1991

A Study Of Vowel Articulation In A Perceptual Space.

Sungbok Lee
University of Alabama at Birmingham

Order Number 9134233

A study of vowel articulation in a perceptual space

Lee, Sungbok, Ph.D.

University of Alabama at Birmingham, 1991

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

A STUDY OF VOWEL ARTICULATION IN A PERCEPTUAL SPACE

by

SUNGBOK LEE

A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biomedical Engineering in the Graduate School, The University of Alabama at Birmingham

BIRMINGHAM, ALABAMA

1991

ABSTRACT OF DISSERTATION
GRADUATE SCHOOL, UNIVERSITY OF ALABAMA AT BIRMINGHAM

Degree Ph.D. Major Subject Biomedical Engineering

Name of Candidate Sungbok Lee

Title A STUDY OF VOWEL ARTICULATION IN A PERCEPTUAL SPACE

Vowel articulation was studied in a perceptually-oriented space with the aid of magnetic resonance imaging and a PARAFAC tongue model (Harshman et al., 1977). The study was based on the hypothesis that articulatory information contained in acoustic data can be more explicitly represented in a perceptually-oriented space than in the traditional formant space. The two-dimensional working space called a modified auditory-perceptual space (MAPS) was constructed based on the auditory-perceptual theory of Miller (1989). Perceptual vowel target zones of nine monophthongal American English vowels were constructed in MAPS. The vowel target zones of MAPS classified a vowel data base by Peterson and Barney (1952) with 89% accuracy. It was shown that the transformation of formant frequencies based on the formant-ratio theory elaborated by Miller can extract appropriate articulatory information contained in acoustic data. It redistributed the data points to align with the extracted articulatory dimensions, the location and degree of the tongue constriction and the lip opening area, in MAPS. This allows a more explicit phonetic description of vowel articulation in MAPS. It was also shown that MAPS is a valid space for the comparison of vowel systems across languages. Therefore, it was concluded that MAPS could be an ideal phonetic space. Based on the vocal tract shapes derived from the PARAFAC tongue model, acoustic stability introduced by the quantal theory of speech (Stevens, 1989) was studied in terms of the location and degree of the tongue constriction and the lip opening area. It was shown that acoustically stable regions in the vocal tract are not fixed but interactively determined by the combination of three articulatory parameters, which makes the notion of acoustic stability complicated. A question that remains is how the roles of articulatory parameters and acoustic stability can be integrated for the selection of places of articulation.

Abstract Approved by: Committee Chairman Martin J. McCutcheon

Program Director C. M. Stokely

Date May 22, 1991 Dean of Graduate School [signature]

PS-5744

ACKNOWLEDGEMENTS

I express sincere gratitude to the persons who offered support, advice, and their valuable time during the completion of my study at the University of Alabama at Birmingham. I especially thank my adviser, Dr. Martin J. McCutcheon, for his support, encouragement, and patience throughout my graduate study.

I thank my committee members, Dr. James E. Flege, Dr. A. V. Lakshminarayanan, Dr. Samuel G. Fletcher, Dr. Linda C. Lucas, and Dr. Ernest M. Stokely, for guidance and insight in all phases of this dissertation research. I cannot too warmly express my gratitude to Dr. Samuel G. Fletcher and Dr. Larry E. Adams who allowed me to work at the Department of Biocommunication at the University of Alabama at Birmingham. I would like to thank Steve Smith and Sherry Sutphin in the Biocommunication lab for their help and understanding.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF ABBREVIATIONS

APS	Auditory-perceptual space
DOC	Degree of tongue constriction
Fn	nth Formant frequency
LOC	Location of tongue constriction
MAPS	Modified auditory-perceptual space
MR	Magnetic resonance
PARAFAC	Parafactor analysis
RC	Rate of change
/IY/	Vowel [i] in the word "heed"
/IH/	Vowel [I] in the word "hid"
/EH/	Vowel [ɛ] in the word "head"
/AE/	Vowel [æ] in the word "had"
/AH/	Vowel [ɑ] in the word "hod"
/AW/	Vowel [ɔ] in the word "hawd"
/UH/	Vowel [ʌ] in the word "hud"
/UU/	Vowel [U] in the word "hood"
/UW/	Vowel [u] in the word "who'd"
/ER/	Vowel [ɝ] in the word "heard"

CHAPTER 1. INTRODUCTION

Vowels can be described in three different domains: articulatory, acoustic, and perceptual. In the articulatory domain, vowels have traditionally been represented by the static or quasi-static vocal tract shapes which are determined by the positions of speech articulators such as the tongue, jaw, lips, and velum. From the stand point of a speaker, these vocal tract shapes can be considered as the goal states (or targets) of intended vowel sounds. In the acoustic domain, vowels are acoustic waveforms (sounds). It has long been known that perceptually different vowel sounds have different positions of predominant peaks in the spectral envelope of the sounds. The peaks have traditionally been called formant frequencies or formants. Formants correspond to the resonance frequencies of the vocal tract. Among them, the first two or three formants have been known to be the most important acoustic cues for listeners to identify vowel sounds. In this context, vowel sounds have traditionally been represented as points in an acoustic space whose axes are the first two (F1/F2) or three (F1/F2/F3) formant frequencies. The formant space has long been used to represent acoustic and perceptual properties of the vowel sounds as well as for the phonetic description of vowel articulation in terms of "height" and "backness" of the tongue (e.g., Ladefoged, 1975). In the perceptual domain, vowels may be regarded as "perceived forms" of the acoustic signals through auditory-perceptual processes by listeners. For a convenient description of speech, they can be embodied as phonetic symbols. Such phonetic symbols may be regarded as perceptual vowel targets which should be commonly referred to by both speakers and listeners to be able to communicate with each other. In principle, the three different descriptions of a vowel can be unified through the concept of the vowel target. However, such an integration requires the existence of a unique relationship (or mapping) between the vowel targets in the three different domains because they are different facets of one object, the vowel.

A fundamental aspect of speech production is that there is no unique way to produce a vowel. Different speakers (or even a single speaker) may use different articulatory shapes or strategies to produce the same vowel as long as listeners permit such articulatory freedoms. Therefore, the resulting acoustic signals are highly variable among speakers and such acoustic variations make any attempt difficult to define a unique articulatory-acoustic or acoustic-perceptual relationship of a vowel. This situation is well illustrated by the classic work of Peterson and Barney (1952), which has shown that perceptually distinctive vowels produced by multiple speakers overlap in the formant space. The sources of acoustic variations causing the overlap have been known to be: (1) personal differences in sex, age, vocal tract anatomy, and articulatory habit or style, usually called inter-speaker differences; (2) coarticulation (or consonantal context effects) and the rate and stress of speech, called intra-speaker variations.

To eliminate the overlap in the formant space, several speaker or vowel normalization procedures have been proposed (Gerstman, 1968; Lobanov, 1970; Bladon et al., 1984; Syrdal and Gopal, 1986; Neary, 1989; Miller, 1989). Irrespective of techniques used, a common idea underlying the normalization procedures is to separate different vowels into unique and non-overlapping regions in a "perceptual (vowel) space" derived by a normalization procedure used. This dissertation is based on a hypothesis that articulatory aspects of vowels will be more accurately represented in a perceptual space than in the traditional formant space because acoustic variations caused by the inter-speaker differences may not exceed vowel-specific acoustic variations caused by inherent articulatory differences between vowels. We may further hypothesize that the more we eliminate the acoustic variations caused by inter-speaker differences from the acoustic signal, the better we may extract inherent articulatory properties of vowels. That may be what listeners do for the perception of a vowel produced by multiple speakers. In that sense, we agree with the view that speech perception take place in a specialized phonetic mode which is narrowly adapted for the efficient production and perception of phonetic structure, and the phonetic mode is not auditory but gestural (Liberman and Mattingly, 1985). The "motor theory" of speech perception seems to suggest a way to avoid dealing with the inherent complexity of speech signals by integrating both speech production and perception into a single specialized mode in the brain, although a question that remains is how the gestural information embedded in acoustic signals is decoded by the auditory system of humans. The motor theory also may suggest that what we have to seek in acoustic signals is not acoustic invariances but gestural information. In terms of the motor theory, this dissertation is the description of a procedure to extract articulatory information embedded in acoustic data. We believe that such a procedure may provide a more natural way to deal with the complexity of acoustic signals caused by the inter- and intra-speaker variations in speech production.

Along the line of thinking mentioned above, the major purpose of this dissertation is to develop a phonetic (or perceptual) space in which the articulatory and perceptual descriptions of vowels can be integrated. Such an integration will be desirable for the study of phonetic description of vowel articulations. It is also an attempt to find a more unique relationship between vowel articulations and their acoustic properties.

In chapter 2, as a preliminary step, acoustic performances of three midsagittal width-to-cross-sectional area conversion algorithms are compared using a tube model of the vocal tract constructed from midsagittal vocal tract images acquired by magnetic resonance (MR) imaging technique. The purpose of the comparison study is to assay the accuracy of currently available conversion algorithms and to select the most appropriate one for the purpose of this dissertation. Also, the detailed procedure used to obtain the vocal tract images utilizing the MR imaging technique is described in this chapter because the visualization of the vocal tract during speech production is basic in the study of speech production and perception. In chapter 3, vowel articulation is studied in a two-dimensional modified auditor-perceptual space (MAPS) based on the tongue shapes generated by the PARAFAC tongue model (Harshman et al., 1977). The purpose of the study is to provide a phonetic space in which articulatory and perceptual properties of vowels are better represented than in the conventional formant space. It is also shown that MAPS is a useful space for the comparison of vowel systems between languages. In chapter 4, the acoustic stability, which is a key concept of the quantal theory of speech (Stevens, 1989), is discussed in terms of three articulatory parameters, the location and degree of the tongue constriction and the lip opening area, extracted from the simulated vocal tract shapes of vowels. Finally, in chapter 5, the results of this dissertation research are summarized.

CHAPTER 2. COMPARISON OF AREA CONVERSION METHODS USING MAGNETIC RESONANCE (MR) IMAGING

I. Introduction

Perceptual identity of the sounds of speech are basically determined by the vocal tract shapes at the moment of sound production. Therefore, images of the vocal tract during speech production are necessary to fully understand the articulatory-acoustic relationships and their perceptual implications in speech. Data on the vocal tract shape and dimension are also essential for the development and testing of theories or hypotheses regarding speech production (e.g., Lindblom, 1986; Stevens, 1989) as well as for quantitative modeling of articulatory movements (e.g., Mermelstein, 1973; Coker, 1976). Because speech production is a dynamic process, rapid, accurate, and non-invasive imaging of the vocal tract is required to obtain the necessary data.

Lateral cine x-ray projection techniques have traditionally been used to acquire data on the vocal tract in the midsagittal plane (Chiba and Kajiyama, 1941; Fant, 1960; Heinz and Stevens, 1964; Perkell, 1969; Harshman et al., 1977). However, due to the inherent limitations imposed by the radiation dosage limit and x-ray projection geometry, only a limited amount of shape information is available using this method. Also, the superimposition of medial and lateral bone and soft tissue structures in radiographic images obscures the detailed anatomy of important articulators and makes the quantitative measurements of vocal tract dimensions difficult. Computed tomography (Kiritani et al., 1977; Johansson et al., 1983) seems to be an effective tool for obtaining cross-sectional data on the vocal tract. However, the limitation in radiation dosage and the restricted maneuverability of subject positions with respect to the imaging plane of the system makes the full utilization of this technique difficult. Its application in speech is currently limited to the imaging of only a few cross-sections of the pharyngeal region. The x-ray microbeam system (Fujimura et al., 1973; Stone, 1990) examines the movements of speech articulators by tracking gold pellets attached to selected points on the articulators using on-line computer control of the focused x-ray beam deflection. The microbeam technique minimizes the inherent radiation hazard of x-ray, and it is a good tool to study movements of the tongue, lips and jaws in the midsagittal plane. However, it does not provide detailed shapes of the entire vocal tract needed to fully investigate articulatory-acoustic correlations of such movements.

Recently, several methods which do not use ionizing radiation have been introduced for the study of the vocal tract shape and articulatory movements. The magnetometer (Sonoda, 1974) is a point-tracking method which can be applied to the study of tongue movements in the oral cavity, but it does not provide information about tongue shape. Opto-electronic instrumentation has also been employed for study of tongue movement (Chuang and Wang, 1975; Flege et al., 1986). However, this method provides information only pertaining to the midsagittal tongue-palate distances in the palatal region. Ultrasound has also been employed to obtain information on the tongue configuration (Stone et al., 1988; Stone, 1990), but its application has been limited to the study of the shape of the surface of the tongue or its movement. The ultrasound technique can not provide information about tongue-palate distances because of the air gap in the vocal tract cavity.

MR imaging (Pykett et al., 1982) appears to be free of many of the limits and disadvantages associated with the above-mentioned methods. The MR imaging system is built around a strong magnet with a large round tunnel in which subjects are positioned. Inside the magnet are gradient coils that function as radiofrequency antennas to transmit and detect radiofrequency signals needed to produce MR images. MR images are obtained by manipulating the protons (mostly in water molecules) of the human body without using any ionizing radiation. When MR images are displayed for viewing, intense signals are displayed as white, weak ones as black, and intermediate signals appear as shades of gray. The intensity of the signal is related to the density of protons within a tissue being imaged. MR imaging is unique in that it detects soft tissues because of their high density of protons, rather than the calcified structures (bones) detected by x-ray.

Since the vocal tract is composed of soft tissues such as the tongue, lips, velum, hard and soft palates, and pharynx, such organs can be detected by MR imaging. In addition, various cross-sectional (e.g., sagittal, coronal, or transaxial) images can be obtained without altering subject's position. Also, the strong static or gradient magnetic field and radiofrequency signals used for MR images have no known hazardous effects on humans. It has already been demonstrated that MR imaging could be a useful tool in visualizing the three dimensional shape of the vocal tract (Rokkaku et al., 1986), and in estimating the cross-sectional area in the pharyngeal region (Baer et al., 1987). Lakshminarayanan et al. (1989) have also shown that the midsagittal section of entire vocal tract during the vowel or consonant production can be clearly visualized with the MR imaging. Also, the accuracy of MR measurements of the midsagittal tongue-palate distances in the oral cavity have been tested using acrylic tongue-palate spacers to establish a known tongue configuration and location during MR imaging, and the overall measurement errors were found to be less than 1 mm (McCutcheon et al., 1990).

The MR imaging offers several advantages over conventional radiographic techniques in vocal tract imaging. There are nevertheless several limitations in the current MR imaging method. MR imaging is time consuming and the combined time-space resolution of MR imaging is currently insufficient to acquire dynamic images of the vocal tract. The teeth contain few protons and can not be imaged. Thus, without appropriate pre-preparation, we may lose an important landmark in the vocal tract. It is possible to record utterances simultaneously during MR imaging. However, the background (gradient) noise induced by the rapid switching of the magnetic fields during imaging interferes with the utterances. A simple experiment with the MR imaging system used in the current study revealed that the third formant and bandwidth of all formants were affected by the gradient noise and resonances of the tunnel as illustrated in Fig. 1. The vowel used in the experiment was /IY/ produced by an adult male speaker. Although there currently exist several limitations in the use of MR imaging for the study of the vocal tract during speech production as mentioned here, it seems likely that many of these limitations will be overcome in the near future.

Since work by Fant (1960) on the acoustic theory of speech production in which the vocal tract has been viewed as an acoustic tube of varying cross-sectional areas and implemented as an electric transmission line analog, the conversion of midsagittal widths to cross-sectional areas has been a great concern in the literature because of its importance in a construction of the tube model of vocal tract for the study of articulatory-acoustic relationship. Midsagittal

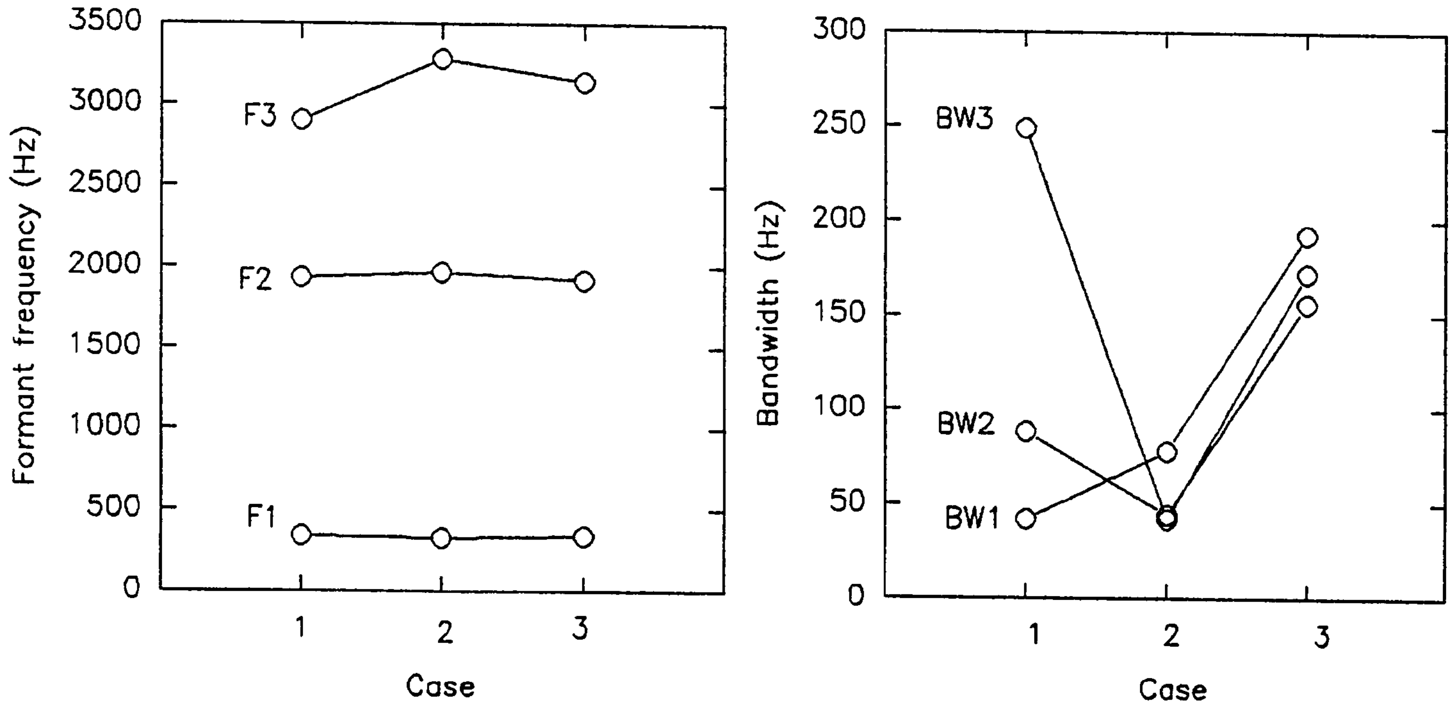


Fig. 1. Environmental effect of MR imaging system on the acoustic properties of the vowel /IY/. The vowel was recorded in a sound-proof room (case 1), in the tunnel without gradient noise (case 2), and with gradient noise (case 3), respectively.

widths can be measured from x-ray profiles of the vocal tract. However, the corresponding cross-sectional area is difficult to estimate. It requires a transformation from a one dimensional quantity to a two-dimensional one which is irregular in shape. Both the cross-sectional dimensions and the volume of the vocal tract are necessary to establish a reliable relation between the midsagittal widths and cross-sectional area. However, this requires cross-sectional and three dimensional data on the vocal tract, which are limited in quantity due to the limited capability of currently available imaging methods.

Despite the lack of necessary data, some conversion algorithms have been introduced by several authors (Heinz and Stevens, 1964; Ladefoged et al., 1971; Lindblom and Sundberg, 1971; Mermelstein, 1973; Rubin et al., 1981). They have been shown to be reasonably successful for the estimation of the acoustic properties of the vocal tract. However, the applicability of such algorithms to arbitrary speakers from various native language backgrounds is uncertain because most of them have been derived from midsagittal images of the vocal tracts of a limited number of speakers and sounds drawn from different languages. Therefore, it may be desirable to test and compare the acoustic performances of currently available algorithms with other speakers. The comparison study may illuminate some problems embedded in the algorithms. Another aim of the comparison study is to select the most accurate conversion algorithm for the studies described in chapters 3 and 4. These studies are based on the computation of formants of given midsagittal vocal tract shapes using a tube model of the vocal tract. The validity of computed formants, and thus the validity of these studies, is largely dependent on the accuracy of width-to-area conversion.

For the purposes mentioned above, midsagittal MR images of the vocal tract were acquired during the sustained production of five American English vowels by two adult male subjects. Midsagittal widths and the length of the vocal tract were measured from the MR images. Corresponding cross-sectional areas were calculated by employing three different width-to-area conversion methods. The first three formant frequencies were computed using a tube model of the vocal tract. The computed formant frequencies were compared with acoustically measured values from the subjects' utterances. Also, because the use of MR imaging for the acquisition of the vocal tract data is relatively new and promises to be a good alternative to the traditional radiographic methods, detailed methods and outcomes are described in this chapter.

II. Methods

A. Acquisition of vocal tract data using MR imaging

1. Subjects and speech materials

The MR images were collected from two volunteer adult male subjects, MJM and LSB, with no known speech or hearing disorders. MJM is a native speaker of American English and LSB is a native speaker of Korean. Neither subject had received any formal phonetic training in the production of American English vowels.

The vowels examined were /IY/, /EH/, /AE/, /AH/, and /UW/ contained in the monosyllabic words "heed", "head", "had", "hod", "who'd", respectively. Before the MR session, the subjects practiced uttering the words and sustaining each vowel sound for the required duration.

2. MR images acquisition

MR scans were performed on a 0.5T Picker Vista MR 2055 system. Each subject was positioned supine in the standard head-coil. The subject's head was centered using laser positioning devices on the scanner to ensure proper midsagittal section imaging. In addition, transaxial pilot images were obtained to enable accurate positioning of the section. Calibration imaging with a phantom of known dimensions was also performed before each imaging session using the same parameters employed in the actual imaging. The size measurements from the test image were used to calibrate the output of the gradient drives to ensure the accuracy of the distance measurements. The onset of acoustic noise generated by the gradient coils during the calibration phase of each scan was used to cue the subject to begin an utterance.

Initially, a spin-echo sequence with short repetition time (TR) of 83 ms and short echo time (TE) of 20 ms was used. Data were acquired using 256 phase-encoding steps and 512 samples (2 x oversampling) to cover the Fourier space. The scan field of view was 25 cm and the image was reconstructed on a 256 x 256 matrix yielding a pixel resolution of 1mm. The total imaging time was 20 seconds.

The imaging strategy just described was subsequently abandoned for two reasons. It was necessary to improve the signal-to-noise ratio considerably for the images to be useful for the vocal tract measurements. For this purpose, a special coil was designed and constructed. The geometry of the coil is shown in Fig. 2. In use, the coil was wrapped around the lower part of the face as shown in Fig. 3. The Q of the coil-preamplifier combination was measured to be 150 at the operating frequency of 21.25 MHz. Also, it was difficult for subjects to maintain a steady state of the vocal tract for the required 20-second interval. Therefore, the use of rapid gradient-echo sequences was investigated.

Both Fourier acquired steady state (FAST) (Gyngell et al., 1986) and variable flip angle, gradient-spoiled sequences similar to FLASH (fast low-angle shot) (Haase et al., 1985) were employed. The flip angle was optimized by observing the signal intensity as a function of flip angle and choosing the angle corresponding to the maximum signal intensity. An echo time (TE) of 10 ms and a repetition time (TR) of 33 ms gave an imaging time of only 4 seconds for a 128-sample image covering half the Fourier plane. The imaging time of 4 seconds was short enough to be comfortable for sustaining a vowel sound and sufficiently long to obtain good-quality images.

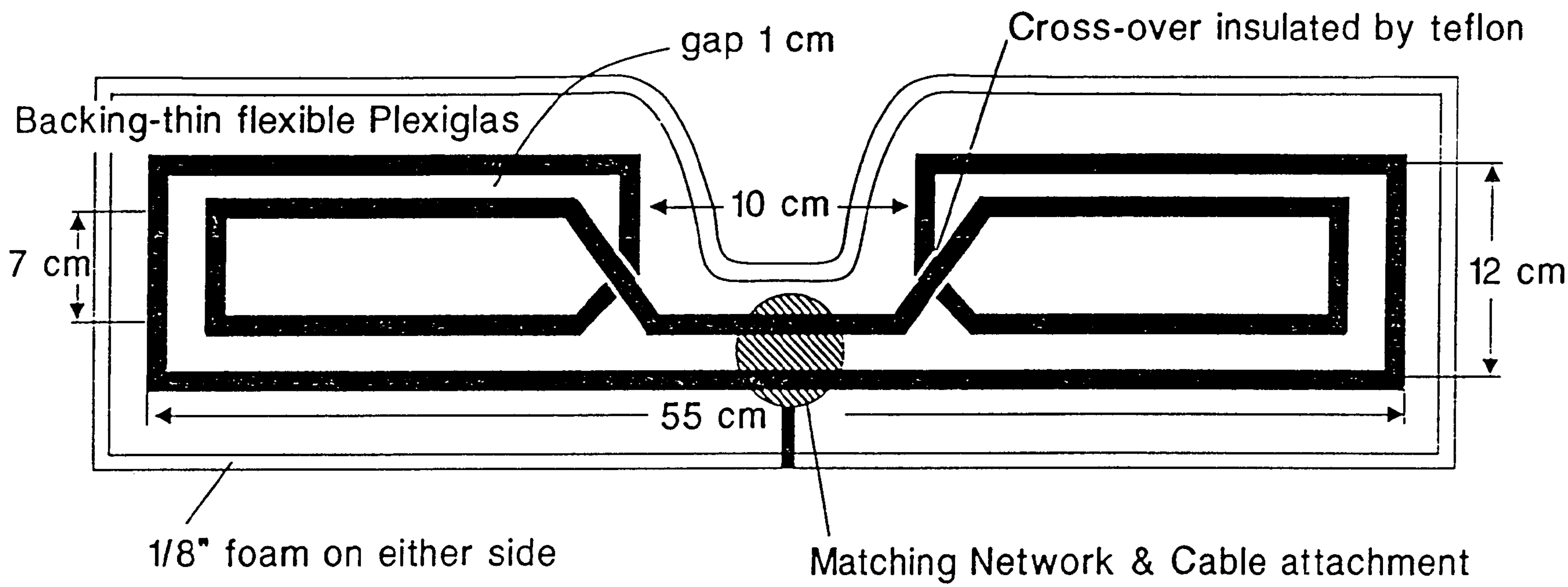


Fig. 2. Schematic diagram of the special MR receive-only coil. The two-turn double-loop coil is made from 0.5 inch (1.3 cm)-wide copper tape (CCH-XX-101-0050; Chometrics). The speech coil is wrapped around a subject's face to form a cylinder and matched to its preamplifier by balanced series capacitor of 300 pF. The cable is a 100-ohm Belden 9207, with an appropriate Lemo connector.

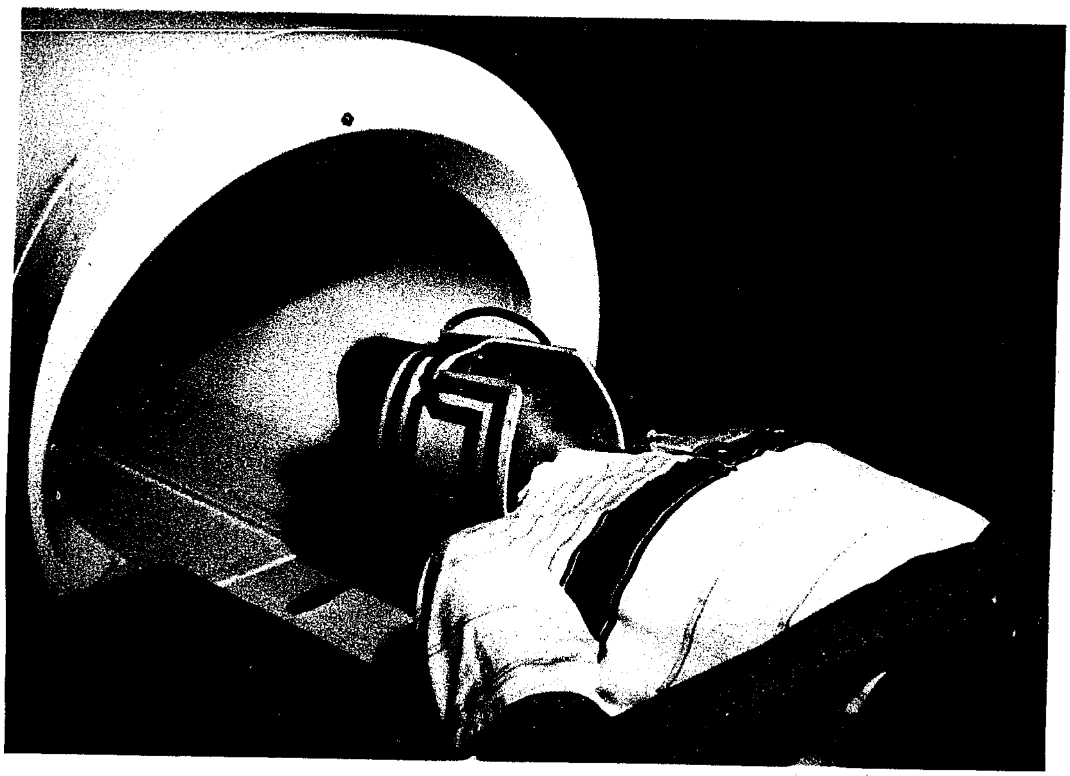


Fig. 3. Photograph of a subject with the speech coil wrapped around the lower part of the face for imaging.

3. Measurement of MR images

Vocal tract contours were manually traced three times from each hard copy (the original film of the image). The difference between tracings was close to the pixel resolution of the MR images (1mm). A master copy was made by eye examination of overlapped tracings. Each master copy was magnified by the factor of two and then used for the measurements of vocal tract dimensions.

To consistently measure the vocal tract dimensions of vowels spoken by different subjects, it is convenient to use a grid or coordinate system (e.g., Heinz and Stevens, 1964; Ladefoged et al., 1971; Lindblom and Sundberg, 1971; Mermelstein, 1977; Engstrand, 1988). Irrespective of the grid system used, the first step in vocal tract measurements is to determine the origin and a midline (or center line) of the vocal tract along which it is assumed that sound waves propagate from the glottis to the lips. The midline is necessary to realize a tube model of the vocal tract in which cross-sections along the vocal tract are defined to be normal to the midline.

Because the junctions between the cervical vertebrae and between the hard and the soft palates are clearly observable in the MR images, they were used to determine the origin of the grid system. A reference line (A-B) was drawn to be approximately aligned with the rear wall of the pharynx through the junctions of the first and second and the second and third cervical vertebrae. The y-axis (C-D) was drawn from the junction of the hard and the soft palates to be parallel to the reference line A-B. The x-axis (E-F) was drawn to be normal to the y-axis from the center position between the junctions of the first and second and the second and third cervical vertebrae. The intersection of the x and y axes was taken as the origin of the grid system. A half-circle was drawn from this origin. Its circumference included the roof of the mouth. Straight grid-lines were drawn from the origin in the polar region and from the y-axis in the rectangular region. The angle between radial grid-lines was 5^0 and the distance between the horizontal grid-lines was .39 cm. The angle from the x-axis of the first four grid-lines was 90^0 and the distance between the vertical lines was also .39 cm. The grid system is shown in Fig. 4.

Center positions of the grid-line segments delimited by the superior-posterior and inferior-anterior contours of the vocal tract were determined using a scale. A midline of the vocal tract was then composed of line segments which were formed by connecting the center points of adjacent grid-line segments from the glottis to the lips. Length of the midline was taken as the length of the vocal tract. Since the midline of the vocal tract is affected by the tongue and soft palate contours, larynx-lowering, and the lip protrusion, the vocal tract length was different across the vowel sounds.

As mentioned earlier, midsagittal widths should be determined along the lines which are normal to a midline of the vocal tract. To follow the guideline, midsagittal widths

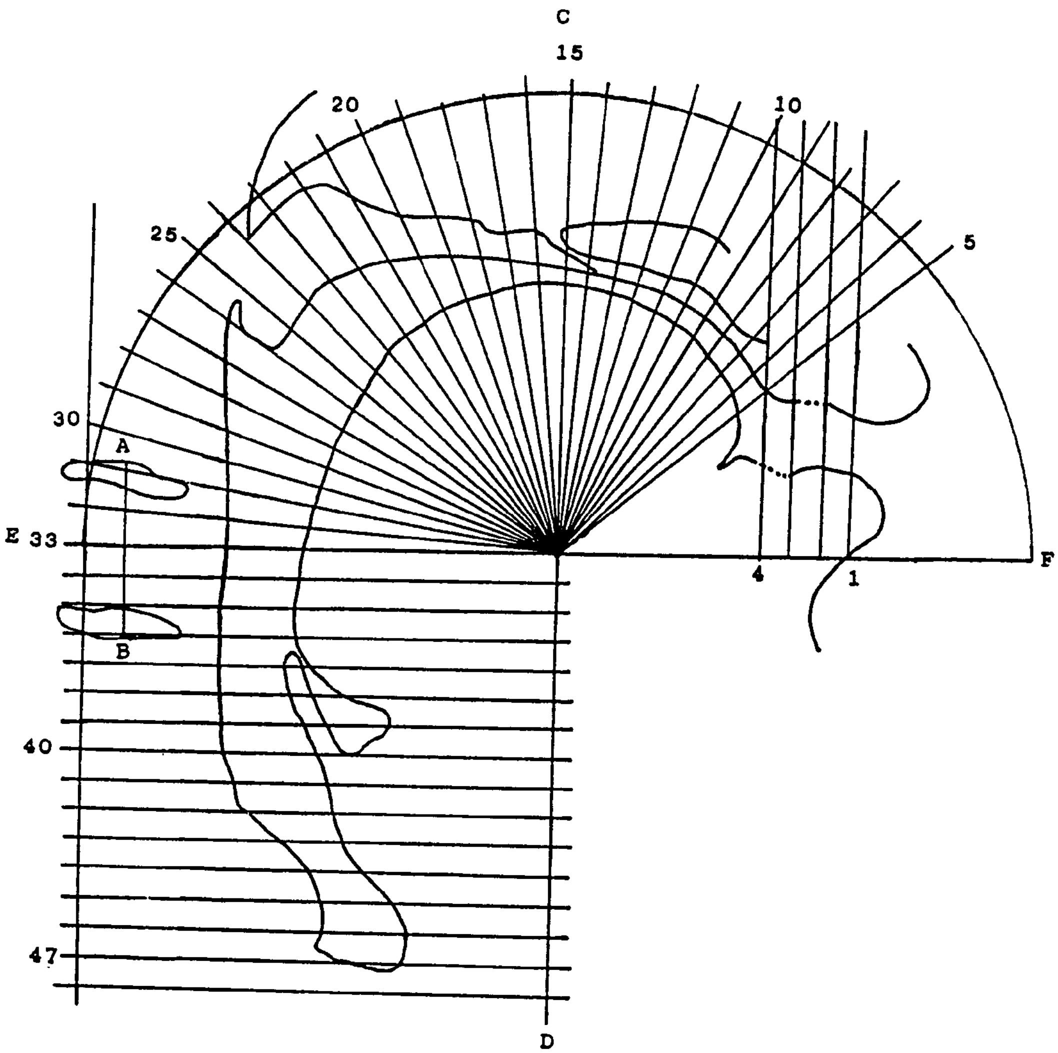


Fig. 4. Grid system used to measure the vocal tract dimensions. The illustrated vocal tract shape is of /IY/ produced by subject LSB.

were measured by the following procedures depicted in Fig. 5. A reference line R1 was drawn by connecting center positions of (n-1)th and (n+1)th grid-line segments. Another reference line R2 was drawn to be normal to R1 at the center position of the nth grid-line segment. The midsagittal width at the center position of the nth grid-line segment was then measured along R2. These procedures were based on an assumption that the direction of R2 along which the midsagittal width is measured is normal to the midline of the vocal tract. This assumption can be validated if the distances between sampling positions are small enough (They were about 0.35 cm apart in the current study). It should be noted that if the midsagittal widths are simply measured along the grid-lines, the result may be erroneous in terms of the tube model. The measurements would thus yield improper acoustic outputs if they are used to estimate acoustic properties of the vocal tract.

Vocal tract lengths and midsagittal widths measured from the MR images of two subjects are listed in Appendix A. The resulting midsagittal widths were converted to corresponding cross-sectional areas using the three different conversion algorithms by Ladefoged (1988), Rubin et al. (1981), and Lindblom and Sundberg (1971) by the procedures described in the next section.

B. Application of area conversion algorithms to MR data

The 17-section (from the glottis to around the alveolar ridge) area conversion table constructed by Ladefoged (1988)

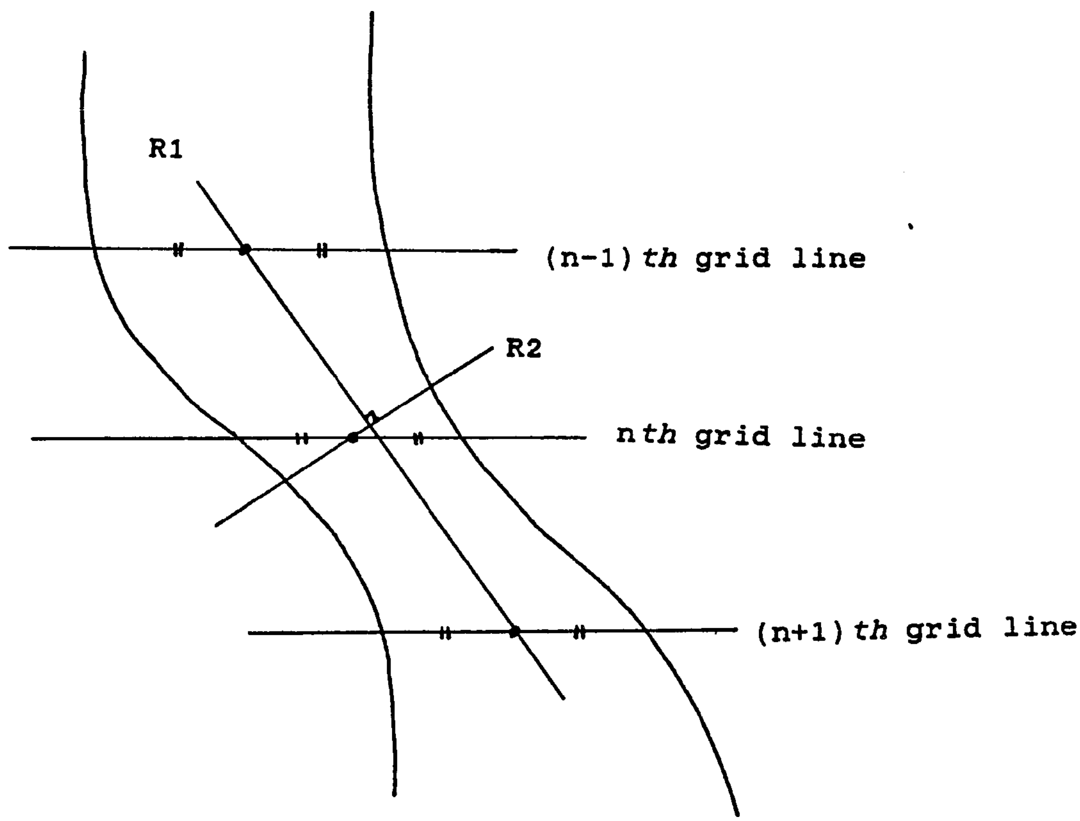


Fig. 5. Schematic representation of a method used to measure midsagittal widths. The midsagittal width at the center position of the nth grid line is measured along the reference line R2 which is normal to the reference line R1 connecting the center positions of the (n-1)th and (n+1)th grid lines.

is listed in Appendix B. The table was derived from a study by Ladefoged et al. (1971) in which casts of the oral and the pharyngeal region of a single subject were made. There are 17 rows in the table and there are three lines in each row. The numbers in each row represent cross-sectional area with the midsagittal widths varied from 0.1 cm to 3.0 cm. Each row corresponds to the region in the vocal tract referred to by the reference lines shown in Fig. 6.

In the current study, the original table was expanded to a 34-section table by a linear interpolation between rows. The 17-section approximation of the vocal tract may be reasonable for the comparison of the tongue shapes, but it is somewhat crude for the estimation of acoustic outputs of the vocal tract. Another reason for the expansion was to take advantage of the large number of midsagittal widths (42-50 sections) measured in the current study. To use the expanded area conversion table, traced vocal tracts were converted into a 35-section (including the lips) tube using a cubic spline interpolation (Press et al. 1986).

