1668

Journal of Phonetics (1996) **24,** 399–421

An MRI-based study of pharyngeal volume contrasts in Akan and English

Mark K. Tiede

ATR Human Information Processing Research Laboratories, 2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02, Japan

Received 2nd May 1995, and in revised form 4th June 1996

Characteristic differences in pharyngeal volume between Akan ± Advanced Tongue Root (ATR) vowel pairs have been investigated using Magnetic Resonance Imaging (MRI) techniques, and compared to the Tense/Lax vowel distinction in English. Two subjects (one for each language) were scanned during steady-state phonation of three pairs of contrasting vowels. Analysis of the resulting images shows that it is the overall difference in pharyngeal volume that is relevant to the Akan vowel contrast, not just the tongue root advancement and laryngeal lowering previously reported from X-ray studies. The data also show that the ATR contrast is articulatorily distinct from the English Tense/Lax contrast.

---

1. Introduction

1.1. *Akan vowel harmony*

Several West African languages have a phonological process of Vowel Harmony that splits the distribution of their vowels into two congruent and mutually exclusive sets. The Akan language, a member of the Niger-Congo Kwa group (Stewart, 1971) spoken in Ghana, is typical of this pattern. The non-nasal vowels of Akan may be grouped as follows:

(1)

Set 1 (+ATR)			Set 2 (−ATR)	
i	u		ɪ	ʊ
e	o		ɛ	ɔ
		a		

In general, vowels in an Akan word harmonize; that is all vowels are drawn exclusively from one set or the other. The low mid vowel /a/ may occur in a word with vowels from either set. (In Clements' (1980) analysis of Akan vowel harmony /a/ is treated as an opaque vowel that induces the [−ATR] feature on any subsequent vowels.) Some dialects of Akan have an additional [+ATR] front vowel /ɜ/ that alternates with /a/; however, the exact distribution of this vowel is

0095-4470/96/040399 + 23 $25.00/0

unclear, and it is ignored here. Affixes have two forms for compatibility with stems having set 1 or set 2 vowels. For example (from Dolphyne, 1988: 15):

(2)	mI	+	di	→	midi	"I eat"
	[1st Sg]		[eat]			
	O	+	di	→	odi	"he eats"
	[3rd Sg]		[eat]			
	mI	+	dɪ	→	mɪdɪ	"I am called"
	[1st Sg]		[be called]			
	O	+	dɪ	→	ɔdɪ	"he is called"
	[3rd Sg]		[be called]			

Drawing on a cineradiographic study of the related language Igbo by Ladefoged (1964), Stewart (1967) proposed an analysis of Akan in which corresponding vowels from the two harmony groups are distinguished articulatorily by differences in tongue root position. Cineradiographic studies of Akan undertaken by Lindau (1975, 1979) confirmed the primacy of tongue root position in the harmony mechanism, but also showed correlated variation with larynx height: advanced tongue root was combined with lowered larynx, and retracted root with raised larynx (these results were consistent for all five of the subjects examined). In reporting this work Lindau suggested that the relevant contrast was not just the relative positions of these articulators but rather the overall difference in pharyngeal volume produced by their cooperative positioning, and proposed the feature "Expanded" to describe it. A contrast based on differences in pharyngeal volume suggests that vowels from the two groups may also differ in the left-to-right or lateral dimension of the pharynx, assuming that this is subject to voluntary control, but studies based on cineradiography are inherently unable to explore this possibility, since the technique collapses all lateral information into a flat (sagittal) image. The purpose of the study discussed here was to investigate the predicted difference in pharyngeal volume using Magnetic Resonance Imaging (MRI), which is not subject to the same limitation.

1.2. *English tense/lax*

The English Tense/Lax distinction eludes precise articulatory description, but is similar in many ways to the Akan contrast: tense vowels are generally articulated with an advanced tongue root, and sometimes a lowered larynx. English vowels comparable to those used in Akan may be grouped as in (1) above:

(3)	Tense		Lax	
	i	u	ɪ	ʊ
	e	o	ɛ	ɔ

Previous attempts to identify the English distinction with the same mechanism used in the Akan ATR contrast include a proposal by Halle & Stevens (1969), and a cineradiographic study by Perkell (1971). However, a more extensive cineradiographic study conducted by Ladefoged, DeClerk, Lindau & Papcun (1972) showed that tongue root advancement is just one of several complementary strategies used to implement the Tense/Lax contrast, and is not used consistently by all speakers.

Similarly, electromyographic data reported by Raphael & Bell-Berti (1975) showed differences between subjects in patterns of muscle tension used to distinguish between tense and lax vowels. Factor analysis of X-ray-derived tongue shapes by Harshman, Ladefoged & Goldstein (1977) showed that tongue position for English tense/lax pairs can be predicted with reference to just two parameters along which each of the tongue root and tongue dorsum positions covary, whereas a similar analysis of Akan by Jackson (1988) based on Lindau's data found three parameters necessary for tongue shape specification. While it is probably the case that different mechanisms are involved in each language, they are similar enough to make direct comparison feasible and interesting, and so they have been treated in parallel in the current study.

1.3. *MRI*

The magnetic resonance technique has only recently become a viable imaging alternative. Developed primarily for medical diagnostic purposes, MRI exploits the behavior of hydrogen nuclei in a magnetic field to construct an image correlated with the concentration of hydrogen in the scanned tissue (see e.g., Bradley, Newton & Crooks, 1983 for discussion). Because different tissue types have differing hydrogen densities and bondings, MR images provide soft tissue definition over a range inaccessible to X-ray techniques. Two additional advantages make MRI an especially attractive imaging modality: there are currently no known health risks for the subject associated with the technique, and because the imaging plane may be reoriented without moving the subject, three-dimensional data collection is possible.

Despite these advantages, the MRI technique has substantial drawbacks in its potential for phonetic research, chief of which is the tradeoff between imaging time and resulting image quality. Although image acquisition rates continue to drop as the technology evolves, they are currently still too slow by an order of magnitude for capturing dynamic speech. In addition, three-dimensional scanning requires multiple passes through the same volume, further increasing acquisition time. This limits the current usefulness of MRI in phonetics to studies involving static vocal tract shapes and sustainable patterns of phonation.

The current study was able to proceed under these constraints. Although vowels sustained for many times their normal speaking duration represent an admittedly artificial source of data, all of the vowels examined (except English lax vowels) can occur in open syllables, and are therefore artificial only in duration, and not in syllable structure. Moreover, the fact that sung vowels of constant pitch and quality and extended duration occur in the musical traditions of both Ghanaian and American cultures provided some context for the task required of the subjects. There is also precedent for use of the MRI technique applied to measurement of static vowel shapes in the work of Baer, Gore, Gracco & Nye (1991) for example, who successfully used it to obtain vocal tract area functions of four English vowels for two subjects.

2. Method

In this experiment, the contrast in cross-sectional area was examined at adjacent levels through the pharynx for corresponding expanded and constricted vowels. (To facilitate cross-language comparison in the following discussion I freely apply the terms "expanded" and "constricted" to both Akan and English, but do not mean to imply by this that the same feature mechanism is involved in both languages; nor am I referring to Lindau's term for the Akan feature. The terms are simply meant as shorthand physical descriptions of the respective contrasts, so that "expanded" for example should be understood as [+ATR] in the context of Akan, and [+Tense] in the context of English.) The experiment involved two subjects, both male, in their early thirties. Subject AO is a native speaker of the Asante dialect of Akan; subject MT is a native speaker of Midwestern American English. The vowels selected for comparison were /i : ɪ/, /e : ɛ/, and /u : ʊ/, chosen because of their reasonable similarity across the two languages. Prior to the experiment a stimulus tape was prepared for each subject by making an audio recording of two repetitions of each of the target vowels, extracting the vowel portion using digital editing techniques, then recording it as a continuous 'utterance' by concatenation. The following target words were used:

(4)	Vowel	Akan		English	
	[i]	pi	"many"	hid	"heed"
	[ɪ]	fɪ	"to vomit"	hɪd	"hid"
	[e]	ʃɥe	"empty"	heɪd	"hayed"
	[ɛ]	ɔʃɥɛ	"he looked at"	hɛd	"head"
	[u]	bu	"to break"	hud	"who'd"
	[ʊ]	bʊ	"to be drunk"	hʊd	"hood"

These words were also recorded for acoustic analysis:

[o]	ako	"parrot"	hod	"hoed"
[ɔ/ɑ]	kɔɔ	"red"	hɑd	"hawed"
[a/æ]	daa	"everyday"	hæd	"had"

The experiment was performed on a General Electric Signa machine installed at the Yale New Haven Hospital. The Signa system consists of a toroidal superconducting electromagnet developing a 1.5 Tesla flux density, placed in a scanning room designed to minimize interference from external electromagnetic noise. The magnet is controlled from an operator's console outside the scanning room, and an attached computer is used for image reconstruction, collation, and storage.

