opening adds 15-25 Hz to F1.

Discussion and conclusions

The results show that variation of the pharyngeal cavity yields a relatively small contribution to the spectral difference between tense and lax vowels. The very much larger contribution from variation of the degree of constriction is almost sufficient in itself for the spectral contrast, at least for the spread-lip vowels. In the case of the rounded vowels there is an equally large contribution from lip variation between well rounded and slightly rounded. For the pharyngovelar vowels, tongue root variation is not involved in the creation of the degree of constriction, but it is necessary to keep the lower pharynx open and thus avoid confusion with the low pharyngeal vowels. Any tendency for the tense pharyngovelar vowel to have the wider constriction means there is a spectral penalty from the point of view of this contrast.

It is also clear that the terms _tense_ and _lax_ need to be more precisely defined. In particular, the traditional notion that lax vowels have more "central" tongue positions is irrelevant and unacceptable in view of the inadequacies and inaccuracies of the tongue-arching model (Wood 1975). Are there any features that are common to all tense-lax pairs?

Fant has observed that the vocal tract is less deformed (nearer to the uniform tube) for lax vowels. As a generalization this is true, except perhaps for the [o, ɔ]-like vowels. The area functions at Figs. 6 and 10 show this resonator difference (although these are model configurations, they are the result of realistic articulatory manoeuvres based on observations of real speech). The details of how and where the vocal tract is less deformed vary from pair to pair.

Tongue root advancement and consequent pharyngeal expansion have been observed for tense vowels. This difference is most obvious for the palatal and palatovelar vowels and can be clearly seen in the examples at Figs. 1 and 2. Raphael and Bell-Berti (1975) have found corresponding differences in EMG activity in the genioglossi for American English /i-ɪ, e-ɛ, u-ʊ/. The results reported at Fig. 7 are that pharyngeal expansion contributes relatively little to these spectral contrasts whereas varying the degree of constriction yields the greatest spectral difference. However, it is generally accepted that advancing the tongue root has the secondary effect of raising the tongue body. This manoeuvre therefore also perticipates in control of the degree of constriction in this set of vowels and remains very much acoustically relevant. For the pharyngovelar vowels, the tongue root also tends to be further forward for tense [o] than lax [ɔ], widening the small cavity below the constriction. The spectral consequence of this is small (Fig. 8) but it is the right direction. There has so far been no data published regarding any correlated EMG activity in the genioglossi for this pair of vowels. For the low pharyngeal vowels, the relationshiop is reversed - narrower lower pharynx for tense [ɑ]. The advanced tongue root rule cannot apply in this case.

It is also frequently said that tense gestures are more precise and have greater extent. Regarding precision, it is fascinating to watch a motion X-ray film and see the level of precision achieved for all vowels, tense and lax. In view of the magnitude of spectral difference that can be achieved by widening the constriction, the amount of widening is critical and the ranges given at Table 1 must be respected. Regarding the extent of the tongue gestures (which are in the direction of the tongue constriction) the degree of constriction is narrower for the tense vowel in all pairs except [o, ɔ]. Figs. 1 and 2 show how the tongue is raised further towards the hard palate for tense [i, e], further towards the soft palate for tense [u] and further into the lower pharynx for tense [ɑ]. The results reported at Fig. 7 revealed that this narrowing of the constriction is the major single lingual factor contributing to the spectral contrast.

For the palatal and palatovelar vowels, the genioglossi aid the raising of the tongue body. The differences of EMG activity in tense and lax vowels reported by Raphael and Bella-Berti are therefore in a muscle that is actively involved in the basic tongue gesture of these vowels. For the palatovelar vowels, the styloglossi are also involved to draw the tongue back to the soft palate. But Raphael and Bella-Berti reported no noteworthy difference of activity between tense [u] and lax [ʊ] in this pair of muscles. For all three pairs they also reported a clear difference of activity in the inferior longitudinal muscle, an intrinsic muscle that depresses the tongue blade and helps bunch the tongue. The consequence of this can be seen at Figs. 1 and 2 for these vowels. For the rounded vowels, this can yield an F_2 difference of 100-200 Hz (Figs. 7 and 8).

The corresponding active extrinsic muscle for the low pharyngeal vowels is the hyoglossus. There are no EMG investigations reported for this muscle but I would expect more hyoglossal activity for the narrower constriction of tense [ɑ] than for the wider constriction of lax [a]. X-ray tracings for many subjects also show a more depressed tongue blade for [ɑ] indicating that the same difference of inferior longitudinal activity probably applies to this pair too (e.g. Figs. 1 and 2).