The conversion method used by Rubin et al. (1981) was a summary of previously published data. Based on a study by Heinz and Stevens (1964), the cross-sectional area in the pharyngeal region was approximated as an ellipse with a midsagittal width (w) as one axis and the other increasing from 1.5 to 3 cm as one moves upward from the larynx tube to the velopharynx. For the oral region, the cast measurement data by Ladefoged et al. (1971) was approximated as follows:

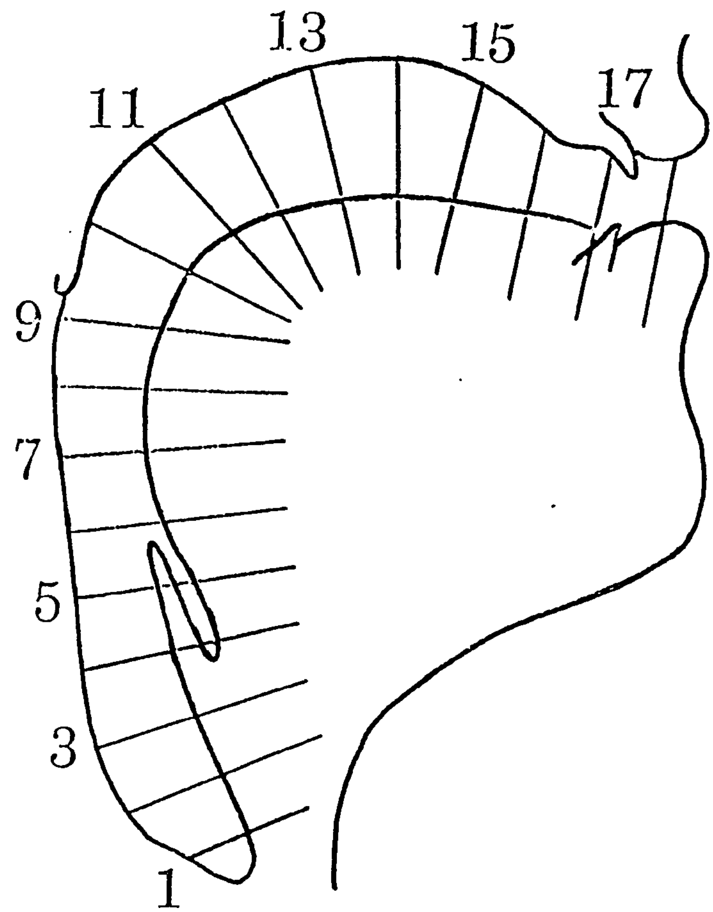


Fig. 6. Grid system used by Ladefoged et al. (1971) to measure midsagittal widths and cross-sectional area. Each reference line corresponds to each row number in the area conversion table listed in appendix B (from Ladefoged et al., 1971).

In the soft palate region the area was taken as $2w^{1.5}$, in the hard palate region as $1.6w^{1.5}$, and between the alveolar ridge and incisors as 1.5w for $w<0.5$, $0.75+3(w-0.5)$ for $0.5<w<2$, and $5.25+5(w-2)$ for $w>2$. In the labial region the area was assumed to be elliptical and the distance between the corners of the mouth was given by $2+1.5(s_l-p_l)$, where p_l is the lip protrusion and s_l is the vertical lip separation (Mermelstein et al., 1971). This formula was not used in the present study because the lip protrusion length p_l could not be measured due to the invisibility of the teeth in the MR images. The conversion method was applied to our two subjects as follows (see Fig. 4 to refer to the positions of grid lines): Grid lines 31-47 (subject MJM) and 27-46 (subject LSB) were regarded as the pharyngeal region; Grid lines 16-30 (MJM) and 15-26 (LSB) as the soft palate region; Grid lines 6-15 (MJM) and 5-14 (LSB) as the hard palate region; Grid lines 3-5 (MJM) and 3-4 (LSB) as the region around the alveoli; and Grid lines 1-2 as the labial region for both subjects.

The conversion method used by Lindblom and Sundberg (1971) was based on a power function approximation between area and midsagittal width. To derive areas in the pharyngeal region, these authors used tomographic data obtained from a Swedish subject examined earlier by Fant (1964). In the lower pharyngeal region (from the upper point of epiglottis to the bottom part of the pharynx) the area was taken as $1.1w^{2.21}$. In the upper pharyngeal region (i.e., from the uvula down to the upper point of epiglottis) the area was taken as $1.68w^{1.9}$. In the oral region the area was given as $2.2w^{1.38}$, based on a study by Heinz and Stevens (1964). This conversion method was applied as follows: Grid lines 36-47 (MJM) and 36-46 (LSB) as the lower pharyngeal region; Grid lines 31-35 (MJM) and 27-35 (LSB) as the upper pharyngeal region; Grid lines 3-30 (MJM) and 3-26 (LSB) as the oral region; and Grid lines 1-2 as the labial region for both subjects.

Because the area conversion for the labial region was not given or not applicable, the lip opening area was estimated from the lip opening height H (i.e., vertical separation of the lips) which were measured from MR images as the midsagittal width at the center position of the first grid-line segment in Fig. 4. To find a probable relation between the lip opening area (A) and the lip opening height (H), a lip data base collected from eight American subjects (Linker, 1982) during the production of nine American English vowels was analyzed using a linear regression method. The results are shown in Fig. 7. The best predictor of A was found to be the product of H and W (r=.98) where W is the lip opening width (i.e., horizontal distance between the corners of the mouth). The slope of the regression line was 0.80. A similar result was reported by Fromkin (1964) with a value of 0.7. Also reported by Fromkin, our analysis showed that the area of lip opening (A) can be approximated by the area of an ellipse whose major and minor axes are W and H.

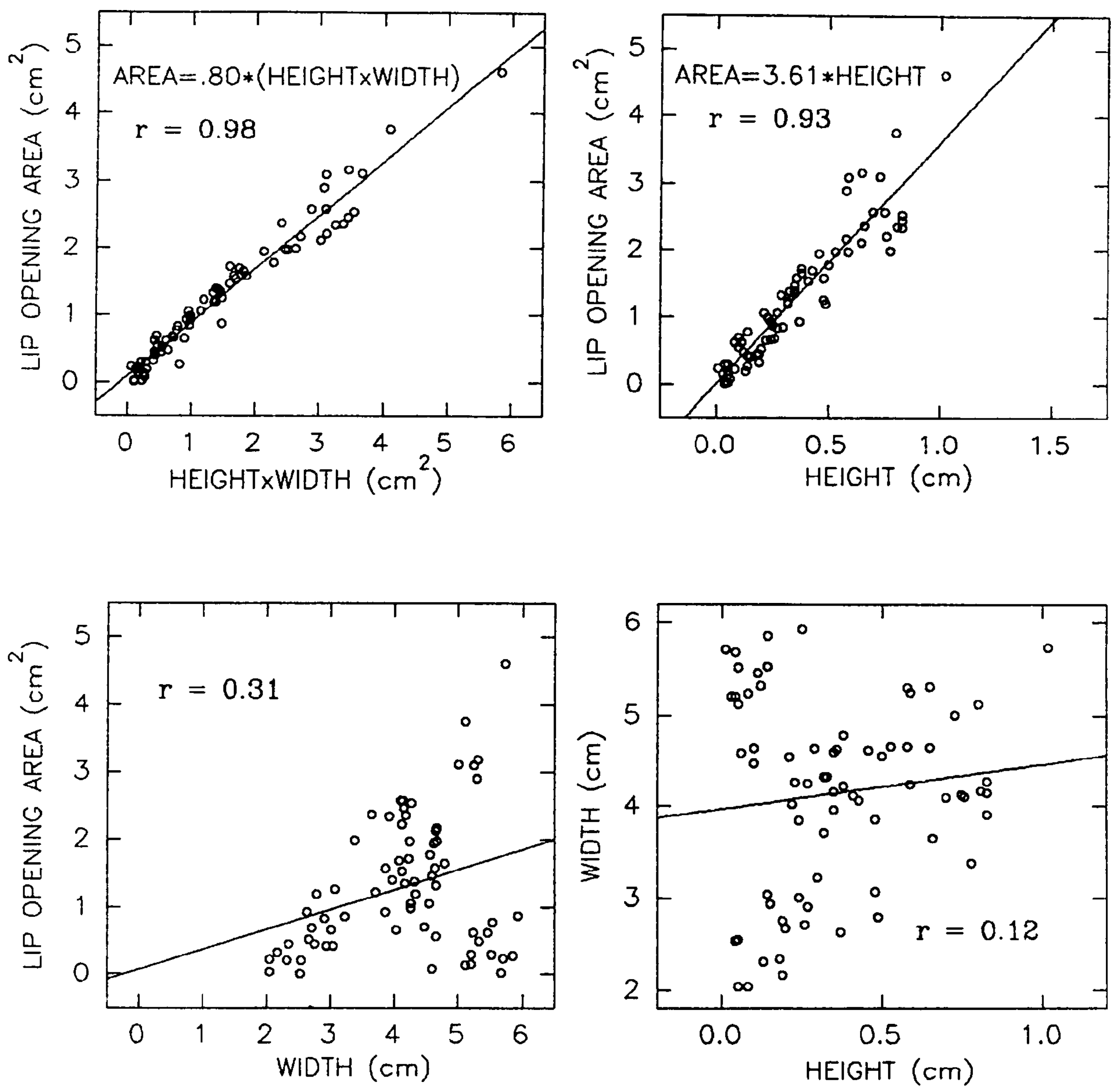


Fig. 7. Result of a regression analysis between the lip parameters (r: correlation coefficient). It shows that the lip height (H) alone is a good predictor of the lip opening area (top right).

However, this does not mean that the lip opening shape itself is elliptical. The regression analysis also revealed that the lip opening height (H) alone is a reasonable predictor of A (r=0.93) across vowels and talkers. The slope of regression line was 3.61. This value was used to estimate A from the lip height measurements obtained from the MR images using the equation

$$A = 3.61*H. \qquad (1)$$

It is interesting to observe that W is poorly correlated both with A (r=0.31) and with H (r=0.12). This suggests that W is not an important parameter for the determination of the lip opening area among American English speaker talkers, although it enhances the correlation between A an H. This may imply that W can be replaced with another lip parameter such as the lip protrusion. Also, it is uncertain what kind of constraints make such a strong correlation between A and one-dimensional quantity H possible. Further study may be needed.

By including the above relation for the determination of the lip opening area, it was possible to determine the area function of the entire vocal tract using only midsagittal dimensions. The resulting cross-sectional areas were used to estimate formant frequencies using a tube model of the vocal tract described in the next section.

C. Computation of formants using a tube model

1. Tube model of the vocal tract

In order to analyze the acoustical performance of the vocal tract it is convenient to treat the vocal tract as an acoustic tube of continuously varying cross-sectional areas with a straight axis[1] (Fant, 1960; Flanagan, 1972). The cross-sectional area at each point of the tube is defined as a function of the distance from the glottis, which is usually called an area function. The tube is excited at the glottal end of the straightened tube, and sound pressure is radiated at the opposite end (i.e., the mouth opening). The entire tube is usually approximated as a concatenation of a large number of uniform cylindrical tubes.

Because the average length of the vocal tract (about 17 cm in men) is comparable to the wavelength of sound in the frequency range interest[2], and because the cross-sectional dimensions of the vocal tract are usually small compared to the wavelength, we can assume that only plane waves propagate in the tube along its length dimension. With this assumption of one-dimensional wave propagation in the vocal tract, we can apply a simple boundary condition that pressure and volume velocity are continuous across the boundaries between adjacent uniform tubes. This greatly facilitates mathematical details of sound wave propagation in the tube.

---

[1] Sondhi (1986) has shown that the shift in the resonance frequencies of the bent vocal tract of a lossless uniform rectangular cross-section from those of the straight vocal tract is in the range of 2%-8% for frequency below 4KHz for typical dimensions of the vocal tract.

[2] It is usually up to 4 KHz. For frequencies above 4KHz, the auditory resolution, expressed in terms of the width of the critical bands (Zwicker and Terhardt, 1980), is poor. Consequently, listeners are relatively insensitive to the detailed spacing of spectral peaks at these high frequencies.

The propagation of sound along the tube can then be described by an electrical transmission line analog (Dunn, 1950; Fant, 1960; Flanagan, 1972), partial differential equations with appropriate boundary conditions (Sondhi, 1974; Portnoff, 1973; McGowan, 1987), or digital networks (Kelly and Lochbaum, 1962; Maeda, 1982; Meyer et al., 1989). The transmission line analog of the vocal tract provides a relatively simple but reasonable way to estimate the acoustic properties of the static vocal tract. The differential equation approach has theoretical interest for the solution of the wave equation of the time-varying vocal tract. The digital network representation is ideal for the hardware implementation of real-time speech synthesis.

Of these approaches, the electrical transmission line analog was selected for the present study. In the transmission line analog of the vocal tract, each uniform cylindrical tube can be represented by an electrical four-pole network. This analogy comes from the fact that sound pressure and volume velocity for plane wave propagation in the uniform tube satisfy the same wave equation as do voltage and current on the uniform transmission line (e.g., Pipes and Hovanessian, 1969). The transmission (or ABCD) matrix of the entire tube can then be represented by a product of such matrices of each uniform tube. Also, it has been well known that energy losses during sound wave propagation along the vocal tract affect the formant frequencies and bandwidth of sounds (Fant, 1960; Flanagan, 1972). To account for the attenuation effects, five sources of energy loss are usually considered: (1) viscous loss in a boundary layer at the inner surface of the tube; (2) loss due to heat conduction through the vocal tract walls; (3) loss due to radiation impedance at the mouth opening; (4) loss due to glottal impedance; (5) loss due to yielding of the vocal tract walls.

In the present application, transmission line analog of each uniform tube consisted of three parts: a lossless uniform section, a frequency-dependent series resistance R_a to represent the viscous losses, and a frequency-dependent shunt conductance G_a to represent the heat conduction losses. Following a suggestion by Badin and Fant (1984), the loss due to yielding walls Y_w was inserted at two locations in the vocal tract, near the root of epiglottis and near the mouth, after being weighted by a factor $(6/A_m)^{0.5}$ where A_m is the mean area over a 4 cm part of the area function centered around the insertion locations. It has been shown that this prevents large F1 shift (lowering) for back vowels, especially for /AH/ (Badin and Fant, 1984). The last section (the lip opening) includes a terminating impedance Z_r to represent the radiation load at the lips. In this study, Z_r was represented by the electrical network suggested by Stevens et al. (1953). The glottal impedance Z_g was inserted at the first section of the tube (i.e., the glottal end) using a value given by Wakita and Fant (1978). A transmission line analog of the vocal tract implemented in the current study is schematized in Fig. 8.

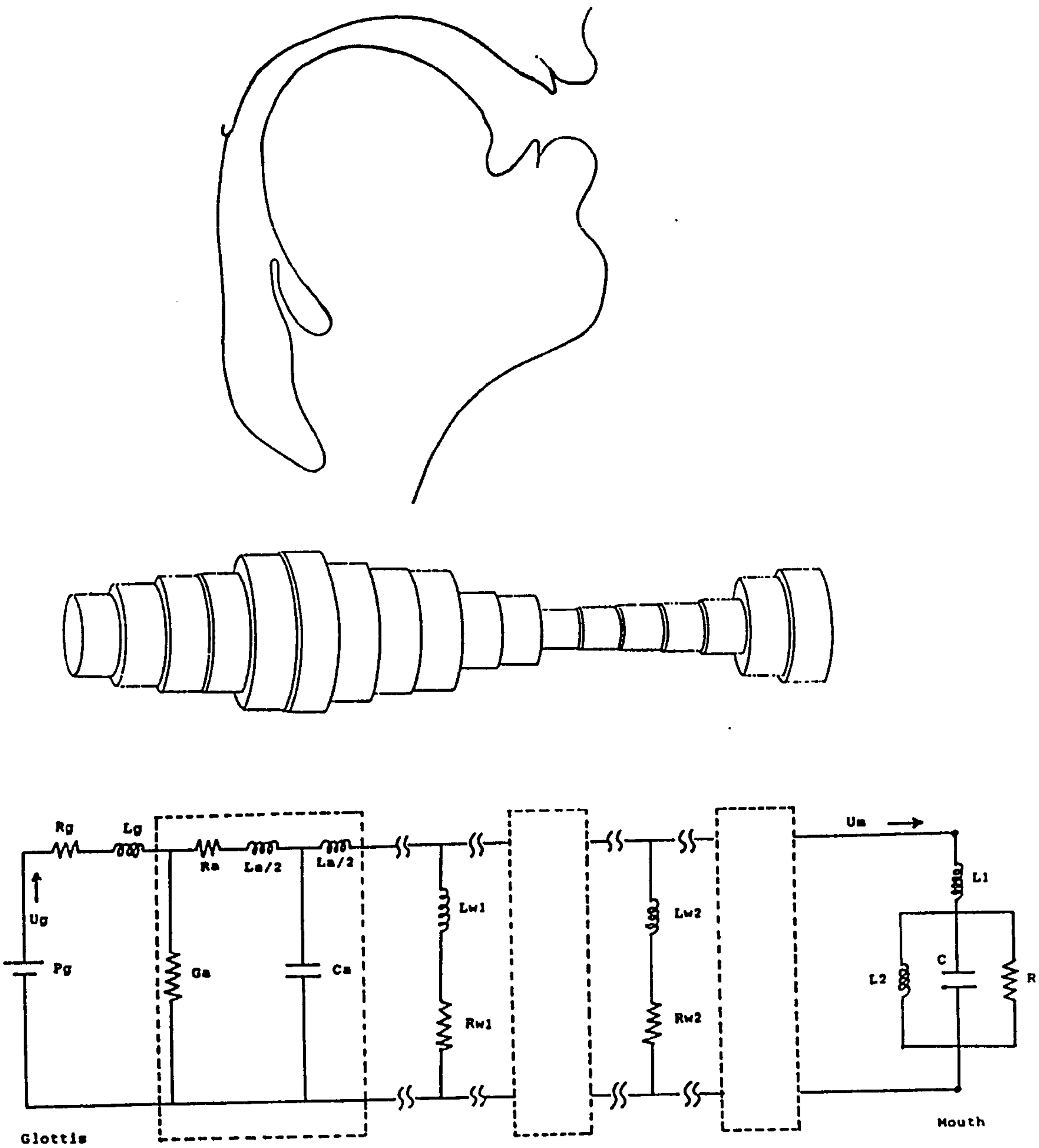


Fig. 8. From the top to bottom, a midsagittal vocal tract and corresponding tube model (from Ladefoged, 1971) and a transmission line analog of the vocal tract implemented in the current study (see the text).

2. Computation of formant frequencies

For voiced sounds like vowels, the transfer function F(s) is defined as the ratio of the volume velocity at the lips to the volume velocity at the glottis (Atal et al., 1978). The transfer function F(s) can be expressed as

$$F(s) = F_z(s) \prod_{K=1}^{\infty} [s_k s^*_k/((s-s_k)(s-s^*_k))], \tag{2}$$

where $s_k = \sigma_k + j\omega_k$ is the complex frequency of the kth pole, s^*_k is the complex conjugate of s_k, and $F_z(s)$ contains zeros of F(s) which are not important for the computation of formant frequencies. From each pole s_k, the formant frequency and bandwidth can be calculated by the relation

$$F_k = \omega_k/2\pi$$

and

$$B_k = -(\sigma_k/\pi). \tag{3}$$

Therefore, in terms of formant frequencies and bandwidths, the acoustic characteristics of the vocal tract can be represented by the pole frequencies of the transfer function.

The numerical computations of formants were carried out in two steps: (1) Calculation of the transfer function F(s) at the complex frequency s for a given vocal tract configuration specified in terms of cross-sectional areas and lengths of each uniform section, and (2) Determination of formants and their bandwidth by finding zeros of 1/F(s).

It is well known that input-output relationships of a four-terminal network are described by a matrix equation of the form

$$\begin{bmatrix} P_i(s) \\ \\ U_i(s) \end{bmatrix} = [T(s)] \begin{bmatrix} P_o(s) \\ \\ U_o(s) \end{bmatrix} , \qquad (4)$$

where P_o, U_o are the pressure and volume velocity at the output of the network, respectively, P_i, U_i are the corresponding quantities at the input of the network, and T(s) is the transmission matrix of the network. The transmission matrix for a lossless uniform tube with cross-sectional area S and length L is given by (Flanagan, 1972)

$$T_u(s) = \begin{bmatrix} \cosh(SL/c) & (dc/S)\sinh(SL/c) \\ \\ (S/dc)\sinh(SL/c) & \cosh(SL/c) \end{bmatrix} \qquad (5)$$

where c is the velocity of sound in air = 35,000 cm/s and d is the density of air = 0.00114 gm/cm^3. The transmission matrix for the four-pole network representation of the viscous and the heat conduction losses is given by (Pipes and Hovanessian, 1969)

$$T_l(s) = \begin{bmatrix} 1 & R_a \\ \\ G_a & 1 + R_aG_a \end{bmatrix} . \qquad (6)$$

The transmission matrix for the entire tube is then the product of the individual transmission matrices. Thus the relation between pressures and volume velocities at the glottis P_g, U_g and at the mouth P_m, U_m is given as

$$\begin{bmatrix} P_g \\ \\ U_g \end{bmatrix} = \begin{bmatrix} A & B \\ \\ C & D \end{bmatrix} \begin{bmatrix} P_m \\ \\ U_m \end{bmatrix} , \qquad (7)$$

where

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ Z_g & 1 \end{bmatrix} \prod_{i=1}^{n} T_l^{(i)} T_u^{(i)}, \quad (8)$$

and Z_g is the glottal impedance and n is the total number of uniform tubes. From Ohms law,

$$P_m = Z_r U_m. \quad (9)$$

By solving Equations (7) and (9) we obtain the transfer function $F(s) = U_m/U_g$ which yields

$$F(s) = U_m/U_g = 1/(CZ_r+D). \quad (10)$$

C and D are computed from the matrix product in Equation (8).

Next, the formants and bandwidth were determined by finding the zeros of 1/F(s). The inverse of the transfer function was taken to prevent overflows during computations. The inverse of transfer function was divided into the real part and the imaginary part. We first examined the real part by increasing the imaginary part of the complex frequencies s_n in 200 Hz step until a sign reversal is found. When a sign reversal had been detected, we searched a zero crossing within the interval using the Newton-Raphson method (Press et al., 1986) with an accuracy of 0.1 Hz. About five to ten iterations were enough to find each root (i.e. a pole) from which the formant and bandwidth can be computed by the Equation (2).

The above algorithm was tested using six well known area functions of Russian vowels that were measured in a study by Fant (1960). The Russian vowels considered in his study were

/ɨ/, /i/, /e/, /a/, /o/, and /u/. The computed formants using the current algorithm were compared with both the spectrographically measured values and computed values by an electrical line analog (LEA) given in the Fant's study as shown in Fig. 9. The average deviation of computed formants from the acoustically measured values was of the order of 9% in F1, and 5% in both F2 and F3. The maximum absolute deviation in F1 was occurred in /i/ (60 Hz). It was occurred in /e/ for F2 (117 Hz) and F3 (250 Hz). The results show that the tube model of the vocal tract and the current algorithm can correctly reflect articulatory states of the given vocal tract with reasonably accurate acoustic outputs. The sources of error are likely to be inaccurate estimations of the constants representing the energy loss terms and difference between actual and measured cross-sectional area.

The above algorithm was used to compute formant frequencies from the area functions determined by the three different conversion algorithms. The source code of the algorithm is listed in Appendix C with the numerical values used to represent the energy loss terms and other constants. The computed formants were compared with acoustically measured values determined by the procedures described in the next section.

D. Analysis of acoustic data

Due to the acoustic noise produced by the gradient coils, it was difficult to estimate formant frequencies of the utterances simultaneously recorded with MR imaging.

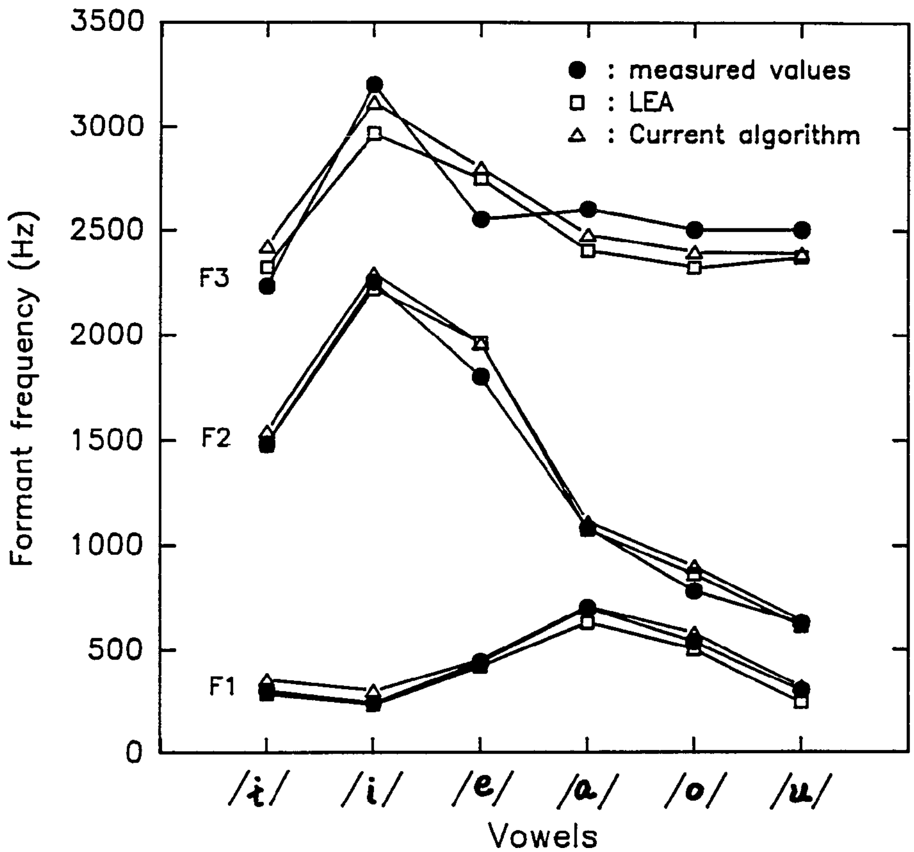


Fig. 9. Computed formants from Fant's area functions using the current algorithm (triangles). They are compared with both spectrographically measured values (filled circles) and computed values using LEA (squares) by Fant (1960).

Therefore, speech samples were collected separately in a sound-controlled room immediately prior to the MR session. It was assumed that the differences in tongue positions used in forming vowels at the two times would not be large.

The same list of five words "heed" (/IY/), "head" (/EH/), "had" (/AE/), "hod" (/AH/), and "who'd" (/UW/) was given to the subjects. Each word was pronounced five times, yielding a total of 25 utterances for each subject. The utterances were recorded on a tape and analyzed using an linear predictive coding (LPC) autocorrelation method (Markel and Gray, 1973, 1976) with the following procedures. For each utterance, the most stationary 200-ms portion of waveform were selected and digitized at a 10 KHz sampling frequency. After high-frequency preemphasis, the autocorrelation method was applied to the data with a 20 ms Hamming window and 20 ms window shift. A 20th order polynomial was fit to each frame, and the roots of the polynomial were solved using the Laguerre's method (Press et al., 1986). The formant frequencies and bandwidths were estimated from the roots of the polynomial. The above algorithm usually yielded eight or nine pairs of frequencies and bandwidths. Selection of the first three formants was straightforward in most cases. This is because bandwidths were usually much smaller for formants than non-formants. However, when two frequencies existed within a 200-300 Hz range around a probable formant value which could be inferred from the results of previous frames, the mean value of the two candidates was taken as the formant frequency. The results from five utterances of each vowel (total 50 observations per vowel) were averaged to produce a set of mean formants. The mean formant values measured from the acoustic signals and the formants computed from the area functions using the three different conversion algorithms are listed in Table I. They are graphically compared in Fig. 10.

III. Results and discussion

The MR images of the two subjects obtained during the production of sustained vowel sounds /IY/, /EH/, /AE/, /AH/, and /UW/ are shown in Fig. 11 (subject MJM) and Fig. 12 (subject LSB). The corresponding cross-sectional areas estimated by the three different conversion methods from the MR images are graphically represented in Fig. 13 (MJM) and Fig. 14 (LSB).

The FLASH-type scan with an optimized flip angle of 50^0 was found to yield the best images with reasonable signal-to-noise ratio and sufficient contrast to observe the entire vocal tract shape distinctly. The lingual contours, velum position, contours of the hard and soft palate, pharynx and the lips were well defined in the MR images.

A. Qualitative description of MR images

Differences in the tongue shapes between vowels could be clearly observed in the MR images. The corresponding MR images of both subjects showed that the vowels /IY/, /AH/, and /UW/ were distinguished by the location of the tongue constriction (LOC) or the tongue body position in the oral

Table I. List of acoustically measured formant frequencies by the LPC analysis of subjects' utterances and computed formant frequencies from area functions estimated by the three different area conversion methods.

Vowel	Subject MJM			Subject LSB		
	F1	F2	F3	F1	F2	F3 (Hz)
/IY/						
LPC:	290	2351	2853	302.	1997.	3213.
Ladefoged:	345.	2348.	2836.	355.	1803.	3295.
Rubin et al.:	368.	2416.	2890.	371.	1885.	3495.
Lindblom and: Sundberg	382.	2358.	2837.	417.	1839.	3262.
/EH/						
LPC:	480.	1933.	2668.	470.	1736.	2599.
Ladefoged:	440.	2109.	3062.	480.	1730.	2864.
Rubin et al.:	460.	2154.	3171.	498.	1779.	3019.
Lindblom and: Sundberg	509.	2057.	2838.	544.	1756.	2827.
/AE/						
LPC:	583.	1857.	3027.	641.	1753.	2659.
Ladefoged:	519.	2114.	2813.	545.	1676.	2847.
Rubin et al.:	523.	2236.	2951.	553.	1770.	3071.
Lindblom and: Sundberg	595.	2100.	2915.	625.	1689.	2850.
/AH/						
LPC:	671.	1243.	2745.	699.	1144.	2840.
Ladefoged:	691.	1476.	2742.	818.	1592.	2596.
Rubin et al.:	726.	1536.	2900.	828.	1731.	2753.
Lindblom and: Sundberg	803.	1530.	2726.	897.	1862.	2579.
/UW/						
LPC:	326.	1107.	2183.	332.	976.	2408.
Ladefoged:	428.	1236.	2209.	448.	837.	1980.
Rubin et al.:	455.	1293.	2304.	524.	964.	2044.
Lindblom and: Sundberg	484.	1265.	2216.	529.	812.	1936.

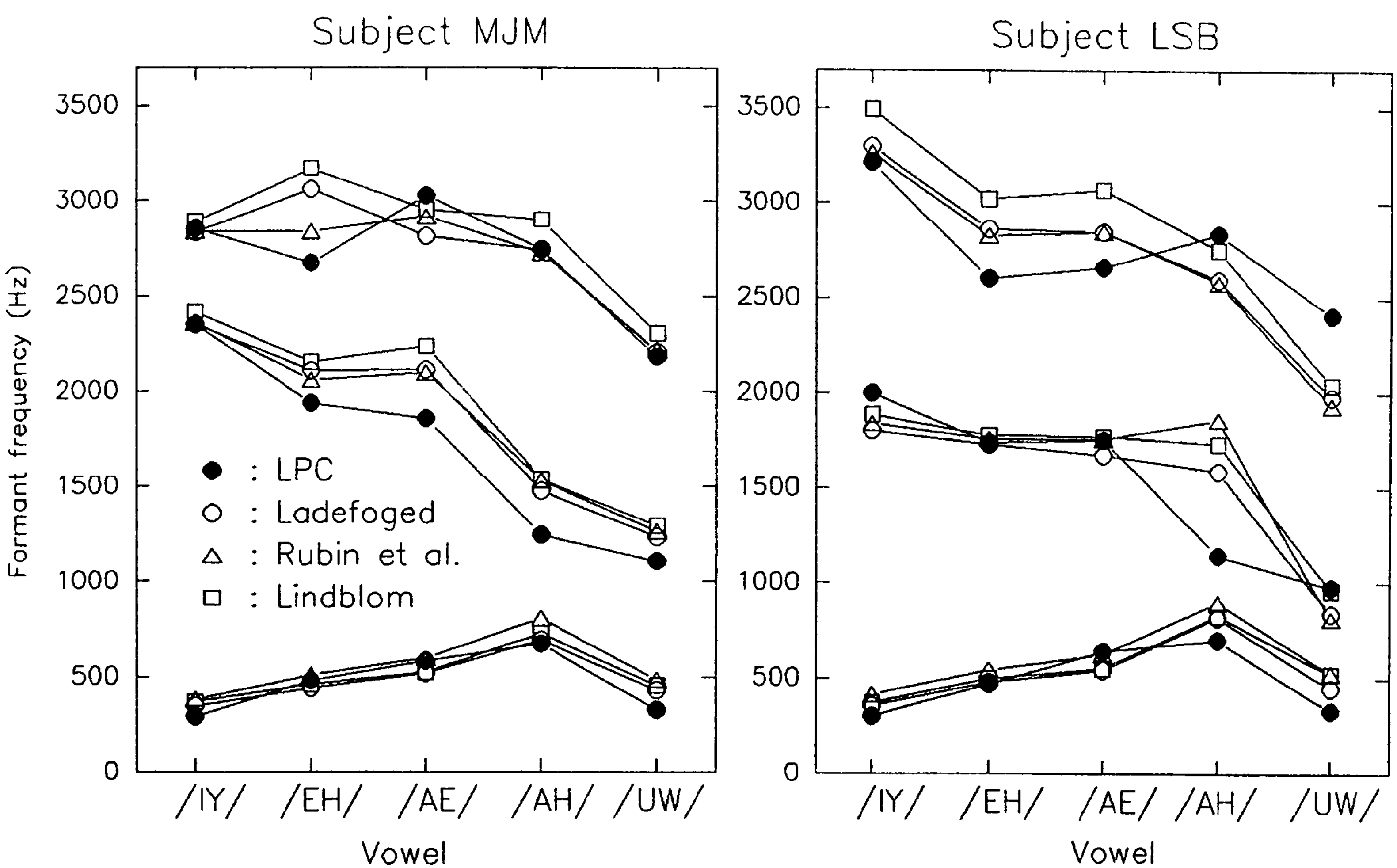


Fig. 10. Formant frequencies estimated by the three different area conversion algorithms are compared with each other. Acoustically measured values (filled circles) are plotted as standard values.

Fig. 11. MR images of the vocal tract of subject MJM obtained during sustained vowel sounds: /IY/ (top left), /EH/ (top right), /AE/ (middle left), /AH/ (middle right), and (e) /UW/ (bottom left).

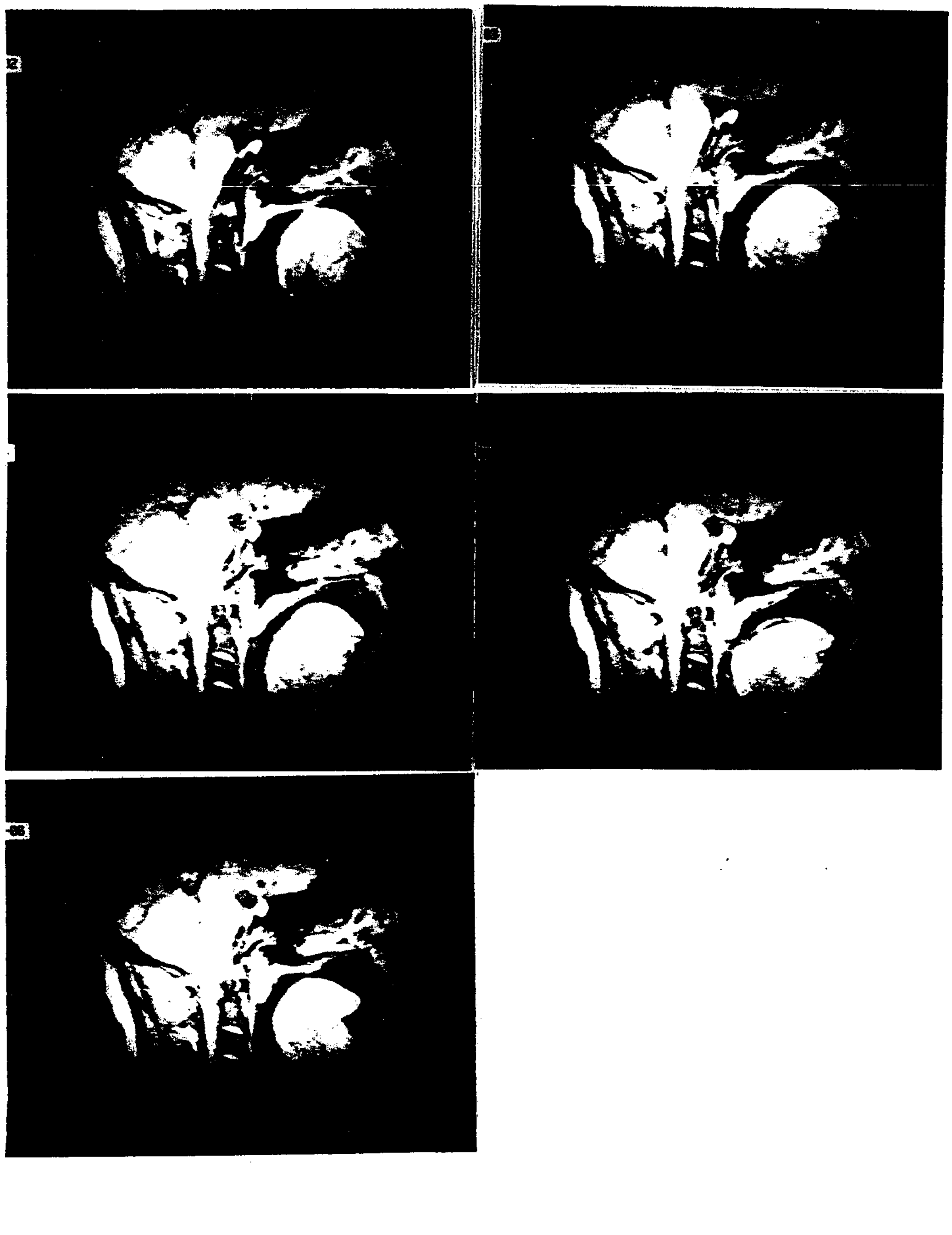


Fig. 12. MR images of the vocal tract of subject LSB obtained during sustained vowel sounds: /IY/ (top left), /EH/ (top right), /AE/ (middle left), /AH/ (middle right), and (e) /UW/ (bottom left).


-12


-10


-11

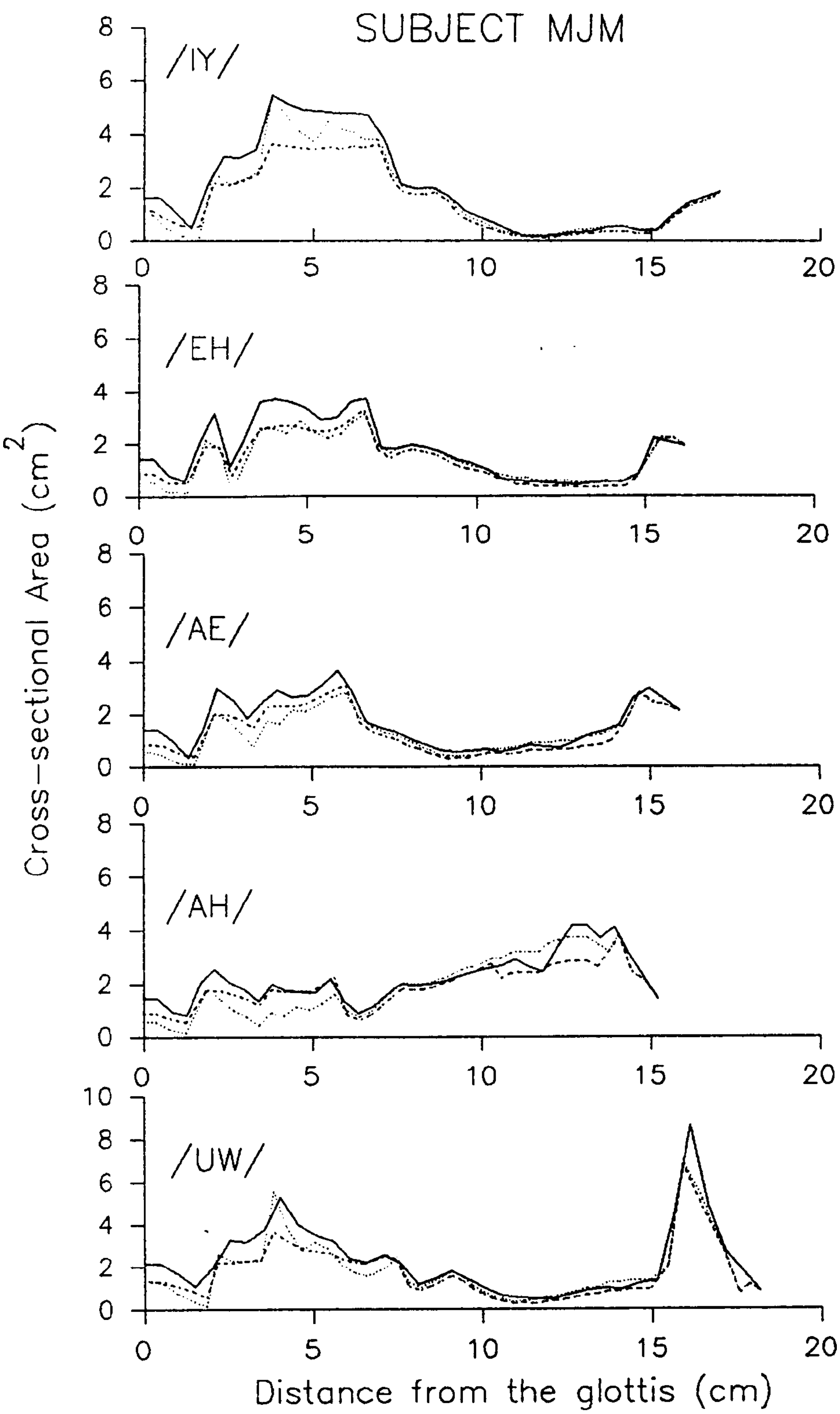


Fig. 13. Area functions of subject MJM determined by the three area conversion algorithms (solid line: Ladefoged, dashed line: Rubin et al., dotted line: Lindblom and Sundberg).

