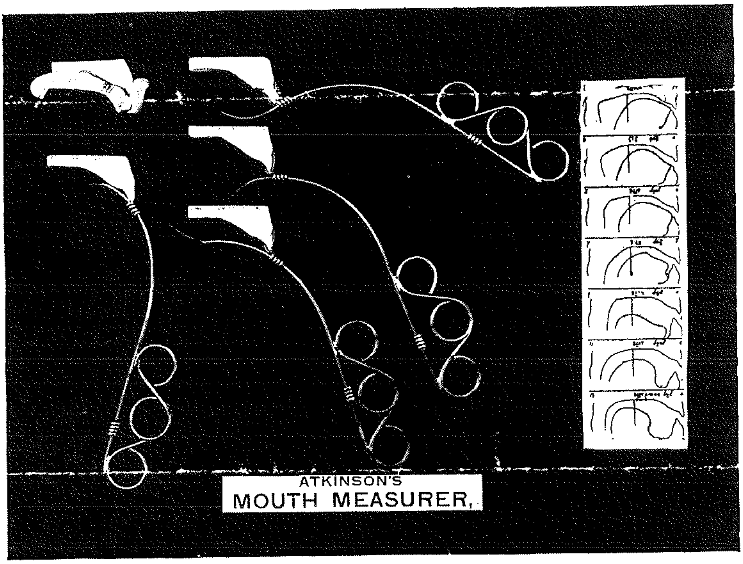
CONTENTS

ERRATUM

line	for	read
p 5, l 3 from bottom	Ghosh 1838	Ghosh 1938
p 21 bottom line	Liberman	Lieberman
p 37 bottom		add: waves are found.
p 152, l 5 from bottom	[i] and [e]	[ɪ] and [e]
p 155, l 3	[ɪ, e]	[i, e]
Bibliography	add:	Kaneko N (1957). Specific x-ray observations on the movement of the lower jaw, lips and tongue in pronunciation. *Study of Sounds* 8: 1-18 and 450-451. Tokyo

Atkinson's Mouth Measurer.

FOR THE PRACTICAL STUDY OF PHONETICS.



The two instruments are about 7 inches = 17.5 cm. long, and are made of nickel-plated German silver.

Place three fingers through the rings, then the coil of wire below can be moved by the thumb, thus pushing out or drawing back from the other end of the long narrow tube a fine wire sliding inside it. In one instrument the wire comes out in a curve following that of the tube, in the other in a reversed curve. Each of the four tooth stops can be used on either instrument. They can be moved to different positions on turning them to the inside of the curve of the tube.

The instrument is placed in the mouth and the wire pushed forward until the end just touches the tongue. This point is recorded by laying the instrument against a diagram, wooden model or plaster cast of the teeth and hard palate. From other positions of the tooth stop, or by using other tooth stops, a series of points is obtained showing the shape of the tongue. The shapes of the soft palate may be similarly measured. *Pronounce the sound when taking a measurement.*

To obtain the shape of the teeth, hard palate, and front of the gums, an impression is taken with the moulding composition supplied, softened in water warmed to not more than 140 deg. F. = 60 deg. C. Two or three impressions should be taken and the results compared to ensure that the shape obtained is correct. A piece of card or wood is cut to fit this, or a plaster cast is prepared from it. (See top left-hand corner of photograph.) The plates of the tooth stops should be bent to fit the teeth. The tube may be bent a bit with the fingers to suit the height of the palate. The position of the end of the hard palate must be determined and the card, wood or plaster cut to the correct length.

The relative positions of the jaws may be measured by pressing a piece of softened composition against the upper and lower teeth. The shape of the lower jaw may be obtained in the same way as that of the upper, and included in diagrams as in the examples shown.

Dr. Henry Sweet, University Reader in Phonetics, Oxford .

"Mr. Atkinson's apparatus is, as far as my knowledge goes, the best that has yet been devised for determining objectively the position of the tongue in forming the different speech sounds."

Mr. Daniel Jones, Lecturer on Phonetics at University College, London, and Assistant Editor of the "Maître Phonétique" :—

"I have made frequent use of Mr. Atkinson's 'Mouth Measurer' for determining the tongue positions of vowel sounds, etc., and find it gives excellent results. It is a most ingenious little instrument, and should prove of great value to students and teachers of phonetics."

Mr. Bernard MacDonald, Lecturer in Voice Production and Elocution at Mansfield College, Oxford; Queen's College, London, etc. :—

"I wish to tell you how very useful I have found your instruments for taking mouth measurements. By their use I have been able to clear up many previously doubtful points as to vowel positions. The instruments have been invaluable to my work."

Prof. D. L. Savory, Prof. of French and Romance Philology at the University of Belfast :—

"I consider that the instrument invented by Mr. Harold W. Atkinson, M.A., for measuring the position of the tongue in the formation of speech sounds is exceedingly valuable for the purpose of scientific research, and further as a demonstration apparatus which should be in the hands of every teacher of modern languages. He would then be in a position to show clearly to his pupils the differences in articulation of English, French and German sounds, and such a practical demonstration is infinitely more valuable than any kind of theoretical description."

X-RAY AND MODEL STUDIES OF VOWEL ARTICULATION

1. INTRODUCTION

The basic principle of Alexander Bell's high-low front-back tongue arching model (which introduced the revolutionary invention of a class of central vowels) is still firmly established in phonetics and phonology despite severe criticism. There was never any real opportunity to test the physiological basis of the model before the introduction of radiology in the 1890s provided the first real 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 brings in question the *predictive capability* of the model. The tongue does not always do what the Bell model says it should do. Current knowledge of vocal tract acoustics and the neuromotor level of speech production show that the theory of articulation underlying the Bell model is not only inaccurate but also so irrelevant to the processes of speech production that the model lacks *explanatory power* and *physiological realism*.

There has always been a school of phonetics throughout this same period that has been sceptical of articulatory descriptions of speech. Some, especially towards the end of the 19th century when there was bitter rivalry between "acoustic" and "organic" schools, rejected articulation altogether as a legitimate area of study.

X-ray tracings of vowel articulations are examined here in the light of the criticisms of the Bell model. These pictures corroborate that the Bell model fails to prescribe tongue position correctly and to relate articulation to the spectral characteristics of the speech wave. The implications of this failure are discussed. The constancy of vocal tract configurations in relation to standing wave phenomena, compared to the ambiguity and acoustical irrelevance of the tongue arch position, points to a more suitable type of model in which physiologically relevant manoeuvres are coordinated to shape the vocal cavities, tuning the vocal tract resonances for the desired spectral contrast.

The five papers presented here explore various aspects of such a model, demonstrating how the articulatory manoeuvres analysed from x-ray films can be related back to speech programming and forward to the encoding of the speech wave. The acoustical consequences of these manoeuvres are investigated in model experiments by manipulating mid-sagittal profiles to simulate the manoeuvres and then calculating and evaluating the vocal tract resonances.

Papers I and II are concerned with constriction locations for vowels. Paper I reports experiments on the sensitivity of palatal and velar vowel spectra to variation of the constriction location. In Paper II, the constriction locations used for vowels are analysed from x-ray profiles. Monophthong vowels are divided into four classes by constriction location - palatal, velar, upper pharyngeal and lower pharyngeal. Acoustical, physiological and phonological aspects of constriction locations are discussed.

Papers III and IV are concerned with how vowels are differentiated at one of the relevant constriction locations - the palatal vowels. Paper III reports a series of model experiments where the contribution of each individual articulatory component of rounded palatal vowels is assessed. While lip rounding generates the phonological contrast, other components, such as larynx depression, stabilize resonance conditions and ensure that the resonances respond similary to lingual variation in both rounded and spread-lip palatal vowels. Physiological implications are discussed, especially compensation and motor control. In Paper IV, jaw, lip and tongue manoeuvres in spread-lip palatal vowels are analysed from x-ray profiles. Palatal vowels are divided into open and close classes by mandible position and tense and lax classes by tongue posture. The lips are spread more for tense than for lax vowels.

Paper V deals with an example of one particular articulatory contrast - tenseness and laxness - across all vowel classes (palatal, velar, upper pharyngeal and lower pharyngeal).

This thesis overview outlines the historical background to the Bell model, gives details of x-ray and modelling procedures, summarizes the five reports and presents some general conclusions.

2. THE HISTORICAL BACKGROUND

The tongue arch model

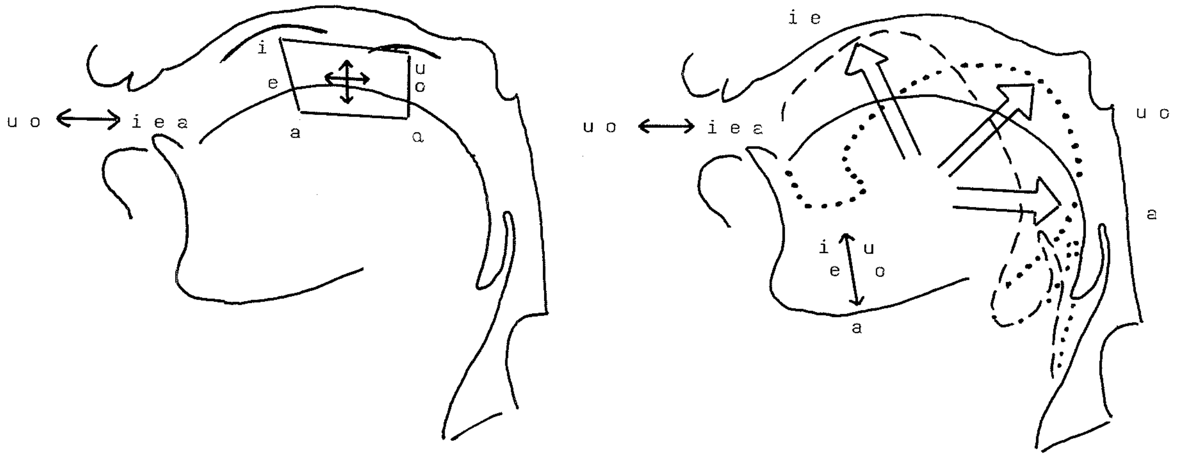
Alexander Melville Bell (1867) recounts how, after much puzzling over the articulation of the vowel of *sir*, he got the idea of the tongue not only being raised to the hard or soft palate but also medially between them. The tongue arching model, as Russel (1928) called it, was born and with it the revolutionary innovation of a class of vowels with alleged central tongue position. Publicized by among others Sweet (1877), the model was rapidly adopted by the new movements that dominated the latter part of the century - the neogrammarians, the language teaching reformers and the IPA.

Much of the terminology of Bell's original version has been discarded (*narrow-wide, inner-outer* etc.) , but *high* and *low* are still used. The basic principle of the model remains unchanged: a vowel is defined by the position of the tongue arch in the mouth. Each vowel has its own position. Raising, lowering, advancing or retracting the tongue yields a new vowel. See Fig. 1a. This principle has provided the main framework for vowels in phonetics and phonology for more than 100 years. The model is probably more familiar today in some more recent version, like Daniel Jones's (1967, nine revised editions and numerous reprints since 1918).

To avoid confusing Bell's original version with the general class of model, I shall refer to the former as the *Bell model* and the latter as the *tongue arch model*.

The ancient palatal-labiovelar- pharyngeal-jaw type of model

The Bell model displaced the hitherto current model that had largely remained unmodified since before the times of Pāṇini. The ancient Indians described vowel production in terms of



(a) The tongue arch model

(b) The ancient model

Fig. 1 The tongue arch model assigns a position of the top of the tongue arch to each vowel. Raising, lowering, advancing and retracting the tongue arch anywhere within the quadrilateral yields a new vowel. The ancient model recognized three basic tongue manoeuvres (palatal, velar, pharyngeal), jaw position and lip position.

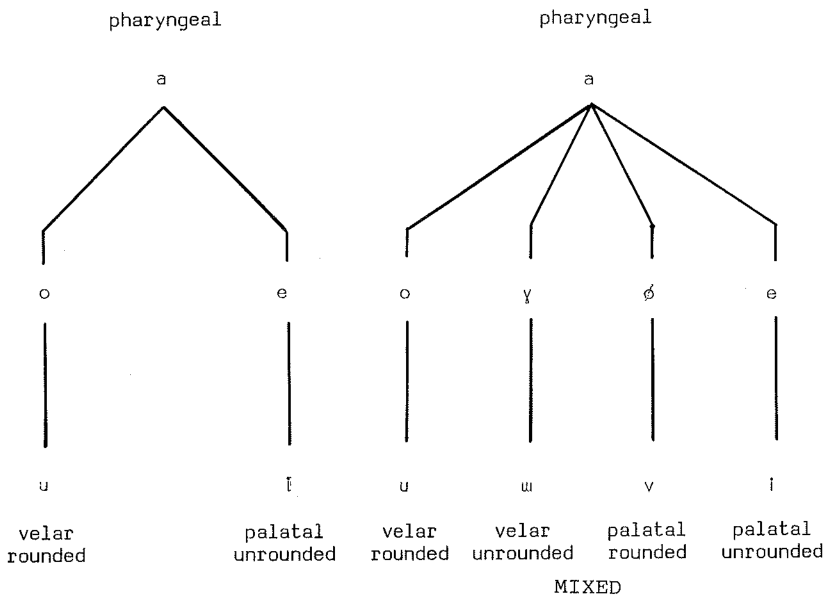


Fig. 2 In the early years of the 19th century the ancient model was commonly portrayed as trees with palatal and velar series branching off from a basic pharyngeal configuration. Vowels in each branch are differentiated by open vs close jaw position. The mixed branches combine the spread lips of the palatals with the velar tongue manoeuvre and the rounded lips of the velars with the palatal tongue manoeuvre.

pharyngeal [a]-like vowels, palatal [i-ɛ]-like vowels and labiovelar [u-o]-like vowels (Varma 1929, Allen 1953). See Fig. 1b. These correspond to three of the four relevant constriction locations reported in Paper II.

Through Budhism and cultural contact with Islam, sanskrit grammar, and with it the vowel model, spread east to China and Japan and west to Arabia (Staal 1972, Fleisch 1957, Gairdner 1935). Examples of the long tradition of the ancient model in phonetics are the 6th and 5th century BC Indian recitation manuals (Whitney 1862, 1871, Regnier, Ghosh 1838), the Roman grammarians Terentianus Maurus and Marius Victorinus (Keil Vol. 6), Arab grammarians such as Avicenna (Ibn Sīnā) and Ibn Ginni (Bravmann

1934, Semaan 1963, Fleisch 1958). In post-renaissance Europe there are examples like John Hart (1569, Danielsson 1955) and Jacob Madsen af Aarhus (1589, Møller and Skautrup 1930) and later still Hellwag (1781). Bell had himself, only a year or two before in (1863), still retained this ancient model in a new edition of a handbook first published in 1849.

The two types of model

How was it possible for the ancient model to be ousted so abruptly in the latter half of the 19th century? Most likely because the time was ripe. Philologists and Christian missionaries were having increasing difficulty in accommodating within the framework of the ancient model the unfamiliar vowel qualities that were being dicovered in languages around the world. The Bell model, with its completely new category of central vowels appeared far more attractive. Controversy between schools supporting the rival models lingered on but the ancient model hardly survived into the 20th century. The rare exceptions were new editions of earlier works, mainly books of a technical or therapeutical rather than linguistic nature. For example, Helmholtz (1863) had referred to the ancient model and a 6th edition appeared in 1913.

There is a fundamental conceptual difference between the two types of model. The ancient model recognized distinct palatal, velar and pharyngeal tongue manoeuvres. In the early years of the 19th century the model was commonly portrayed in the form of a tree (not a triangle) with velar and palatal series branching off from a basic pharyngeal configuration (Fig. 2).

Since Hellwag's treatise (1781) the tree had been augmented with additional branches, first for rounded palatals and later also for unrounded velars. The insertion of these extra branches between velar and palatal never implied intermediate tongue positions between front and back. The additional branches were known as *mixed* because they combined the tongue manoeuvre of one basic branch with the lips of the other (rounded or unrounded). Much later Bell (1867) used the term *mixed* to denote the *central* tongue positions of his new model. As the typographical layout of the trees became more complex they came to look like pyramids of ABC blocks. The visual impact became triangular, but with [a] still at the apex.

In contrast, central tongue positions between front and back are explicit in Bell's model. In addition, the front, central and back positions could be advanced or retracted, making 9 x 9 classes in all. It was this feature, alien to the ancient model, that made it so attractive by providing a rich system of matrix cells for finer or unusual contrasts.

The Bell model also seemed to tie up with acoustics to build a comprehensive theory. It had long been recognized that vowel quality had something to do with resonance. The tongue was at first seen as the limit of the buccal cavity in which a characteristic vowel tone was formed. For example, Isaac Newton (Elliott 1954) had likened it to the note produced by a flagon being filled with water. After Helmholtz's (1863) discovery of at least two vocal tract resonances, it was believed that the tongue arch constituted a neck between a buccal and a pharyngeal cavity, each with its own resonance. Tongue *height* and *retraction* were said to regulate the resonance of each cavity. One spoke of the *mouth formant* and the *throat formant*. The tongue arch model fitted neatly into this theory.

Now that the acoustics of the vocal tract are much better understood (Chiba and Kajiyama 1941, Stevens and House 1961, Fant 1960, 1980, Stevens 1972) it is clear that this role of cavity boundary assigned to the top of the tongue arch was a misconception. There are several resonances and for each one there is a standing wave that fills the entire vocal tract. The frequency of a resonance can therefore be modified by manoeuvres in various parts of the vocal tract (by the lips, the tongue blade, the tongue body, the tongue root and sphincter muscles) in relation to the nodes and antinodes of the standing wave. The three tongue manoeuvres and the jaw and lip positions of the ancient model fit more closely to modern knowledge and theory than the tongue arch model ever did.

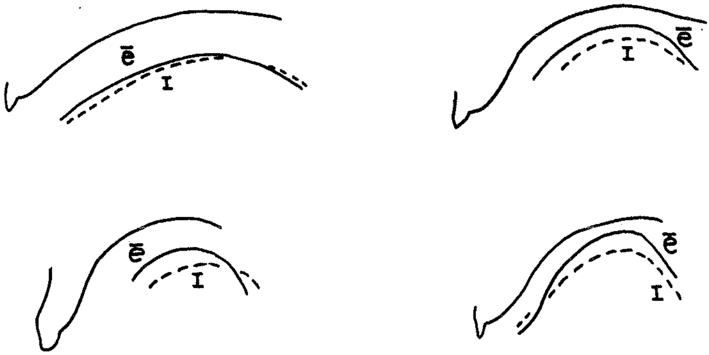
The crisis

From the 1860s to the turn of the century there had been no way of testing the Bell model. It was discredited at the first real test (Meyer 1910). It never was validated. On the contrary, x-ray studies have continuously revealed tongue positions that are anomalous with reference to tongue arch positions.

Grandgent (1890) had devised a novel method of fitting different-sized discs into various parts of the vocal tract to measure its cross-section. He was one of the first to point out the large pharyngeal cavity of palatal vowels. The numerous repetitions of a vowel articulation necessary for his method meant that measurements were very coarse and they did not therefore show up the anomalous tongue positions later reported from x-ray investigations. Atkinson (1898) had used a similar probing method.

Meyer (1910) suspended fine strips of metal foil from a false palate. These were then deformed by the tongue and retained an imprint of its contour. He found that the tongue was lower for lax /ɪ/ than for tense /e:/ (German, Dutch and Swedish subjects), contrary to the predictions of the Bell model. This was the first evidence against it. Vietor (1914) agreed it showed earlier ideas about tongue articulation were erroneous, but Chlumsky (1913) had less success with Meyer's method and was more cautious.

The first x-ray inspection of vowel articulation had been performed before the turn of the century (Scheier 1909). Meyer (1907) had taken the first photographs of vowel profiles from the x-ray screen. Both authors used tense German vowels and missed the opportunity to observe the heights of /e:/and /ɪ/. But Kruisinga (1925) noted that /e:/ and /i:/ were equally high on Meyer's radiograms. Russel (1928) took his first radiograms in order to demonstrate the tongue arch model to his students, but failed to obtain a convincing set of tongue positions. After taking several thousand radiograms from over 400 subjects he concluded the model was fallaceous. In addition to [ɪ] being lower than [e] he also found that [ɔ] was often lower than [a]. Carmody (1937) also reported irregular tongue positions in Holbrook's radiograms (for example, that "low back vowels depend mostly on lip positions 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". Unlike Russel, he continued to believe in the model.



American English German

Fig. 3 Four cases of lax [ɪ] with lower tongue height than tense [e].

The evidence of x-ray profiles

38 sets of x-ray tracings of vowel articulations in 15 different languages have been collected from the literature (for details see the section on materials and procedures). They have been published at various times since the beginning of the century and constitute the evidence that has been available during that period. What do they show?

Meyer's and Russel's criticisms had mainly concerned tongue height, expecially that the tongue was lower for [ɪ] than for [e] and lower for [ɔ] than for [a] . This is confirmed by the collection of profiles and there are even examples of [o] lower than [a]..

The problem of [ɪ] and [e] occurs in languages with four palatal quality contrasts [i,ɪ,e,ɛ] (phonemes or allophones). American English and German are the only languages in the collection with examples of all four vowels. In all 7 possible sets /e/ was higher than /ɪ/. Four examples are illustrated at Fig. 3. The higher /e/ than /ɪ/ has also been reported by Ladefoged et al.

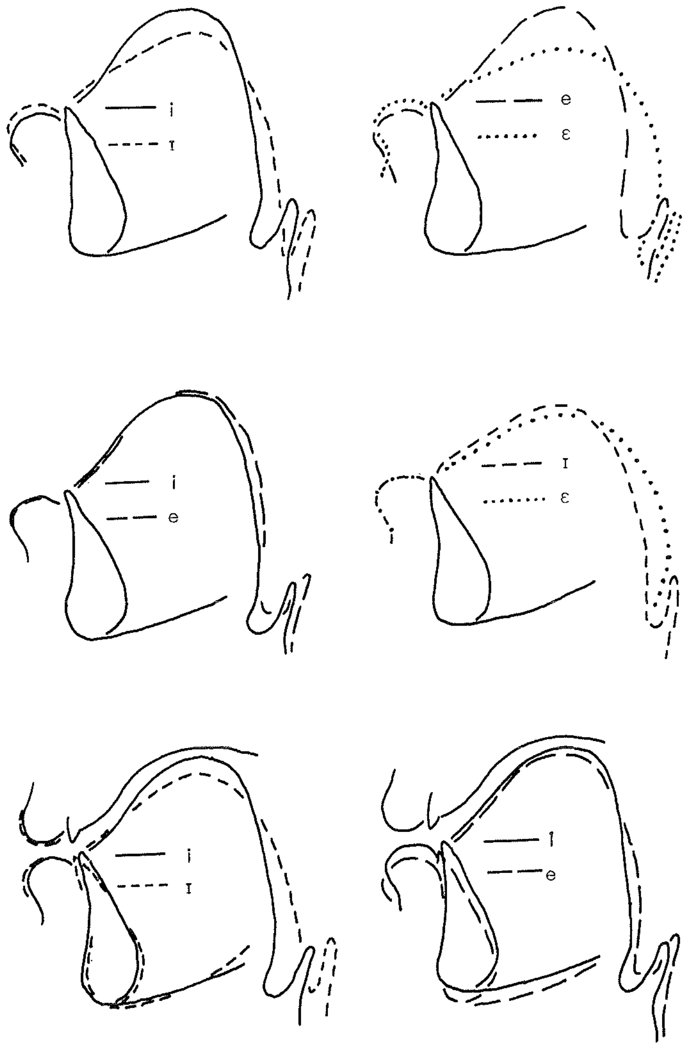


Fig. 4 Tongue posture and jaw openings in [i ɪ e ɛ] (Chiba and Kajiyama's German) showing tense tongue and lip postures for [i e] and lax for [ɪ ɛ], and higher jaw position for [i ɪ] and lower for [e].

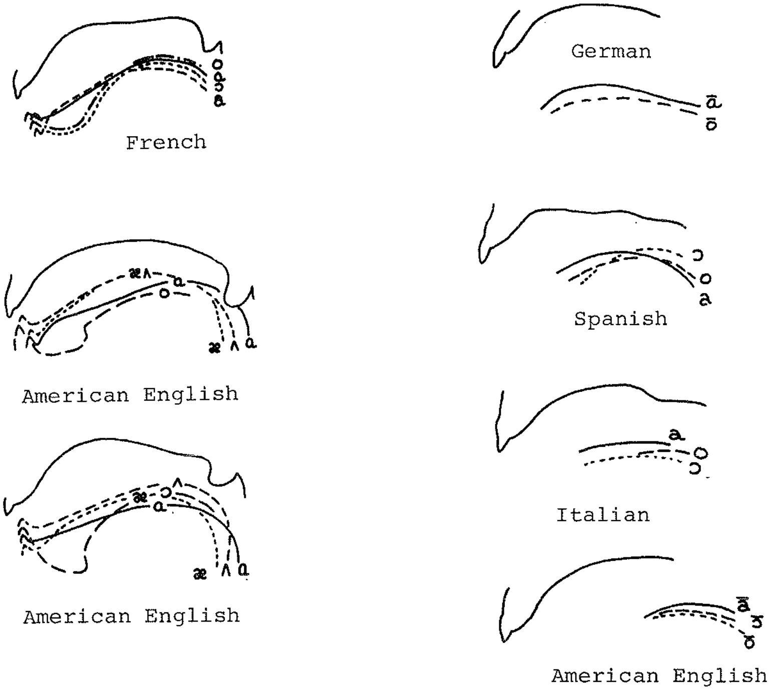


Fig. 5 Confusing tongue heights of so-called mid and low back vowels in various languages.

(1972) for 4 out of 6 subjects (and 1 with the same height and 1 lower).

One might think that the failure of the tongue arch model to get the tongue heights right for [e] and [ɪ] was an easily rectifiable mistake. These vowels need only be put in their proper places. But this would be a superficial remedy that solved nothing. As Jesperson (1913 p.54) put it: "Die ganze Theorie der Vokale ist also zurzeit ins Schwanken gebracht" (after reporting Meyer's results in a revised 2nd edition of his textbook).

	[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)	1	1	1		1
English (S Br)	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	

Table I. Comparison of tongue heights of [o ɔ ɑ a]-like vowels from a large number of sets of published x-ray tracings.

Inverting the tongue heights of [ɪ] and [e] in the model would still not help it capture the true articulatory relationship between "tense" and "lax" palatal vowels. As an example, Chiba and Kajiyama's (1941) German profiles are analysed in Fig. 4. They show that there is a typical tense posture of the tongue relative to the mandible for [i,e] and a typical lax posture for [ɪ,ɛ].. These postures are combined with a narrower jaw opening for [i,ɪ] and a wider jaw opening for [e,ɛ] .

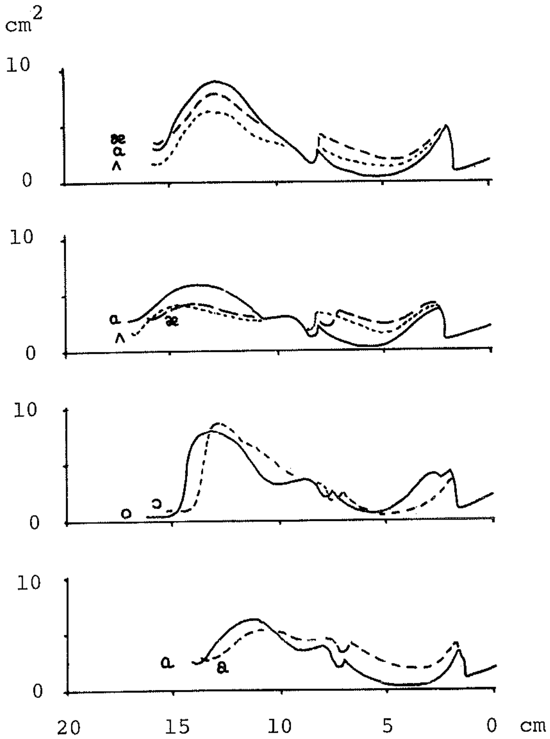


Fig. 6

Area functions of the English and French vowels to the left in Fig. 5 showing that [o]-like vowels have a constriction a little higher in the pharynx than the [a]-like vowels.

This relationship is examined more closely in Papers IV and V in this thesis.

The balance of combining the flatter lax tongue posture and close jaw position for [ɪ] and the bunched tense tongue posture and open jaw position for [e] is that "tongue height" usually comes out lower for [ɪ].

The second situation, conflicting tongue heights for [ɔ] and [a] was reported for languages with two quantitatively different sounds [o] and [ɔ], whether allophones or phonemes. It is the "lower" vowel [ɔ] that was reported to be lower than [a]. Figure 5 illustrates examples of confusing tongue heights for these and related vowels. Ladefoged (1964) has similar examples of confusing tongue heights among back vowels in African language material.

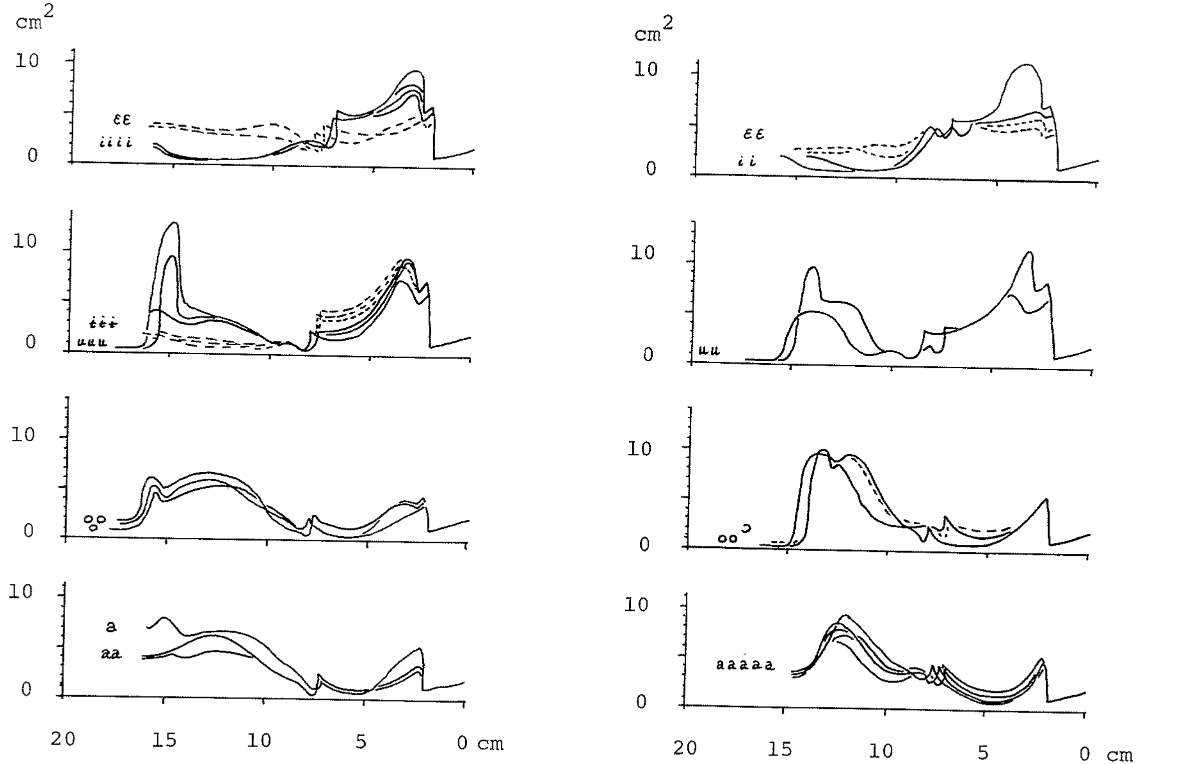


Fig. 7 Area functions: Russian (Koneczna and Zawadowski) and Spanish (Parmenter and Trevino)

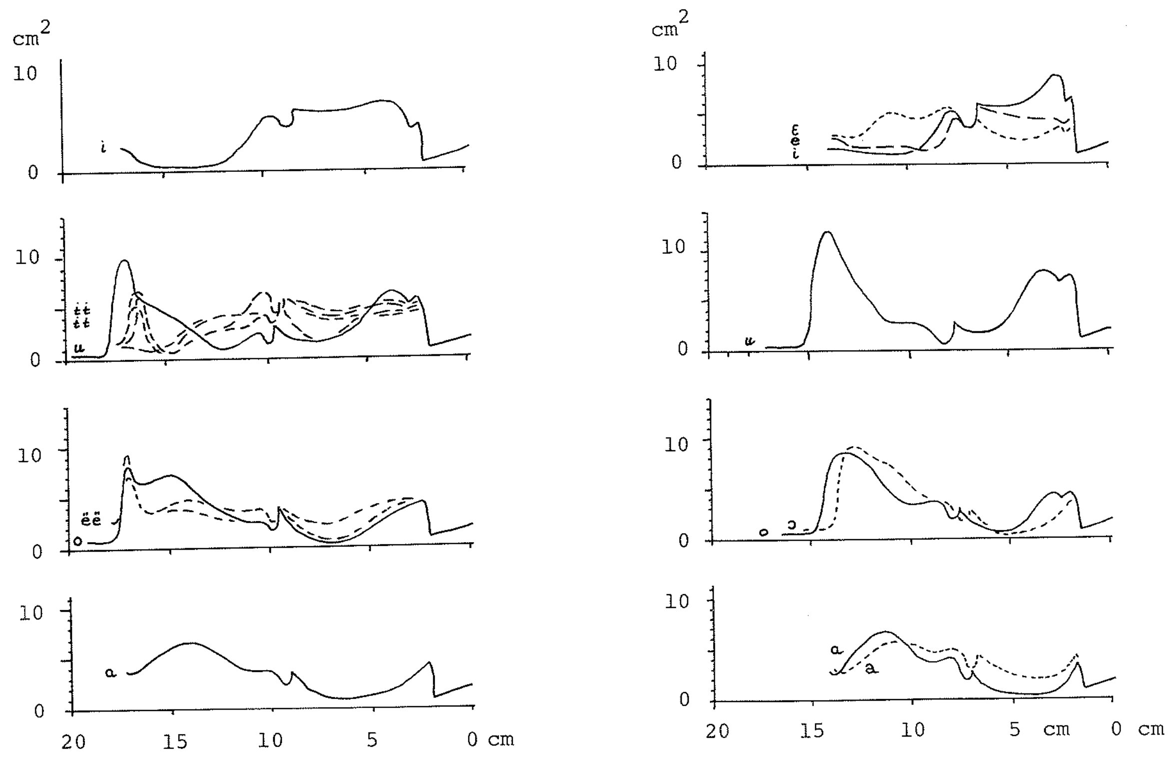


Fig. 8 Area functions: Chinese (Ohnesorg and Svarny) and French (Holbrook)

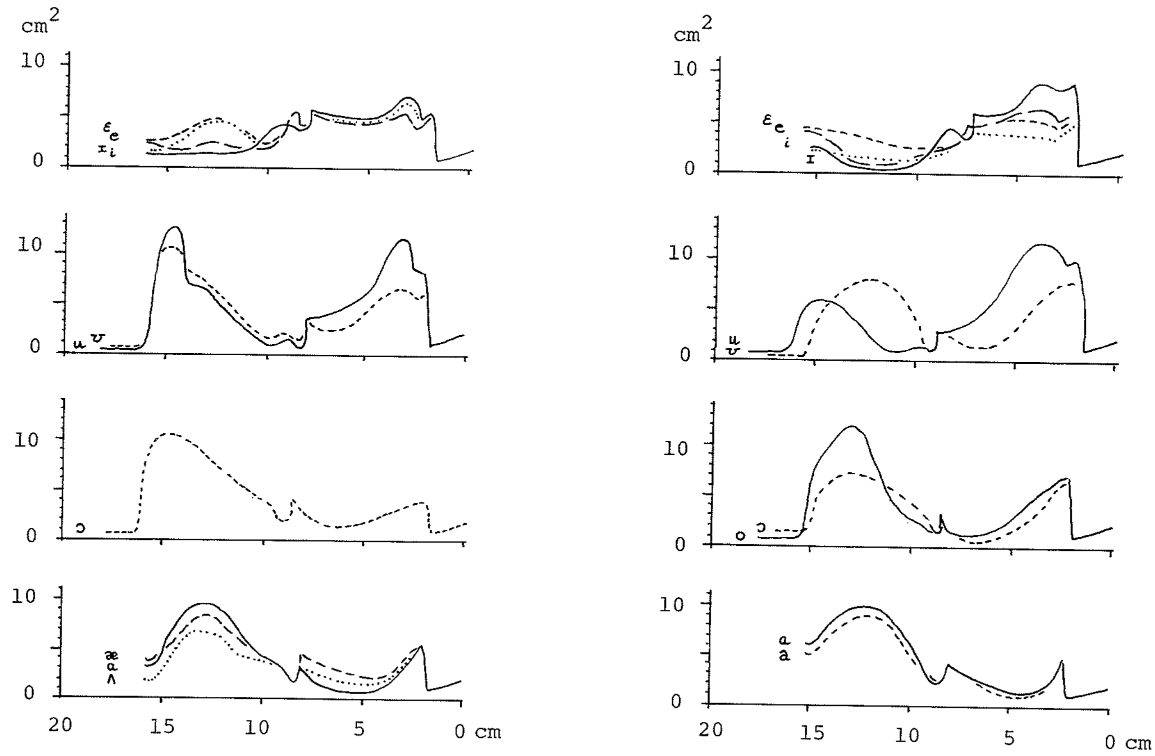


Fig. 9 Area functions: American English (Holbrook) and German (Wängler)

The relative heights for all examples of [o], [ɔ] and [ɑ, a] in the collection of profiles are presented in Table I.

Notwithstanding the random character of the tongue heights of [o]-like and [a]-like vowels, comparison of the area functions of these vowels shows that [o] and [ɔ] always have a place of constriction a little higher in the pharynx than [ɑ, a] while [ɑ, ʌ, a, æ] have a common low pharyngeal constriction location and are differentiated by having a progressively wider pharyngeal constriction from about 0.5 cm^2 for [ɑ] to about 2.5 cm^2 for [æ]. Some examples are given at Fig. 6.

This regularity of constriction location is the subject of Paper II in this thesis. The tongue and the superior pharyngeal constrictors narrow the upper pharynx for [o, ɔ] while the tongue and the middle and superior pharyngeal constrictors narrow the lower pharynx for [ɑ, a, æ]. Whatever the orders of tongue heights in sets of vowels, examination of all the sets in the collection shows that this relation of constriction locations for the two vowel classes is a very strong constant of vowel articulation. Some examples are illustrated in Figs. 7 to 9.

The x-ray pictures also show that the lower tongue height for [o, ɔ] than for [ɑ, a] is the natural result of larynx depression for [o, ɔ] (the larynx is regularly lower for rounded vowels than for spread-lip vowels, lengthening the pharynx by up to 15 mm, see Paper III).

The reports of anomalous tongue heights were therefore correct. The randomness of [o,ɔ,ɑ,a] tongue heights (Table I, Fig. 5) compared with the constancy of vocal tract configurations (Figs. 6 to 9) shows that tongue height is hardly a physiologically relevant parameter of vowel articulation. The physiological aspects are discussed in Paper II.

Reactions to the crisis

Although the tongue arch type of model has been discredited for more than half a century, it has never been completely disavowed. It still occupies a strong position for both teaching and research as well as for phonology as a glance through phonetics and linguistics journals and manuals will show.

Meyer's and Russel's results were nevertheless embarrassing and reactions were varied. Some rejected the tongue arch model outright (in some cases rejecting articulation too), some still believed in it with unbroken faith, most just continued to use it (with perhaps an apologetic disclaimer).

Experimental design and methods were queried. Chlumsky (1913) had failed to reproduce Meyer's results. Others feared that contrast chains and sustained utterances distorted the articulation of x-ray subjects despite the reassurances of practitioners like Russel or Gutzmann (1930). There were even public demonstrations by S Jones (1929), who pronounced the name of the Welsh village

Llanfairpwllgwyngyllgogerychwyrndrobwllllantisiliogogoch

with one silver chain along the tongue and another through the nose and down over the velum.

There had been opposition to the Bell model right from the beginning, not only from supporters of the ancient model as would be expected but also from those who insisted that since speech consisted of sounds it should only be described in acoustic or auditory terms. Lloyd (1890) deplored the rivalry between the "organic" and the "acoustic" schools. He pleaded "it is evident to a dispassionate observer that there is no true place for partisanship, that neither line of investigation ought rightly to exclude or overlook the other, but that each is necessary for the other's completeness". The acoustical tradition has continued, gaining impetus both from advances in instrumentation and the discrediting of the Bell model.

Russel believed that the acoustic school's impressionistic analysis of speech was the better way. Russel's book (1928) was itself an assault on the Bell model and he continued to attack the tongue arch type of model at every opportunity (e.g. 1935). As an alternative he proposed impressionistic features describing sound quality.

Joos (1948) insisted that phoneticians who believed they could feel tongue positions kinesthetically were deceiving themselves. On the basis of spectrographic evidence he held that they were really judging the vowels by auditory sensation. Further evidence is given by Ladefoged (1967, chapt. 2.). Subjective judgments of height are related to the frequency of F_1 and judgments of advancement or retraction to the frequency of F_2 (or to the difference between F_1 and F_2) (Joos 1948, Delattre 1951, Ladefoged 1975). This has been incorporated in the feature system developed by Lindau (1978).

With improved instrumentation, attention turned to gathering data on the acoustical structure of speech segments in different languages, to discovering the coding of the speech signal and to exploring speech perception. One product of this was the set of acoustic features of Jakobson et al. (1952) based on perceptually relevant spectral characteristics of speech segments.

There were others who still believed in the model despite the evidence. For example, Jespersen (1913) revised his textbook to report Meyer's results but he did not modify the model. He felt that there was a subjective affinity between [i] and [ɪ] and that [ɪ] consequently belonged between [i] and [e]. Carmody (1937) had himself reported anomalous tongue heights but he defended the model firmly. He dismissed criticism because it came "unfortunately from teachers acquainted with phonetics only at second hand".

The main reason for not giving up the tongue arch model as easily as the ancient model had been discarded in the 19th century is the lack of a substitute. As Ladefoged (1971) puts it: "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". Authors were not deaf to the crisis and they often acknowledged their awareness of the problem. One cannot help feeling sorry for Bloomfield (1969) who after giving the following neat and correct account of vowel production, "The differences between [vowels] seem to be largely differences of tongue-position, and to consist, acoustically, of differences in the distribution of overtones", went on to add, "Even these principles are disputed".

The tongue arch model has continued to provide a convenient abstract classifying system that fulfills a foremost requirement during this period. Any set of features would have done equally well. A classifying system is not in itself affected by factual errors in the classifying labels providing the categories remain intact. Scholars, whose only requirement has been a classification system, have been able to continue, deaf to the theoretical crisis surrounding tongue articulation.

3. THE WEAKNESSES OF THE TONGUE ARCH MODEL

The published x-ray tracings confirm the anomalous tongue positions that contradict the tongue arch model. Close [ɪ] is lower than half-close [e]. The heights of half-open [ɔ] and open [a] are random. In only two thirds of the cases was half-close [o] higher than open [a]. Tongue arch position is ambiguous with regard to resonator shaping and consequently to the spectrum of the vowel generated in a particular configuration. The vocal tract configuration is dependent on a number of other factors, information on which is not readily available from the tongue arch position.