After divesting themselves of all ferrous material, subjects were fitted with an earphone, microphone, and neck RF transceiver imaging coil, and positioned on their backs inside the bore of the magnet. The earphone was the terminus of a length of plastic tubing connected to a small speaker placed as far away from the magnet as possible (because the strength of the magnetic field tended to overwhelm the speaker coil). The speaker was driven by an amplifier outside the scanning room, and was used to communicate with the subject and to play the prerecorded target stimuli during scanning. The microphone was used to record subject phonation immediately prior to and following scanning; while phonation during scanning was also recorded, it was unusable for analysis because of the intensity of the noise produced by the machine.

Prior to each run the experimenter verified the target vowel by reminding the subject of the characteristic word containing it. Playback of the target vowel was then initiated through the earphone, and scanning was begun shortly after the subject began phonating. Subjects were instructed to produce the target vowel with

TABLE I. Sagittal MR imaging parameters.

Subject AO (Akan)
256 × 256 pixels × 2 passes (NEX)
GRE/30 Multi (Grass Echo)
28 cm field of view
3 mm thickness
0 mm interscan skip
8 images

Vowel	TR	TE	Time
i	32	13	2:25
ɪ	37	11	2:36
e	39	13	2:44
ɛ	39	12	2:44
u	37	11	2:36
ʊ	37	11	2:36

Subject MT (English)
256 × 256 pixels × 2 passes (NEX)
GRE/30 Multi (Grass Echo)
28 cm field of view
3 mm thickness
0 mm interscan skip
8 images

Vowel	TR	TE	Time
i	36	10	2:32
ɪ	36	10	2:32
e	37	11	2:32
ɛ	37	11	2:32
u	36	10	2:32
ʊ	37	11	2:32

steady pitch and uniform quality, to take shallow breaths without changing vocal tract configuration, and to refrain from head movement and swallowing as best they could.

Two scans were made for each vowel. The first took approximately two and a half minutes to complete and produced eight adjacent sagittal images (bisecting face) at 3 mm intervals (see Table I for imaging parameters used). The second scan took approximately three minutes; it produced 28 adjacent images at 5 mm intervals in the axial orientation perpendicular to the pharynx (see Table II). The scanning process resolved a given volume into a flat 256 × 256 pixel image. Sagittal scanning excited a 280 mm × 280 mm × 3 mm volume resulting in a resolution of 1.09 mm/pixel. Half the axial scans were done using a 280 mm × 280 mm × 5 mm volume resulting in the same 1.09 mm/pixel resolution, and half were done using a 300 mm × 300 mm × 5 mm volume giving a resolution of 1.17 mm/pixel.

3. Analysis

The images obtained were converted from 16 bit Signa format to an 8 bit (256 gray level) format compatible with display and analysis software. The axial images were normalized for a standard image density so that a given pixel magnitude

Table II. Axial MR imaging parameters.

Subject AO (Akan)
256 × 128 pixels × 1 pass (NEX)
SPGR/45 Volume (Spoiled grass)
28 cm field of view
5 mm thickness
0 mm interscan skip
28 images
TR 45
TE 5
Time 3:05
Subject MT (English)
256 × 128 pixels × 1 pass (NEX)
SPGR/45 Volume (Spoiled grass)
30 cm field of view (i, ɪ, u, ʊ)
28 cm field of view (e, ɛ)
5 mm thickness
0 mm interscan skip
28 images
TR 45
TE 5
Time 3:05

represented the same value across all series; this was done so that air–tissue boundary measurements could be made consistently. Image analysis was performed on a Macintosh computer using the public domain NIH Image program developed at the U.S. National Institutes of Health (Rasband, 1990).

3.1. *Sagittal images*

The sagittal images were used to replicate the cineradiographic results. Using the top of the second cervical vertebra as the reference, measurements of relative height for the tongue dorsum, jaw, and larynx were obtained for each vowel from the image judged to be closest to sagittal midline for that scan set. Dorsum height was measured from the point on the tongue that maximized the vertical distance from the reference (see Fig. 1); jaw height was measured from a point at the base of the root of the lower incisors; and larynx height was measured from its vertical distance to the reference. Relative tongue root advancement was measured as the length of a line drawn from the bottom of the vallecular sinuses to the rear pharyngeal wall. Fig. 1 illustrates how measurements were obtained, and shows an overlay of the vocal tract outlines obtained for the comparison between Akan /i/ and /ɪ/. Full sagittal measurements are given in Table III.

3.2. *Axial images*

The axial images were used to obtain data for the left-to-right lateral dimension inaccessible to the cineradiographic studies. For purposes of comparison, 11 sequential images were chosen from each axial scan set. The reference used for

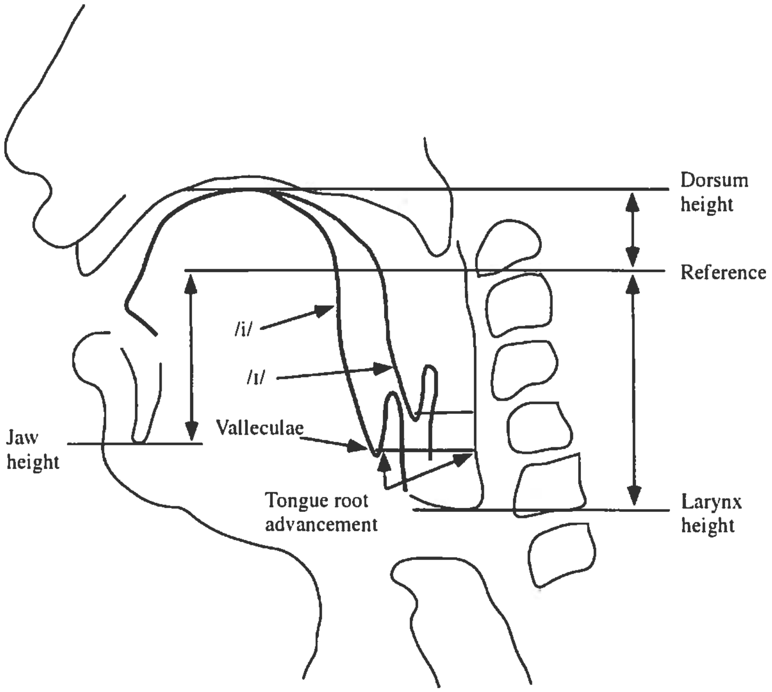


Figure 1. Sagittal measurements, showing Subject AO (Akan) MRI-derived midsagittal profile for /i/ overlain by /ɪ/.

TABLE III. Sagittal measurements (all values in mm).