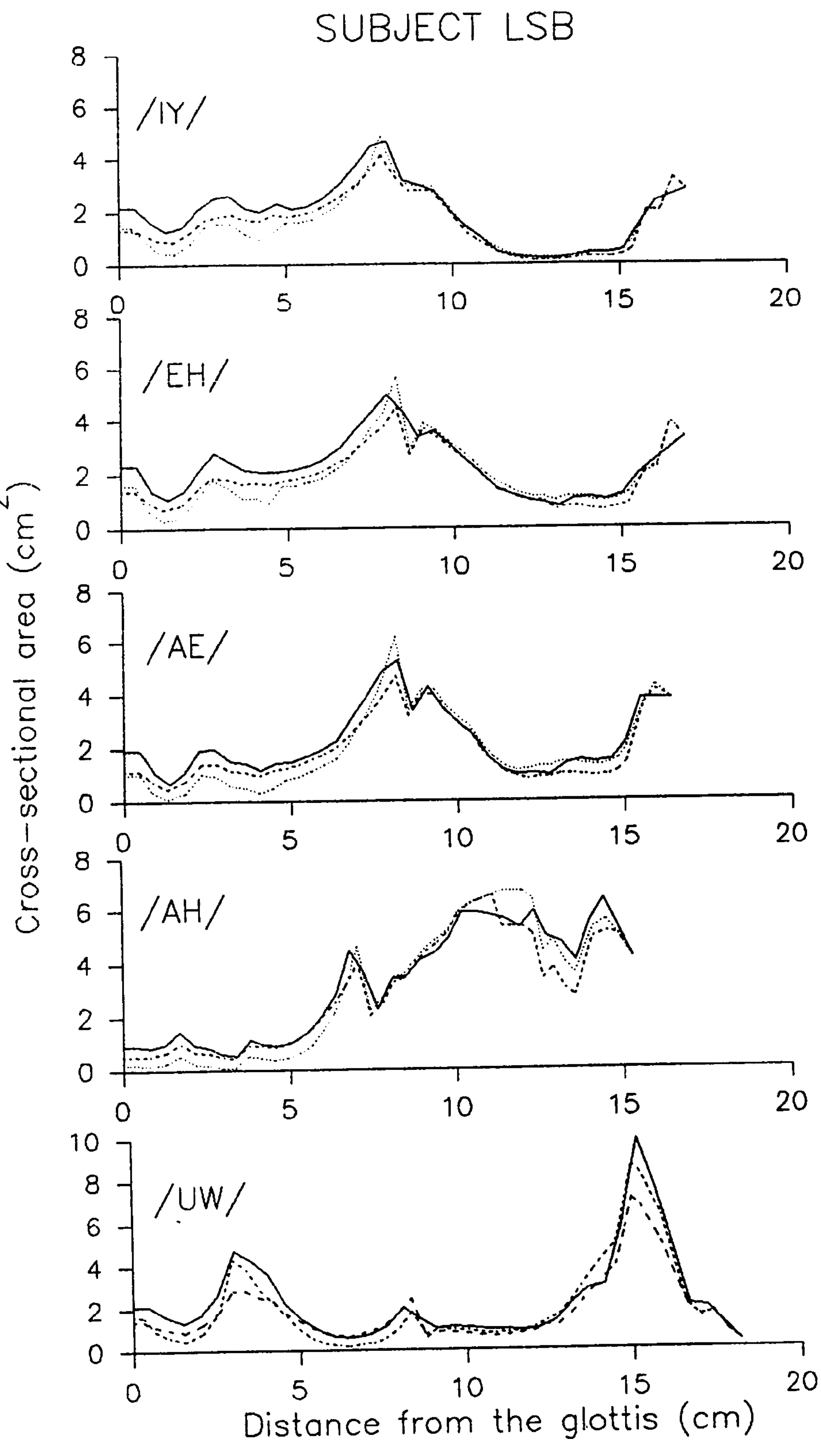


Fig. 14. Area functions of subject LSB determined by the three area conversion algorithms (solid line: Ladefoged, dashed line: Rubin et al., dotted line: Lindblom and Sundberg).

cavity. For /IY/, the tongue body was positioned in a superior-anterior region, creating a narrow constriction in the hard palate region. The degree of tongue-palate constrictions (DOC) were comparable between the two subjects. For /AH/, the tongue body was positioned in an inferior-posterior position with respect to the oral cavity. This created a narrow constriction in the upper pharyngeal region above the epiglottis. However, DOC was somewhat different between the two subjects for the vowel /AH/. For /UW/, the tongue body was located in a superior-posterior position for both subjects. However, LOC was different between the subjects: It was located in the soft palate region in case of subject MJM, and the upper pharyngeal region in case of subject LSB. The origin of such a difference might be the difference in combined muscle recruitment by the subjects or native language. However, it is interesting to note that despite the substantial articulatory differences a similar perceptual quality was evident. Such inter-subject variations in a vowel articulation may be a nature of speech production in that, for speakers, they might allow articulatory freedoms only governed by listeners' responses or self-feedback or language-specific speech learning.

For the three front vowels (/IY/, /EH/, and /AE/), the corresponding MR images of both subjects showed that the tongue body was positioned anteriorly with different tongue-palate distances. The overall tongue-palate distance increased in the order of /IY/, /EH/, and /AE/ for both subjects. As the tongue was lowered progressively from /IY/ to /EH/ to /AE/, the distance between the tongue and the upper pharyngeal wall decreased. For subject MJM, this pharyngeal narrowing was clearest between /IY/ and /EH/. For subject LSB, it was most evident between /EH/ and /AE/. It was also observed that subject MJM mainly changed LOC to contrast /EH/ and /AE/ whereas subject LSB employed the pharyngeal narrowing gesture. These observations on the front vowels suggest that although the tongue-palate distance is the primary articulatory factor in the differentiation of the front vowels, the pharyngeal narrowing is another articulatory factor that may also be considered.

The velum position and soft palate contour also differed substantially among vowels. The position of the velum was lowest for /AE/ in case of subject MJM. It was lowest for /AH/ in case of subject LSB, however. A velum lowering gesture may become prominent whenever there is an effort to narrow the pharyngeal region irrespective of the tongue body position. The MR images also showed that the nasal cavity was not utilized during the production of the vowels. This can be observed from the close contact of the velum with the back wall of the pharynx, which effectively decoupled the oral and nasal cavities. This result was consistent with the fact that no nasal sounds were used in this study. However, there may have been a slight coupling during /AH/ by subject LSB. The vocal tract length was also different across the vowels. It ranged from 15.2 cm (/AH/) to 18.2 cm (/UW/) in the case of subject MJM, and from 15.3 cm (/AH/) to 18.2 cm (/UW/) in the case of subject LSB.

The above observations illustrate well how a speaker manipulates the tongue and other articulators to produce intended vowel sounds. We have also employed a cine loop display of the vocal tract MR images. This "video-display" dynamically visualizes distinctive articulatory features between the various vowels.

B. Acoustic analysis of MR images

From Fig. 10, it can be observed that the relative formant frequency relationships among vowels obtained from the tube model using the three different conversion methods generally agreed well in most instances with those of acoustically measured formant values. However, the formant frequencies calculated from MR data deviated up to 718 Hz (F2 of /AH/ produced by subject LSB) from those of actual acoustic measurements. It is not unusual for the third and higher formant values, but not F1, to differ substantially for repeated productions of the same vowel. However, it was the first formant frequency in general that has the greatest relative errors with respect to the acoustically measured formants. The relative errors of F1, F2, and F3 averaged for all vowels and all conversion algorithms were 18%, 12%, and 5% in the case of subject MJM. They were 23%, 15%, and 10% in the case of subject LSB.

The discrepancies may reflect the inaccuracy of the width-to-area conversion method as well as parameters used in the tube model. Some errors may also result from errors in acoustic measurements. Another source of errors may be variations in tongue position between the sound recording session and image acquisition. It may have occurred in the case of /AH/ produced by subject LSB during MR imaging session. Unusually large deviation in computed F2 (718 Hz) from the acoustically measured value can be observed in Fig. 10.

Of the many possible sources of error, the main source of error is likely to have been inaccuracy of the area conversion algorithms. This results from a lack of knowledge about the relationship between the cross-sectional dimension and the midsagittal measurement upon which the area conversions were based. Therefore, the algorithms considered in this study may be considered incomplete in that they are using only one parameter, without any constraint, to estimate a two-dimensional quantity which is irregular in shape. For these and other reasons, extension of MR imaging to acquire multi-sectioning sagittal, coronal, and transaxial, or volume data, is highly desirable.

C. Comparison of area conversion methods

1. Comparison of estimated cross-sectional area

From area functions shown in Fig. 13 (subject MJM) and Fig. 14 (subject LSB), we can observe that derived cross-sectional area by the three conversion methods were substantially different in the pharyngeal region. They agreed relatively well in the oral region, however. The use of

Ladefoged's area table yielded the largest cross-sectional area for both subjects in the entire region of the vocal tract. In most instances, the algorithm by Lindblom and Sundberg yielded the least area in the pharyngeal region and the algorithm of Rubin et al. yielded intermediate area values.

The relatively good agreement in the oral region may result from the fact that two of the algorithms (Ladefoged's area table and the algorithm of Rubin et al.) were derived from the same data (Ladefoged et al. 1971), and all three algorithms were based on the data measured from speakers of American English for area conversions in the oral region. It may be noted that the agreement does not mean necessarily that the estimated area values were accurate, only that they were relatively reliable.

However, for both subjects, there were discrepancies in the oral region for the vowels /AH/ and /UW/. These vowels were usually characterized by large tongue-palate distance in the front of the oral cavity, as can be observed from the corresponding MR images. The other vowels, /IY/, /EH/, and /AE/, were characterized by relatively smaller tongue-palate distances. Therefore, it seems that when the tongue-palate distance is large in the oral region, the algorithms behave differently or inaccurately.

The discrepancy in the pharyngeal region may be attributed to the fact that each area conversion for the pharyngeal region was derived from different data sets and speakers with different language backgrounds. At the current stage of knowledge, it is not possible to say much about the main source of the discrepancy because there are too many sources of variability (e.g., variations in the size of the pharynx, differences in pharyngeal maneuvers due to personal or language difference, and differences in the scheme used to analyze the data). However, it is clear that the behavior of a conversion algorithm in terms of cross-sectional area was strongly dependent on the data used to derive the algorithm. To overcome this data dependency of the conversion algorithms, i.e. to improve the generality of the algorithms, we obviously need more reliable data from a large number of speakers.

2. Comparison of acoustic performance

The test of acoustic performance of the three conversion algorithms was based on relative errors of computed formants with respect to the acoustically measured values. The overall relative errors were computed by averaging the relative errors of the first three computed formants. They are listed in Table II and plotted in Fig. 15 for both subjects.

The use of Ladefoged's area table yielded the smallest (average) relative errors except in the cases of /EH/ and /AE/ for subject MJM, and /AE/ for subject LSB. In these cases, the algorithm of Lindblom and Sundberg yielded the best estimation of formants. The formants of front vowels were better estimated than those of back vowels, /AH/ and/UW/, for both subjects, irrespective of the algorithm used.

Table II. Relative errors of computed formants with respect to the acoustically measured values.

	Relative error(%)				
	Subject MJM				
Conversion method	/IY/	/EH/	/AE/	/AH/	/UW/
Ladefoged	6.7	10.7	10.7	7.3	14.7
Rubin et al.	10.3	11.3	11.0	12.7	21.0
Lindblom and Sundberg	11.0	6.0	6.3	14.7	21.3
	Subject LSB				
Ladefoged	10.3	4.0	8.6	21.6	22.3
Rubin et al.	12.7	8.0	10.0	24.0	24.7
Lindblom and Sundberg	16.0	8.6	4.3	33.3	32.0

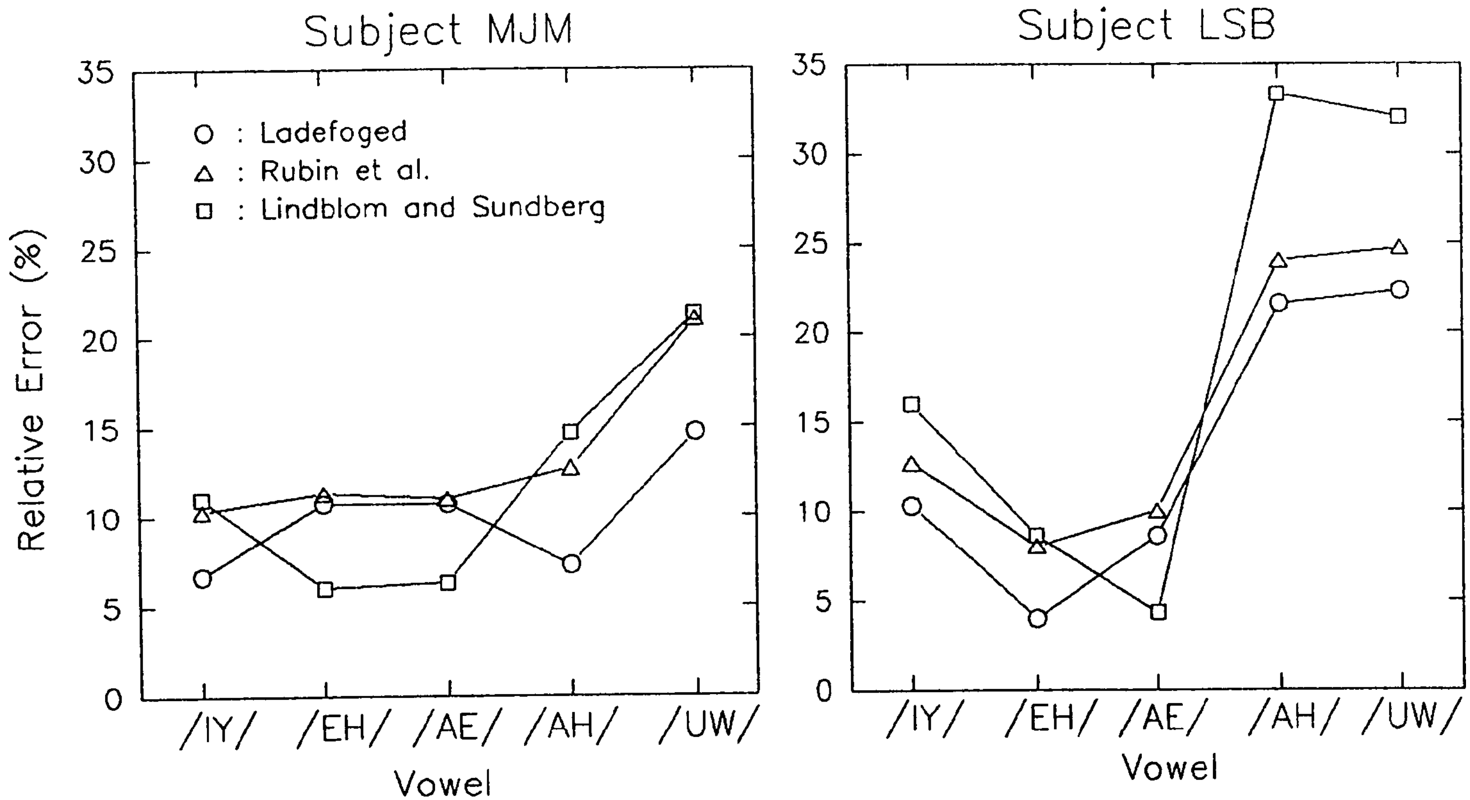


Fig. 15. Comparison of relative errors between the computed formants estimated by the three area conversion algorithms. The errors are computed with respect to the acoustically measured values.

It is interesting to note that there was a tendency for the overall acoustic performance of the conversion algorithms to be proportional to the magnitude of the cross-sectional area in the pharyngeal region. This suggests that the algorithms underestimated the cross-sectional area in the pharyngeal region. Also, the relatively poor estimates of the formants for the back vowels, especially /UW/, may be attributed to the fact that the cross-sectional area at the pharyngeal region, which is an important region for the vowel, were poorly determined by the algorithms.

Although the same order of inaccuracy also exist at the pharyngeal region for the front vowels, the acoustic errors induced by the inaccuracy in the estimated cross-sectional area in the pharyngeal region may be compensated by the relatively well determined cross-sectional area at the palatal region which is an important region for the production of front vowels. Therefore, we may tentatively conclude that the current algorithms inaccurately estimate cross-sectional areas in the pharyngeal region, and their accuracy becomes poorer even in the palatal region when tongue-palate distances are large.

It is possible to derive a new table or algorithm which yields more accurate estimations of formants for the current subjects with an iterative procedure. It may not be so much meaningful to do so, however, because the validity of such a procedure might be limited to the current subjects. Clearly, a large amount of reliable vocal tract data from multiple subjects are needed to improve the accuracy and generality of current area conversion algorithms.

Based on the above results, it was decided to use the Ladefoged's area table for the purposes of studies described in chapter 3 and 4 because it yields least relative errors among the algorithms tested. However, this does not mean at all that Ladefoged's area table will always yield the best estimation of formants for various speakers of different language backgrounds.

IV. Conclusions

The use of a spoiled-gradient gradient-echo sequence in combination with a special receiving-coil provided fast (4 seconds), high-quality midsagittal MR images of the vocal tract. The spatial configurations of the different speech articulators (the lips, jaw, tongue, velum, and larynx tube) can be clearly observed. The images showed that, for the vowels /IY/, /AH/, and /UW/, each vowel has a specific tongue configuration which can be clearly characterized by tongue body position in the oral cavity, or by the location of the tongue constriction. For the vowels /IY/, /EH/, and /AE/, the tongue-palate distance is the main articulatory difference between them. This suggests that the so-called tongue height is a reasonable articulatory feature in the cases of front vowels. However, it can be equally well described by the degree of the tongue constriction in the oral cavity. In addition to the tongue-palate distance, the pharyngeal narrowing gesture seems to be another factor to be considered in the production of front vowels.

The MR images have sufficient resolution and contrast for direct measurement of the dimension of the entire vocal tract, enabling the construction of reasonable acoustic tube models of the tract. The formant frequencies calculated from the tube model agreed relatively well with those obtained from acoustic measurements. However, there were some substantial differences in values between the two sets. It is likely that the discrepancy stems from the inaccuracy of the width-to-area conversion due to the lack of constraint between cross-sectional and midsagittal dimension. Clearly, a more refined conversion algorithm based on direct measurement of cross-sectional or three-dimensional vocal tract is needed to establish reliable articulatory-acoustic relationships.

The cross-sectional areas estimated by the three area conversion algorithms compared in the study agreed relatively well in the palatal region for the two subjects examined. However, when the tongue-palate distance became large in the palatal region, the discrepancy between the cross-sectional areas estimated by the algorithms were large. Also, the algorithms tested seemed to underestimate the cross-sectional areas in the pharyngeal region. Among the conversion algorithms tested, the use of Ladefoged's area conversion table yielded the largest cross-sectional area in the pharyngeal region and relatively better estimations of formant frequencies for the two subjects.

This study was perhaps the first attempt to image the entire midsagittal vocal tract during vowel production. It demonstrated the usefulness of MR imaging technique in speech production research. With improved techniques this technology may also be a potential tool for assessment of speech disorders.

CHAPTER 3. STUDY OF VOWEL ARTICULATION
IN A PERCEPTUAL SPACE

I. Introduction

Vowels have traditionally been defined as points in an acoustic space whose axes are the first two (F1/F2) or three (F1/F2/F3) formants. The formant space has long been used for the representation of acoustic or perceptual properties of vowels. The space has also been known to be useful for the phonetic description of vowel articulation. For instance, F1 is known to be related to the tongue height and F2 (or the difference between F2 and F1) to the backness of the tongue in the oral cavity (Ladefoged, 1975).

However, it is well known that two talkers can produce vowel sounds which are perceptually similar but substantially different in formant values. They may also produce perceptually different vowel sounds which have remarkably similar formants (Peterson and Barney, 1952). Therefore, it may happen that both articulatorily and perceptually distinctive vowels produced by multiple speakers overlap in a formant space. Such overlaps occur even when data are limited to those vowels which were unanimously and correctly classified by a group of listeners (Peterson and Barney, 1952). This might obscure perceptual identities of vowels as well as phonetic descriptions of vowel articulation in the formant space.

The major sources of acoustic variations which cause overlaps in the formant space are known to be: (1) differences in sex, age, vocal tract dimension, and articulatory habits; (2) consonantal context effects (or coarticulation) and speech rate and stress. The first factors are usually referred to as inter-speaker differences and the second factors as intra-speaker variations. To overcome the overlap caused by the inter-speaker differences, several speaker or vowel normalization procedures has been proposed (Gerstman, 1968; Lobanov, 1971; Harshman and Papcun, 1976; Neary, 1977; Bladon et al. 1984; Syrdal and Gopal, 1986; Miller, 1989). Irrespective of the techniques used, a common idea in these normalization procedures is to separate vowels produced by multiple speakers into unique, non-overlapping regions. More formally, the idea is to maximize phonetic contrast between vowel categories by factoring out systematic -- but phonetically nondistinctive -- covariations in the acoustic signals of the vowels, thereby revealing more invariant acoustic patterns between the vowel sounds (Neary, 1989). However, it should be noted that the normalization schemes are data analytic procedures, not truly perceptual ones. This is because they deal with the separation of vowel categories based on acoustic data only, not listeners' performance. Therefore, a question that remains to be answered is how much they reflect actual auditory-perceptual process by listeners.

Normalization procedures can be divided into two categories: extrinsic and intrinsic normalization. Extrinsic normalization procedures (Gerstman, 1968; Lobanov, 1971; Harshman and Papcun, 1976; Neary, 1977) reduce the inter-speaker variations in acoustic data by rescaling formant values to minimize overlap in the scaled formant space. They use the minimum and maximum values of the first and second formants (Gerstman, 1968), a multidimensional factor analysis technique (PARAFAC) (Harshman and Papcun, 1976), mean and deviation (Lobanov, 1971), or a speaker-dependent scaling constant (Neary, 1977) for the scaling. Such procedures thus require a whole data set or prior knowledge about individual speaker's vowel space to derive scaling factors. This is an undesirable feature for the on line recognition of vowels produced by various speakers. Also, although such procedures may be effective in vowel categorization, the scaling of formant values may eliminate (or distort) phonetically relevant articulatory information contained in the acoustic data, as discussed by Disner (1980).

Recently, intrinsic speaker normalization procedures using auditorily based approaches have been proposed (Bladon et al., 1984; Syrdal and Gopal, 1986; Miller, 1989). They assume that all information necessary to identify a vowel is contained within the vowel spectrum, and that when certain transformations of the fundamental frequency and formants or the whole spectrum are employed, variations due to the inter-speaker differences will be removed in the transformed space.

These procedures are strictly based on individual acoustic data and thus they do not require any prior knowledge about speakers. In addition, because formant values are not altered during such normalization procedures, the possibility of distorting articulatory information contained in acoustic data is minimized. For terminological convenience, we will use the term "perceptual (vowel) space" to refer to a space which is derived by a transformation of acoustic data.

This study is based on the belief that articulatory properties of vowels can be more accurately represented in a perceptual space than in the conventional formant space. This is because variations in acoustic data caused by inter-speaker differences can be largely removed in the perceptual space. We can expect that what mainly remains in the space might be the vowel-specific articulatory differences, which is the very subject we want to study, and the variations due to personal articulatory style or intra-speaker differences acting as perturbations. The purpose of this study is thus to develop a phonetic space in which articulatory and auditory (or perceptual) properties of vowels can be directly related with each other. This is also an attempt to integrate both articulatory and acoustic descriptions of vowels in a single perceptual space. Such an integration is desirable for the phonetic description of vowel articulation and the comparison of the vowel systems across languages.

For the purpose mentioned above, Miller's approach (1989) was selected because the use of the formant-ratio theory (Lloyd, 1890) suggests that it may have the best articulatory characteristic among the auditorily-based intrinsic normalization procedures. Although the bark-difference approach (Syrdal and Gopal, 1986) has been shown to be effective in binary feature classifications of vowels, it requires a statistical procedure (i.e., a linear discriminant analysis) for vowel categorization. I believe that the use of statistical procedures should be avoided in vowel categorization (and generally in perception research) because there is no evidence that listeners recognize vowels based on statistical procedures. A more physically or physiologically based approach is desirable.

Because this study is based on the formant-ratio theory elaborated by Miller (1989), brief reviews of the theories and some problems are described below.

A. Formant-ratio theory

The formant ratio-theory (Lloyd, 1890) states: (1) like articulations produce perceptually similar vowel qualities and (2) like articulations produce like formant ratios. Lloyd called his theory "the relative resonance theory" and stated that vowel quality depends on the interval between the formants, not their absolute values. The first statement may be regarded as a hypothesis of the formant-ratio theory. In fact, it is a fundamental assumption in modeling of speech production because of the anatomical similarity of the human vocal organs. The second statement can be considered as a practical application of the hypothesis for acoustic data.

However, it was observed that the F2/F1 and F3/F2 ratios are very similar within the vowel groups [/AH/ and /AW/] and [/UU/ and /UW/] in American English (Potter and Steinberg, 1950). A similar observation was duplicated with a large number of subjects of different sexes and age (Peterson and Barney, 1952). The observations illustrate that there is no unique relationship between formant ratios and perceptual qualities of vowels. Peterson (1961) concluded that only to a first approximation phonetically equivalent vowels have similar formant ratios, and the major strength of formant-ratio theory is its ability to reduce inter-speaker differences due to sex and age (i.e., adult versus child) in the formant representation of vowels. Peterson also described the underlying idea of the formant-ratio theory as:

> "Since vocal tract cavities of men, women, and children are similar in general shape rather than in size, it seems reasonable that vowels produced with similar articulatory shapes should have formant frequencies which do not correspond in absolute magnitudes, but which do have similar frequency ratios. Thus in the development of speech, for example, a child acquires certain articulatory neuromuscular coordinations. By employing a system of formant ratios, his articulatory patterns need not alte appreciably as his vocal cavities increase in size." (Peterson, 1961; p.26).

This statement illustrates well how the formant-ratio theory can be related to both articulatory and perceptual aspects of vowel production.

B. Auditory-perceptual theory

The auditory-perceptual theory proposed by Miller (1989) is a descendent of the formant-ratio theory. To present his theory graphically (or geometrically), Miller constructed a three-dimensional auditory-perceptual space (APS). With the improved version of formant-ratio space and several assumptions about the stages of human vowel recognition, he tried to explain several important topics in vowel perception such as speaker normalization, diphthongization, and coarticulation effects.

Based on a observation by Potter and Steinberg (1950) that strong correlations exist between the fundamental frequency and formant values within each vowel category, Miller developed the concept of a "sensory reference" (SR) to accommodate the correlation into his theory. SR was given by

$$SR = 168(GMF0/168)^{1/3}, \qquad (11)$$

where GMF0 is the geometrical mean of the fundamental frequency of a subject during the production of a vowel. The value of 168 was selected as a geometrical mean of the fundamental frequency of adult male (125 Hz) and female (225 Hz). The inclusion of the exponent (1/3) was based on an observation that formants of children are about 30% higher than those of adult males, while fundamental frequencies differ by about 100% (Ainsworth, 1975). Thus, the role of SR was to establish an absolute reference point against which the position of the first formant is to be judged such that the ratio of the first formant and SR is almost independent of the talker's sex and age.

APS was constructed with a following coordinate system:

$$x = \log (SF3/SF2),$$
$$y = \log (SF1/SR), \qquad (12)$$
$$z = \log (SF2/SF1),$$

where SF1, SF2, and SF3 represent the first three (sensory) formant frequencies and SR is the sensory reference. The x and z coordinates are adapted from the original formant-ratio theory. Miller used log formant-ratios based on evidence that the logarithmic frequency scale is quite fundamental to hearing (Greenwood, 1961). Using the sensory reference, Miller largely removed the overlaps between the vowels /AH/ and /AW/, and the vowels /UU/ and /UW/.

After examining the distribution of points in APS with a data base consisting of 435 data points for the monophthongal ten American English vowels /IY/, /IH/, /EH/, /AE, /AH/, /UH/, /AW, /UU/, /UW/, and /ER/, Miller found that a vowel slab exists in APS as illustrated in Fig. 16. The equation of the middle plane of the vowel slab was given by

$$x + y + z = 1.22 \pm 0.135, \tag{13}$$

which means that the vowel slab has a thickness of about 0.156 log units. This implies that most of the variability of the distribution of the vowel data points in APS is along the height and width dimensions of the vowel slab, not in the depth dimension, which is well illustrated by the side view of the vowel slab shown in Fig. 16.

To bring the vowel slab to the vertical with respect to a viewer, Miller transformed the coordinates of APS using the equations given by:

$$\begin{aligned} x' &= 0.7071(y - x), \\ y' &= 0.8162z - 0.4081(x + y), \\ z' &= 0.5772(x + y + z). \end{aligned} \tag{14}$$

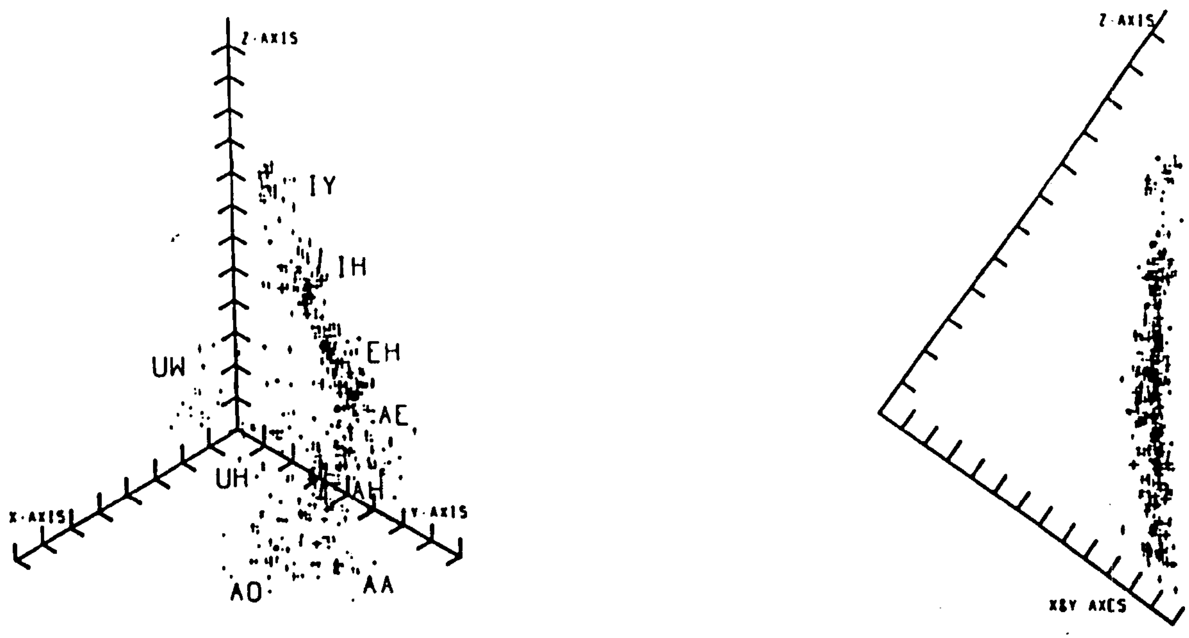


Fig. 16. Miller's auditory-perceptual space (APS) (left) and the existence of the vowel slab (right) (from Miller, 1989).

This is a transformation using the Euler angles. It does not change the origin, but the view point. The resultant space was called SLAB space. Miller constructed perceptual vowel target zones for the vowels by hand with resolution of 0.01 log unit in SLAB space with aids of computer graphics. The target zones are shown in Fig. 17. Vowel zone for /ER/ is absent because it is below the vowel slab. Miller claimed that the vowel target zones classify the ten vowels with 93% accuracy and there is no overlap between target zones.

C. Problems

When one combines the two statements of the original formant-ratio theory mentioned earlier, the theory may be understood as "like ratios of formants have perceptually similar vowel qualities". It seems that this is the way Miller interpreted the format-ratio theory. However, such an interpretation lacks the articulatory side of the theory. To my understanding, what was emphasized in the original formant-ratio theory is not acoustic but articulatory aspects of vowels.

An implicit hypothesis of the original formant-ratio theory seems to be a uniform scalibility of the vocal tract size between men, women, and children. However, it is known that the vocal tract is not uniformly scalable because of the relative length difference between the pharynx and the mouth cavities among speakers. Fant (1966) has noted that the main difference between male and female vocal tracts is in the pharyngeal length, which is shorter relative to the mouth in

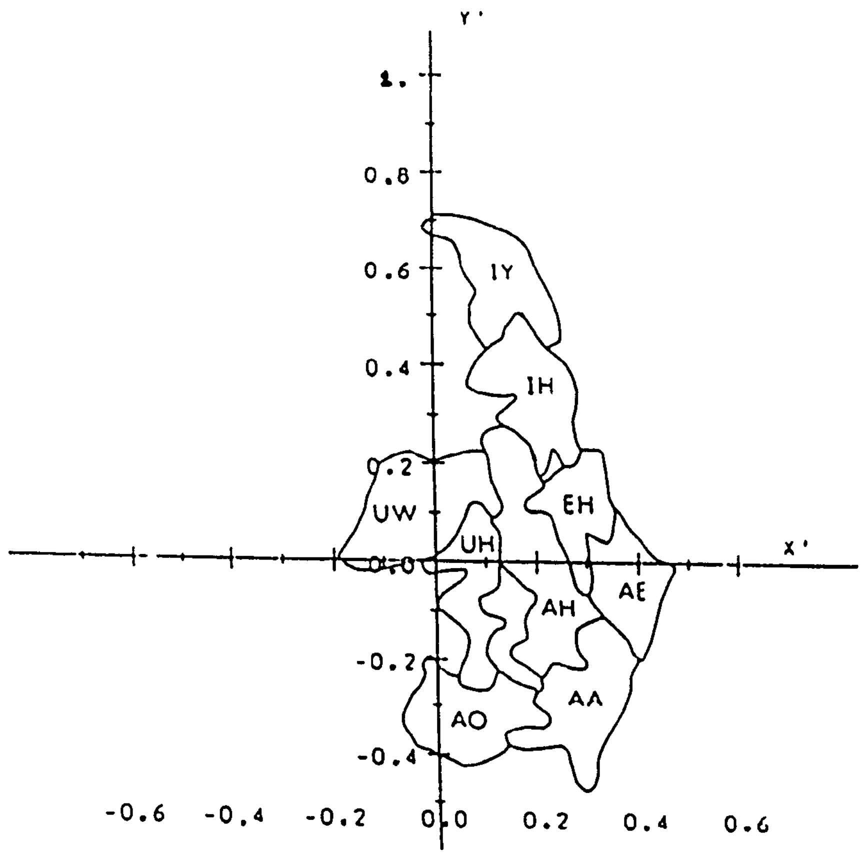


Fig. 17. SLAB space and vowel target zones constructed by Miller (from Miller, 1989).

females. There might also be a difference in cross-sectional dimension (e.g., Nordstrom, 1975). The formant-ratio theory can not account for this kind of anatomical constraints. This may act as an error source in testing predictions of the theory, and may partially explain why the distribution of data points in APS is represented as a vowel slab with the depth dimension, not a plane.

Several uncertainties exist in the procedures used by Miller to determine the vowel zones. Because the key idea in the construction of vowel target zones is to resolve overlaps between vowel categories by imposing boundaries between them, the perceptual validity of the vowel target zones is largely dependent on a criterion used to determine the boundaries. However, Miller did not mention explicitly what criterion was used. Also, the number of data points (viz., 435) is too small for one to claim that each vowel zone was correctly determined. Miller regarded APS as a perceptual space, and the vowel zones as perceptual targets. However, the use of the formant-ratio theory suggests that it might also have articulatory characteristics. Therefore, in the current study, articulatory aspects of the vowel zones will be equally emphasized. This is because I believe that a perceptually relevant space should have explicit articulatory characteristics. After all, the ultimate purpose of articulation is to induce perceptual contrasts for speech communication.

II. Modified auditory-perceptual space (MAPS)

A. Construction of MAPS

Miller argued that the depth dimension is certainly related to the differences in vocal tract size, fundamental frequency, and the third formant. Although a further study is needed to reveal quantitative relationships between these factors and the position of a data point in APS, we tentatively assume that they are the sources of deviations from the predictions by the formant-ratio theory (or the auditory-perceptual theory) due to imperfect normalization of the inter-speaker differences. Analysis of the vowel data set by Peterson and Barney (1952) conducted in the current study also showed that the depth dimension was almost uniformly presented among vowels or groups of speakers. For instance, the middle position of the vowel slab in APS was 1.203 (± 0.106) for 33 men, 1.191 (±0.084) for 28 women, 1.241 (± 0.102) for 15 children. Numbers in parentheses are depth dimensions expressed by two-sigma values. The differences in the middle positions of the vowel slabs might result from differences in the vocal tract anatomy among speaker groups.

The depth dimension seems to come from combined effects of the inter-speaker differences and it is currently not possible to separate them in APS. Therefore, based on the fact that variations in depth dimension are small compared with those of height and width dimensions, the degeneracy along the depth dimension (z′ axis in SLAB space) was eliminated by a transformation:

$$X = x'/d,$$
$$Y = y'/d, \qquad (15)$$
$$Z = z'/d,$$

where d is the Euclidean distance from the origin to a point (x',y',z') in the SLAB space. The transformation corresponds to a unit vector representation of the vector (x',y',z') in SLAB space, yielding a constraint $X^2 + Y^2 + Z^2 = 1$. The resulting spherical surface was projected onto the X-Y plane to obtain a two-dimensional working space. The new space was called a modified auditory-perceptual space (MAPS). Because of the nature of the transformation, the coordinates of MAPS are equivalent with those of SLAB space.

The transformation merely redistributes the variations along the depth dimension onto a plane. However, it has two practical advantages. First, it yields a two-dimensional space without any loss of characteristics of the three-dimensional APS. Second, the resulting space (i.e., MAPS) is more flexible than APS in that another dimension can be added if necessary. For example, the dimension can be the intrinsic duration of vowels, third formant, or any physical quantity which can be considered as a determinant of articulatory or perceptual properties of vowels.

B. Determination of vowel target zones in MAPS

Vowel target zones for nonretroflex, monophthongal nine American English vowels, /IY/, /IH/, /EH/, /AE/, /AH/, /AW/, /UH/, /UU/, and /UW/ were determined in MAPS. Construction of the vowel target zones was necessary to test the perceptual validity of MAPS in terms of vowel categorization. Also it can be used as an automatic vowel recognizer. In the current study, another important usage of the vowel target zones was to classify simulated vocal tract shapes into appropriate vowel categories for the articulatory interpretation of MAPS.

1. Data

Our vowel formant data mainly came from data sets published by Peterson and Barney (1952) and by Peterson (1961). Also, we added data measured in our laboratory which have been almost correctly identified by listeners (98%). Strictly speaking, only data points which are unanimously identified by listeners should be used to claim a perceptual validity of the vowel zones. Number of data points in each vowel group were 147 for /IY/, 147 for /IH/, 151 for /EH/, 146 for /AE/, 156 for /AH/, 154 for /AW/, 166 for /UH/, 150 for /UU/, 160 /UW/. The total number of data points were 1,377. The data base is listed in Appendix D.

2. Procedure

Near the edge of a vowel zone, the data points were usually distributed sparsely. Thus, with visual inspection only, it was not easy to determine whether a data point should be included in the vowel zone or not. This was especially true when a data point exists on a region where overlap occur between vowel zones. Without a decision criterion, there exists the possibility that a vowel zone will contain different vowels, or occupy unreasonable regions.

To provide such a criterion, it was necessary to have enough data points, especially near the edge of each vowel zone. As a way to do this, we selected two data point and added the averaged value to the original data set (i.e., $N_{tot} = {}_nC_2 + n$, where N_{tot} = total number of data points in the new data set, n = number of original data points, ${}_nC_2$ = number of derived data points). For instance, in case of /IY/, there were 147 data points in the original data. We obtained 10,878 (= ${}_{147}C_2$ + 147) data points by the procedure. Although the interpolated data points were not tested perceptually, we can assume that they have the same perceptual qualities of the original data points because they are always surrounded by positions of the original data points.

After the data generation procedure, the position (X,Y) of each data point in MAPS was determined with a resolution of 0.01 log unit using the equations (12), (14), and (15), subsequently. It was observed that many data points which widely separated in formant space fell into the same cell in MAPS. This was interpreted as a result of the reduction of inter-speaker differences in MAPS. Also, the number of entries in each cell gradually decreased from the center to the edge of a vowel region. To quantify the observation, a number P_{ij} is assigned to the i^{th} cell in a vowel zone j as

$$P_{ij} = 100\ (n_{ij}\ /\ MAX_j\ ), \qquad (16)$$

where n_{ij} is number of entries in the i^{th} cell and MAX_j is the maximum entry number among cells in a vowel zone j. P_{ij} is determined to be an integer and the range of P_{ij} is from 1 to

100. Boundaries among adjacent vowel target zones where overlap occurs are determined by comparing magnitude of P_{ij}. For example, when a cell was occupied by two vowel zones, the cell was assigned to one of the vowel zones which had a larger magnitude of P_{ij}. The comparison of P_{ij} provides a reasonable criterion for the determination of boundaries of vowel target zones, although it should be confirmed by perceptual experiments with human listeners.

P_{ij} was also used to divide each vowel zone into "core" and "boundary" zones with an assumption that the perceptual identity of a data point is degraded when its position is far from the center of a vowel region. The current value of P_{ij} used as a criterion for the division is 20, which means that the core zone is composed of cells whose entry number (n_{ij}) exceeds 20% of the maximum number of entry (MAX_j) in each vowel zone. There was no special condition to choose the criterion except that there were no overlaps among the core zones of vowels when P_{ij} was equal or greater than 20.

3. Results

The vowel target zones and core zones in MAPS are shown in Fig. 18. Each core zone was represented by a two-sigma-radius ellipse with an assumption that data points in each vowel zone are distributed normally. The two-sigma ellipse encloses approximately 95% of the population along the principal axes.