The tongue arch model is *physiologically unsound* since it was based on a misconceived idea 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 pharyngeal constriction for [a]-like vowels and its significance was underlined by Russel. The shaping and acoustical function of the pharyngeal cavity are outside the domain of the tongue arch 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 narrowing of the lower pharyngeal cavity. The meaning of such secondary concepts as "uvularization" and "pharyngealization" is not clear. The extrinsic muscles of the tongue (which do the major share of tongue shaping and positioning in vowels) all contract in the pharyngeal region, and whatever other task these muscles may be performing they always immediately and directly alter the pharyngeal cavity (Fig.10).

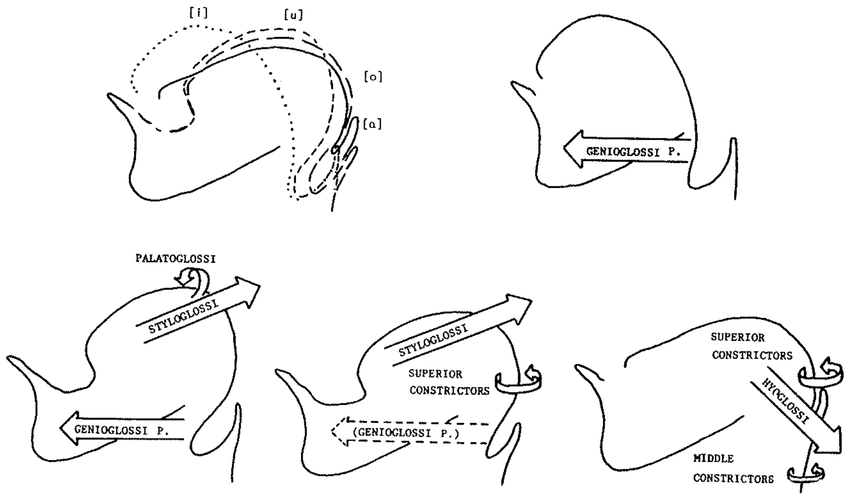


Fig. 10 The direction of contraction of the extrinsic muscles of the tongue and of the pharyngeal constrictors, arranged according to their activity for the formation of the four constrictions

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 were hampered by their inadequate anatomical knowledge. Avicenna (Ibn Sīnā, text in Bravmann 1934 and Semaan 1963) had made a detailed description of the muscular structure of the tongue, but he was unable to relate it to the tongue manoeuvres he observed during vowels*. Yet these manoeuvres can be related to the muscles of the tongue. It is emphasized in Paper II that the extrinsic musculature of the tongue is ideally

* This was because he, or Galen, had dissected the tongue of the ape and not of man (Singer 1957). 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). A necessary element of the phylogeny 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. This permits a wide range of cavity-shaping manoeuvres by the tongue, which has been stressed particularly by Liberman (1972).

situated both for making the basic constrictions for vowels and for other local widening and narrowing manoeuvres in relation to the standing waves.

In contrast, the anatomy of the tongue was well known in the 19th century and this knowledge was available to Bell. Since at least the treatise of Hellwag (1781), there has been virtual agreement about the role of the extrinsic muscles of the tongue for directing lingual manoeuvres. The hyoglossi draw the tongue bodily downwards to narrow the lower pharynx (Fig. 10). The posterior genioglossi pull the tongue root forward to widen the pharynx and push the tongue body up and forward. The glossopharyngei (fibres of the superior 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 palatoglossi have a sphincter-like action that narrows the faucal isthmus (provided the palatine levators are contracting simultaneously). In addition, the superior and middle pharyngeal constrictors narrow the pharynx locally. The involvement of these muscles in the basic constriction-forming manoeuvres in vowels is discussed in Paper II.

A type of model based on constriction location is compatible with observable motor activity. But specific muscle activity is not unambiguously and exclusively related to the raising or lowering, advancing or retracting of the tongue arch. Consequently the tongue arch model cannot link neuromotor activity to articulation. It offers a bewildering physiological framework, not least for EMG investigations of the tongue in speech.

Predictive capability

One aspect of the weak predictive capability of the tongue arch model has already been described. The tongue positions that can be observed in speech are not always those prescribed by the model. Table I and Figs. 3, 5 to 9 show that while the coordinates defining tongue arch position are irrelevant for the articulation of vowels, a type of model based on 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 on the tongue arch quadrilateral by assigning the "correct" coordinates. It would still

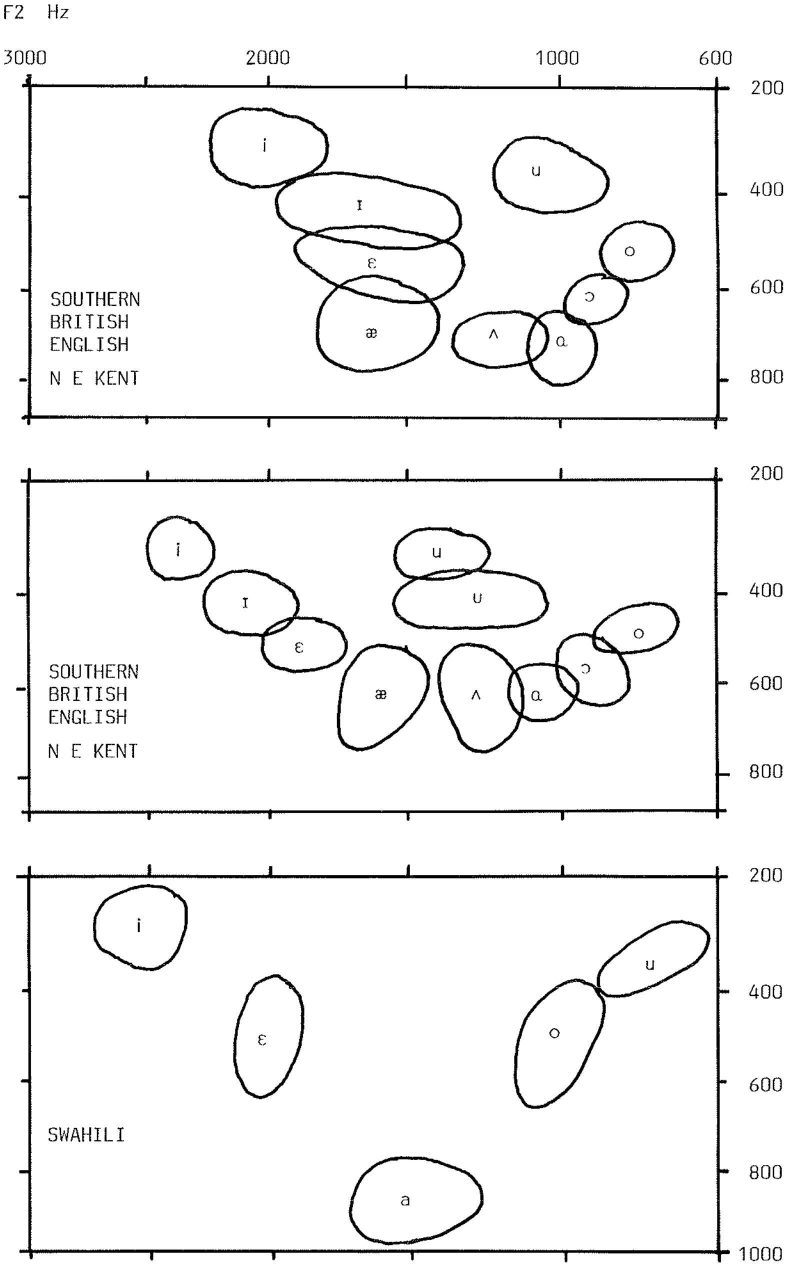


Fig. 11 Frequencies of F1 and F2 from stressed vowels in continuous speech by three adult males.

be impossible to predict the resonator configuration, and hence the spectrum. It would still be impossible to predict the underlying motor activity.

Explanatory power

In view of its unsound physiological basis and ambiguous relation to vocal tract shaping and resonance properties, and its consequently unsatisfactory predictive capability, the tongue arch model fails to provide a smooth and direct link between articulation and acoustics. It cannot therefore explain the relationships between the successive stages 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 declined when compared with the progress made in the study of speech acoustics and perception. The bewildering relationship between the parameters of the tongue arch model and the spectral character of vowels has hampered the construction of a comprehensive view embracing and integrating all phases of speech production: neuromotor, articulatory, acoustic, spectral, perception.

4. AN ALTERNATIVE ARTICULATORY MODEL

Although there are large differences between speakers in the actual frequencies of the vowel formants, the spectral character of vowels is fairly constant within the speech of one speaker (Peterson and Barney 1952). See Fig. 11. Differences of formant frequency range between speakers due to differences of vocal tract scale are regular and predictable. Spectral variations within the same speaker's speech are regular and can be related to such factors as coarticulation, degree of stress, style or temporal constraints (Tiffany 1959, Stevens and House 1963, Lindblom 1963). Much of this variation is an integral part of the coding of the speech signal (Liberman et al. 1967). The relative spectral contrasts utilized for phonemic distinctions (Jakobson et al. 1952) are probably universal. Comparing these spectral regularaties to the variability and confusion of tongue arch positions, and knowing that there are theoretically more than one possible resonator shape for a given spectrum, it was natural that many phoneticians preferred to believe there was no constancy at all in articulation and that the speaker's only concern was to produce the correct spectrum by any means available.

Figures 6 to 9 show that there are constant characteristics of vocal tract configuration. It is then reasonable to look also for constancy in the manner of forming these configurations. Is there, for example, a simple set of manoeuvres that are combined in various ways to achieve the desired configurations? The example of tense-lax vowels [i, ɪ, e, ɛ] discussed above and illustrated in Fig.4 suggests there is. This possibility is examined in Papers II, IV and V. Other examples, concerning the component manoeuvres of rounded palatal vowels, are studied in Paper III.

There has been a growing tendency in recent years to look to the shaping of the vocal tract for a substitute articulatory model. Stevens and House (1955) concluded 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 constrictions and by the degree of constriction". Kaneko (1957) compared the American English and Japanese vowel systems with reference to the vocal tract configurations. Lindblom and Sundberg described the Swedish vowel system (1969a) and the cardinal vowels (1969b) using the place and degree of constriction to define the place of articulation. Lindblom and Sundberg (1971), simulating physiological factors that determine the vocal tract area function, explored and described the spectral consequences of individual manouevres. I have myself described the West Greenlandic Eskimo vowels system using the constriction location as a phonological feature that can be used in generative rules (Wood 1971).

Vocal cavity configuration and the manoeuvres forming it can provide the basis for an alternative model that permits each phase of speech communication - motor control, articulation, speech signal, perception - to be uniquely and unambiguously linked.

5. MATERIALS AND PROCEDURES

X-ray profiles

The investigations are based on mid-sagittal x-ray tracings of vowel articulation from a collection of published tracings and 3 x-ray motion films.

38 sets from 15 different languages were collected from the literature. Their authors covered a wide range of interests such as language teaching, linguistics theory, dialectology, speech acoustics, speech therapy, laryngology and so on.

The following sets were collected:

Meyer (1907): German
Scheier (1909): German
Polland and Hála (1926): Czech
Parmenter and Treviño (1932): Spanish
Carmody (1936): Holbrook's German
Carmody (1937): Holbrook's French (3), Spanish, Italian, Portuguese, American English (2), S British English, Russian, Polish
Chlumsky et al. (1938): French
Sovijärvi (1938): Finnish
Chiba and Kajiyama (1941): Japanese, German
Mazlová (1949): Zábřeh dialect of Czech
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 British English
Fant (1960): Russian
Wängler (1961): German
Malmberg (1966): Strenger's Spanish
Perkell (1969): American English
Perkell (1971): American English
Pétursson (1974): Icelandic

Brichler-Labaeye (1970) has been consulted but was not obtained in time to be included in the above.

The 3 x-ray motion films (Southern British English, Cairo Arabic and West Greenlandic Eskimo) were made at the angiocardiographic unit of the Lund University Hospital.

Cinefluorography procedures

Each subject was limited to one reel of 35 mm film, which ran for about 40 seconds at 75 frames/second.

A highly unidirectional microphone placed close to the subject's mouth reduced the background noise level recorded from the rotating anode of the x-ray generator and the camera shutter. A sound treated booth was not available.

The camera delivered a synchronising pulse which appeared as a dot on every tenth frame. These pulses were also recorded synchronously with the speech signal on a parallel track after transposition to 6000 Hz to avoid the lower frequency region of the speech spectrum. Spectrograms were subsequently made after mixing and balancing the speech and pulse signals. It was then possible to identify the speech spectrum corresponding to any film frame.

A copper wire with blobs of solder at 10 mm intervals was taped to the subject's head and nose in the mid-sagittal plane to provide a scale for life-size reproduction of the image.

A radio-opaque paste was applied to the midline of the tongue and lips.

The subject's head was steadied against a small rest.

The camera speed was 75 frames/second (one picture every 13.3 milliseconds). The radiation was pulsed with one 3 millisecond burst to each frame. Finegrain 35 mm negative cine film was used (Fuji x-ray cine film FS). Positive copies were made for viewing and analysis (Agfa Gevaert S 99).

The subjects wore a protective apron and the the radiation was coned down to protect the eyes and the thyorid gland.

The estimated exposure and dose for each reel of film were as follows (anode voltage 120 kV, cathode current 64 mA, 4 mm aluminium and 1 mm copper filter, focus-skin distance 1200 mm, field 100 × 100 mm, 3200 frames/reel):

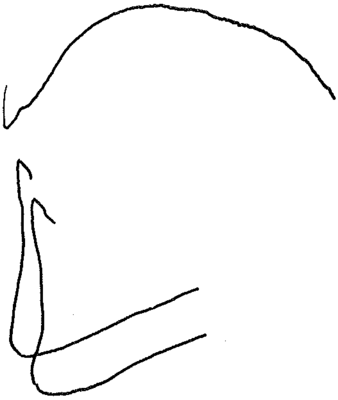
exposure at the image intensifier	15 mr/frame, 50 r/reel
exposure at the skin	230 mr/reel
absorbed dose	60-200 mrad/reel
gonad dose	negligible
bone marrow dose	negligible

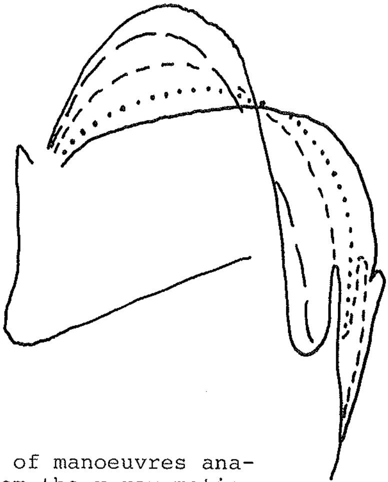
Radiation intensity is measured in *roentgens* (a measure of the ionization of the air by radiation). A radiation dose is measured in *rads* (the amount of radiation energy absorbed by biological material, 1 rad is 100 ergs/gram). A rough relation is that exposure to diagnostic x-radiation of 1 roentgen gives a dose of about 1 rad. A third unit, the *rem*, is defined as a comparable radiation dose that has the same biological effect on man as exposure to 1 roentgen of x-rays. Some examples will put these figures in perspective. The average adult annual dose in southern Sweden from cosmic, natural and environmental radiation sources is 100 mrad. A diagnostic x-ray examination of the abdomen can give doses up to 5 rad for about a dozen still pictures. The safety limit for personnel employed in x-ray units and nuclear power stations is 5 rad/annum.

Safety requirements have become increasingly stringent during the past two decades. 25 years ago a maximum exposure of 20-25 roentgens was said to be acceptable (Lusted and Miller, 1956). This is approaching the exposures used in clinical treatment, where the intention is to damage malignant tissue. Cineradiography involved high radiation doses in those days owing to the need for continuous radiation (as though one were filming with light) (Subtelny et al. 1957). In 1960, Moll estimated his radiation exposures at 24 frames/second as 3.9 roentgens for a three minute film. A finer grain (but less sensitive) film emulsion for better picture quality would have required 7.5 roentgens. Faster filming rates to capture rapid movements would have needed more intense radiation (or alternatively a more sensitive but coarser film emulsion at the same exposure, with consequent loss of picture quality). Improved image intensifiers require less radiation and currently emit enough light for fine-grain film emulsions and faster camera speeds. Intermittent radiation synchronized with the camera shutter, as used in this laboratory, obviates the need for continuous radiation and reduces the dose considerably. An even smaller dose can be obtained by x-ray microbeam systems (Kiritani et al. 1975a, 1975b). This method is more suited to tracing the movement of individual structures, often defined by a small metal pellet, than to outlining the contours of the vocal cavities.

The transparency of the body to x-rays depends on the spectral quality of the radiation which is in turn governed by the voltage applied to the anode of the x-ray tube. Radiation arises from the abrupt arrestation of electron movement, whether in an x-ray tube or in the atom nucleus of a radioactive substance. The radiation frequency depends on the force of impact. The higher the anode voltage, the greater the acceleration of electrons towards the anode, the greater the impact, and the higher the frequency of radiation emitted. The skeleton absorbs more radiation energy from low frequency radiation and consequently the shadows appear heavier on the radiograms. This is disadvantageous for a balanced picture of bone and soft tissue. Higher frequency radiation is more suitable for pictures of speech movements in order to lighten the shadow of especially the mandible relative to the soft tissues of the tongue, lips, velum and pharynx. This is achieved by using a high anode voltage, in the present case 120 kV. The copper and aluminium filters are used to further improve the spectral quality by removing low frequency radiation. The cathode current determines the intensity of the radiation by controlling the electron flow towards the anode and is set according to the intensity required to expose the film.

The camera speed is not without importance. A slow speed fails to arrest movement on each frame. A fast speed consumes more film, resulting in less talking time for the same radiataion dose. A speed of 24 frames/second fails to resolve and delimit manoeuvres with sufficient precision. This gives the impression that the speech organs are never still. A lot can happen during 40 milliseconds in speech: an articulator can turn round and be on its way back again, an occlusion can be missed altogether between frames at this speed. At least 50 frames/second is needed to resolve speech movements. The present films were made with a camera speed of 75 frames/second. The articulators are seen to complete their excursions and then wait until called on to move again.





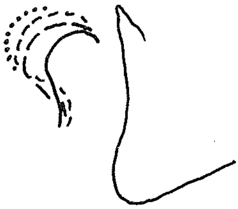


Fig. 12

Examples of manoeuvres analysed from the x-ray motion films - the mandible relative to the maxilla, the tongue relative to the mandible and the lower lip relative to the mandible.

Tracing procedures

The profiles collected from the literature were photographed from the publication and enlarged to life-size with the help of any information available about scale. Where there were no scales or measurements for guidance, the profiles were enlarged to give an overall vocal tract length in the range 15-19 mm (depending on the vowel) for male speakers and somewhat shorter for females. Comparison of features like cervicle segments, incisors, mandible, maxilla and hyoid bones ensured that all profiles in one set were reproduced to the same scale.

The x-ray films were projected frame by frame onto the back of tracing paper placed on a glass screen. The image was adjusted to life-size with the aid of the copper wire scale.

The shapes of hard structures like the mandible and maxilla were averaged over several frames.

There was a brief period in the middle of the stressed vowels where the articulators were stationary. All the component manoeuvres required for the vowel were completed but the new manoeuvres needed for the oncoming consonant had not yet commenced. There was one occasional exception to this: the Arabic subject sometimes waited until the transition to the post-vocalic consonant before completely spreading the lips after a prevocalic [b] . This brief period between consonant transitions was taken to represent the intended vowel configuration (it was the configuration that was least perturbed by the surrounding consonants). At the same time, the individual component manoeuvres were analysed by following them through from the consonant to the completed vowel configuration and then on to the next consonant.

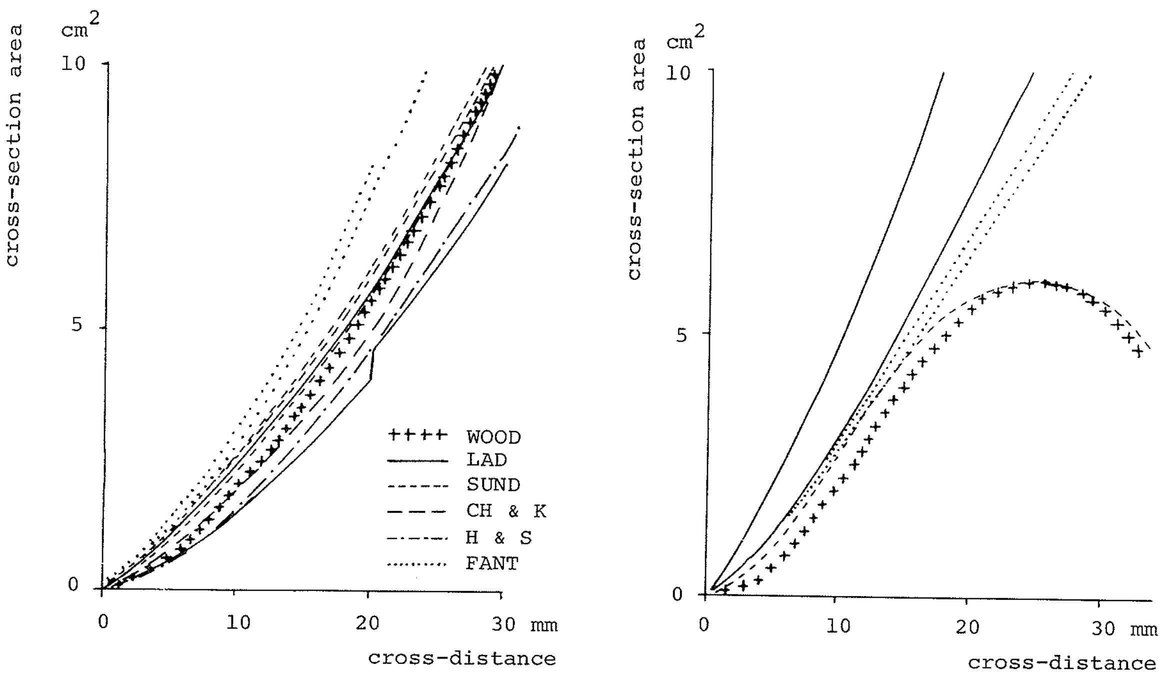
Model experiments

Papers I, III, IV and V include model experiments designed to assess the contribution of individual component manoeuvres to the vowel spectrum, to a spectral contrast or to spectral variation.

First, the manoeuvre is studied on film sequences. Some examples are illustrated in Fig. 12. Note that a manoeuvre is defined with reference to the structure on which the articulator is situated. For example, the mouth opening has two components - mandibular and labial. The mandible carries the lips up and down. The lip muscles modify the lip posture with reference to the mandible. A lip manoeuvre, such as protrusion or spreading, is defined narrowly in the latter sense, the movement of the lips relative to the mandible as a result of lip muscle activity. Similarly, there are two components of tongue movement - mandibular and lingual. The mandible carries the tongue up and down. The tongue muscles modify the tongue posture with reference to the mandible. A tongue manoeuvre is a movement of the tongue relative to the mandible as a result of tongue muscle activity.

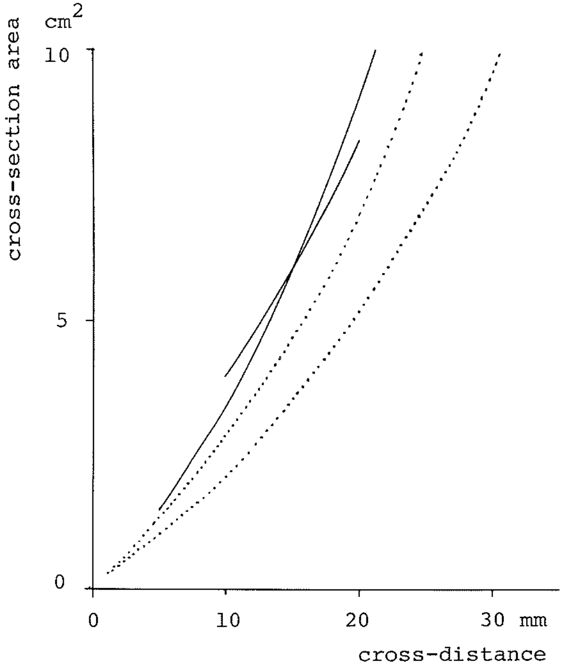
The vocal tract profile is then selectively manipulated to reproduce the desired manoeuvre. Sometimes it is convenient to perform the manoeuvre in more than one step.

Finally, the spectrum is calculated at each step from the area



(a) Mouth region

(b) Upper pharynx



(c) Lower pharynx

Fig. 13

The relationship between mid-sagittal cross-distances and the cross-section area in three different parts of the vocal tract. LAD: Ladefoged et al. (1971). H & S: Heinz and Stevens (1964), SUND: Sundberg (1969). CH & K (Chiba and Kajiyama) and FANT (Fant 1960) are esti-estimated from their x-ray profiles and area functions. WOOD: see text.

function for the configuration and evaluated.

Procedures for deriving area functions and calculating resonances are outlined below.

Deriving area functions

The vocal tract area function specifies the cross-section area along the vocal tract from the lips to the glottis. It is needed as a representation of vocal tract shape when calculating the vocal tract resonances.

The vocal tract is divided into 5 mm segments from the central incisors to the glottis. The cross-distance in each segment is measured perpendicular to the axis of the vocal tract and transformed into the cross-section area of the segment.

The x-ray profiles show only the mid-sagittal plane. Transverse distances cannot be seen, nor can the shape of a transverse cross-section. Just how complex the shapes of transverse sections of the vocal tract are can be seen in, for example, Fant (1960, Fig. 2.2-5). The problem is how to find the cross-section area when only the mid-sagittal cross-distance is known. The problem has been discussed by Chiba and Kajiyama (1941), Fant (1960, 1965), Heinz and Stevens (1964), Sundberg (1969) and Ladefoged et al. (1971).

For the buccal cavity, all these authors made oral casts which were then sectioned and measured. Heinz and Stevens expressed the cross-section area as a power function of the mid-sagittal cross-distance. This is illustrated in Fig. 13a (-··-). Sundberg's results (from 3 subjects) confirmed this; his results showed little variation between the subjects, see Fig. 13a (----). Ladefoged et al. give slightly different functions for two different parts of the mouth, one close to Heinz and Stevens's and one close to Sundberg's, Fig. 13a (——). Fant's and Chiba and Kajiyama's curves are estimated by comparing their vocal tract profiles with their published area functions. All these results are fairly similar, with Sundberg (3 cases), Ladefoged (1 of 2) and Chiba and Kajiyama (1) about average for all these results. I have used Sundberg's results with one modification, Fig. 13a (++++). Experience suggested that the cross-areas were too large at small cross-distances, so the curve was flattened a little at the low end.

The functions for Sundberg's subjects are:

$A = 2.07\ a^{1.47}$ $A = 2.35\ a^{1.34}$ $A = 2.63\ a^{1.33}$

A = cross section area
a = midsagittal cross-distance

My modified function is:

$A = 1.93\ a^{1.50}$

The pharyngeal cavity is less accessible. Chiba and Kajiyama inspected the pharynx by endoscope and found it to be eliptical in section. Heinz and Stevens also treated the cross-section as an elipse. Anthony used calipers to measure the transverse distance (1964, Ladefoged et al. 1971). Ladefoged et al. made a cast in vivo (using rapid setting dental compound) for one state, the vowel [ə]. They also inspected the pharynxes of cadavers. Fant used frontal tomograms of the pharynx. These showed that the transverse width of the pharynx (the major axis of the eliptical cross-section) decreased substantially for narrow passages. Chiba and Kajiyama's endoscopic examination revealed the same lateral narrowing for the [a] constriction. Sundberg observed on Fant's tomograms that the lateral pharyngeal walls tended to approach each other when the mid-sagittal cross-distance between the back of the tongue and the posterior pharyngeal wall exceeded 19 mm or so with the result that further pharyngeal widening actually diminished the cross-section area. See Fig. 13b (----). The other authors expected the cross-section area to continue to increase as the cross-distance increased. Lindqvist and Sundberg (1971) made a fibrescopic examination of the pharyngeal cavity and estimated cross-section areas from photographs. These agreed closely with Sundberg's previous result. I have therefore followed Sundberg, approximating his curve with a cosine function (Fig. 13b).

Lindblom and Sundberg (1971) did not observe this narrowing effect at larger cross-distances in the lower pharynx. Like them, I have used Fant's curve for the lower pharynx (Fig. 13c). This was based on tomographic evidence

The cross-section area of the lip opening is modelled according to Lindblom and Sundberg (1971). A brief account is given here for reference. For more details see their article. Figure 14 illustrates the lip opening.

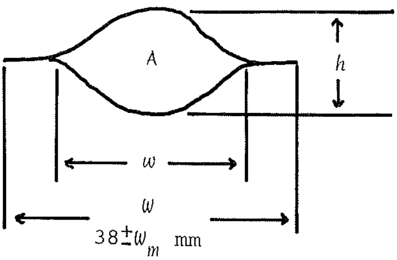


Fig. 14

The lip opening parameters of Lindblom and Sundberg's (1971) lip model, used for estimating the area of the lip section.

$$\text{Area } A = h\,w\,\frac{p}{p+1} \qquad (1)$$

p is a personal constant, 2 is recommended

h, the vertical lip separation, is measured on the tracing

w, the horizontal distance between the corners of the opening, is estimated from h

The transverse distance w is a function of h:

$$w = W \sqrt[p]{\frac{h}{h+2}} \qquad (2)$$

where

$$W = W_o + W_m \qquad (3)$$

W_o is a constant, 38 mm is recommended

W_m is a muscular increment such that

$$0 < W_m \leq +10 \text{ mm for spread lips}$$

$$W_m = 0 \text{ mm for neutral lips}$$

$$0 > W_m \geq -10 \text{ mm for rounded lips}$$

For simulating lip articulation in model experiments, h is divided into a jaw-dependent component h_o and a muscular component H_m

$$h = h_o + H_m \qquad (4)$$

where

$h_o = J - 4$ mm (J is the jaw opening)

$H_m = 2$ to 4 mm for lax spread lips

$H_m = 5$ to 8 mm for tense spread lips (see Paper IV)

$H_m = 0$ mm for neutral lips

$H_m = -1$ mm for lax rounded lips

$H_m = -2$ mm for tense rounded lips

The length of the lip section is estimated according to Lindblom and Sundberg and transformed into 5 mm segments.

A standard wedge-shaped larynx is fitted to all configurations, widening from 1 cm^2 to 2 cm^2.

Calculation of resonance frequencies

Two methods have been used.

At first, the area function was set up on an electric line analogue (LEA at the Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Fant 1960 pp. 100-101). See Fig. 15. The analogue is swept with a sine wave and the frequency of spectral peaks read off on the frequency meter.

Subsequently a digital computer was used. Webster's equation was used to calculate the pressure distribution along the configuration at successively higher excitation frequencies until standing

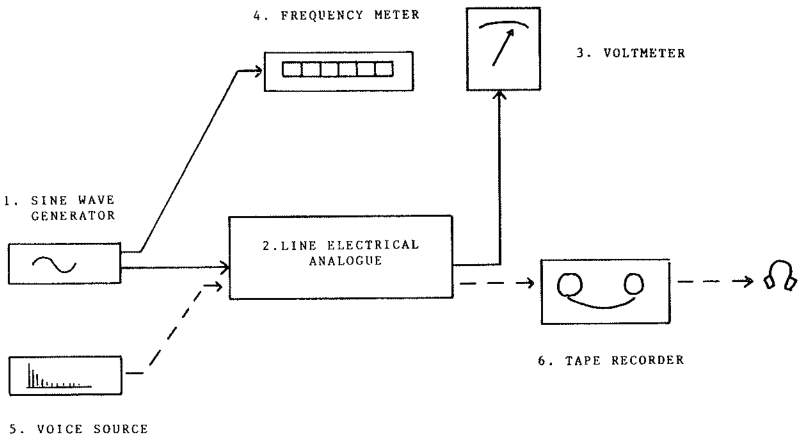


Fig. 15 For sweeping and measuring resonances, a sine wave from the signal generator (1) is passed through the line electric analogue LEA (2) to a voltmeter (3). Voltage maxima occur at resonance frequencies which can be read off from the frequency meter (4). For monitoring and recording synthetic vowel spectra, a voice spectrum from a voice source (5) is filtered by the analogue and recorded on a tape recorder (6).

The resonance frequencies calculated from Fant's Russian area functions (1960 p. 115) are compared in Fig. 16 with the corresponding frequencies given by Fant (1960 Table 2.31-1) measured on LEA, calculated by BESK (one of the first Swedish experimental computers) and measured from the subject's speech.

The difference between kinetic and potential energy distributions along the vocal tract is used as a measure of the sensitivity of a resonance mode to a local area perturbation (Fant 1975, 1980, Fant and Pauli 1975; see also Papers II and III). Complete sets of sensitivity functions are presented in Paper II for the subjects of the English and Arabic films. Similar sets have previously been published for Russian by Fant (preceding references) and for French by Mrayati and Carré (1976).

The first step is to calculate the volume velocity segment by segment from the pressure distribution:

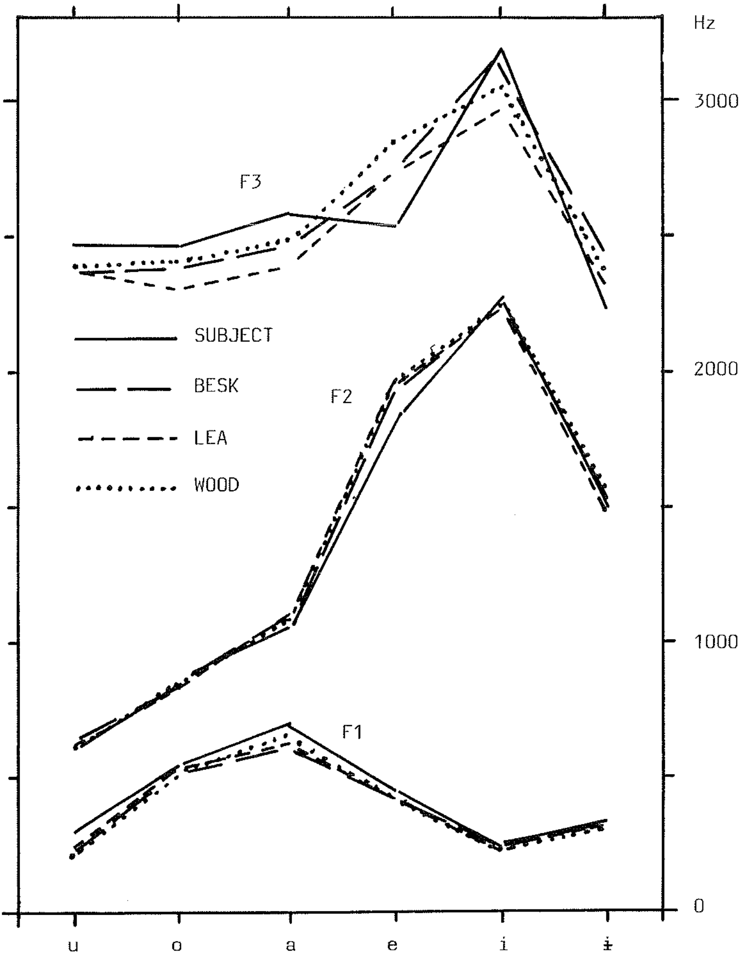


Fig. 16 A comparison of calculated formant frequencies for Fant's area functions (1960, p. 115) and the formant frequencies of Fant's subject.

$$U_i = \frac{P_{i+1} - P_i}{\ell} \times \frac{A_i}{\rho \times \omega} \quad (5)$$

where

ℓ = segment length

ρ = density of air, 1.2^{-3}

A = cross-section area of segment i

ω = $2 \times \pi \times F$

F = frequency

The kinetic energy in a segment is proportional to the volume velocity:

$$Kin_i = \left(\frac{U_i}{A_i}\right)^2 \times \ell \times A_i \times \rho \quad (6)$$

and the potential energy in a segment is proportional to the pressure:

$$Pot_i = P_i^2 \times \ell \times \frac{A_i}{\rho \times c^2} \quad (7)$$

where c = velocity of sound, 35 000 cm/sec

To permit comparisons of sensitivity in different vocal tract configurations the local energy difference is normalized by expressing it as a proportion of the total energy in the tube (i.e. by setting the total energy to 1).

The vocal tract transfer function is calculated according to Fant (1960, pp. 49-56).

Speech material

The desire to keep exposures low and yet use a relatively fast camera speed to capture fast movements and brief events limits the useful time available to about 40 seconds/reel at 75 frames/second. Silence between utterances is wasted time. In order to exploit the time available as effectively as possible the periods of silence were minimized by using longer utterances. Each test vowel was repeated several times in a test sentence.

The subject of the English film read a randomized list of sentences of the form ['pVti 'pVtə 'p*V*ti,pVtə] where different English vowels were substituted for [V] and where [*V*] indicates the focus (intonation nucleus) of the sentence. A model for this sentence is *Petty put a patty-potter.* Examples of actual sentences used are *Petty pet a petty-petter*, *Potty pot a potty-potter*, *Pawty pawt a pawty-pawter* etc. The sentences were read at two different rates. They were read once at about 4.5 syllables/second (a little slower than normal average rate) and once at about 6.5 syllables/second (a little faster than average). This yielded a total of 8 renderings of each of 11 different stressed vowels. Articulatory variation was induced by the different rates and by the different positions in the sentence (varying degrees of stress in relation to sentence focus).

The Cairo Arabic subject read similar sentences with the form ['bVti 'bVti 'bVta 'b*V*ti] and ['bVṭa 'b*V*ṭi] (with "emphatic" [ṭ]). These were read once each and yielded 4 renderings of each vowel. A model for the Arabic sentences is *Naadi šaari tiina bunni* (Nadi/has bought/figs/brown). Different vowels were substituted for [V] as above yielding *Baati baati baata baati*, *Bitti bitti bitta bitti* etc. Here again [*V*] indicates the sentence focus.

For both subjects, the alternation between labial and lingual consonants and between the intervening weak [i] and [ə] or [a] ensured that all test vowels were created anew each time and that each articulation was therefore truly independent.

The number of tokens of each test vowel is important for statistical treatment (published sets of tracings rarely contain more than one token of each vowel).

6. THE FIVE PAPERS: SUMMARIES AND CONCLUSIONS

I. TONGUE RETRACTION IS NOT SO USEFUL AFTER ALL

As a result of experiments with a simple cylindrical model Stevens (1972) found three zones where vowel spectra are insensitive to small shifts of the constriction location, corresponding to palatal, velar and pharyngeal constrictions. This paper reports the results of some similar model experiments based on the vocal tract itself, using palatal and velar configurations. They confirm that the spectra are not very sensitive to location shifts around the palatal and velar zones but they are very sensitive to location shifts in between.

II. A RADIOGRAPHIC ANALYSIS OF CONSTRICTION LOCATIONS FOR VOWELS

Vocal tract area functions estimated from 38 sets of x-rayed vowel articulations collected from the literature and from new x-ray motion films of English and Arabic speech reveal four constriction locations: along the hard palate, along the soft palate, in the upper pharynx and in the lower pharynx. Each location is appropriate for a definable class of vowel qualities, confirming the quantal nature of at least this aspect of vowel articulation. The acoustical, physiological and phonological implications are discussed. In a given phonotactic environment the precision of the constricting tongue manoeuvre was good. The only truly language specific difference was a preference for either the mid-palatal or pre-palatal location for palatal constrictions. The tongue muscles are found to be admirably situated for creating constrictions at the four locations.

Six examples of area functions from the collected material, additional to the examples illustrated in the article, are given in Figs. 7 to 9 above. These also show the four locations.

The following conclusions were drawn.

1. There are four different places where the vocal tract is

narrowly constricted by the tongue for vowels - along the hard palate, along the soft palate, in the upper pharynx and in the lower pharynx (Figs 1, 2). This finding confirms Stevens's hypothesis that we seek to constrict the vocal tract for vowels at those places where F_1 and F_2 are least sensitive to variability of constriction location.

2. The vowels produced at these locations fall into distinct families: [i-ε, y-ø]-like, [u-u, ɨ]-like, [o-ɔ, ɣ]-like and [ɑ-a-æ]-like respectively (Figs 1-5). This supports Stevens's conclusion regarding the quantal nature of vowels.

3. The tongue assumes characteristic postures relative to the mandible that correspond to the four constriction locations (Figs 6-10). Lingual movements to these postures can be unambiguously referred to muscular activity (Fig 11).

4. There are documented examples of languages preferring either the pre-palatal or mid-palatal locations for the palatal constrictions. However, the sphincteral function of the palatoglossi and the pharyngeal constrictors leaves little opportunity to vary the locations of the other three constrictions.

5. In a given consonant environment there is good precision of the constricting movements. There was no evidence that the direction of the constricting gesture is modified to compensate for random vertical larynx movement. But the structure of the epiglottis and the thyroid cartilage ensures that the low pharyngeal constriction automatically remains at about the same distance from the glottis.

6. When vowels are rounded, there is an increase in the distance from the glottis to each of the four regions where F_1 and F_2 are insensitive to small shifts of the constriction location. This is allowed for by depressing the larynx considerably for rounded vowels. This is investigated further in Paper III.

7. Articulatory features for use in phonology should reflect the preference for four constriction locations and the unique relationship between constricting tongue gestures, muscle situation and the degrees of sensitivity of vocal tract resonances to area perturbations at different parts of the vocal tract.

8. The approach to tongue articulation outlined here facilitates the building of a comprehensive description of speech production in which each of the successive stages (neuromotor, articulation, cavity shaping, spectral output) are unambiguously related to each other. In contrast, the features of tongue arch height and fronting of the established model are ambiguous in these respects and constitute a capricious medium for relating the different phases of speech production.

III. THE ACOUSTICAL CONSEQUENCES OF TONGUE, LIP AND LARYNX ARTICULATION IN ROUNDED PALATAL VOWELS

According to available data, languages contrasting [y] and [i] have a prepalatal rather than midpalatal constriction for these vowels. The tongue blade is sometimes raised a little for [y] and the tongue body is usually lower than for [i]. The lips are less rounded for [y] than for [u]. The larynx is lower for [y] than for [i]. The acoustical consequences of these manoeuvres are studied in a series of model experiments and the advantage of the combination is found to be threefold: it yields a maximum plain-flat /i/-/y/ spectral contrast, it ensures similar spectral sensitivity to tongue displacement (for example from coarticulation) in both spread-lip and rounded palatal vowels and it contributes to stable resonance conditions in the vocal tract. These results are valid for vocal tracts of any size and can be generalized to all speakers. Compensatory manoeuvres are also examined. An indirect form of compensation is provided by local manoeuvres that control sensitivity elsewhere in the vocal tract. Finally, physiological and phonological implications are discussed.

The following conclusions were drawn.

1. Only the prepalatal tongue position is favourable for the plain-flat contrast. Prepalatal [i] has a high F_3 that is lowered more than F_2 by lip rounding so that F_2 and F_3 are close together in [y] (spectral flattening). The midpalatal position is unsuitable for this contrast. The frequency of F_3 is already low in midpalatal [i] and is lowered less than F_2 by lip rounding so that F_2 and F_3 would actually diverge (no spectral flattening).

2. Only moderate lip rounding is suitable for [y]. With more

than moderate lip rounding, the spectral flattening of [y] becomes increasingly sensitive to lingual perturbations, F_2 and F_3 responding differently to tongue location shifts. The spectral consequences become increasingly unpredictable when lip rounding exceeds the critical limit.