Subject AO (Akan) Vowel	Root advancement	Dorsum height	Laryngeal height	Jaw height
i	22.97	19.69	68.91	48.13
ɪ	17.50	19.69	57.97	45.94
e	21.88	25.16	54.69	41.56
ɛ	19.69	24.06	51.41	47.03
u	32.81	21.88	74.38	43.75
ʊ	17.50	18.59	67.81	41.56
Subject MT (English) Vowel	Root advancement	Dorsum height	Laryngeal height	Jaw height
i	22.97	30.63	55.78	30.63
ɪ	21.88	28.44	47.03	32.81
e	21.88	27.34	51.41	36.09
ɛ	19.69	25.16	49.22	32.81
u	30.63	30.63	55.78	33.91
ʊ	18.59	26.25	54.69	33.91

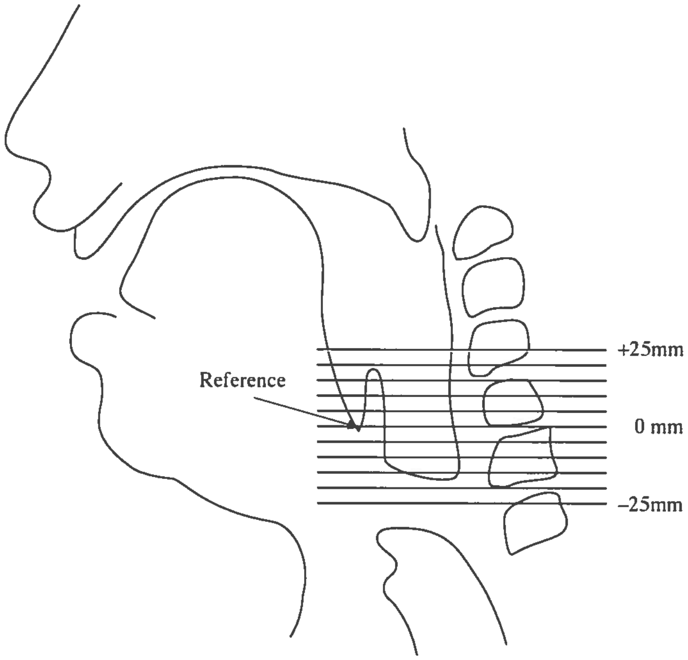


Figure 2. Approximate location of axial scans showing epiglottal reference (vallecular sinus).

comparison was the lowest image in each series in which the epiglottis was visible as a detached entity; this point corresponds to the bottom of the vallecular sinuses used in the sagittal images to identify the tongue root position. This image was used as the pivot for choosing five sequential images below that level in the pharynx, and five more above, for the total of 11. The approximate locations of axial scans measured are shown overlain on a sagittal outline in Fig. 2. The reason for using the epiglottis to coordinate inter-vowel comparison was to minimize any effects due to laryngeal height differences, so that comparison would be made at corresponding anatomical points for both vowel conditions.

Three measurements were obtained from each axial image: the distance from the anterior to the posterior boundaries of pharyngeal airspace corresponding to tongue root advancement, which will be referred to as "depth"; the novel left-to-right lateral distance across pharyngeal airspace, which will be referred to as "width"; and a measure of the cross-sectional area of pharyngeal airspace at that level; these are illustrated in Fig. 3. The two distance measurements were made along a straight line at the point of widest distension for each dimension. The area measurement was computed by determining a perimeter of standard image intensity corresponding to the pharyngeal air–tissue boundary, counting the number of enclosed pixels, and converting to area measure. Except when determining the reference or pivot level, the epiglottis was ignored in all measurements; in particular at levels above the reference (those in which a detached epiglottis was visible), the anterior wall was taken to be the base of the tongue, and area measurements included any visible epiglottal tissue.

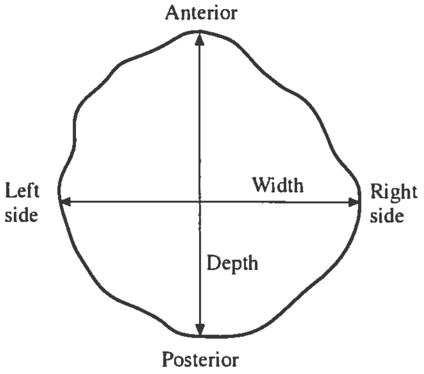


Figure 3. Axial cross-section showing orientation of axial measurements.

Axial image measurement below the epiglottis is complicated by the trifurcation of the airspace into three tubes by the aryepiglottic folds: the central laryngeal passageway and two dead-end sidechannels, the piriform sinuses. The approach taken for these images was to obtain an overall area measurement by adding the cross-section measured for the central larynx tube to those obtained for each of the piriform sinuses (i.e., aryepiglottic tissue was not included). The width measurement was made from the left side of the left sinus to the right side of the right sinus, but the depth measurement was made across the point of widest distension of the central larynx tube only. Full axial measurements are given in Table IV.

3.3. *Acoustics*

The acoustic recordings of pre- and post-scan phonation were used to verify that subject vowel quality remained on target during scanning. The resulting utterances were digitized at a sampling frequency of 10 kHz. Formant values for each token as well as the original vowel target utterances were obtained by a linear-predictive-coding autocorrelation analysis, using a 20 ms Hamming window shifted 10 ms to fit a 14th-order polynomial, with averaging of successive frames for which the first three formants were available.

4. Results

4.1. *Images*

Fig. 4 is representative of the resulting images. The left side of the figure shows a midline sagittal view and mid-pharynx view from the expanded (+ATR) variant of the Akan vowel /i/, and on the right the corresponding constricted views (/ɪ/). The orientation of the axial images corresponds to a slice through the neck, perpendicular to the pharynx, presented so that the top of the image shows the front of the face or neck. Notice the difference in pharyngeal airspace: the sagittal view on the left shows a wider opening in the lower pharynx than the corresponding view for the constricted variant on the right. Similarly, the axial view of the expanded vowel

Table IV. (a) Axial measurements for subject AO (Akan).

			Area (mm^2)			
Level (mm)	i	ɪ	e	ɛ	u	ʊ
−25	309.84	133.98	69.38	107.67	508.42	245.24
−20	230.88	137.57	137.57	205.76	534.74	290.70
−15	296.68	230.88	174.66	293.09	807.50	363.67
−10	419.90	373.24	294.29	327.78	820.65	504.83
−5	492.87	313.43	328.98	311.04	788.35	520.39
0	656.76	422.29	476.12	297.88	1233.37	878.08
5	750.07	451.00	488.09	374.44	1165.19	764.43
10	787.16	455.79	462.96	380.42	721.36	480.91
15	722.56	437.84	429.47	418.70	646.00	381.62
20	714.18	421.09	404.35	374.44	640.01	349.32
25	734.52	403.15	410.33	362.48	543.12	238.06

			Width (mm)			
Level (mm)	i	ɪ	e	ɛ	u	ʊ
−25	35.00	28.44	28.44	28.44	33.91	29.53
−20	32.81	29.53	29.53	29.53	35.00	30.63
−15	32.81	32.81	31.72	30.63	39.38	31.72
−10	35.00	33.91	33.91	29.53	38.28	31.72
−5	36.09	31.72	35.00	30.63	37.19	36.09
0	32.81	29.53	33.91	25.16	37.19	38.28
5	30.63	24.06	29.53	24.06	35.00	32.81
10	30.63	26.25	22.97	25.16	35.00	30.63
15	28.44	28.44	28.44	24.06	35.00	27.34
20	31.72	25.16	24.06	22.97	35.00	27.34
25	30.63	22.97	25.16	17.50	33.91	20.78

			Depth (mm)			
Level (mm)	i	ɪ	e	ɛ	u	ʊ
−25	19.69	18.59	20.78	12.03	37.19	26.25
−20	25.16	15.31	20.78	14.22	31.72	21.88
−15	27.34	18.59	19.69	15.31	31.72	24.06
−10	25.16	17.50	24.06	15.31	29.53	22.97
−5	17.50	14.22	14.22	14.22	30.63	20.78
0	33.91	24.06	27.34	17.50	43.75	30.63
5	33.91	18.59	27.34	20.78	42.66	28.44
10	31.72	18.59	18.59	22.97	32.81	21.88
15	29.53	24.06	20.78	21.88	27.34	18.59
20	32.81	22.97	24.06	24.06	27.34	15.31
25	31.72	24.06	24.06	25.16	25.16	12.03

on the left shows a larger cross-sectional area than the corresponding constricted vowel on the right.

In subjective terms the images obtained varied from fair to good quality compared to others of this type, representing a reasonable compromise between image noise and required scanning time (image quality was somewhat inferior to that obtained by Baer *et al.* (1991); however, those researchers used custom-made cephalostats

TABLE IV. (b) Axial measurements for subject MT (English).