In terms of vowel production, the large and irregular shapes of vowel targets zone might be related to articulatory

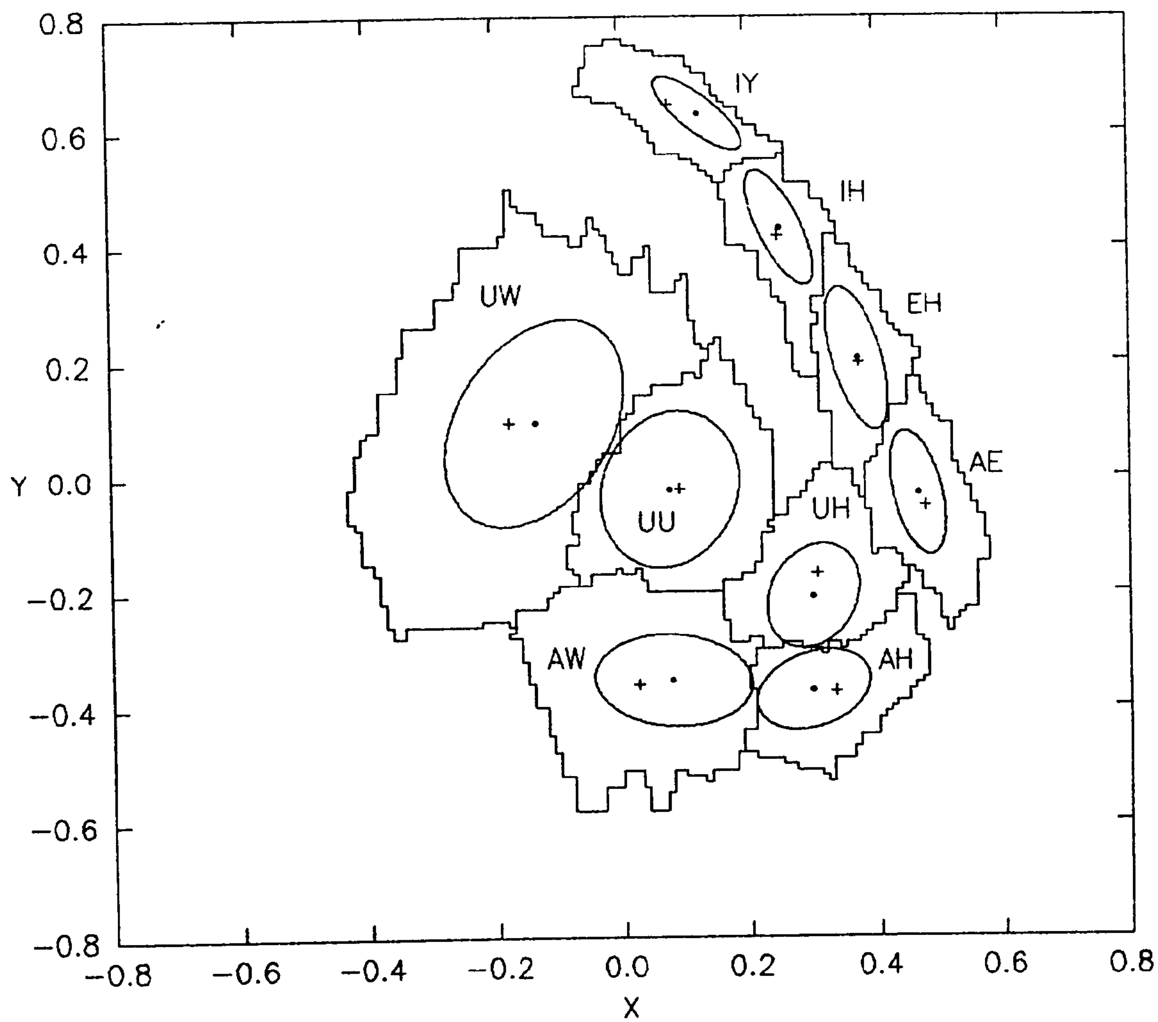


Fig. 18. Vowel target zones and core zones determined in MAPS. The core zones are represented as 2-sigma radius ellipses. The cross-marks and the dots are the centers of the entire vowel target zones and the core zones, respectively.

variations (or articulatory freedoms) among speakers in the production of the vowels. However, the boundaries between the target zones may act as constraints for speakers which prevent excessive deviations from the articulatory targets of the intended vowels. From a perceptual standpoint, the entire area of a vowel zone may represent the range of acoustic variations which can be tolerated by listeners. Therefore, we hypothesize that the boundaries between the vowel target zones are both articulatory and perceptual boundaries (or constraints) imposed by speakers and listeners to be able to communicate with each other.

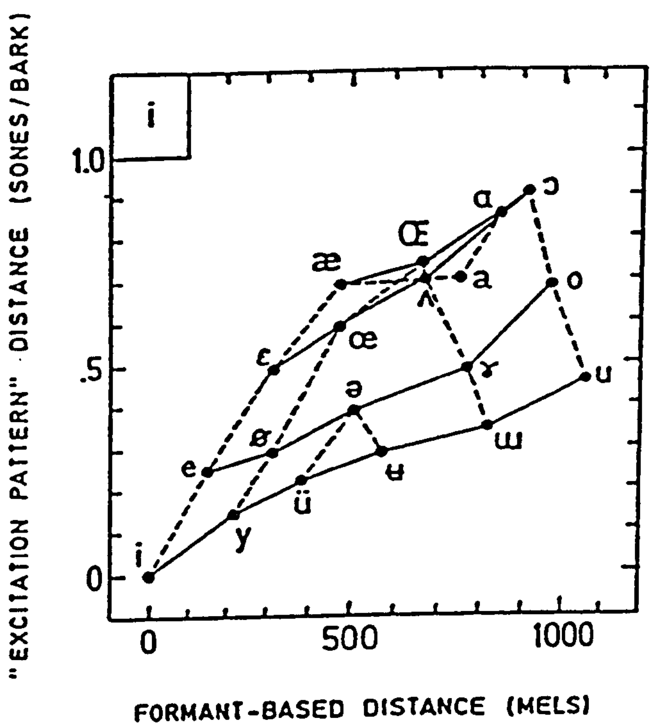
It is interesting to compare vowel target zones constructed in APS by Miller (Fig. 17) and in MAPS (Fig. 18). It can be observed that there exist two unoccupied regions in APS (Fig. 17), a small region between /IH/ and /EH/ and a larger region between /UH/ (/UU/ in Fig. 18) and /AH/ (/UH/ in Fig. 18). This might come from the use of an insufficient number of data points (total 435 in the study by Miller). In the current study, there were many data points falling into these regions. It was also observed that the /AH/ vowel zone in APS (/UH/ in Fig. 18) was determined too large in APS. In the data set used by Miller, there was only two data points which occupy this region in APS. In the data set used in the current study, there were no data points approximately above the x-axis in the /UH/ (/AH/ in Fig.18) vowel zones in MAPS. It is uncertain why Miller has so much extended this vowel zone in APS.

There exists an apparent gap or discontinuity between vowel zones [/IY/, /IH/, /EH/] corresponding to front vowels and [/UU/, /UW/] corresponding to high-back vowels in MAPS. It may result from an insufficient number of data points used. However, a more probable reason may be that this region is not preferred, articulatorily or perceptually, by speakers of American English. We expect that the "quantal theory" of speech (Stevens, 1989) be helpful to explain why this region is remained empty, if it should be.

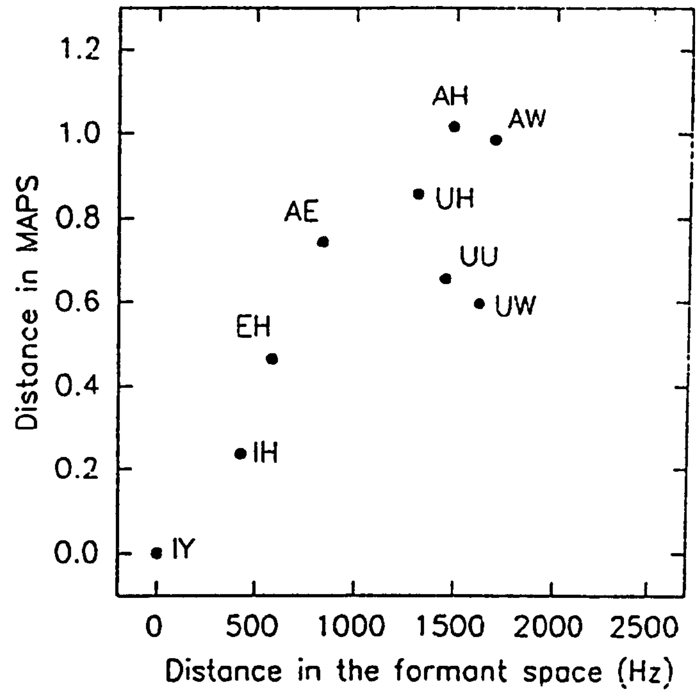
III. Properties of MAPS

A. Auditory characteristic of MAPS

Lindblom (1986) has adopted a perceptual distance measure based on an auditory transformation of vowel spectra to predict vowel systems occurred in natural languages. He has used the auditorily based distance measure to correct the unreasonable prediction of vowel positions (e.g., vowels between /IY/ and /UW/) by the "adaptive dispersion" hypothesis which initially employed a formant-based (mel) distance measure (Liljencrants and Lindblom, 1972). The comparison of the perceptual distances of vowels between the formant-based measure (ordinate) and the auditorily based measure (abscissa) given by Lindblom (1986) are shown in Fig. 19(a). The distance between /IY/ ("i" in Fig. 19(a)) and /UW/ ("u" in Fig. 19(a)) was greatly reduced in the auditorily based measure, making the probability of the occurrence of improbable vowels between /IY/ and /UW/ smaller. For the purpose of comparison, the (Euclidean) distances were computed with



(a)



(b)

Fig. 19. (a) Perceptual distances between vowels in an auditorily-based measure (ordinate) and in a formant-based measures (abscissa) (from Lindblom, 1986), and (b) those in MAPS (ordinate) and in F1-F2 space (abscissa). It shows that the distance relations between vowels in MAPS are comparable with those of the auditorily-based measure.

respect to vowel /IY/ in MAPS. They were also computed in the F1-F2 space. The result is shown in Fig. 19(b). The distance relationship in MAPS (the ordinate in Fig. 19(b)) was comparable to that of the auditorily based measure shown in Fig. 19(a). Although we did not attempt in the current study, it might be interesting to test the adaptive dispersion hypothesis in MAPS.

B. Reduction of inter-speaker differences in MAPS

To compare the degree of reduction of the inter-speaker differences in MAPS and the formant space, formant data of /IY/, /AE/, /AH/, and /UW/ produced by men, women, and children (Peterson and Barney, 1952) were plotted in MAPS and the formant space as shown in Fig. 20. Each vowel cluster was delimited by a 2-sigma-radius ellipse in both spaces.

Large distances between the centers of ellipses corresponding to men, women, and children in each vowel category in the formant space shows that acoustic variations caused by the inter-speaker differences may exceed variations of vowel articulation itself. Whereas, the extensive overlap between the ellipses of speaker groups in MAPS illustrate that the inter-speaker differences were largely removed. Other vowels showed the same tendency. The overlaps might be ascribed to the articulatory similarity in vowel production irrespective of speakers' sexes and age. Therefore, it allows more explicit articulatory interpretation of MAPS, which is an advantage over the formant space.

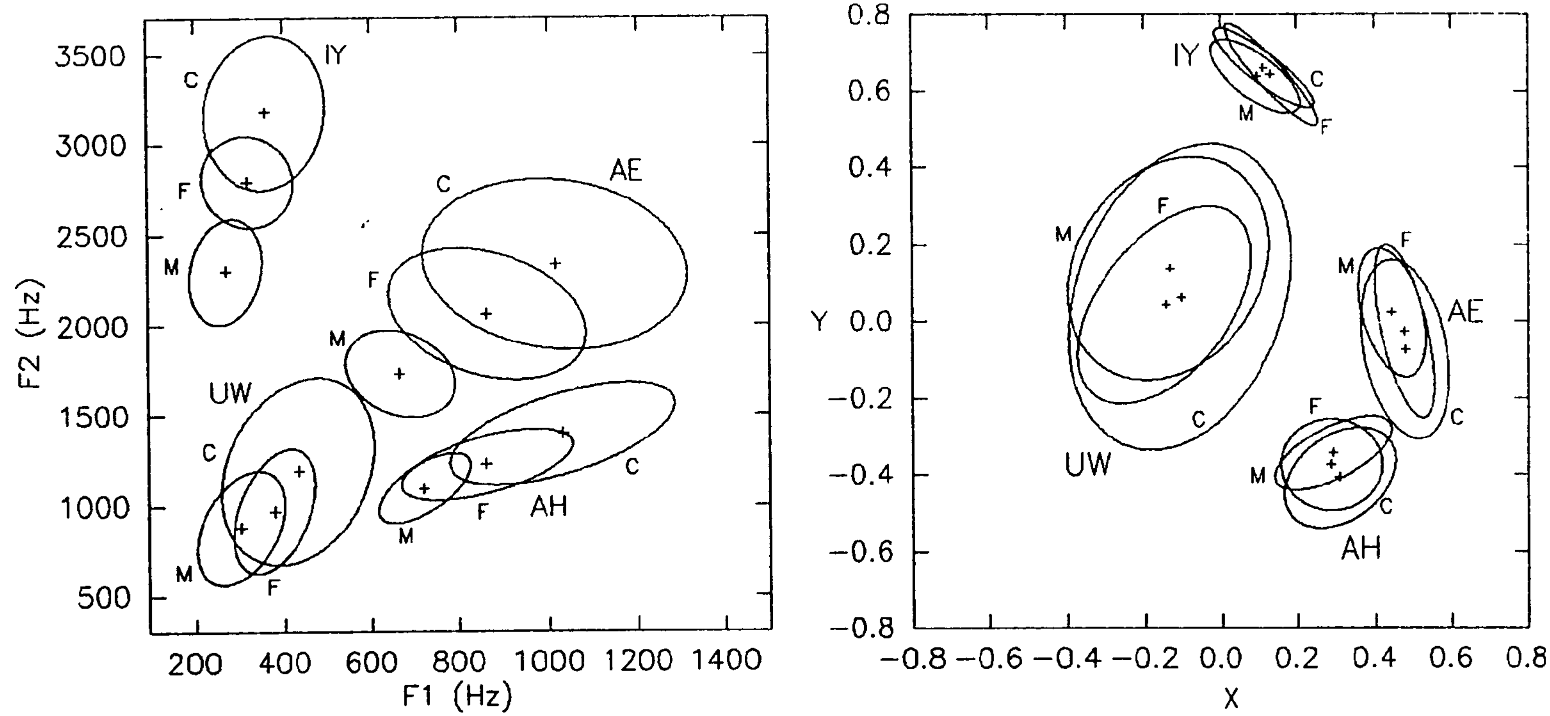


Fig. 20. Comparison of the reducibility of the inter-speaker differences between the formant space (left) and MAPS (right). (M: male, F: female, C: child).

Based on the observation, we hypothesize that the center positions of vowel target zones in MAPS are determined by the intrinsic articulatory differences between vowels, and positions of data points within a vowel zone are determined by personal variations in a vowel articulation (e.g, differences in tongue positioning and lip opening in the production of the vowel).

C. Vowel classification in MAPS

From a perceptual point of view, one measure of the goodness of a speaker normalization scheme is how well it separates various vowels into unique, non-overlapping regions. For that reason, a formant data of American English vowels by Peterson and Barney (1952) were classified in MAPS.

1. Data

The Peterson and Barney data has obtained by spectrographic measurements of vowels spoken twice by 32 men, 28 women, and 15 children during the production of vowels /IY/, /IH/, /EH/, /AE/, /AH/, /AW/, /UH/, /UU/, /UW/, and /ER/. Each speaker has produced each vowel twice, so there was a total 152 data points in each vowel category. The data are included in the vowel data base listed in Appendix D. Vowel /ER/ was not considered in the classification because it is retroflexed.

2. Procedure

After the determination of positions of 152 data points in each vowel category in MAPS by applying the equations (2), (4), and (5), subsequently, the data were classified by a simple computer graphics technique: (1) A grid system was drawn on a monitor by drawing x and y grid-lines with a resolution of 0.01 log unit; (2) A vowel target zone is displayed on the grid system by filling each cell with a (white) color; (3) Data of a vowel category corresponding to the vowel target zone were displayed on the monitor, while increasing a counter by one whenever cells in the vowel target zone are occupied by one of the 152 data points in the vowel category.

By the above procedures we can compute the number of data points which occupy the vowel target zone. By repeating the procedures with the whole data set, we can count how many data points in a vowel category occupy the corresponding vowel target zone and how many data points come from other vowel categories.

3. Result and discussion

The result of classification is shown in Table III. The overall correctness of classification is 89.0%, ranging from 96.1% (/IY/) to 81.6% (/IH/). The classification rate was reasonably comparable with the observation that human listeners identify English vowels with approximately 95% accuracy (Peterson and Barney, 1952). Also, the overall accuracy was better than that of a previous classification result (84.9%) of the same data set by Syrdal and Gopal (1986) in the bark-difference space. They used a linear discriminant analysis for the classification. The current classification rate seems to be less than that (93%) by Miller in APS. However,

Table III. Number of tokens of each vowel enclosed by each vowel target zone. The number of tokens in a vowel category are 152.

	target zones									Percent correct
Vowel	/IY/	/IH/	/EH/	/AE/	/AH/	/AW/	/UH/	/UW/	/UU/	
/IY/	146	6	0	0	0	0	0	0	0	96.1
/IH/	19	124	9	0	0	0	0	0	0	81.6
/EH/	0	17	134	1	0	0	0	0	0	88.2
/AE/	0	0	16	135	0	0	1	0	0	88.8
/AH/	0	0	0	0	136	9	7	0	0	89.5
/AW/	0	0	0	0	5	142	4	0	1	93.4
/UH/	0	0	1	2	12	4	133	0	0	87.5
/UU/	0	0	0	0	0	3	3	128	18	84.2
/UW/	0	0	0	0	0	2	0	10	140	92.1
Total	165	147	160	138	153	160	148	138	159	89.0

when we compare total number of data points considered (i.e., 1,368 data points in the current study versus 406 data points (excluding /ER/) in the study by Miller), the current result is not unreasonable.

Syrdal and Gopal (1986) argued that the bark-difference scale is a more accurate representation of the spatial coding of frequency by the auditory system than the log ratio scale. However, Hillenbrand and Gayvert (1987) has shown that the use of spatial coding of frequency by the auditory system does not, by itself, mean that the bark transformation of formants will necessarily solve any problems related to inter-speaker differences. Miller (1989) has also shown that the use of differences, ratios, or log ratios of any of the suggested auditory scales, Mels (Fant, 1973), Koenigs (Koenig, 1949), and Barks (Zwicker and Terhardt, 1980), as well as Hertzs, should be about equally effective in clustering the vowels. The extensive overlaps between ellipses of different speaker groups as shown in Fig. 20, and the higher rate of classification also suggest that the log ratio scale used in the current study is equally effective or better than the use of the bark transformation in terms of the inter-speaker differences reducibility and the vowel categorization. In addition, the vowel classification procedure used in the current study was strictly based on an individual formant data, not a statistical analysis. It is a more desirable feature for the application of such algorithm to the automatic recognition of vowels produced by various speakers.

D. Interpretation of core zone

We have shown that it is possible to divide each vowel target zone into the core and boundary zones using the number P_{ij}. The core zones were so determined that there were no overlaps between them as shown in Fig. 18.

From the figure, we observe that the nine vowel target zones are almost evenly distributed in MAPS. However, it can be observed that the core zones are divided into three groups [/IY/, /IH/, /EH/, /AE/], [/UU/, /UW/], and [/UH/, /AH/, /AW/] which have traditionally been described as front, low-back, and high-back vowels, respectively.

Notice the center of each core zone (dot) and the direction of deviation from the center of the entire target zone (cross-mark) in each of the three vowel groups. The centers of core zones in each vowel group become closer as if there exists a kind of affinity between the vowels. Because the three vowel groups have traditionally been differentiated by the tongue body position in the oral cavity, it may be reasonable to interpret the observation as articulatory as well as perceptual contrasts among the groups in MAPS. Therefore, we hypothesize that the core zones are unified vowel targets among articulatory, acoustic, and perceptual domain in the sense that they are more refined versions of the vowel target zones, and thus they may be used as representatives of vowels across the domains. We expect that the use of the core zones may provide a simple and meaningful way to compare vowel of a single language or between languages.

E. Inter-speaker variability in MAPS

We have shown that the inter-speaker differences are substantially removed in MAPS (see Fig. 20). However, the deviations of the center positions of ellipses between groups of speakers in a vowel category suggests that there still exist some variations due to the inter-speaker differences in MAPS. Although the deviation may partially come from the difference in fundamental frequency between the groups of speakers, the main source of the deviation might be the imperfect normalization of vocal tract size (i.e., length and cross-sectional dimension) between men, women, and children.

Although the consideration was limited to the vocal tract length, Fant (1975) introduced a non-uniform formant scaling method to reduce the acoustic variation between men, women and children. He showed that the transformation from male to female involves a larger scaling of the pharynx than of the mouth whereas, in the relation from female to child, the scaling is uniform. This suggests that the variations in MAPS can be further reduced by preprocessing the formant data with a normalization technique which can accommodate the anatomical constraint among speaker groups of different sex or age. This might allow more explicit explanation of articulatory and perceptual aspects of vowels in MAPS. To treat this problem rigorously, however, we eventually need a large amount of midsagittal and cross-sectional vocal tract data acquired from speakers of different sex and age. The MR imaging technique described in chapter 2 might be helpful to obtain such vocal tract data conveniently.

IV. Study of vowel articulation in MAPS

We have shown that MAPS is a more perceptually oriented space than the formant space in terms of the reducibility of the inter-speaker differences and the goodness in vowel categorization. We have also hypothesized that the boundary between the vowel target zones can be regarded as both perceptual and articulatory boundaries. Then it may be interesting to ask a question about what articulatory factors determine the position of the vowel target zones and the boundaries between them. The purpose of this study is thus to describe vowel articulation in MAPS in terms of three articulatory parameters, the location and degree of tongue constriction and the lip opening area, as a preliminary attempt to answer the question. The study also provides a way to interpret articulatory aspects of vowels in MAPS.

A. Procedures

We generated midsagittal vocal tract shapes using an articulatory model, and extracted three articulatory parameters from the generated vocal tract shapes. The corresponding formant frequencies are computed using the tube model of the vocal tract and corresponding positions in MAPS are determined by applying Equation (2), (4) and (5), subsequently. By these procedures, a vocal tract shape is mapped to a point in MAPS. The distribution of points are analyzed in terms of the three articulatory parameters.

1. Tongue model

Several articulatory models have been proposed for the generation of the tongue shapes for vowels (Stevens and

House, 1950; Lindblom and Sundberg, 1971; Harshman et al. 1977; Atal et al., 1978). Among them, a tongue model by Harshman et al. (1977) was selected because: (1) the model has been derived from the vocal tract data of the largest number of subjects (five American speakers); (2) two parameters are enough to generate realistic midsagittal tongue shapes relevant to most American English vowels; (3) the two parameters accounts for 96% of the variance in the data. The two-parameter tongue model used front and back rasing components which were extracted using a factor analysis technique (PARAFAC) from midsagittal x-ray data on ten vowels of American English produced by five speakers. The measurement system used in their study is illustrated in Fig. 21. The system was so determined that the position of a reference line always refers to the same location in the vocal tract irrespective of the length of the vocal tract. It was used to refer to a region in the vocal tracts generated in the present study.

The philosophy of the PARAFAC tongue model is that the articulatory parameters which are the speaker-independent tongue displacement modes, when weighted by their contributions to specific vowels and speakers, can generate midsagittal tongue shapes that occur in all nonretroflex American English vowels. This can be expressed by the equation

$$x_{ijk} = n_i + d_{ijk}, \qquad (17)$$

where the indices i, j, and k refer to a location in the vocal tract, a vowel, and a speaker, respectively. x_{ijk} is

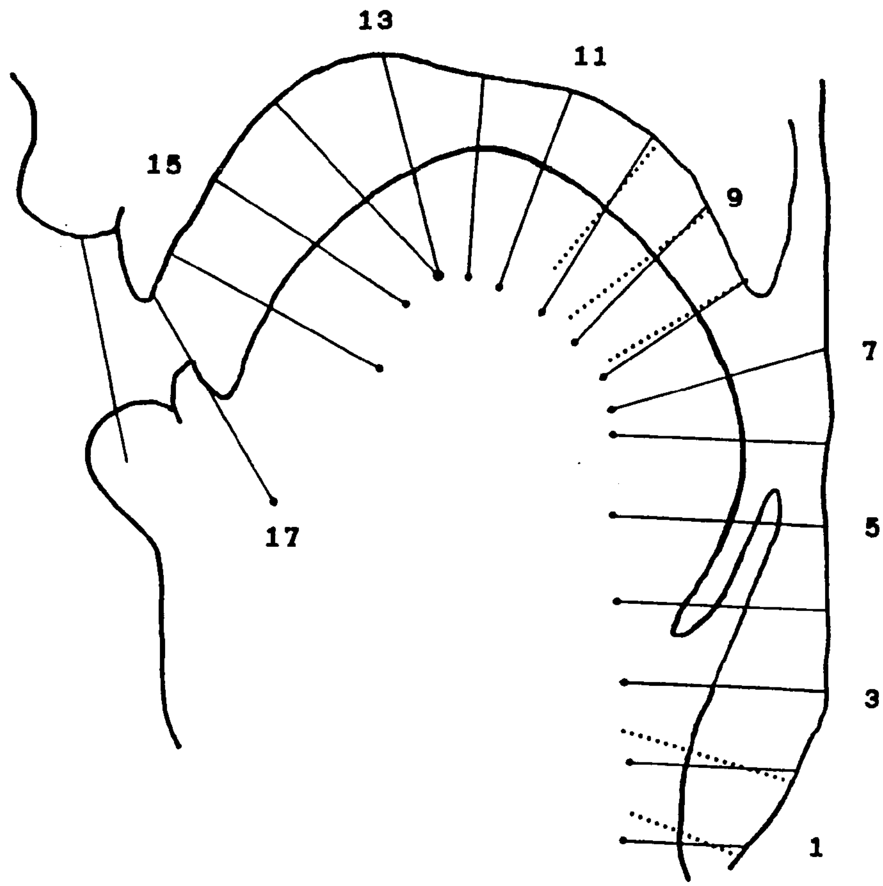


Fig. 21. Grid system used by Harshman et al. (1977) to measure midsagittal widths from x-ray profiles. The data were used to derive PARAFAC tongue model (from Harshman et al., 1977).

the estimated distance in centimeters between the tongue and the upper surface of the vocal tract (i.e., midsagittal width), and n_i is a reference distance measured in centimeters at location i with respect to the upper surface of the vocal tract. d_{ijk} is displacement from n_i. The range of index i is from 1 to 16, which represents a location along the vocal tract from just above glottis (reference line 1 in Fig. 21) to near the alveolar (reference line 16). The displacement d_{ijk} is determined by the relation

$$d_{ijk} = (t_{1i}v_{1j}s_{1k}) + (t_{2i}v_{2j}s_{2k}), \qquad (18)$$

where v_{1j} and v_{2j} are the weights for a vowel j on the components t_1 and t_2, respectively, and s_{1k} and s_{2k} are speaker-dependent scaling factors with respect to components t_1 and t_2, respectively.

2. Generation of the tongue shapes

The PARAFAC tongue model was adopted to generate tongue shapes of vowels for the current study. The s_{1k} and s_{2k} were set to one with an assumption that speakers employ similar tongue shapes to produce a vowel because of similar vocal tract anatomy. As a result, tongue shapes generated in this study were of a hypothetical speaker whose n_i is given by the average value of ten vowels of five speakers. Index j, which designates vowels, was also not needed because the purpose is to generate all possible tongue shapes, not particular ones. Then the Equation (18) can be simplified as

$$d_i = v_1t_{1i} + v_2t_{2i}. \qquad (19)$$

Next, we calculated d_i by varying v_1 and v_2 within appropriate ranges to generate all possible displacements with respect to n_i for the hypothetical speaker. The upper and lower limits of both v_1 and v_2 were chosen as 1.2 and -1.2, respectively, and an increment step of 0.04 was used. The distance between the upper and lower walls of the vocal tract at location i is then given by

$$x_i = n_i + d_i. \qquad (20)$$

The values of t_{1i}, t_{2i}, and n_i were adopted from a work by Ladefoged et al. (1978). The vocal tract length was fixed as 16.0 cm.

To avoid the possibility of generating impossible tongue shapes (e.g., closure in the vocal tract by the contact of the tongue and upper wall of vocal tract, or an inflection in the curvature of the middle of the tongue) as cautioned by Ladefoged et al. (1978), two constraints were imposed: (1) the midsagittal dimension of all sections is equal or larger than 0.1 cm, and (2) the midsagittal dimensions of section 8, 9, and 10 (see Fig. 21) are less than 3.0 cm.

Initially, 1,660 tongue shapes were generated. However, there were many similar tongue shapes which yielded almost the same formant frequencies. Therefore, the components v_1 and v_2 which yielded similar tongue shape and formants were averaged, and a smaller set of tongue shapes were generated with the new components. The final number of tongue shapes were 321.

To parameterize the generated vocal tract shapes, the degree of tongue constriction (DOC) and the location of the tongue constriction (LOC) were extracted from the generated tongue shapes because it is believed that they best characterize the tube model of the vocal tract. Also, it is believed that the LOC and DOC are more appropriate articulatory parameters for a description of articulatory-acoustic relationship (Stevens and House, 1955) than the putative height and backness of the tongue. LOC is usually defined as the distance from the glottis to the point of minimum midsagittal width in the region spanned by the tongue and DOC is defined as the midsagittal width at LOC. The resulting range of LOC was varied 5.38 - 12.63 cm in 0.25 cm step. These ranges approximately correspond to the region between the reference line 6 and the midposition of the reference lines 13 and 14 in Fig. 21. DOC was varied in 0.1 cm step, but the range was dependent on LOC. The range of DOC was 0.1 - 1.5 cm in the maximal case. It was 0.1 - 0.5 cm in the minimal case.

3. Lip configuration

Lip configuration is also important in vowel production (Lindblom and Sundberg, 1971). The above procedures only yield tongue shapes from section 1 to 16 shown in Fig. 3.6. To determine midsagittal widths of the entire vocal tract, we need to know the widths of section 17 (the teeth separation) and section 18 (the lips).

The section length corresponding to the lips was fixed as 1 cm. The lip opening width, W, (distance between the corners of the mouth) was fixed as 4 cm which is an average value of eight American speakers estimated from a lip data base (Linker, 1982). The vertical lip separation H (x_{18}) was varied from 0.1 cm to 2 cm with an increment step of 0.1 cm. The dimension x_{17}, corresponding roughly to the distance between the upper and lower teeth, was taken as the average value of x_{16} and x_{18}, as suggested by Ladefoged et al. (1978). Finally the lip opening area A was computed as

$$A = 0.75*W*H, \quad (21)$$

where the weight 0.75 was estimated by an average value of 0.7 (Fromkin, 1964) and 0.8 which was estimated by a regression analysis using the data by Linker (1982) as described in chapter 2. The resulting range of lip opening area was from 0.3 cm^2 to 6.0 cm^2 in 0.3 cm^2 step, yielding 20 different lip opening areas with length of 1 cm.

The lip configuration simulated in this study can not generate the lip gesture for rounded vowels but generates only the resultant effective lip opening area. This is because the lip opening width (W), lip opening height (H), and lip protrusion should be considered simultaneously to simulate such lip gestures. The separated simulation of tongue shape and lip opening is also not realistic because the lips, jaw, and tongue does not move independently. However, we believe that the above tongue and lip simulation would be enough to capture basic articulatory features in terms of LOC, DOC and the lip opening area.

4. Computation of formant frequencies

Total 6,420 (321 tongue shapes and 20 steps in lip opening areas in each tongue shape) midsagittal vocal tract shapes were generated by the above procedures. Corresponding formants were computed using the tube model of the vocal tract described in chapter 2.

For the formant computations, each 17-section tongue shape (excluding the lips) was converted into 34 sections using a cubic spline interpolation method (Press et al., 1986) because we need at least 20 sections to reasonably estimate formant frequency using a tube model (Atal et al, 1978). Area functions were estimated using Ladefoged's area conversion table listed in Appendix B.

B. Results

1. Description of vowel articulation in MAPS

The computed formants from the generated vocal tract shapes were plotted in F1-F2 plane and in MAPS as shown in Fig. 22. Many overlapped points in the formant space and in MAPS illustrate that there exist several vocal tract shapes which give the same formants or the same perceptual quality. It may be interpreted as the possibility of compensatory articulation (Steven and House, 1955; Atal et al., 1978) or personal variations in the production of a vowel.

Several articulatory trajectories were plotted in MAPS as shown in Fig. 23(a) - (f). For the ease of visual observation, the nine vowel target zones were also plotted as a background in a plot. In each plot, LOC was fixed, and DOC and the lip opening area were varied. A number at the left or right side of a curve represents DOC in cm unit. Each curve consists of 20 data points representing the lip opening

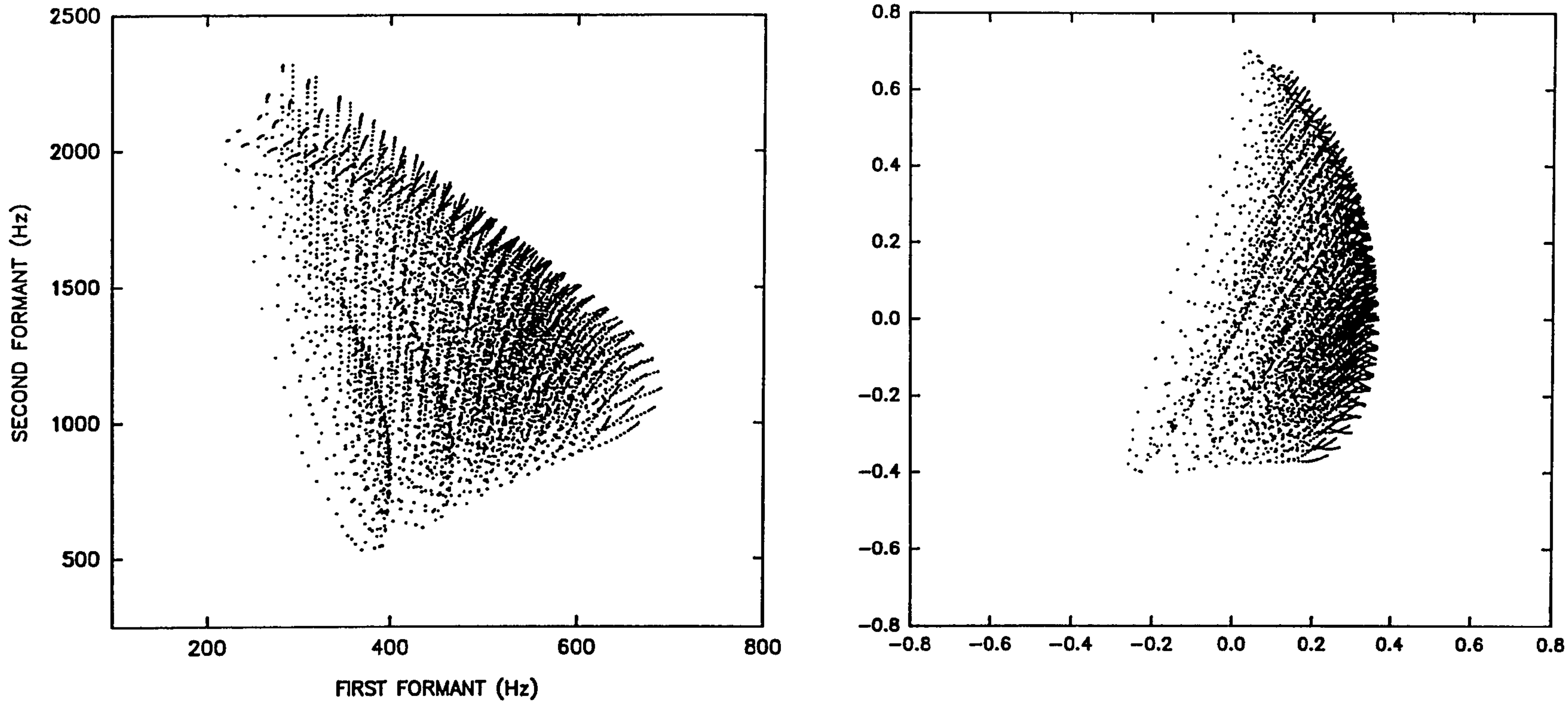


Fig. 22. Scatter plot of 6,420 vocal tract shapes in the formant space (left) and in MAPS (right).