3. Depressing the larynx for [y] keeps the zone, where F_2 is least sensitive to tongue location perturbations, within the prepalatal region. This zone is shifted away from the glottis by lip rounding, but simultaneous larynx depression (i.e. lowering the glottis) brings it back to the prepalatal region.

4. Depressing the larynx for moderately rounded [y] also ensures that F_2 and F_3 respond similarly to constriction location perturbations throughout the prepalatal region. This safeguards the spectral flattening of [y] against location perturbations, provided the lips are not more than moderately rounded.

5. Depressing the larynx for [y] ensures similar spectral consequences of disturbances of the degree of palatal constriction and variations of pharyngeal width for both rounded and spread lip palatal vowels. Similar lingual movements would otherwise produce contrary spectral consequences for [y] and [i].

6. Resonance conditions are further stabilized by raising the tongue blade in [y]. However, only a limited amount of tongue blade elevation is tolerable since this also raises F_3, thereby weakening the spectral flattening. The tolerable limit is small in, for example, Danish, Dutch, French and German where a fully flattened [y] is contrasted with [i]. Swedish requires more tongue blade elevation in order to provide a higher F_3 for the partially flattened /y:/ which has to contrast with both /i:/ and /ʉ:/.

7. Tongue body lowering raises F_1 and cancels the fall that would have followed from the lip rounding of [y]. F_1 remains much the same as in [i]. This provides a common reference that enhances the flattening of F_2 and F_3.

8. Larynx depression cannot compensate for labial undershoot in [y]. The laryngeal contribution to the lowering of F_3 for spectral flattening is much smaller than the labial contribution. Lip rounding and larynx depression are complementary

rather than mutually compensating.

9. The articulation of [i] and [y] does not need to be modified during growth in order to compensate for progressive scalar changes of vocal tract morphology. The relative consequences of lingual and labial manoeuvres and larynx height remain constant irrespective of scalar differences. Absolute resonance frequencies are shifted downward during growth, but experience shows that this is tolerated in speech perception.

IV. RADIOGRAPHIC AND MODEL STUDIES OF THE PALATAL VOWELS

Jaw openings, tongue positions and lip positions of palatal vowels in German, English, French, Russian and Arabic are analysed from x-ray profiles collected from the literature and from two x-ray films. A jaw opening of 8 or 9 mm divided a closer class [i, ɪ] from a more open class [e, ɛ]. The tongue assumed a more bunched and elevated posture relative to the mandible for tense vowels [i, e] and a lower, flatter posture for lax vowels [ɪ, ɛ]. The lips were spread more for tense vowels than for lax vowels. The acoustical consequences of these articulatory differences (especially the tolerances for variation and overlapping of manouevres and the acoustical efficiency of compensations) are investigated in a series of model experiments. Finally, some implications for phonology and speech production theory are discussed.

The following conclusions were drawn.

1. A jaw opening of 8 or 9 mm divides the palatal vowels into a close class [i, ɪ] and an open class [e,ɛ]. Within each class the jaw opening ranges of different vowels overlap almost completely (usually about 5-9 mm and 9-15 mm or more respectively). These ranges are similar across languages irrespective of the number of palatal vowels in each language.

2. With lingual compensation for mandibular variation, as in natural speech, the vowel spectra are hardly sensitive to a wide range of jaw opening variation. The consequent spectral variation is small compared to the total spectral variation of a vowel that can be observed in speech. There is no need for labial compensation for mandibular variation in spread-lip palatal vowels.

3. There was no evidence of labial compensation for mandibular variation in spread-lip palatal vowels in the two x-ray films.

4. There are differences of lip spreading relative to the mandible between tense and lax vowels. The lips are spread more for tense vowels, less for lax vowels.

5. The degree of palatal narrowing is very similar for similar vowels across languages, indicating that this is a constant factor of vowel articulation irrespective of the number of vowels in a system. In particular [i] and [ɛ] are very similar across languages and no modification is necessary to make room for additional vowels.

6. The lower tongue height for [ɪ] than for [e] is confirmed.

7. The tongue assumes a typical bunched and elevated posture relative to the mandible for tense vowels [i, e]. This narrows the palatal passage and widens the pharynx. The tongue assumes a typical flatter and lower posture for lax vowels [ɪ, ɛ]. This widens the palatal passage and narrows the pharynx. The two tongue postures are found across languages irrespective of the number of vowel contrasts in a system.

8. The palatal vowels are differentiated by combining the two degrees of jaw opening with the tense or lax tongue posture. In a model experiment, Russian profiles for close tense [i] and open lax [ɛ] were manipulated to simulate close lax [ɪ] and open tense [e]. The two new combinations generated spectra that did not clash with the original [i, ɛ], spectra confirming that [i] and [ɛ] did not need modifying to make room for [ɪ, e]. The lip, jaw and tongue positions were assigned tolerances that correspond to articulatory variation observed in speech. The resulting spectral variation was comparable to actual variation observable in speech.

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

This paper reviews articulatory, physiological and spectral differences between tense and lax vowels and reports a series of model experiments that were designed to assess how much of the spectral difference within pairs of tense and lax vowels can be attributed to the degree of constriction and how much to the difference in the pharynx. Midsagittal profiles of the vocal tract were systematically modified, the corresponding area functions set on an electrical vocal tract analogue and the resonance frequencies found and measured.

The following conclusions were drawn.

1. The difference in the degree of constriction contributed most to the tense-lax contrast.

2. For rounded vowels, about half the spectral contrast comes from the lips, which are less rounded in lax vowels.

3. For low pharyngeal [ɑ, a]-like vowels, all the spectral contrast lies in the degree of constriction.

4. Although it is spectrally advantageous for tense [o] to have a narrower constriction than lax [ɔ], just as in all other tense-lax pairs, x-ray profiles show instead it usually has a wider constriction. This is possibly a consequence of the more advanced tongue root for [o] . The small spectral penalty to the contrast is more than made up by the difference of lip protrusion.

5. The difference in tongue root position for [o, ɔ] adds little to the contrast but it is still necessary in order to keep the lower pharynx open (maintaining the F_1 contrast with [a]-like vowels).

6. Tense vowels are not necessarily more precise than lax vowels. Precision is equally necessary for both.

7. There is not one common articulatory or physiological component for all tense-lax pairs. The one general common attribute is a difference in the degree of contraction of some muscle or muscles already actively involved in creating the basic configuration, with a consequent difference of position of the corresponding articulator.

8. Although tense vowels usually tend to be longer than lax vowels, the tense-lax contrast should not be confused with quantity contrasts. One of these contrasts is often redundant, but not always. A necessary part of the tense-lax contrast is a muscular difference leading to different articulator positions, different vocal tract configurations and consequently different spectra and timbre.

7. GENERAL CONCLUSIONS

The goal that was set up was to explore various aspects of an articulatory model in which physiologically relevant manoeuvres are coordinated to shape the vocal cavities for the intended spectral contrast. Such a model would relate manoeuvres back to speech programming and motor control and forward to the encoding of the speech signal. It would thus be possible to provide a realistic and unambiguous link between physiology, acoustics and perception.

The five papers have been concerned with various aspects of such a model. Papers I and II are devoted to where and how a major constriction is formed by the tongue in the vocal tract. Various physiological and acoustical properties of these constrictions are discussed. Paper III is devoted narrowly and in depth to the contribution of each individual articulatory component of a very small number of vowels to the plain-flat spectral contrast. Paper IV analyses how various manoeuvres are combined to achieve the spectral differentiation of spread-lip palatal vowels. Paper V deals with tenseness and laxness. Each of these aspects of vowel production is examined in both a physiological and acoustical perspective.

The model has also provided a useful framework for discussing various general problems of speech production: the quantal nature of speech (Papers I, II, III, IV), articulatory precision and tolerances (Papers II, IV), motor control and compensations (Papers I, II, III, IV), the resonance properties of the vocal tract (Papers II, III), phonological implications (Papers II, III, IV, V).

TONGUE RETRACTION IS NOT SO USEFUL AFTER ALL

Sidney Wood

According to a century-old tradition in phonetics, small adjustments of tongue fronting or retraction provide an active and useful means of modifying vowel quality. However, acoustic theory does not support this tradition in every case. Stevens (1972) has demonstrated that spectral sensitivity to constriction location perturbations is not continuous along the vocal tract. His experiments with simple tubes indicated that when the anterior (palatal) portion, the mid (velar) portion (with lip-rounding) or the posterior (pharyngeal) portion are constricted vowel spectra are hardly affected by moderate displacement of the constriction location within those regions.

This can also be seen by studying three-parameter model nomograms (Fant 1960, Stevens and House 1955) which show that there are in fact four such regions where vowel spectra are relatively insensitive to location perturbations - along the hard palate and in the lower pharynx for spread-lip [i-ɪ]-like and [ɑ-æ]-like vowels, and along the soft palate and in the upper pharynx for rounded [u-ʊ]-like and [o-ɔ]-like vowels. The same four regions were deduced from a spectrographic study of eskimo vowels (Wood 1971) and X-rayed vowel articulations confirm that these regions are exclusively used for vowels in speech (Wood 1977).

Advancing and retracting palatal and velar vowels means that the constriction is displaced along the hard and soft palates. Theoretically, this should yield but little spectral advantage since these are two of the four regions mentioned above. Three model experiments were designed to repeat Steven's experiments, this time to test the sensitivity of vowel spectra to perturbations of the constriction location in natural human vocal tract configurations. Mid-sagittal vocal tract profiles were systematically modified by retracting the tongue body from palatal [i] and [ɛ]-like configurations and by advancing the tongue body from a velar [u]-like configuration. The resonances of each new configuration were found by sweeping a line electric analogue (LEA at the Royal Institute of Technology, Stockholm, by courtesy of Gunnar Fant), Fig. 1.

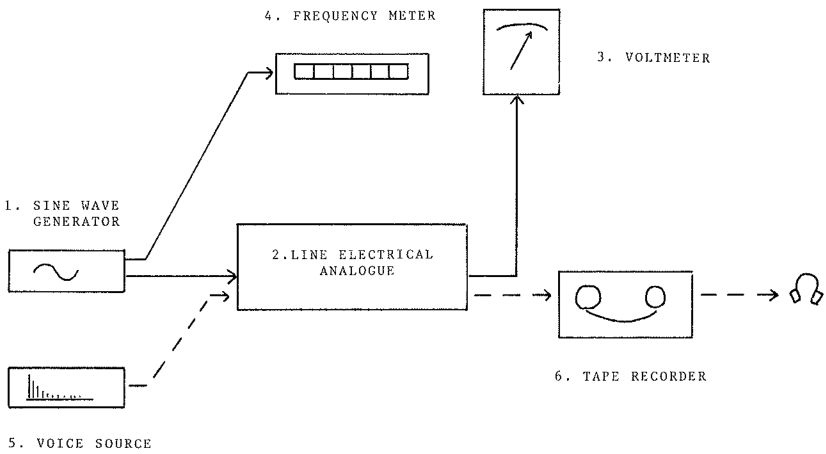


Fig. 1. The configuration to be measured is set upon the analogue (2). A sine wave from the generator (1) is passed through the analogue to a voltmeter (3). Voltage maxima occur at resonance frequencies which can be read off from the frequency meter (4). For monitoring and recording synthetic vowel qualities,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

The tongue body was retracted 5 mm in 1 mm steps from palatal [i]-like and [ɛ]-like configurations. This amounted to 20 mm retraction of the constriction along the domed roof of the mouth. Finally, the tongue was advanced 6 mm in 2 mm steps from a velar [u]-like configuration. This amounted to 20 mm advancement of the constriction along the soft palate. The degree of constriction was kept constant in each experiment (cross-section area at the constriction 0.65 cm^2 for [i] and [u], 2.6 cm^2 for [ɛ] so the only variable was the constriction location with consequent modifications to the front and back cavities.

The modification of the [i] profile is illustrated to the left in Fig. 2. A similar modification was made for the [ɛ] configuration. The range of movement represents retraction from a prepalatal constriction through midpalatal to postpalatal. X-rayed vowel articulations confirm universal language-specific preferences for either the prepalatal or midpalatal posi-

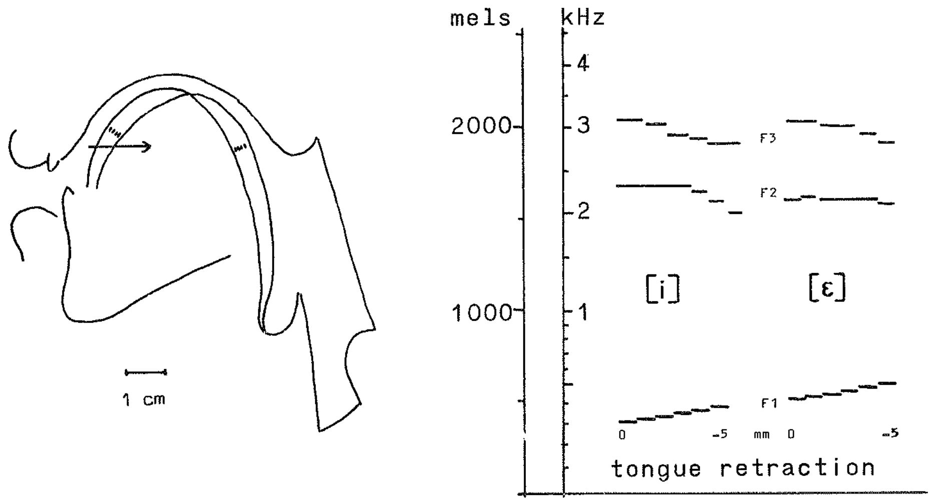


Fig. 2. Retraction of the tongue along the hard palate from palatal [i] (left). A similar modification was made for [ɛ]. Spectral consequences of these modifications (right).

tions in speech (Wood 1977). The spectral consequences of the modifications are given to the right in Fig. 2.

F_3 fell continuously in both [i] and [ɛ] for each retracted step. This is why the prepalatal [i] of say Swedish or Russian sounds sharper than the midpalatal [i] of English (Fant 1960) and why Swedish [e] and English [i] sound alike to Swedes. F_2 was hardly affected by retraction through the prepalatal and midpalatal locations and did not begin to fall until the end of each series. It fell appreciably in the last two (postpalatal) steps from [i] and had just started to fall at the last step from [ɛ]. F_1 rose gradually. The trend of these results is predictable from the pressure and volume velocity standing waves and from the energy distributions (Chiba and Kajiyama 1941, Fant and Pauli 1975). The consequences of widening the prepalatal part and narrowing the postpalatal part are (i) contrary and largely self-cancelling for F1, (ii) negligible for F2 and (iii) cumulatively negative for F3 (on account of the latter's prepalatal volume velocity node and velar pressure node).

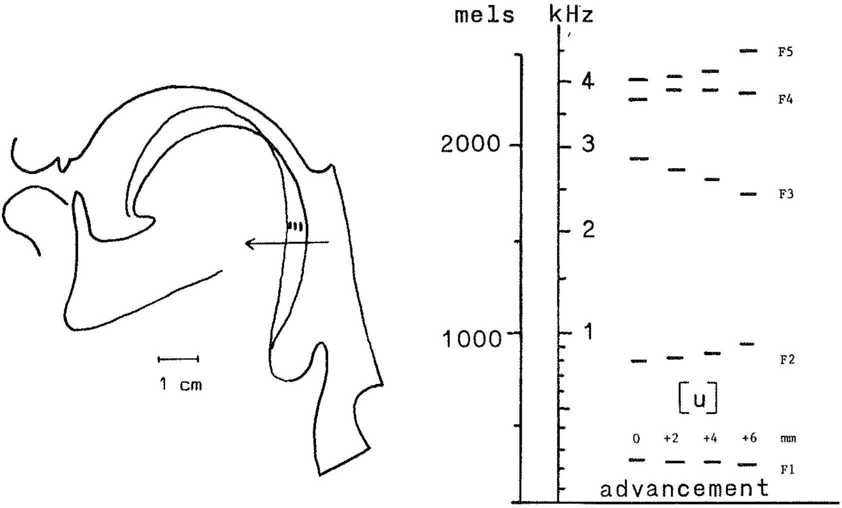


Fig. 3. Advancement of the tongue along the soft palate from velar [u] (left) and the spectral consequences (right).

The lack of change in F_2 from the **prepalatal to midpalatal locations is** expected from Stevens's finding. The postpalatal location is apparently beyond the region where formant frequencies are relatively insensitive to variation of the constriction location. F_2 changes rapidly when the constriction is retracted beyond the midpalatal position and would require considerable articulatory control in a monophthong. In actual speech vowel constrictions occur only fleetingly in this region, either during transitions to or from adjacent consonants or during diphthongs where the main information is in the gliding formant.

The modification of the [u] constriction and the spectral consequences are given at Fig. 3.

The main result of advancing the tongue body along the soft palate was that F_3 fell sharply, a cumulative consequence of widening the upper pharynx where F3 of an [u]-like configuration has its largest pressure maximum and a considerable excess of potential over kinetic energy and simultaneously narrowing the postpalatal region. Where F3 has a pressure node and an excess of kinetic energy. F_2 was hardly affected by advancing the constriction at the posterior end of the soft palate, which agrees with Stevens's finding. F_2 began to rise when the constriction was near the middle of the soft palate and it had risen about 120 Hz when the constriction was advanced to the front end. This is within the range of F_2 variation found in natural speech for [u]-like vowels.

Although it has been known for many years that the conceptual basis and assumptions of the traditional model were largely false, the implications have not yet been fully drawn. At first attention turned away from articulation and towards the composition of the acoustical signal and to the reactions of listeners to acoustical cues. But it is still highly relevant to ask what the speaker is doing when he produces those acoustical cues that we now know the listener needs. It is especially important in view of the growing interest in articulatory programming and the motor control of speech. Analysis of X-ray films (Wood 1977) showed that for the articulation of vowels the tongue aims to narrow the vocal tract at one of the four règions mentioned above, a simpler task than had hithertoo been envisaged. The tongue musculature was found to be admirably situated for creating the four constrictions and the sphincter function of the palatoglossi and the pharyngeal constrictors ensure the accuracy of all but the palatal manoeuvres. The experiments reported above confirm that vowel spectra are relatively insensitive to location perturbations in those regions. Not only is voluntary displacement of the constriction location physiologically unlikely, there is little spectral advantage to be gained from doing it anyway.

References

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

Fant, C.G.M. (1960), The acoustic theory of speech production. The Hague: Mouton

Fant, C.G.M. and Pauli, S. (1975), Spatial characteristics of vocal tract resonance modes. Speech Communication 2, 121-132. Stockholm: Almqvist & Wiksell

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, 51-66

Stevens, K.N. and House, A.S. (1955), Development of a quantitative description of vowel articulation. Journal of the Acoustical Society of America 27: 464-495

Wood, S. (1971), A Spectrographic Study of allophonic variation and vowel reduction in West Greenlandic Eskimo. Working Papers 4: 58-94, Department of Linguistics, Lund University

Wood, S. (1977), A radiographic analysis of constriction locations for vowels. Working Papers 15: 101-131, Department of Linguistics, Lund University.

Journal of Phonetics (1979) 7, 25–43

A radiographic analysis of constriction locations for vowels

Sidney Wood

Lund University, Institute of Linguistics, Department of General Phonetics, Kävlingevägen 20, 222 40 *Lund, Sweden*

Received 17*th October* 1977

Abstract: Vocal tract area functions estimated from 38 sets of X-rayed vowel articulations collected from the literature and from new X-ray motion films of English and Arabic speech reveal four constriction locations: along the hard palate, along the soft palate, in the upper pharynx and in the lower pharynx. Each location is appropriate for a definable class of vowel qualities, confirming the quantal nature of at least this aspect of vowel articulation. The acoustical, physiological and phonological implications are discussed. In a given phonotactic environment the precision of the constricting tongue manoeuvre was good. The only truly language specific difference was a preference for either the mid-palatal or pre-palatal location for palatal constrictions. The tongue muscles are found to be admirably situated for creating constrictions at the four locations.

Introduction

Since the second half of the nineteenth century, work on vowel articulation has largely been based on a model that prescribes for each vowel a unique tongue position in terms of height and fronting of the tongue arch. This model, initiated by Bell (1867), was rapidly adopted long before it could ever be subjected to experimental verification. It has never been validated, but it has been contradicted. For example, Meyer's (1910) plastopalatograms revealed unexpected tongue heights. The Bell model was conclusively discredited by Russel's massive X-ray study (1928) which failed to corroborate the predicted tongue arch positions. This was followed by several decades of advances in acoustical analysis and psycho-acoustical experimentation, which has led to proposals for purely auditory or integrated acoustical–auditory systems for describing vowels (Jakobson *et al.*, 1952; Ladefoged *et al.*, 1972; Lindau, 1975; Lieberman, 1976). During the same half century articulation was largely disavowed for vowel description. The confusing picture of vowel articulation obtained from X-ray studies and the theoretical possibility of producing the same sound in a variety of ways were seen as apparent proofs of the inconstancy of articulation and the multiplicity of compensatory adjustments. However, I suggest in this paper that the alleged inconsistencies are due to observation of the wrong articulatory variables (height and fronting) rather than to articulatory irregularity. There are important and relevant regularities in vowel articulation (Wood, 1975a), one of which will be reported here: the number of locations where the vocal tract is constricted by the tongue.

The location of a major internal narrowing is a fundamental parameter of vocal tract configuration for vowels (Stevens & House, 1955; Fant, 1960). These authors, and Lind-

0095-4470/010025+19 $02.00/0

blom & Sundberg (1971), treated the length of the vocal tract as a continuum of constriction locations for model explorations of vocal tract resonances. On the other hand, the articulatory model universally accepted from ancient India until the nineteenth century had divided vowels into [i-ɛ]-like palatals, [u-o]-like labiovelars and [ɑ-a]-like pharyngeals. These are precisely the three regions where Stevens (1972) found vowel spectra to be relatively insensitive to moderate displacements of constriction location. Examination of three-parameter model nomograms (such as Fant, 1960, Fig. 1.4–11) discloses four locations with this property: at the hard palate and in the lower pharynx for spread-lip vowels and at the soft palate and in the upper pharynx for rounded vowels (Gunnilstam, 1974). I have previously inferred the same four locations from the formant transition frequencies of West Greenlandic Eskimo vowels (Wood, 1971): (i) along the hard palate for the [i, e]-like allophones of /i/ and the [ɛ]-like allophone of /a/, (ii) along the soft palate for the [u]-like allophone of /u/, (iii) in the upper pharynx for the [o, ɔ]-like allophones of /u/

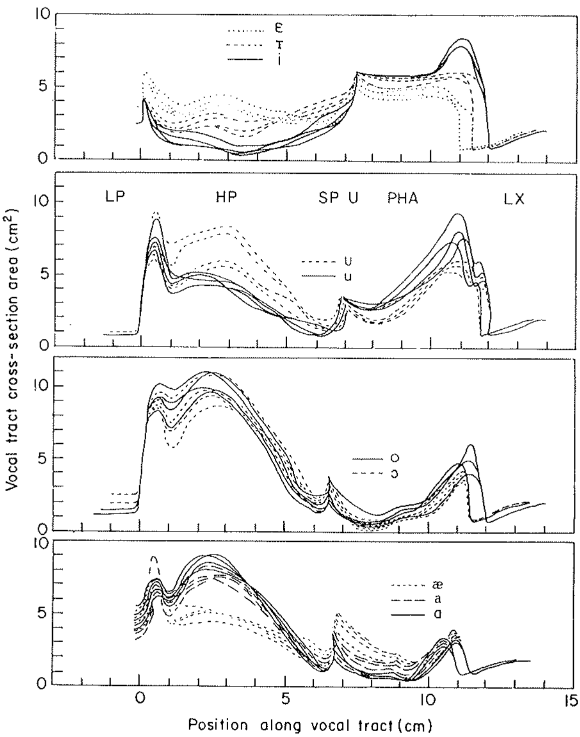


Figure 1 The configurations of the vocal tract for vowels by the Southern British English subject, classed according to constriction location (hard palate, soft palate, upper pharynx, lower pharynx). The area functions are lined up from the central incisors (coordinate 0 cm). The letters identify parts of the vocal tract: LP lips; HP hard palate; SP soft palate; U uvula; PHA pharynx; LX larynx.

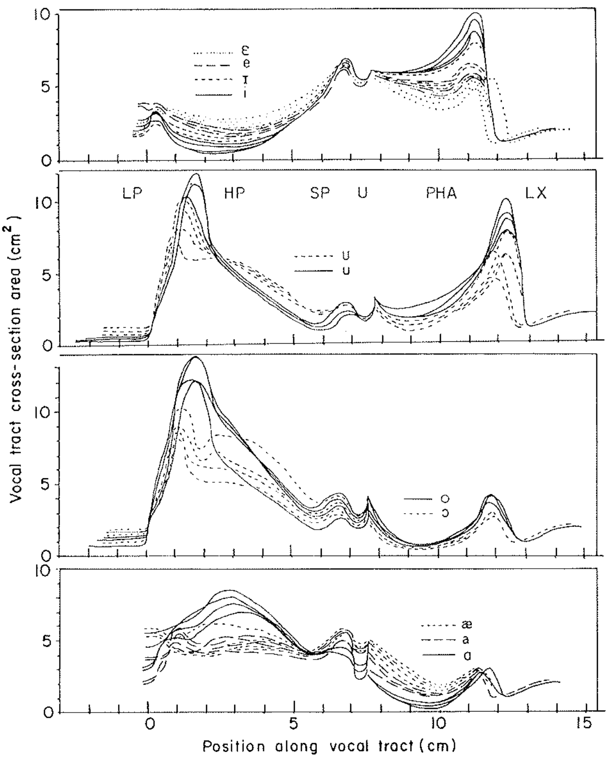


Figure 2 Area functions for vowels by the Egyptian Arabic subject. *Cf.* Fig. 1.

and the [ɣ]-like allophone of /i/, and (iv) in the lower pharynx for the [ɑ, a, æ]-like allophones of /a/. All this varied evidence points to the use of a small number of discrete locations for the constriction. The analysis of X-rayed vowel articulations reported below confirms these four constriction locations without exception by 40 subjects in 13 languages.

This finding has important implications for the study of speech physiology and the evolution of the speech organs. If it is true that it is not acoustically relevant to utilize more than four constriction locations for vowels, and if speakers universally confine themselves to these four locations, then the tongue must have a far simpler task than has hitherto been assumed. We should expect to find manoeuvres directed towards each of the four constriction targets in the vocal tract and a suitable arrangement of tongue musculature for this purpose.

Radiographic material

This investigation is based on 38 sets of mid-sagittal vocal tract tracings for 12 languages that have been reported in the literature during the past 75 years (Wood, 1975*a*) and on two new X-ray motion films of Southern British English and Egyptian Arabic speech. Tracings from a third film (of West Greenlandic Eskimo speech) are also used to illustrate

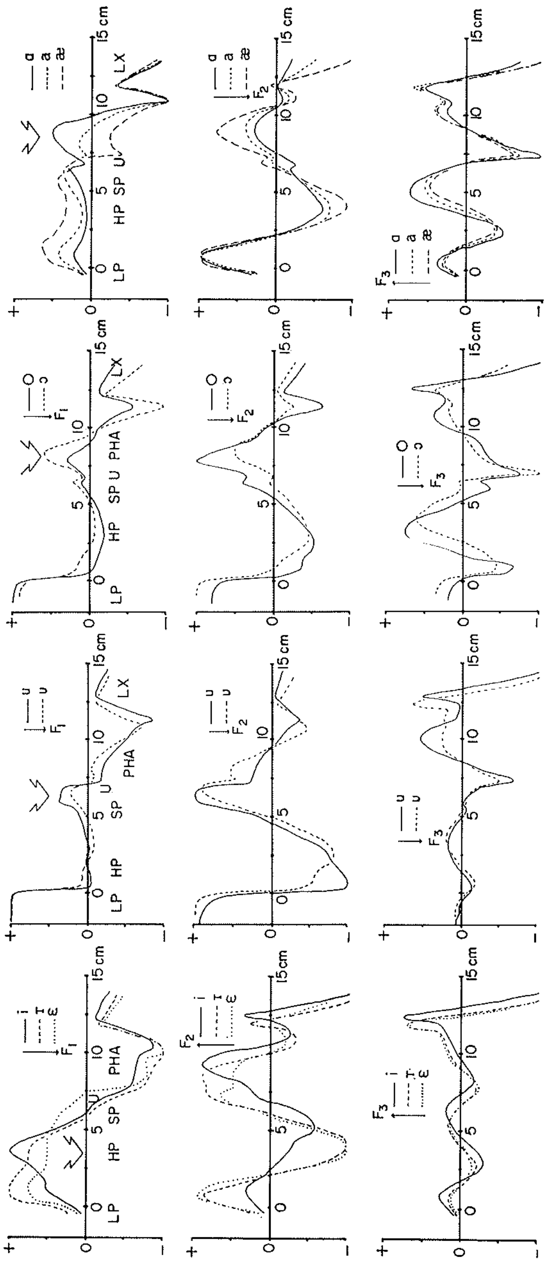


Figure 3 The difference between the kinetic and potential energy distributions along the vocal tract, indicating the relative sensitivity of vocal tract resonances to local area perturbations. Local narrowing causes a resonance to fall or rise in proportion to the excess of kinetic or potential energy respectively (conversely for local expansion). English subject (*Cf.* Fig. 1).

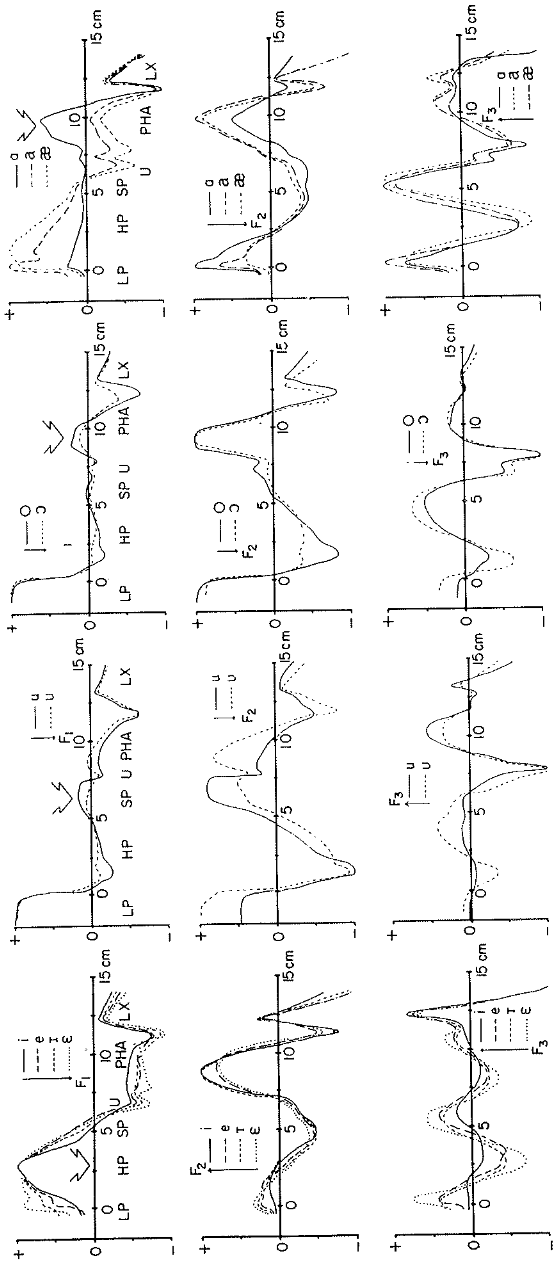


Figure 4 Vocal tract resonance sensitivity functions, *cf.* Fig. 3. Egyptian Arabic subject (*cf.* Fig. 2).

part of this report. These films were made at the angiocardiographic unit of the Lund University Hospital. The subjects were limited to one reel of 35-mm film each, which provided 40 s at 75 frames/s. The X-rays were delivered in one 3 ms pulse per frame. The absorbed radiation dose was 60–200 mrad/reel.

The English subject read a randomized list of test sentences of the form ['pVti 'pVtə ˌp*V*ti 'pVtə], where different test vowels were substituted for [V] and where [*V*] indicates the focus (intonation nucleus) of the sentence. The sentences were read at two different rates, yielding a total of eight renderings of each of ten different vowels. The Egyptian Arabic subject's sentences had the form ['bVti 'bVti 'bVta 'b*V*ti] and ['bVṭa 'b*V*ṭi] (with "emphatic" ṭ). These were read once and yielded four renderings of each vowel. For both subjects, the intervening lingual consonant and the weak vowels ensured that all test vowel articulations were independent. The number of tokens of each test item is important for statistical treatment (published sets of tracings rarely contain more than one token of each vowel).

The vocal tract area function for each rendering was estimated according to Sundberg (1969) and Lindblom & Sundberg (1971). Measurement of vocal tract resonance modes on a line electric analogue (LEA at the Speech Transmission Laboratory of the Stockholm Institute of Technology) indicated that estimated cross-section areas at the faucal isthmus in velar constrictions were too large. This will be discussed below.

Constriction locations

The striking tendency in all sets was that the tongue narrowed the vocal tract at one of four locations for vowels: (i) along the hard palate for [i-ɛ]-like and [y-ø]-like vowels, (ii) along the soft palate for [u-ʊ] and [ɨ]-like vowels, (iii) in the upper pharynx for [o-ɔ] and [ɤ]-like vowels, and (iv) in the lower pharynx for [ɑ-a-æ]-like vowels. Figs 1 and 2 show area functions for my English and Arabic subjects. Area functions for the material collected from the literature were similar.

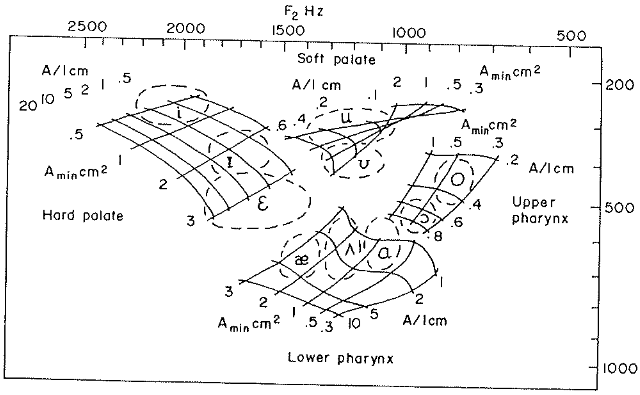


Figure 5 The frequencies of F_1 and F_2 generated by the three-parameter model for the four preferred constriction locations, based on nomograms by Stevens & House (1955) (distance from source to constriction 12 cm for hard palate, 8·5 cm for soft palate, 6·5 cm for upper pharynx, 4·5 cm for lower pharynx). The superimposed vowel areas are from a sample of Southern British English speech recorded from the radio.

These are precisely those locations, referred to in the introduction, where resonance modes are insensitive to some displacement of the constriction. The fact that all subjects restricted themselves to these locations suggests that this preference is universal.

Acoustical considerations

The combinations of F_1 and F_2 produced by constricting the vocal tract at each of the four locations can be studied by referring to three-parameter model nomograms. Figure 5 is based on the Stevens & House (1955) nomograms and gives the frequencies of F_1 and F_2 generated by varying the degree of mouth-opening and the degree of constriction at each of the four constriction locations. Figure 5 also contains the frequencies of F_1 and F_2 of stressed vowels in a sample of Southern British English speech recorded from the radio. The four constriction locations conveniently divide the entire F_1/F_2 space into four relatively unambiguous areas, each enclosing a definable family of vowel qualities. Within each area, the different spectra are obtained by varying the degree of constriction (corresponding to lingual and mandibular articulation) and the degree of mouth-opening (corresponding to labial and mandibular articulation). In natural speech the formant frequencies are also determined by tongue root movement in the lower pharynx, by tongue blade movement in the buccal cavity and by vertical larynx movement. These movements can not be simulated by a three-parameter model independently of the three model parameters. Independent simulation is only possible by direct manipulation of the area function to reproduce the desired articulatory modification (Lindblom & Sundberg, 1971; Mermelstein, 1973).

The vocal tract is a single non-homogeneous pipe whose resonance modes are sensitive to local narrowing or expansion. All parts of the vocal tract contribute in varying degrees to each mode. The following laws help us understand the spectral consequences of each of the four preferred constriction locations. Firstly, local narrowing of the vocal tract will cause a resonance mode to rise or fall according as the perturbation is made in the vicinity of a sound pressure or volume velocity maximum in the standing wave for that mode (Chiba & Kajiyama, 1941). The converse is true for local expansion. Secondly, the sensitivity of a resonance mode to a local area perturbation is related to the difference between the kinetic and potential energy at that point (Fant, 1960, 1975; Schroeder, 1967; Fant & Pauli, 1975). A local expansion will cause a resonance mode to rise or fall in proportion to the excess of kinetic or potential energy respectively. Chiba and Kajiyama (1941), Fant (1960, 1975) and Mrayati & Carré (1976) give the volume velocity and sound pressure distributions for sets of Japanese, Russian and French vowels. In addition, Fant (1975) and Mrayati & Carré (1976) give the energy distributions for their Russian and French sets. The sensitivity functions for my English and Arabic subjects are given at Figs 3 and 4. All this data from different sources is strikingly similar for similar vowel qualities, as should be expected from the gross similarities of the corresponding area functions and the apparent universality of the four preferred constriction locations. It is therefore possible to generalize with confidence from this data to similar cases in other languages.

It is possible to predict from the volume velocity, sound pressure and energy distributions that limited but specifiable portions of the full F_1 and F_2 frequency ranges can be exploited at each of the four constriction locations (as exemplified at Fig. 5). These distributions also show why the formants are insensitive to slight constriction shifts at the four locations. The sound pressure and volume velocity maxima are not narrowly localized but range over extended zones, consequently, resonance mode sensitivity to local narrowing or expansion does not alter appreciably through these zones (Figs 3 and 4).

Since all parts of the vocal tract contribute to each resonance mode, vowel qualities within each of the four classes are not determined exclusively by varying the degree of constriction. It is particularly striking to note that in each of the four basic configuration types the tongue blade, the tongue body and the tongue root are in appropriate positions to narrow or expand the vocal tract precisely at sensitive parts of the F_1 and F_2 standing waves (Figs 1–4). Further, the extrinsic tongue muscles and the pharyngeal constrictors are admirably situated for this purpose (Fig. 11).

The precision of the constricting gestures and their relation to larynx height

Stevens (1972) has pointed out the possible articulatory advantage of utilizing the regions

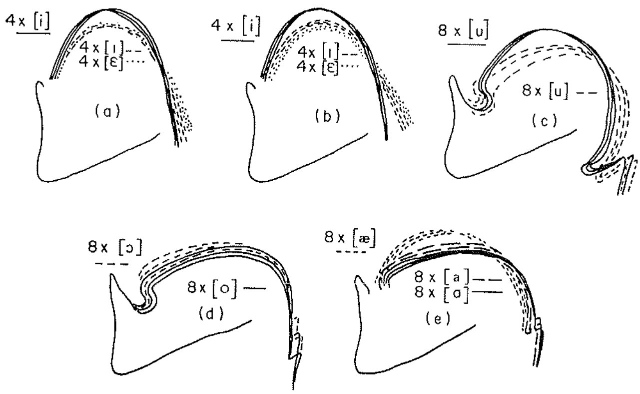


Figure 6 Positions of the tongue relative to the mandible for stressed vowels by the Southern British English subject, classed by constriction location. (a) slower rate, (b) faster rate, (c)–(e) both rates. The three profiles for each vowel represent the range of variation and the average position. *Cf.* Fig. 1. Note the mid-palatal posture for the palatal vowels (*cf.* Fig. 7).

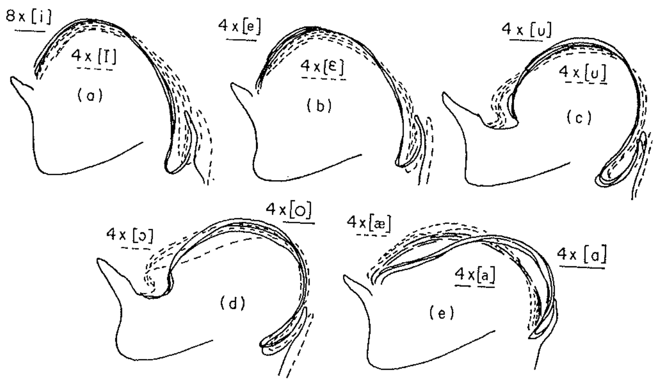


Figure 7 The positions of the tongue relative to the mandible for stressed vowels by the Egyptian Arabic subject. *Cf.* Fig. 2. Note the pre-palatal posture for the palatal vowels (*cf.* Fig. 6).

of the vocal tract where formant frequencies are insensitive to some variation of constriction location. Such variation might arise from the coarticulatory constraints of normal connected speech. We can imagine the speaker may strive to maintain the same distance from the glottal source to the constriction, implying lingual compensation for vertical larynx movement by fine adjustment of the direction of the constricting tongue gesture. Alternatively, the speaker may constantly constrict the same part of the vocal tract, implying one location target disregarding the acoustical consequences of any variation of the vertical position of the layrnx.

There is no evidence in the material reported here that speakers attempted to keep a constant distance from the glottal source to the constriction by making compensatory adjustments of the constricting gesture.

Figures 1 and 2 are lined up on a fixed anatomical landmark, the central incisors. Variation of larynx height for any one vowel type is indicated by non-alignment at the glottal end. For these two subjects, the tongue was directed towards the same part of the vocal tract for each token of a vowel type with little variation between tokens and irrespective of any variation of larynx height. A possible exception was the low pharyngeal constriction of the [ɑ, a, æ]-like class which tended to keep a constant distance to the glottis. This is probably because the upper part of the epiglottis contributes to this constriction (observed visually by Chiba & Kajiyama, 1941; Lindqvist & Sundberg, 1971). The epiglottis is linked to the thyroid cartilage so that some of the vertical movement of the larynx is transmitted to it. Consequently, the constriction remains at about 4 or 5 cm from the glottal source.

Individual constricting tongue movements are illustrated at Figs 8–10. These show the movement of the tongue from one vowel constriction to another. The similarities between the three subjects (representing three unrelated languages) should be noted.

It may be more pertinent to ask if there is any opportunity to vary the constriction locations. Some of the muscles involved have a sphincteral function with a localized constricting effect (the palatoglossi and the pharyngeal constrictors, see Fig. 11). Only in the case of the palatal constriction, where there are no muscles contracting across the vocal tract to pull the tongue upwards, does there seem to be any freedom for varying the location target. There is further discussion of this below.

The range of larynx heights between separate tokens of each vowel type by both subjects was about $\pm 2{\cdot}5$ mm. The consequent variation of the distance from the glottal source to the constriction is within the acoustically tolerable limits and has little effect on the formant frequencies. The consequent variation of the overall length of the vocal tract also has little effect, about $\pm 2\%$ for F_2 of palatal vowels and much less for other vowels (Lindblom & Sundberg, 1971; Wood, 1975*b*). I assume that this random variation of larynx height from token to token reflects varying tensions of the extrinsic laryngeal musculature.