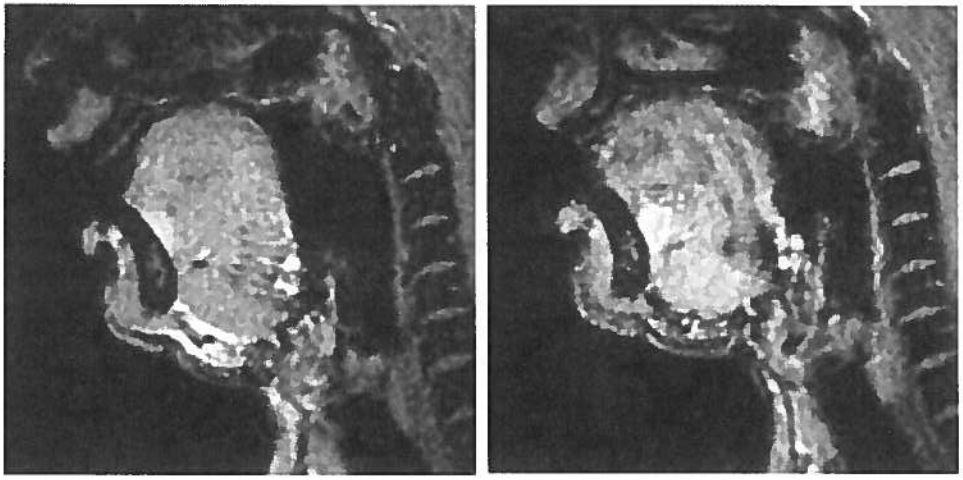
			Area (mm^2)			
Level (mm)	i	ɪ	e	ɛ	u	ʊ
−25	438.08	508.12	470.14	486.89	753.94	618.48
−20	517.73	519.10	411.52	514.40	699.01	617.29
−15	527.34	616.61	391.19	614.89	806.12	608.91
−10	661.93	649.57	576.61	590.97	862.43	752.47
−5	606.99	585.02	523.97	644.80	896.76	718.97
0	914.61	659.18	608.91	532.35	1031.34	848.17
5	951.69	630.34	653.17	368.46	958.56	802.71
10	942.08	611.11	575.42	295.48	944.82	724.95
15	948.94	597.38	517.99	324.19	855.56	628.05
20	940.70	649.57	526.37	337.35	811.61	541.92
25	961.30	708.62	525.17	394.78	752.56	485.69

			Width (mm)			
Level (mm)	i	ɪ	e	ɛ	u	ʊ
−25	32.81	33.98	33.91	33.91	33.98	32.81
−20	37.50	36.33	36.09	36.09	36.33	33.91
−15	39.84	39.84	38.28	38.28	39.84	36.09
−10	41.02	39.84	39.38	41.56	41.02	39.38
−5	41.02	41.02	39.38	40.47	41.02	40.47
0	38.67	38.67	40.47	38.28	39.84	39.38
5	37.50	38.67	39.38	37.19	38.67	39.38
10	42.19	36.33	36.09	36.09	39.84	39.38
15	43.36	41.02	35.00	33.91	41.02	31.72
20	44.53	38.67	36.09	35.00	41.02	36.09
25	45.70	42.19	33.91	33.91	39.84	38.28

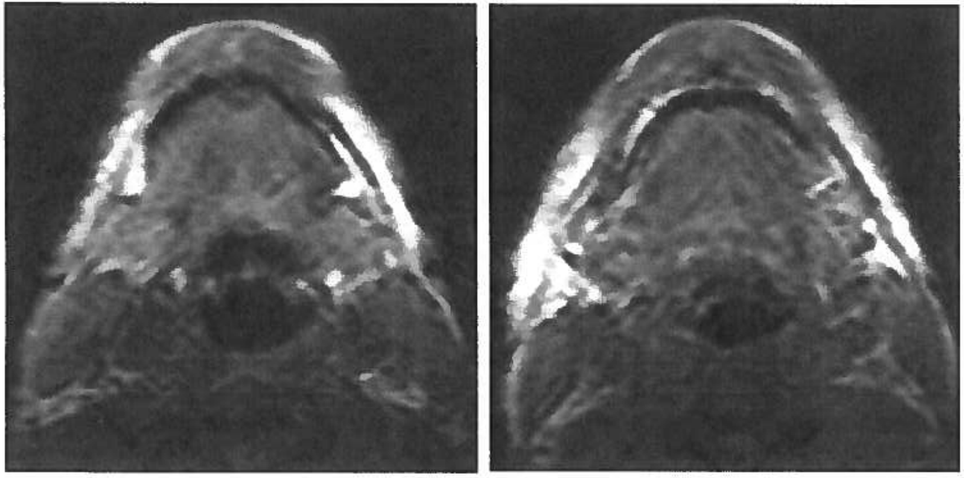
			Depth (mm)			
Level (mm)	i	ɪ	e	ɛ	u	ʊ
−25	26.95	29.30	29.53	26.25	30.47	30.63
−20	21.09	22.27	21.88	22.97	28.13	28.44
−15	19.92	19.92	21.88	24.06	29.30	26.25
−10	21.09	19.92	17.50	19.69	24.61	29.53
−5	17.58	17.58	17.50	17.50	24.61	24.06
0	31.64	23.44	24.06	17.50	31.64	26.25
5	30.47	22.27	22.97	15.31	30.47	27.34
10	30.47	21.09	20.78	14.22	29.30	19.69
15	30.47	18.75	20.78	12.03	25.78	18.59
20	31.64	23.44	18.59	12.03	24.61	16.41
25	29.30	23.44	20.78	17.50	21.09	15.31

and longer scanning times in obtaining their best results). Measurements were based on air/tissue interfaces, which tended to image well, even if resolution of tissue-internal anatomical details was non-optimal.

To determine whether normalization of measurements for different head sizes was appropriate, subject tract lengths were measured using the midline sagittal image obtained for the vowels /i/ and /ɪ/ in each language. The measurement was made


Subject AO (Akan) vowel /i/
Mid-sagittal

Subject AO (Akan) vowel /ɪ/
Mid-sagittal


Mid-pharynx

Mid-pharynx

Figure 4. Representative images.

from the level of the vocal folds, around the curve of the tongue, to the midpoint of a line bisecting the narrowest lip approximation, presented here in centimeters:

(5) Vocal Tract Lengths (cm)

Vowel	Subject AO (Akan)	Subject MT (English)
i	17.4	16.9
ɪ	15.7	15.3

These values were considered to be sufficiently close as to make normalization unnecessary.

Errors affecting measurements may have been introduced in two different ways: from distortion of the image due to subject movement during scanning, or from subject head tilt. Both possibilities would affect sagittal scanning more than axial. Sagittal scans were resolved individually one after another, whereas the axial scans were obtained using a method that computed all images concurrently. Subject movement in a sagittal scan set resulted in blurring of the single image being acquired at that moment, making it unusable for measurement, while movement in an axial set caused only a loss of definition across all images in the set, leaving them all still viable for measurement.

The effect of any head tilt or rotation was to cause the scanning plane to intersect with the subject at an oblique angle, rather than producing a true bisection of the head. This was not a problem for axial images, since any tilting would have been too slight to cause measurable distortion in that orientation. For the sagittal images however, even slight tilting was sufficient to make determination of the head midline difficult; for example, one image might show midline at the level of the larynx, yet be off center at the level of the palate, thus affecting tongue dorsum height measurements. Therefore all sagittal tongue measurements were verified on images adjacent to the one chosen as midline.

4.2. *Sagittal orientation*

Fig. 5 shows the (expanded–constricted) difference between sagittal measurements obtained for corresponding vowel pairs. In each case the Akan results confirm the cineradiographic studies mentioned above: tongue root advancement was larger for the expanded (+ATR) variant of each vowel pair, larynx height was lower for each expanded variant as well, and tongue dorsum height showed either no difference (ɪ:ɪ) or slightly higher values for expanded alternates (e:ɛ, u:ʊ). The results for English were very similar: for each tense variant root advancement was larger, larynx height was lower, and tongue dorsum higher. However, the English measurements show smaller differences in magnitude for both root advancement and laryngeal lowering than Akan, and greater differences in tongue dorsum height. Results obtained for jaw height showed small and inconsistent differences in both languages.