How do the pharyngovelar vowels fit into this pattern? The tongue root is more advanced for tense [o] than for lax [ɔ]. The difference recorded at Table 3 is typical, but for this pair in particular the absolute measure depends on the size of the valleculae which can vary considerably between subjects (cf. Figs. 1 and 2). The active muscles for tongue root advancement (the posterior genioglossi) are not involved in the creation of the upper pharyngeal constriction (which requires tongue retraction, not raising). It is nevertheless necessary to keep the lower cavity open in order to avoid confusion with the low pharyngeal [ɑ, a]-like vowels. The tongue root advancing manoeuvre is therefore an essential component for the pharyngovelars. The constriction itself is presumably formed by the pharyngeal constrictors (including the glossopharyngeal fibres that insert into the sides of the tongue). As for all other tense vowels, it is spectrally advantageous for [o] to have as narrow a constriction as possible. Paradoxically, [o] tends to have a slightly wider constriction than [ɔ] and therefore suffers a slight spectral penalty (Fig. 8) that is disadvantageous to the contrast. This may be due to the partial antagonism between the forward movement of the tongue root and the rearward movement of the tongue body. Finally, the tongue blade is more depressed for tense [o] than for lax [ɔ]. Here too, we should once again expect to find differences in inferior longitudinal activity.

It is therefore very likely that the physiological and articulatory difference between tense and lax vowels lies in varying the degree of contraction of a muscle that is already actively involved for a pair of vowels - such as the posterior genioglossi for the constriction of the palatal and palatovelar vowels and for keeping the lower pharynx open in the pharyngovelar vowels, and the hyoglossi for the constriction of the low pharyngeal vowels. The spectral consequences are always in the right direction for the contrast, very much so for the differences of degree of constriction, less so for the differences of pharyngeal cavity size. There are also differences of tongue blade depression and tongue bunching for all pairs which can be ascribed to differences of inferior longitudinal contraction (this may be what Sweet meant by the "convexity" of the tongue). For the palatal and palatovelar vowels the bunching aids in controlling the constriction against the roof of the mouth. For the palatovelar and pharyngovelar vowels, tongue blade depression enlarges the bucchal cavity and lowers F_2. Both these effects are favourable to the spectral contrast.

In addition, there is the difference of lip-rounding - more for tense vowels and less for lax vowels. Fig. 7 indicates that this can account for half the spectral contrast between [u] and [ʊ] and between [o] and [ɔ]. This is coupled to similar differences of laryngeal depression. The spectral consequences of this are relatively small (Fig. 7 and Lindblom and Sundberg 1971) but they are in the right direction.

Notes

1. See Fant (1960). I am endebted to Professor Gunnar Fant and Dr. Johan Sundberg of the Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, for permission to use LEA and for assistance.

2. These films were made at the Röntgen Technology Unit of the University Hospital with the consent of Professor Olof Norman and the assistance of Dr. Thure Holm, Radiophysicists Gunnila Holje and Gudmund Swahn and Technician Rolf Schöner. The angiological laboratory was specially equipped for observing events in soft tissue such as blood vessels, and was therefore admirable for our purpose. In addition, the camera provides a synchronizing pulse that flashes on every tenth frame and which also appears alongside a patient's cardiogram. We recorded this pulse on magnetic tape together with the speech signal, on separate tracks. The film speed was 75 frames/second. X-rays were emitted in brief bursts, 3 msec per frame, which kept the radiation dose within the range 60-200 mrad per reel of film. Each subject was limited to one reel (40 seconds). I am endebted to Gösta Bruce and Per Lindblad for permission to include their films.

3. This summary is of necessity very scanty. More details will be given in a forthcoming thesis on the articulation of vowels.

References

Fant C.G.M. 1960. The acoustic theory of speech production. The Hague

Gunnilstam O. 1974. The theory of local linearity. Journal of Phonetics 2:91-108

Halle M. and K.N. Stevens. 1969. On the feature Advanced Tongue Root. QPR 94:209-215. MIT Research Lab. of Electronics

Lindau M. 1975. Vowel features. Working Papers No. 11, Phonetics Laboratory, Department of Linguistics, University of Lund

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

Lindblom B. and J. Sundberg. 1971. Acoustical consequences of lip, tongue, jaw and larynx movements. Journal of the Acoustical Society of America 50:1166-1179

Perkell J.S. 1971. Physiology of speech production: a preliminary study of two suggested revisions of the features specifying vowels. QPR 102:123-139. MIT Research Lab. of Electronics

Raphael L.J. and F. Bella-Berti. 1975. Tongue musculature and the feature tension in English vowels. Phonetica 32:61-73

Stevens K.N. and A.S. House. 1955. Development of a quantitative description of vowel articulation. Journal of the Acoustical Society of America. 27:484-495

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

Sundberg J. 1969. On the problem of obtaining area functions from lateral X-ray pictures of the vocal tract. Appendix to Articulatory differences between spoken and sung vowels in singers. STL-QPSR 1/1969, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm

Sweet H. 1906. Primer of phonetics. 3rd edn. Oxford

Wood S. 1975. The weaknesses of the tongue-arching model of vowel articulation. Working Papers No. 11. Phonetics Laboratory, Department of Linguistics, University of Lund