The larynx position does nevertheless appear to play an important role in tuning the vocal tract. It is well known that the larynx is lower for rounded vowels than for spread lip vowels (Sundberg, 1969). In the English and Arabic films the average larynx height was about 10 mm lower for "tense" [u, o] and about 5 mm lower for "lax" [ʊ, ɔ] than for spread-lip [i, ɛ]-like vowels. In the tracings collected from the literature, the larynx tended to be lower for [u, o, y, ø]-like vowels than for [ɨ, ɣ, i, ɛ]-like vowels. By lengthening the vocal tract overall, this larger laryngeal depression contributes to the lowering of F_2 of the rounded vowels, especially those with a palatal constriction. But more important is the adjustment of the distance from the glottal source to the constriction. Fant's three-parameter model nomograms clearly show that when the lip-opening is narrowed there is a lengthening of the distance from the glottal source to each of the four regions where formant frequencies are not sensitive to small shifts of constriction location. The invariant tongue constriction locations still coincide with these regions in rounded vowels thanks to the lower larynx. Statistically, these adjustments are quantal. The larynx is depressed for rounded vowels (more so for "tense" than for "lax" vowels) and elevated for spread-lip vowels.

Language-specific differences of constriction location

The English and Arabic subjects had strikingly different constriction locations for palatal vowels (Figs 1, 2, 6, 7). The English subject centred his constrictions midway along the hard palate about 35 mm behind the central incisors. The Arabic subject's constrictions were more anterior, about 27 mm behind the central incisors. This difference is reflected in the directions of their constricting tongue movements for palatal vowels (Figs 6–9). The American and British subjects in the collection of published tracings all had mid-palatal constrictions centred at up to 40 mm from the incisors. Fant (1965) has pointed out a similar difference between "the Russian and Scandinavian [i] vowels which are pre-palatal whereas the [i] is articulated more towards the mid-palatal region in English", giving the former a sharper quality. The consequence of the less anterior mid-palatal constriction is a wider pre-palatal part and narrower post-palatal part, which will both yield a lower F_3. The F_3 standing wave has a pre-palatal pressure maximum and a post-palatal volume velocity maximum. Figures 3 and 4 show that F_3 is sensitive to modifications at both places in palatal vowels. It is interesting to note that most cases of pre-palatal location in the collected material are from languages contrasting [i] with [y] or [ɨ] qualities. These contrasts are enhanced by having a maximally high F_3 in [i].

The estimated area functions also indicated a difference between the velar constrictions for [u] by these two subjects (Figs 1 and 2). The maximum narrowing estimated for the English subject was opposite the uvula, the vocal tract widening out anteriorly along the soft palate. The Arabic subject appeared to narrow the dorsovelar passage near the front end of the velum. F_2 should not vary much for constriction shifts within this region but a model experiment on the analogue indicated that F_2 would rise about 120 Hz when this constriction is fronted along the whole length of the velum. Yet paradoxically the English subject's F_2 was the higher, 1000–1200 Hz against 750–900 Hz by the Arabic subject. F_2 is usually higher than 1000 Hz for [u]-like vowels in English (*cf.* Fig. 5) and lower than 1000 Hz for the graver [u]-like vowels of some other languages such as Swedish. The English subject had less close lip-rounding and a less depressed tongue blade, which according to analogue simulation would together more than counter any negative effect on F_2 from a slightly retracted constriction. But why is the Arabic subject's F_2 so much lower? Other analogue experiments indicated that the constricted dorsovelar passage needs to be extremely narrow if F_2 is to be lowered into the 700–1000 Hz range of a very grave [u] quality. Close lip-rounding alone is not sufficient. Narrowing of the dorsovelar passage is accompanied by narrowing of the top of the pharynx by the back of the tongue. The F_2 standing wave has a volume velocity maximum extending right through the dorsovelar [u] constriction, with its peak in the pharynx immediately below the uvula. F_2 is therefore lowered not only by further narrowing of the dorsovelar passage but also by narrowing behind the constriction in the vicinity of the uvula. F_2 of [u] has a considerable excess of kinetic energy throughout this region (Figs 3 and 4) so that its frequency is extremely sensitive to narrowing all the way from the faucal isthmus to the top of the pharynx. A cross-section area in the constricted dorsovelar passage of at most 0·5 or 0·7 cm^2 would be necessary to keep F_2 below 1000 Hz. The area given for the Arabic subject has probably been overestimated. It is notoriously difficult to assess cross-section areas accurately in the velar region from cross-distances alone. The palatoglossal muscles (in the anterior faucal pillars) can function sphincterally, which would narrow the faucal isthmus transversely and contribute to the degree of constriction (*cf.* Fig. 11). If these muscles are active for [u]-like vowels (the evidence will be discussed below), then they could help provide the extra narrowing needed for the Arabic subject to lower his F_2 to well below 1000 Hz.

Of these two cases, the difference of location within the palatal region is most likely to provide a difference of vowel quality, by modifying F_3. For the velar region, any F_2 variation attributable to differences of constriction location is much smaller than any opposite variation from other articulatory variables (degree of constriction, degree of lip-rounding, tongue blade depression, larynx depression) and is therefore easily cancelled by them. The true difference between the [u] vowels of these two subjects is the degree of constriction throughout the velar-uvular region.

The amount of freedom available for varying the constriction location is again a crucial problem in this context. The sphincteral mode of the palatoglossi will narrow the vocal tract locally in [u]. If the styloglossi and palatoglossi (Fig. 11) together guide the tongue into position for [u], little freedom should remain for varying the constriction location. For the palatal vowels on the other hand, there are no muscles that contract across the anterior part of the vocal tract to pull the tongue up into position. The tongue has to be pushed up from below, which leaves greater freedom for determining the target of the movement.

Physiological considerations

This section is devoted to how the four constrictions may be achieved. The discussion is based entirely on published material—anatomical descriptions supplemented by reported EMG results. The conventional view, extending back at least to Hellwag (1781), is that the tongue body is positioned for vowels by its extrinsic muscles. Hellwag listed the genioglossi (for palatals), the styloglossi (for velars) and the hyoglossi (for pharyngeals). In addition, the vocal tract is shaped by the palatoglossi (linking the tongue to the velum), the superior pharyngeal constrictors (including the glossopharyngeal fibres) and the middle pharyngeal constrictors. There will also be a contribution from the intrinsic muscles, especially the inferior longitudinals for tongue blade depression. The directions of contraction ascribed to these muscles are illustrated in Fig. 11. EMG data is still limited and has largely become available during the past decade.

Figures 6 and 7 show tracings of the tongue profile relative to the mandible for a selection of vowels at each location. This presentation highlights similarities and differences of purely lingual articulation. There was little variation between different tokens of the same vowel by the same speaker.

Figures 8–10 show the movement of the tongue to the different constriction locations from a preceding weak [i], [ə] or [a] through [p] or [b] occlusions. These are examples of the lingual manoeuvres that form the constrictions. The manoeuvres by each speaker are strikingly similar although they are for three unrelated languages.

Palatal vowels

There are no muscles that pull the tongue up towards the hard palate. The palatal tongue position is generally ascribed to contraction of the genioglossi, especially the posterior fibres. This is now being confirmed by EMG investigation (Harris, 1971; Smith, 1971;

Raphael & Bell-Berti, 1975; Miyawaki *et al.*, 1975). These fibres pull the tongue root forward, widening the lower pharynx where both F_1 and F_2 are sensitive (Figs 3 and 4). This contraction is also believed to push the tongue body upwards towards the hard

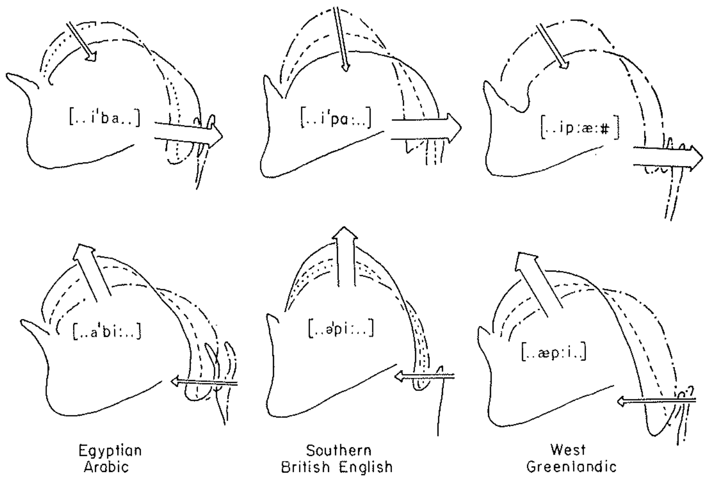


Figure 8 The movement of the tongue relative to the mandible from a palatal [i], through [p, b] occlusion to low pharyngeal [ɑ, a, æ] (above) and vice versa (below). The large arrows indicate the movement of the tongue body into the low pharyngeal or palatal constrictions respectively. The movement was sampled four times: first vowel ——, [p, b] occlusion · · · ·, [p, b] release – – – –, second vowel ——.

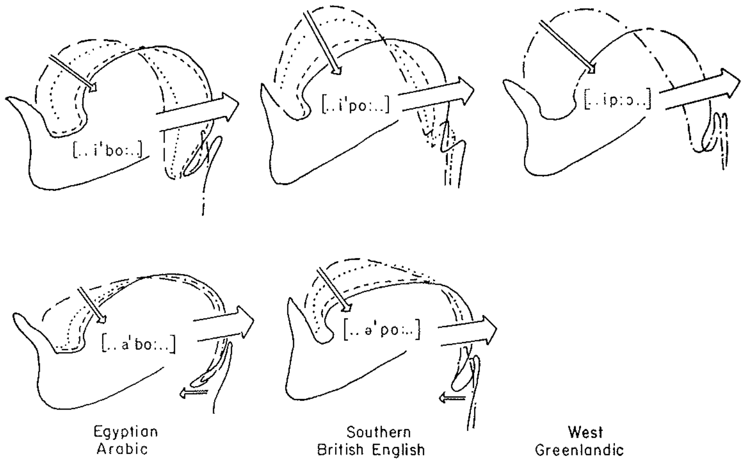


Figure 9 The movement of the tongue relative to the mandible through [p, b] occlusion to upper pharyngeal [o] from a preceding [i] (above) and [a] (below). *Cf.* Fig. 8.

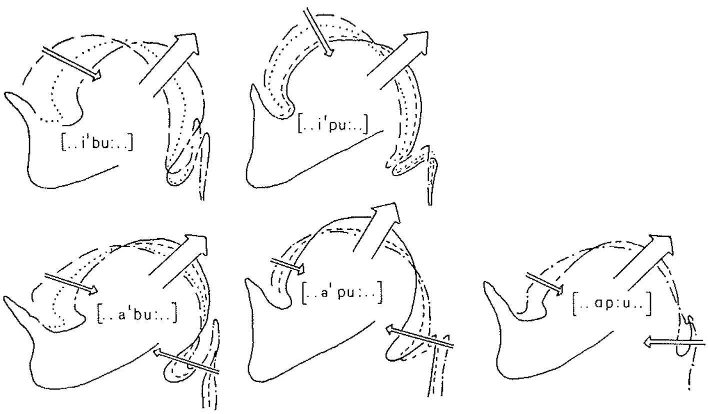


Figure 10 The movement of the tongue relative to the mandible through [p, b] occlusion to velar [u] from a preceding palatal [i] (above) and low pharyngeal [a] (below). *Cf.* Fig. 8.

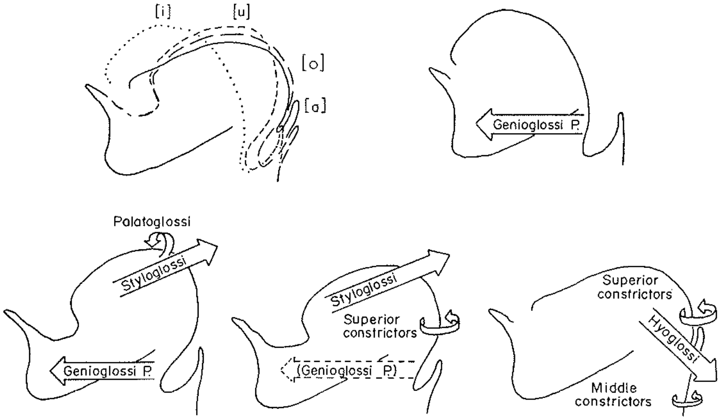


Figure 11 The directions of contraction of the extrinsic muscles of the tongue and of the pharyngeal constrictors, arranged according to their presumed activity for the formation of the four constriction locations.

palate where it narrows the palatal passage. The formants are also sensitive to the degree of palatal constriction, especially F_1. The mylohyoid muscles will also elevate the tongue body by pushing it up from the floor of the mouth but there is little agreement among authors as to whether they are used for vowels. They may, for example, provide the extra lift for a consonantal obstruction against the roof of the mouth. Any additional help in shaping the tongue in this class must come from intrinsic lingual muscles.

Labiovelar vowels

The [u, ʊ]-like and [o, ɔ]-like vowels constitute two separate classes according to constriction location, but it is convenient to discuss them together here. Figs. 6, 7, 11 show how the tongue is lower and more retracted relative to the mandible for [o] than for [u]. The same difference is true for all subjects in the collection of published tracings. The wider jaw-opening for [o, ɔ] will help lower the tongue away from the soft palate but it cannot unaided transfer the constriction to the upper pharynx. This must be done by the tongue itself.

The styloglossi draw the tongue upwards and rearwards towards the uvula. EMG investigations by Harris (1971), Smith (1971), Raphael & Bell-Berti (1975) have found the styloglossi active for [u]-like vowels. The similar bunched-up tongue postures of [u] and [o] suggest the styloglossi are also active for [o]-like vowels, but the data is very limited. Smith found styloglossal activity in cardinal [o]: but hardly any in cardinal [ɔ:]. Miyawaki *et al.* (1975) also studied the styloglossi but their results were unfortunately marred by artefacts.

The styloglossi can determine the general upward and rearward direction of lingual movement. But how is the tongue elevated towards the soft palate for [u] or retracted into the pharynx for [o]? For elevation there are two possibilities—the genioglossi pushing from below and the palatoglossi pulling from above. For retraction there are the superior pharyngeal constrictors. The hyoglossi, which draw the tongue down into the lower pharynx, have also been proposed in the literature for [o].

Harris (1971), Smith (1971), Raphael & Bell-Berti (1975) and Miyawaki *et al.* (1975) have all found the posterior genioglossal fibres to be active for [u]. Such activity can be deduced from the very wide lower pharynx for [u]. The posterior genioglossi should be less active for [o] since the tongue body is not elevated towards the velum and the tongue root is less advanced. At the same time there should be some activity for [o, ɔ] in these fibres in order to keep the lower pharynx sufficiently open for F_1 to stay in the 350–550 Hz range. There is also a difference of tongue root position between "tense" [o] and "lax" [ɔ]-like vowels (Wood 1975*b*) which indicates a difference of genioglossal activity between them. The posterior fibres of the genioglossi have so far been studied for [o] by only one group of investigators, Miyawaki *et al.* (1975). They found that the same electrode that picked up considerable activity for [u] also detected activity for [o].

Let us now turn to the palatoglossi. These are a very slender pair of muscles linking the dorsal part of the tongue and the soft palate. They lie in the anterior faucal pillars. They appear to have several modes of activity depending on what other structures or muscles happen to be doing (Lubker, 1975), including (i) to draw the tongue towards the soft palate when the latter is firmly elevated by the palatine levators, and (ii) to act like a sphincter to narrow the faucal isthmus transversely. Both of these modes are relevant for the articulation of [u]-like vowels by (i) guiding the velar (styloglossal) movement anteriorly towards the soft palate, and (ii) controlling the degree of constriction there. It is well known that the velum tends to be most tightly closed against the posterior pharyngeal wall during [u]-like vowels. An acoustic and perceptual explanation has usually been offered for this tight closure in an [u]-like configuration, based on the particular sensitivity of its oral resonances to nasal excitation and on the fact that contrastive nasality is rare for [u]-like vowels. However, a velum tightly closed by the palatine levators is a necessary condition for the sphincteral mode of palatoglossal function, which means there is a firm physiological constraint against nasalizing [u]. Any EMG activity recorded from the palatoglossi during [u] can only be due to the two modes quoted above from Lubker. Other modes are not applicable for [u] (opening the oronasal passage when the palatine levators are not contracting, or adjusting muscle length to accommodate a downward movement of the tongue). Fritzell (1969) found strong bursts of potentials in the palatoglossi during the transition from [f] to [u] in *foolish* and in [u] in the sequences [bu, du, gu] and [mu, nu, ŋu] (all with strong simultaneous levator contraction). Of Bell-Berti & Hirose's (1973) two subjects, one revealed no palatoglossal activity whatsoever, neither for [u] nor for nasals. The other had palatoglossal activity for all nasal openings and [k] and the largest bursts of all for [u]. The balance of all this data is in favour of palatoglossal involvement in the articulation of [u] with simultaneous contraction of the palatine levators.

For the retraction of the tongue into the pharynx for [o, ɔ], the most likely muscles are the superior pharyngeal constrictors (including the glossopharyngeal fibres) in view of their upper pharyngeal situation. Smith (1971) found the glossopharyngei active in [o, ɔ]-like vowels. MacNeilage & Sholes (1964) concluded from the activity picked up by surface electrodes from the back of the tongue that the hyoglossi were involved for [o]. But surface activity at that location could equally well have come from the glossopharyngei. The hyoglossi have so far not been investigated with electrodes inserted into the muscles themselves and their involvement is open to speculation in the absence of further data. Two factors probably rule out the hyoglossi for [o. ɔ]: (i) the upper pharyngeal [o, ɔ] constriction occurs where the glossopharyngei and the superior pharyngeal constrictors embrace the pharynx, and (ii) the moderately large cavity in the lower pharynx for [o, ɔ], necessary to avoid an unduly high F_1, precludes hyoglossal activity.

I conclude that the tongue is most likely drawn upwards and rearwards by the styloglossi for both the [u]-like and the [o]-like classes, and that this common movement is deflected towards the soft palate by the palatoglossi and the posterior genioglossi for [u]-like vowels, and into the upper pharynx by the superior pharyngeal constrictors and the glossopharyngei for [o]-like vowels. The genioglossi are active to widen the lower pharynx and assist in raising the tongue, more so for [u]-like vowels than for [o]-like vowels. The sphincteral function of the palatoglossi and the pharyngeal constrictors probably leaves little freedom for variation of these constriction locations.

Low pharyngeal vowels

For the [ɑ, a, æ]-like vowels, there is considerable general narrowing of the lower pharynx where F_1 has high pressure and F_2 a sensitive volume velocity maximum. The maximum constriction is in the vicinity of the epiglottis. The acoustical consequences are that F_1 is maximally high for all vowels in this class (at least 600–700 Hz with the large mouth-openings that are also typical of these vowels) and that the individual vowel qualities are distinguished by F_2 according to the degree of constriction (*cf.* Fig. 5). In contrast, the lower pharynx is less narrow for [ɛ, ɔ] configurations for which F_1 does not exceed about 550 Hz even with exaggerated mouth-openings.

The extrinsic tongue muscles that draw the tongue down into the lower pharynx are the hyoglossi. The inaccessibility of this pair has so far deterred investigators from attempting to insert electrodes. The pharyngeal constrictors are also frequently mentioned in conjunction with these vowels. The superior constrictors (including the glossopharyngeal fibres) retract the tongue and narrow the pharynx above the constriction. The middle constrictors, arising from the corni of the hyoid bone, narrow the pharynx at and below the constriction. Both Chiba & Kajiyama (1941) and Lindqvist & Sundberg (1971) have observed transversal narrowing of the constriction at the level of the epiglottis, indicating contraction of the constrictors. Smith (1971) found considerable activity from the glossopharyngeal fibres for cardinal [bɑ:] and [ba:]. Minifie *et al.* (1974) recorded the activity of the superior and middle constrictors to investigate differences of "tenseness" and "laxness" in [p, b] and reported no apparent differences between the vowels [i] and [a]. The activity they observed during [a] is expected for constrictor involvement in this vowel, but that observed for [i] is unexpected for a palatal vowel. A previous ultrasonic scan (Minifie *et al.*, 1970) had shown considerable inward displacement of the lateral pharyngeal walls (3–4 mm by the left wall) during the low pharyngeal vowels [ɑ, ʌ, æ] but little or nothing during [i, u], indicating constrictor activity for the former but not for the latter. Minifie *et al.* (1974) suggest that "had the EMG signals been integrated, discernable differences would have emerged" but they did not pursue the question further.

The availability of EMG data is least satisfactory of all in this class of vowels—nothing for the hyoglossi and occasionally contradictory for the pharyngeal constrictors. But the direct (visual) and indirect (X-ray, ultrasound) observations of movements are unanimous: the tongue is drawn into the lower pharynx and the pharynx is narrowed transversely by inward movement of the lateral walls. This points strongly to involvement of both the hyoglossi to depress and retract the tongue and of the pharyngeal constrictors to narrow the pharynx generally and especially at and below the constriction. Narrowing of the lower pharynx is predictable from acoustic theory to raise F_1 to its maximum, above 600 or 700 Hz.

Phonological considerations

All the data and discussion so far presented have clear implications for phonology. The preference for the four constriction locations is apparently universal. These locations are acoustically and physiologically significant. They divide the spectral space into four vowel quality families. The extrinsic tongue muscles are located just where they are needed to provide these constrictions and to narrow or expand the vocal tract at the sensitive parts of resonance mode standing waves. Phonetic processes that claim to describe phonemic contrasts, allophonic distributions, vowel shifts, morphophonemic alternations and so on in articulatory terms should reflect this unique way of utilizing the innate acoustical and anatomical properties of the vocal tract. The established tongue articulation model, which professed to describe tongue articulation in terms of the height and fronting of the tongue arch, failed to do this.

Notice also that the boundaries between the four families (Fig. 5) constitute the basic phonemic contrasts. In two-phoneme systems such as Kabardian (Halle, 1970) there is a contrast between (i) [A]-like vowels produced with low pharyngeal locations and (ii) non-[A]-like vowels produced at the other constriction locations. In three-phoneme systems there are contrasts between (i) [A]-like vowels with low pharyngeal constrictions, (ii) spread-lip [I-E]-like vowels with palatal constrictions, and (iii) rounded [U-O]-like vowels with velar and upper pharyngeal constrictions. Five-phoneme systems have contrasts between (i) [A]-like vowels with low-pharyngeal constrictions, (ii) spread-lip [I]-like and (iii) spread-lip [E]-like vowels with palatal constrictions, (iv) rounded [U]-like vowels with velar constrictions and (v) rounded [O]-like vowels with upper pharyngeal constrictions. No further contrasts are obtained by varying the constriction location. Other modifications are utilized for additional contrasts, for example the degree of constriction, the degree of mouth-opening and the degree of lip-rounding.

An interesting phonological problem is how far and under what conditions the phoneme boundaries just outlined may be transgressed. For example, in Arabic and Eskimo there are spread-lip [ɨ] or [ɣ]-like allophones of /i/ with velar or pharyngeal constrictions carried over from adjacent consonant articulations. Again, umlauting involves an interchange of constriction location between (i) low pharyngeal [a], upper pharyngeal [o] and velar [u] and (ii) palatal [ɛ, ø, y] respectively. Similar interchanges of constriction locations are involved in instances of vowel harmony.

This approach to vowel articulation provides a new perspective to these types of phonological problem, whereas the established model cannot capture their true nature in terms of the height and fronting of the tongue arch. The ambiguities in the latter model regarding the relationships between parameter values, physiological activity, resonator configuration and spectral output means that it is powerless to explain central areas of speech production (Wood, 1975*a*). The explanatory power of phonological theory will be greatly enhanced if the features of tongue articulation were instead derived from a comprehensive model based on the shaping of the entire vocal tract in a manner that directly relates physiology, articulation and acoustics.

Conclusions

(1) There are four different places where the vocal tract is narrowly constricted by the tongue for vowels—along the hard palate, along the soft palate, in the upper pharynx and in the lower pharynx (Figs 1, 2). This finding confirms Stevens's hypothesis that we seek to constrict the vocal tract for vowels at those places where F_1 and F_2 are least sensitive to variability of constriction location.

(2) The vowels produced at these locations fall into distinct families: [i-ɛ, y-ø]-like, [u-ʊ, ɨ]-like, [o-ɔ, ɤ]-like and [ɑ-a-æ]-like respectively (Figs 1–5). This supports Stevens's conclusion regarding the quantal nature of vowels.

(3) The tongue assumes characteristic postures relative to the mandible that correspond to the four constriction locations (Figs 6–10). Lingual movements to these postures can be unambiguously referred to muscular activity (Fig. 11).

(4) There are documented examples of languages preferring either the pre-palatal or mid-palatal locations for the palatal constrictions. However, the sphincteral function of the palatoglossi and the pharyngeal constrictors leaves little opportunity to vary the locations of the other three constrictions.

(5) In a given consonant environment there is good precision of the constricting movements. There was no evidence that the direction of the constricting gesture is modified to compensate for random vertical larynx movement. But the structure of the epiglottis and the thyroid cartilage ensures that the low pharyngeal constriction automatically remains at about the same distance from the glottis.

(6) When vowels are rounded, there is an increase in the distance from the glottis to each of the four regions where F_1 and F_2 are insensitive to small shifts of the constriction location. This is allowed for by depressing the larynx considerably for rounded vowels.

(7) Articulatory features for use in phonology should reflect the preference for four constriction locations and the unique relationship between constricting tongue gestures, muscle situation and the degrees of sensitivity of vocal tract resonances to area perturbations at different parts of the vocal tract.

(8) The approach to tongue articulation outlined here facilitates the building of a comprehensive description of speech production in which each of the successive stages (neuromotor, articulation, cavity shaping, spectral output) are unambiguously related to each other. In contrast, the features of tongue arch height and fronting of the established model are ambiguous in these respects and constitute a capricious medium for relating the different phases of speech production.

I gratefully acknowledge the assistance of Gudmund Swahn, Gunnila Holje and Rolf Schoener of the Roentgen Technology Section of the Lund University Hospital for the X-ray films, of Professor Gunnar Fant and Dr Johan Sundberg of the Speech Transmission Laboratory at the Stockholm Institute of Technology for the use of the line analogue LEA, of Sven Norén of the Lund University Computer Centre for programming and of Eva Gårding, James Lubker and Nina Thorsen for valuable discussion.

References

Bell, A. M. (1867). *Visible Speech.* London.
Bell-Berti, F. & Hirose, H. (1973). Patterns of palatoglossal activity and their implications for speech. *SR* **34**, 203–209. Haskins Laboratories.
Chiba, T. & Kajiyama, M. (1941). *The Vowel—Its Nature and Structure.* Tokyo.
Fant, C. G. M. (1960). *The acoustic theory of speech production.* The Hague: Mouton.
Fant, C. G. M. (1965). Formants and cavities. *Proceedings of the 5th international congress of phonetic sciences*, pp. 120–141. Basle: Kargar.
Fant, C. G. M. (1975). Vocal tract area and length perturbations. *STL-QPSR* 4/1975: 1–14. Speech Transmission Laboratory, Institute of Technology, Stockholm.
Fant, C. G. M. & Pauli, S. (1975). Spatial characteristics of vocal tract resonance modes. *Speech communication*, Vol. 2, pp. 121–132. Stockholm: Almqvist & Wiksell.
Fritzell, B. (1969). The velopharyngeal muscles in speech. *Acta Oto-laryngolica*, Suppl. 250.
Gunnilstam, O. (1974). The theory of local linearity. *Journal of Phonetics* **2**, 91–108.
Halle, M. (1970). *Foundations of Language* **6**, 95–103.
Harris, K. (1971). Action of the extrinsic tongue musculature in the control of tongue position. *SR* 25/26: 87–96, Haskins Laboratories.
Hellwag, C. F. (1781). *Dissertatio inauguralis physiologico-medico de formatione loquelae.* Tübingen. Reprinted by Vietor (1886), Heilbronn.
Jakobson, R., Fant, C. G. M. & Halle, M. (1952). *Preliminaries to speech analysis.* Cambridge, Mass: M.I.T. Press.
Ladefoged, P., Declerk, J. L., Lindau, M. & Papçun, G. (1972). An auditory-motor theory of speech production. *UCLA Working Papers in Phonetics* **22**, 48–75.
Lieberman, P. (1976). Phonetic features and physiology: a reappraisal. *Journal of Phonetics*: **4**, 91–112.
Lindau, M. (1975). Vowel Features. *Working Papers* **11**, 1–42. Department of Linguistics, Lund University.
Lindblom, B. F. & Sundberg, J. (1971). Acoustical consequences of lip, tongue, jaw and larynx movements. *Journal of the Acoustical Society of America* **50**, 1166–1179.
Lindqvist, J. & Sundberg, J. (1971). Pharyngeal constrictions. *STL-QPSR* 4/1971: 26–31. Speech Transmission Laboratory, Institute of Technology, Stockholm.
Lubker, J. (1975). Normal velopharyngeal function in speech. *Clinics in Plastic Surgery* **2**, 249–259.
MacNeilage, P. & Sholes, G. N. (1964). An electromyographic study of the tongue during vowel production. *Journal of Speech and Hearing Research* **7**, 211–232.
Mermelstein, P. (1973). Articulatory model for the study of speech production. *Journal of the Acoustical Society of America* **53**, 1070–1082.
Meyer, E. (1910). Untersuchungen über Lautbildung. *Festchrift Wilhelm Vietor* pp. 166–248 (special number of *Die Neueren Sprachen*).
Minifie, F. D., Abbs, J. H., Tarlow, A. & Kwaterski, M. (1974). EMG activity within the pharynx during speech production. *Journal of Speech and Hearing Research* **17**, 497–504.
Minifie, F. D., Hixon, T. J., Kelsey, C. A. & Woodhouse, R. J. (1970). Lateral pharyngeal wall movement during speech production. *Journal of Speech and Hearing Research* **13**, 584–594.
Miyawaki, K., Hirose, H., Ushijima, T. & Sawashima, M. (1975). A preliminary report on the electromyographic study of the activity of the lingual muscles. *Annual Bulletin of the Research Institute of Logopedics and Phonetics* **9**, 91–106. Tokyo.
Mrayati, M. & Carré, R. (1976). Relations entre la forme du conduit vocal et les traits charactéristiques acoustiques des voyelles françaises. *Phonetica* **33**, 285–306.
Raphael, L. J. & Bell-Berti, F. (1975). Tongue musculature and the feature of tension in English vowels. *Phonetica* **32**, 61–73.
Russel, O. (1928). *The vowel.* Columbus, Ohio.
Schroeder, M. R. (1967). Determination of the geometry of the human vocal tract by acoustic measurements. *Journal of the Acoustical Society of America* **41**, 1002–1010.
Smith, T. (1971). A phonetic study of the function of the extrinsic tongue muscles. *UCLA Working Papers in Phonetics* 18.
Stevens, K. N. (1972). The quantal nature of speech: evidence from articulatory-acoustic data. In (David, E. E. & Denes, P. B., Eds.), *Human Communication, a Unified View*, pp. 51–66.

Stevens, K. N. & House, A. S. (1955). Development of a quantitative description of vowel articulation. *Journal of the Acoustical Society of America* **27**, 484–495.

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

Wood, S. (1971). A spectrographic study of allophonic variation and vowel reduction in West Greenlandic Eskimo. *Working Papers* **4**, 58–94, Department of Linguistics, Lund University.

Wood, S. (1975*a*). The weakness of the tongue-arching model of vowel articulation. *Working Papers* **11**, 55–108. Department of Linguistics, Lund University.

Wood, S. (1975*b*). Tense and lax vowels—degree of constriction or pharyngeal volume? *Working Papers* **11**, 109–134. Department of Linguistics, Lund University.

THE ACOUSTICAL CONSEQUENCES OF TONGUE, LIP AND LARYNX ARTICULATION IN ROUNDED PALATAL VOWELS

Sidney Wood

ABSTRACT

According to available data, languages contrasting [y] and [i] have a prepalatal rather than midpalatal constriction for these vowels. The tongue blade is sometimes raised a little for [y] and the tongue body is usually lower than for [i]. The lips are less rounded for [y] than for [u]. The larynx is lower for [y] than for [i]. The acoustical consequences of these manoeuvres are studied in a series of model experiments and the advantage of the combination is found to be threefold: it yields a maximum plain-flat /i/-/y/ spectral contrast, it ensures similar spectral sensitivity to tongue displacement (for example from coarticulation) in both spread-lip and rounded palatal vowels and it contributes to stable resonance conditions in the vocal tract. These results are valid for vocal tracts of any size and can be generalized to all speakers. Compensatory manoeuvres are also examined. An indirect form of compensation is provided by local manoeuvres that control sensitivity elsewhere in the vocal tract. Finally, physiological and phonological implications are discussed.

INTRODUCTION

The model experiments reported here were undertaken in order to investigate the acoustical consequences of the individual articulatory components of [y]. These are moderate but not close lip rounding, prepalatal but not midpalatal tongue position, depressed larynx, lower tongue body than in [i], sometimes an elevated tongue blade. The evidence for these components is reviewed below. Their prime purpose is to create the flattened spectrum of [y] that contrasts with the sharper [i] (a downward shift of the formants in [y], F_3 coming close to F_2, while in [i] F_3 is high and closer to F_4; see Table I and Jakobson et al. 1951). It has always been recognized that this is the role of lip rounding in [y]. But what are the other manoeuvres doing?

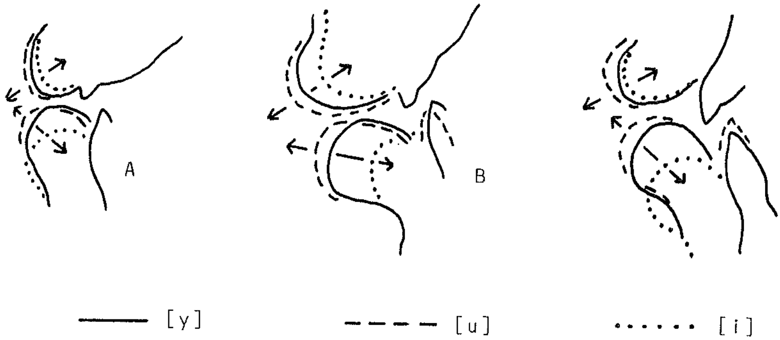
Group (a)	F1	F2	F3	F1	F2	F3
DANISH						
1. Several men	225-250	2000-2600	2800-3600	225-275	1800-2100	2000-2400
2. 10 men	233	2123	3009	240	1846	2077
1 man	274	2187	2988	263	1898	2144
9 women	278	2587	3397	247	2041	2383
6 children	304	2921	3522	281	2081	2605
DUTCH						
3. 50 men	294	2208	2766	305	1730	2208
FRENCH						
4. Subject 1	233	2225	3325	241	1767	2283
Subject 2	267	1893	2983	300	1700	2250
Subject 3	275	2283	3100	300	1817	2300
Subject 4	283	2116	3133	300	1817	2257
GERMAN						
5. Subject 1	254	2359	2975	260	1825	2207
Subject 2	238	2411	2956	236	1840	2265
Subject 3	201	2222	2767	238	1571	1925
Subject 4	227	2367	2904	222	1625	2021
Subject 5	228	2121	2825	231	1567	1963
Subject 6	266	2004	2763	276	1621	1975

Group (b)	F1	F2	F3	F1	F2	F3
SWEDISH						
6. 24 men	255	2190	3150	260	2060	2675
7. Subject 1	320	2130	3000	310	2050	2650
Subject 2	300	2000	3100	300	1900	2590

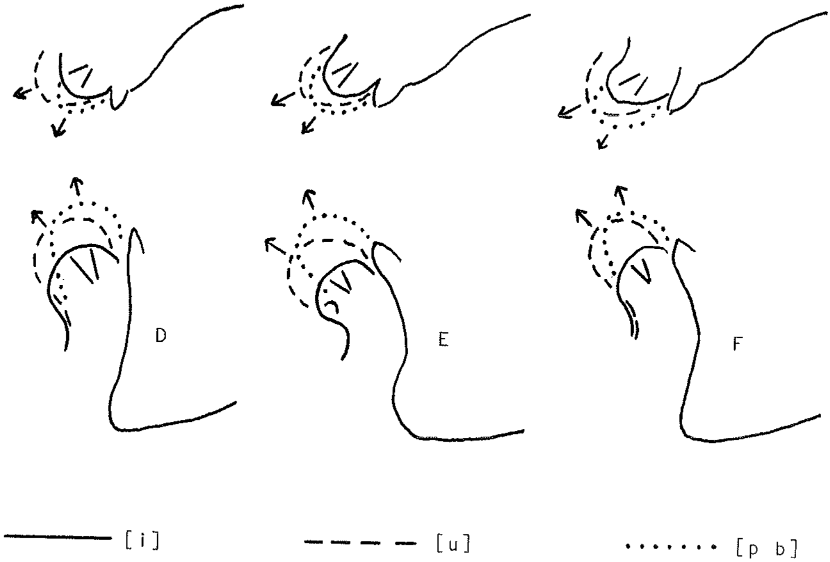
Table 1 A selection of formant frequencies for "tense" /i-y/ pairs in several languages: Group (a) with F_3 close to F_2 in [y] and group (b) with F_3 midway between F_2 and F_4 in [y]. (1) Fischer-Jørgensen (1954), (2) Frøkjaer-Jensen (1966), (3) Pols, Tromp & Plomp (1973), (4) Riordan (1976), (5) Jörgensen (1969), (6) Fant (1973, chapt. 5), (7) Sundberg & Nordström (1976)

One possible reason why the articulators assume these particular positions for [y] is that there may be quantal relations between manoeuvres and resonance conditions. Stevens (1972) has demonstrated that a quantal relationship exists between articulatory manoeuvres and speech wave characteristics under certain conditions. For example, he found zones in the vocal tract where F_2 is hardly sensitive to small shifts of constriction location. Wood (1979) compared area functions and x-ray profiles of vowels in several different languages and found that there are four such zones, that monophthong constrictions are located to these zones and that the tongue musculature is admirably situated for modifying the vocal tract at places where the standing waves of the lowest resonance modes are sensitive to local area perturbations. Further confirmation is given by Fujimura and Kakita (1979). On the other hand, Gunnilstam (1974) and Ladefoged et al. (1977) have pointed out that the degree of constriction is an exception. Formant frequencies are very sensitive to small changes in the degree of constriction. Perkell and Nelson (1981) have found that the motor control system responds differently to these two different sensitivities. Their pellet-tracking microbeam x-ray study of vowel articulation revealed greater precision in making the degree of palatal constriction (a parameter to which F_2 is very sensitive) and less precision in locating the palatal constriction along the vocal tract (a parameter to which F_2 is hardly sensitive).

Some of the articulatory components of [y] have properties that suggest there might be acoustical discontinuities that are quantally related to articulation. For example, the data reviewed below indicates that close lip rounding and mid-palatal tongue positions are avoided for [y].

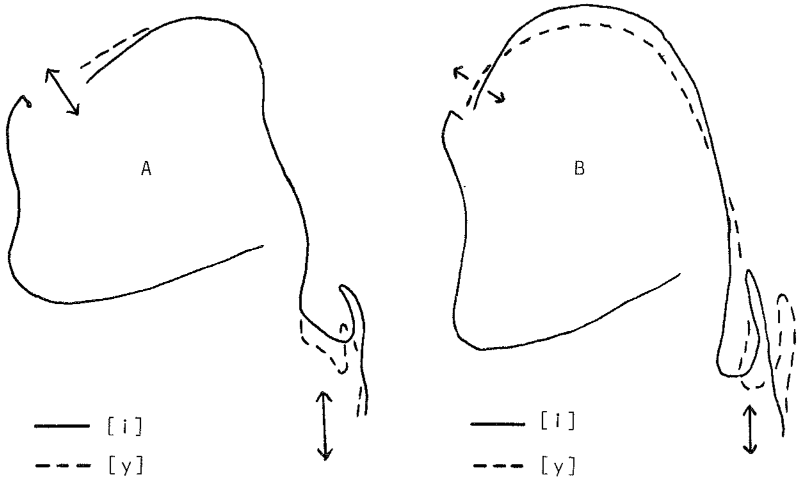


(a) Lip articulation for "tense" [i, y, u] by two Danish subjects (A, B) and a German subject (C) showing less lip rounding for [y] than for [u].

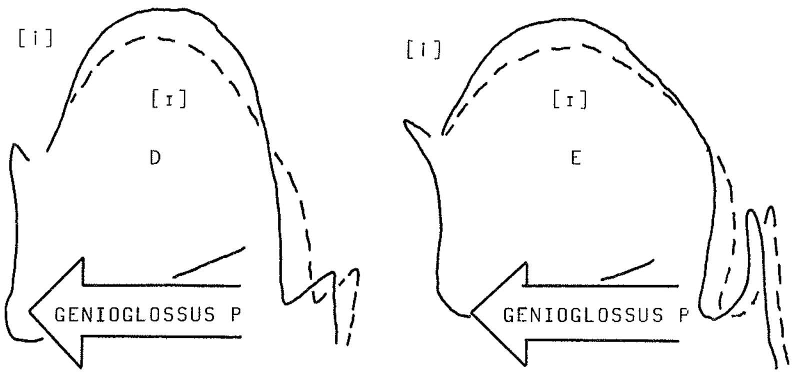


(b) Lip spreading for [i], lip rounding for [u] and lip approximation for [p, b] by Southern British English (D), Cairo Arabic (E) and West Greenlandic (F) subjects.

Fig. 1. Lip rounding and lip spreading for vowels and lip approximation for labial stops. Subjects A-C by courtesy of Professor Eli Fischer-Jørgensen, Copenhagen. Subjects D-F are the same subjects as in Wood (1979).



(a) Tongue profiles relative to the mandible for "tense" [i] and [y] (Danish) showing tongue blade raising and larynx lowering for [y]. Subject B's tongue was also less raised relative to the mandible for [y].



(b) Tongue profiles relative to the mandible for "tense" [i] and "lax" [ɪ] in Southern British English (D) and Cairo Arabic (E).

Fig. 2. Tongue articulation relative to the mandible for palatal vowels. Subjects A and B are traced from x-ray photographs by courtesy of Professor Eli Fisher-Jørgensen, Copenhagen. Subjects D and E are the same subjects as in Wood (1979).

Further, three parameter model nomograms (Stevens and House 1955, Fant 1960) show that the zones where F_2 is least sensitive to constriction location shifts are dependent on the degree of mouth opening, such that they are displaced anteriorly away from the glottis when the lips are rounded (Gunnilstam 1974). Does simultaneous lowering of the larynx (i.e. of the glottis) ensure that the anterior zone remains in the palatal region just as for [i]? In addition, a raised tongue blade in [y] prolongs the prepalatal constriction anteriorly. Is the speaker striving to keep this insensitive zone within striking distance of the tongue?

Each manoeuvre is studied with respect to its effect on formant frequencies, spectral flattening (proximity of F_3 to F_2), spectral sensitivity to constriction shifts and sensitivity of F_2 to area perturbations in the corresponding part of the vocal tract. The acoustical consequences are studied by manipulating [i] and [u] vocal tract configurations traced from x-ray films (the articulatory components of [y] are selectively added or subtracted) and then calculating and evaluating the spectral contribution at each step.

PARAMETERS STUDIED

The lips in [y] and [u]

Whenever differences of lip articulation have been reported for [y] and [u] they have always revealed less close lip rounding for [y] (McAllister et al. 1974, Hadding et al. 1976, Lubker et al. 1977 for Swedish; Brichler-Labaeye 1970, Benguerel and Cowan 1974, Riordan 1976 for French; private communication from Eli Fisher-Jørgensen for Danish and German; see Fig. 1a). For Swedish, this difference has traditionally been denoted outrounding (for [y, ø]) and inrounding (for [ʉ, u, o]).