4.3. *Axial orientation*

4.3.1. *Akan*

The left side of Fig. 6 illustrates the difference in cross-sectional area between Akan vowel pairs. In each graph in this and subsequent figures the height of each bar shows the difference in area between expanded and constricted variants at corresponding levels of the pharynx, increasing in height at 5 mm intervals from five measurement levels below the base of the epiglottis to five levels above. Observe that the pharyngeal airspace is larger at all measured levels for the expanded (+ATR) variants of vowels /i/ and /u/, and for all levels of expanded /e/ above the epiglottal pivot. An estimate of pharyngeal volume was obtained by summing

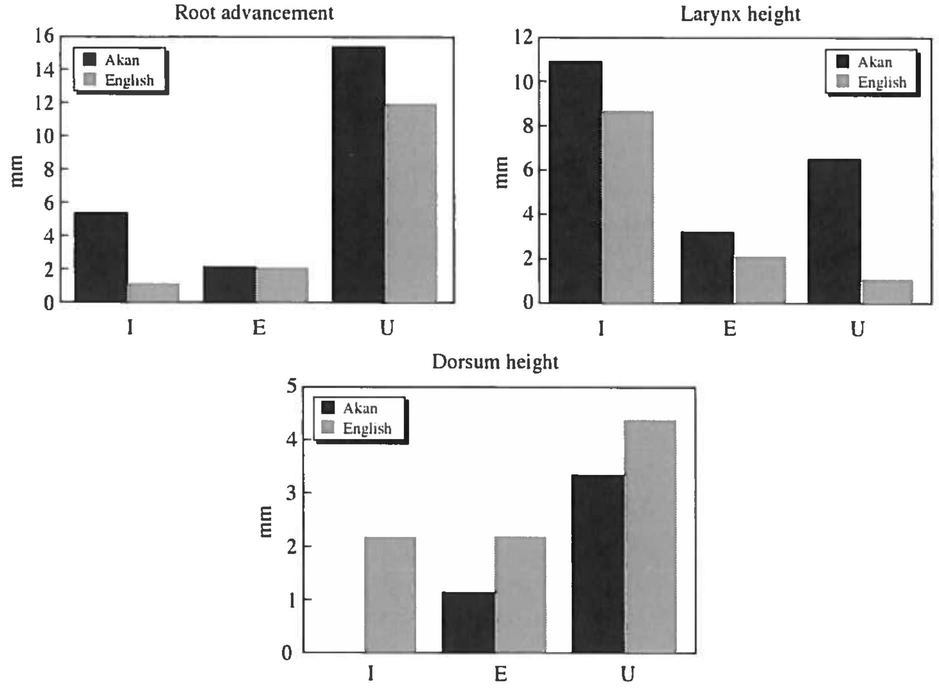


Figure 5. Sagittal results for Akan and English vowels. Bars indicate [expanded-constricted] differences obtained for corresponding vowel pairs. I refers to /i : ɪ/, E to /e : ɛ/, U to /u : ʊ/.

the measured cross-sections weighted by the (5 mm) scan thickness, representing that portion of the tract delimited by ±25 mm from the base of the epiglottis:

(6) Derived pharyngeal volumes for subject AO (Akan) (cm^3)
(I = /i : ɪ/; E = /e : ɛ/; U = /u : ʊ/

	I	E	U
+ATR	30.58	18.38	42.04
−ATR	18.90	17.27	25.09

The larger volumes observed for the "expanded" variants of each vowel show that Lindau's term is an apt name for the feature being contrasted. It is interesting that the ratio of expanded to constricted volume is nearly the same for vowels /i/ and /u/ (1.62 *vs.* 1.68), and the absence of a difference of similar magnitude for /e/ suggests that the data obtained for that vowel should be viewed with caution (although they agree with the sagittal results, obtained during a separate scanning sequence, which also showed the smallest root advancement difference for /e/). The smaller differences observed for /e : ɛ/ may be due to the fact that the inherently lower tongue constriction location characteristic of these vowels leaves less room in the lower pharynx for effecting the contrast.

The observed difference in pharyngeal volume is consistent with the tongue root advancement observed in the sagittal images and cineradiographic studies, but

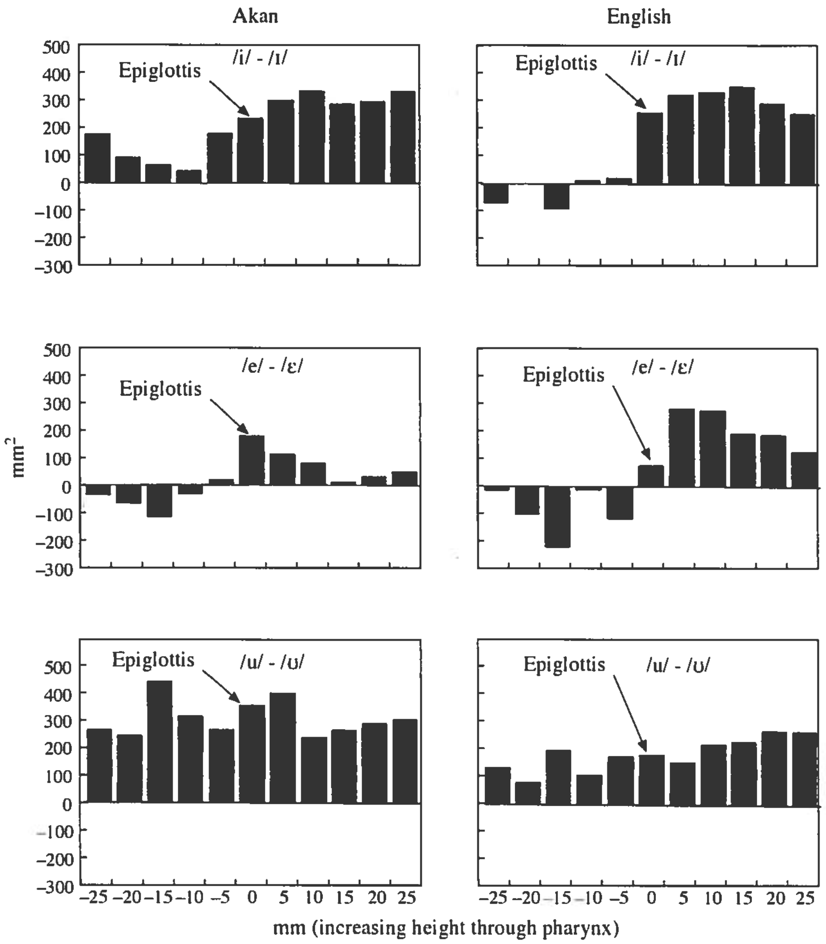


Figure 6. Cross-sectional area for Akan and English vowels. Bars indicate [expanded–constricted] difference in area at each measured pharynx level.

leaves open the question of whether the difference is due entirely to root position, or whether pharyngeal lateral width contributes as well. The relative contributions of width and depth to the area values obtained for Akan are illustrated in Figs. 7 and 8, respectively (left side). Recall that width is a measure of side-to-side or lateral distance, and depth is a measure of anterior/posterior distance corresponding to root advancement. Notice that while somewhat noiser than the area data, the overall trend for both width and depth is for consistently larger values for the expanded variant of each vowel. (Though it is true that the uppermost axial levels measured for Akan /e : ɛ/ show (slightly) larger depth values for the contracted variant, reversing the pattern found everywhere else, it is possible that at these levels scanning has moved out of the region manipulated for the ATR contrast, and into the relatively invariant region of the tongue constriction characterizing the vowel.) Notice also that the observed differences in lateral width are almost as large as those observed for anterior/posterior depth, showing that for this speaker at least control of the lateral dimension is an important part of the mechanism used to produce the