(a)

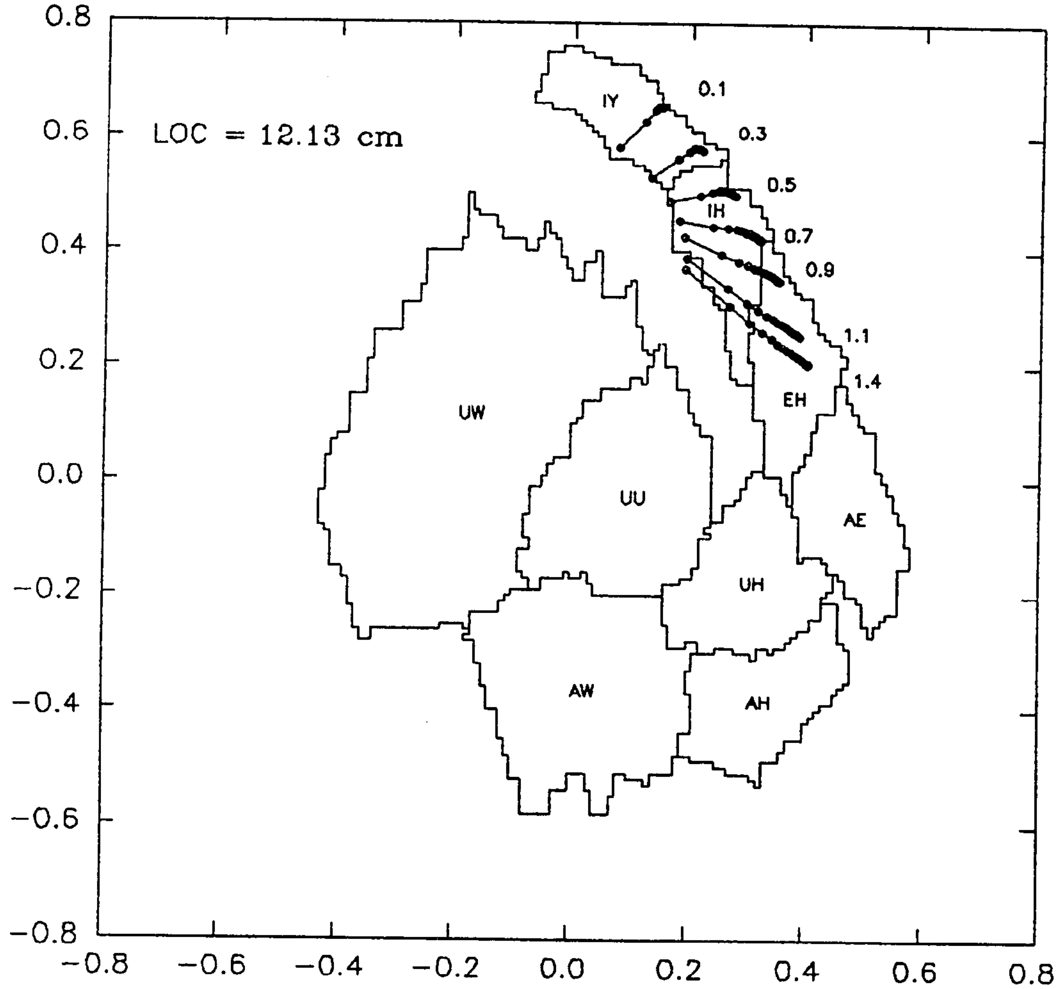
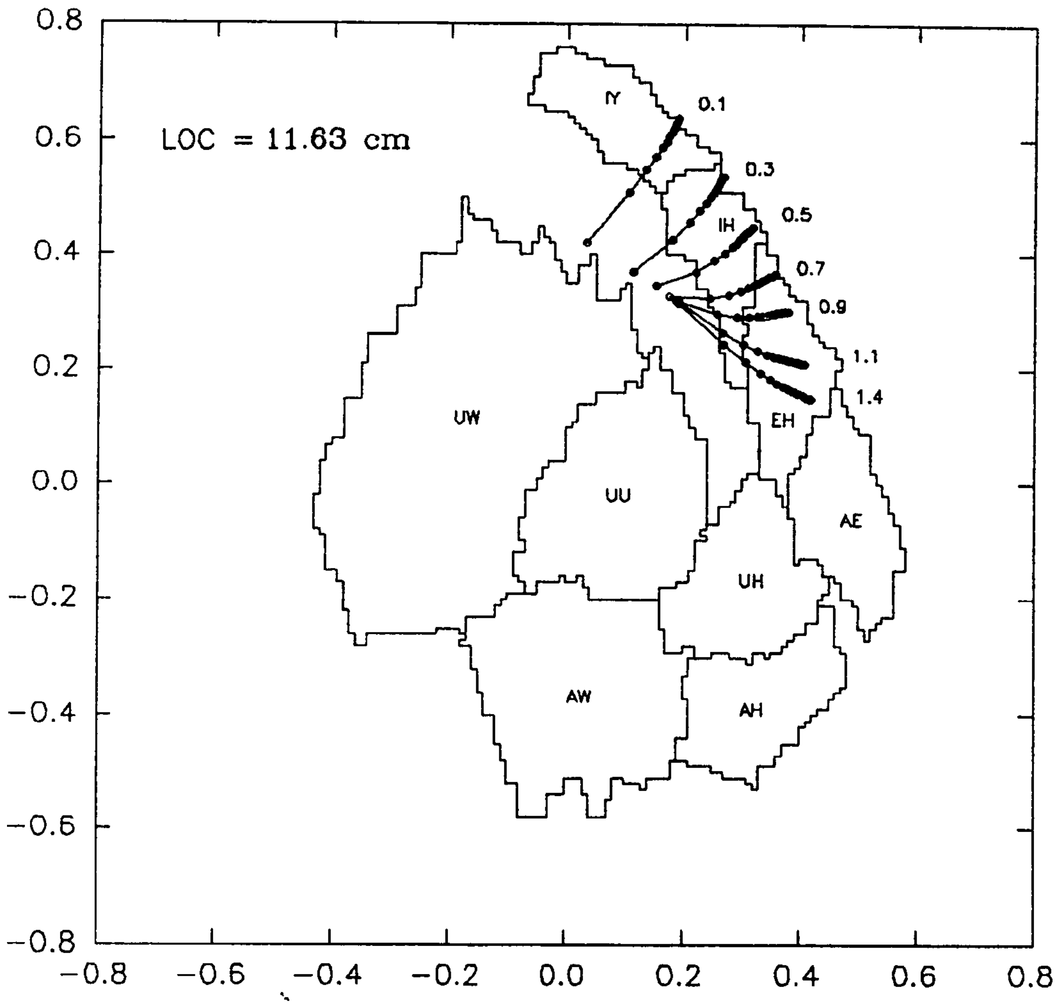
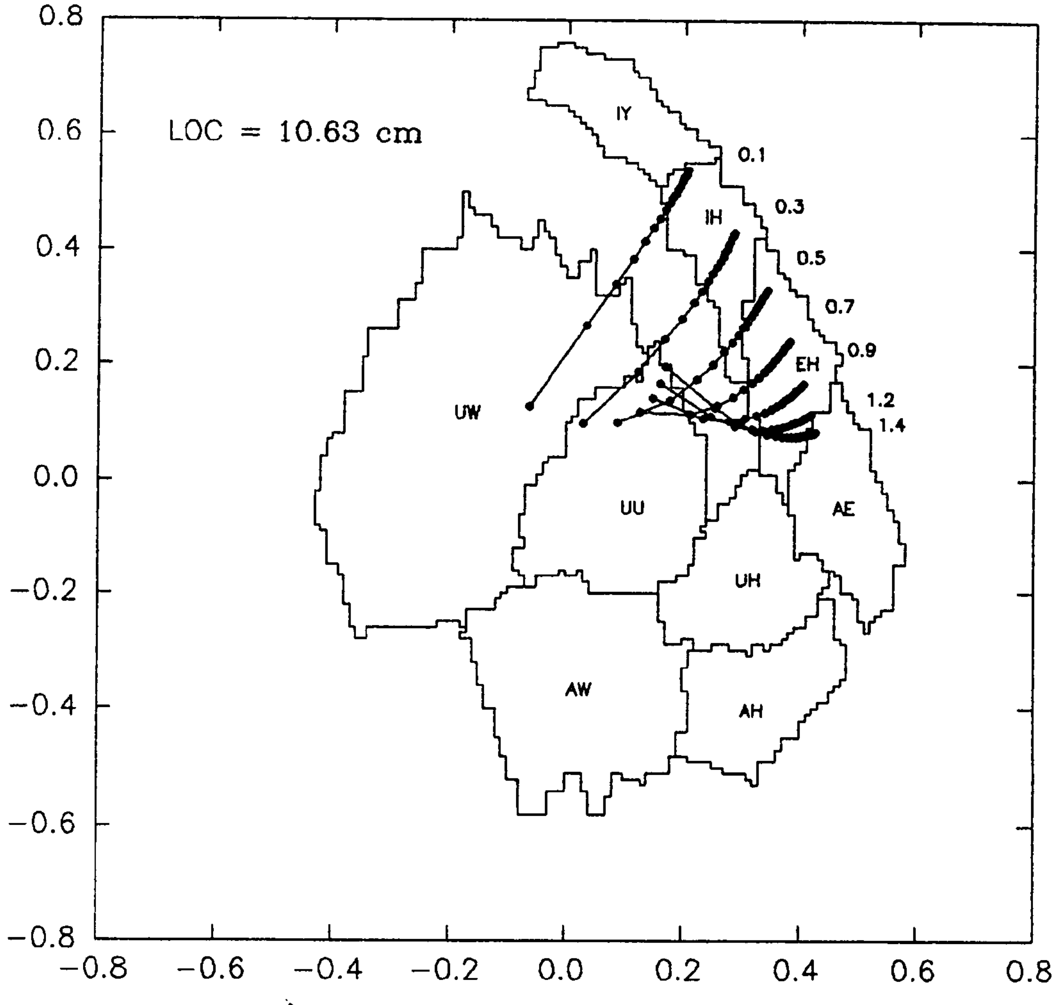


Fig. 23. Articulatory trajectories in MAPS. (a) LOC=12.13 cm, (b) 11.63 cm, (c)10.63 cm, (d) 9.63 cm, (e) 7.63 cm, and (f) 5.63 cm, respectively. Numbers at left or right end of curve is DOC in cm units. The lip opening area is varied from 0.3 cm^2 (circle at left end of curve) to 6.0 cm^2 (right end).

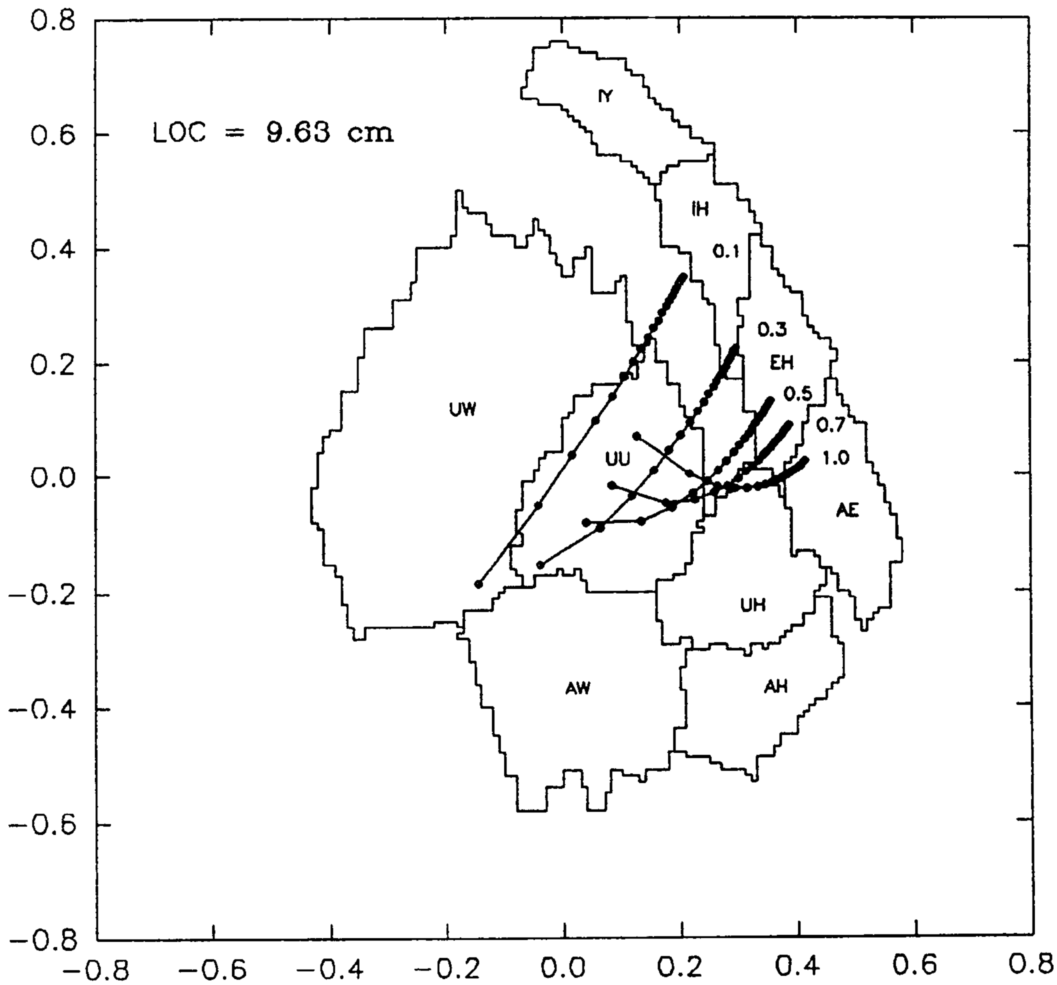
(b)


LOC = 11.63 cm
IY
IH
EH
AE
UW
UU
UH
AW
AH
0.1
0.3
0.5
0.7
0.9
1.1
1.4
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
−0.8
−0.6
−0.4
−0.2
0.0
0.2
0.4
0.6
0.8

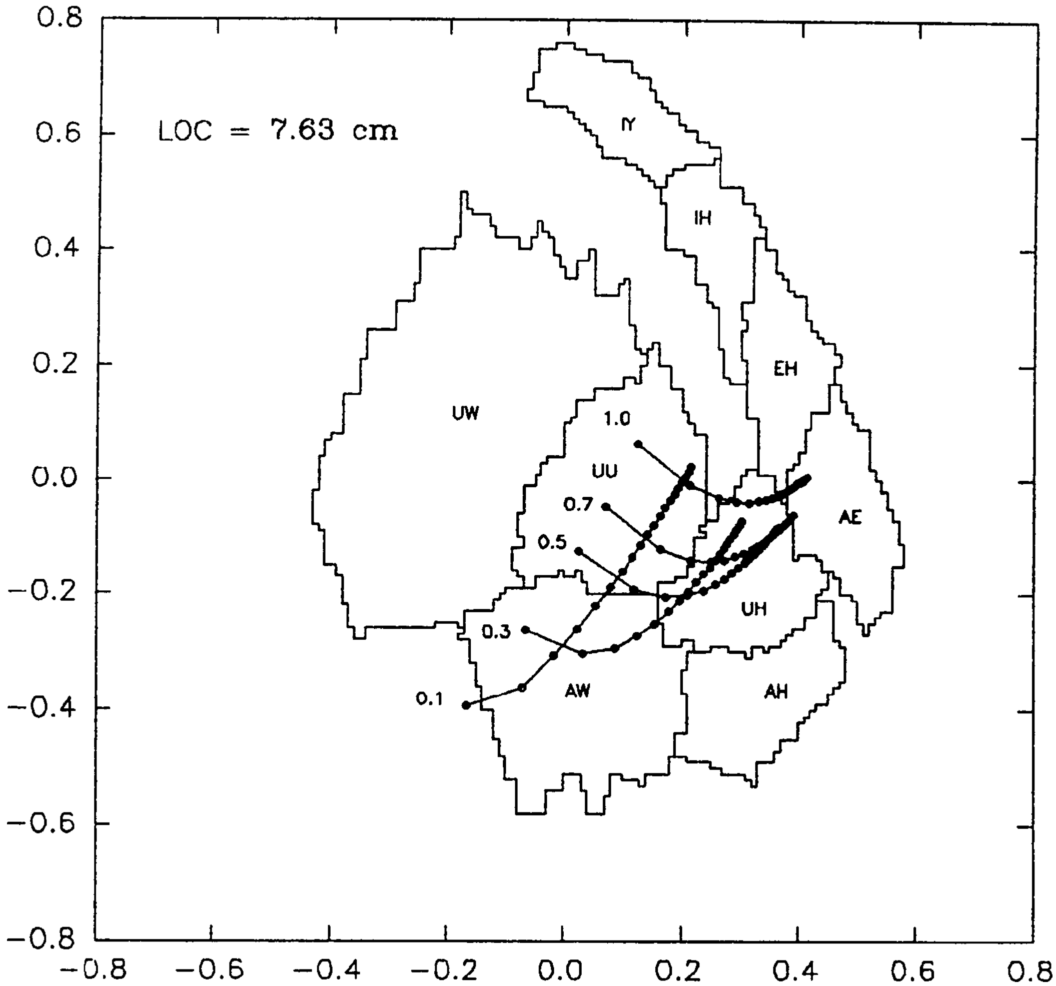
(c)


LOC = 10.63 cm
IY
IH
EH
UW
UU
AE
UH
AW
AH
0.1
0.3
0.5
0.7
0.9
1.2
1.4
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
−0.8
−0.6
−0.4
−0.2
0.0
0.2
0.4
0.6
0.8

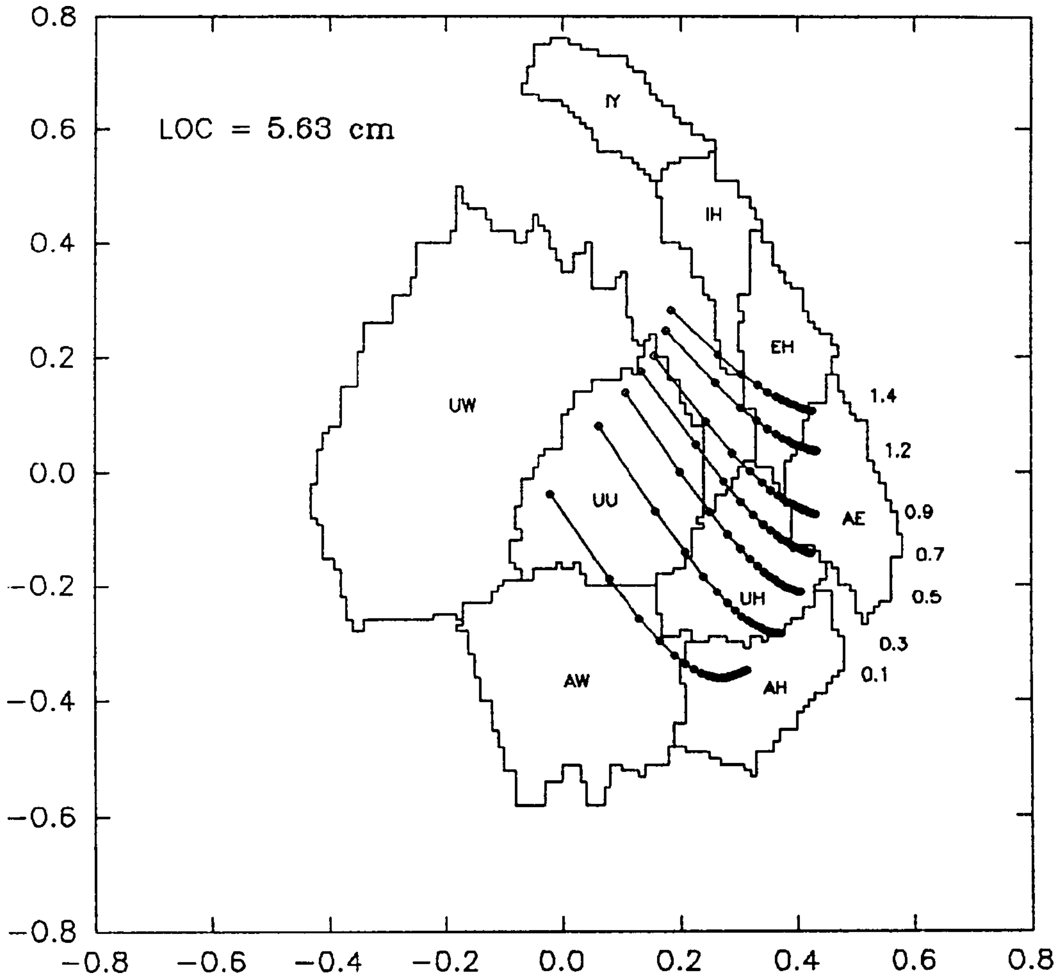
(d)


LOC = 9.63 cm
IY
IH
0.1
0.3
EH
0.5
0.7
1.0
UW
UU
AE
UH
AW
AH
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
−0.8
−0.6
−0.4
−0.2
0.0
0.2
0.4
0.6
0.8

(e)


LOC = 7.63 cm
IY
IH
EH
UW
UU
AE
UH
AW
AH
1.0
0.7
0.5
0.3
0.1
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
−0.8
−0.6
−0.4
−0.2
0.0
0.2
0.4
0.6
0.8

(f)


LOC = 5.63 cm
IY
IH
EH
UW
UU
AE
UH
AW
AH
1.4
1.2
0.9
0.7
0.5
0.3
0.1
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
−0.8
−0.6
−0.4
−0.2
0.0
0.2
0.4
0.6
0.8

area which is varying from 0.3 cm^2 (left end of a curve) to 6 cm^2 (right end) in 0.3 cm^2 step.

From the series of figures, it is evident that the boundaries of the vowel target zones can be crossed by varying the three articulatory parameters. For convenience of description, the 6,420 generated vocal tract shapes were classified into each vowel category using the same procedure described in the section III.C.2. LOC, DOC, and the lip opening area of vocal tract shapes in each vowel category were averaged. The results are listed in Table IV. It is also graphically represented in Fig. 24. In the figure, the radius of a circle was determined to be proportional to the lip opening area.

Vowels /IY/ and /AH/ were distinguished from the other vowels by the narrow DOC and relatively restricted range of LOC. Vowel /IH/ and /EH/ were distinguished from /IY/ mainly by DOC irrespective of LOC (see Fig. 23(a) - (c)). The role of the lip opening area may not be important for the perceptual contrast of the three vowels. Vowel /UW/ was distinguished from /UU/ by the relatively small DOC and lip opening area (see Fig. 23(c) - (d)). Also, vowel /UW/ has the least magnitude of the lip opening area among all the vowels. Vowels /AW/ and /UH/ were mainly distinguished from /AH/ by the larger DOC and more broad range of LOC (see Fig. 23(e) - (f)). The effect of the lip opening area could not be ignored for the perceptual contrast between these back vowels. It was observed that vowel /AE/ was characterized by the biggest DOC and the lip opening area among vowels.

Table IV. Averaged values of articulatory parameters of vocal tract shapes classified into each vowel category. Numbers in parenthesis are standard deviations.

Vowels	LOC (cm)	DOC (cm)	Lip opening area (cm^2)
/IY/	12.2 (0.33)	0.18 (0.07)	3.2 (1.58)
/IH/	11.8 (0.67)	0.58 (0.27)	3.3 (1.61)
/EH/	10.8 (2.00)	1.04 (0.26)	4.2 (1.23)
/AE/	5.6 (0.16)	1.16 (0.07)	5.3 (0.47)
/AH/	5.9 (0.20)	0.40 (0.07)	4.9 (0.78)
/AW/	7.3 (0.94)	0.32 (0.16)	2.7 (1.53)
/UH/	6.8 (1.08)	0.77 (0.20)	4.0 (1.28)
/UU/	8.6 (1.30)	0.54 (0.30)	2.8 (1.79)
/UW/	9.5 (1.76)	0.33 (0.29)	0.9 (0.95)

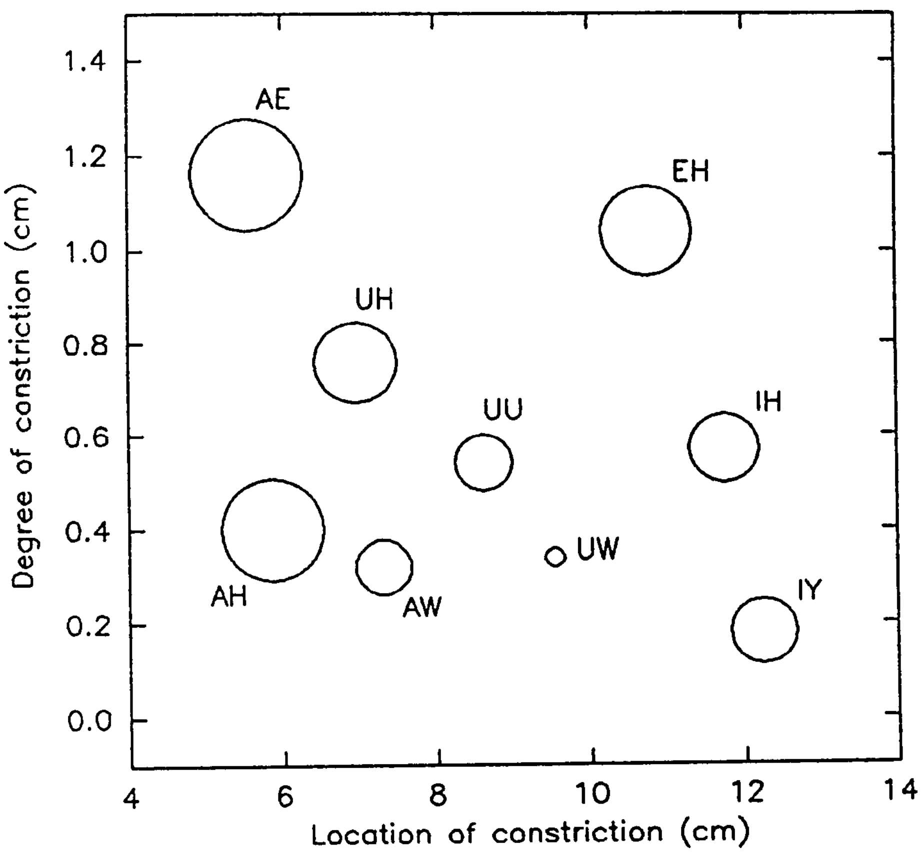


Fig. 24. Average values of LOC, DOC, and the lip opening area of the simulated vocal tract shapes classified into each vowel target zone. Sizes of circles were determined to be proportional to the lip opening.

To test the above observations, we measured the three articulatory parameters from the MR images obtained from two male subjects during the production of vowels /IY/, /EH/, /AE/, /AH/, and /UW/ (see chapter 2). The result is shown in Fig. 25 with the average values of three articulatory parameters obtained from simulated vocal tract shapes classified into the corresponding vowel categories. Because the vocal tract lengths measured from MR images were different across the vowels, they were normalized to have the same lengths of 16 cm. It can observe that the overall relationships between LOC, DOC, and the lip opening area of simulated vocal tract shapes reasonably agree well with those of the actual vocal tract shape, particularly for the vowels /IY/, /AH/, and /UW/. This suggests that the simulated vocal tract shapes were reasonably classified into corresponding vowel categories by the vowel target zones. This supports the hypothesis that the vowel target zones can be interpreted as articulatory targets zones as well as perceptual ones.

2. Interpretation of vowel articulation in MAPS

To interpret the position of a data point in MAPS in terms of the three articulatory parameters, we need to know the correlation between the articulatory parameters and the coordinate values (X,Y) of a data point. As a way to do this, we performed a regression analysis with the data of 6,420 vocal tract shapes. The result is shown in Table V.

From the table, we can observe that the three articulatory features of vowels, LOC, DOC and the lip

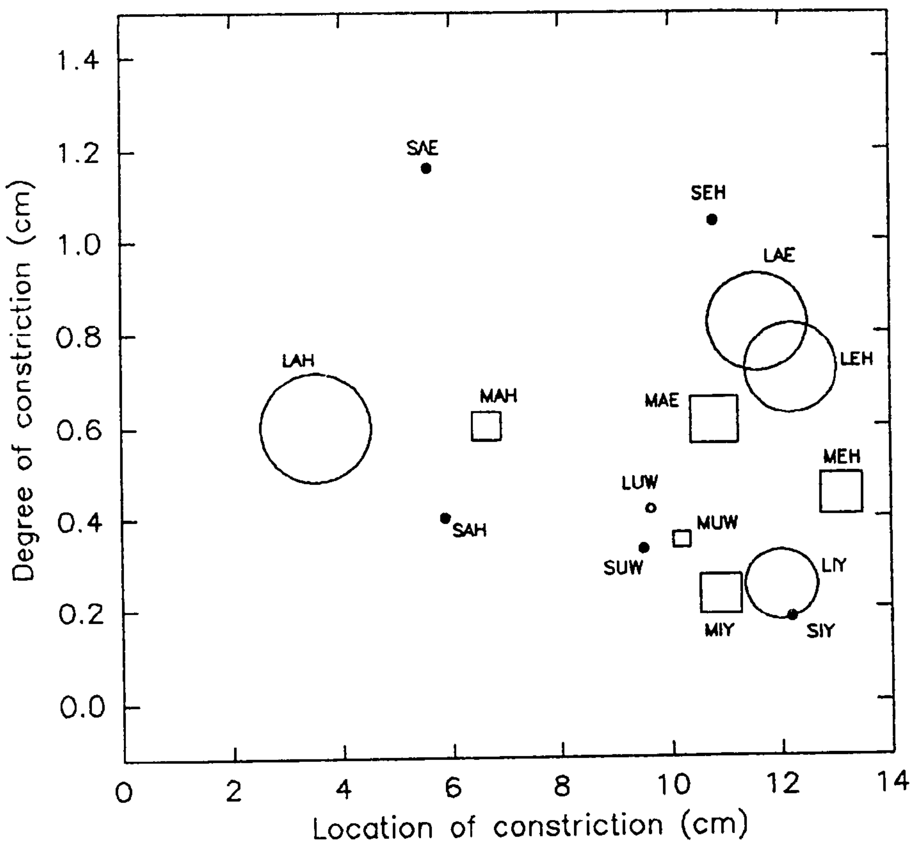


Fig. 25. Comparison of articulatory parameters between simulated vowel articulations and actual vowel articulations obtained from MR images. Dots, squares, and circles represents vowels corresponding to the simulated vocal tracts (S), subject MJM (M), and subject LSB (S), respectively. The sizes of circles and squares were determined to be proportional to the lip opening area.

Table V. Correlation between articulatory parameters and the X and Y coordinates in MAPS. Correlation between the articulatory parameters and the formant frequencies are also shown. Numbers are correlation coefficients between variables.

Articulatory parameters	MAPS		Formant space		
	X	Y	F1	F2	F3
LOC	.17	.80	-.66	.69	-.25
DOC	.67	.09	.30	.22	-.62
Lip opening area	.50	.02	.42	.36	.10

opening area were relatively well separated in MAPS. The LOC was mainly correlated with the Y-axis (r=0.80), not the X-axis (r=0.17). The DOC was mainly correlated with the X-axis (r=0.67) and there was almost no correlation between the DOC and the Y-axis (r=0.09). The lip opening area was weakly correlated with the X-axis (r=0.5) and there was almost no correlation between the lip opening area and the Y-axis (r=0.02). Thus it is relatively simple to interpret positions of data points in MAPS in terms of the three articulatory parameter: the coordinate value of the Y-axis determines the LOC, the coordinate value of the X-axis determines the DOC, and the lip opening area. Therefore, the articulatory characteristic of a position in MAPS can be reasonably determined by the combination of the two coordinate values, although they can not explain all the variations in vowel articulation in MAPS in terms of the three articulatory parameters.

For the purpose of comparison, it may be interesting to investigate the correlation between the three articulatory parameters and formant frequencies. We have observed that LOC was strongly correlated with the coordinate value of the Y-axis (r=0.80). This suggests that one parameter may explain major portions of variations of LOC among vowels in MAPS. Whereas, in the formant space, the LOC was almost equally correlated with F1 (r= -0.66) and F2 (r=0.69). This implies that the acoustic effect of LOC is split into the two parameters, and either one can not be a good representative of the LOC in the formant space. We have also observed that the DOC was moderately correlated with the X-axis in MAPS (r=0.67). However, in the formant space, it was mainly correlated with F3 (r= -0.62), not F1 (r=0.30) or F2 (r= 0.22). This implies that the tongue height of front vowels can not be represented properly in the F1-F2 space. At best, the tongue height is more correlated with F1 (r=0.30) than F2 (r=0.22).

From the above results, we may conclude that MAPS is best characterized by the LOC, DOC, and the lip opening area, and it is a better articulatorily-oriented space than the F1-F2 space in terms of the three articulatory parameters. Also, this suggests that the series of transformations used to construct MAPS emphasize the articulatory information contained in formant data.

C. Discussion

We have inductively shown that relevant articulatory information contained in acoustic data can be extracted by eliminating the acoustic variations caused by inter-speaker differences with an appropriate transformation of formant data. This is interesting because the original purpose of the transformation was to reduce the overlap in the formant space and thus to improve perceptual contrasts between vowels in the transformed space. The reason for doing so resides in the fact that the transformation also extracts appropriate articulatory dimension contained in the formant data and redistributes the data points to align them with the extracted articulatory dimensions in the transformed space. Such a trend can be clearly observable in the Fig. 22 in which the distribution of the simulated vocal tract shape were represented in the formant space and MAPS. There are two differences: (1) In MAPS, the right edge of the distribution is approximately parallel to the Y-axis and the down edge to the X-axis, and what we have shown is that the Y-axis is strongly correlated with the LOC, and the X-axis is moderately correlated with the DOC. There is no such a trend in the formant space; (2) In MAPS, the vowel /UW/ occupies a mid position of the left edge and it yielded a proper position of LOC with respect to vowels /IY/ and /AH/ along the Y-axis. As shown previously, it also yielded proper perceptual distances between vowels. However, in the formant space, /UW/ occupies the lower left-hand corner and we can not obtain such an ordering without appropriate transformations.

The above observations may confirm our hypothesis that the articulatory information contained in acoustic data of vowels can be more explicitly represented in a perceptual space than in a formant space. What we have also shown was that the format-ratio theory and its elaborated version by Miller (1989) have an inherent capability of extracting both articulatory and perceptual characteristic of vowels from acoustic data by reducing the acoustic variations caused by the inter-speaker differences. This supports our hypothesis that the more we eliminate the inter-speaker differences in acoustic signals, the better we can extract inherent articulatory information of vowels. In that sense, we agree with the view that speech perception take places in a specialized phonetic mode which is narrowly adapted for the efficient production and perception of phonetic structure, and the phonetic mode is not auditory but gestural (Mattingly and Liberman, 1986), although a question that remains is that how the gestural information contained in acoustic signals is decoded by the auditory system of humans. We also believe that what we have to seek in acoustic signals of speech is not acoustic invariance but gestural information. However, it is interesting to think about how the two different views on speech perception by humans can be integrated. This is because they emphasize different levels on the process of human speech perception.

It has well been known that when we represent formant frequencies of vowels measured from a single speaker in F1-F2 space, there is almost no overlap (Potter and Steinberg, 1950; Peterson and Barney, 1956). The observations illustrate the usefulness of formant space for the representation of acoustic properties of vowels. However, in terms of vowel articulation, the fact may not guarantee that the formant space can correctly reflect articulatory aspects of vowels because the axes themselves of the formant space are not good representatives of articulatory dimensions as shown in this study. The extrinsic normalization procedures (e.g., Gerstman, 1968; Lobanov, 1970; Neary, 1978) rescale the formant space to maximize the distances between vowel groups in the new scaled formant space. However, for the reason mentioned above, the rescaling of the axes may not contribute much to the articulatory contrasts among vowels but rather distort or eliminate phonetically relevant acoustic variations (i.e., vowel-inherent acoustic differences). That may be the reason why they are not effective for the comparison of vowel systems between languages (Disner, 1980).

Miller (1989) has also attempted to relate a position in APS with the tongue height and backness by performing another rotation of the xyz dimensions of APS. The underlying idea seems to demonstrate the versatility of APS. However, such a rotation may not be needed because the axes in SLAB space have already been aligned with appropriate articulatory dimensions (i.e., LOC and DOC) and, therefore, another rotation is likely to change correct articulatory dimensions into inappropriate ones.

V. Linguistic validity of MAPS

In order to compare vowels within or between languages, it is necessary to factor out the acoustic variations due to the inter-speaker differences before comparing. This is the only way to minimize a possibility that inter-speaker variations exceed phonetically relevant vowel- or language-specific variations. The reduction of acoustic variations is usually accomplished by speaker normalization procedures. However, there is a possibility that the normalization procedure is so effective that they even remove phonetically relevant vowel or language specific acoustic variations. This would be undesirable in a cross-language study. After evaluation of linguistic validity of several extrinsic normalization procedures, Disner (1980) has also shown that such procedures had poor linguistic validity because of the use of scaling factors which remove or distort phonetically relevant acoustic variations between languages. Therefore, as Disner concluded, it should be ensured that variations which remain in the normalized data are truly linguistic factors, not artifacts of the normalization technique itself.

The normalization procedure used in the current study is naturally free of such a defect. This is because it is strictly based on individual acoustic data, not any scaling factor. Also, a space derived by the procedure, MAPS, has explicit articulatory characteristics in that positions in MAPS are mainly determined by the three articulatory parameters, DOC, LOC, and the lip opening area. This is important because the vowel or language specific acoustic variations are nothing other than systematic differences in vowel articulations in a single language or between languages.

MAPS was constructed based on the observation by Miller (1989) that a vowel slab exists in the vowel system of American English, such that variation in the depth dimension is much smaller than that of the height and width dimensions (see Fig. 16). Therefore, to use MAPS as a tool for the between-language comparison, there should exist similar constraints in other languages.

In general, a vowel slab can be expressed by a constraint

$$x + y + z = C \pm d, \quad (22)$$

where x, y, and z are the coordinates of a data point in APS (Miller, 1989), C is a middle position of the vowel slab, and d is deviation (or depth dimension). To verify this condition, the vowels of five Germanic languages were selected and tested: English (Peterson and Barney, 1952), German (Jorgensen, 1969), Dutch (Pols et al., 1973), Swedish (Fant et al., 1969), and Danish (Fischer-Jorgensen, 1972). Vowel data sets for each language consisted of the first three formants measured from multiple male speakers: 33 men in English; 6 men in German; 50 men in Dutch; 6 men in Swedish; and 8 men in Danish. When the fundamental frequency was not given in a data set, the value of 132 Hz was used, which was an average value of male speakers in English data set. The result is shown in Table VI.

Existence of a vowel slab seems to be universal, that is, the depth dimensions (d) is always much smaller than the height and width dimensions. The middle positions of the vowel slabs (C) were not much different across languages. It is expected that if the number of subjects were enough in each language, the vowel slabs could almost overlap with each other in APS. The existence of vowel slabs across languages may be attributed to a fundamental fact that speakers have similar vocal organs and use similar articulatory gestures to produces similar vowel sounds irrespective of languages spoken. Therefore, the elimination of the depth dimension

Table VI. Equations of the middle positions of vowel slabs in APS. Numbers in parentheses are the depth dimen sions of the vowel slabs expressed by two-sigma values.

Language	Equation of vowel slab
English	$x + y + z = 1.204$ (0.105)
German	$x + y + z = 1.177$ (0.117)
Dutch	$x + y + z = 1.204$ (0.090)
Danish	$x + y + z = 1.198$ (0.132)
Swedish	$x + y + z = 1.204$ (0.094)

which was used for the construct MAPS, can be generalized to other languages. This supports the validity of using MAPS to compare vowel systems across languages.

We demonstrated below that MAPS is a useful space for the within- and between-language comparison of vowels with a few examples. For the comparison, we used the core zones rather than the entire vowel target zones because we believe that the core zones are good representatives of the entire zones. To represent core zones, the 2-sigma radius ellipse was used.

It has been noted that the front vowels of Danish are unevenly distributed in the vowel space, with [/IY/, /E/, and /EH/] relatively high and close to each other, and the vowel /AE/ comparatively low and further away (Fischer-Jorgensen, 1972). Another observation is that Danish /IY/ is tenser than the English /IY/ (Walshe, 1965; p.90). To test these two observations in MAPS, the corresponding vowel data sets were represented in MAPS as shown in Fig. 26. We can observe that the three Danish vowels, /IY/, /E/, and/EH/ are tightly distributed and the Danish vowel /AE/ is well separated from the other three vowels. This supports the observation by Fischer-Jorgensen. We interpret the term "tenser" to mean a smaller DOC and a more advanced LOC in the case of front vowels. Then the relative positions of the tenser Danish /IY/ and the less tenser English /IY/ is well supported in Fig. 26.

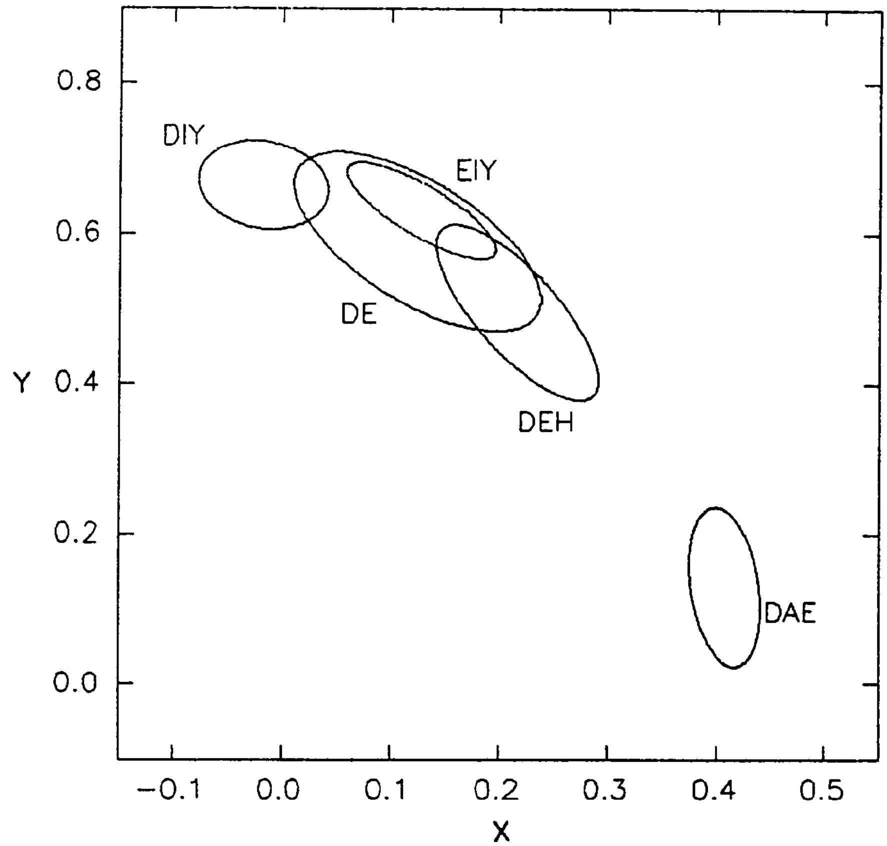


Fig. 26. Comparison of the tongue positions between Danish vowels /IY/ (DIY), /E/ (DE), and /EH/ (DEH). Danish vowel /IY/ (DIY) is also compared with English vowel /IY/ (EIY).

It has also been noted that Dutch /EH/ is more open than English /EH/ and intermediate between English /EH/ and /AE/ (Koolhoven, 1968; p.5). The vowels are represented in MAPS as shown in Fig. 27. If we interpret the term "open" as larger DOC because they are front vowels, Dutch /EH/ is well located between English /EH/ and /AE/.

As a final example, consider German /IY/ and /Y/. The vowel /Y/ is a front-rounded vowel absent in American English. We represent the vowels in the formant space and in MAPS as shown in Fig. 28. In the formant space, the positions of the vowels are almost the same in the ordinate (F1 dimension) and are well separated in the abscissa (F2-F1 dimension). In terms of the height and backness of the tongue, this means that there is no difference in the tongue height (i.e., no difference in DOC) and the tongue body is retracted for the vowel /Y/. However, after examination of several published x-ray profiles of /Y/ including German as shown in Fig. 29, Wood (1986) concluded that the tongue body position of /Y/ is not retracted but slightly lower than /IY/, yielding a wider tongue-palate distance than /IY/. It is evident that the observation by Wood is correctly reflected in MAPS, not in the formant space.

From the above examples, we may conclude that at least the tongue positions of vowels in terms of the location and degree of tongue constriction are correctly represented in MAPS. It is thought that another dimension is needed to accommodate the effect of the lip configuration in MAPS.

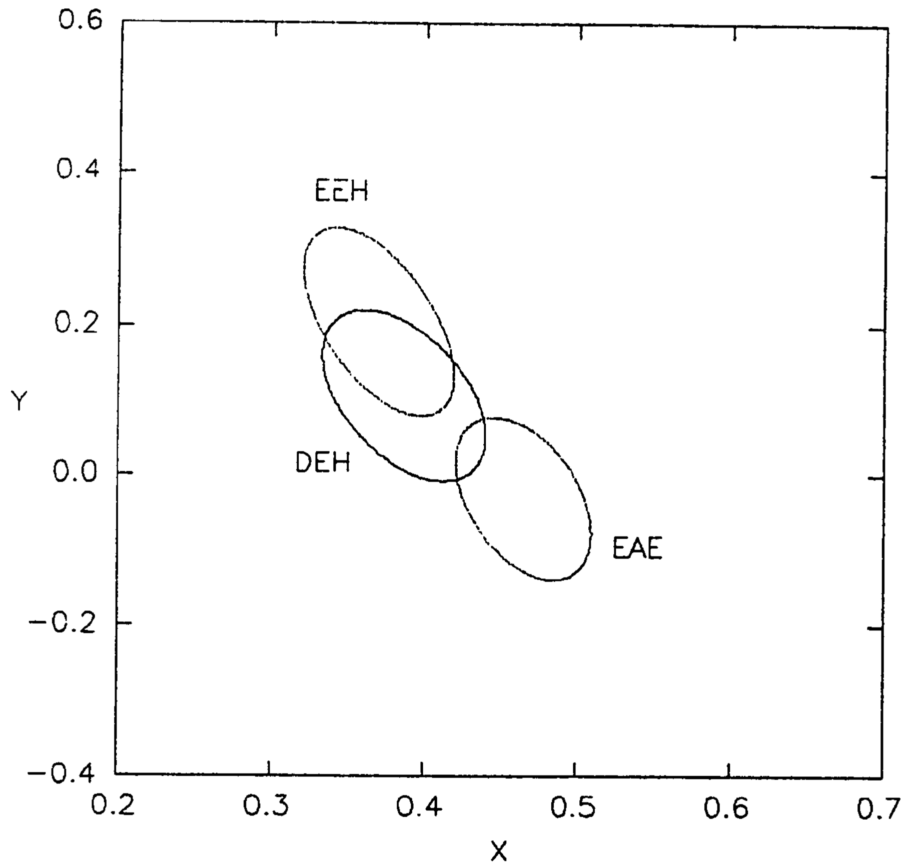


Fig. 27. Dutch vowel /EH/ (DEH) is compared with English vowels /EH/ (EEH) and /AE/ (EAE) in MAPS.