Lyttkens and Wulff (1885) described outrounding as more protrusion with rounding from the sides, while inrounding also involved more approximation. This was interpreted for the modelled lip positions as *protrusion without approximation* versus *protrusion with approximation*. The Danish and German lip profiles in Fig. 1a had less protrusion and less approximation for [y], more protrusion and more approximation for [u]. Similarly, the Swedish data illustrated by Lubker et al. (1977) exhibits the same tendency - less of both factors for [y], more of both for [u], just like Fig. 1a. Swedish [ʉ] was not represented in this data. It was noted above that the zone where F_2 is least sensitive to constriction location perturbations is shifted away from the glottis by lip rounding. The closer the lips are rounded for [y], the greater is the anterior shift of this zone. One possible reason for avoiding close lip rounding for [y] is that beyond a critical degree of lip rounding this zone will advance beyond the prepalatal constriction, causing the palatal tongue gesture to have increasingly erratic spectral consequences.

The acoustical consequences of lip activity were modelled by simulating four different lip positions (spread, neutral, moderately rounded and closely rounded). Lindblom and Sundberg's (1971) lip model was used. The vertical muscular component H_m was set to +5 mm for spread lips, 0 mm for neutral lips and -2 mm for full approximation. The horizontal muscular component W_m was set to +10 mm for spread lips, 0 mm for neutral lips and -10 mm for full protrusion. These settings are based on lip data obtained from x-ray films and from Lindblom and Sundberg's recommendations. The moderate degree of rounding was set at full protrusion but no approximation, the close degree of rounding was set at full protrusion and full approximation.

As mentioned above, lip data published by Lubker et al. (1977), like the Danish and German examples in Fig. 1a, show instead less protrusion and less approximation for [y], more protrusion and more approximation for [u]. However, both approaches yielded similar lip conductivity indices (A/ℓ cm, Stevens and House 1955, Fant 1960), so this detail should not have influenced the results. Resonance conditions for intermediate positions can be interpolated from the results.

The tongue in [y] and [i]

Published x-ray tracings reveal language-specific preferences for either prepalatal or midpalatal tongue positions for palatal vowels (Wood 1975, 1979). Languages contrasting [y] with [i] preferred the prepalatal position. F_3 is higher and the timbre sharper for prepalatal [i] (e.g. Swedish) than for midpalatal [i] (e.g. English) (Fant 1960). Both the prepalatal and the midpalatal constrictions are formed in the zone where F_2 of palatal vowels is least sensitive to location shifts. The F_2 frequency of [i] is very similar at either location and is consequently stable against location shifts. But the prepalatal location is close to an F_3 pressure antinode and the midpalatal location is close to an F_3 pressure node, hence the large difference in F_3 frequencies between the two [i] sounds.

There may be additional narrowing anteriorly to the prepalatal constriction for [y], for example in Danish (private communication from Eli Fisher-Jørgensen and Nina Thorsen, Copenhagen). This is achieved by raising the tongue blade rather than by fronting the tongue body (Fig. 2a). Two of the ten profiles in Fig. 3 exhibit a slightly raised tongue blade for [y]. Whatever the benefit of this manoeuvre, there is a cost. Tongue blade elevation narrows the vocal tract at an F_3 pressure antinode and will consequently raise the frequency of F_3, thereby endangering the flattening of [y].

All ten x-rayed profiles in Fig. 3 show the tongue body to be slightly lower for [y] than for [i]. This lingual difference is largely relative to the mandible (see Fig. 2a, subject B). The result is a less constricted palatal passage and a narrower pharynx, yielding a higher F_1 and lower F_2. The lower tongue body for [y] has also been reported for French by Brichler-Labaeye (1970) and for Dutch by Raphael et al. (1979). Raphael et al. studied differences in EMG related to differences of palatal constriction in [y] and [i] by their Dutch informant and found weaker posterior genioglossal activity for [y], confirming that this tongue height difference is primarily *lingual* (i.e. tongue articulation relative to the mandible) and not *mandibular* (i.e. carrying the tongue down by lowering the mandible).

Tongue retraction for [y] is sometimes said to compensate for labial undershoot by lowering F_2. Published x-ray tracings provide no evidence for retraction in [y] (Fig. 3). Nor is there theoretical support for tongue retraction to lower F_2 within the palatal region, where location shifts hardly effects F_2.

Variation of constriction location, especially the prepalatal and midpalatal locations and retraction towards the velum, was modelled by systematically altering a prepalatal tongue body configuration into a velar configuration in ten equal steps (Fig. 4). A prepalatal [i] configuration and a velar [u] configuration were selected from an x-ray film. Both had the same jaw opening and the same degree of constriction (cross-section area of the constricted passage 0.5 cm^2). The high larynx position of the [i] was copied onto the [u]. The two configurations were superposed to provide a composite model. Seven other tongue profiles were then interpolated with constrictions located at even intervals between the original prepalatal and velar configurations. Finally, a tenth profile was extrapolated anteriorly to the original prepalatal constric-

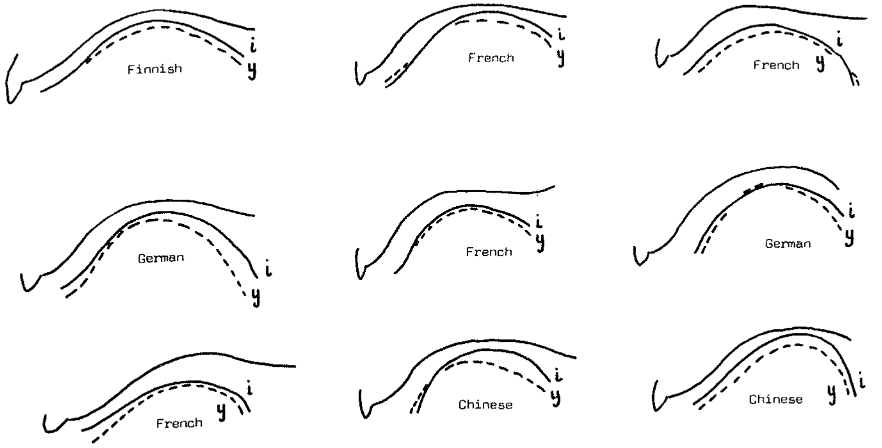
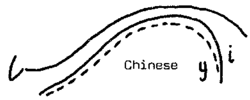


Fig. 3. A selection of [i] and [y] profiles collected from the literature (Wood 1975a) showing that the tongue is not more retracted for [y] (but it is slightly lower).



tion. The cross-section area at the constriction was constant for all profiles, 0.5 cm^2. Numbering from the front, positions nos. 2 (prepalatal), 4 (midpalatal) and 10 (velar) are attested by x-ray films as natural tongue body targets for monophthongs in speech (Wood 1979). The other positions occur as momentary passing places during lingual diphthongs or in transitions between vowels and consonants.

Tongue blade elevation in [y] was modelled by raising the tongue blade on the model profile to narrow the palato-alveolar region in even steps until the cross-section area was uniformly the same as in the constricted palatal passage. For this [y] configuration, the lips were moderately rounded, the constriction prepalatal and the larynx low (i.e. the full complex of [y] manoeuvres).

Two different tongue heights for [y] were modelled by performing the tongue blade experiment with two different degrees of pre-

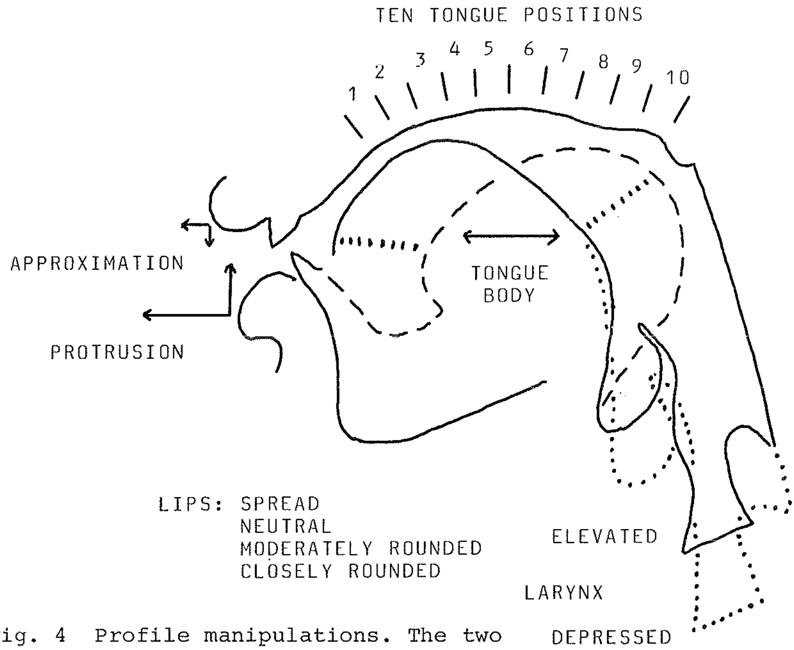


Fig. 4 Profile manipulations. The two extremes of tongue position modelled. Position no. 2 is prepalatal and no. 4 midpalatal. Intermediate positions were interpolated in even steps. The degree of constriction was constant for all profiles, cross-section area 0.5 cm^2. Profiles 2 and 10 are taken from an x-ray film.

palatal constriction cross-section area 0.5 cm^2 and 1.0 cm^2. This and the associated pharyngeal difference were modelled by lowering the tongue on the modelled profile relative to the mandible (cf. subject B in Fig. 2a, subjects D and E in Fig. 2b). This adjustment also conforms to the posterior genioglossal data of Raphael et al. (1979).

The larynx in rounded vowels

The larynx is lower for rounded vowels than for corresponding spread lip vowels (Brücke 1856, Perkell 1969, Sundberg 1969). Several investigators, reporting correlations between larynx height and voice fundamental frequency (reviewed by Barbier 1978) have data that also confirms the larynx to be lowest for rounded vowels. X-ray tracings indicate a range of 15 mm for vertical larynx movement in speech, while trained singers (who have learned to control their larynx position voluntarily) may utilize a larger range (Sundberg and Nordström 1976).

One possible reason for depressing the larynx in [y] has already been proposed above, namely that the zone where F_2 is least sensitive to constriction location shifts in palatal vowels should remain in the prepalatal region despite the lip rounding in [y] (the extra distance from the glottis to the least sensitive zone being taken up at the laryngeal end).

Riordan (1977) has proposed that the formant frequency lowering associated with larynx depression compensates for labial undershoot. This proposal is contradicted by acoustical theory (since lip rounding and larynx depression affect resonance conditions differently) and by the negative findings of Tuller and Fitch (1980). This will be discussed further in connection with the results.

The depressed larynx position was modelled by lengthening the vocal tract posteriorly by 15 mm (Fig. 4). A 5 mm segment was inserted immediately outside the larynx and another 10 mm distributed over the oropharynx above the epiglottis. This matched vertical larynx movement observed on x-ray films.

Scale factor

A different issue concerns scalar differences, for example between children and adults. Is the complex relationship between the individual articulatory components of [y] and vocal tract resonance conditions constant during life? Or does a speaker need to revise the articulatory programme for [y] during growth in order to accommodate progressive changes of scale and proportions? Similarly, do different people with different vocal tract sizes require different strategies? Further, is there a paradox in the fact that a shorter tract with depressed larynx can have the same overall length as a longer tract with elevated larynx and yet exhibit the different resonance conditions that are more suitable for [y]?

All the experiments were repeated with a shorter and a longer vocal tract (corresponding to a young adolescent and a long-necked man). When the larynx was depressed in the shorter model the overall vocal tract length was the same as in the standard model with elevated larynx. When the larynx was depressed in the standard model the overall length was the same as in the longer model with elevated larynx.

The length of the standard model (elevated larynx) was 14.5 cm from the glottis to the central incisors (i.e. excluding the lips). The three x-rayed male subjects reported in Wood (1979) had 14.5 cm or 15 cm for [i]. For the longer model an additional 15 mm was distributed over the oropharynx above the epiglottis, giving a corresponding length of 16 cm for [i]. This is appropriate for an adult male with a long neck. Fant's (1960) Russian subject, Chiba and Kajiyama's (1941) Japanese subject and Perkell's (1969) American English subject had 15.5 - 16.5 cm for [i]. The shorter tract was modelled by removing 10 mm from the oropharynx and 5 mm from the mouth region, giving a length of 13 cm for [i]. This is appropriate for a young adolescent (see Nordström 1975 for typical juvenile vocal tract dimensions).

ACOUSTICAL THEORY

The entire vocal tract participates in the tuning of each resonance. The effect of perturbations in all parts of the tract must therefore be taken into account when discussing the effect of a manoeuvre on a particular spectral feature in the speech wave. In the case of [y] there are the lips (that are

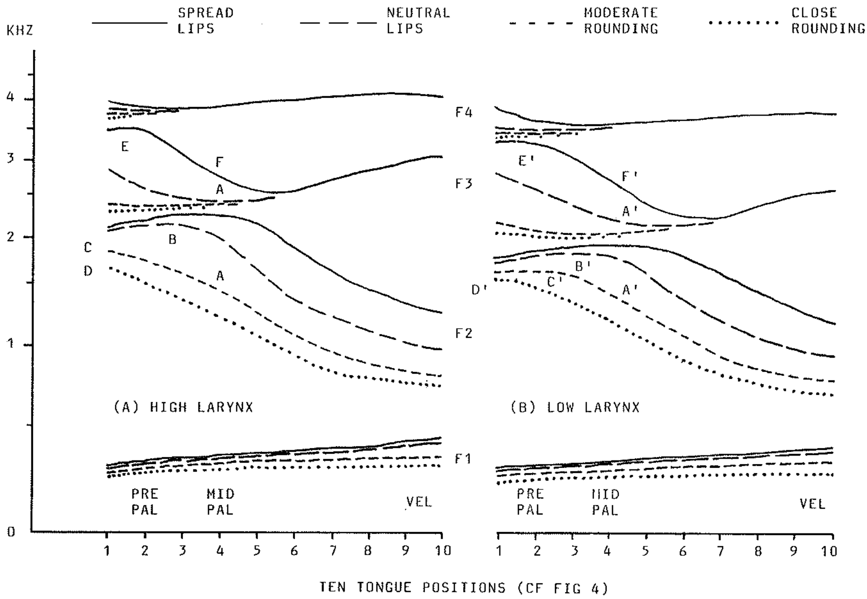


Fig. 5. Formant frequencies for the 10 tongue positions and for four lip positions with high larynx (left) and low larynx (right). The letters denote features referred to in the text.

constrained to moderate rounding), the tongue blade (that regulates the passage past the tooth ridge), tongue body position (that is constrained to the prepalatal constriction location), tongue body height (that regulates the passage along the constriction and in the pharynx) and larynx height (that is constrained to the low position).

These manoeuvres cause local area and length perturbations of the vocal tract and a fruitful way of quantifying their effect on the vocal tract transfer function is to relate them to the standing waves of the resonance modes.
Resonance frequencies are modified by perturbing the cross-section area locally at or near the nodes or antinodes of the standing waves, or by perturbing the length (Chiba and Kajiyama

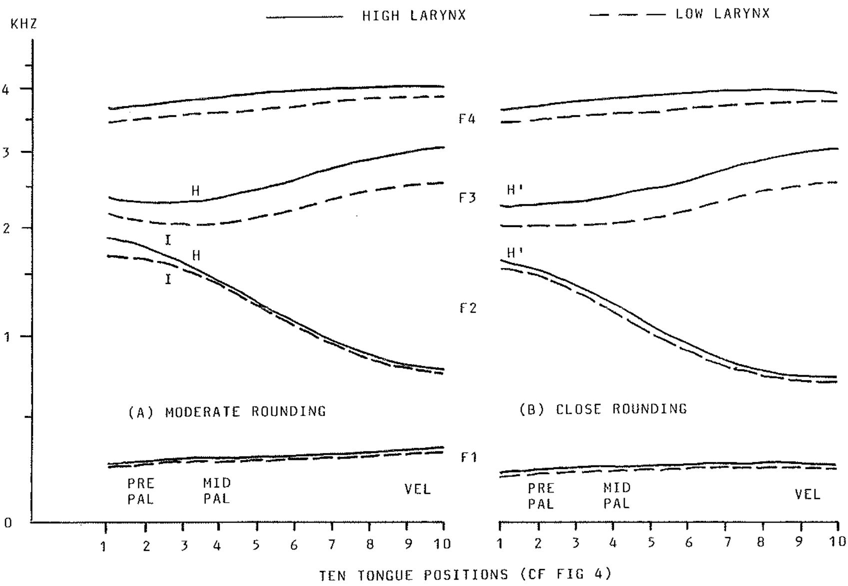


Fig. 6. Formant frequencies for the ten tongue positions and high (———) and low (----) larynx positions for moderate rounding (left) and close rounding (right).

1941). The sensitivity of a resonance mode to such a local perturbation is proportional to the local difference between kinetic and potential energies (Fant 1975, 1980; Fant and Pauli 1975). The local energy difference thus provides a useful measure of the sensitivity of a resonance mode to manoeuvres that widen or narrow the vocal tract locally, viz. lip approximation, tongue blade elevation, tongue body movement in the mouth and in the pharynx. The sensitivity of a resonance mode to a local length perturbation is proportional to the total reactive energy stored locally (i.e. the local sum of potential and kinetic energies). The total energy distribution is thus a measure of sensitivity to manoeuvres that lengthen the vocal tract locally, viz. lip protrusion and larynx depression.

The frequency of F_2 is a useful indicator of vowel quality in a pair like [i]-[y], where part of the contrast lies in the F_2 difference. Further, an essential part of spectral flattening is to bring F_3 down close to F_2; I assume there is a tolerable limit for how far F_3 can be permitted to rise away from F_2 without endangering the plain-flat contrast. Much of the discussion of the spectral consequences of the component manoeuvres of [y] and of compensations will therefore be concerned specifically with F_2 and F_3.

Vocal tract area functions, resonance frequencies and kinetic and potential energy distributions are computed as in Wood (1979).

RESULTS

Only the results for the standard model are given in the figures. The results for the longer and shorter models are not illustrated, with the exception of one example at Fig. 9. The results for all three models were essentially the same apart from a proportionate shift in absolute frequencies (as in natural speech). The conclusions drawn in the following discussion are valid for all three models.

Constriction location and the plain-flat contrast

How is spectral flattening (proximity of F_3 to F_2) related to tongue position (prepalatal or midpalatal)? For [i]-like configurations (spread lips and high larynx, continuous line in Fig. 5a), F_3 was higher with the prepalatal tongue position (E in Fig. 5a) and lower with the midpalatal position (F in Fig. 5a). This agrees with natural speech, for example the "sharper" Swedish /i:/ with its high F_3 compared with the "duller" English /i:/ with its lower F_3.

Lip rounding (dashed and dotted lines) lowered F_3 more than F_2 with the prepalatal tongue position (by about 1000 Hz and 500 Hz respectively, relative to spread lips), bringing F_3

close to F_2 for the flattened spectrum of [y]. With the midpalatal tongue position the situation was the reverse: F_3 was already low in [i] (F in Fig. 5a) and lip rounding actually shifted F_2 and F_3 farther apart (A-A in Fig. 5a). This was also the case when the larynx was lowered (F', A'-A' in Fig. 5b).

Consequently, lip rounding cannot flatten the spectrum with the midpalatal tongue position. Only the prepalatal position is suitable for the plain-flat contrast. Further, this means that tongue retraction from prepalatal [y] would be disastrous for the plain-flat contrast.

Tongue position, degree of lip rounding, larynx height and the sensitivity of F2 to location perturbations

How sensitive is the frequency of F_2 to tongue location perturbations? Is F_2 of [y] more resistant to location perturbations when the larynx is lowered, as hypothesized in the introduction?

Figure 5a gives the results for elevated larynx. For spread lips ([i]-like configuration, solid line), F_2 varied by as little as 5 Hz per mm of constriction displacement within the prepalatal or midpalatal region (locations 2 and 4, Fig. 5a). This is expected from Stevens (1972). For both prepalatal and midpalatal tongue positions the F_2 frequency of [i]-like vowels is very similar and is not particularly sensitive to location perturbations within that region. For more posterior tongue body positions (locations 6, 7 etc.), F_2 was very sensitive and fell by about 50 Hz per mm of retraction.

With neutral lips (broken line), the least sensitive zone was more anterior (B in Fig. 5a), as anticipated for the narrower mouth opening. For the two rounded conditions, the least sensitive zone was considerably advanced (C and D, off the diagram in Fig. 5a). This confirms that this zone advances away from the glottis when the lips are rounded.

Fig. 5b gives the corresponding results for depressed larynx. The least sensitive zone remained in the palatal region for the prepalatal configuration (location 2) for neutral and moderately rounded lips (B', C' in Fig. 5b) but not for close rounding (D'). On the other hand, for the midpalatal configuration (location 4), F_2 was still very sensitive to location perturbations with any degree of lip rounding. Larynx depression did not help in this case.

The prepalatal tongue position, moderate lip rounding and depressed larynx thus combine to provide stability of F_2 in [y] against location perturbations. This is the combination of manoeuvres preferred for [y] in natural speech. The midpalatal position is less favourable for [y] with respect to location perturbations, even with depressed larynx. As noted, the midpalatal position is avoided for [y] in speech. With close lip rounding, the least sensitive zone is very advanced even with depressed larynx. With respect to location perturbations, close lip rounding is less favourable for [y] at either tongue position and with any larynx height. Again, close lip rounding is avoided for [y] in speech.

Tongue position, degree of lip rounding, larynx height and the sensitivity of spectral flattening to location perturbations

An important part of spectral flattening is that F_3 is very close to F_2. Some variation of actual frequencies may be acceptable provided F_2 and F_3 remain close together, i.e. provided they rise and fall together. How do the components of [y] affect the sensitivity of spectral flattening to location perturbations?

The results for moderate lip rounding with high and low larynx are illustrated in Fig. 6a. With high larynx (solid line), any retraction of the tongue from the prepalatal position (location 2) caused F_2 and F_3 to diverge (H-H in Fig. 6a), i.e.

spectral flattening is weakened. When the larynx is depressed for [y] (broken line), F_2 and F_3 remain close together and parallel in the prepalatal region (I-I in Fig. 6a). Consequently location perturbations would have little effect on the spectral flattening of prepalatal moderately rounded [y] within this region provided the larynx is depressed.

The results for close lip rounding are illustrated in Fig. 6b. Tongue retraction anywhere in the palatal region (locations 1 to 4) caused F_2 and F_3 to diverge (H'-H' in Fig. 6b) whether the larynx was high or low. Consequently, location perturbations would lead to variation in the degree of spectral flattening if the lips were closely rounded.

Moderate lip rounding with the larynx depressed is thus the only favourable combination for preserving the flattened spectrum of prepalatal [y] against location perturbations. Once again, this is the combination preferred in natural speech. With elevated larynx or closely rounded lips, the degree of flattening is sensitive to location perturbations.

Degree of lip rounding, larynx height and the sensitivity of F2 to area perturbations

How does the combination of manoeuvres for [i] and [y] affect the sensitivity of F_2 to local area perturbations (from e.g. variation of lip opening or tongue blade elevation or tongue height along the hard palate or tongue root movement in the pharynx)? The discussion will be limited to the prepalatal configuration as the midpalatal configuration has already been found unsuitable for the plan-flat contrast.

The sensitivity of F_2 to area perturbations in a prepalatal vowel is illustrated in Fig. 7. Figure 7a is for high larynx, Figure 7b for low larynx. The various curves are for different lip positions.

With the spread lips and high larynx of [i] (solid line, Fig.

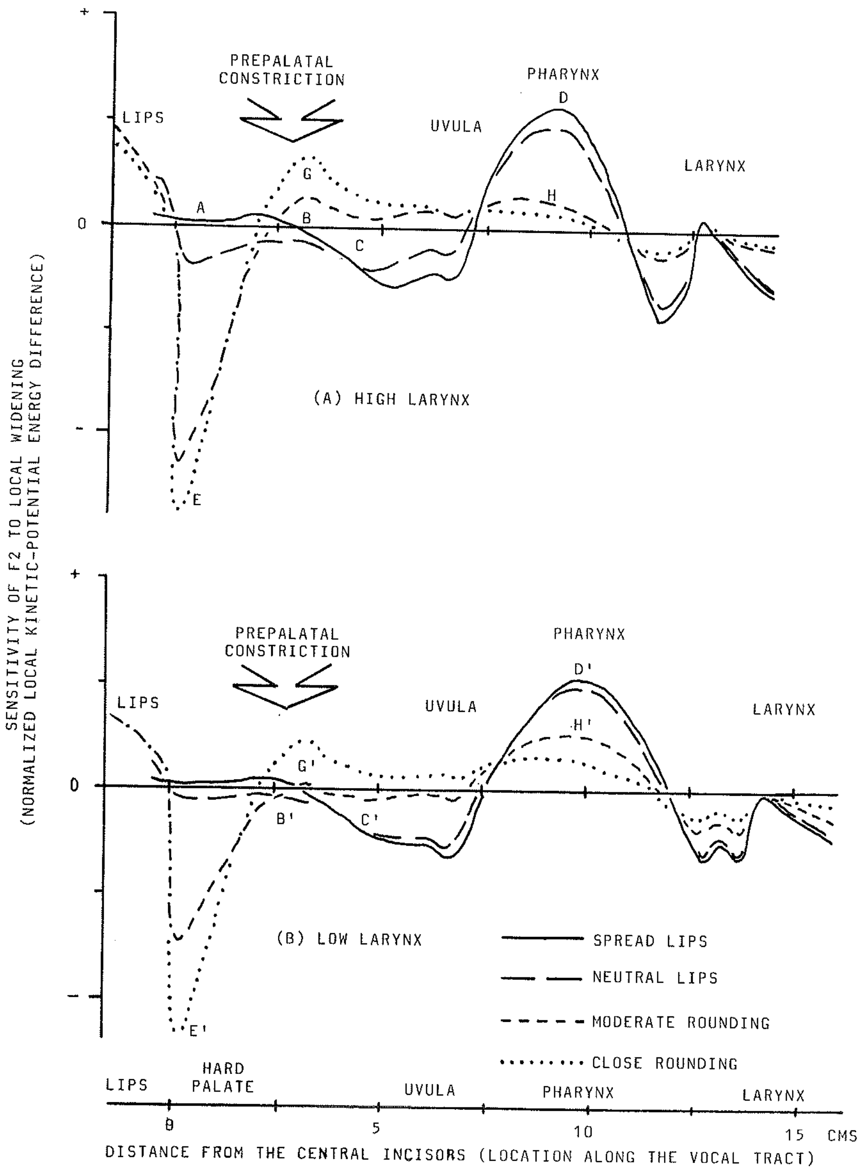


Fig. 7. The effect of lip position on the sensitivity of F_2 to local area perturbations (prepalatal vowel; high larynx (a), low larynx (b)). Local widening shifts F_2 in the direction of the sensitivity sign (vice versa for local narrowing) in proportion to the deviation from 0.

7a), F_2 is hardly sensitive to area perturbations anteriorly to the prepalatal constriction (A in Fig. 7a) or at the constriction itself (B in Fig. 7a). It is moderately sensitive posteriorly to the constriction (C) and very sensitive in the pharynx (D). F_2 of [i] is thus more sensitive to variation of pharyngeal width than to the degree of palatal constriction, as pointed out by Fant and Pauli (1975). In actual speech, tongue lowering of the type illustrated in Fig. 2 combines both perturbations simultaneously.

Before looking at what happens to the sensitivity of F_2 when the lips are rounded and the larynx lowered for [y] as in natural speech, it is instructive for this comparison to consider first the imaginary case that the larynx is high while the lips are rounded (Fig. 7a, dashed and dotted lines).

- The frequency of F_2 would become extremely sensitive to area perturbations in the alveolar and palatoalveolar region (E in Fig. 7a), i.e. to tongue blade elevation or depression
- Sensitivity would increase at the prepalatal constriction, but with a change of polarity relative to spread lips (G in Fig. 7a). This means that a given change in the degree of palatal constriction would shift F_2 in opposite directions with spread lips and rounded lips respectively
- In the pharynx, sensitivity to tongue root movement would diminish considerably with rounded lips (H in Fig. 7a)

Thus, if the larynx were high for [y] the sensitivity of F_2 to similar area perturbations (similar manoeuvres) would be different for [i] and for [y]. This would call for different control mechanisms (e.g. different compensation measures) in response to similar perturbations.

However, the larynx is normally depressed when the lips are rounded for [y]. How does this affect the sensitivity of F_2 to area perturbations when the lips are rounded? The results for a prepalatal configuration with depressed larynx are illustrated in Fig. 7b.

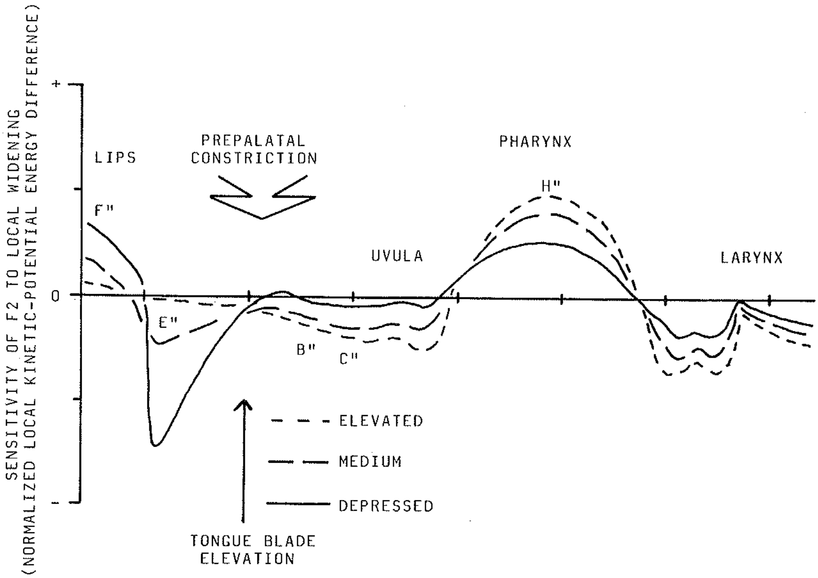


Fig. 8. The effect of tongue blade elevation on the sensitivity of F_2 to local area perturbations (configuration: [y]-like vowel with prepalatal constriction, moderate lip rounding, low larynx).

- The increase in sensitivity in the alveolar and palatoalveolar region is not so large when the larynx is depressed (cf. E' in Fig. 7b with E in Fig. 7a), especially with only moderate rounding (dashed lines)
- At the prepalatal constriction (G' in Fig. 7b), larynx depression prevents a change of polarity when the lips are moderately rounded (cf. G in Fig. 7a) but not when they are closely rounded (dotted line). Thus, when the larynx is depressed for [y] a perturbation of the degree of palatal constriction shifts F_2 in the same direction for [i] and *moderately* rounded [y], but the frequency shift would still be in opposite directions for [i] and *closely* rounded [y].
- In the pharynx (H' in Fig. 7b), sensitivity is a little higher for [y] when the larynx is low than when it is high (cf. H in Fig. 7a), especially for moderate rounding.

The effect of larynx depression on the sensitivity of F_2 to local area perturbations in [y] is thus to attenuate some of the increased sensitivity to tongue blade movement (especially

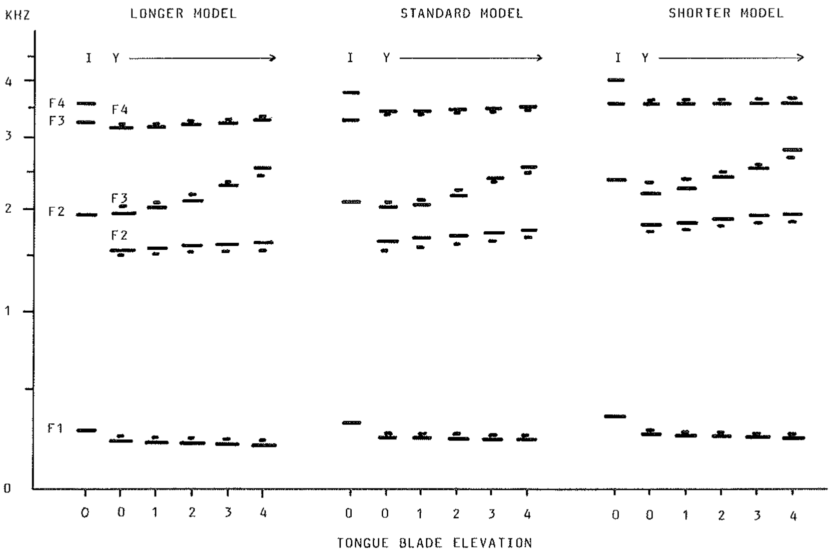


Fig. 9 Four steps of tongue blade elevation (1-4) from fully depressed (0) in a prepalatal [y]-like vowel (moderate rounding, depressed larynx) compared with a prepalatal [i]-like vowel (spread lips, high larynx). Cross-section area in the palatal constriction 0,5 cm^2 (———) and 1.0 cm^2 (----).

when only moderately rounded), to prevent a change of polarity of sensitivity at the prepalatal constriction (for moderate rounding but not for close rounding) and to restore sensitivity to tongue root movement in the pharynx (more so for moderate rounding than for close rounding). Larynx depression for moderately rounded [y] thus ensures similar F_2 shifts as for [i] in response to similar manueuvres in the palatal and pharyngeal regions.

Note that larynx depression does not yield these benefits for closely rounded [y]. Irrespective of larynx position, close lip rounding leads to considerable sensitivity of F_2 to tongue blade movement, sensitivity of opposite polarity along the hard palate (compared with [i]) and very reduced sensitivity in the pharynx (compared with [i]). Once again, close lip rounding is less suitable for [y].

This experiment highlights another important feature of resonance conditions in the vocal tract: a manoeuvre in one part of the tract can attenuate or enhance spectral sensitivity to some other gesture somewhere else in the vocal tract.
For example, lip rounding for [y] reduces the sensitivity of F_2 to tongue root movement in the pharynx; simultaneous larynx depression restores that sensitivity. This will be taken up again in a discussion of articulatory compensation and motor control.

Tongue blade elevation

How does the spectrum of [y] respond to tongue blade elevation, and how is the spectral contrast with [i] affected?

The influence of tongue blade elevation on the sensitivity of F_2 to local area perturbations in a moderately rounded [y]-like configuration with low larynx is illustrated in Fig. 8. Fig. 9 shows how the frequencies of the first four formants were modified by tongue blade elevation in [y]. Figure 10 shows the calculated frequency/amplitude envelopes for the two extremes of tongue blade movement modelled for [y], compared with [i].

Tongue blade elevation affected resonance conditions in [y] as follows:

- Sensitivity of F_2 weakened at the lips as the tongue blade was raised in [y] (F" in Fig. 8). This means that the degree of lip rounding becomes less critical and the spectrum of [y] less susceptible to perturbations to lip rounding. This is another example of how a local manoeuvre can influence the sensitivity of a resonance to local perturbations elsewhere in the tract.
- Sensitivity weakened in the palatoalveolar region (E" in Fig. 8). This is complementary to the weakening already provided by depressing the larynx (cf. E, E' in Fig. 7). F_2 becomes less sensitive to further tongue blade movement.
- Sensitivity increased at and posteriorly to the prepalatal lingual constriction (B", C" in Fig. 8) with the same po-

larity as in [i] (cf. B, C in Fig. 7a, spread lips). This is complementary to the similar effect already provided by depressing the larynx for [y] (cf. G', C' in Fig. 7b, moderately rounded lips) and offers further protection against a change of polarity that would otherwise have resulted from lip rounding. With the tongue blade elevated and the larynx lowered in [y] the sensitivity of F_2 to variation of the degree of palatal constriction is very similar for both [y] and [i].

- Sensitivity increased in the pharynx (D" in Fig. 8). This restores most of the sensitivity otherwise lost there by lip rounding (cf. D, H in Fig. 7a and D', H' in Fig. 7b). With the tongue blade raised and larynx lowered in [y] the sensitivity of F_2 to variation of pharyngeal width by the back of the tongue is very similar for both [y] and [i].

The general effect of raising the tongue blade in a moderately rounded prepalatal vowel with low larynx is thus to maintain similar sensitivity of F_2 to similar area perturbations in the palatal and pharyngeal regions for both rounded and spread lip palatal vowels. The previously noted advantages of larynx depression for [y] are reinforced. In addition F_2 of [y] becomes less sensitive to labial perturbations.

There is a cost, however. With the tongue blade depressed in [y], F_2 is very sensitive to tongue blade movement (E" in Fig. 8, solid line). The very act of elevating the tongue blade will raise F_2. Further, F_3 will also rise (there is an F_3 volume velocity node opposite the tongue blade). These consequences will shift the [y] spectrum towards [i].

How serious is this penalty?

Figure 9 confirms the rise in F_2 and F_3. It also demonstrates that the major spectral consequence of tongue blade elevation in [y] is that F_3 rises more sharply as the palatoalveolar passage becomes increasingly narrow. The interval between F_2 and F_3 gradually widens as the tongue blade is raised in [y], weakening the spectral flattening.

A rough indication of how far F_3 can tolerably rise away from F_2 in [y] in different languages without endangering the plain-flat contrast can be gleaned from the selection of published formant frequency data for /i/-/y/ contrasts in several lan-

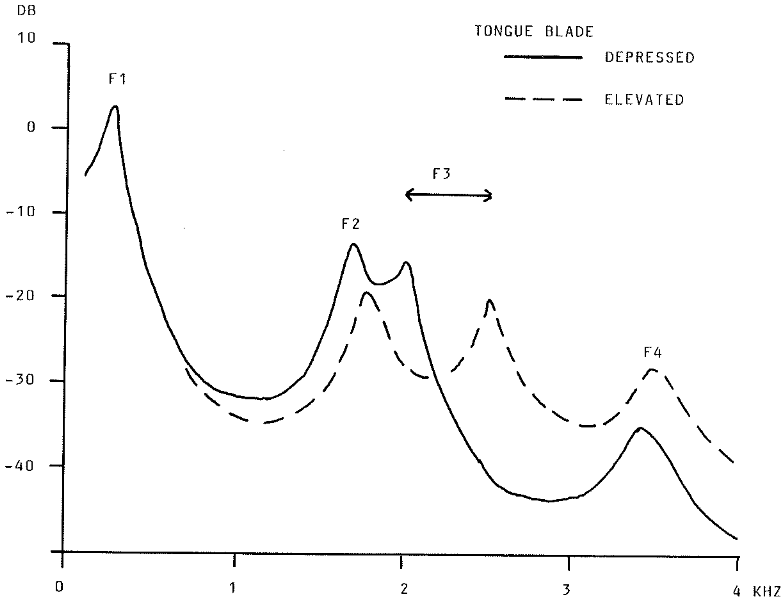


Fig. 10 The calculated spectrum envelopes of the two extremes of tongue blade elevation modelled, cf. Fig. 9 (configuration: [y]-like vowel, prepalatal constriction, moderate lip rounding, low larynx, constriction 1.0 cm^2).

guages quoted in Table I. These examples fall into two groups: (i) Danish, Dutch, French and German with F_3 of [y] (2000-2200 Hz for adult males) very close to F_2, and (ii) Swedish with F_3 of [y] (around 2600 Hz for adult males) midway between F_2 and F_4.

This much higher F_3 of Swedish [y] reflects a definitely elevated tongue blade corresponding at least to the final step in the modelled series (Fig. 9). Swedish /y:/ is produced as a lingual diphthong in most dialects; for both /i:/ and /y:/ the tongue is raised in a continuous movement towards the hard palate, often narrowing the prepalatal passage so much that tha vowel ends with a fricative off-glide, [ij] versus [yj] (for more details of the Swedish vowel system see Fant 1973, chapts. 5, 11, 12 and Lindblom and Sundberg 1969). The [yj] diphthong is spectrally similar to the sequence in the modelled series, a sharp rise in F_3 and a slight rise in F_2, indicating that tongue blade elevation is a part of the lingual manoeuvre (cf. Figs. 9, 10, 11).

The weakened flattening of Swedish /y:/ yields a more slender spectral contrast with /i:/, adequate for a native speaker but troublesome for a non-native learner. The slender spectral difference is related to the fact that phonologically Swedish /i:/ and /y:/ have to fit into a threeway contrast with another labial palatal vowel /ʉ:/ with lower F_2 and F_3 (about 1600-1700 Hz and 2100-2200 Hz respectively for adult males, see Fig. 11). This contrast is reinforced by /ʉ:/ having a strong inrounded labial diphthong component that causes F_2 and F_3 to fall, whereas the lingual diphthong component of /y:/ causes F_2 and F_3 to rise (Fig. 11).

Tongue body height

Published radiograms usually show a slightly lower tongue body position for [y] than for [i] (subject B in Fig. 2a and Fig. 3). In addition to widening the constricted palatal passage, this modification also narrows the pharynx. What does this contribute to the spectral contrast?

Figure 9 shows that the lower tongue body position for [y] (broken line) raised F_1, diminishing the F_1 contrast with [i]. Note that the F_1 frequency did not rise to that of [i] in this experiment, but in the data quoted in Table I 12 of the 19 examples had F_1 higher for [y] than for [i].
Whenever F_1 of [y] is roughly the same as or higher than F_1 of [i], this most likely reflects a more open palatal passage and narrower pharynx for [y] due to a lower tongue body.

The frequency of F_2 fell slightly when the tongue body was lowered, reinforcing the lowering already provided by lip rounding and augmenting the F_2 contrast with [i]. Note that this shift in [y] is in the same direction as would also be expected in [i], thanks to the low larynx position and only moderately rounded lips, as discussed above.

The frequency of F_3 was raised somewhat by lowering the tongue body for [y], provided tongue blade elevation was only slight. With more tongue blade elevation, as in Swedish, F_3 was slightly lower with the lower tongue body. In this case the more slender plain-flat contrast is reinforced.

The general consequence of lowering the tongue body for [y] relative to [i] are thus to retain the same F_1 as for [i] and to reinforce the lowering of F_2 for [y]. The interval between F_1 and F_2 is shortened for [y], reinforcing the flattening. Without tongue body lowering, F_1 would be lowered by the lip rounding of [y] and the total effect might then be considered as a general downward transposition of the entire spectrum rather than spectral flattening.

Laryngeal compensation for perturbed lip protrusion

Riordan (1977) has proposed that the formant frequency lowering associated with larynx depression compensates for labial undershoot in rounding. This conclusion was drawn from the negative correlation she found between larynx depression and upper lip protrusion. However, Tuller and Fitch (1980), in a similar experiment to Riordan's, did not find compensatory larynx movement when vowels were produced with varying perturbations to lip protrusion.