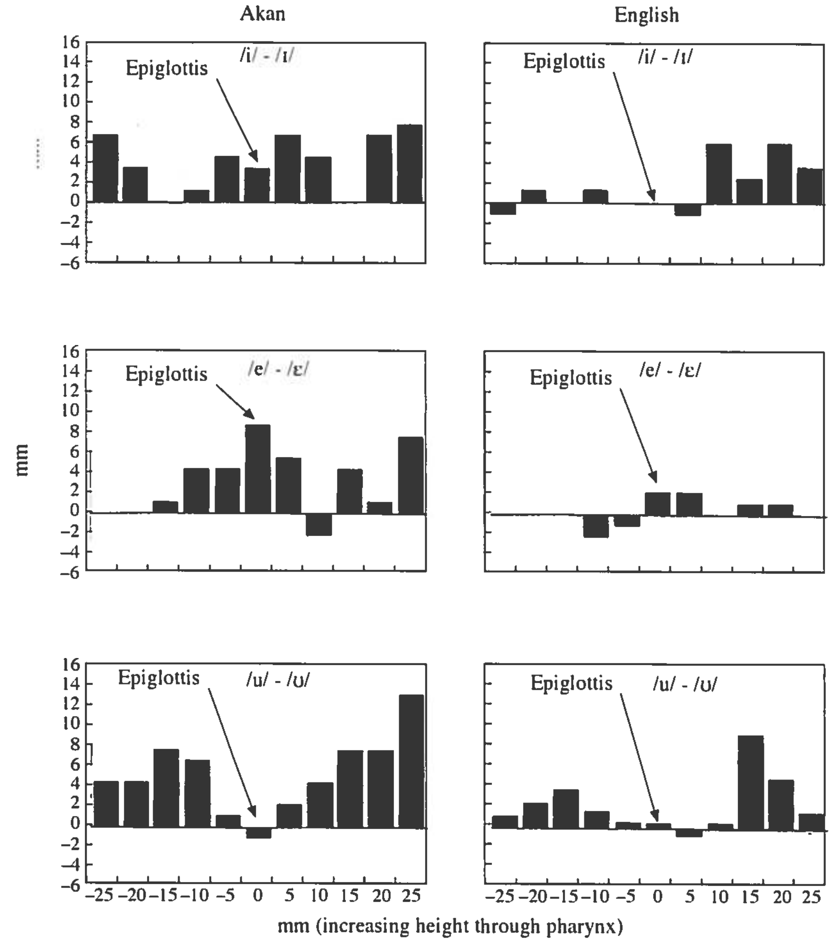


Figure 7. Comparison of pharyngeal width. Bars indicate [expanded–constricted] difference in width at each measured pharynx level.

contrast, thereby confirming Lindau's position that speakers of Akan seek to maximize the difference in overall pharyngeal volume in effecting the contrast.

4.3.2. *English*

Figs. 6, 7, and 8 provide a comparison of the English area, width, and depth results with those of Akan. As before, the height of each bar shows the difference between expanded and constricted variants at corresponding levels through the pharynx. Combining area cross-sections as for Akan, the following approximations to pharyngeal volume were obtained:

(7) Derived pharyngeal volumes for subject MT (English) (cm^3)

	I	E	U
Tense	37.25	26.28	43.10
Lax	30.13	23.55	34.31

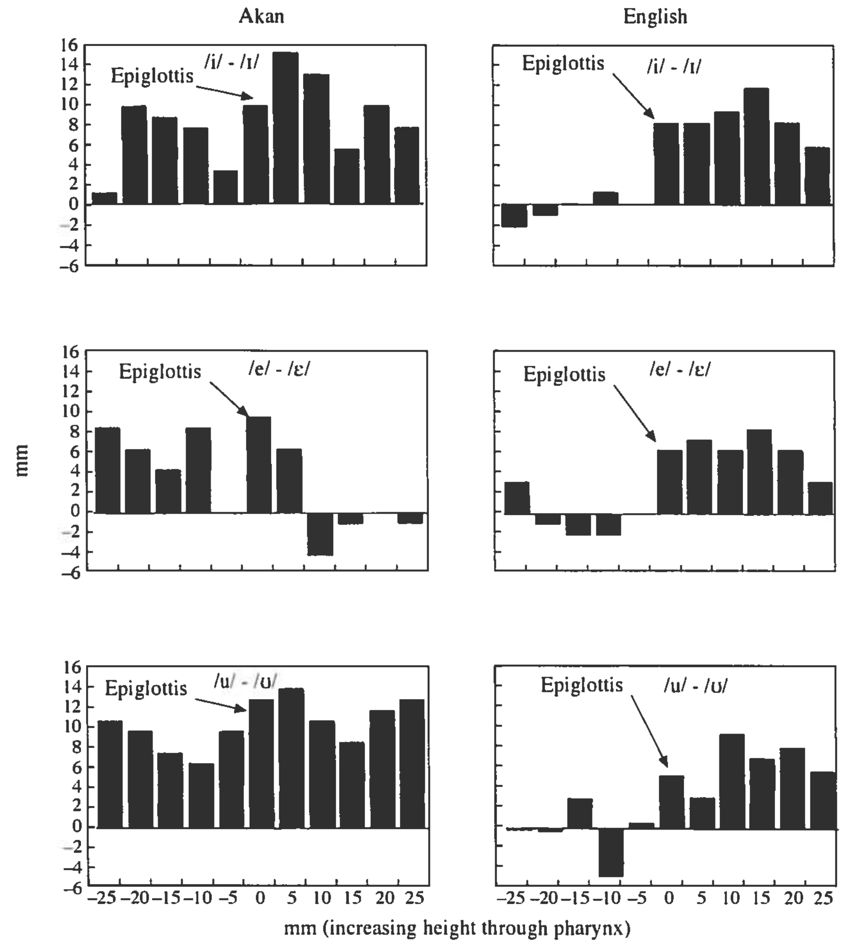


Figure 8. Comparison of pharyngeal depth. Bars indicate [expanded–constricted] difference in depth at each measured pharynx level.

As in Akan, the expanded (Tense) variants show consistently larger overall pharyngeal volume than their lax counterparts. Again, it is interesting to note that the ratio of expanded to constricted volume is nearly the same for vowels /i/ and /u/ (1.24 *vs.* 1.26), although the magnitude of the difference is considerably less than that observed for Akan.

4.3.3. *Akan and English compared*

Akan and English show different patterning of axial data at measured levels below the epiglottis. With one exception (area measured at the three lowest levels of /e/), the area, width, and depth measurements obtained for Akan show consistently larger values for expanded (+ATR) variants at all measured levels, above and below the epiglottis. But while the English data also show consistently larger values above the epiglottal pivot, at levels below that point differences between tense and lax variants are inconsistent in sign and considerably smaller in magnitude. The change occurs abruptly, and is evident to some degree in all three parameters measured.

An attempt was made to quantify the significance of this divergence by performing an analysis of variance on each measurement parameter, treating the individual levels as repeated measures nested within an upper or lower grouping factor. For purposes of symmetry, the uppermost level was dropped from the analysis, so that the five sub-epiglottal levels constituted the "lower" group, and the epiglottal pivot and its four succeeding levels the "upper." The following ANOVA design was used (levels as repeated measures nested in group): Level (measurement level, 1–5); Group (below/above epiglottis); Vowel (I/E/U); ATR (expanded/constricted); Language (Akan/English). The Level × Vowel × ATR × Language interaction was used as the error term (df = 16). Under this design, inter-language differences in behavior above and below the epiglottis between expanded and constricted configurations are encompassed in the Group × ATR × Language interaction (df = 1). Results of the analysis for the width parameter were not significant for this interaction ($F = 1.12$), but showed moderate significance for depth ($F = 9.48$, $p < 0.05$), and strong significance for area ($F = 15.84$, $p < 0.005$), reflecting the difference in sub-epiglottal behavior between the two subjects.

Further evidence of divergent behavior is provided by a post-hoc regression analysis of width against depth for corresponding measurement levels, illustrated in Fig. 9. The results show a significant ($r^2 = 0.196$, $p < 0.01$) positive correlation between Akan width and depth across all measurement levels (df = 58), but no

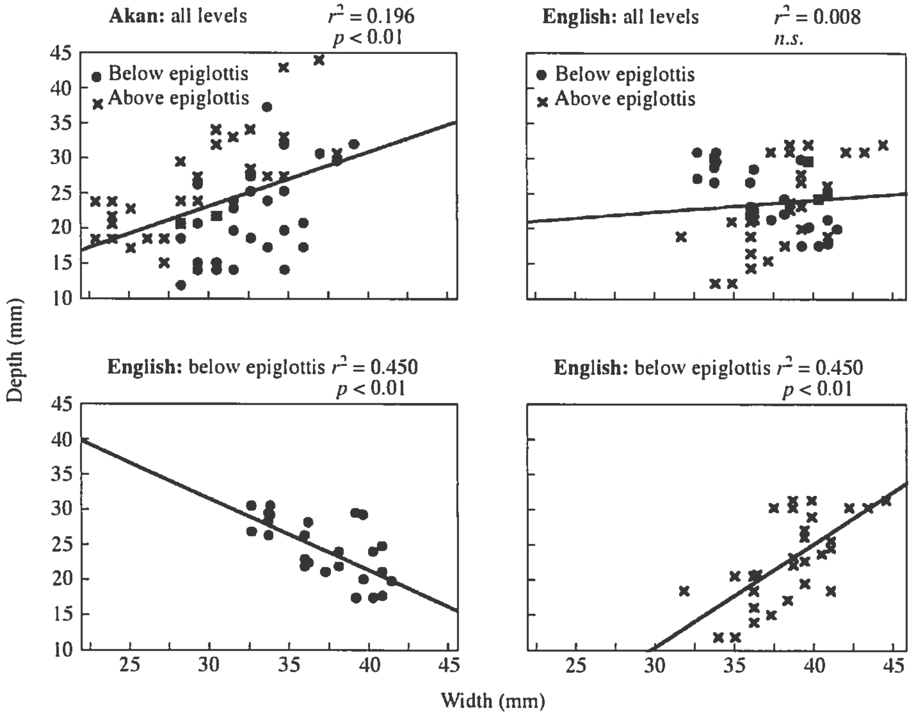


Figure 9. Regression of axial width against depth.