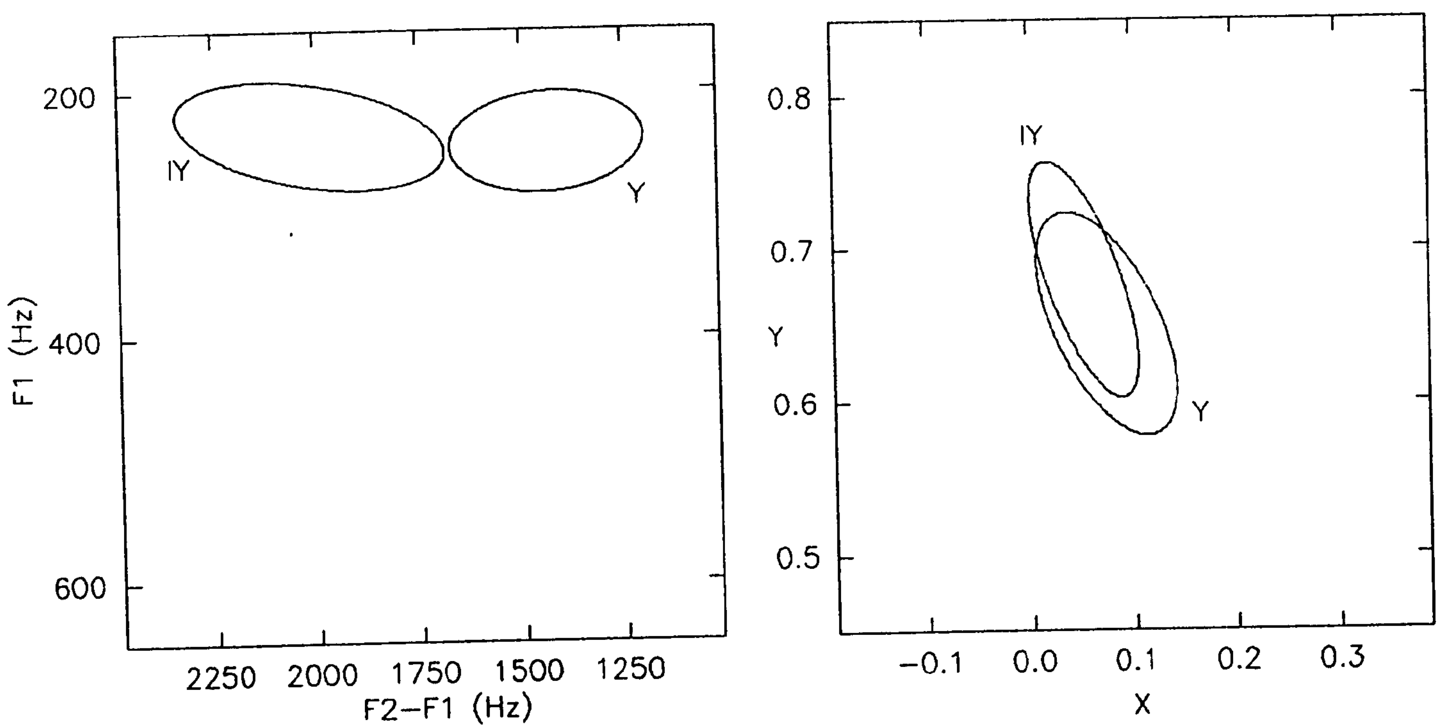


Fig. 28. German vowels /IY/ and /Y/ are compared. It shows that the tongue position of the vowel /Y/ is correctly reflected in MAPS.

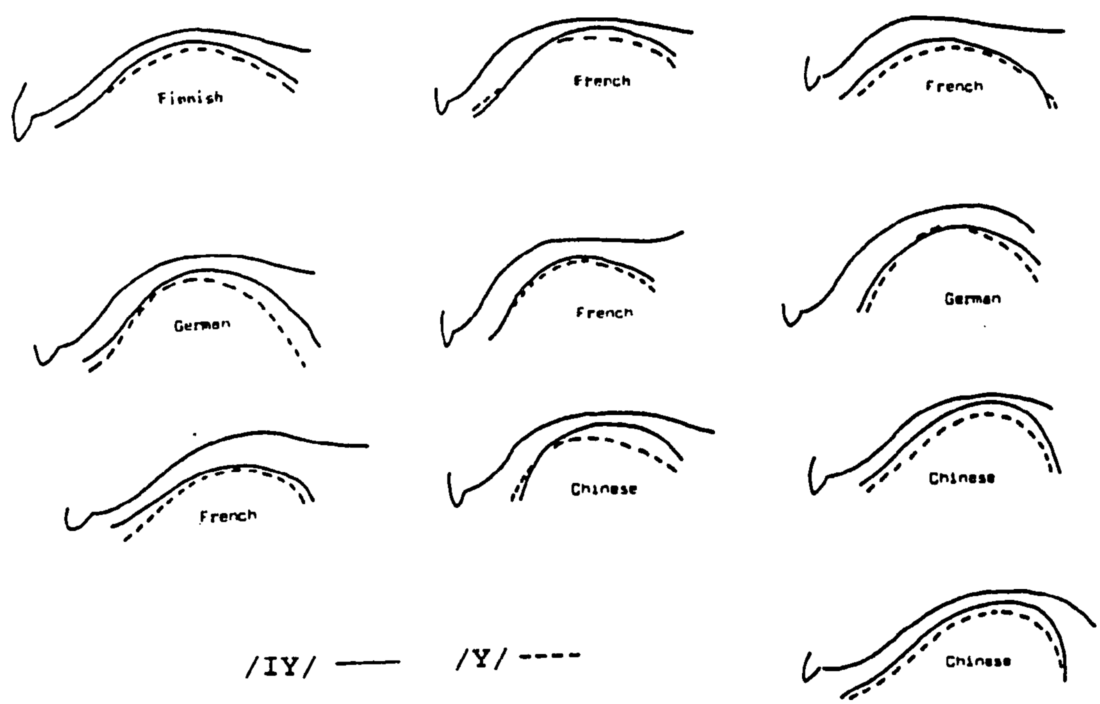


Fig. 29. X-ray images of /Y/ obtained from several languages including German. The images suggest that the tongue position is not retracted but slightly lower than /IY/ (from Wood, 1982).

VI. Conclusions

Articulatory information contained in acoustic data were effectively extracted by the use of the format-ratio theory elaborated by Miller (1989) which eliminates acoustic variations caused by the inter-speaker differences. This may be due to the fact that the transformation also extracts appropriate articulatory features embedded in acoustic data and distributes the data points to align with the extracted articulatory features. This is an important difference between MAPS and the traditional formant space. MAPS was best characterized by the location and degree of tongue constriction and the lip opening area.

It was shown that MAPS is a useful space for the comparisons of vowel systems between languages. This results from the fact that the articulatory differences between vowels are emphasized by representing corresponding acoustic data in MAPS, and that MAPS has inherent articulatory characteristics which are conserved across languages. MAPS could be an ideal phonetic space.

CHAPTER 4. ON THE QUANTAL NATURE OF VOWEL ARTICULATION

I. Introduction

Quantal theory of speech (Stevens, 1989) states that articulatory-acoustic relations are quantal in the sense that an acoustic pattern shows a change from one state to another as an articulatory parameter is monotonously varied through a range of values. This implies that there are regions in the vocal tract where perturbation in an articulatory parameter results in relatively small changes in acoustic parameters (e.g., formants), and other regions where relatively small perturbation of the parameter induces substantial acoustic changes. To illustrate this situation, Stevens (1989) used a schematic diagram shown in Fig. 30. The figure shows that there are regions, I and III, where an acoustic parameter remains relatively stable (i.e., the curve is relatively flat in those regions) when small modifications are made in an articulatory parameter, and the two regions are separated by the intermediate (or transitional) region II where abrupt changes of the acoustic parameter occur. The main point of the figure, and therefore the quantal theory, is that the stable regions are separable with respect to each other, that what is important is not nominal changes in value of the acoustic parameter between regions I and III but

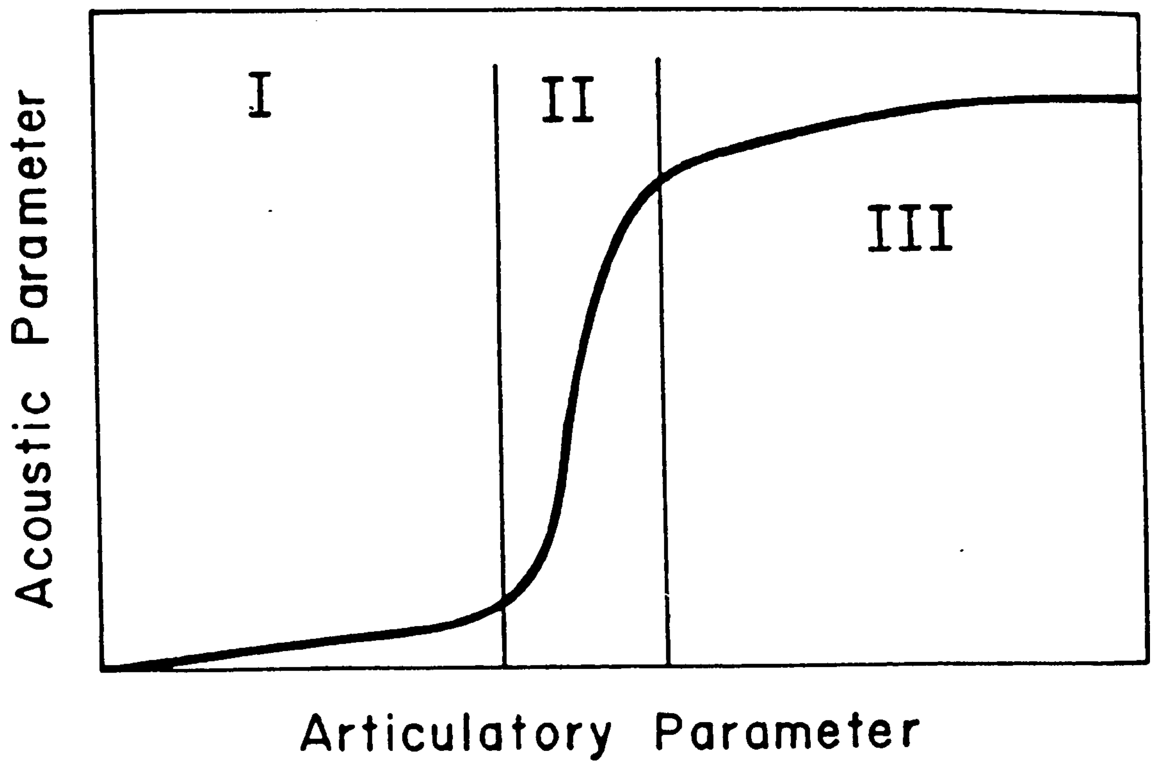


Fig. 30. Schematic diagram by Stevens (1989) to illustrate the quantal nature of speech. Regions I and III are stable regions, and region II is a transitional region (from Stevens, 1989).

auditory (or perceptual) contrasts accompanied by such changes, and that the stable regions in the vocal tract may be preferred by speakers because the stability may allow less articulatory precision at such regions, although it is uncertain how such preferences can be linked to the muscular anatomy of vocal organs or physiological convenience. The quantal theory hypothesizes that stable regions play important roles in the selection of places of articulations or distinctive features in languages. Based on the argument which is later extended to acoustic-auditory (or perceptual) relations, the quantal theory aims to answer a fundamental question in phonetics about what articulatory, acoustic, and auditory (or perceptual) attributes determine the inventory of sounds in languages, and where they come from.

The view of the quantal theory is quite reasonable in the sense that the ultimate purpose of articulation in speech is to change acoustic patterns of sound that will induce perceptual distinctions or contrasts. Such contrasts are not much meaningful if there do not exist some kind of acoustic or perceptual boundaries between the sounds. However, there have been some critics about the way in which Stevens introduced the concept of acoustic stability.

A typical vocal tract configuration offered by Stevens to demonstrate acoustic stability is shown in Fig. 31(a). The simple tube model of the vocal tract is composed of three uniform sections and the total length of the tube is 16 cm. In one example, while keeping the length of the constriction

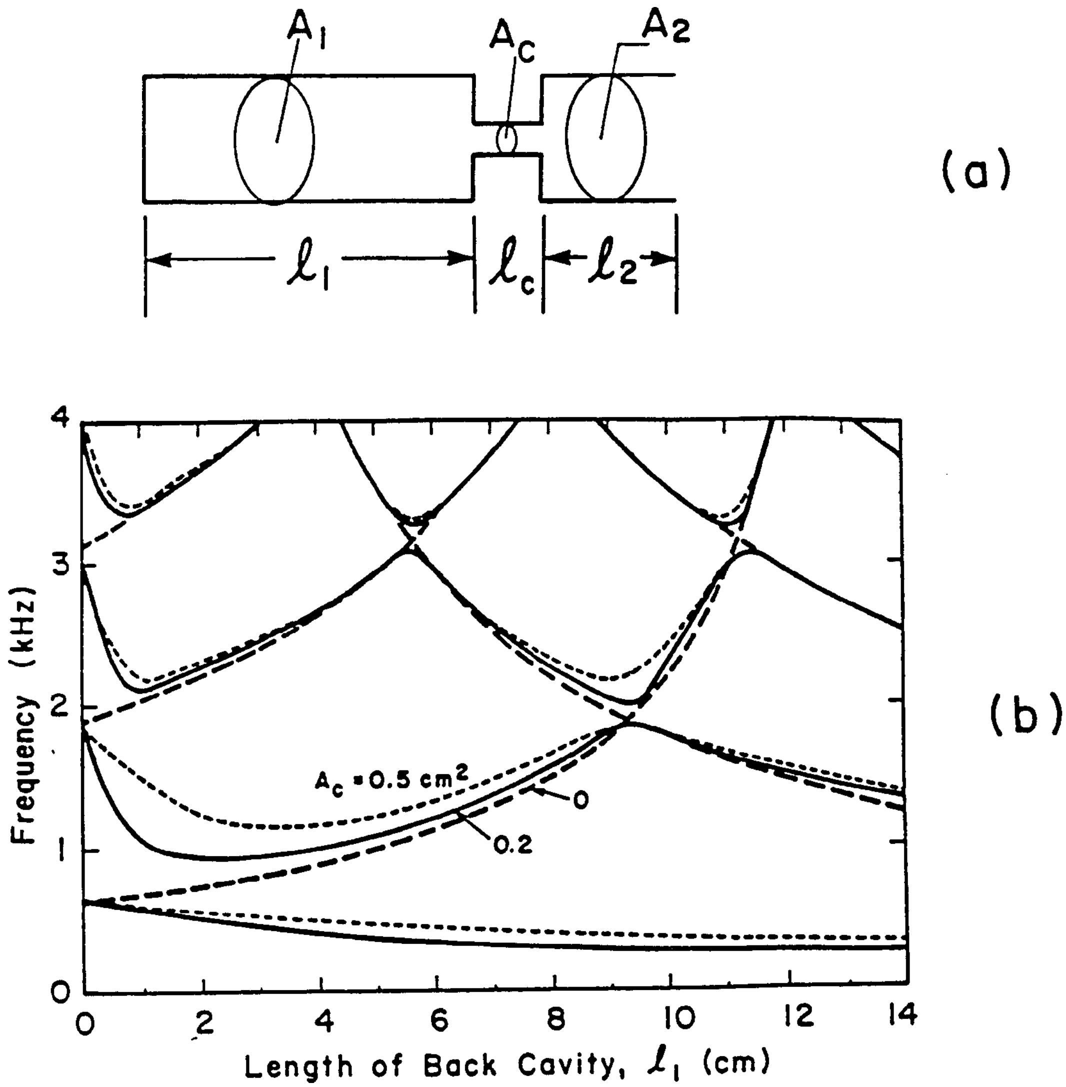


Fig. 31. (a) A tube model of the vocal tract used to demonstrate the acoustic stability. The total length $l_1+l_2+l_c$ = 16 cm, the constriction length l_c = 2 cm, and the cross-sectional areas A_1 and A_2 are 3 cm^2. (b) Computed formant frequencies for the tube model during the manipulation of the back cavity length l_1. The solid and short-dashed line are for A_c = 0.2 and 0.5 cm^2, respectively. According to the quantal theory, stability is achieved where the proximity of adjacent two formant frequencies occurs (from Stevens, 1989).

(l_c) and the total length (l_1+l_c+l_2) constant, he increased the back cavity length (l_1) monotonously and computed corresponding formants based on a double Helmholtz resonator model (Fant, 1960). The manipulation is like that the constriction position moves forward from the glottis toward the lip. The resulting formant values are shown in Fig. 31(b) as a function of the back cavity length l_1. Stevens claimed that there exist three acoustically stable regions where the second and third formant achieve a maximum and minimum value (at l_1 = 9.3 cm), respectively, and the third formant achieves minimum values (at l_1 = 5.5, 11.2 cm). He performed similar experiments with several vocal tract configurations corresponding to several vowels by changing the tube parameters (i.e., length and cross-sectional area) of each uniform section of the tube model and obtained similar results. Based on the results, Stevens concluded that the acoustic parameters (formants) are relatively insensitive to small change in constriction positions at the constriction locations where particular formants show minimum or maximum values or the proximity of adjacent two formant frequencies occurs:

> Thus as the articulatory states undergo a continuous sequence of maneuvers toward and away from the target value, the acoustic parameter resulting from this articulatory gesture may remain relatively stable over some part of this sequence. Furthermore, the precision with which the target articulatory state is achieved may be rather lax." (p.5, Stevens, 1989)

A common question raised by the conclusion has been "How stable is stable?". The question mainly comes from the fact that Stevens has been primarily concerned with one articulatory parameter, the back cavity length or location of the tongue constriction, to demonstrate acoustic stability. Clearly, even in the simplified tube model of the vocal tract as illustrated in Fig. 31(a), there are two other parameters, the cross-sectional area of the front tube (A_2) corresponding to the lip opening area and the length of constriction (l_c), which could affect the locations where acoustic stability occurs. Lindblom and Engstrand (1989) have shown that the second and third formants are substantially varied with the change in the constriction length l_c at a position $l_1 = 9.3$ cm (see Fig. 31(b)) where, as demonstrated by Stevens, the acoustic stability has been achieved during the manipulation of the back cavity length.

As can be noticed from the above example, a major criticism of the quantal theory has been that acoustic stability is observed as long as one examines variations along a single articulatory dimension. It may disappear when other dimensions are introduced. This indicates the possibility that there is no absolutely stable regions in the vocal tract under the perturbation of all possible combination of the articulatory parameters. Another problem is that since no quantitative definition of stability is given, there is some ambiguity as to how the selection criterion of the acoustic stability should be interpreted.

As an attempt to clarify the above situations, the aspects of acoustic stability in vowel production were addressed in this chapter in terms of the LOC, DOC, and the lip opening area. In addition, instead of using vague terms such as "formant maxima" or "formant minima", the rate of change in acoustic parameters (e.g., formants) was used to locate the regions where acoustic stability occurs.

II. Method

A. Data

In chapter 3, total 6,420 midsagittal vocal tract shapes (the vocal tract length is 16.0 cm) were generated based on the PARAFAC tongue model by Harshman et al. (1977). The vocal tract length was 16.0 cm. Two articulatory parameters, the location of constriction (LOC) and the degree of constriction (DOC) were extracted from the simulated vocal tract shapes. The resulting range of LOC was from 5.38 cm to 12.63 cm in 0.25 cm step. It approximately covers the region between the reference line 6 and the midposition between the reference lines 13 and 14 in the grid system shown in Fig. 32. DOC was varied in 0.1 cm step, but the range was dependent on LOC. The range of DOC was 0.1 - 1.5 cm in the maximal case and it was 0.1 - 0.5 cm in the minimal case. The lip opening area was separately varied from 0.3 cm^2 to 6.0 cm^2 in 0.3 cm^2 step, yielding 20 different lip opening area.

Corresponding formants were computed with a tube model of the vocal tract described in chapter 2. The formant data are graphically represented in Fig. 33 as a function of LOC with selected values of four DOC (0.1, 0.3, 0.5, 0.9 cm) and four lip opening areas (0.3, 0.9, 2.4, 6.0 cm^2), yielding 16

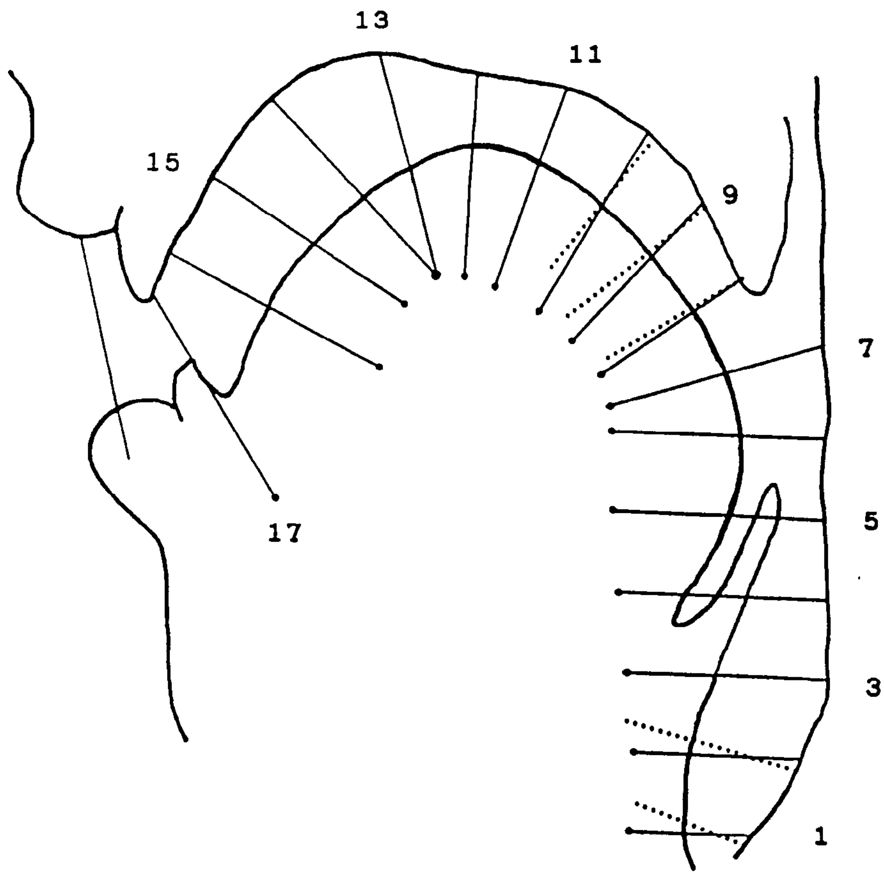
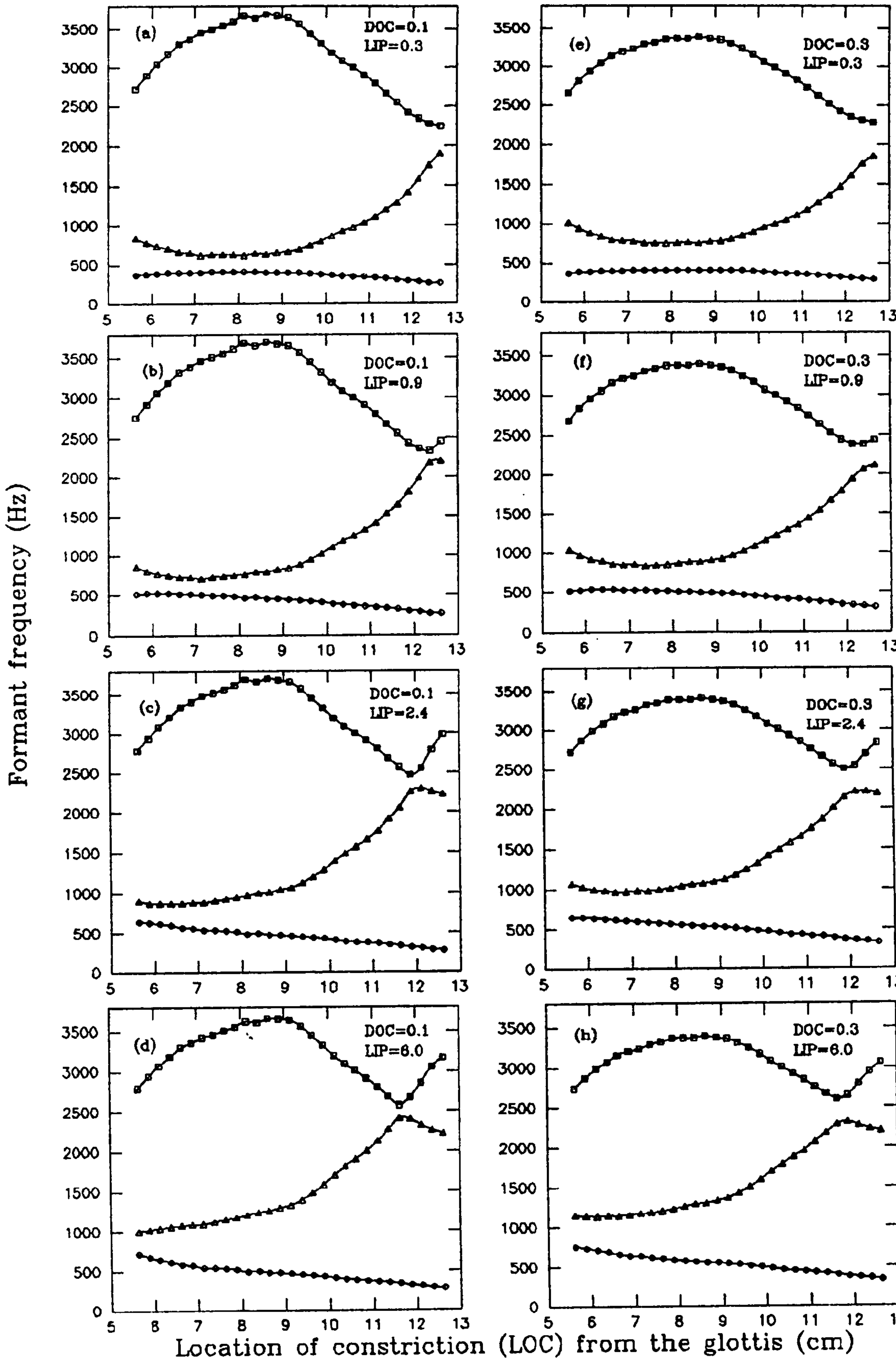
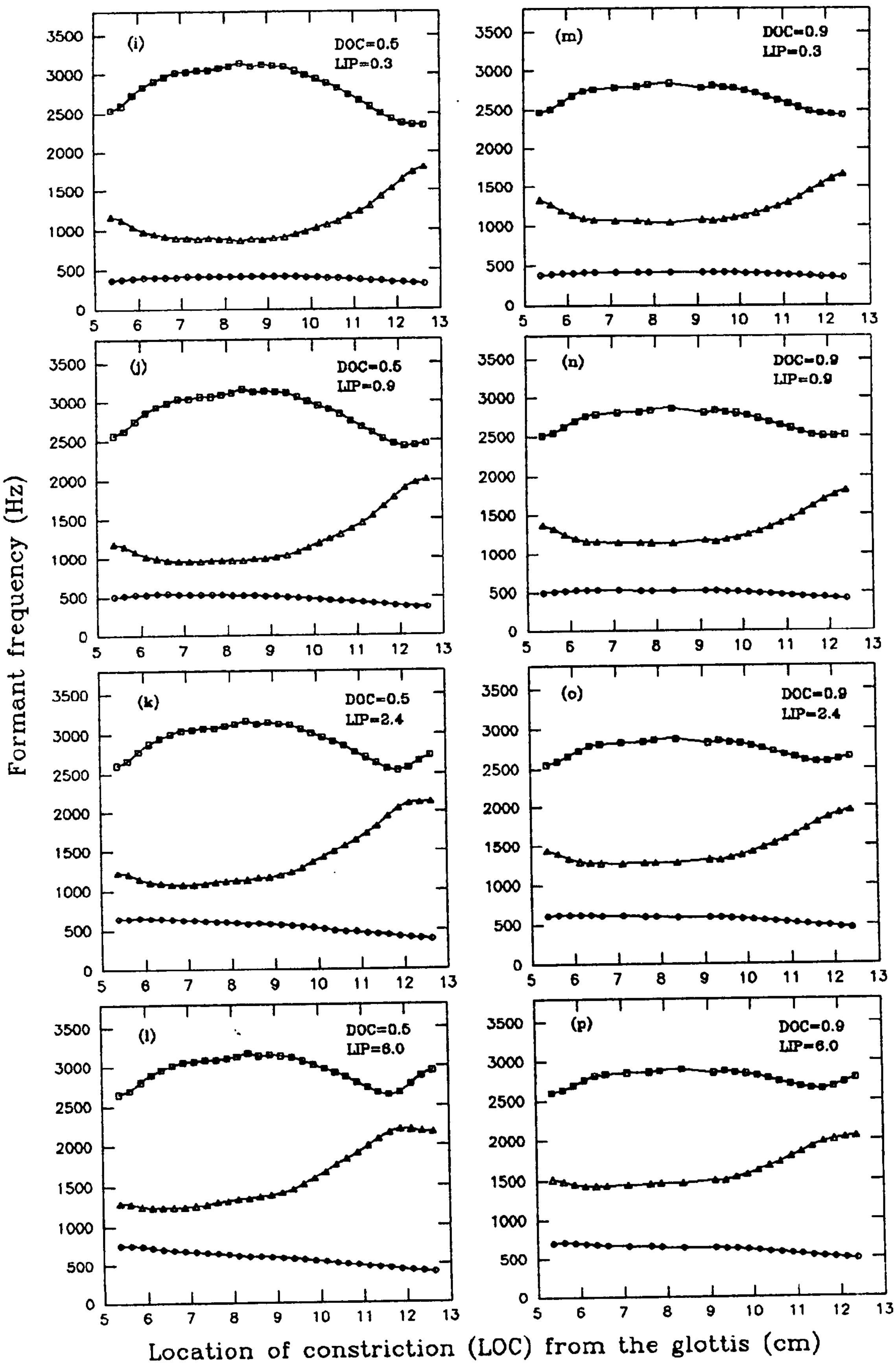


Fig. 32. Grid system used by Harshman et al. (1977). The grid system was used to refer to a region in simulated vocal tracts.

Fig. 33. Formant data obtained from simulated vocal tract shapes are plotted as a function of LOC (circle: F1, triangle: F2, square: F3). There are 16 plots ((a)-(p)). DOC (unit: cm) and the lip opening area (unit: cm^2) are fixed in each plot.


(a)
DOC=0.1
LIP=0.3
(b)
DOC=0.1
LIP=0.9
(c)
DOC=0.1
LIP=2.4
(d)
DOC=0.1
LIP=6.0
(e)
DOC=0.3
LIP=0.3
(f)
DOC=0.3
LIP=0.9
(g)
DOC=0.3
LIP=2.4
(h)
DOC=0.3
LIP=6.0
Formant frequency (Hz)
Location of constriction (LOC) from the glottis (cm)



(i) DOC=0.5 LIP=0.3
(m) DOC=0.9 LIP=0.3
(j) DOC=0.5 LIP=0.9
(n) DOC=0.9 LIP=0.9
(k) DOC=0.5 LIP=2.4
(o) DOC=0.9 LIP=2.4
(l) DOC=0.5 LIP=6.0
(p) DOC=0.9 LIP=6.0
Formant frequency (Hz)
Location of constriction (LOC) from the glottis (cm)

plots. The figure is similar to Fant's nomogram (Fant, 1960) except that the formants corresponding to each set of DOC and the lip opening area are plotted separately. In a plot, the distance between adjacent data points is 0.25 cm along the dimension of LOC (abscissa). For example, the value of LOC (abscissa) of the first data point in Fig. 33 (a)-(h) is 5.63 cm from the glottis and that of the next data point is 5.88 cm. Plots in each column or in each row in the figure have the same value of DOC or lip opening area. By observing the figures by column or by row, we can evaluate the effects of lip opening area or the DOC, respectively, on the formant frequencies. The formant data was used as a data base for the current study.

Because the degree of tongue constriction gradually changes at the center position of tongue constriction, there is a problem about how to determine the constriction length (e.g., l_c in Fig. 31(a)) there. Therefore, the constriction length has not been extracted from the simulated vocal tract shapes. However, to my knowledge, there has been no evidence that the constriction length is an important articulatory parameter for the production of vowels. There is also a question as to how much the manipulation of the constriction length is physiologically plausible with respect to the manipulations of LOC, DOC, or the lip opening area. Although the parameter has been used to demonstrate or criticize acoustic stability, its role on the perceptual contrasts between vowels may be minor in actual vowel articulations.

For these reasons, constriction length was not considered in the current study.

B. A selection criterion of acoustic stability

To determine the range or positions of acoustically stable regions more quantitatively, the rate of change (RC) was used. RC was defined as

$$RC = dF/dA, \qquad (23)$$

where dF is perturbations in acoustic parameters (e.g., formants) induced by an articulatory perturbation dA.

From the above definition, we can expect that RC nears zero, or is relatively small, at stable regions and it may change rapidly at other regions. Thus the center positions of acoustically stable regions (region I and III in Fig. 30) can be located by finding positions where RC achieves the absolute minimum or local minima during a manipulation of an articulatory parameter. Regions between the absolute minimum or local minima of RC can be regarded as transitional regions (e.g., region II in Fig. 30). We also redefined the stable positions as the positions of LOC where RC reaches the absolute minimum or local minima.

C. Investigation of acoustic stability during the manipulation of the constriction position

The purpose of the present analysis was to find regions in the vocal tract where acoustic stability occurs during a manipulation of LOC under the perturbations of DOC and the lip opening area. For the task, the rate of change $RC1_{i,j}$ at a current position j of LOC was defined as

$$RC1_{i,j} = ABS(F_{i,j+1} - F_{i,j-1})/5.0 \text{ (Hz/mm)}, \qquad (24)$$

where i (= 1, 2, 3) is an index for the first three formants and $F_{i,j+1}$ and $F_{i,j-1}$ represent the formants at the next and previous positions of LOC, respectively. The denominator represents the distance in mm between previous and next positions of LOC.

The manipulation of LOC was repeated with different values of DOC and the lip opening area. Selected values of DOC were 0.1 cm, 0.3 cm, 0.5 cm, and 0.9 cm, and those of the lip opening area were 0.3 cm^2, 0.9 cm^2, 2.4 cm^2, and 6.0 cm^2. Therefore, total 16 manipulations (4 DOC x 4 lip opening area) of LOC were performed and $RC1_{i,j}$ was computed at each position of LOC during each manipulation. To prevent local fluctuation of $RC1_{i,j}$, absolute values of $RC1_{i,j}$ were taken and smoothed by applying a three-point average technique once.

D. Investigation of acoustic stability during the manipulation of the degree of constriction

The purpose of this analysis was to find regions where formants are relatively insensitive to the variations of DOC. For the task, $RC2_{i,j}$ at a current position j of LOC was defined as

$$RC2_{i,j} = ABS(Fi_{max} - Fi_{min})/(Dj_{max} - Dj_{min}), \text{ (Hz/mm)} \qquad (25)$$

where Fi_{max} and Fi_{min} (i = 1, 2, 3) represent the first three formant when DOC is the maximum value (Dj_{max}) and the minimum (Dj_{min}) at LOC j, respectively.

To investigate the effect of the lip opening area to the acoustic stability achieved by the manipulation of DOC, the above computation was repeated four times with four different lip opening area of 0.3 cm_2, 0.9 cm_2, 2.4 cm_2, and 6.0 cm_2.

III. Results and discussion

A. Analysis of formant data

We first examined the formant plots shown in Fig. 33. It was observed that, in each plot, there exists at least one region where the proximity of F1 and F2 or the proximity of F2 and F3 occurred. In terms of the quantal theory, these regions are stable under the perturbations of LOC. However, when we increased the lip opening area while keeping the DOC constant (e.g., Fig. 33(a) - 33(d), DOC = 0.1 cm), the positions of LOC where the proximity of F1 and F2 (e.g., around LOC = 8.0 cm in Fig. 33(a)) and the proximity of F2 and F3 (e.g., LOC = 12.38 cm in Fig. 33(b)) occur moved toward the glottis. Other figures also showed the same tendency. This indicates that the positions where acoustic stability occurs during the manipulations of LOC are relatively determined depending on the lip opening area.

When the DOC was increased while keeping the lip opening area constant (e.g., Fig. 33(a), (e), (i), and (m), LIP = 0.3 cm^2), the positions of LOC where F2 reaches maximum or F3 reaches minimum (e.g., around LOC = 8.0 cm in Fig. 33(a)) were not changed much and thus stability occurred at such positions. However, in this instance, it was observed that F2 increased rapidly at those positions than at the others. F3 also decreased fast at those positions. This implies that, although stability was achieved at such positions of LOC under the perturbations of DOC, such positions were unstable in terms of formant values themselves.

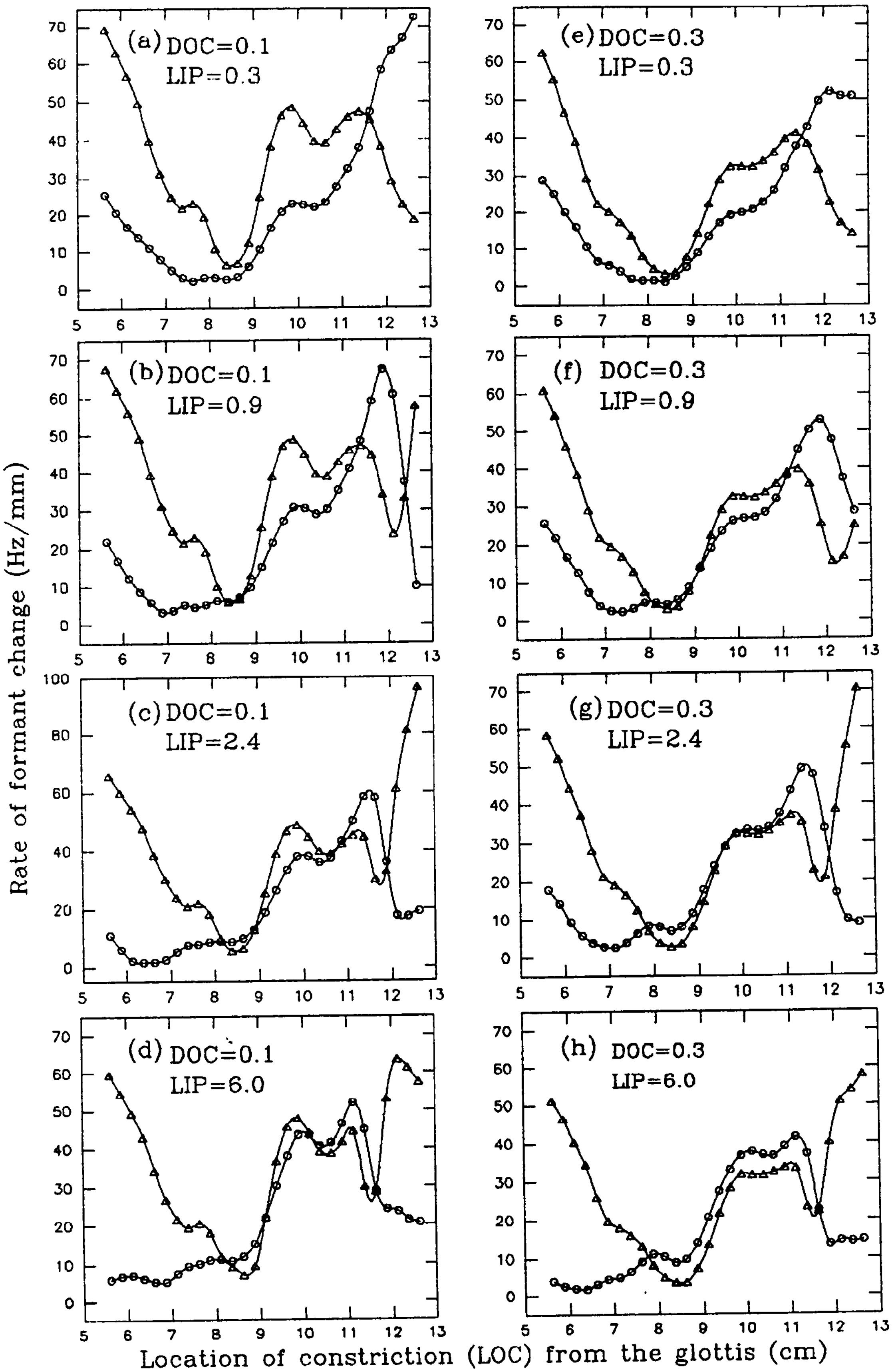
The above observations suggest that, although there exist stable regions, the locations of the stable regions changes when other articulatory parameters -- particularly the lip opening area -- vary.