According to acoustical theory, lip rounding affects mainly resonances that depend on standing wave phenomena immediately behind the lips while larynx depression will mainly affect resonances dependent on the pharynx (Fant 1960, pp. 63-64). Lip protrusion and larynx depression will thus usually affect the vowel spectra differently. Only in the special case where the energy of a mode is stored uniformly at the lips and in the lower pharynx can lip protrusion and larynx depression be mutually compensating. What do the energy distributions look like

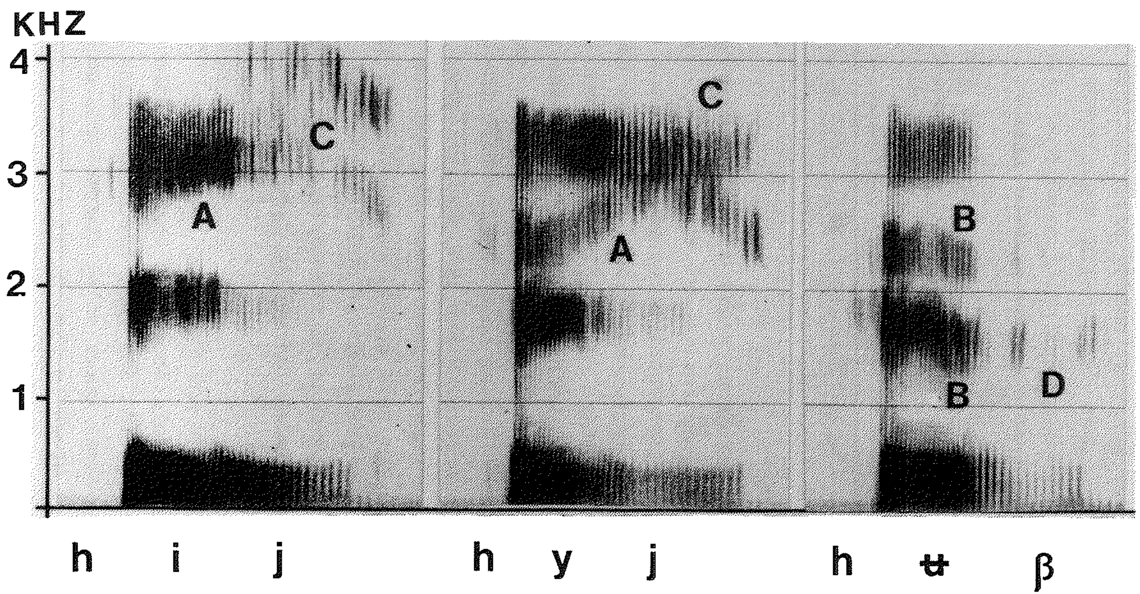


Fig. 11 Spectrograms of /hi:/, /hy:/ and /hʉ:/ by a speaker of a central Swedish dialect. Note that /i:/ and /y:/ are lingual diphthongs (rising F_3 from continuous tongue blade elevation A) and that /ʉ:/ is a labial diphthong (falling F_2 and F_3 from progressive rounding B). Note also the turbulent terminations of these vowels, with turbulence above 2500 Hz in /i:/ and /y:/ from the constricted palatal passage (C) and around 1500 Hz from the closely approximated lips of the "inrounded" /ʉ:/ (D).

at the lips and in the lower pharynx? Energy distributions for a set of Russian vowels have been published by Fant (1975, 1980), for a set of French vowels by Mrayati and Carré (1976) and for sets of English and Arabic vowels by Wood (1979).

These show only F_1 (for both [i] and [y]) to be equally sensitive to both manoeuvres. For [i], lip activity will mainly affect F_3 and larynx displacement mainly F_2 and F_4. For [y], lip activity will mainly affect F_2 and larynx displacement mainly F_3 and F_4. This can also be seen in Figs. 5 and 6. The two manoeuvres are thus complementary rather than mutually compensating, for both vowels.

My x-ray films (Wood 1979) and other model experiments (report forthcoming) show that, physiologically and acoustically, the lower lip alone very capably compensates for perturbations to the degree of rounding. Figures 1a and 1b illustrate how the lower lip performs the major share of labial articulation. Labial compensation can also be seen in the data of Hughes and Abbs (1976) and Lubker et al. (1977). Note that Riordan was constrained in her experiment design to observing only the upper lip. The acoustical correction she recorded can most likely be attributed to the uncontrolled lower lip and not to larynx movement. Note also the data quoted in this report that larynx depression is concomitant with lip rounding and not a substitute for it, and the findings of the experiments reported here that it is acoustically advantageous to lower the larynx when the lips are rounded.

Scale factor

The results from all three models show that each responded in the same way to each perturbation and that the conclusions are therefore general for any vocal tract size. One example comparing all three models is given in Fig. 9. It is not necessary for a speaker to modify the complex of [y] manoeuvres in order to compensate for changes in vocal tract size and proportions. The articulations developed during childhood are still appropriate for the growing adolescent and for the fully grown adult.

The striking scale difference - that resonance frequencies are inversely proportional to vocal tract length - is well known. This is tolerated in speech and is not compensated for. For examples see Table I.

Note that a shorter model with depressed larynx can have the same overall length as a longer model with raised larynx, and yet the resonance conditions are not the same, the acoustical effects of larynx lowering being present in the former case but not in the latter. On the other hand, the resonance conditions of each model, irrespective of absolute size, responded similarly to larynx depression. Scalar differences between individuals occur mainly above the epiglottis, whereas larynx depression also disproportionately lengthens and widens the lower pharynx. Sundberg (1977) has pointed out that this gives rise to an acoustical mismatch between the larynx and the remainder of the vocal tract, such that the quarterwave resonance of the larynx is radiated freely into the pharynx and is largely independent of the remainder of the vocal tract. This appears in the speech wave as a low fourth formant.

GENERAL DISCUSSION AND IMPLICATIONS

The lips

Lip rounding is usually attributed to the superior and inferior orbicularis oris muscles (which draw the corners of the mouth towards each other) and the mentalis muscle (which bends the lower lip outwards).

The orbicularis oris muscles have been investigated by Riordan (1976) for French and by McAllister et al. (1974), Hadding et al. (1976) and by Lubker et al. (1977) for Swedish. Lubker and Gay (1982) have compared coarticulation of lip rounding in American English and Swedish.

The lower tongue body for [y] than for [i] has been reported by several investigators. It is clearly not simply the mechanical consequence of depressing the larynx. A downward drag from larynx depression could easily be countered by more palatal raising of the tongue body (increased posterior genioglossal activity, cf. Fig. 2b), yet this well documented compensatory manoeuvre (Huizinga and Moolenaar-Bijl 1941, Lindblom et al. 1979, Gay et al. 1981) is obviously absent in this case. On the contrary, x-rayed profiles (such as subject B in Fig. 2a) show that there is usually less posterior genioglossal activity for [y] than for [i] (a lower tongue hump relative to the mandible due to a less advanced tongue root). This is confirmed by EMG data of Raphael et al. (1979) for Dutch. They found that the major contribution to the tongue height difference came from weaker genioglossal activity for [y], similar to that usually observed for the "tense-lax" difference between [i] and [ɪ] (cf. Fig. 2b). For comparison, differences in the mylohyoids (whose involvement in tongue elevation for vowels is somewhat controversial) and in the anterior belly of the digastric (a major mandible depressor in speech and consequently lowerer of the tongue with the mandible) were negligible between [i] and [y] (but very slightly weaker for [y] than for [i]).

Compensations and motor control

Compensatory mechanisms fall naturally into at least three classes: (i) restoration of critical features of vocal tract shape, (ii) substitution of alternative manoeuvres with the same spectral effect as the disturbed manoeuvres, and (iii) implementation of some manoeuvre that restores or nullifies spectral sensitivity to a manouevre elsewhere in the vocal tract.

The data reported by all these investigators shows an earlier and small EMG peak for [y] than for [u], indicating that the lip rounding activity was broken off at an earlier and lower level for [y]. This reflects the weaker lip rounding for [y] than for [u]. In view of the necessity for weaker lip rounding for [y], this difference is presumably universal.

The corresponding lip positions illustrated by Lubker et al. (1977) for Swedish agree with those illustrated here for Danish and German (Fig. 1a); the lips were more protruded and more approximated for [u], less protruded and less approximated for [y]. These examples do not suggest that the difference between *inrounding* and *outrounding* is *protrusion with approximation* versus *protrusion without approximation*.

All the investigators of Swedish quoted above reported a second orbicularis oris peak EMG amplitude in the inrounded set (i.e. including [u] but excluding [y]). This marks the additional activity needed for the extra labialized offglide of these vowels in Swedish. A similar second peak is not reported for [u] in the French or English data. The second peak in the Swedish inrounded vowels is therefore language specific.

The tongue

For the palatal tongue gesture there are no muscles that can pull the tongue up into a palatal position. It has to be pushed up from below, in contrast to the velar and pharyngeal vowel classes where it can be drawn into position by the styloglossi, glossopharyngei or hyoglossi and where the constriction can be located precisely by the sphincter function of the palataglossi or the pharyngeal constrictors (Wood 1979). The palatal gesture is generally attributed to contraction of the posterior fibres of the genioglossi, which draw the tongue root forward and force the tongue body upwards, widening the pharynx and narrowing the palatal passage. The intrinsic musculature of the tongue will also contribute to the shaping of the tongue and to the direction of the gesture (prepalatal or midpalatal). The prepalatal position is necessary for the plain-flat contrast.

The best documented examples of compensation are of the first type. For example, tongue elevation relative to the mandible (mainly posterior genioglossal activity) is balanced against mandible position in palatal vowels in order to maintain an ideal degree of palatal constriction (Huizinga and Moolenaar-Bijl 1941, Lindblom et al. 1979, Gay et al. 1981). The articulatory manoeuvre is the same as that in Fig. 2b. The consequence is instantaneous and precise (the first glottal pulse of a bite block vowel had the same formant frequencies as the corresponding unperturbed vowel). The precision is also illustrated by the small statistical variation of the degree of palatal constriction reported by Perkell and Nelson (1981). A similar example is the lower lip, which compensates for disturbed mandible position and restores the ideal degree of mouth opening in rounded vowels (Hughes and Abbs 1976, Wood forthcoming). The manoeuvre is similar to the examples in Fig. 1a. Both these compensations are important for [y] on account of the slightly lower tongue position than for [i] and of the critical degree of lip rounding (moderate, not close). In addition, if balancing the tongue against jaw position is to work for both [y] and [i] the resonances must respond similarly to the tongue movement in both rounded and spread lip palatal vowels. It was found above that sensitivity to area perturbations in the palatal region is similar thanks to the low larynx of [y].

Lindblom et al. (1979) point out that the immediacy of the tongue adjustment precludes utilising auditory feedback; only tactile and proprioceptive information is available. Similarly, Perkell (1980) has proposed that the motor goals should be expressed in orosensory terms. The common features defining this class of compensation are (i) the correction involves something more or something less of a movement that is already activated, and (ii) the movement is already actually involved in creating a critical degree of constriction. Examples of motor equivalence (Hughes and Abbs 1976) are usually taken from this class.

Regarding the second class (substitution of a different manoeuvre in a different part of the vocal tract with similar acoustical consequences to the disturbed manoeuvre), both tongue retraction and larynx lowering have been proposed as compensations for labial undershoot in [y]. The tongue profiles reproduced in Fig. 3 do not show any retraction for [y] and the model experiments reported here have shown that even just a little retraction would be disastrous for the plain-flat contrast, the midpalatal position being already unsuitable for [y]. The constriction must be prepalatal and it may even be prolonged anteriorly by raising the tongue blade. Larynx lowering as a compensation for labial undershoot is contradicted by acoustical theory and the experiments reported above have shown that larynx lowering is necessary during lip rounding in order to control resonance conditions. Even though these particular examples are invalid, this class of compensation may by useful in other situations. The common features that distinguish this class are (i) a different (even inactive) articulator takes over from the disturbed articulator, and (ii) the substitute articulator operates in a different part of the vocal tract, i.e. on a different part of the standing wave. This is a very different motor reorganization task from that involved for the first class, where on-going synergy between a disturbed manoeuvre and a companion manoeuvre ensures that a desired constriction is achieved. It should not be overlooked that while examples of the first class are well documented, no convincing examples of this second class have yet been reported.

The third class involves a manoeuvre that perturbs one part of the vocal tract in order to increase or reduce spectral sensitivity to another manoeuvre in some other part of the tract.

Examples found in the experiments reported above are (i) lip rounding for [y] changed the polarity of sensitivity of F_2 to the degree of palatal constriction and reduced the sensitivity of F_2 to variation of pharyngeal width (i.e. the spectral consequences of a given tongue movement would be contrary for [y] and [i]; (ii) simultaneous larynx depression restored the change of polarity at the palatal constriction and restored the lost sensitivity in the pharynx; (iii) simultaneous tongue blade elevation reinforced the effect of larynx lowering. These examples from [y] are a conceptually different type of compensation. The first two classes imply correction in response to an external disturbance that impedes the on-going activity. But lip rounding for [y] is not a disturbance in the same sense. It is an inherent part of the vowel, albeit with some undesireable side effects. The countermeasure for the side effects (larynx depression, and perhaps also some tongue blade elevation) is a simultaneous and integral component of the articulatory complex for [y]. This is the distinguishing feature of this class. The compensation is an integral part of the original motor programme and not a subsequent correcting response.

It is possible that the acoustical effects of larynx depression are continously correlated with the degree of lip rounding, e.g. the more the lips are rounded the lower the larynx, providing gradual correction. This has not been tested in the present experiments. On the other hand there is an acoustical discontinuity that indicates larynx depression is quantal. With the larynx depressed, the orifice of the larynx is wide enough for the quarterwave resonance of the larynx tube (F_4) to be radiated freely into the pharynx and be largely independent of the remainder of the vocal tract. Hence the paradox that two equally long vocal tracts can have different resonance conditions because the larynx is depressed in the one but elevated in the other. If larynx depression is quantal, the exact level would not be important. All that is necessary is to be above or below a critical position. Other factors in speech that tend to correlate with variations of larynx height, for example voice fundamental frequency, would not then interfere with the spectral effects of having the larynx high versus low for spread lip vowels versus rounded vowels.

A fruitful way of surveying the motor control of the vowels studied in this report is to review the programming task facing a child who is in the process of mastering the production of the plain-flat contrast in its mother tongue.

The tongue body will be raised prepalatally (not midpalatally) - French-speaking children will acquire a different lingual programme to English-speaking children in order to ensure the slight difference in direction of the palatal manoeuvre. The French child will discover that a midpalatal constriction is unsuitable for [y] since the plain-flat contrast cannot be produced there. The tongue body will not be quite so high as for [i] - the posterior fibres of the genioglossi will be programmed differently for each vowel. The critical limit of lip rounding for [y] will be found and the lip muscles will be programmed differently for [y] and for [u]. The child will learn to lower the larynx during [y] to prevent the tongue causing F_2 and F_3 to fluctuate unpredictably, and to ensure similar sensitivity of F_2 to area perturbations at the palatal constriction and in the pharynx for both [i] and [y] (otherwise F_2 would shift in opposite directions, depending on the amount of lip rounding, and different compensation responses would be needed to correct the aberration under different lip conditions). With the larynx depressed in [y], the child will find that balancing posterior genioglossal activity against mandible position to maintain the correct degree of palatal constriction works equally well for both [y] and [i]. The child may also discover that the critical degree of lip rounding can be mitigated somewhat by elevating the tongue blade, at the expense of raising F_3 and losing some spectral flattening.

Phonological considerations

The articulatory contrast between [y] and [i] has traditionally been captured with one feature, rounding. But the present problem has concerned the physiology of speech production as it governs the spectral coding of the speech wave. These are two separate levels of description that must not be confused.

What does the speaker do in order to put certain spectral characteristics into the speech wave? As reported above, the acoustical plain-flat contrast is produced with a complex of articulatory manoeuvres embracing not only the lips but also the larynx, the tongue body and the tongue blade. While the speaker has to execute the entire package of manoeuvres, the phonologist aims to give a non-redundant abstraction of speech.

These manoeuvres fall into two distinct processes. There is a primary process: the lips modify the vocal tract transfer function in order to produce the major formant shifts that constitute the linguistic spectral contrast, that is, the lips do the encoding for [y]. In addition, there is a stabilizing process: the remaining manoeuvres optimize resonance conditions to make the plain-flat contrast possible. Observe that this is not a question of redundancy in the usual sense of the term. Both processes are necessary, indispensible and not interchangeable. But since the supporting manoeuvres are automatic, it is not incorrect for the phonologist to take them for granted and concentrate on the encoding manoeuvre, lip rounding.

But the full complex of manoeuvres cannot be entirely disregarded, either by the speech scientist or by the phonologist.

The unwary experimenter who confuses the two levels of description runs the risk of generating unsound hypotheses. For example, one cannot necessarily expect more precise lip rounding for /y/ than for /u/ simply on the ground that /i/ is distinguished from /y/ only by the labial dimension whereas /u/ is differentiated from /i/ by both the labial and the lingual dimensions. Physiological detail cannot be uniquely predicted from a simple count of phonological features. At the physiological level [y] is distinguished from [i] by all the manoeuvres investigated above and the experimenter must be prepared to find effects from all of them. An alternative hypothesis predicting the same outcome, for example, is that the moderate [y] rounding may be more precise than the [u] rounding owing to the need to respect the limitation imposed by the acoustically critical degree of rounding. The problem is not made easier by the conflicting outcomes of different statistical measures of precision and by possible language-specific effects. For example, Riordan (1976) concluded from the smaller statistical variation of lip EMG that there was better precision for /y/ than for /u/ in French. On the other hand, Lubker and Gay (1982), using a different statistical measure, concluded that lip activity was more precise for /u:/ than for /y:/ in Swedish. Switching statistical measures between these sets of data would also entail switching conclusions.

Similarly, the phonologist who neglects physiological detail may overlook the possible mechanisms underlying alternations or sound change. For example, why do sound changes so often go in the direction *o-u-y*? Surely not for perceptual reasons since the change is into the crowded palatal area where spectral contrasts are already fine. And what is the next move for these vowels - *y, i, ɨ, iu, iy* (cf. French, Welsh, Chinese, Icelandic, English)? And how is the Swedish dialect /i:/ variant, the *Viby i,* produced? The complete shift from [y] to [i] not only means spreading the lips but also elevating the larynx and increasing the posterior genioglossal contraction. Failure to do the latter will lead to [ɪ]. An evolution to some other vowel may well involve carrying over some secondary manoeuvre from [y] to the evolved vowel.

CONCLUSIONS

1. Only the prepalatal tongue position is favourable for the plain-flat contrast. Prepalatal [i] has a high F_3 that is lowered more than F_2 by lip rounding so that F_2 and F_3 are close together in [y] (spectral flattening). The midpalatal position is unsuitable for this contrast. The frequency of F_3 is already low in midpalatal [i] and is lowered less than F_2 by lip rounding so that F_2 and F_3 would actually diverge (no spectral flattening).

2. Only moderate lip rounding is suitable for [y]. With more than moderate lip rounding, the spectral flattening of [y] becomes increasingly sensitive to lingual perturbations, F_2 and F_3 responding differently to tongue location shifts. The spectral consequences become increasingly unpredictable when lip rounding exceeds the critical limit.

3. Depressing the larynx for [y] keeps the zone, where F_2 is least sensitive to tongue location perturbations, within the prepalatal region. This zone is shifted away from the glottis by lip rounding, but simultaneous larynx depression (i.e. lowering the glottis) brings it back to the prepalatal region.

4. Depressing the larynx for moderately rounded [y] also ensures that F_2 and F_3 respond similarly to constriction location perturbations througout the prepalatal region. This safeguards the spectral flattening of [y] against location perturbations, provided the lips are not more than moderately rounded.

5. Depressing the larynx for [y] ensures similar spectral consequences of disturbances of the degree of palatal constriction and variations of pharyngeal width for both rounded and spread lip palatal vowels. Similar lingual movements would otherwise produce contrary spectral consequences for [y] and [i].

6. Resonance conditions are further stabilized by raising the tongue blade in [y]. However, only a limited amount of tongue blade elevation is tolerable since this also raises F_3, thereby weakening the spectral flattening. The tolerable limit is small in, for example, Danish, Dutch, French and German where a fully flattened [y] is contrasted with [i]. Swedish requires more tongue blade elevation in order to provide a higher F_3 for the partially flattened /y:/ which has to contrast with both /i:/ and /ʉ:/.

7. Tongue body lowering raises F_1 and cancels the fall that would have followed from the lip rounding of [y]. F_1 remains much the same as in [i]. This provides a common reference that enhances the flattening of F_2 and F_3.

8. Larynx depression cannot compensate for labial undershoot in [y]. The laryngeal contribution to the lowering of F_3 for spectral flattening is much smaller than the labial contribution. Lip rounding and larynx depression are complementary rather than mutually compensating.

9. The articulation of [i] and [y] does not need to be modified during growth in order to compensate for progressive scalar changes of vocal tract morphology. The relative consequences of lingual and labial manoeuvres and larynx height remain constant irrespective of scalar differences. Absolute resonance frequencies are shifted downward during growth, but experience shows that this is tolerated in speech perception.

RADIOGRAPHIC AND MODEL STUDIES OF THE PALATAL VOWELS

Sidney Wood

SUMMARY

Jaw openings, tongue positions and lip positions of palatal vowels in German, English, French, Russian and Arabic are analysed from x-ray profiles collected from the literature and from two x-ray films. A jaw opening of 8 or 9 mm divided a closer class [i,ɪ] from a more open class [e,ɛ] . The tongue assumed a more bunched and elevated posture relative to the mandible for tense vowels [i,e] and a lower, flatter posture for lax vowels [ɪ,ɛ] . The lips were spread more for tense vowels than for lax vowels. The acoustical consequences of these articulatory differences (especially the tolerances for variation and overlapping of manouevres and the acoustical efficiency of compensations) are investigated in a series of model experiments. Finally, some implications for phonology and speech production theory are discussed.

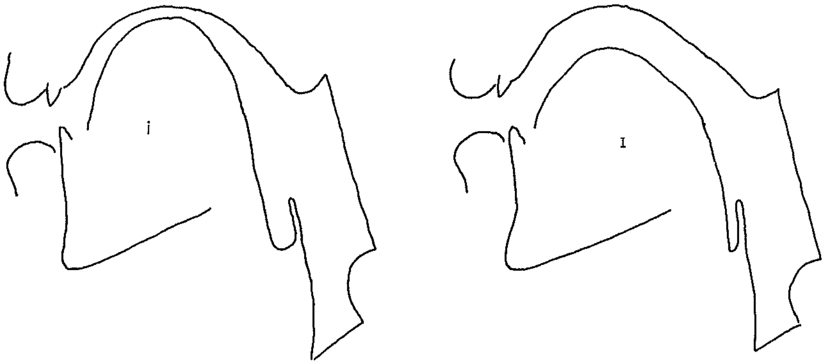
INTRODUCTION

In Wood (1979) it was found that the tongue narrows the vocal tract at one of four locations for monophthong vowels: (i) at the hard palate for [i-ɛ] and [y-œ]-like vowels, (ii) at the velum for [u-ʊ] and [ɯ]-like vowels, (iii) in the upper pharynx for [o-ɔ] and [ɣ]-like vowels and (iv) in the lower pharynx for [a-æ]-like vowels.

The present report is devoted to the palatal class of vowels, especially to x-ray and model studies of how spread-lip palatal vowels are differentiated (the articulation of rounded palatal vowels, exemplified by [y], and the language specific preference for either prepalatal or midpalatal constrictions were studied in Wood (1982).

Figures 1 and 2 illustrate typical profiles of palatal vowels traced from x-ray films of Southern British English and Cairo Arabic. Figures 3 and 4 give the F_1 and F_2 frequencies of all examples in this material.

The parameters available for tuning the resonances of the vocal tract for a palatal vowel are (i) the degree of mouth opening (determined by the lips and mandible), (ii) the degree of palatal narrowing (determined by the tongue and



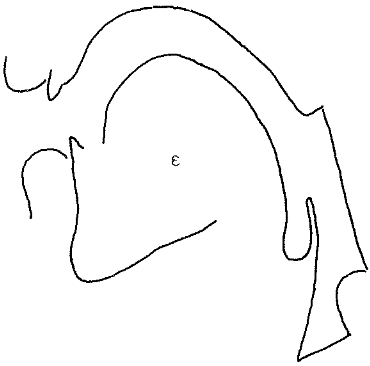


Fig. 1

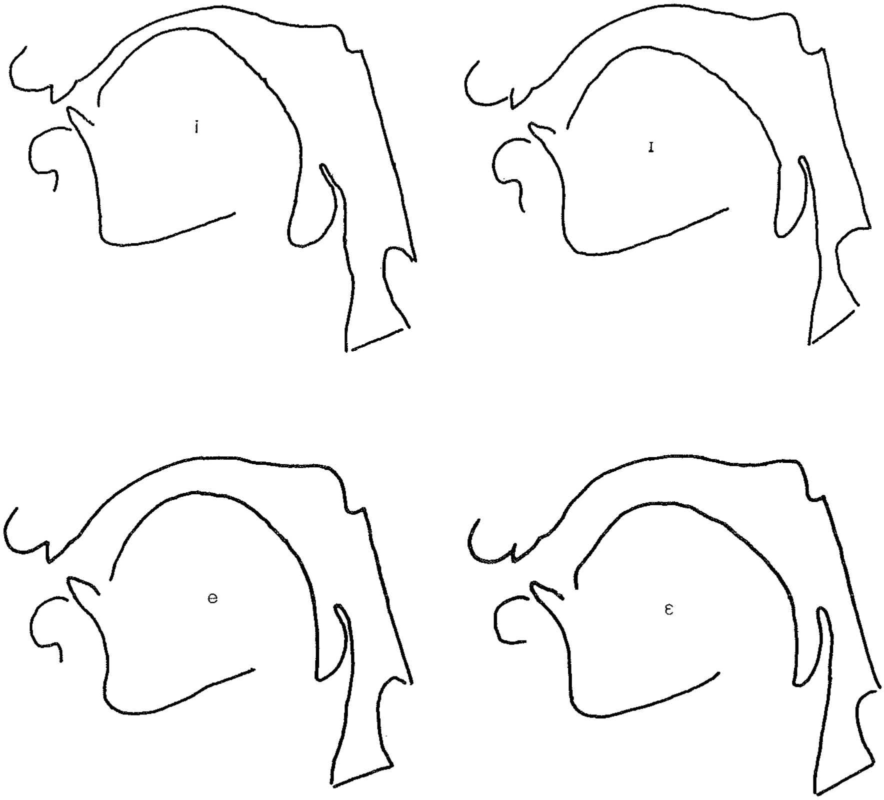
Profiles of S British English palatal vowels

Fig. 2 (opposite)

Profiles of Cairo Arabic palatal vowels

mandible), (iii) the degree of pharyngeal widening (determined by the back of the tongue), and (iv) larynx height (the larynx is usually high for spread-lip vowels).

Figures 3 and 4 give a rough indication of the relationship beteween articulation and the frequencies of F_1 and F_2 in palatal vowels. A grid of F_1 and F_2 frequencies generated by the Stevens and House (1955) three parameter model (for a constriction located at d_o=12 cm from the glottis) is placed over the observed palatal vowel F_1 and F_2 frequencies of the two subjects. This shows how F_1 and F_2 are related to the degree of mouth opening (A/ℓ cm) and the degree of palatal narrowing (A_{min} cm^2). Both formants rise as the mouth opening is enlarged (A/ℓ cm). F_1 rises and F_2 falls as the palatal



passage widens (A_{min} cm^2). The contribution of individual physiological parameters cannot be distinguished in these diagrams. A/ℓ corresponds to the combined effect of lip position and mandible position. A_{min} corresponds to the combined effect of tongue position and mandible position. Lindblom and Sundberg (1971) have studied the general effects of individual physiological parameters on vocal tract resonances. The model experiments desribed in the present report are concerned specifically with the manouevres analysed from x-ray profiles.

A point of particular interest is the quantization of the scales of tongue, mandible and lip position.

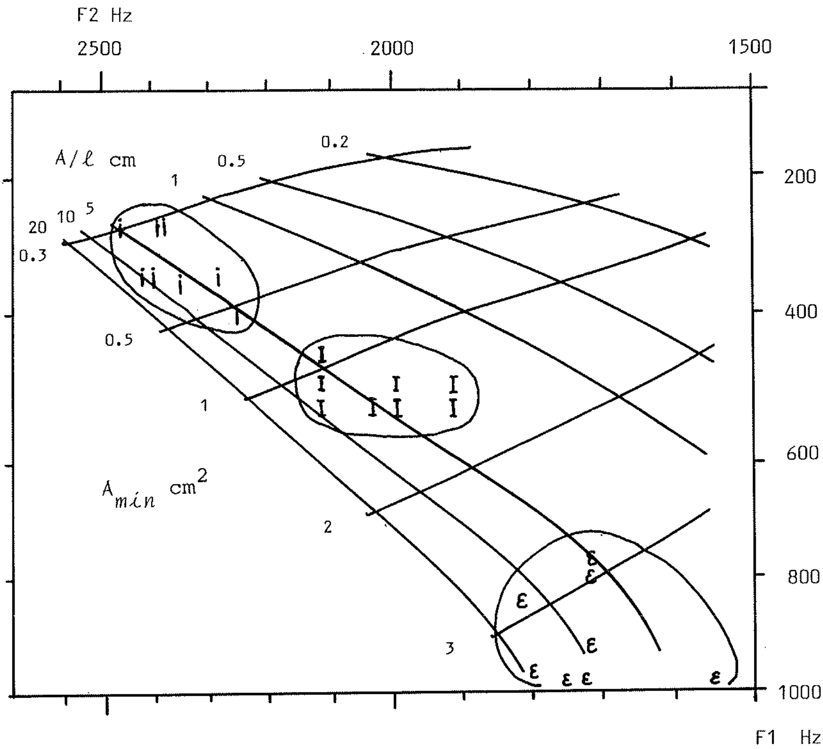


Fig. 3 F_1 and F_2 of stressed palatal vowels from the subject of the English film. The grid is based on the Stevens and House (1955) three parameter model nomograms. It shows how F_1 and F_2 of palatal vowels (palatal constriction, d_0=12 cm) are related to the mouth opening (A/ℓ cm) and the degree of palatal narrowing (cross-section area of the palatal passage A_{min} cm^2).

Stevens (1972) has demonstrated that in certain situations quantal relationships exist between articulation and acoustical features of the speech wave. Such a relation exists, for example, between the four constriction locations mentioned above and vowel formant frequencies (see also Wood 1979). Further examples, concerning larynx depression and the degree of rounding in rounded palatal vowels are given in Wood (1982). Regarding the degree of palatal narrowing, there is no quantal relationship of this type (Gunnilstam

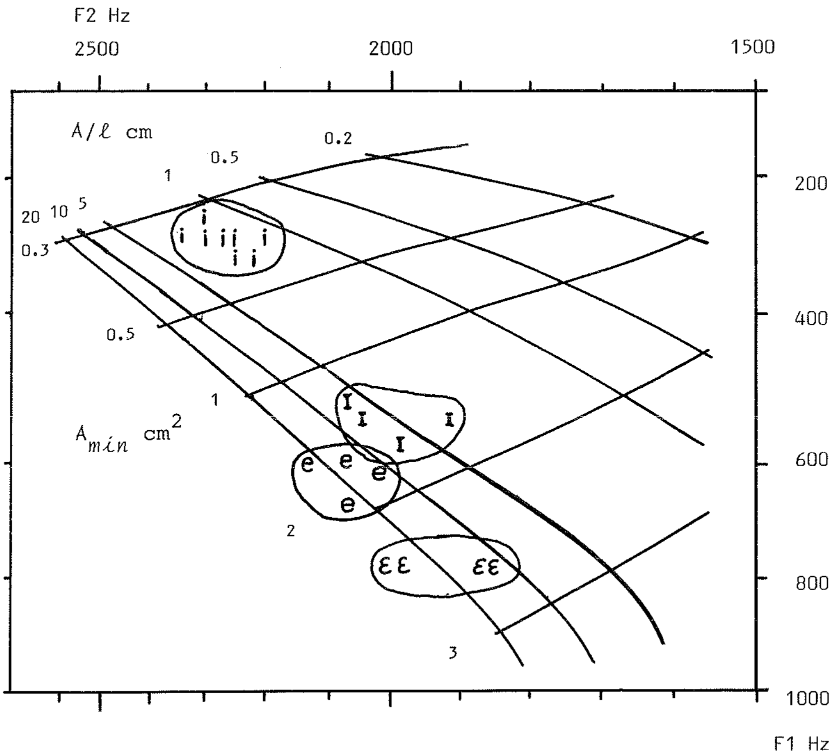


Fig. 4 F_1 and F_2 of stressed palatal vowels from the subject of the Arabic film. For details of the grid, see Fig. 3.

1974, Ladefoged et al. 1977, Perkell and Nelson 1981). The formant frequencies change continuously as the palatal passage widens in palatal vowels.

Traditionally, the continuous scale of "closeness" or "tongue height" is divided ad hoc into 2, 3 or 4 discrete divisions, depending on the language in question. The example of Chiba and Kajiyama's (1941) German subject analysed in Wood (1975a) suggested an alternative solution, see Fig. 5. The mandible was raised (narrower mouth opening and narrower palatal passage) for "close" [i, ɪ] and lowered (wider mouth opening and wider palatal passage) for "open" [e,ɛ] . The tongue was raised higher relative to the mandible (narrower palatal passage and wider pharynx) for "tense" [i, e] and lower (wider palatal passage and narrower pharynx) for "lax" [ɪ, ɛ] . A similar pattern in the English and Arabic x-ray films presented in this report can be seen in Fig. 6. These manouevres (close versus open mandible position combined with tense versus lax tongue posture) have been studied on mid-sagittal x-ray profiles and their acoustical contributions computed in model experiments.

A further point of interest is the extent of lip-mandible and tongue-mandible coordination and how far the lips or tongue can and do compensate for shifts of the mandible. It is well documented that the tongue compensates for shifts of the mandible in palatal vowels and thereby maintains an ideal degree of palatal narrowing (Huizinga and Moolenaar-Bijl 1941, Lindblom et al. 1979, Gay et al. 1981). Hughes and Abbs (1976) have reported similar coordination between the lips and the mandible to maintain an ideal degree of mouth opening. Attention has therefore also been given in this report to evidence of coordination and compensation in the x-ray profiles studied here.

PROCEDURES

The study is based on sets of midsagittal x-ray profiles of German, American English, French and Russian collected from the literature and 2 cinefluorographic films of Southern British English and Arabic, made in the angiocardiographic unit of the Lund University Hospital. References, methods of reproduction and normalization and details of the filming techniques are given in Wood (1975a, 1979).

The profiles collected from the literature were studied first. These constitute the evidence that has accumulated since the turn of the century. Subsequently, the two motion films were made under controlled conditions. These provide a means of testing hypotheses generated from the collection of published profiles.

In order to make comparisons, the material is grouped as follows:

(a) Published x-ray profiles of [i, ɪ, e, ɛ] in German

(b) Published x-ray profiles of [i, ɪ, e, ɛ]
in American English

(c) Published x-ray profiles of [i, e, ɛ]
in French

(d) Published x-ray profiles of [i, ɛ]
in Russian

(e) The English film with [i, ɪ, ɛ]

(f) The Arabic film with [i, ɪ, e, ɛ]

Experimental error will cause some variation in traced profiles, especially in the collection of published material. A further source of variation comes from pooling the speakers of the collected material to form the data groups (a-d) outlined avobe. There are instances of very few examples of some vowels, even after pooling speakers, which makes statistical treatment uncertain. The English film is better in these respects - 8 examples of each vowel by the same speaker. Despite these difficulties many of the tendencies observed in the pooled collected data were confirmed in the fresh one-speaker data.

The standard error of the difference between two means is used to test the statistical significance of differences between average jaw openings in different sets. The probability that the difference between two means is (i) at least twice its standard error is about 5%, (ii) at least three times its standard error is about 0.25%.

The acoustical consequences of a manouevre were studied by manipulating observed profiles (for example by altering the tongue, lip or jaw position as in Figs. 8, 11 and 16) and then calculating and evaluating the spectral contribution at each step. To ensure the realism of the manipulations, the manoeuvres were studied on the x-ray films. For example, tongue movement between the palatal tense and lax postures can be seen in Fig. 6. The resonance frequencies were computed from the area function of each configuration, using Webster's equation.

THE MANDIBLE

A traditional view is that the mandible is progressively lower (the mouth progressively more open) for vowels along

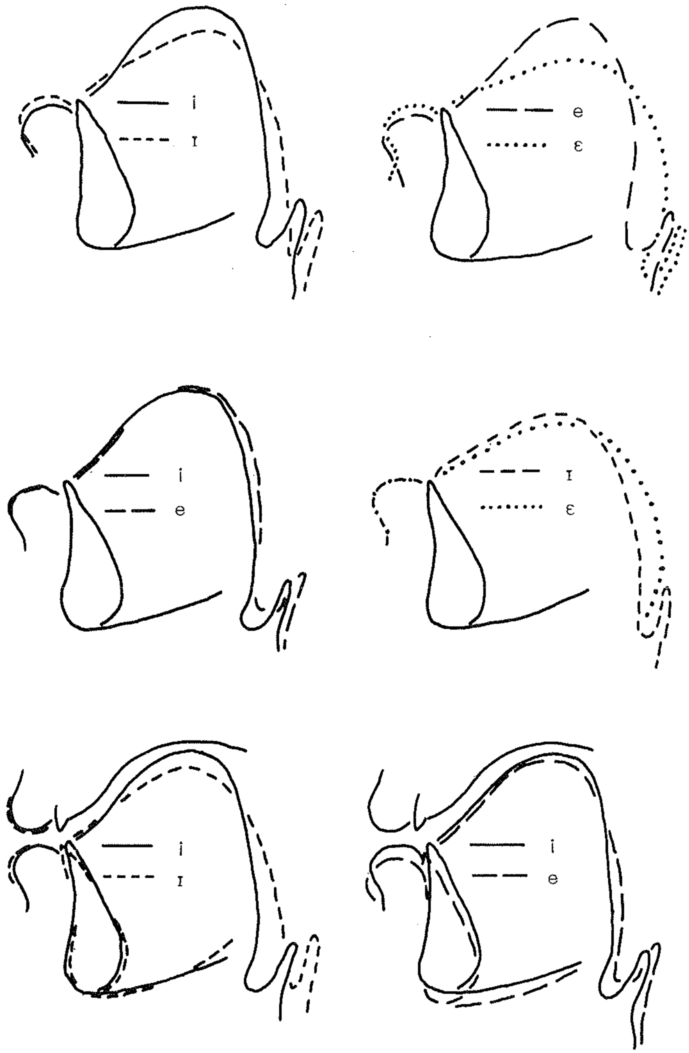


Fig. 5 Tongue posture and jaw openings in [i ɪ e ɛ] (Chiba and Kajiyama's German subject) showing tense tongue and lip postures for [i e] and lax for [ɪ ɛ], and higher jaw position for [i ɪ] and lower for [e ɛ].

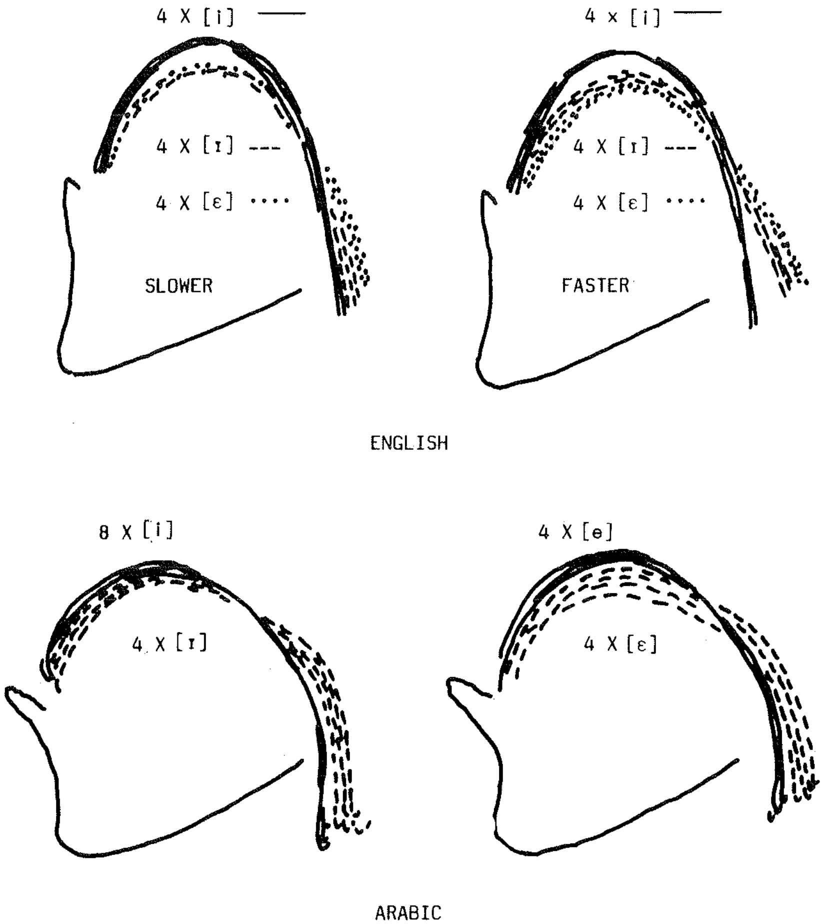


Fig. 6 Tense and lax tongue postures of the palatal vowels from the English and Arabic x-ray motion films.

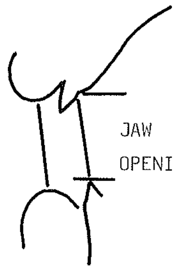


Fig. 7

Definition of lip separation and jaw opening

the scale [i, ɪ, e, ɛ], the number of steps depending on the language. Alternatively, as illustrated in Fig. 5, the mandible tends to be higher for [i, ɪ] and lower for [e,ɛ] .

The jaw openings were measured on lifesize reproductions of the x-rayed profiles, Fig. 7. The material is grouped as outlined above under procedures.

Table I records the means and ranges of jaw opening in each of the six groups. The following can be noted:

- Within each group the mean jaw openings are ranked in the order [i, ɪ, e, ɛ]
- The range of variation in each group and vowel is considerable
- The wide variation and overlapping ranges means that the degree of jaw opening cannot contribute to contrastive differences within [i, ɪ] and [e, ɛ] pairs, where from 50% to 100% of the jaw openings overlapped
- In contrast to the considerable overlapping within [i,ɪ] and [e, ɛ] pairs, the ranges of [i+ɪ] were very well separated from the [e+ɛ] ranges; not surprisingly, the data for the subjects of the two films (e) and (f) showed the best separation, 100% and 97% respectively of [i+ɪ] and [e+ɛ] being separated from each other; yet despite the larger experimental error in the collected data groups (a-d) the separation was considerable here too, 77%, 84%, 83% and 94% respectively; in all data groups except the English film (e), the boundary between the two classes is a jaw opening of about 8 or 9 mm
- The jaw openings by the subject of the English film (e) tended to be larger than the remainder of the material
- In all the remaining material, similar vowels tend to have similar ranges, irrespective of the number of vowel contrasts in each group.