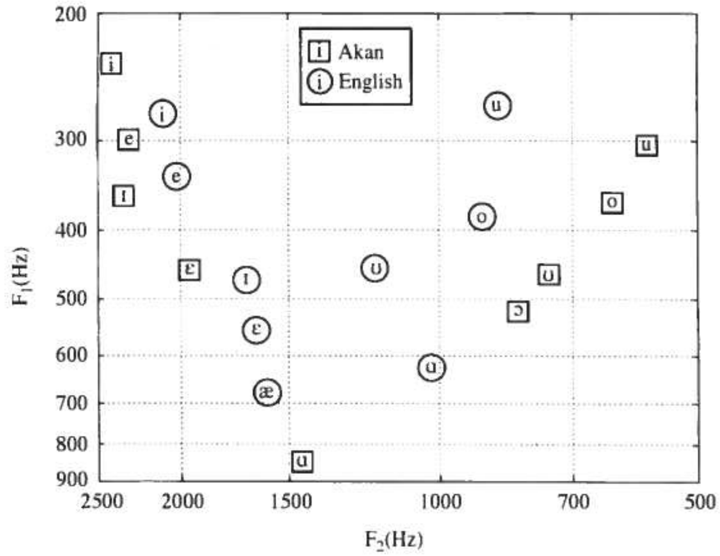


Figure 10. Combined vowel space for Akan and English subjects.

significant correlation for the same analysis on English ($r^2 = 0.008$). When the English values are analyzed separately by upper/lower grouping, however (28 df each), the results show a significant ($r^2 = 0.450$, $p < 0.01$) positive correlation between width and depth for measurement levels above the epiglottal pivot, and a significant ($r^2 = 0.462$, $p < 0.01$) *negative* correlation at levels below it; in other words, below the level of the epiglottis in English an increase in anterior/posterior depth is accompanied by a decrease in lateral width.

4.4. *Acoustics*

Based on the limited data used in this study, the primary acoustic effect of the expanded/constricted contrast in both languages is a raised F_1 for expanded vowels relative to their constricted counterparts, which is consistent with the results obtained (for Akan) by Lindau (1979) and Hess (1992). This is illustrated in Fig. 10, which shows a vowel space chart combining vowels from each language in which the average values of the LPC-derived formants obtained from the two repetitions of each stimulus token have been used to plot F_1 against F_2. The effects of the contrast on F_2 are smaller in terms of percentage differences, and differ in direction for front and back vowels (constricted front vowels have a lower F_2, constricted back vowels a higher F_2, for a net effect of centralization relative to their expanded counterparts). Except for English /u : ʊ/, values measured for F_3 were lower for constricted variants in both languages.

The two subjects representing their respective languages differed somewhat in their realization of expanded/constricted contrasts. In particular, although the overall vowel space used by the English subject is smaller, the observed formant differences between Tense/Lax pairs were larger than those between the ±ATR contrasting pairs of the Akan subject. Expressed in terms of relative differences, it is also interesting that the first formant of Akan [−ATR] vowels was a consistent 1.5

times the value of its [+ATR] counterpart (I = 1.52, E = 1.52, U = 1.52), while the English Lax:Tense ratios were both larger and less consistent (I = 1.71, E = 1.64, U = 1.69).

It may be recalled that the primary purpose of collecting acoustic data was to verify that subjects were in fact producing the target vowel configuration throughout the scanning sequence. Two analyses of variance were performed to evaluate production consistency: one grouped recorded tokens into source groupings of target (the original stimulus words), sagittal scan, and axial scan; the second analysis grouped them into target, pre-scan, and post-scan source groupings (recall that machine noise precluded analysis of phonation recorded during scanning). Ideally the formant values measured for a given vowel should be independent of its source category; in practice this expectation was unrealistic, given that the target vowels were separately recorded under acoustically ideal conditions, while pre- and post-scan tokens were recorded with subjects supine and encased by several tons of intimidating high technology. Effects due to token source on measured formant values were assessed by computing Wilks' Λ (a measure of multivariate correlation) from the separate formant analyses. The results obtained showed that token source category was in fact a significant source of variance for both subjects, both as a main effect, and in interactions with vowel and ATR factors, but the results were mixed depending on how the tokens were grouped. When analyzed by type (i.e., target, pre-scan, post-scan), the Akan subject showed a marginally significant ($F(6, 20) = 3.11$, $p < 0.05$) interaction of source with ATR (the feature of primary interest), but no significance for this interaction when analyzed by run (i.e., target, sagittal, axial). This suggests that while subject AO drifted slightly from the initial target over time, the way in which he did so was consistent across runs, at least with respect to ATR effects. The English subject showed the reverse pattern: the interaction of source with the Tense/Lax parameter was marginally significant ($F(6, 28) = 3.02$, $p < 0.05$) for the analysis of tokens grouped by run, but not significant grouped by type, indicating that while this subject produced slightly different vowels across runs, he did not drift significantly during scanning.

These results indicate that due possibly to nervousness or fatigue subjects were not in fact entirely consistent in vowel production. However, the F-values derived using Wilks' Λ obtained for vowel type, ATR, and their interaction were larger than those obtained for any term involving token source category, for both languages and both analyses. If these values are regarded as reflecting the relative importance of the individual factors, then this suggests that while the token source factor is significant, it is less important relative to the other parameters determining vowel configuration, and does not indicate that subject phonation drifted so much as to seriously skew the imaged tract configurations.

5. Discussion

The English sagittal measurements showing smaller differences in magnitude for both tongue root advancement and laryngeal lowering than Akan, and greater differences in tongue dorsum height, suggest a relatively more signficant role for tongue height in maintaining the English contrast. It should be noted, however, that root advancement and dorsum height may be intrinsically linked. An electromyographic study by Baer, Alfonso & Honda (1988) claimed that the same contraction of the posterior genioglossus (GGP) muscle effecting tongue root advancement also forces the tongue dorsum upwards, while contraction of the separately controlled anterior genioglossus (GGA) pulls the dorsum forward and down. The results obtained here suggest that GGA activity is greater in Akan than in English; active control of the GGA in Akan works against the dorsum-raising effect of the GGP, resulting in smaller net dorsum height differences than those observed for English. This predicts that the correlation of dorsum height with root advancement should be stronger in English than in Akan, which is confirmed by regression analysis ($\mathrm{df} = 4$): English shows significant correlation of dorsum height regressed against root advancement ($r^2 = 0.611$, $p < 0.05$), while the corresponding regression for Akan is not significant ($r^2 = 0.051$). Assuming the pattern of muscle activity mentioned above, one apparent difference, therefore, between the ATR and Tense/Lax mechanisms lies in how each language treats the inherent dorsum-raising effect of root advancement: Akan seeks to neutralize its effect, whereas English (at least in this subject's dialect) appears to exploit it.

Another difference is apparent from the patterning of axial data at measured levels below the epiglottis. Recall that Akan showed consistently larger values for expanded (+ATR) variants above and below the epiglottal reference. But while the English results also showed consistently larger values for expanded (Tense) counterparts above the reference, below it contrasting tense and lax vowels were inconsistent in sign and considerably smaller in magnitude. The fact that this divergence in behavior appears to occur exactly at the level of the base of the epiglottis suggests an anatomical explanation. It can be seen from the sagittal images that the base of the epiglottis is at approximately the same height as the hyoid, which serves as an anchor for the medial pharyngeal constrictor. The primary function of this muscle is to control pharyngeal aperture during swallowing in the region of the pharynx between the level of the hyoid and the upper larynx. Since this is where the Akan and English patterns diverge, the source of the cross-language difference in sub-epiglottal behavior (for these two subjects at least) may be the active involvement of this muscle in effecting the ATR contrast, and its lack of involvement in the Tense/Lax distinction: active control of pharyngeal aperture would account for the consistently larger (expanded–constricted) difference values observed for Akan, and its presumed absence would explain the small and inconsistent differences observed for English.