B. Analysis of RC1 plots

RC1 values of F2 and F3 are plotted in Fig. 34 as a function of LOC. Those of F1 are plotted separately in Fig. 35. Plots in each column or in each row in each figure have a same value of DOC or the lip opening area. By observing the figures by column or by row, we can evaluate the effects of lip opening area or the DOC, respectively, on the regions where stability occurs during a manipulation of LOC.

The series of RC1 plots (Fig. 34(a) - (p)) were different with the corresponding formant plots (Fig. 33(a) - (p)) in several aspects. First, the exact position or range of the stable regions, which is hard to see in the formant plot, are easily located in the RC plots. Second, the positions of formant maxima and minima in the formant plots are divided into two groups in the RC plots: absolute minimum and local minima. Thirdly, the transitional regions (e.g., region II in Fig. 30) which are hardly to observe in the formant plots can be easily located in the RC plots. As an example, let's compare Fig. 34(b) and Fig. 33(b). It can be observed that the position where F3 reaches a maximum in the formant plot can be found exactly in the RC plot and it appears as the position of absolute minimum (LOC = 8.38 cm in Fig. 34(b)), and the position where F3 reaches minimum in the formant plot

Fig. 34. RC1 plots of F2 (circle) and F3 (triangle). There are 16 plots ((a) - (p)). DOC and the lip opening area are fixed in each plot.


(a) DOC=0.1
LIP=0.3
(b) DOC=0.1
LIP=0.9
(c) DOC=0.1
LIP=2.4
(d) DOC=0.1
LIP=6.0
(e) DOC=0.3
LIP=0.3
(f) DOC=0.3
LIP=0.9
(g) DOC=0.3
LIP=2.4
(h) DOC=0.3
LIP=6.0
Rate of formant change (Hz/mm)
Location of constriction (LOC) from the glottis (cm)

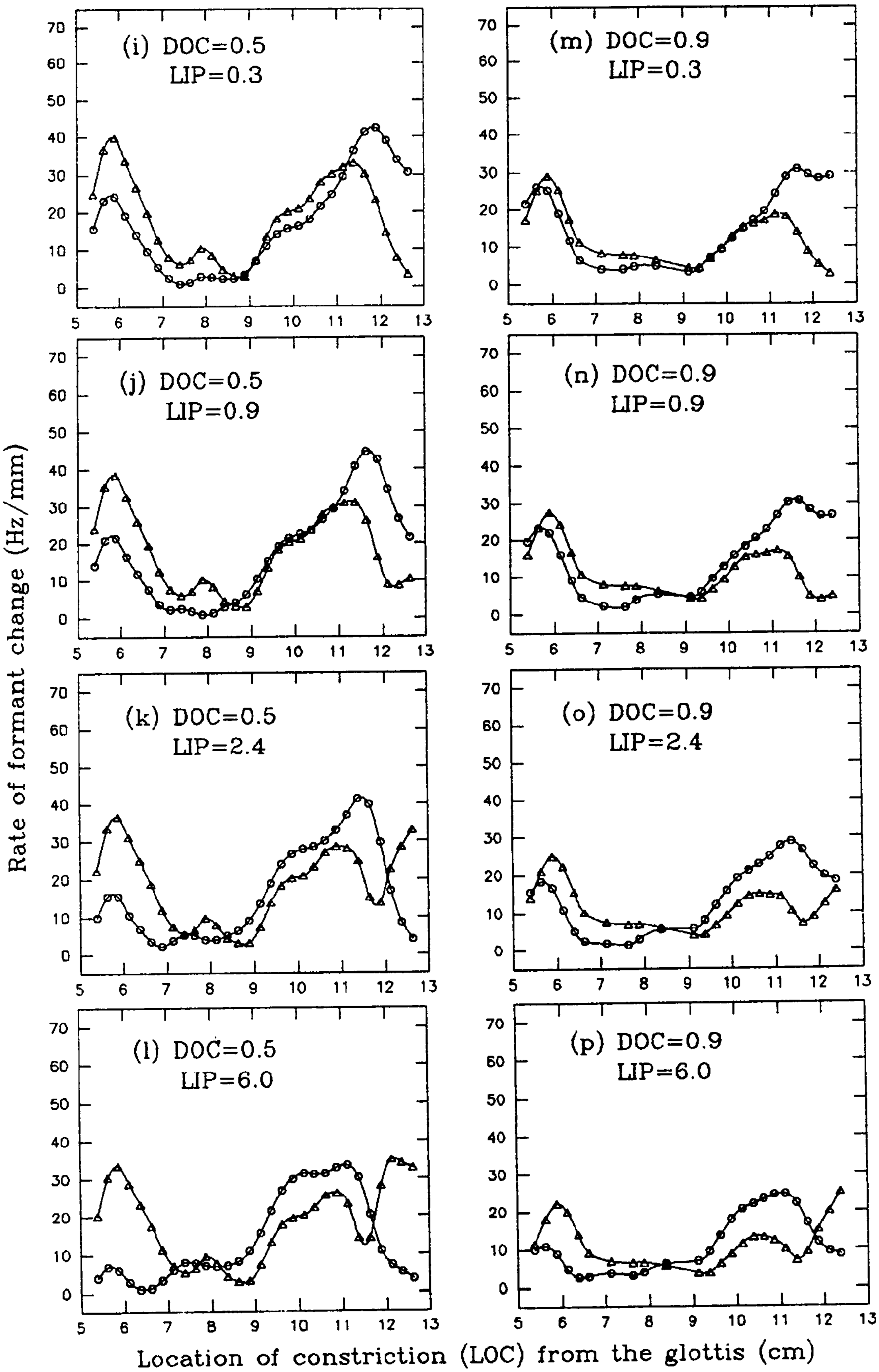
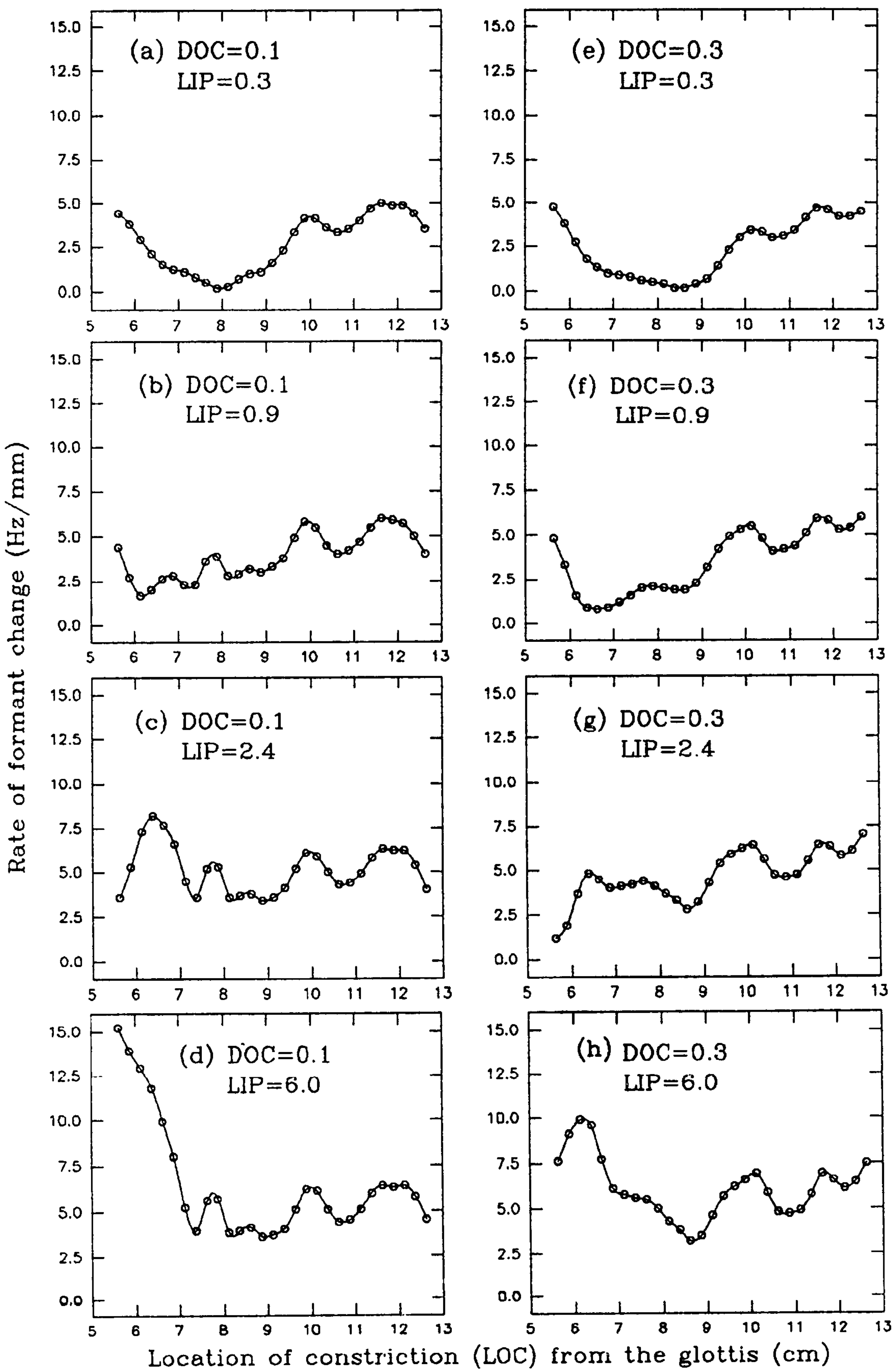
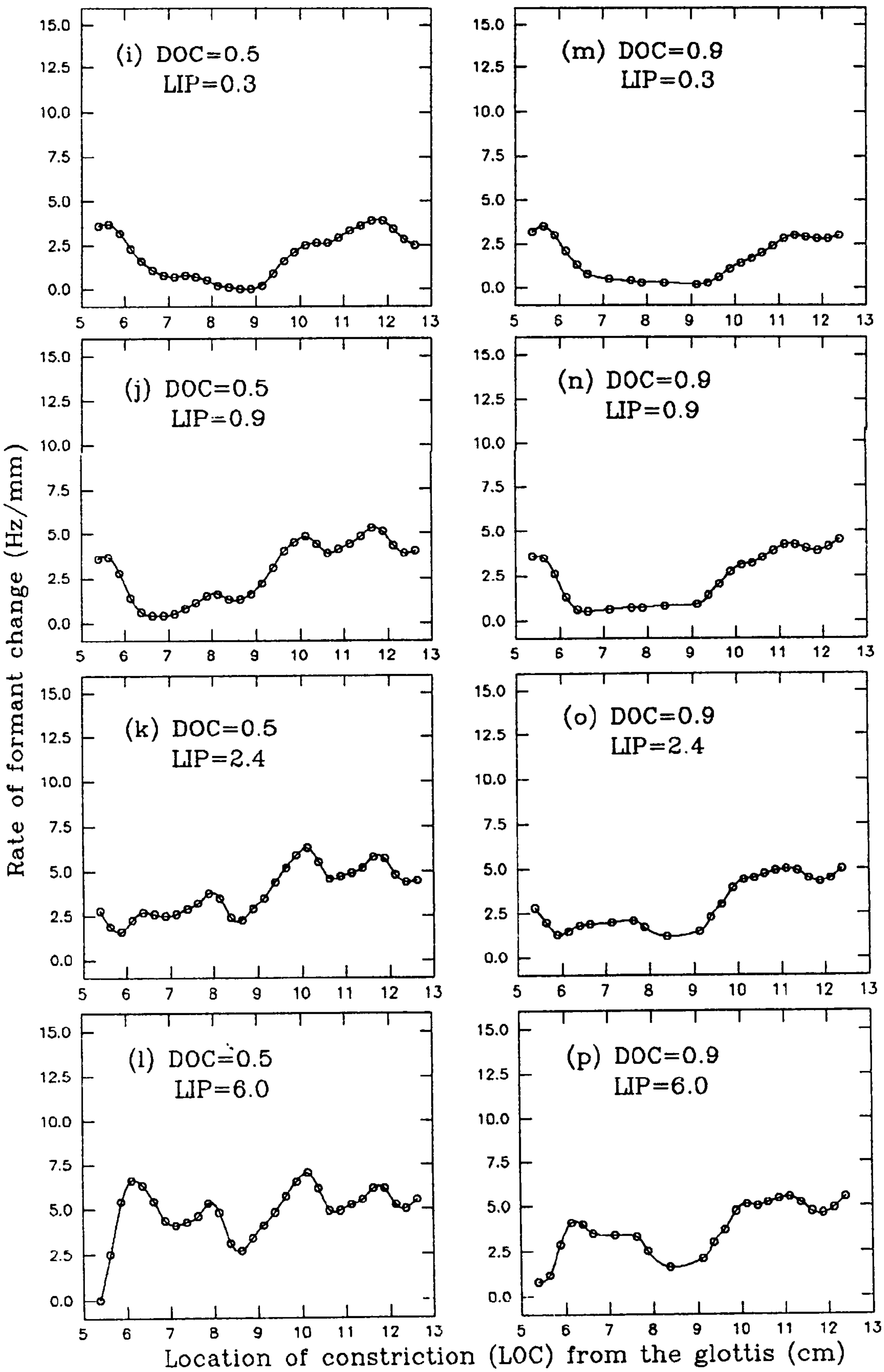

(i) DOC=0.5
LIP=0.3
(m) DOC=0.9
LIP=0.3
(j) DOC=0.5
LIP=0.9
(n) DOC=0.9
LIP=0.9
(k) DOC=0.5
LIP=2.4
(o) DOC=0.9
LIP=2.4
(l) DOC=0.5
LIP=6.0
(p) DOC=0.9
LIP=6.0
Rate of formant change (Hz/mm)
Location of constriction (LOC) from the glottis (cm)

Fig. 35. RC1 plots of F1. There are 16 plots ((a) - (p)). DOC and the lip opening area are fixed in each plot.


(a) DOC=0.1
LIP=0.3
(b) DOC=0.1
LIP=0.9
(c) DOC=0.1
LIP=2.4
(d) DOC=0.1
LIP=6.0
(e) DOC=0.3
LIP=0.3
(f) DOC=0.3
LIP=0.9
(g) DOC=0.3
LIP=2.4
(h) DOC=0.3
LIP=6.0
Rate of formant change (Hz/mm)
Location of constriction (LOC) from the glottis (cm)


(i) DOC=0.5 LIP=0.3
(m) DOC=0.9 LIP=0.3
(j) DOC=0.5 LIP=0.9
(n) DOC=0.9 LIP=0.9
(k) DOC=0.5 LIP=2.4
(o) DOC=0.9 LIP=2.4
(l) DOC=0.5 LIP=6.0
(p) DOC=0.9 LIP=6.0
Rate of formant change (Hz/mm)
Location of constriction (LOC) from the glottis (cm)
0.0
2.5
5.0
7.5
10.0
12.5
15.0
5
6
7
8
9
10
11
12
13

appears as the position of a local minimum (LOC = 12.13 cm in Fig. 34(b)) in the RC plot. It is also observed that the region where F2 achieves absolute minimum is somewhat broad (LOC = 6.88 - 8.63 cm in Fig 34(b)) in the RC plot. The RC plot also shows that there exists a new stable position (LOC = 10.38 cm in Fig. 34(b)) and a transitional region (around LOC = 8 - 9 cm in Fig. 34(b)) which are absent in the formant plot.

From the series of the RC plots (Fig. 34(a) - (p)), we can observe that the positions where RC1 of corresponding formant frequencies reaches absolute minima or local minima change depending on DOC or the lip opening area. However, the mode of variations seems to be dependent on the position along the LOC dimension and formant frequency considered. For example, when we increase the lip opening area while keeping the DOC constant (e.g., Fig. 34(a) - (d), DOC = 0.1 cm), we can observe that the positions of LOC where RC1 of F2 reaches the absolute minimum (LOC = 6.88 - 8.33 cm in Fig. 34(a)) and RC1 of F3 reaches a local minima (LOC = 12.13 cm in Fig. 34(a)) move toward the glottis. Therefore, such positions of LOC are not stable under the perturbation of the lip opening area. However, the positions where RC1 of F3 reaches the absolute minimum (LOC = 8.38 cm in Fig. 34(a)) and the local minimum (LOC = 10.63 cm) and RC1 of F2 reaches the local minimum (LOC = 10.63 cm) are varied little. In these cases, the positions are stable under the perturbation of the lip opening area. Other cases also show the same tendency under the same condition.

When we increase the DOC while keeping the lip opening area constant (e.g., Fig. 34(a), (e), (i), and (m), LIP = 0.3 cm^2), we can observe that the region where RC1 of all the three formants reaches the absolute minimum (around at LOC = 8 cm in Fig. 34(a)) are varied not much. However, the range of the stable region became wider. Other cases also showed the same tendency. This implies that unstable positions can become stable by changing an articulatory parameter.

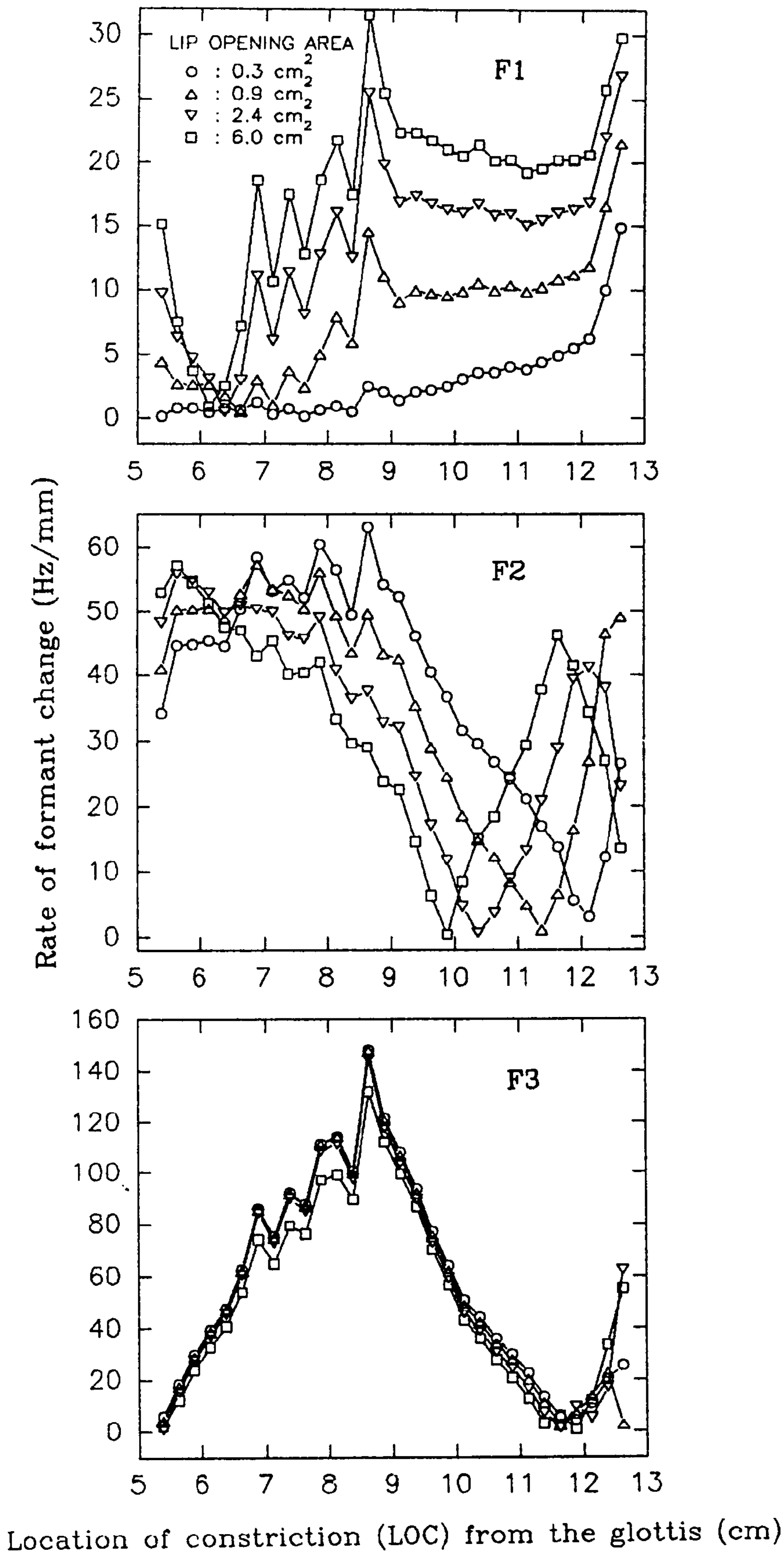
The above observations also suggest that the positions where stability occur can be varied depending on the configuration of other articulatory parameters. Therefore, it seems that there exist no absolutely stable regions in the vocal tract and the selection of places of articulation are interactively determined by the combination of the three articulatory parameters. However, it is interesting to note that the region about 9 - 10 cm along the LOC dimension in any plot in Fig. 34, which can be regarded as a transitional region, always existed there regardless of values of DOC and the lip opening area. It may be the only region which is absolutely stable. In terms of vowel articulation, the stable position existing posteriorly to the transitional region is a probable position for vowel /UW/ (e.g., LOC = 8.63 cm in Fig. 34(a)). The position approximately corresponds to the reference line 10 in the grid system shown in Fig. 32. Anterior to the transitional region, there exist another stable position where RC1 of F3 reaches a local minimum (e.g., LOC = 10.63 cm in Fig. 34(a) - (d)) or where RC1 of F3 reaches the local minimum (e.g., LOC = 11.88 cm in Fig. 34(k)). The position approximately corresponds to the reference line 12 or 13 in the grid system shown in Fig. 32, which might be a probable tongue position of palatal vowels (i.e, front vowels). The transitional region then exists between the probable tongue positions of vowel /UW/ and front vowels. Although it is currently not clear what the role of such a transitional region in vowel articulation is, the transitional region may be assumed as an articulatory boundary between back and front vowels. Such a region may be non-preferred by speakers, either physiologically or perceptually, to distinguish front from back vowels. In chapter 3, we showed that there exist an apparent discontinuity in MAPS between the high-back vowels /UW/ and /UU/ and the front vowels /IY/, /IH/, and /EH/. If the assumption just reached is correct, it may be a clue to answer the question about why the region in MAPS has remained empty.

C. Analysis of RC2 plots

RC2 values are plotted in Fig. 36 as a function of LOC during the manipulation of DOC (i.e., constriction size) under the perturbation of the lip opening area. There are three RC2 plots corresponding to each formant. In a plot, each curve corresponds to the lip opening area of 0.3 cm^2, 0.9 cm^2, 2.4 cm^2, 6.0 cm^2, respectively.

RC2 plot can be interpreted as a measure of effectiveness to induce changes in formant values by adjusting the same amount of DOC at different positions of LOC along the vocal

Fig. 36. RC2 plots of corresponding formant frequencies. The lip opening area is varied in each plot.



LIP OPENING AREA
o : 0.3 cm²
Δ : 0.9 cm²
∇ : 2.4 cm²
□ : 6.0 cm²
F1
F2
F3
Rate of formant change (Hz/mm)
Location of constriction (LOC) from the glottis (cm)

tract. Stevens (1989) did not have much concern about the effect of DOC on the acoustic stability with an assumption that the formant frequencies are usually not strongly sensitive to constriction size under the most condition. However, it can be observed that there exists an apparent sensitivity, i.e., a quantal nature of the acoustic stability, during the manipulation of DOC under the perturbation of the lip opening area.

From the RC2 pattern of F1, we can observe that the stable position exists at the LOC of 6.38 cm (approximately the upper pharyngeal region), and the position is stable under the perturbation of the lip opening area. In the case of F2, the stable position exists at the palatal region of the vocal tract. However, it is observed that the position is unstable under the perturbation of the lip opening area, i.e., it moves anteriorly when the lip opening area decreases. In the case of F3, it is observed that the RC2 pattern itself is almost the same regardless of the formants and lip opening area considered. A stable position exists at the LOC of 11.63 cm (and probably at the LOC of 5.38 cm) and the most unstable position exists at the LOC of 8.63 cm.

The above observations show that the position where acoustic stability occurs is determined depending on the formant being considered, and that acoustic stability of each formant behaves differently under the variation of DOC. This suggests that the acoustic stability achieved by the manipulation of DOC should be considered selectively. It is not much meaningful to consider the proximity of a pair of formant frequency in this case. An interesting question is then, that if speakers utilize the acoustic stability for the selection of place of articulation, as assumed by the quantal theory, how they organize the stability which is varied with the formant being considered. For instance, the stable position of F1 (LOC = 6.38 cm), which may be a probable position of back vowels such as /AH/, corresponds to the most unstable region of F2. Therefore, if a speaker selects the position to produce the vowel /AH/, he might be able to take advantage of the acoustic stability of the position. However, it is limited to F1. He may be cautious to adjust the DOC because F2 is strongly sensitive to the DOC there. It may weaken the hypothesis of the quantal theory.

Another interesting observation is that the position of LOC of 8.63 cm is most unstable, particularly in the cases of F3 and partially F1 and F2, with respect to other positions in the sense that larger changes in the formants can be achieved there. In terms of vowel articulation, the above position of LOC may be a probable position for the vowel /UW/. The vowel is then located at the most unstable position in terms of the rate change of the formants values, i.e., it is located at a position where speakers can induce the largest changes in F3 (also, partially in F1 and F2) by the equal amount of manipulation of DOC along the length dimension of the vocal tract. If we consider the fact that the position is almost absolutely stable during the manipulation of LOC because of the benefit that it exists just anterior to the transitional region, the above observation implies that the acoustic stabilities achieved by the manipulations of the two articulatory dimensions, LOC and DOC, are contradictory with each other in the case of the vowel /UW/.

In summary, there may exist a duality in acoustic stability in the sense that a stable position for a particular formant corresponds to the unstable position of other formants, or an even almost absolutely stable position under the perturbation of an articulatory parameter is shown to be most unstable in terms of the rate change of formants themselves. This is another factor which may weaken the role of acoustic stability on the selection of the places of articulations.

IV. Conclusions

Positions where acoustic stability occur during the manipulation of LOC were varied by the perturbation of other articulatory parameters, DOC and the lip opening area. This suggests that the selection of places of articulation may be interactively determined by the combination of the three articulatory parameters and there are no absolutely stable regions in the vocal tract for specific vowels. However, a transitional region which appears between the probable positions of vowel /UW/ and front vowels seems to be the only region which is almost absolutely stable. The transitional region was hypothesized as an articulatory (or perceptual) boundary between back vowels and front vowels. It was shown that a quantal nature of acoustic stability also exists during the manipulation of DOC. The acoustic stabilities achieved by the LOC and DOC showed a contradictory feature in the case of vowel /UW/. A question that remains is that how the roles of the acoustic stability achieved by the manipulation of LOC and DOC, as well as the lip opening area, can be optimally combined for the selection of the places of articulations.

CHAPTER 5. SUMMARY

The major purpose of this dissertation was to provide a more phonetically relevant space than the traditional formant space. This study was based on the auditory-perceptual theory of vowel recognition by Miller (1989).

In chapter 2, as a preliminary step toward the goal, acoustic performances of three midsagittal width to cross-sectional area conversion algorithms were compared (Lindblom and Sundberg, 1965; Rubin et al., 1981; Ladefoged, 1988). The purpose of the comparison study was to select the most accurate conversion algorithm for the computation of formant frequencies using a tube model of the vocal tract. The comparison study was based on the vocal tract data measured from midsagittal vocal tract images acquired by magnetic resonance (MR) imaging.

The cross-sectional areas estimated by the three area conversion algorithms compared in the study agreed relatively well in the palatal region for the two subjects examined. However, when the tongue-palate distance became large in that region, the discrepancy between the cross-sectional areas also became large. Also, they underestimated the cross-sectional areas in the pharyngeal region. Clearly, a more refined conversion algorithm based on direct measurement of cross-sectional or three-dimensional vocal tract is needed to establish reliable articulatory-acoustic relationships.

Among the conversion algorithms tested in this study, the use of Ladefoged's area conversion table yielded the largest cross-sectional area in the pharyngeal region and relatively better estimations of formant frequencies for the current subjects.

In chapter 3, a two-dimensional modified auditory-perceptual space (MAPS) was derived from the three-dimensional auditory-perceptual space (APS) suggested by Miller (1989). MAPS was a unit vector representation of APS. It was shown that the use of the intrinsic normalization procedure based on the formant-ratio theory allows a substantial reduction of acoustic variations caused by the inter-speaker differences in MAPS. It is expected that a further reduction of inter-speaker differences is possible by employing a non-uniform transformation of formant data which can take account of the anatomical differences of the vocal tract between speaker groups of different sex and age (e.g., Fant, 1975).

Perceptual vowel target zones of nine monophthongal vowels of American English were constructed in MAPS by the comparison of number P_{ij} which was regarded as a measure of perceptual similarity. The use of P_{ij} provided a reasonable way to determine the boundaries between vowel target zones which overlap. It was hypothesized that the vowel target zones can be interpreted as both articulatory and perceptual targets. The vowel target zones classified a vowel data base by Peterson and Barney (1952) with 89% accuracy. Each vowel target zone was divided into the core zone and the boundary zone using the number P_{ij}. The core zone was regarded as a unified vowel target among articulatory, acoustic, and perceptual descriptions of vowels.

Vowel articulation was studied in MAPS based on the simulated midsagittal vocal tract shapes and computation of the formant frequencies using a tube model. The tongue shapes of vowels were generated by using the PARAFAC tongue model (Harshman et al., 1977) which yielded realistic tongue shapes of most monophthongal vowels in American English. Three articulatory parameters, the location and degree of tongue constriction and the lip opening area, were extracted from the simulated vocal tract shapes.

It was shown that the normalization procedure used in the current study has the inherent capability of extracting relevant articulatory information contained acoustic data, and redistributes the data points to align them with the extracted articulatory dimensions, the location and the degree of the tongue constriction and the lip opening area, in MAPS. Also, it was shown that two articulatory dimensions, the location and degree of the tongue constriction, are well split into the X and Y axes in MAPS and are almost equally distributed into the F1 and F2 axes in the F1-F2 spaces. This suggests that articulatory properties of vowels can be more explicitly represented in MAPS than in the traditional formant space.

It was shown that the use of core zones in MAPS provide simple and meaningful way for the comparison of vowels in a single language or across languages. This resulted from the facts that the articulatory information contained in acoustic data are emphasized by the intrinsic normalization procedure used in the study, and that MAPS has explicit articulatory characteristics which are best represented by the location and degree of the tongue constriction. Therefore, it was concluded that MAPS is a more phonetically relevant space than the conventional formant space.

In chapter 4, the acoustic stability which is a key concept in the quantal theory of speech was discussed in terms of the three articulatory parameters, the location and degree of tongue constriction and the lip opening area, which were extracted from simulated vocal tract shapes of vowels.

Acoustic stability existed during the manipulation of constriction location. However, the positions where acoustic stability occurs were varied by introducing the perturbation of the degree of tongue constriction and the lip opening area. This suggests that the selection of places of articulation may be interactively determined by the combination of the three articulatory parameters, and there are no absolutely stable regions in the vocal tract for specific vowels. However, a transitional region which appears between the probable positions of vowel /UW/ and front vowels seems to be the only region which is almost absolutely stable. The transitional region was hypothesized as an articulatory (or perceptual) boundary between back and front vowels.

It was also shown that a quantal nature of acoustic stability also existed during the manipulation of DOC. Therefore, a question that remains to be answered by the quantal theory may be that of how the different aspects of acoustic stability achieved by the manipulation of LOC and DOC, as well as the lip opening area, can be optimally combined for the selection of the places of articulations.

REFERENCES

Ainsworth, W. (1975). "Intrinsic and extrinsic factors in vowel judgements," in Auditory analysis and Perception of Speech, edited by G. Fant and M. Tatham (Academic, London) pp. 103-113.

Atal, B.S., Chang, J.J., Mathews, M.V., and Tukey, J.W. (1978). "Inversion of articulatory-to-acoustic transformation in the vocal tract by a computer-sorting technique," J. Acoust. Soc. Am. 63, 1535-1555.

Badin, P. and Fant, G. (1984). "Notes on vocal tract computation," Speech transmission Lab. Quart. Progr. Status Report 2-3, 53-108.

Baer, T., Gore, J.C., Boyce, S., and Nye, P.W. (1987). "Application of MRI to the analysis of the speech production," Magn. Reson. Imag. 5, 1-7.

Bladon, R.A.W. and Lindblom, B. (1981). "Modeling the judgement of vowel quality differences," J. Acoust. Soc. Am. 69, 1414-1422.

Bladon, R.A.W., Henton, C.G., and Pickering, J. B. (1984). "Towards an auditory theory of speaker normalization," Lang. Commun. 4, 59-69.

Chiba, Ts. and Kajiyama, M. (1941). The vowel: Its Nature and Structure (Tokyo-Kaisseikan, Tokyo)

Chuang, C.K. and Wang, W.S. (1975). "A distance-sensing device for tracking tongue configuration," J. Acoust. Soc. Am. Suppl. 1, 59, S11.

Coker, C.H. (1976). "A model of articulatory dynamics and control," Proc. IEEE 64, 452-460.

Disner, S.F. (1980). "Evaluation of vowel normalization procedures," J. Acoust. Soc. Am. 67, 253-261.

Dunn, H.K. (1950). "The calculation of vowel resonances and an electrical vocal tract," J. Acoust. Soc. Am. 22, 740-753.

Engstrands, O. (1988). "Articulatory correlates of stress and speaking rate in Swedish VCV utterances," J. Acoust. Soc. Am. 83, 1863-1875.

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

Fant, G. (1964). "Formant and Cavities," in Proceedings of the 5th International congress of Phonetic Sciences, Munster 1964 (Verlag S. Karger AG, Basel), pp. 120-141.

Fant, G. (1966). "A note on vocal tract size factors and non-uniform F-pattern scaling," Speech Transmission Lab. Quart. Progr. Status Report 4, 22-30

Fant, G., Hennigsson, G., and Stalhammar, U. (1969). "Formant frequency of Swedish vowels," Speech Transmission Lab. Quart. Progr. Status Report 4, 26-31.

Fant, G. (1973). Speech Sounds and Features (MIT, Cambridge, MA).

Fant, G. (1975). "Non-uniform vowel normalization," Speech Transmission Lab. Quart. Progr. Status Report 2-3, 1-19.

Fischer-Jorgensen, E. (1972). " Formant frequencies of long and short Danish vowels," in Studies for Einar Haugen, edited by E.S. Firchow et al. (Mouton, The Hague).

Flanagan, J.L. (1972). Speech Analysis Synthesis and Perception. 2nd ed. (New-York: Spring-Verlag)

Flege, J.E., Fletcher, S.G., McCutcheon, M., and Smith, S. (1986). "The physiological specifications of American English vowels," Language and Speech 29, 361-388.

Fromkin, V. (1964). "Lip positions in American English vowels", Language and Speech 7, 215-225.

Fujimura, O., Kiritani, S., and Ishida, H. (1973). "Computer controlled radiography for observation of movements of articulatory and other human organs," Comput. Med. Biol. 3, 371-384.

Gerstman, L.H. (1968). "Classification of self-normalized vowels," IEEE Trans. Audio. Electroacoust. AU-16, 78-80.

Greenwood, D.D. (1961). "Auditory masking and the critical band," J. Acoust. Soc. Am. 33, 484-501.

Gyngell, M.L., Palmer N.D., and Eastwood, L.M. (1986). "The application of steady-state free precession (SFP) in 2D-FT MR imaging," SMRM abstract 1986, 666-667.

Haase, A., Frahm J., Matthaie, D., Hanicke, W., and Merboldt, K.D. (1985) "Rapid images and NMR movies," SMRM abstract 1985, 980-981.

Harshman, R. and Papcun, G. (1976). "Vowel normalization by linear transformation of each speaker's acoustic space," J. Acoust. Soc. Am. Suppl. 1, 59, S71(A).

Harshman, R. and Ladefoged, P. and Goldstein, L. (1977). "Factor analysis of tongue shapes," J. Acoust. Soc. Am. 62, 693-707.

Heinz, J.M. and Stevens, K.N. (1964). "On the derivation of area functions and acoustic spectra from cineradiographic films of speech," J. Acoust. Soc. Am. 36, 1037

Hillenbrand, J. and Gayvert, R. (1987). "Speaker-independent vowel classification based on fundamental frequency and formant frequencies," J. Acoust. Soc. Am. Suppl. 1 81, S93.

Johansson, C., Sundberg, J., Wilbrand, H., and Ytterbergh, C. (1983). "From sagittal distance to area: A study of transverse, cross-sectional area in the pharynx by means of computer tomography," Speech Transmission Lab. Quart. Progr. Status Report 4, 39-49.

Jorgensen, H.P. (1969). "Die gespannten und ungespannten Vokale in der norddeutschen Hochsprache mit einer spezifischen Untersuchung der Struktur ihrer Formantenfrequenzen," Phonetica 19, 217-254.

Kelly, J.L. and Lochbaum, C.C. (1962). "Speech synthesis," Proc. 4th Int. Congr. Acoust. Paper G42, 1-4.

Kiritani, S., Tateno, Y., Iinuma, T., and Sawashima, M. (1977). "Computed tomography of the vocal tract," in Dynamic Aspects of Speech Production edited by M. Sawashima and F. Cooper (Univ. of Tokyo Press, Tokyo) pp. 203-206.

Koenig, W. (1949). "A new frequency scale for acoustic measurements," Bell Labs. Rec. 27, 299-301.

Koolhoven, H. (1968). Dutch (The English University Press, London)

Ladefoged, P., Anthony, J., and Riley D. (1971). "Direct measurements of the vocal tract," J. Acoust. Soc. Am. 49, 104(A).

Ladefoged, P. (1975). A course in Phonetics (Harcourt Brace Jovanovich, New York)

Ladefoged, P., Harshman, R., Goldstein, L., and Rice, L. (1978). "Generating vocal tract shapes from formant frequencies," J. Acoust. Soc. Am. 64, 1027-1035.

Ladefoged, P. (1988). Personal communication.

Lakshminarayanan, A.V., Lee Sungbok, and McCutcheon, M.J. (1989). "MR imaging of the vocal tract during vowel production," Radiology 173(P), 411.

Liberman, A.M. and Mattingly, I.G. (1985). "The motor theory of speech perception revised," Cognition 21, 1-36.

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

Lindblom, B. and Sundberg, J. (1971). "Acoustic consequences of lip, tongue, jaw and larynx movement," J. Acoust. Soc. Am. 50, 1166-1179.

Lindblom, B. (1986). "Phonetic universals in vowel systems," in Experimental Phonology, edited by J. J. Ohala and J.J. Jaeger (Academic Press). pp. 13-44.

Lindblom, B., and Engstrand, O. (1989). "In what sense is speech quantal?," J. Phon. 17, 107-121

Linker, W. (1982). "Articulatory and acoustic correlates of labial activity in vowels: A cross-linguistic study," UCLA Working Papers Phonet. 56, 1-134.

Lloyd, R.J. (1890). Some Researches into the Nature of the Vowel-Sound (Turner and Dunnett, Liverpool, England)

Lobanov, B.M. (1971). "Classification of Russian vowels spoken by different speakers," J. Acoust. Soc. Am. 49, 606-608.

Maeda, S. (1982). "A digital simulation method of the vocal tract system," Speech Commun. 1, 199-229.

Markel, J.D. and Gray, A.H., Jr. (1973). "On autocorrelation equations as applied to speech analysis," IEEE Trans. Audio. Electroacoust. AU-21, 69-79.

Markel, J.D. and Gray, A.H., Jr. (1976). "Linear Prediction of Speech (Springer, New-York)

McCutcheon, M.J., Lee Sungbok, Lakshminarayanan, A.V., and Fletcher, S.G. (1990). "A comparison of glossometric measurements of tongue position with magnetic resonance images of the vocal tract," J. Acoust. Soc. Am. Suppl 1 87, S122.

Mcgowan, R.S. (1987). "Articulatory synthesis: Numerical solution of a hyperbolic differential equation," Haskins Lab. Status Report on Speech Research, SR-89/90, 69-79

Mermelstein, P., Maeda, S., and Fujimura, O. (1971). "Description of tongue lip movement in a jaw-based coordinate system," J. Acoust. Soc. Am. 49, 104.

Mermelstein, P. (1973). "Articulatory model for the study of speech production," J. Acoust. Soc. Am. 53, 1070-1082.

Meyer, P., Wilhelms, R., and Strube, H.W. (1989). "A quasiarticulatory speech synthesizer for German language running in real time," J. Acoust. Soc. Am. 86, 523-539.

Miller, J.D. (1989). "Auditory-perceptual interpretation of the vowel," J. Acoust. Soc. Am. 85, 2114-2134.

Nadler, R. and Abbas, J. (1988). "Use of the x-ray microbeam system for the study of articulatory dynamics," J. Acoust. Soc. Am. Suppl 1 84, S124.

Neary, T.M. (1977). "Phonetic feature systems for vowels," unpublished doctoral dissertation, University of Connecticut, Storrs, CT.

Neary, T.M. (1989). "Static, dynamic, and relational properties in vowel perception," J. Acoust. Soc. Am. 85, 2088-2113.

Nordstrom, P.E. (1975). "Attempts to simulate female and infant vocal tracts from male area functions," Speech Transmission Lab. Quart. Progr. Status Report 2-3, 20-33

Perkell, J.S. (1969). Physiology of Speech Production: Results and Implications of a Quantitative cineradiographic Study (MIT, Cambridge, MA)

Peterson, G.E. and Barney, H.L. (1952). "Control methods used in a study of the vowels,". J. Acoust. Soc. Am. 24, 175-184.

Peterson, G.E. (1961). "Parameters of vowel quality," J. Speech Hear. Res. 4, 10-29.

Pipes, L.A. and Hovenessian, S.H. (1969). Matrix-Computer Methods in Engineering (Wiley, New-York).