Table I Mean jaw openings and ranges

Data group	[i]	[ɪ]	[e]	[ε]
(a) Collected German				
Mean (mm)	8	8.5	11.6	16.6
Range (mm)	6-11	7-10	6-19	8-23
60% of [i] openings overlapped 100% of [ɪ] openings				
80% of [e] openings overlapped 60% of [ε] openings				
A jaw opening of 9.5 mm separated 77% [i, ɪ] from 77% [e, ε]				
(b) Collected American English				
Mean (mm)	6	7.3	9	9
Range (mm)	4-8	6-9	7-12	8-12
50% of [i] openings overlapped 75% of [ɪ] openings				
60% of [e] openings overlapped 100% of [ε] openings				
A jaw opening of 8 mm separated 84% [i, ɪ] from 84% [e, ε]				
(c) Collected French				
Mean (mm)	6.3	-	9.4	10.3
Range (mm)	5-9	-	6-13	10-11
50% of [e] openings overlapped 100% of [ε] openings				
A jaw opening of 7.5 mm separated 83% [i] from 83% [e, ε]				
(d) Collected Russian				
Mean (mm)	6.9	-	-	13.4
Range (mm)	6-8	-	-	6-20
A jaw opening of 8 mm separated 94% [i] from 94% [ε]				
(e) English film				
Mean (mm)	9.8	10	-	14.3
Range (mm)	8-11	9-11	-	12-17
88% of [i] openings overlapped 100% of [ɪ] openings				
A jaw opening of 11.5 mm separated 100% [i, ɪ] from 100% [ε]				
(f) Arabic film				
Mean (mm)	7.3	8	10.5	10.3
Range (mm)	6-9	7-9	9-12	9-11
75% of [i] openings overlapped 100% of [ɪ] openings				
75% of [e] openings overlapped 100% of [ε] openings				
A jaw opening of 9 mm separated 97% [i, ɪ] from 97% [e, ε]				

Table II Jaw openings: significance of [i] versus [ɪ] and of [e] versus [ɛ] (the probability that the observed difference is at least this much larger than the standard error)

Data group	Number	Mean (mm)	Difference (mm)	Standard error (mm)	Significance
(a) Collected German					
[i]	5	8	0.5	1.15	66%
[ɪ]	4	8.5			
[e]	5	11.6	5	3.29	14%
[ɛ]	5	16.6			
(b) Collected American English					
[i]	4	6	1.3	1.1	24%
[ɪ]	4	7.3			
[e]	3	9	0	1.82	100%
[ɛ]	4	9			
(c) Collected French					
[e]	5	9.4	0.9	1.19	48%
[ɛ]	4	10.3			
(e) English film					
[i]	8	9.8	0.2	0.39	61%
[ɪ]	8	10			
(f) Arabic film					
[i]	8	7.3	0.7	0.55	20.4%
[ɪ]	4	8			
[e]	4	10.5	0.2	0.80	80%
[ɛ]	4	10.3			

Table III Jaw openings: significance of [i+ɪ] versus [e+ɛ]

Data group	Number	Mean (mm)	Difference (mm)	Standard error	Significance
(a) Collected German					
[i+ɪ]	9	8.2	5.9	1.87	0.16%
[e+ɛ]	10	14.1			
(b) Collected American English					
[i+ɪ]	8	6.6	2.4	0.97	1.36%
[e+ɛ]	7	9.0			
(c) Collected French					
[i]	6	6.3	3.5	0.89	0.008%
[e+ɛ]	9	9.8			
(d) Collected Russian					
[i]	11	6.9	6.5	1.82	0.02%
[ɛ]	7	13.4			
(e) English film					
[i+ɪ]	16	10.0	4.1	0.67	0.000 000 2%
[ɛ]	8	14.1			
(f) Arabic film					
[i+ɪ]	12	7.5	2.9	0.47	0.000 000 2%
[e+ɛ]	8	10.4			

Table II gives the result of a statistical test on the average jaw opening difference between [i] and [ɪ] and between [e] and [ɛ] respectively. This shows that the jaw opening differences within [i,ɪ] and [e,ɛ] pairs respectively are not significant (a difference between jaw opening means that is close to the standard error is highly probable).

Table III shows that the difference between the [i+ɪ] mean jaw opening and the [e+ɛ] mean jaw opening is highly significant (the differences being about 3 to 6 times larger than their standard errors in the various data groups, a result for which the probability is extremely small).

It is therefore concluded that there is an overwhelming tendency to use jaw openings larger than 8 or 9 mm for [e] and [ɛ] and smaller than 8 or 9 mm for [i] and [ɪ], dividing the palatal vowels into an "open" class and a "close" class.

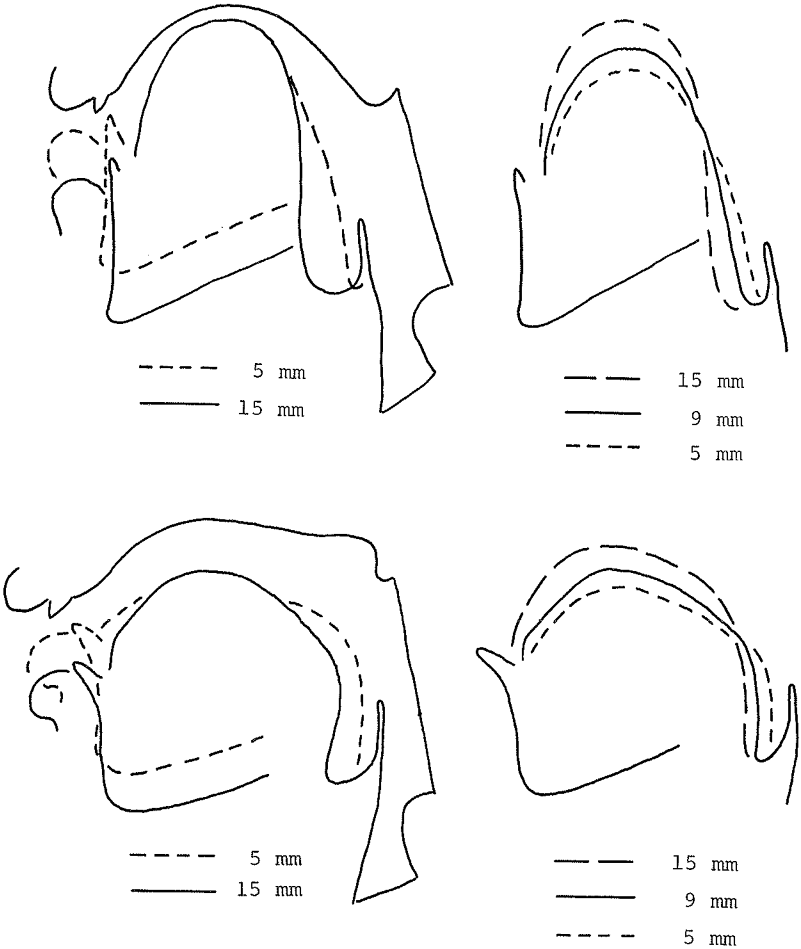


Fig. 8 Examples of modelled variation of jaw opening with perfect lingual compensation; English profile and [i] (above) and Arabic profile and [ε] (below). The extreme jaw positions are illustrated (left) and the corresponding lingual adjustments for compensation (right). As the mandible is lowered the tongue is pushed correspondingly higher by drawing in the tongue root. The result is good restoration of the area function along the palatal passage while the pharynx becomes progressively wider as the mandible is lowered.

Table IV Changes in formant frequency when the mandible is lowered from 5 to 15 mm in 2 mm steps with internal (lingual) compensation

		5-10 mm		10-15 mm	
Model		F1	F2	F1	F2
(e)	[i]	275-273	2411-2453	(273-267)	(2453-2465-2458)
	[ɪ]	415-433	2032-2146	(433-441)	(2146-2203)
	[ɛ]	(523-558)	(1678-1877)	558-561-555	1877-2009
(f)	[i]	263-261	2189-2240	(261-259)	(2240-2258)
	[ɪ]	414-427	1949-2080	(427-429)	(2080-2140)
	[e]	(437-447)	(2002-2107)	447-440	2107-2176
	[ɛ]	(454-497)	(1775-1959)	497-505	1959-2048

Values in () are at jaw openings that are normally out of range for the vowel in question (too open or too close, an opening of about 9 mm separating the two classes)

In short, the mandible is either "up" (for [i] and [ɪ]) or "down" (for [e] and [ɛ]). The variation is considerable, so that "up" means about 5-9 mm and "down" means about 9-15 mm.

There is clearly considerable latitude for variation for the jaw opening. Apart from pipe-smoking and phoneticians' bite blocks, factors influencing variation of mandible position are care of articulation and speaking rate. Lindblom (quoted by Sundberg 1969) has demonstrated how the absolute magnitude of the mouth opening is proportional to speaking effort. The larger jaw openings by the subject of the English film (e) previously noted in Table I may reflect more deliberate articulation and more effort, perhaps in response both to the experimental situation and to the background noise of the x-ray unit. There is also a speaking rate effect by the English subject (e). This material was repeated once a little slower than normal speaking rate at about 4.5 syllables/second and once a little faster at about 6.5 syllables/second. The jaw openings for [ɛ] averaged 15 mm at the slower rate against 12.5 mm at the faster rate.

How does this jaw variation affect the spectra of the vowels? To test this, the mandible was displaced on profiles selected from the English and Arabic films (see examples at Fig. 8). Without any lingual or labial compensation, increasing the jaw opening will widen the palatal passage and mouth opening and narrow the pharynx slightly. This is equivalent to simultaneously increasing A_{min} and A/ℓ values in Figs. 2 and 3. However, the tongue compensates for mandible shifts in palatal vowels to maintain an ideal degree of palatal narrowing. The modelled manouevres illustrated in Fig. 8 assume perfect lingual compensation along the palatal passage. Consequently, the pharynx becomes wider as the mandible is lowered (the tongue root is pulled further forward to raise the tongue body higher relative to the mandible). The jaw opening was varied from 5 to 15 mm in 2 mm steps for each vowel. The three narrower steps (5-9 mm) are normally in range for [i,ɪ] and out of range for [e,ɛ] . The three larger steps (11-15 mm) are normally out of range for [i,ɪ] and in range for [e, ɛ].

The results are given in Table IV. There is very little effect on F_1 - the acoustical consequences of simultaneously increasing the mouth opening and narrowing the pharynx largely cancel each other. F_2 rises as the jaw is opened (increasing the mouth opening and widening the pharynx both raise F_2). These variations are small compared with the total variation recorded in Figs. 2 and 3.

In real speech the lingual compensation is presumably not absolutely perfect. This will be discussed below. The experiment did not assume labial compensation. If this had been included there would have been even less acoustical variation, but its absence is clearly not critical for spread lip palatal vowels. The effects of labial articulation will be discussed in the section on the lips.

THE TONGUE

Bell's (1867) innovation of the tongue arching model, introducing the notion of "high", "mid" and "low" tongue positions was rapidly adopted during the latter part of the 19th century. Interest shifted from the jaw opening to tongue height. Now it was the tongue that was said to be progressively lower in the order [i, ɪ, e, ɛ]. Meyer (1910) and later Russel (1928) had objected that [e] in fact had a higher tongue height than [ɪ]. All possible examples in the

Table V Degree of palatal narrowing (cross section area of the palatal passage)

Data group	[i]	[ɪ]	[e]	[ɛ]
(a) Collected German				
Number	4	3	4	4
Range (cm^2)	0.5-0.8	1.2-2	0.7-1.2	1.6-2.7
Mean (cm^2)	0.65	1.47	0.93	2.1
(b) Collected American English				
Number	4	4	3	3
Range (cm^2)	0.5-1.6	2.2-2.8	1.7-2.6	1.9-3.4
Mean (cm^2)	1	2.45	2	2.7
(c) Collected French				
Number	6	-	5	2
Range (cm^2)	0.4-1.2	-	0.7-1.6	1.9-2.7
Mean (cm^2)	0.68	-	1.12	*
(d) Collected Russian				
Number	11	-	-	4
Range (cm^2)	0.3-0.7	-	-	1.5-2.7
Mean (cm^2)	0.53	-	-	2.43
(e) English film				
Number	8	8	-	8
Range (cm^2)	0.5-1.0	2.0-2.6	-	2.6-3.2
Mean (cm^2)	0.73	2.15	-	2.8
(f) Arabic film				
Number	8	4	4	4
Range (cm^2)	0.5-1.0	1.0-2.2	1.6-2.0	2.0-3.2
Mean (cm^2)	0.76	1.8	1.8	2.6

collection of x-ray profiles examined in Wood (1975a) had [e] higher than [ɪ] . This can be seen in Table V, which gives the ranges of the cross section area of the narrow palatal passage (a measure of the degree of palatal constriction, inversely proportional to palatal tongue height; the areas are given rather than the cross-distances for ease of comparison with Figs. 3 and 4). The following can be noted:

- The narrowed palatal passage is progressively wider in the order [i, e, ɪ, ɛ], demonstrating that [ɪ] has a lower tongue height than [e]
- The ranges for each individual vowel are so similar across languages that they suggest the degree of constriction is a very constant factor of vowel articulation. The constrictions are smaller than 1.0 cm^2 for an [i]-like quality, about 1.5 to 2.0 cm^2 for [e] and [ɪ] (slightly narrower for [e] and slightly wider for [ɪ]) and more than 2.5 cm^2 for [ɛ].

Table VI Pharyngeal width at the epiglottis

Data group	[i]	[ɪ]	[e]	[ε]
(a) Collected German				
Number	4	3	4	3
Full range (mm)	22-30	16-23	20-23	17-24
Mean (mm)	26	(20)	22	(20)
(b) Collected American English				
Number	4	4	3	4
Full range (mm)	23-29	19-23	16-22	16-21
Mean (mm)	26	20	(19)	18
(c) Collected French				
Number	6	-	5	4
Full range (mm)	24-32	-	18-24	14-23
Mean (mm)	28	-	21	17
(e) English film				
Number	8	8	-	8
Full range (mm)	22-27	16-22	-	15-20
Mean (mm)	25	19		17
(f) Arabic film				
Number	8	4	4	4
Full range (mm)	27-32	21-25	22-23	16-20
Mean (mm)	29	22	22	19

- The [ɪ]-like constriction ranges are very similar in all data groups, as are also the [ε]-like constriction ranges. This suggests that no modification is necessary to these vowels to "make room" for additional contrasts involving [e] and/or [ɪ].

- The [i] set is clearly distinguished from both lax sets [ɪ] and [ε]; the [ε] set is clearly distinguished from both tense sets [i] and [e]; there is the greatest risk of confusion between lax [ɪ] and tense [e].

That [ɪ] has a lower tongue position than [e] has constituted a crisis for phonetics and phonology since the turn of the century (Wood 1975a). The tongue has obviously not been doing what rules and language descriptions say it should have been doing.

Pharyngeal widening is correlated with palatal narrowing. Table VI gives the ranges of pharyngeal width at the widest part (level with the epiglottis). Pharyngeal width is very similar for similar vowels across languages, matching the previously observed similarity at the palatal narrowing. Typical pharyngeal widths are 25-30 mm for [i], 20-23 mm for [ɪ] and [e], 15-20 mm for [ε]. The correlation between pharyngeal width and palatal narrowing is usually attributed to posterior genioglossal activity, which pulls the tongue root forwards (widening the pharynx) and pushes the tongue bodily upwards (narrowing the palatal passage).

The average degree of palatal narrowing for a vowel might be considered as a target constriction. The targets would then lie along a four-grade scale of palatal narrowing for [i, e, ɪ, ɛ] . This is the traditional view apart from the inversion of [e, ɪ].

Unfortunately, a graded scale of palatal targets does not bring out any specific difference between the traditional categories "tense" and "lax". These were originally related to subjective assessments of muscular tension. The same contrast was denoted "narrow" and "wide" by Bell and Sweet. Other proposed differences, summarized by Jakobson and Halle (1964), have been the degree of "centralization" of the tongue and the degree of deformation of the vocal tract. More recently there has been the advanced tongue root debate (Stewart 1969, Perkell 1971, Lindau et al. 1972, Raphael and Bell-Berti 1975, Wood 1975b, Lindau 1978).

A further difficulty associated with a scale of palatal targets is its possible implications for theories of motor control. Does the brain actually select one out of a stock of say four possible palatal constrictions (tongue heights) as its goal for the tongue, or is the graded scale simply the phonetician's convenient arrangement that is nevertheless not immediately relevant for motor control?

Table V shows that the palatal constriction is narrower for the tense vowel [i] or [e] than for the corresponding lax vowel [ɪ] or [ɛ] respectively. For palatal vowels at least, this is the sense in which Bell and Sweet meant narrow and wide. Similarly the pharynx is wider for each tense vowel [i] or [e] than for the corresponding lax vowel [ɪ] or [ɛ] respectively. This is the basis of the advanced tongue root proposal. Both these factors together confirm that the vocal tract is more deformed for the tense than for the lax vowel in each pair. But this does not clarify why all four vowels are ranked in a particular order for a given property or why

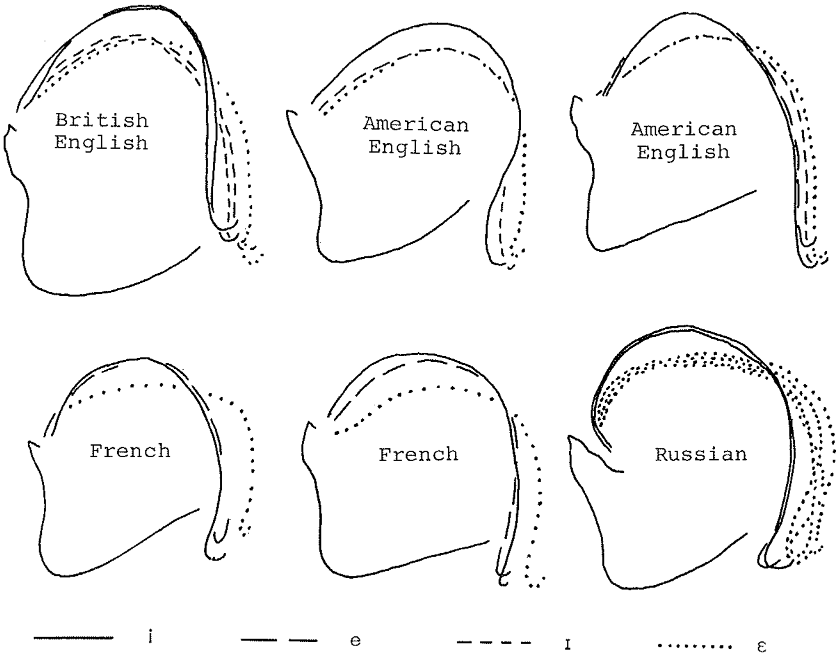


Fig. 9 Six profiles confirming a tense tongue posture for [i, e] and a lax posture for [ɪ, ɛ].

there is ambiguity between [ɪ] and [e] regarding the degree of palatal narrowing and pharyngeal widening.

Reference has already been made to Figs. 5 and 6 which show a tendency for the tongue to assume one posture relative to the mandible for the tense vowels [i] and [e] and another somewhat lower for the lax vowels [ɪ] and [ɛ] (German, English and Arabic subjects). Groups (a-c) in the collection of published x-ray material contains twelve possible sets that are relevant for this point. In addition there were two further German sets that could not be used because the mandible had not been traced. The extent to which this material confirms the two tongue positions relative to the mandible for tense and lax palatal vowels respectively is as follows:

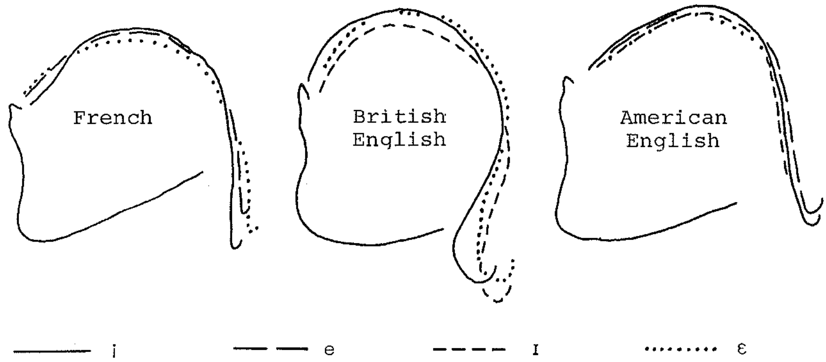


Fig. 10 Three profiles with similar tongue postures for both tense and lax palatal vowels.

German	2 sets	1 confirms 1 has random differences
Am. Eng.	4 sets	2 confirm 1 partly confirms 1 has same position for all examples
S.Br.Eng.	2 sets	1 confirms 1 partly confirms
French	4 sets	3 confirm 1 partly confirms

Taken together, there were 7 sets that confirm, 3 that partly confirm and 2 that exhibit no systematic difference between tense and lax positions. The 7 confirming examples are reproduced in Figs. 5 and 9. The English and Arabic subjects in Fig. 6 also confirm. The Russian profiles (one example in Fig.9) also had different tense-lax-postures for [i] and [ɛ]. Figure 10 gives examples of partial confirmation and of contradiction (the same tongue posture for all 4 vowels).

Clearly 9 out of 14 sets confirm a tense posture for [i, e] and a lax posture for [ɪ, ɛ] while 3 provide partial confirmation. There can be several possible reasons why there are a few exceptions:

- Each speaker has his own individual method for coordinating the tongue and mandible to achieve the desired vocal cavity configuration; Ladefoged et al. (1972a, 1972b) reached this

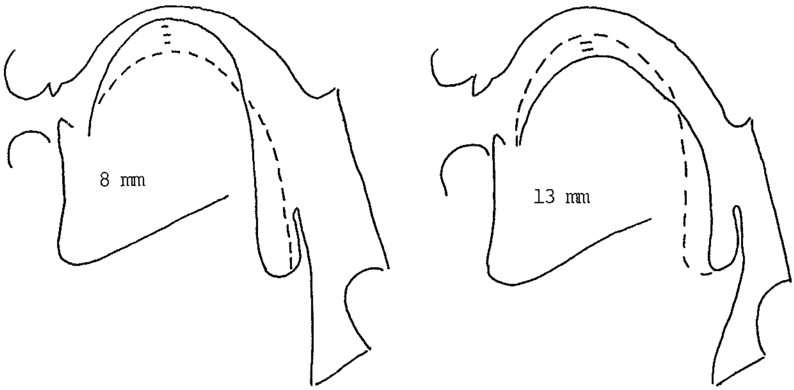


Fig. 11 Profile manipulations for modelling tongue lowering in 2 mm steps from the tense posture to the lax posture at two different jaw openings, 8 mm and 13 mm, representing [i-ɪ] and [e-ɛ] respectively. The solid profile is an original x-rayed configuration.

conclusion after performing factor analyses on data from x-ray motion films for 6 subjects; however, they record that 3 of the 6 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 in Figs. 5 and 6; note that 5 of their 6 subjects had [e] higher than [ɪ], 3 of the 6 confirmed the tense-lax difference outlined above and the remaining 3 disagreed both with that pattern and with each other.

- It is well documented that the tongue can and does compensate for variation of mandible position in palatal vowels; the actual tongue height relative to the mandible on a given occasion will reflect this compensation.

- the English subject (e) also exhibits a speaking rate effect; Fig. 6 shows how the English [ɛ] profiles coincided with the [ɪ] profiles at the slower rate (average [ɛ] jaw opening 15 mm), while they were slightly lower than [ɪ] at the faster rate (average [ɛ] jaw opening 12.5 mm); Fig. 6 shows how the English subject adjusted the tongue to this mandibular difference in [ɛ] .

- Experimental error will account for some variation in the traced profiles.

The conclusion is that the tongue is higher and more bunched relative to the mandible for the tense vowels [i, e] and lower, flatter and bulging further into the pharynx for the lax vowels [ɪ,ɛ] . This difference is performed with the

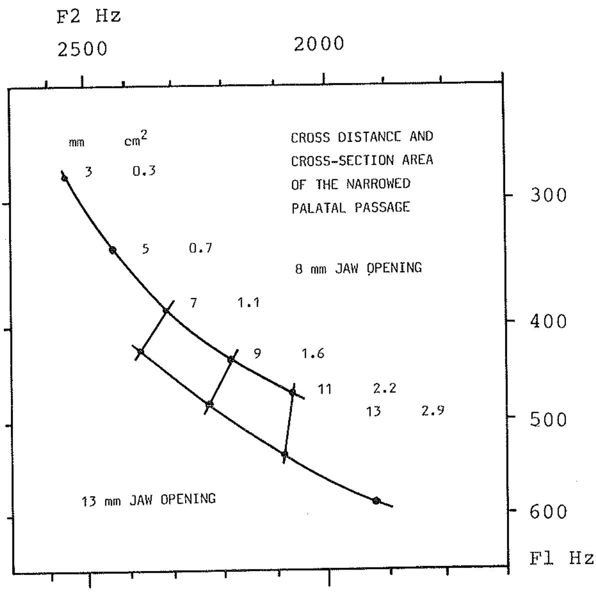


Fig. 12

The resulting F_1 and F_2 frequencies for the configurations modelled in Fig. 11. Cf. these results with the subject's own speech, Fig. 3.

mandible raised (for close [i, ɪ]) and lowered (for open [e,ɛ]). Given average speaking rate and effort the tongue will tend to assume a typical tense posture for [i, e] and a typical lax posture for [ɪ, ɛ]. Most of the sets examined conformed to this pattern. Whenever circumstances cause the jaw to vary more, the tongue will compensate and there will be corresponding variation of tongue position relative to the mandible.

The apparent ambiguity of the degree of palatal narrowing and pharyngeal widening for [ɪ,e] can now be explained. There are two components of pharyngeal width for palatal vowels - lingual (posterior genioglossal contraction draws the tongue root forward, widening the pharynx) and mandibular (as the mandible is depressed the tongue tends to bunch at the root and bulge into the pharynx). Both components combine to make the pharynx wide for close tense [i] and narrow for open lax [ɛ]. But the components oppose each other for [ɪ] and [e] - higher mandible with lax tongue posture for [ɪ] and lower mandible with tense tongue posture for [e]. The balance of these conflicting manoeuvres is that the palatal passage tends to be slightly wider for [ɪ] than for [e] (tongue height lower) and that the pharynx tends to have roughly the same width for both vowels.

How does variation of the degree of palatal narrowing affect the formant frequencies of palatal vowels? Some idea can be gained from the three parameter model nomograms (e.g. Figs. 3 and 4) by widening the palatal passage (increasing A_{min}) for a constant degree of mouth opening (A/ℓ). Figure 11 illustrates a model experiment in which palatal vowel profiles selected from the English film were modified by varying the tongue position. The tongue was lowered on the model profiles in 2 mm steps from the bunched, fully tense posture to the flatter, fully lax posture (cf. Fig. 6). With a close jaw opening of 8 mm this covered the range from [i] to [ɪ] (width across the palatal passage 3-11 mm, corresponding to cross-section areas 0.3-2.2 cm^2). With an open jaw opening of 13 mm, the range covered was from [e] to [ɛ] (width across the palatal passage 7-13 mm, corresponding to 1.1 to 2.9 cm^2). Note that the vowel [e] is included in the experiment by combining the tense tongue posture with the open mandible position. This vowel does not occur as a phoneme in the speech of subject (e), cf. Figs. 1 and 3.

Figure 12 gives the results. These are similar to those given by the Stevens and House nomograms (cf. Fig. 3). Clearly F_1 and F_2 are very sensitive to variations of palatal narrowing and pharyngeal width (the consequences of varying tongue height) in palatal vowels. A 2 mm tongue height step causes a larger spectral shift than the entire 5-15 mm mandible range of the preceding experiment (cf. Table IV). Figure 12 shows how necessary lingual compensation for mandibular variation is for palatal vowels - a 1 mm difference in tongue height means 25 Hz for F_1 and 60 Hz for F_2. This is presumably a contributory factor to the spectral variation recorded in Figs. 3 and 4 (if, for example, the precision of tongue height is $\pm$1 mm there will be a variation of about 50 Hz in F_1 and 120 Hz in F_2).

A further point to note is that the term "place of maximum constriction" is ambiguous in natural speech. Figures 1, 2 and 11 show how prominent features of vocal tract morphology can more or less permanently narrow some parts of the vocal tract. For example, the velum of the English subject (e) is not so firmly elevated as that of the Arabic subject (f). Consequently, the English subject's vocal tract tapers off behind the narrowed palatal passage towards the velum while the Arabic subject's widens towards the velum. The important point is that both subjects have a tongue manoeuvre directed at the palatal passage (cf. Fig. 6). The English subject's velar narrowing is a personal characteristic. The postpalatal tapering is particularly evident in the lax tongue postures where the back of the tongue bulges into the pharynx. Again, the postpalatal tapering is exaggerated by the midpalatal tongue manoeuvre; it would have been less evident with a pre-palatal tongue manoeuvre.

Another example is provided by the tongue blade and the tooth ridge. The bunched tongue body of the tense posture is not entirely due to the genioglossi drawing in the back of the tongue. It is also due to depression of the tongue blade to compress the tongue anteriorly (Wood 1975b). In the lax posture the tongue blade is less depressed and if the tooth ridge happens to be fairly prominent the vocal tract may be narrowed anteriorly to the narrowed palatal passage.

THE LIPS

The previous sections have shown how the jaw opening tends to be less than 8 or 9 mm for "close" [i, ɪ] and wider for "open" [e, ɛ] (Tables II and III). The variation is considerable (Table I), but the resulting variation of the mouth opening has little effect on F_1 and F_2 providing the tongue adjusts completely to the mandibular variation to restore the degree of palatal narrowing (Fig. 8, Table IV). There is therefore no urgent need to adjust the lips in spread-lip palatal vowels in order to compensate and restore an ideal degree of opening. In contrast, the spectra of rounded vowels are very sensitive to variations of the mouth opening (Wood 1982 and data and model experiments on rounded back vowels in the same x-ray material studied here, report forthcoming). However, Hughes and Abbs (1976) have presented data that suggests the lips do compensate

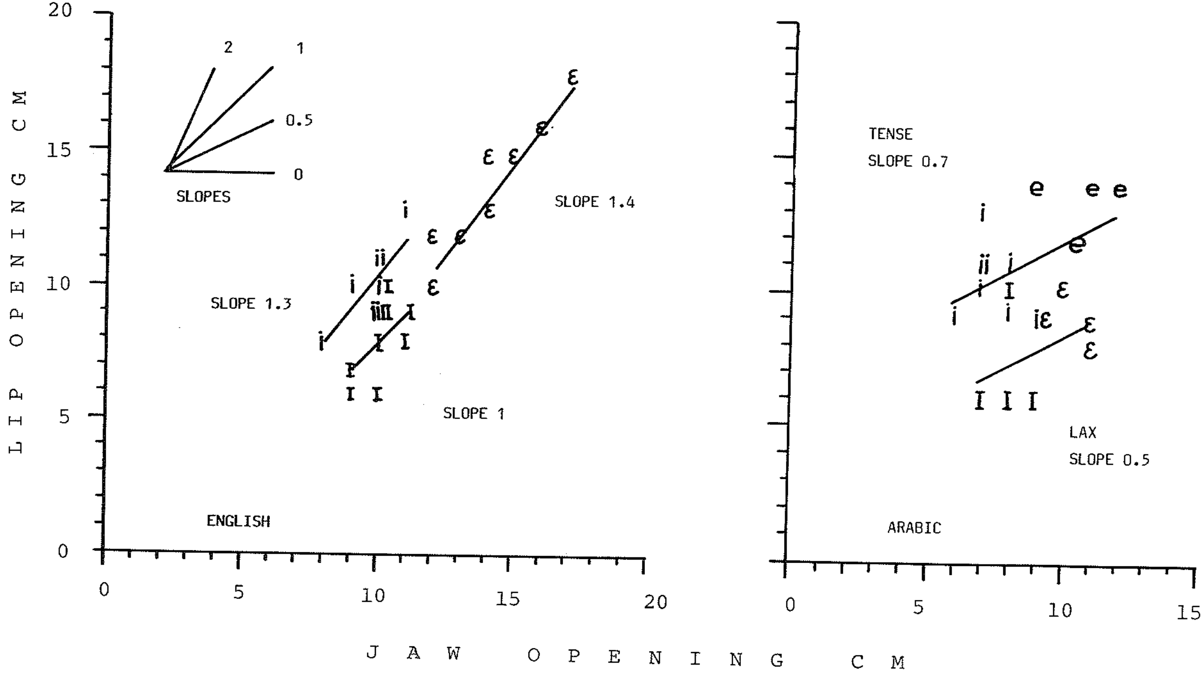


Fig. 13 Correlation between lip opening and jaw opening in (a) English palatal vowels, left, and (b) Arabic palatal vowels, right. A regression slope of 0 means full labial compensation (the lip opening is the same at all jaw openings), a slope of 1 means no labial compensation (the lip opening increases in step with the jaw opening). The Arabic vowels are pooled into a tense set and a lax set as there are so few of each individual vowel.

for variation of jaw position in spread-lip vowels.

Figure 13 shows the extent to which the lips and mandible were coordinated in the two x-ray films. The vertical lip and jaw separations were measured on life-size reproductions of the x-ray profiles (Fig. 7). The slope of the regression between the lip and jaw separations is a measure of the amount of labial compensation:

- A slope around 0, i.e. a horizontal regression (a constant lip opening at different jaw openings) means complete labial compensation
- A slope around 1 (the lip opening varies by the same amount as the jaw opening) means that there was no compensation at all.

The slopes of the English subject's vowels (Fig. 13a) were from 1 to 1.4, which means the lip separation increased in step with the jaw opening. There was no labial compensation for jaw variation.

The slopes of the Arabic subject's vowels (Fig. 13b) were 0.5-0.7, which hardly represents compensation. The vowels are pooled into tense and lax sets as there are so few examples of each (cf. the English subject's lax [ɪ] and [ɛ] which follow a common regression). The Arabic data is also confounded by variation in the amount of remnant coarticulation from the labial consonant [b] that preceded the vowels. In some syllables the lips were fully spread during the vowel itself, while in others they were not fully spread until the transition to the following [d]. This explains the wide scatter of lip openings for each vowel. In the English data each vowel conforms fairly closely to the regression (correlations r=0.41-0.87). In the Arabic data the individual vowels are more randomly distributed and the regressions apply to tense and lax sets.

Figure 13 also illustrates another aspect of lip articulation - the lips were more spread for tense vowels than for lax vowels (by about 2 mm in the English data and about 4 mm in the Arabic data). The same difference can be seen in Fig. 5. Figure 14 gives examples of tense and lax lip positions from the English and Arabic films. A corresponding difference has been reported for Swedish rounded vowels by McAllister (1980). He found that the lips were more protruded for tense than for

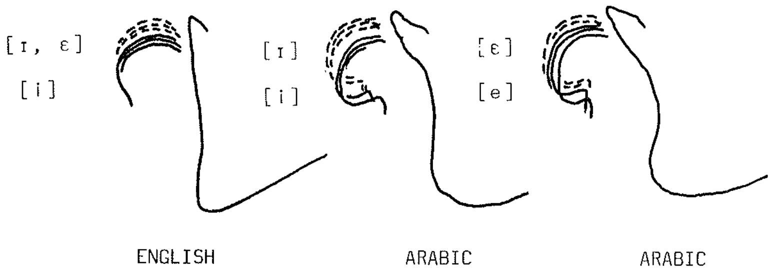


Fig. 14 Tense and lax lip postures of spread-lip palatal vowels in English and Arabic. Cf. also the German example in Fig. 5.

lax rounded vowels. The mean vertical lip separations, relative to the jaw opening, were:

	English subject	Arabic subject	
Tense:	+0.4 mm	+3.0 mm	lip spreading related to jaw opening
Lax:	-1.3 mm	-1.1 mm	

This means that the vertical spreading increments were as follows (assuming, with Lindblom and Sundberg 1971, that the jaw-dependent neutral vertical lip separation h_0 is 4 mm narrower than the jaw opening):

	English subject	Arabic subject	
Tense:	+4.4 mm	+7.0 mm	lip spreading related to neutral lips
Lax:	+2.7 mm	+2.9 mm	

This corresponds to Lindblom and Sundberg's (1971) vertical muscular component H_m (the increment in vertical lip separation due to muscular activity, positive for spreading and negative for rounding).

How do these differences of lip position affect the vowel spectra? To test this, one example of each vowel was selected from both films and the mandible positions adjusted to a constant 8 mm for close [i, ɪ] and 13 mm for open [e, ɛ]. The lip positions were then altered systematically from neutral, through lax to tense. This was done by increasing the value of H_m from 0 mm (for neutral lips) in 2 mm steps to 8 mm (fully tense). The Lindblom and Sundberg lip model also includes two components of horizontal opening, jaw-dependent w and a muscular increment W_m (positive for spreading, negative for rounding). Lindblom and Sundberg suggested +10 mm increment for spreading. The horizontal increment is presumably also subject to tense-lax differences corresponding to the vertical differences observed above. The horizontal distances cannot be seen on the mid-sagittal projection of the films. For the experiments, the horizontal muscular component W_m was assumed to vary in proportion to the vertical component H_m as follows: 0 mm (neutral) to 12 mm (fully tense) in 3 mm steps, i.e. assuming $W_m = 1.5\ H_m$. These values provided a perfect match for the area data on Swedish vowels given by Lindblom and Sundberg (1971). The following combinations were used for the experiment:

		Neutral	Lax 1	Lax 2	Tense 1	Tense 2
Vertical increment	H_m mm	0	2	4	6	8
Horizontal increment	W_m mm	0	3	6	9	12

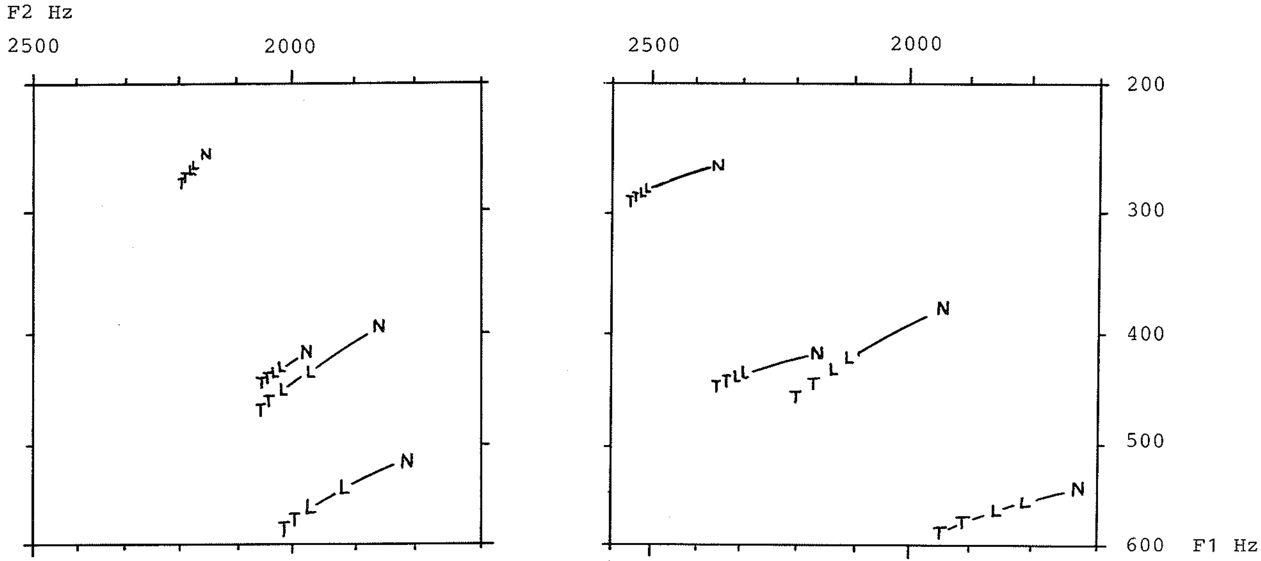
with otherwise constant conditions:

Jaw opening	J	= 8 mm for	[i, ɪ]
		= 13 mm for	[e, ɛ]
Jaw-dependent lip opening (vertical)	h_o	= J-4 mm	
Jaw-dependent lip opening (horizontal)	w_o	= f(J)	

(see Lindblom and Sundberg 1971)

A tense [e] vowel was constructed for the English model by combining the tense tongue posture with the open jaw position, cf. Fig. 11.

The results are given in Fig. 15. Each vowel, in both models, is most sensitive to variation of lip spreading in the interval from neutral lips to the first 2 mm of spreading. The lax vowels [ɪ, ɛ] are more sensitive than the tense vowels [i, e]. Variation of lip spreading, e.g. remnant coarticulation from a labial consonant, may account for up to 30-100 Hz of F_2



N neutral lips L lax lips T tense lips

Fig. 15 Results of a model experiment in which the lips were varied from neutral through lax to tense for each of four basic configurations: tense and lax tongue posture with a jaw opening of 8 mm (corresponding to [i, ɪ]) and with a jaw opening of 13 mm (corresponding to [e, ɛ]). The model to the left is based on the Arabic subject and that to the right on the English subject (cf. Figs. 3 and 4).

variation, depending on vowel and severity (cf. actual formant frequency variation in Figs. 3 and 4).

PHONOLOGICAL CONSIDERATIONS AND SPEECH PRODUCTION

Analysis of the x-ray profiles has shown that the degree of jaw opening divides palatal vowels into an open set [e, ɛ] with jaw openings larger than 8 or 9 mm and a close set [i, ɪ] with jaw openings smaller than 8 or 9 mm.

It was also found that the tongue posture relative to the mandible divided palatal vowels into a tense set [i, e] with the tongue more bunched and elevated relative to the mandible and a lax set [ɪ, ɛ] with the tongue flatter and lower.

This confirms the hypothesis derived from Chiba and Kajiyama's German profiles in Fig. 5 and outlined in the introduction. The following feature system is proposed, in place of the continuous scales of tongue height:

		JAW OPENING (MANDIBLE POSITION)	
		CLOSE	OPEN
TONGUE POSTURE	TENSE	i	e
	LAX	ɪ	ɛ

The experiments illustrated in Figs. 12 and 15 confirm that these combinations of jaw and tongue position generate spectra that match the spectra of natural speech (cf. Figs. 3 and 4).

Tenseness also defines the lip position relative to the mandible (Figs. 5 and 14). The lips are more spread for tense [i,e] and less spread for lax [ɪ, ɛ]. The phonologist will not usually need to specify this separately, but anyone studying raw speech data must keep it in mind.

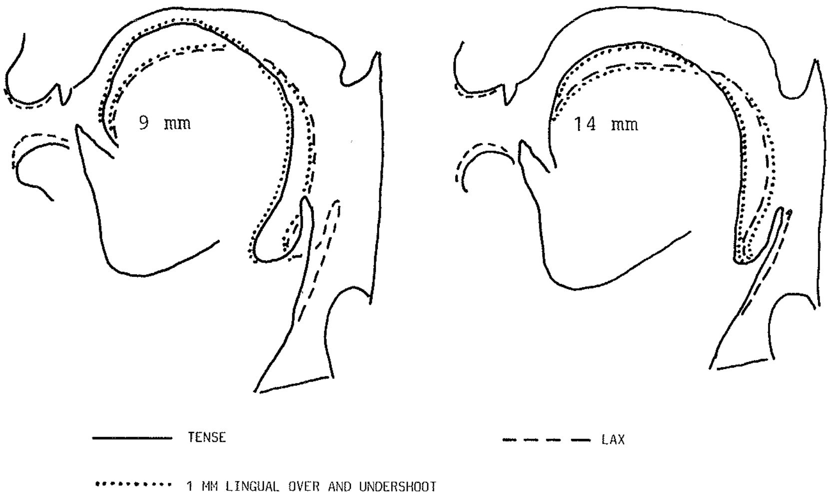


Fig. 16 Examples of manipulations of Russian profiles in order to simulate other vowels. The lax tongue posture of [ɛ] was added to the close jaw position of [i] and the tense tongue posture of [i] was added to the open jaw position of [ɛ], producing [ɪ] and [e]. The close jaw opening was varied from 5 to 9 mm and the open jaw opening from 9 to 14 mm. Lingual compensation was allowed a tolerance of ±1 mm and the tense and lax lip postures ±2 mm. The examples illustrated show the tense and lax tongue postures at each of two of the jaw openings and 1 mm overshoot (left) and undershoot (right). See Fig. 17 for the results.