The trading relation observed between English width and depth also supports the hypothesized non-involvement of the medial constrictor in the Tense/Lax mechanism: when GGP action on the tongue pulls it up and forward in Tense configurations, the relaxed constrictor does not oppose it, and so the lateral pharyngeal walls collapse slightly below the epiglottis as a consequence of the advanced tongue root. In Akan, however, active control of the constrictor results in the pharyngeal walls being tensed in [+ATR] configurations, preventing similar deformation of the lateral sides. Hardcastle (1976) refers to two modes of pharyngeal muscle contraction: an "isotonic" mode inducing sphincteral narrowing of pharyngeal aperture, and an "isometric" mode serving to tense pharyngeal walls without reducing aperture. Under this account of the differences in sub-epiglottal behavior between English and Akan, English does not actively control the pharyngeal constrictor, whereas Akan uses isotonic contraction of the pharyngeal constrictor to minimize pharyngeal aperture in [−ATR] configurations, and isometric contraction in [+ATR] configurations to prevent deformation of pharyngeal lateral walls. Pharyngeal isometric tension may be significant in a different context: Hardcastle (1973) has suggested that it may be important in the production of tensed initial stops in Korean.

Assuming that the patterns observed for these two subjects are in fact representative of Akan and English, it is evident that the ATR and Tense/Lax distinctions are only superficially similar. In producing [+ATR] vowels Akan speakers enlarge the pharyngeal cavity by advancing the root of the tongue, lowering the larynx, and maintaining tension in the pharyngeal walls. In [−ATR] configurations the root is retracted, pharyngeal aperture is constricted, and the larynx is raised, minimizing pharyngeal volume. Because speakers appear to make adjustments to maintain relatively constant dorsum height across the two configurations, the ATR distinction is essentially a contrast in pharyngeal volume. Tense vowels in English are also articulated with an advanced tongue root, enlarging the pharyngeal cavity as a consequence; but below the level of the epiglottis this enlargement is counteracted by the deformation of the lateral walls from lack of constrictor tension. In lax configurations pharyngeal volume is smaller, but it is unclear whether this is due to actual constriction by the medial constrictor, or simply the relaxed position of the tongue root. The dorsum-raising effect of root advancement is not adjusted for in English, and instead constitutes an integral part of the contrast.

Generalizations of this sort are somewhat presumptuous given the limited scope of this study, and those made for English in particular should be viewed in the context of the Ladefoged *et al.* (1972) and Raphael & Bell-Berti (1975) findings mentioned above, showing that different English speakers produce the Tense/Lax contrast differently. Separate studies by Lindau in 1975 (one subject) and 1979 (four subjects) found consistent articulatory implementation of the ATR mechanism, so the generalizations made here for Akan are perhaps on firmer ground, especially as the sagittal results of this study replicated her findings. In any case, while different speakers of English may approximate more or less closely the Akan ATR articulatory mechanism, the point is that the Tense/Lax contrast is not identical to it, and must therefore be represented by a different feature. Furthermore, because the Akan distinction appears to involve two muscles that the English contrast does not exploit (the medial pharyngeal constrictor and possibly the anterior genioglossus), from an articulatory standpoint it appears to be more complex with respect to the active control and coordination needed to produce it.

6. Conclusions

Although exploratory in scope, this study did succeed in achieving certain objectives. It demonstrated that despite its inherent drawbacks, Magnetic Resonance Imaging can be a useful technique for obtaining information about static vocal tract configurations, one that will undoubtedly increase in importance as the technology is further refined. Measurements from the sagittal images collected here successfully replicated previous results from cineradiographic studies showing the significance of tongue root advancement and larynx height in effecting the Akan ATR contrast, and those obtained from axial images extended these results for the first time into the dimension of lateral width. The axial measurements confirm Lindau's position that it is the overall difference in pharyngeal volume that is relevant to the Akan vowel contrast, not just relative larynx or tongue root positions. Finally, the parallel analysis of the English Tense/Lax contrast gave results showing that despite superficial similarities with the Akan ATR distinction, the two contrasts are not the same.

I am grateful to the many people who helped me with this project, particularly Cathe Browman, Louis Goldstein, Robin Greene, Alfred Opoku, Eric Vatikiotis-Bateson, and all those who agreed to be magnetized for Science. Thanks as well to Michel Jackson, Peter Ladefoged, and Pam Speeter Beddor for careful reading and helpful suggestions; goofs and garbles remain of course my responsibility. This work was supported by NIH grant DC-00121 while the author was associated with Haskins Laboratories.

References

Baer, T., Alfonso, P. & Honda, K. (1988) Electromyography of the tongue muscles during vowels in /əpVp/ environment, *Annual Bulletin of the Research Institute for Logopedics and Phoniatrics* (University of Tokyo), **22**, 7–19.

Baer, T., Gore, J. C., Gracco, L. C. & Nye, P. W. (1991) Analysis of vocal tract shape and dimensions using magnetic resonance imaging: Vowels, *Journal of the Acoustical Society of America*, **90**(2), 799–828.

Bradley, W. G., Newton, T. H. & Crooks, L. E. (1983) Physical principles of nuclear magnetic resonance. In *Modern neuroradiology: advanced imaging techniques* (T. H. Newton & D. G. Potts, editors), pp. 15–62. San Francisco: Clavadel Press.

Clements, G. N. (1980) Vowel harmony in a nonlinear generative phonology: an autosegmental model, *Indiana University Linguistics Club.*

Dolphyne, F. A. (1988) *The Akan (Twi-Fante) Language.* Accra: Ghana Universities Press.

Halle, M. & Stevens, K. N. (1969) On the feature Advanced Tongue Root, *Quarterly Progress Report*, **94**, Research Laboratory of Electronics, Massachussetts Institute of Technology, 209–215.

Hardcastle, W. J. (1973) Some observations on the tense-lax distinction in initial stops in Korean, *Journal of Phonetics*, **1**, 263–272.

Hardcastle, W. J. (1976) *Physiology of speech production.* New York: Academic Press.

Harshman, R., Ladefoged, P. & Goldstein, L. (1977) Factor analysis of tongue shapes, *Journal of the Acoustical Society of America*, **62**, 693–707.

Hess, S. (1992) Assimilatory effects in a vowel harmony system: an acoustic analysis of advanced tongue root in Akan, *Journal of Phonetics*, **20**, 475–492.

Jackson, M. T. (1988) Phonetic theory and cross-linguistic variation in vowel articulation, *Working Papers in Phonetics*, **71**, University of California, Los Angeles.

Ladefoged, P. (1964) *A phonetic study of West African languages.* Cambridge: University Press.

Ladefoged, P., DeClerk, J., Lindau, M. & Papcun, G. (1972) An auditory-motor theory of speech production, *Working Papers in Phonetics*, **22**, University of California, Los Angeles, 48–75.

Lindau, M. (1975) [Features] for vowels, *Working Papers in Phonetics*, **30**, University of California, Los Angeles.

Lindau, M. (1979) The Feature Expanded, *Journal of Phonetics*, **7**, 163–176.

Perkell, J. (1971) Physiology of speech production: a preliminary study of two suggested revisions of the features specifying vowels, *Quarterly Progress Report*, **102**, Research Laboratory of Electronics, Massachussetts Institue of Technology, 123–139.

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

Rasband, W. (1990) NIH Image version 1.29q (Public domain computer program developed at the U.S. National Institutes of Health and available on the Internet at **http://rsb.info.nih.gov/nih-image/**).

Stewart, J. M. (1967) Tongue root position in Akan vowel harmony, *Phonetica*, **16**, 185–204.

Stewart, J. M. (1971) Niger-Congo, Kwa. In *Current trends in linguistics* (T. Sebeok, editor), pp. 179–212. The Hague: Mouton.