Note that tenseness and laxness are defined here in terms of differences of muscular tension that lead to different lip and tongue positions. Consequently, the shape, and hence resonance conditions, of the vocal tract are different. The members of a tense-lax pair sound different. This articulatory definition of tenseness and laxness implies differences of sound timbre. It does not imply differences of vowel quantity, although tense vowels often tend to be long vowels. Vowel quantity is a temporal property. See Wood (1975b) for further discussion.

Combining tense-lax tongue and lip positions with open-close mandible positions is not just a convenient arrangement of

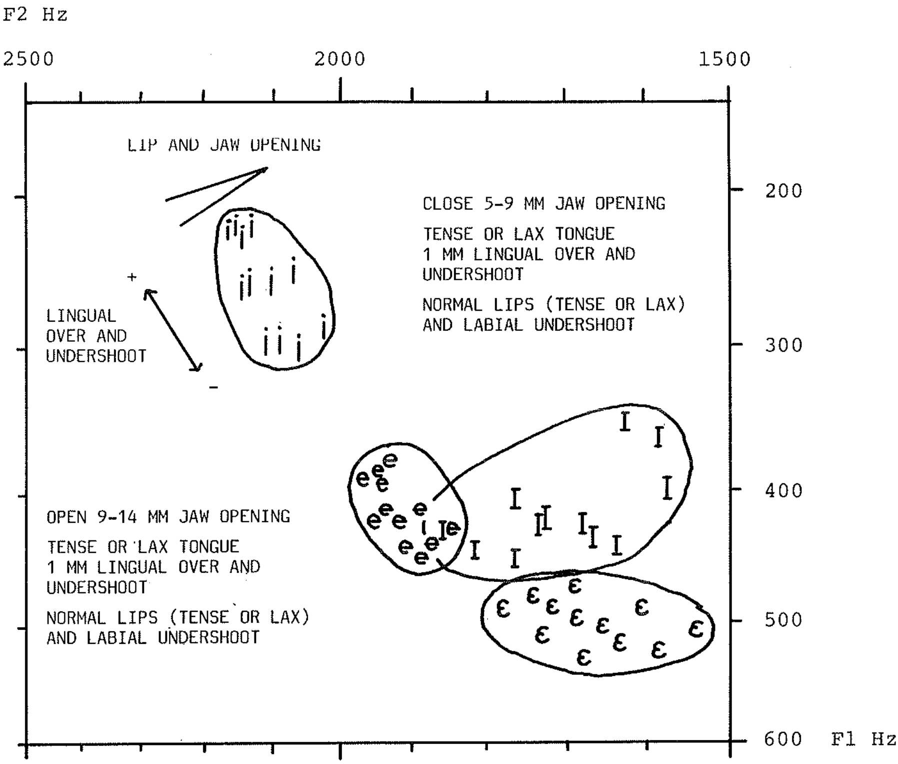


Fig. 17 Variation of F_1 and F_2 generated by varying the jaw position and allowing ±1 mm of lingual precision and ±2 mm of labial position for different combinations of tense and lax tongue postures and open and close jaw positions, as illustrated in Fig. 16.

features to classify sets of vowels. The combination describes the production of these vowels in a way that the ad hoc divisions of tongue height scales never did. For example, transforming a three-height scale into a four-height scale is a complicated business. Note that if a particular combination of tenseness and openness is not utilised in a language the articulation remains vacant (for example, close lax [ɪ] is absent from French, open tense [e] is absent from the dialect of the subject of the English film).

This corresponds to the data recorded in Tables I, V and VI, where the [ɪ] and [ɛ] configurations were similar across languages whether or not there were other contrasts too. It was concluded from that data that it would not be necessary to modify existing articulations to make room for new contrasts. This is demonstrated by an experiment in which a model based on a language with only a contrast between close, tense [i] and open, lax [ɛ] was "taught" close, lax [ɪ] and open, tense [e] (Figs. 16 and 17). This is the equivalent of modelling a vowel system evolving from /i,ɛ/ to /i,ɪ, e,ɛ/.

An [i] and an [ɛ] profile were taken from a published Russian set. The tense tongue posture of [i] was combined with the open jaw position of [ɛ] to produce [e] and the lax tongue posture of [ɛ] was combined with the close jaw position of [i] to produce [ɪ] (cf. Fig. 16 with Fig. 6). These four basic configurations were then further modified in various ways to model the extremes of variation observed in Tables I, II, III, V and VI and in Figs. 6 and 14. The close vowels were modelled with 5 mm and 9 mm jaw openings and the open vowels with 14 mm and 9 mm jaw openings. The tongue was adjusted in each case to provide compensation. In addition to perfect lingual compensation, the tongue was also presumed to overshoot and undershoot compensation by 1 mm. Finally, for each of these configurations the lip spreading was varied to simulate over and undershooting of the tense and lax lip postures relative to the mandible (cf. Fig. 14). There was no labial compensation for the mandibular variation.

The results are given in Fig. 17. The areas for each vowel represent the spectral limits generated by the tolerances included in the experiment - 5 or 6 mm jaw opening variation, tongue height ±1 mm, vertical lip spreading ±2 mm. The following should be noted:

- The modelled combinations of openness and tenseness generate four distinct vowel areas that are comparable with those of Figs. 3 and 4

- The vowels [ɪ] and [e] were inserted without first having to rearrange the articulation of [i] and [ɛ] to make room for them (the original x-rayed [i] and [ɛ] profiles were preserved, subject to the experimental tolerances introduced above)

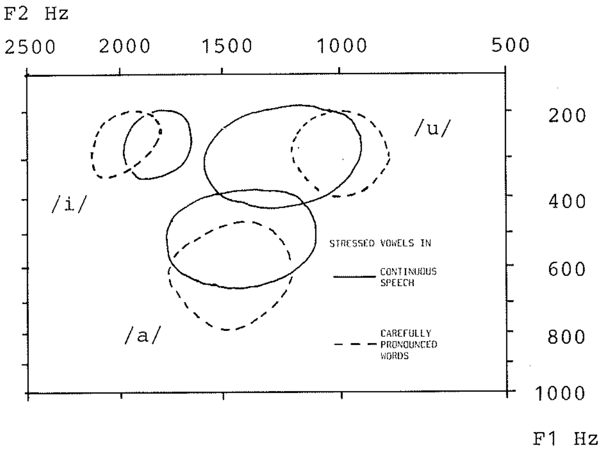


Fig. 18

F_1 and F_2 of stressed vowels in carefully pronounced words and continuous speech, in non-pharyngeal environments.

- The model vocal tract configurations are realistic in that they reproduce such typical aspects of observed articulations as (i) complete overlapping of [i] and [ɪ] jaw ranges (5-9 mm) and [e] and [ɛ] jaw ranges (9-14 mm), (ii) higher tongue height for [e] than for [ɪ]
- The amount of spectral variation within each vowel area is comparable with observed variation (cf. Figs. 3 and 4); the modelled articulatory variation can account for much of the variability of formant frequencies observed in everyday speech
- The vowel areas are distinct and do not overlap; the acoustical variability resulting from the amount of articulatory variability modelled in this experiment is clearly tolerable in speech (these areas are typical of the variability that can be observed in speech)

The view presented here is somewhat mechanical - the jaw is either up or down, the tongue is either bunched or flattened (and the lips more spread or less spread). Nor is there any need to reorganize the articulation of existing vowels to make room for a new vowel. These conclusions clash with the traditional view that, for example, tongue height has to be reorganized to fit in an extra vowel. Spectrographic analyses of say three, five or seven vowel systems usually show the vowel areas spread out and filling the available space. An example from West Greenlandic Eskimo is given in Fig. 18. Can this be reconciled with the view expressed in this report?

Note first that the phoneme areas illustrated in Fig. 18 contain several allophones that have regular distributions and consequently regularly unique articulations. But this does not answer the objection. The contrasts of openness and tenseness described above were not only defined in terms of typical ideal positions. More important, they were also defined as tolerances. For the open jaw position, for example, the mandible is lowered more than 8 or 9 mm. The tense or lax tongue and lip postures can vary within about ±1 or 2 mm. It is possible that these tolerances are less stringent in a language with fewer vowel contrasts. For the palatal vowels, an extra ±1 mm of permissible lingual over and undershooting of the degree of palatal narrowing would add an extra 50 Hz to the F_1 variation of a vowel.

CONCLUSIONS

1. A jaw opening of 8 or 9 mm divides the palatal vowels into a close class [i, ɪ] and an open class [e, ɛ]. Within each class the jaw opening ranges of different vowels overlap almost completely (usually about 5-9 mm and 9-15 mm or more respectively). These ranges are similar across languages irrespective of the number of palatal vowels in each language.

2. With lingual compensation for mandibular variation, as in natural speech, the vowel spectra are hardly sensitive to a wide range of jaw opening variation. The consequent spectral variation is small compared to the total spectral variation of a vowel that can be observed in speech. There is no need for labial compensation for mandibular variation in spread-lip palatal vowels.

3. There was no evidence of labial compensation for mandibular variation in spread-lip palatal vowels in the two x-ray films.

4. There are differences of lip spreading relative to the mandible between tense and lax vowels. The lips are spread more for tense vowels, less for lax vowels.

5. The degree of palatal narrowing is very similar for similar vowels across languages, indicating that this is a constant factor of vowel articulation irrespective of the number of vowels in a system. In particular [i] and [ɛ] are very similar across languages and no modification is necessary to make room for additional vowels. See also (8) below.

6. The lower tongue height for [ɪ] than for [e] is confirmed.

7. The tongue assumes a typical bunched and elevated posture relative to the mandible for tense vowels [ɪ, e]. This narrows the palatal passage and widens the pharynx. The tongue assumes a typical flatter and lower posture for lax vowels [ɪ, ɛ] . This widens the palatal passage and narrows the pharynx. The two tongue postures are found across languages irrespective of the number of vowel contrasts in a system.

8. The palatal vowels are differentiated by combining the two degrees of jaw opening with the tense or lax tongue posture. In a model experiment, Russian profiles for close tense [i] and open lax [ɛ] were manipulated to simulate close lax [ɪ] and open tense [e]. The two new combinations generated spectra that did not clash with the original [i, ɛ], spectra confirming that [i] and [ɛ] did not need modifying to make room for [ɪ, e]. The lip, the jaw and tongue positions were assigned tolerances that correspond to articulatory variation observed in speech. The resulting spectral variation was comparable to actual variation observable in speech.

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 mid-saggital 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. Mid-saggital 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 biologioal 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.

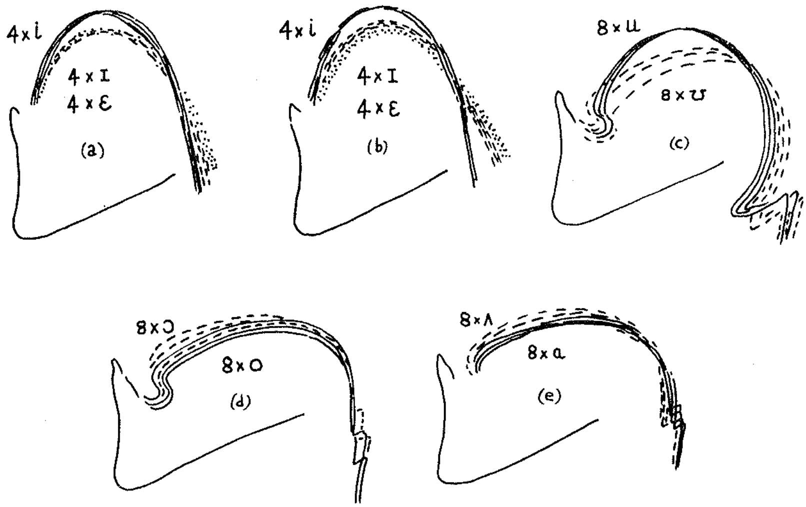


Fig. 1. Sets of tongue profiles for tense-lax pairs by a Southern British English subject. There are 8 examples of each vowel, 4 uttered a little slower than average everyday speech (4.5 syllables/sec) and 4 a little faster (6.5 syllables/sec). The main articulatory consequence of the rate difference 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-

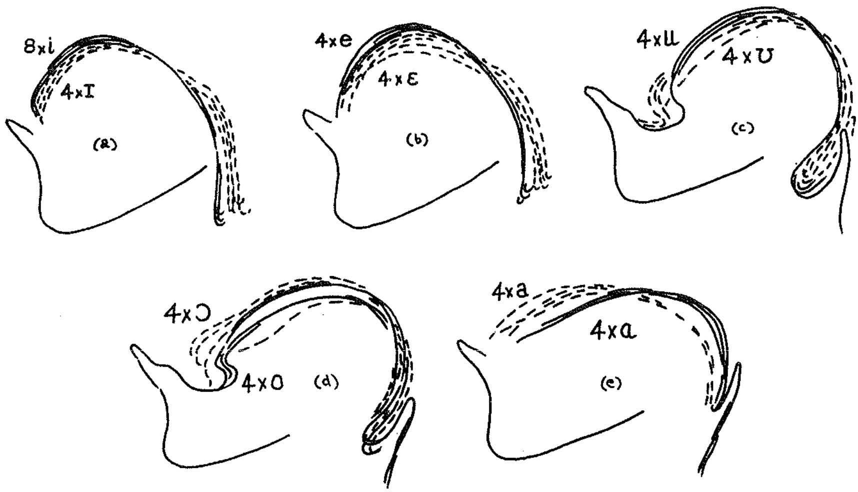


Fig. 2. Sets of tongue profiles for tense-lax pairs by an Egyptian subject. There are four examples of each vowel, except for [i]. The [a] quality represents /a/ in a "non-emphatic" environment while [ɑ] represents 2 x /a/ and 2 x /a+a/ in an "emphatic" environment.

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

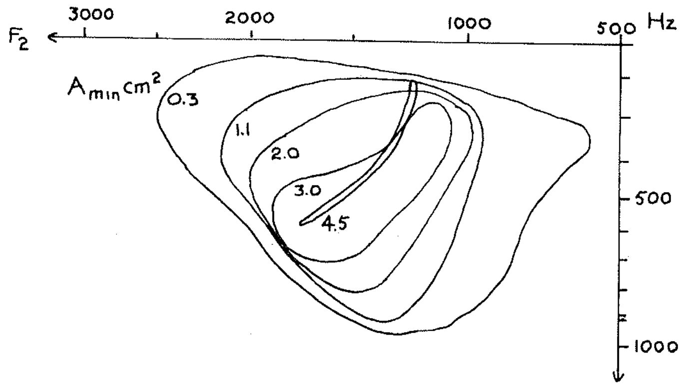


Fig. 3. The maximum possible spectral ranges for F1 and F2 at different degrees of constriction (A_{min} cm^2). This is based on the Stevens and House (1955) three-parameter model nomograms. Each ring encloses the spectra generated by all combinations of constriction location and mouth-opening size for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

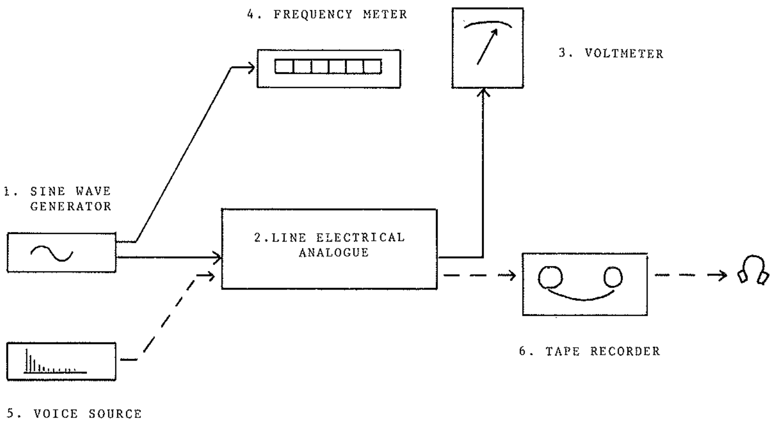


Fig. 4. For sweeping and measuring resonances, a sinus wave from the generator (1) is passed through the analogue LEA (2) to a voltmeter (3). Voltage maxima occur at resonance frequencies which can be read off from the frequency meter (4). For monitoring and recording synthetic vowel qualities, a voice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

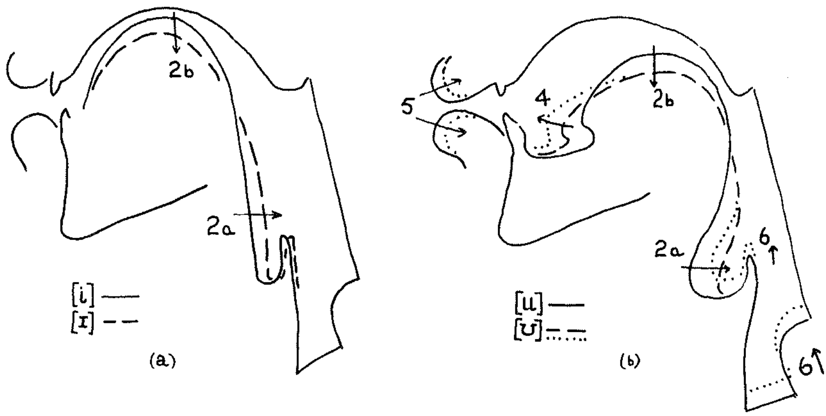


Fig. 5. Modifications made to the model profile for (a) palatal vowels and (b) palatovelar vowels.

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).

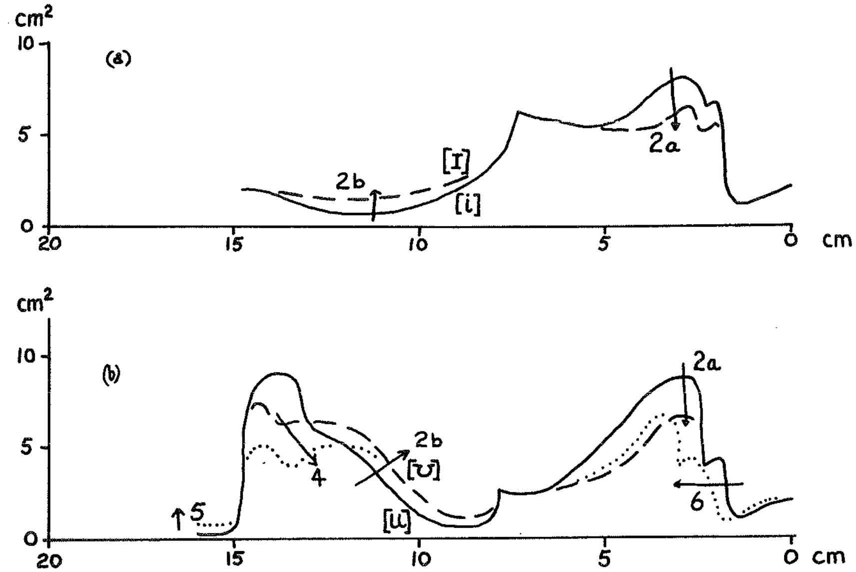


Fig. 6. Area functions for the configurations at Fig. 5(a, b).

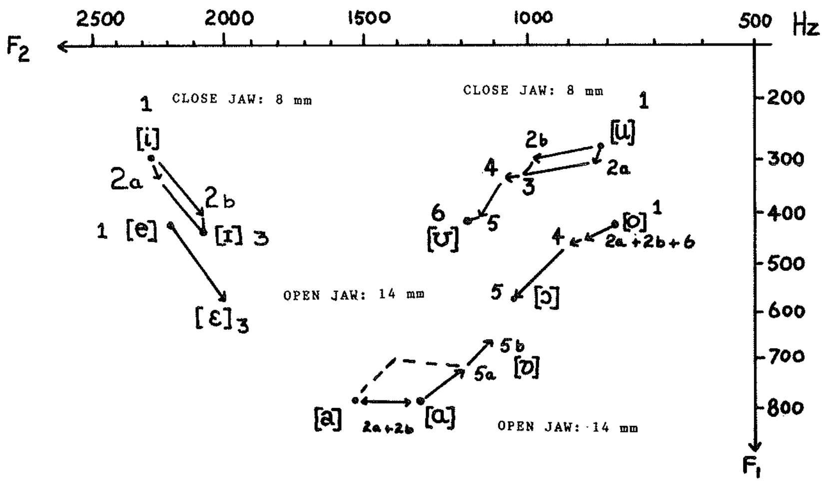


Fig. 7.

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

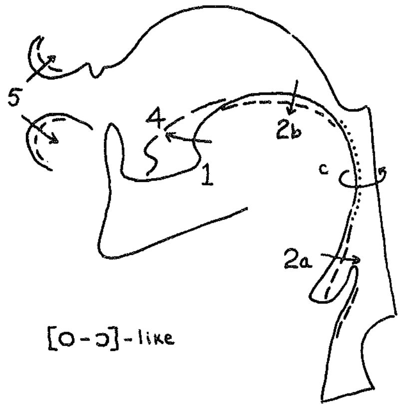
Both are in the right direction, but the contribution of the narrowed pharynx was small compared with that of the widened constriction.

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

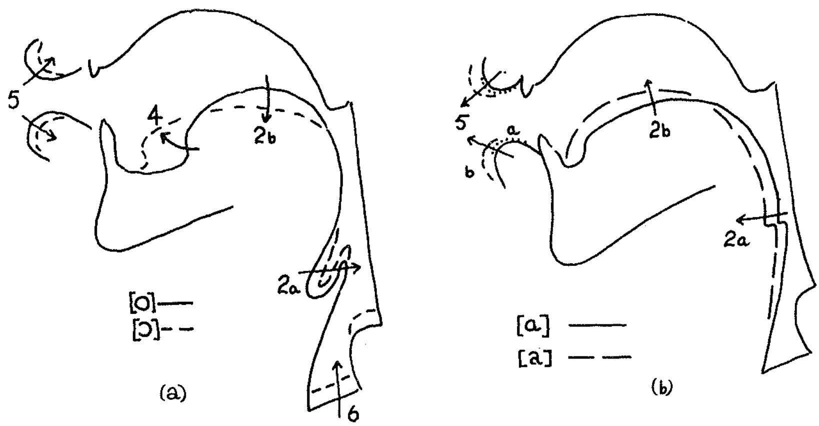


Fig. 9. Modifications made to the model profile for (a) pharyngovelar vowels and (b) low pharyngeal vowels.

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.

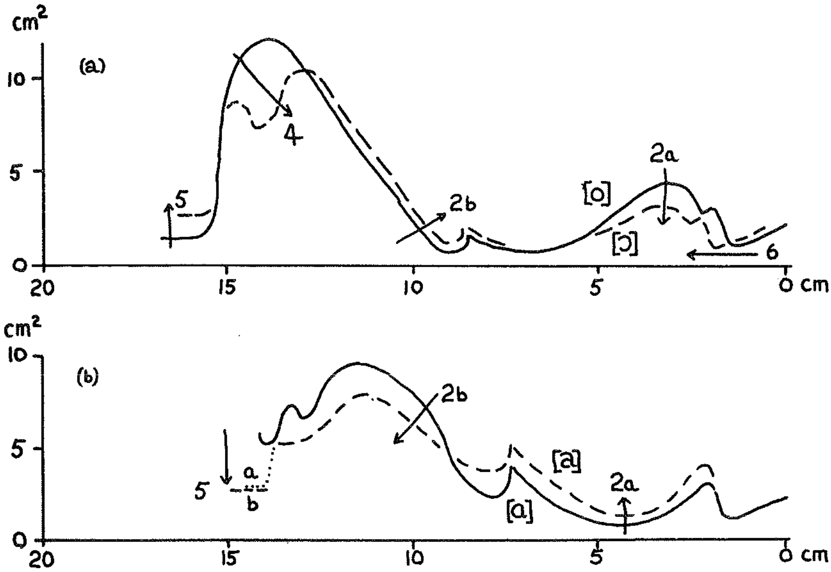


Fig. 10. Area functions for the configurations at Fig. 9(a, b).

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 participates 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

BIBLIOGRAPHY

Allen W S (1953). *Phonetics in Ancient India.* London

Anthony J F K (1964). Replica of the vocal tract. *UCLA Working Papers in Phonetics* 1: 10-14

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

Barbier P (1978). Les mouvements du larynx dans la chaîne parlée en français. *Travaux de l'Institut de Phonétique de Strasbourg* 10: 98-119

Bell A M (1863). *The Principles of Speech and Vocal Physiology; and Dictionary of Sounds.* London and Edinburgh

Bell A M (1867). *Visible Speech.* London

Bell-Berti F & Hirose H (1973). Patterns of palatoglossal activity and their implications for speech. *Status Report on Speech Research* 34: 203-209. Haskins Laboratories

Benguerel A P & Cowan H A (1974). Coarticulation of upper lip protrusion in French. *Phonetica* 30: 41-55

Bloomfield L (1969). *Language.* London (10th impression, first published 1933)

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

Brichler-Labaeye C (1970). *Les voyelles françaises, mouvements et positions articulatoires à la lumière de la radiocinématographie.* Centre de Philologie et de Littératures Romanes de l'Université de Strasbourg. Série A, No 18

Brücke E (1856). *Grundzüge der Physiologie und Systematik der Sprachlaute für Linguisten und Taubstummenlehrer.* Vienna

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

Carmody F (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. 21, No 5

Chiba T & 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, Pauphilet A & Polland B (1938). *Radiographie francouzských samohlásek a polosamohlásek.* Rozpravy České Akademie Věd a Umění, Trida 3, Cislo 75. Prague

Danielsson B (1955). *John Hart's Works.* Part I. Stockholm (including a reprint of Hart's *Orthographie*)

Darembourg C (1854). *Oeuvres anatomiques, physiologiques et médicales de Galien.* 2 Vols. Paris

Delattre P (1951). The physiological interpretation of sound spectrograms. *Publications of the Modern Languages Association of America* 66: 864. Reprinted in Delattre P: *Studies in French and Comparative Phonetics*, The Hague, 1966

Elliott R W V (1954). Isaac Newton as phonetician. *Modern Language Review* 49: 5-12 (including extracts from Newton's notebooks)

Fant C G M (1960). *The Acoustic Theory of Speech Production.* The Hague

Fant C G M (1965). Formants and cavities. *Proceedings of the Fifth International Congress of Phonetic Sciences*, pp 120-141. Basel

Fant C G M (1973). *Speech Sounds and Features.* Cambridge, Mass

Fant C G M (1975). Vocal tract area and length perturbations. *Quarterly Progress and Status Report STL-QPSR* 4/1975: 1-14. Speech Transmission Laboratory, Royal Institute of Technology, Stockholm

Fant C G M (1980). The relation between area functions and the acoustic signal. *Phonetica* 37: 55-86

Fant C G M & Pauli S (1975). Spatial characteristics of vocal tract resonance modes. In Fant C G M (Ed), *Speech Communication* Vol 2, pp 121-132. Stockholm

Fischer-Jørgensen E (1954). Acoustic analysis of stop consonants. *Miscellanea Phonetica* Vol II, pp 42-59

Fleisch H (1957). Esquisse d'une histoire de la grammaire arabe. *Arabica* 4: 1-22

Fleisch H (1958). La conception phonétique des arabes d'aprés le *Sirr Sinā'at al-I'rab* d'Ibn Ginni. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 108: 74-105

Fritzell B (1969). The velopharyngeal muscles in speech. *Acta Oto-Laryngologica* Suppl 250

Frøkjaer-Jensen B (1966). The Danish long vowels. *Annual Report of the Institute of Phonetics University of Copenhagen* 1: 34-45

Fujimura O & Kakita Y (1979). Remarks on quantitative description of lingual articulation. In Lindblom B & Öhman S (Eds), *Frontiers of Speech Communication Research*, pp 17-24. London

Gairdner W H T (1935). The Arab phoneticians on the consonants and vowels. *The Moslem World* 25: 242-257

Galen. *On the Uses of the Parts of the Body.* French translation by Darembourg (1854)

Gay T, Lindblom B & Lubker J (1981). Production of bite-block vowels: Acoustic equivalence by selective compensation. *Journal of the Acoustic Society of America* 69: 802-810

Ghosh M (1938). *Pāṇinīya Śikṣā or the Śikṣā Vedānga ascribed to Pāṇini.* Calcutta

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 (1930). Röntgenaufnahmen von Zunge und Gaumensegel bei Vokalen und Dauerkonsonanten. *Fortschritte auf dem Gebiete der Röntgenstrahlen* 41: 392-404

Hadding K, Hirose H & Harris K (1976). Facial muscle activity in the production of Swedish vowels: an electromyographic study. *Journal of Phonetics* 4: 233-245

Hála B (1959). *English Vowels in Phonetic Pictures*. Československé Akademie Věd, Sešit 5, Ročnik 69. Prague

Halle M (1970). *Foundations of Language* 6: 95-103

Halle M & 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. *Status Report on Speech Research SR* 25/26: 87-96. Haskins Laboratories

Hart J (1569). *An orthographie, conteyning the due order and reason howe to write or paint thimage of mannes voice, most like to the life or nature*. London (reprint in Danielsson 1955)

Heinz J M & Stevens K N (1964). On the derivation of area functions and acoustic spectra from cineradiographic films of speech. Paper read at the 67th Meeting of the Acoustical Society of America. *Journal of the Acoustical Society of America* 36: 1031(A)

Hellwag C F (1781). *Dissertatio inauguralis physiologico-medico de formatione loquelae*. Tübingen. Facsimile reprint by Vietor W (Ed), Heilbronn, 1886

Helmholtz H von (1863). *Die Lehre von den Tonempfindungen*. Braunschweig

Hughes O M & Abbs J H (1976). Labial-mandibular coordination in the production of speech: implications for the operation of motor equivalence. *Phonetica* 19: 271-245

Huizinga E & Moolenaar-Bijl A (1941). Kompensierung in Anzatsstück bei der Bildung von Selbstlauten. *Archives Néerlandaises de Phonétique Expérimentale* 17: 1-8

Jakobson R, Fant C G M & Halle M (1952).*Preliminaries to Speech Analysis*. Cambridge, Mass

Jespersen O (1913). *Lehrbuch der Phonetik*. Leipzig (2nd revised edition, first published 1904)

Jones D (1967). *An Outline of English Phonetics*. Cambridge (9th edition, 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

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

Keil F (1855-1880). *Grammatici Latini*, 7 Vols. Leipzig (Vol 6)

Kiritani S, Itoh K & Fujimura O (1975). Tongue pellet tracking by a computer-controlled x-ray microbeam system. *Journal of the Acoustical Society of America* 57: 1516-1520

Kiritani S, Itoh K, Imagawa H, Fujisaki H & Sawashima M (1975). Tongue pellet tracking and other radiographic observations by a computer controlled x-ray microbeam system. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics* 9: 1-14

Koneczna H & Zawadowski W (1956). *Obrazy rentgenografize glosek rosijskich.* Warsaw

Korlén G & Malmberg B (1959). *Tysk ljudlära.* Lund

Kruisinga E (1925). *A Handbook of Presentday English. Part I. English Sounds.* Utrecht

Ladefoged P (1964). *A Phonetic Study of West African Languages.* Cambridge

Ladefoged P (1967). *Three Areas of Experimental Phonetics.* Oxford

Ladefoged P (1971). *Preliminaries to Linguistic Phonetics.* Chicago

Ladefoged P (1975). *A Course in Phonetics.* New York

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

Ladefoged P, DeClerk J L & Harshman R (1972). Control of the tongue in vowels. *Proceedings of the Seventh International Congress of Phonetic Sciences*, pp 349-354. The Hague.

Ladefoged P, Declerk J L, Lindau M & Papçun G (1972a). An auditory-motor theory of speech production. *UCLA Working Papers in Phonetics* 22: 48-75

Ladefoged P, Harshman R & Rice L (1977). Vowel articulation and formant frequencies. *UCLA Working Papers in Phonetics* 38: 16-40

Liberman A M, Cooper F S, Shankweiller D P & Studdert-Kennedy M (1967). Perception of the speech code. *Psychological Review* 74: 431-461. Reprinted in Pribram K H (Ed), *Brain and Behaviour* Vol 4 (1969)

Lieberman P (1972). *The Speech of Primates.* The Hague

Lieberman P (1976). Phonetic features and physiology: a reappraisal. *Journal of Phonetics* 4: 91-112

Lindau M (1975). Vowel features. *Working Papers* 11: 1-42. Department of Linguistics, Lund. See also Lindau (1978)

Lindau M (1978). Vowel features. *Language* 54: 541-562

Lindau M, Jacobson L & 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, Lubker J & Gay T (1979). Formant frequencies of some fixed-mandible vowels and a model of speech motor programming by predictive simulation. *Journal of Phonetics* 7: 147-161

Lindblom B & Sundberg J (1969a). A quantitative model of vowel production and the distinctive features of Swedish vowels. *Quarterly Progress and Status Report STL-QPSR* 1/1969: 14-32. Speech Transmission Laboratory, Royal Institute of Technology, Stockholm

Lindblom B & Sundberg J (1969b). A quantitative theory of cardinal vowels and the teaching of pronounciation. *Quarterly Progress and Status Report STL-QPRS* 2-3/1969: 19-25. Speech Transmission Laboratory, Royal Institute of Technology, Stockholm

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

Lindqvist J & Sundberg J (1971). Pharyngeal constrictions. *Quarterly Progress and Status Report STL-QPSR* 4/1971: 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

Lubker J (1975). Normal velopharyngeal function in speech. *Clinics in Plastic Surgery* 2: 249-259

Lubker J & Gay T (1982). Anticipatory labial coarticulation. Experimental, biological and linguistic variables. *Journal of the Acoustical Society of America* 71: 437-448

Lubker J, McAllister R & Lindblom B (1977). On the notion of interarticulator programming. *Journal of Phonetics* 5: 213-226

Lusted L B & Miller E (1956). Progress in direct cineroentgenography. *American Journal of Roentgenography* 75: 56-62

Lyttkens I & Wulff F A (1885). *Svenska språkets ljudlära och beteckningslära*. Lund

MacNeilage P & Sholes G N (1964). An electromyographic study of the tongue during vowel production. *Journal of Speech and Hearing Research* 7: 211-232

Madsen J af Aarhus (1589). *De Literis Libri Duo*. Basel. (Facsimile and Danish translation in Møller and Skautrup 1930)

Malmberg B (1966). *Spansk fonetik*

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

McAllister R (1980). Some phonetic correlates of the tense-lax feature in Swedish rounded vowels. *Journal of Phonetics* 8: 39-61

McAllister R, Lubker J & Carlson J (1974). An EMG study of some characteristics of the Swedish rounded vowels. *Journal of Phonetics* 2: 267-278

Mermelstein P (1973). Articulatory model for the study of speech production. *Journal of the Acoustical Society of America* 53: 1070-1082

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*)

Minifie F D, Abbs J H, Tarlow A & Kwaterski M (1974). EMG activity within the pharynx during speech production. *Journal of Speech and Hearing Research* 17: 497-504

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

Miyawaki K, Hirose H, Ushijima T & Sawashima M (1975). A preliminary report on the electromyographic study of the activity of the lingual muscles. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics* 9: 91-106. Tokyo

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

Møller C & Skautrup P (Eds) (1930-31). *Jacobi Matthie Arhusiensis, De literis libri duo, herausgegeben und erläutert von Christen Møller und Peter Skautrup mit einer Dänischen Ubersetzung nebst einer Abhandlung über Text und Quellen von Franz Blatt.* 2 Vols, Aarhus

Mrayati M & Carré R (1976). Relations entre la forme du conduit vocal et les traits caractéristiques acoustiques des voyelles françaises. *Phonetica* 33: 285-306

Nordström P E (1975). Attempts to simulate female and infant vocal tracts from male area functions. *Quarterly Progress and Status Report STL-QPSR* 2-3/1975: 20-23. Speech Transmission Laboratory, Royal Institute of Technology, Stockholm

Ohnesorg K & Švarny O (1955). *Etudes experimentales des articulations chinoises.* Rospravý Československé Akademie Věd, Ročnik 65, Sešit 5

Parmenter C E & Trevino S N (1932). An x-ray study of Spanish vowels. *Hispania* 15: 483-496. Stanford University

Perkell J (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. *Quarterly Progress Report* 102: 123-139. Research Laboratory of Electronics, Massachusetts Institute of Technology

Perkell J (1980). Phonetic features and the physiology of speech production. In Butterworth B (Ed), *Language Production*, Vol 1 *Speech and Talk*, pp 337-372. London

Perkell J & Nelson W L (1981). Articulatory targets and speech motor control: a study of vowel production. Paper read at the Symposium on Speech Motor Control, Stockholm, May 1981

Peterson G & Barney H L (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24: 175-185

Pétursson M (1974). *Les articulations de l'islandais à la lumière de la radiocinématographie.* Société de Linguistique de Paris, Collection Linguistique, No 68

Polland B & Hála B (1926). *Artikulace českych zvuků v roentgenových obrazech.* Nā kladem Ceské Akademie Věd a Uměnī. Prague

Pols L C W, Tromp H R C & Plomp R (1973). Frequency analysis of Dutch vowels from 50 male speakers. *Journal of the Acoustical Society of America* 53: 1093-1101

Raphael L J & Bell-Berti F (1975). Tongue musculature and the feature of tension in English vowels. *Phonetica* 32: 61-73

Raphael L J, Bell-Berti F, Collier R & Baer T (1979). Tongue position in rounded and unrounded front vowel pairs. *Language and Speech* 22: 37-48

Riordan C (1976). Electromyographic correlates of the phonological /y/-/u/ distinction in French. *Journal of Phonetics* 4: 1-16

Riordan C (1977). Control of vocal tract length in speech. *Journal of the Acoustical Society of America* 62: 998-1002

Russel O (1928). *The Vowel.* Columbus, Ohio

Scheier M (1909). Die Bedeutung des Röntgenverfahrens für die Physiologie der Sprache und Stimme. *Archiv für Laryngologie und Rhinologie* 22: 175-208

Schroeder M R (1967). Determination of the geometry of the human vocal tract by acoustic measurements. *Journal of the Acoustical Society of America* 41: 1002-1010

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

Smith T (1971). A phonetic study of the extrinsic tongue muscles. *UCLA Working Papers in Phonetics* 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 & Denes P B (Eds), *Human Communication, a Unified View*, pp 51-66

Stevens K N & 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 & House A S (1961). An acoustical theory of vowel production and some of its implications. *Journal of Speech and Hearing Research* 4: 75-320

Stevens K N & House A S (1963). Perturbation of vowel articulations by consonantal context: an acoustical study. *Journal of Speech and Hearing Research* 6: 111-128

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

Subtelny J D, Pruzansky S & Subtelny J (1957). The application of roentgenography in the study of speech. In Kaiser L (Ed), *Manual of Phonetics*, pp 166-179. Amsterdam

Sundberg J (1969). Articulatory differences between sung and spoken vowels in singers. *Quarterly Progress and Status Report STL-QPSR* 1/1969: 33-45. Speech Transmission Laboratory, Royal Institute of Technology, Stockholm. (With an appendix on the problem of obtaining area functions from lateral x-ray pictures of the vocal tract)

Sundberg J (1977). The acoustics of the singing voice. *Scientific American*, March 1977: 82-91

Sundberg J & Nordström P E (1976). Raised and lowered larynx - the effect on vowel formant frequencies. *Quarterly Progress and Status Report STL-QPSR* 2-3/1976: 35-39. Speech Transmission Laboratory, Royal Institute of Technology, Stockholm

Sweet H (1877) *A Handbook of Phonetics*. Oxford

Sweet H (1906) *A Primer of Phonetics*. 3rd edn. Oxford

Tiffany W R (1959). Nonrandom sources of variation in vowel quality. *Journal of Speech and Hearing Research*. 2: 305-317

Tuller B & Fitch H L (1980). Preservation of vocal tract length - a negative finding. *Journal of the Acoustical Society of America* 67: 1068-1071

Wängler H H (1961). *Atlas der Deutsche Sprachlaute*. 2nd edn. Berlin

Wood S (1971). A spectrographic study of allophonic variation and vowel reduction in West Greenlandic Eskimo. *Working Papers* 4: 58-94. Department of Linguistics, Lund

Wood S (1975a). The weaknesses of the tongue arching model of vowel articulation. Working Papers 11: 55-108. Department of Linguistics, Lund. The thesis summary in this volume comprises most of this article

Wood S (1975b). Tense and lax vowels - degree of constriction or pharyngeal volume? *Working Papers* 11: 110-134. Department of Linguistics, Lund. Paper V in this volume

Wood S (1977). A radiographic analysis of constriction locations for vowels. *Working Papers* 15: 101-131. Department of Linguistics, Lund. Paper II in this volume. See also Wood (1979)

Wood S (1979). Tongue retraction is not so useful after all. *Working Papers* 16: 173-179. Department of Linguistics, Lund. Paper I in this volume

Wood S (1979). A radiographic analysis of constriction locations for vowels. *Journal of Phonetics* 7: 25-43. Paper II in this volume

Wood S (1982). The acoustical consequences of tongue, lip and larynx articulation in rounded palatal vowels. Paper III in this volum.

Wood S (1982). Radiographic and model studies of the palatal vowels. Paper IV in this volume

ADDENDA

Regnier (1856-58). Etudes sur le grammaire vēdique. Prātiçākhya du Rig-vēda. *Journal Asiatique* 5e sērie, Vols 7-12

Russel O (1935). Synchronized 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* pp 198-205. London

ADDENDA

(Cont'd.)

Semaan K I (1963). *Arabic Phonetics. Ibn Sīnā's Risālah on the Points of Articulation of the Speech Sounds.* Translated from the mediaeval Arabic. Arther Jeffery Memorial Monographs No 2. Lahore

Singer C (1957). *A Short History of Anatomy and Physiology from the Greeks to Harvey.* New York

Staal J F (Ed) (1972). *A Reader on the Sanskrit Grammarians.* Cambridge, Mass

Varma S (1929). *Critical Studies in the Phonetic Observations of Indian Grammarians.* London

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

Whitney W D (1862). The Atharva-Veda Prātiçākhya, or Çāunakīyā Caturādhyāyikā" text, translation and notes. *Journal of the American Oriental Society* 7: 355-615

Whitney W D (1871). The Taittirīya Prātiçākhya with its commentary the Tribhāshyaratna: text, translation and notes. *Journal of the American Oriental Society* 9: 1-469

ACKNOWLEDGEMENTS

My sincerest thanks to

- My supervisor Eva Gårding and
- Colleagues in Lund and elsewhere, especially Björn Lindblom, Jim Lubker, Joe Perkell, Johan Sundberg and Nina Thorsen, for fruitful discussions and encouragement
- Gunnar Fant and Johan Sundberg of the Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, for assistance with the line analogue LEA
- Gudmund Swahn, Gunnila Holje and Rolf Schoener of the Roentgen Technology Section, Lund University Hospital, for assistance with the x-ray films
- Eli Fischer-Jørgensen, Department of Phonetics, University of Copenhagen, for the loan of radiograms
- Jørgen Rischel, Department of Linguistics, University of Copenhagen, for assistance with the West Greenlandic film
- Sven Norén of the Lund University Computer Centre for programming
- Eva Körner, Britt Nordbeck, Birgitta Jensen, Lena Svensson and Eva Björkbom for their willing help in getting the manuscript together
- To the University Photocopying Unit and Bloms Tryckeri for prompt service

Thank you all