Pols, L.C.W., van der Kamp, L.J. Th., and Plomp, R. (1969). "Perceptual and physical space of vowel sounds," J. Acoust. Soc. Am. 46, 458-467.

Pols, L.C.W., Tromp, H., and Plomp, R. (1973). "Frequency analysis of Dutch vowels from 50 male speakers," J. Acoust. Soc. Am. 53, 1093-1101.

Portnoff, M.R. (1973). "A quasi-one-dimensional digital simulation for the time-varying vocal tract," Unpublished master's thesis, MIT, Cambridge, Mass.

Potter, R.K. and Steinberg, J.C. (1950). "Toward the specification of speech," J. Acoust. Soc. Am. 22, 807-820.

Press, H.P., Brian, P.F., Saul, A.T., and Vetterling, W.T. (1986). Numerical Recipes (Cambridge University Press, New-York)

Pykett, I.L., Newhouse, J.H., Buonanno, F.S., Brady, T.J., Goldman, M.R., Kistler, J.P., and Pohost, G.M. (1982). "Principles of nuclear magnetic resonance imaging," Radiology 143, 157-168.

Rokkaku, M., Hashimoto, K., Imaizumi, S., Niimi, S., and Kiritani, S. (1986). "Measurements of the three-dimensional shape of the vocal tract based on the magnetic resonance imaging technique," Annual Bulletin, Research Institute of Logopedics and Phoniatrics (University of Tokyo) 20, 47-54.

Rubin, P., Baer, T., and Mermelstein, P. (1981). "An articulatory synthesizer for perceptual research," J. Acoust. Soc. Am. 50, 1180-1192.

Sondhi, M. M. (1974). "Model for wave propagation in a lossy vocal tract," J. Acoust. Soc. Am. 55, 1070-1075.

Sondhi. M. M. (1986). "Resonances of a bent vocal tract," J. Acoust. Soc. Am. 79, 1113-1116

Sonoda, Y. (1976) "A high sensitivity magnetometer for measuring tongue point movements," in Dynamic Aspects of Speech Production edited by M. Sawashima and F. Cooper (Univ. of Tokyo Press, Tokyo) pp. 145-156.

Stevens, K.N., Kasowski, S., and Fant, G. (1953). "An electrical analog of the vocal tract," J. Acoust. Soc. Am. 25, 734-742.

Stevens, K.N. and House, A.S. (1955). "Development of a quantitative description of vowel articulation," J. Acoust. Soc. Am. 27, 484-493.

Stevens, K. N. (1989). "On the quantal nature of speech," J. Phon. 17, 3-45

Stone, M., Shawker, T., Talbot, T., and Rich, A. (1988). "Cross-sectional tongue shape during vowel production," J. Acoust. Soc. Am. 83, 1586-1596.

Stone, M. (1990). "A three-dimensional model of tongue movement based on ultrasound and x-ray microbeam data," J. Acoust. Soc. Am. 87, 1997-2217.

Syrdal, A.K. and Gopal, H.S. (1986). "A perceptual model of vowel recognition based on the auditory representation of American English vowels," J. Acoust. Soc. Am. 79, 1086-1100.

Walshe, M. O'C. (1965). Introduction to Scandinavian Languages (Andre Deutshe, London).

Wakita, H. and Fant, G. (1978). "Toward a better vocal tract model," Speech Transmission Lab. Quart. Progr. Status Report 1, 9-29.

Zwicker, E. and Terhardt, E. (1980). "Analytical expressions for critical band rate and critical bandwidths as a function of frequency," J. Acoust. Soc. Am. 68, 1523-1525.

APPENDIX A

LIST OF MIDSAGITTAL WIDTHS MEASURED FROM MR IMAGES.
W DENOTES MIDSAGITTAL WIDTH AND L DENOTES SECTION LENGTH IN CM UNITS. IT BEGINS FROM THE GLOTTIS.

/IY/ W	/IY/ L	/EH/ W	/EH/ L	/AE/ W	/AE/ L	/AH/ W	/AH/ L	/UW/ W	/UW/ L
Subject MJM									
1.25	0.33	0.95	0.55	0.95	0.23	0.95	0.48	1.42	0.25
0.90	0.49	0.55	0.57	0.85	0.55	0.70	0.50	1.35	0.51
0.55	0.65	0.50	0.75	0.50	0.53	0.50	0.64	1.08	0.53
0.50	0.65	1.75	0.70	0.35	0.68	1.60	0.80	0.80	0.60
1.90	0.55	1.65	0.52	1.70	0.72	1.45	0.50	0.50	0.50
1.75	0.55	0.65	0.50	1.62	0.50	1.25	0.50	1.90	0.40
1.80	0.55	1.45	0.52	1.40	0.50	1.10	0.50	1.80	0.58
1.90	0.55	1.90	0.52	1.10	0.48	0.85	0.48	1.80	0.57
2.65	0.55	1.90	0.52	1.60	0.51	1.20	0.48	1.80	0.50
2.50	0.55	1.85	0.52	1.55	0.50	1.10	0.48	2.70	0.50
2.35	0.55	1.73	0.52	1.48	0.50	1.05	0.48	2.30	0.50
2.25	0.55	1.58	0.52	1.46	0.50	1.00	0.48	1.95	0.56
2.20	0.55	1.50	0.52	1.55	0.50	1.10	0.48	1.80	0.50
2.10	0.45	1.55	0.47	1.65	0.50	1.26	0.48	1.70	0.50
2.08	0.45	1.70	0.47	1.70	0.50	0.70	0.48	1.45	0.50
2.00	0.45	1.80	0.47	1.16	0.56	0.60	0.50	1.30	0.50
2.00	0.53	1.20	0.52	1.00	0.50	0.70	0.41	1.25	0.41
1.45	0.53	1.05	0.45	0.90	0.45	0.90	0.41	1.30	0.41
1.25	0.37	1.15	0.45	0.85	0.45	1.07	0.41	1.40	0.41
1.20	0.46	1.20	0.45	0.72	0.45	1.22	0.41	1.42	0.40
1.20	0.42	1.15	0.45	0.65	0.41	1.20	0.41	0.85	0.45
1.23	0.42	1.10	0.45	0.53	0.40	1.20	0.38	0.75	0.45
1.10	0.43	1.00	0.40	0.45	0.38	1.25	0.38	0.85	0.40
0.85	0.43	0.90	0.40	0.38	0.38	1.30	0.38	0.96	0.40
0.70	0.43	0.85	0.38	0.40	0.39	1.35	0.37	1.10	0.40
0.60	0.43	0.80	0.38	0.40	0.39	1.43	0.37	1.00	0.40
0.50	0.43	0.75	0.38	0.50	0.39	1.47	0.37	0.80	0.40
0.35	0.43	0.65	0.38	0.55	0.39	1.52	0.37	0.65	0.40
0.30	0.43	0.60	0.38	0.55	0.39	1.60	0.37	0.50	0.40
0.25	0.43	0.55	0.38	0.60	0.39	1.60	0.37	0.42	0.40
0.25	0.43	0.55	0.38	0.62	0.39	1.68	0.37	0.35	0.40
0.23	0.43	0.50	0.38	0.70	0.39	1.70	0.37	0.37	0.40
0.30	0.43	0.50	0.38	0.70	0.39	1.70	0.37	0.40	0.40
0.35	0.43	0.50	0.38	0.70	0.39	1.70	0.37	0.50	0.40
0.40	0.43	0.50	0.38	0.72	0.39	1.82	0.37	0.52	0.40
0.42	0.43	0.45	0.38	0.72	0.39	1.86	0.37	0.60	0.40
0.50	0.43	0.45	0.38	0.80	0.39	1.90	0.37	0.65	0.40
0.50	0.43	0.47	0.38	0.82	0.39	1.90	0.37	0.73	0.40
0.50	0.43	0.50	0.38	0.86	0.39	1.90	0.37	0.75	0.40
0.40	0.43	0.48	0.38	0.93	0.39	1.80	0.37	0.85	0.40
0.40	0.43	0.50	0.38	1.05	0.50	1.70	0.42	0.85	0.40
0.41	0.43	0.60	0.38	1.55	0.57	1.95	0.41	0.90	0.40
0.70	0.43	0.96	0.48	1.40	0.50	1.40	0.47	0.90	0.40
0.90	0.62	1.30	0.46	0.85	0.50	0.75	0.50	0.90	0.40
0.97	0.50	0.80	0.46	0.77	0.50	0.50	0.50	1.20	0.65
0.60	0.50	0.70	0.46					3.00	0.60
0.68	0.23							2.40	0.75

/IY/		/EH/		/AE/		/AH/		/UW/	
W	L	W	L	W	L	W	L	W	L
								1.50	0.80
								0.60	0.60
								0.42	0.38
								0.30	0.38

Subject LSB

1.45	0.45	1.53	0.53	1.25	0.60	0.60	0.25	1.55	0.40
1.25	0.47	1.04	0.50	0.80	0.50	0.55	0.75	1.20	0.50
0.90	0.55	0.70	0.60	0.45	0.60	0.65	0.65	1.05	0.50
0.80	0.65	0.80	0.70	0.70	0.63	0.95	0.55	0.85	0.60
1.10	0.65	1.35	0.72	1.25	0.68	0.60	0.60	1.15	0.70
1.45	0.53	1.60	0.50	1.20	0.60	0.55	0.55	1.65	0.62
1.50	0.53	1.50	0.55	0.95	0.55	0.40	0.50	2.40	0.53
1.50	0.53	1.30	0.57	0.90	0.55	0.35	0.50	2.30	0.53
1.30	0.53	1.30	0.55	0.75	0.50	0.70	0.50	2.00	0.56
1.20	0.53	1.20	0.52	0.87	0.54	0.65	0.50	1.80	0.50
1.35	0.53	1.25	0.52	0.89	0.50	0.60	0.50	1.30	0.50
1.25	0.53	1.25	0.52	0.95	0.50	0.65	0.53	1.00	0.50
1.25	0.50	1.30	0.52	1.05	0.50	0.75	0.45	0.75	0.50
1.30	0.50	1.35	0.47	1.15	0.50	0.95	0.45	0.55	0.55
1.37	0.43	1.45	0.47	1.25	0.45	1.20	0.45	0.45	0.45
1.45	0.43	1.55	0.47	1.45	0.45	1.45	0.45	0.42	0.45
1.60	0.45	1.75	0.47	1.70	0.45	1.80	0.45	0.50	0.45
1.75	0.45	1.95	0.50	1.90	0.45	2.20	0.45	0.55	0.45
1.95	0.46	2.10	0.50	2.16	0.45	1.30	0.52	0.70	0.47
2.25	0.50	2.45	0.52	2.57	0.55	1.50	0.45	1.00	0.52
1.80	0.50	1.56	0.52	1.75	0.50	1.80	0.42	1.35	0.54
1.60	0.50	1.95	0.52	2.05	0.50	1.85	0.42	0.60	0.54
1.60	0.45	1.85	0.49	2.05	0.50	2.02	0.43	0.75	0.54
1.60	0.45	1.70	0.47	1.82	0.45	2.20	0.43	0.73	0.53
1.40	0.45	1.57	0.47	1.67	0.45	2.30	0.43	0.70	0.53
1.15	0.50	1.40	0.47	1.55	0.45	2.43	0.43	0.65	0.54
0.90	0.45	1.20	0.49	1.27	0.45	2.65	0.41	0.63	0.51
0.75	0.45	1.05	0.45	1.05	0.45	2.75	0.43	0.65	0.51
0.60	0.45	0.95	0.47	0.88	0.45	2.80	0.43	0.70	0.51
0.45	0.45	0.85	0.45	0.80	0.41	2.85	0.43	0.75	0.51
0.35	0.45	0.80	0.45	0.82	0.47	2.90	0.40	0.90	0.47
0.30	0.45	0.80	0.45	0.87	0.47	2.90	0.40	1.05	0.48
0.25	0.42	0.72	0.48	0.87	0.42	2.90	0.40	1.35	0.48
0.25	0.45	0.80	0.48	0.95	0.45	2.80	0.40	1.60	0.40
0.25	0.45	0.80	0.48	0.95	0.45	2.15	0.40	1.90	0.45
0.30	0.41	0.73	0.45	0.92	0.45	2.30	0.40	2.35	0.90
0.35	0.41	0.70	0.45	0.90	0.47	2.00	0.41	3.00	0.35
0.40	0.41	0.76	0.47	0.92	0.45	1.85	0.42	3.50	0.30
0.40	0.41	0.85	0.47	1.00	0.36	2.45	0.70	3.40	0.30
0.40	0.40	1.20	0.62	1.20	0.35	2.55	0.53	3.18	0.30
0.45	0.40	1.30	0.50	1.87	0.65	1.82	0.45	2.80	0.50
0.62	0.40	1.40	0.50	2.10	0.45	1.50	0.55	1.25	1.00

/IY/		/EH/		/AE/		/AH/		/UW/	
W	L	W	L	W	L	W	L	W	L
1.20	0.53	1.20	0.58	1.40	0.45			1.00	0.51
1.20	0.53			1.37	0.20			1.10	0.52
1.15	0.53							0.32	0.52
1.00	0.50							0.13	0.52

APPENDIX B

LADEFOGED'S AREA CONVERSION TABLE. THERE ARE 17 ROWS AND EACH ROW HAS THREE LINES. FIRST ROW CORRESPONDS TO THE GLOTTIS AND LAST ROW CORRESPONDS TO AROUND THE ALVEOLI. FIRST NUMBER IN EACH ROW CORRESPONDS TO THE CROSS-SECTIONAL AREA OF MIDSAGITTAL WIDTH 0.1 CM AND LAST ONE TO 3.0 CM.

```
0.2  0.4  0.6  0.8  1.0  1.2  1.4  1.6  1.8  2.0  2.2  2.47
2.70 2.95 3.20 3.56 3.92 4.28 4.64 5.00 5.38 5.78 6.18 6.58
7.00 7.60 8.20 8.80 9.4 10.00

0.2  0.4  0.6  0.8  1.0  1.2  1.4  1.6  1.8  2.0  2.2  2.40
2.70 2.95 3.20 3.56 3.92 4.28 4.64 5.00 5.38 5.78 6.18 6.58
7.00 7.60 8.20 8.80 9.4010.00

0.2  0.38 0.58 0.80 1.0  1.20 1.42 1.62 1.84 2.06 2.32 2.61
2.98 3.16 3.46 3.83 4.18 4.50 4.83 5.15 5.48 5.82 6.16 6.48
6.85 7.32 7.78 8.24 8.70 9.15

0.2  0.37 0.57 0.79 1.00 1.21 1.44 1.65 1.88 2.12 2.42 2.75
3.16 3.36 3.72 4.10 4.44 4.72 5.02 5.30 5.58 5.86 6.14 6.41
6.70 7.04 7.36 7.68 8.00 8.30

0.19 0.36 0.56 0.78 1.0  1.22 1.44 1.68 1.92 2.18 2.52 2.89
3.34 3.57 3.98 4.37 4.7  4.94 5.2  5.45 5.68 5.9  6.12 6.34
6.55 6.76 6.94 7.12 7.30 7.45

0.18 0.35 0.55 0.77 1.00 1.23 1.46 1.70 1.96 2.24 2.62 3.03
3.52 3.79 4.24 4.64 4.96 5.18 5.4  5.6  5.78 5.95 6.11 6.27
6.40 6.48 6.52 6.56 6.6  6.65

0.17 0.34 0.54 0.76 1.0  1.24 1.48 1.72 1.98 2.30 2.70 3.16
3.62 4.0  4.50 4.9  5.20 5.40 5.60 5.75 5.88 6.00 6.10 6.20
6.25 6.20 6.10 6.00 5.88 5.75

0.18 0.35 0.55 0.77 1.00 1.23 1.46 1.70 1.96 2.24 2.62 3.03
3.52 3.79 4.24 4.64 4.96 5.18 5.40 5.60 5.78 5.95 6.11 6.27
6.40 6.48 6.52 6.56 6.60 6.65

0.19 0.36 0.56 0.78 1.00 1.22 1.44 1.68 1.92 2.18 2.52 2.89
3.34 3.57 3.98 4.37 4.70 4.94 5.20 5.45 5.68 5.90 6.12 6.34
6.55 6.76 6.94 7.12 7.30 7.45

0.20 0.37 0.57 0.79 1.00 1.21 1.44 1.65 1.88 2.12 2.42 2.75
3.16 3.36 3.72 4.10 4.44 4.72 5.02 5.30 5.58 5.86 6.14 6.41
6.70 7.04 7.36 7.68 8.00 8.30

0.20 0.38 0.58 0.80 1.00 1.20 1.42 1.62 1.84 2.06 2.32 2.61
2.98 3.16 3.46 3.83 4.18 4.50 4.83 5.15 5.48 5.82 6.16 6.48
6.85 7.32 7.78 8.24 8.70 9.15

0.09 0.2  0.33 0.50 0.72 0.94 1.16 1.40 1.71 2.04 2.31 2.53
2.92 3.26 3.62 3.92 4.28 4.68 5.05 5.50 5.95 6.40 6.84 7.28
7.72 8.16 8.60 9.04 9.48 9.92

0.10 0.22 0.36 0.55 0.78 0.96 1.20 1.44 1.76 2.08 2.32 2.54
2.87 3.18 3.52 3.80 4.12 4.50 4.82 5.14 5.47 5.80 6.14 6.54
6.94 7.34 7.74 8.14 8.54 8.94
```

0.10 0.20 0.32 0.50 0.68 0.83 1.00 1.21 1.42 1.65 1.88 2.10
2.33 2.60 2.90 3.13 3.40 3.78 4.12 4.48 4.84 5.20 5.56 5.96
6.36 6.76 7.16 7.56 7.96 8.36

0.15 0.30 0.40 0.60 0.86 1.18 1.43 1.82 2.16 2.45 2.70 3.10
3.50 3.90 4.30 4.70 5.10 5.50 5.90 6.30 6.70 7.10 7.50 8.90
9.30 9.7 10.1 10.5 10.9 11.31

0.15 0.30 0.40 0.60 0.86 1.18 1.43 1.82 2.16 2.45 2.70 3.10
3.50 3.90 4.30 4.70 5.10 5.55 5.90 6.30 6.70 7.10 7.50 8.90
9.30 9.7 10.1 10.5 10.9 11.31

0.15 0.30 0.40 0.60 0.86 1.18 1.43 1.82 2.16 2.45 2.70 3.10
3.50 3.90 4.30 4.70 5.10 5.55 5.90 6.30 6.70 7.10 7.50 8.90
9.30 9.7 10.1 10.5 10.9 11.30

APPENDIX C

SOURCE CODE FOR FORMANT COMPUTATION USING A TUBE MODEL OF THE VOCAL TRACT.

```
C.........................................................
C  THIS PROGRAM READS MIDSAGITTAL WIDTHS AND COMPUTES THE
C  FIRST THREE FORMANT FREQUENCIES USING THE THREE AREA
C  CONVERSION ALGORITHMS.
C
C  /PARAMETERS/
C
C        A : CROSS-SECTIONAL AREA OF A SECTION
C       TL : TUBE LENGTH OF EACH SECTION
C        N : NUMBER OF TUBES
C       YW : LOSS DUE TO WALL IMPEDANCE
C        R : LOSS DUE TO THE VISCOUSITY OF THE BOUNDARY
C            LAYER
C        G : LOSS DUE TO HEAT CONDUCTION
C       RG : LOSS DUE TO THE GLOTTAL RESISTANCE
C       ZM : LOSS DUE TO THE RADIATION IMPEDANCE
C       FQ : INVERSE OF THE TRANSFER FUNCTION
C        S : COMPLEX FREQUENCY
C     FACC : ACCURACY FOR THE ITERATION
C       FF : FORMANT FREQUENCY
C       BW : BAND WIDTH
C.........................................................
      PROGRAM FORMANT
C.........................................................
      INTEGER RTNUM, NUM
      REAL CAREA(17,37), DUMA(34,37)
      REAL SAGDIM(60), TDIM(60), XPOS(60), DZ2(60),
     +     XVAL(60)
      REAL FF(4), BW(4), QS1, QS2
      COMPLEX S1,S2,FACC,ZERO,FREQ,FQ
      CHARACTER*20 FMT, FNAME, QA
      COMMON/DATA/ A(60), TL(60), N
C
      FMT='(1X,(A\))'
999   FORMAT(A)
      WRITE(*,FMT)' File name to save formant: '
      READ(*,999) FNAME
      OPEN(4,FILE=FNAME)
      WRITE(*,FMT)' File name to save area function: '
      READ(*,999) FNAME
      OPEN(5,FILE=FNAME)
1     WRITE(*,*)' 1. Test Fant area functions'
      WRITE(*,*)' 2. Use Ladefoged area conversion table'
      WRITE(*,*)' 3. Use Rubin et al conversion method'
      WRITE(*,*)' 4. Use Lindblom and Sundberg conversion
     +             method'
      WRITE(*,FMT)' Select one: '
      READ(*,*) METHOD
      WRITE(*,*)
      IF(METHOD.EQ.1) THEN
         WRITE(*,FMT)' Fant area file name: '
         READ(*,999) FNAME
         OPEN(3,FILE=FNAME)
```

```
      ELSE
         WRITE(*,FMT)' MR data file name: '
         READ(*,999) FNAME
         OPEN(3,FILE=FNAME)
      ENDIF
C
      IF(METHOD.EQ.1) THEN ! TEST FANT AREA FUNCTION
         READ(3,1000) N
         DO I = 1, N
            READ(3,1100) A(I)
            TL(I) = 0.5
         ENDDO
      ELSE
         READ(3,1000) N ! NUMBER OF SECTIONS
         READ(3,1010) NR1, NR2, NR3 ! INDEX FOR RUBIN'S
                                    ! METHOD
         READ(3,1020) NL1, NL2      ! INDEX FOR LINDBLOM'S
                                    ! METHOD
         VTL = 0.                   ! VOCAL TRACT LENGTH
         DO I = 1, N                ! READ MR DATA FILE
            READ(3,1200) SAGDIM(I), TL(I)
            SAGDIM(I) = SAGDIM(I)*30./19.25/2.0164
            TL(I) = TL(I)*30./19.25/2.0164
            VTL = VTL + TL(I)
         ENDDO
         IF(METHOD.EQ.2) THEN ! TEST LADEFOGED'S TABLE
            OPEN(99,FILE='CAREA.DAT' ) ! READ AREA TABLE
            DO I = 1, 17
               READ(99,1300) (CAREA(I,J), J = 1, 37)
            ENDDO
            CLOSE(99)
            DO J = 1, 17
               DUMA(1,J) = CAREA(1,J)
            ENDDO
            DO I = 1, 16 ! EXPAND AREA TABLE TO 34-SECTION)
               DO J = 1, 37
                  DUMA(2*I,J) = CAREA(I,J)
                  DUMA(2*I+1,J) = (CAREA(I,J)+
     +                            CAREA(I+1,J))/2.
               ENDDO
            ENDDO
            DO J = 1, 37
               DUMA(34,J) = CAREA(17,J)
            ENDDO
C
            TDIM(1) = SAGDIM(1)       ! TDIM AT THE GLOTTIS
            DO I = 2, N
               TDIM(I) = SAGDIM(I-1)
            ENDDO
            TDIM(N+1) = SAGDIM(N)
            XPOS(1) = 0. ! INITIAL POSITION FOR
                         ! INTERPOLATION
            DO I = 2, N+1
```

```
         XPOS(I) = XPOS(I-1) + TL(I-1)
      ENDDO
      CALL NATSPLINE(XPOS,TDIM,N+1,DZ2) ! SPLINE
      NDUM = N+1       ! ORIGINAL # OF SECTIONS
      N = 35           ! CONVERT TO 35 SECTION TUBE
      SLENG = VTL/18.! A SECTION LENGTH
      DO K = 1, N-1
         TL(K) = SLENG/2.
      ENDDO
      TL(N) = SLENG ! TUBE LENGTH FOR LIP
      DO K = 1, N-1
         XVAL(K) = (SLENG/2.)*K
         CALL VALUE(XPOS,TDIM,NDUM,DZ2,XVAL(K),
     +              SAGDIM(K))
         IF(SAGDIM(K).LT.0.) PAUSE
      ENDDO
      CALL AREA (SAGDIM,DUMA)       ! CONVERT TO AREAS
      A(N) = 3.61*SAGDIM(NDUM-1)    ! LIP OPENING AREA
      XORG = 0.0
      WRITE(5,1400) XORG, A(1) ! WRITE AREA FUNCTION
      DO K = 1, N-1
         WRITE(5,1400) XVAL(K), A(K)
      ENDDO
      XVAL(N) = XVAL(N-1) + SLENG
      WRITE(5,1400) XVAL(N), A(N)
   ELSEIF (METHOD.EQ.3) THEN        ! TEST RUBIN METHOD
      DUM = 1.5/(NR1-1)
      DO I = 1, NR1                 ! PHARYNGEAL REGION
          A(I) = 3.14159*(SAGDIM(I)/2.)*(1.5 +
                       DUM*(I-1))/2.
      ENDDO
      DO I = NR1+1, NR2 ! SOFT PALATE REGION
         A(I) = 2.0*SAGDIM(I)**1.5
      ENDDO
      DO I = NR2+1, NR3 ! HARD PALATAL REGION
         A(I) = 1.6*SAGDIM(I)**1.5
      ENDDO
      DO I = NR3+1, N-2  ! AROUND THE ALVEOLI
         IF(SAGDIM(I).LE.0.5) THEN
            A(I) = 1.5*SAGDIM(I)
         ELSEIF(SAGDIM(I).GT.0.5.AND.SAGDIM(I).LE.2.)
     +   THEN
            A(I) = 0.75 + 3.*(SAGDIM(I) - 0.5)
         ELSE
            A(I) = 5.25 + 5.*(SAGDIM(I) - 2.)
         ENDIF
      ENDDO
      A(N-1) = 3.61*SAGDIM(N-1)    ! LIP OPENING AREA
      A(N) = 3.61*SAGDIM(N)
      XVAL(1) = 0.
      WRITE(5,1400) XVAL(1), A(1)
      DO I = 2, N+1
         XVAL(I) = XVAL(I-1) + TL(I-1)
```

```
          ENDDO
          DO I = 2, N+1
             WRITE(5,1400) XVAL(I), A(I-1)
          ENDDO
       ELSE ! TEST LINDBLOM & SUNDBERG METHOD
          DO I = 1, NL1
             A(I) = 1.1*SAGDIM(I)**2.21   ! LOWER PHARYNX
          ENDDO
          DO I = NL1+1, NL2
             A(I) = 1.68*SAGDIM(I)**1.9   ! UPPER PHARYNX
          ENDDO
          DO I = NL2+1, N-2
             A(I) = 2.2*SAGDIM(I)**1.38   ! MOUTH
          ENDDO
          A(N-1) = 3.61*SAGDIM(N-1) ! LIP OPENING AREA
          A(N) = 3.61*SAGDIM(N)
          XVAL(1) = 0.
          WRITE(5,1400) XVAL(1), A(1)
          DO I = 2, N+1
             XVAL(I) = XVAL(I-1) + TL(I-1)
          ENDDO
          DO I = 2, N+1
             WRITE(5,1400) XVAL(I), A(I-1)
          ENDDO
       ENDIF
    ENDIF
1000 FORMAT(I2)
1010 FORMAT(3(I2,1X))
1020 FORMAT(I2,1X,I2)
1100 FORMAT(F5.2,1X)
1200 FORMAT(2(F5.3,1X))
1300 FORMAT(15F5.2/(15F5.2)/(7F5.2))
1400 FORMAT(2(F5.2,1X))
C
    RTNUM = 4
    NUM=0
    FREQ=CMPLX(0.,0.)
    FACC=CMPLX(.1*3.141592,.1*2.*3.141592)
    QS1=REAL(FQ(FREQ))
    DO IR = 1, 25 ! UP TO 5000 Hz IN 200 Hz STEP
       FREQ=FREQ + CMPLX(0.,200.*2.*3.141592)
       QS2=REAL(FQ(FREQ))
       IF(QS1*QS2.LT.0) THEN
          NUM=NUM+1
          S1=FREQ-CMPLX(0.,200.*2.*3.141592)
          S2=FREQ
          CALL ROOT(S1,S2,FACC,ZERO)
          FF(NUM)=AIMAG(ZERO)/(2.*3.141592)
          BW(NUM)=-REAL(ZERO)/3.141592
       ENDIF
       QS1=QS2
       IF(NUM.EQ.RTNUM) GOTO 98
    ENDDO
```

```
98    WRITE(*,1500) METHOD, (FF(J),J=1,RTNUM),
     +      (BW(J),J=1,RTNUM), VTL
      WRITE(4,1500) METHOD, (FF(J),J=1,RTNUM),
     +      (BW(J),J=1,RTNUM), VTL
1500  FORMAT('METHOD:',I1,3X,4(F5.0,1X),4(F4.0,1X),1X,F4.1,
     +                  1X,'cm')
      WRITE(*,FMT)' Another calculation ? (y/n) -> '
      READ(*,999) QA
      IF(QA.EQ.'Y'.OR.QA.EQ.'y') THEN
         CLOSE(3)
         GOTO 1
      ENDIF
      CLOSE(4)
      STOP
      END

C.......................................................
C     CALCULATION OF THE INVERSE OF TRANSFER FUNCTION
C.......................................................
      COMPLEX FUNCTION FQ(X)
C
      REAL Z(60),R(60),G(60), L1, L2, LGG
      COMPLEX TMAT(60,2,2),HCOS(60),HSIN(60),MAT(2),X,
     +        ST(60), RG(2,2),MT(2), ZM
      COMMON/DATA/A(60),TL(60),N
C
      C=35000.
      D=1.14*1E-3
C
      DO I = 1, N
         ST(I) = X*TL(I)/C
         Z(I) = D*C/A(I)
         R(I) = 1.15*.001*TL(I)*(SQRT(A(I))**(-3))*
     +           SQRT(AIMAG(X))
         G(I) = 3.22*1.E-7*TL(I)*SQRT(A(I)*AIMAG(X))
      ENDDO
      L1 = 4.23*1.E-4/SQRT(A(N))  ! RADIATION IMPEDANCE
      R1 = 45.9/A(N)
      C1 = 1.033*1.E-7*(A(N))**(3./2.)
      L2 = 7.11*1.E-5/SQRT(A(N))
      ZM = R1*L1*AIMAG(X)/CMPLX(L1*AIMAG(X),AIMAG(X)
     +     **2.*L1*R1*C1-R1) + CMPLX(0.,AIMAG(X)*L2)
      DO I = 1, N
         HCOS(I) = .5*(CEXP(ST(I))+CEXP(-ST(I)))
         HSIN(I) = .5*(CEXP(ST(I))-CEXP(-ST(I)))
      ENDDO
C
      DO I = 1 ,8
         TMAT(I,1,1) = HCOS(I)+(R(I)/Z(I))*HSIN(I)
         TMAT(I,1,2) = Z(I)*HSIN(I)+R(I)*HCOS(I)
         TMAT(I,2,1) = G(I)*HCOS(I)+(1.+R(I)*G(I))*
     +                 HSIN(I)/Z(I)
         TMAT(I,2,2) = Z(I)*G(I)*HSIN(I)+(1.+R(I)*G(I))
```

```
+                    *HCOS(I)
      ENDDO
      ! YIELDING WALL NEAR GLOTTIS
      ADUM = (A(5)+A(6)+A(7)+A(8)+A(9)+(10)+A(11))/7.
      TMAT(9,1,1) = CMPLX(1.,0.)
      TMAT(9,1,2) = CMPLX(0.,0.)
      TMAT(9,2,1) = 1./(SQRT(6./ADUM)*CMPLX(7.16,AIMAG(X)
     +              *0.015))
      TMAT(9,2,2) = CMPLX(1.,0.)
      DO I = 10 ,30
         TMAT(I,1,1) = HCOS(I-1)+(R(I-1)/Z(I-1))*HSIN(I-1)
         TMAT(I,1,2) = Z(I-1)*HSIN(I-1)+R(I-1)*HCOS(I-1)
         TMAT(I,2,1) = G(I-1)*HCOS(I-1)+(1.+R(I-1)*
     +                 G(I-1))*HSIN(I-1)/Z(I-1)
         TMAT(I,2,2) = Z(I-1)*G(I-1)*HSIN(I-1)+(1.+R(I-1)*
     +                 G(I-1))*HCOS(I-1)
      ENDDO
      ! YIELDING WALL NEAR MOUTH
      ADUM = (A(N-6)+A(N-5)+A(N-4)+A(N-3)
     +       +A(N-2)+A(N-1)+A(N))/7.
      TMAT(31,1,1) = CMPLX(1.,0.)
      TMAT(31,1,2) = CMPLX(0.,0.)
      TMAT(31,2,1) = 1./(SQRT(6./ADUM)*CMPLX(18.,AIMAG(X)
     +               *0.038))
      TMAT(31,2,2) = CMPLX(1.,0.)
      DO I = 32 ,N+2
         TMAT(I,1,1) = HCOS(I-2)+(R(I-2)/Z(I-2))*HSIN(I-2)
         TMAT(I,1,2) = Z(I-2)*HSIN(I-2)+R(I-2)*HCOS(I-2)
         TMAT(I,2,1) = G(I-2)*HCOS(I-2)+(1.+R(I-2)*
     +                 G(I-2))*HSIN(I-2)/Z(I-2)
         TMAT(I,2,2) = Z(I-2)*G(I-2)*HSIN(I-2)+(1.+R(I-2)*
     +                 G(I-2))*HCOS(I-2)
      ENDDO
C
      RG(1,1) = CMPLX(1.,0.)           ! GLOTTAL IMPEDANCE
      RG(1,2) = CMPLX(0.,0.)
      RGG = 270.
      LGG = 16.67*1E-3
      RG(2,1) = CMPLX(RGG*(AIMAG(X)*LGG)**2 + RGG**3,
     +          -AIMAG(X)*LGG*RGG**2 - (AIMAG(X)*LGG)**3)/
     +          (2.*(AIMAG(X)*RGG*LGG)**2
     +          + RGG**4 + (AIMAG(X)*LGG)**4)
      RG(2,2) = CMPLX(1.,0.)
C
      DO I = 1, 2
         MT(I) = (0.,0.)
         DO J = 1, 2
            MT(I) = MT(I)+RG(2,J)*TMAT(1,J,I)
         ENDDO
      ENDDO
C
      TMAT(1,2,1) = MT(1)
      TMAT(1,2,2) = MT(2)
```

```
C
      DO K = 2, N + 2
         CALL MUL(TMAT,MAT,K)
         DO J = 1, 2
            TMAT(K,2,J) = MAT(J)
         ENDDO
      ENDDO
      FQ = MAT(1)*ZM + MAT(2)
      RETURN
      END
C
C.............................................
      SUBROUTINE MUL(A,M,K)
C.............................................
C
      COMPLEX A(60,2,2), M(2)
C
      DO I = 1, 2
         M(I) = (0.,0.)
         DO J = 1, 2
            M(I) = M(I) + A(K-1,2,J)*A(K,J,I)
         ENDDO
      ENDDO
      RETURN
      END
C
C...............................................
      SUBROUTINE ROOT(X1,X2,XACC,XZERO)
C...............................................
C
      COMPLEX X1,X2,XACC,XZERO,DX,FX,DFX
C
      XZERO = 0.5*(X1+X2)
      DO I = 1, 20
         CALL FUNCD(XZERO,FX,DFX)
         DX = FX/DFX
         XZERO = XZERO-DX
         IF(CABS(DX).LT.CABS(XACC)) THEN
            WRITE(*,*) I
            RETURN
         ENDIF
      ENDDO
      PAUSE 'EXCEEDING MAXIMUM ITERATION'
      RETURN
      END
C
C..............................................
      SUBROUTINE FUNCD(SX,F,DF)
C..............................................
C
      COMPLEX SX,DEL,DELS,DF,F,DELF,FQ
C
      F = FQ(SX)
```

```
      DEL = CMPLX(10., 10.)
      DELS = SX+DEL
      DELF = FQ(DELS)
      DF = (DELF-F)/DEL
      RETURN
      END
C.............................................
        SUBROUTINE AREA (DIM,DUMMY)
C.............................................
C
      REAL DIM(60), DUMMY(34,37)
      COMMON/DATA/A(60),TL(60), N
C
      DO I = 1, N-1
         DUM2 = 10.*DIM(I)
         IDUM3 = INT(DUM2)
         DUM4 = DUM2 - FLOAT(IDUM3)
         DUM5 = DUMMY(I,IDUM3+1) - DUMMY(I,IDUM3)
         A(I) = DUMMY(I,IDUM3) + DUM4*DUM5
      ENDDO
      RETURN
      END
```

APPENDIX D

FORMANT DATA BASE (1: /IY/, 2: /IH/, 3: /EH/, 4: /AE/, 5: /AH/, 6: /AW/, 7: /UH/, 8: /UU/, 9: /UW/, AND 10: /ER/)

	F1	F2	F3	F0		F1	F2	F3	F0 (Hz)
1	285.	2521.	3161.	110.	1	250.	2240.	2740.	125.
1	266.	2546.	3174.	108.	1	260.	2450.	3180.	122.
1	237.	2510.	3177.	115.	1	260.	2420.	3200.	123.
1	259.	2538.	3155.	111.	1	250.	2220.	2880.	125.
1	292.	2575.	3251.	109.	1	340.	2700.	3310.	245.
1	316.	2517.	3006.	117.	1	340	2580.	3100.	246.
1	317.	2463.	3105.	111.	1	340.	2560.	3210.	259.
1	271.	2536.	2985.	117.	1	370.	2900.	3540.	264.
1	329.	2514.	2916.	113.	1	380.	3000.	3490.	291.
1	330.	2479.	2857.	116.	1	430.	3370.	4010.	355.
1	332.	2053.	2729.	114.	1	274.	2005.	2910.	132.
1	294.	2043.	2832.	98.	1	264.	2354.	3269.	132.
1	309.	2035.	2731.	99.	1	322.	2247.	3213.	132.
1	318.	1983.	2757.	112.	1	270.	2490.	3186.	132.
1	323.	2031.	2665.	125.	1	285.	2377.	3031.	132.
1	315.	2522.	2905.	87.	1	287.	2459.	3149.	132.
1	311.	2562.	2926.	84.	1	350.	2404.	3039.	132.
1	333.	2452.	2798.	82.	1	269.	2362.	2885.	132.
1	357.	2532.	2917.	84.	1	324.	2450.	3149.	132.
1	325.	2398.	2991.	85.	1	247.	2240.	2730.	108.
1	295.	2246.	3119.	128.	1	260.	2298.	2668.	124.
1	275.	2283.	2930.	129.	1	215.	2205.	3235.	148.
1	266.	2252.	2973.	129.	1	235.	2050.	2915.	120.
1	269.	2249.	2952.	130.	1	235.	2500.	3220.	121.
1	270.	2299.	2977.	134.	1	285.	2280.	2835.	136.
1	336.	2674.	3263.	143.	1	255.	2195.	2720.	129.
1	298.	2726.	3117.	147.	1	300.	2408.	3265.	150.
1	352.	2693.	3303.	140.	1	273.	2040.	2795.	137.
1	294.	2723.	3340.	138.	1	306.	2365.	2950.	120.
1	354.	2627.	3353.	134.	1	260.	2340.	2820.	173.
1	350.	3049.	3627.	171.	1	190.	2675.	3225.	132.
1	368.	3121.	3687.	184.	1	257.	2233.	2765.	128.
1	355.	3150.	3547.	183.	1	335.	2295.	3100.	181.
1	352.	3113.	3762.	183.	1	366.	2465.	2860.	122.
1	343.	3122.	3727.	182.	1	244.	2300.	2920.	145.
1	387.	3023.	3664.	200.	1	267.	2450.	3150.	107.
1	293.	2989.	3769.	224.	1	255.	2230.	2823.	127.
1	377.	2934.	3609.	214.	1	295.	2340.	2860.	134.
1	287.	2726.	3650.	219.	1	243.	2200.	3088.	143.
1	302.	2898.	3600.	222.	1	294.	2195.	2950.	129.
1	370.	2780.	3540.	189.	1	224.	2350.	2890.	116.
1	390.	2690.	3328.	199.	1	293.	2415.	2860.	147.
1	368.	2714.	3544.	192.	1	254.	2085.	2890.	106.
1	374.	2795.	3521.	175.	1	265.	2238.	2938.	160.
1	374.	2786.	3445.	194.	1	279.	2435.	3125.	140.
1	390.	2541.	3087.	200.	1	215.	2090.	3135.	127.
1	390.	2511.	3071.	188.	1	290.	2420.	3275.	153.
1	372.	2455.	3016.	190.	1	330.	2283.	3115.	132.
1	371.	2517.	3087.	182.	1	290.	2360.	2935.	153.
1	384.	2477.	3093.	191.	2	432.	2036.	2884.	85.
1	370.	3200.	3730.	272.	2	448.	1909.	2746.	98.

1	310.	2790.	3310.	235.	2	473.	2009.	2772.	118.
1	360.	2960.	3475.	240.	2	439.	1991.	2739.	105.
1	305.	2790.	3150.	266.	2	462.	1979.	2727.	110.
1	380.	2930.	3305.	253.	2	478.	1988.	2626.	105.
1	388.	2725.	3163.	233.	2	455.	1951.	2676.	83.
1	213.	2805.	3600.	213.	2	449.	1593.	2442.	112.
1	350.	2755.	3120.	204.	2	466.	1522.	2428.	118.
1	344.	2835.	3230.	181.	2	468.	1550.	2482.	122.
1	339.	2710.	3250.	229.	2	459.	1518.	2371.	117.
1	298.	2900.	3450.	286.	2	469.	1473.	2321.	127.
1	260.	2800.	3305.	250.	2	486.	1772.	2703.	84.
1	312.	2370.	3015.	202.	2	463.	1738.	2683.	77.
1	308.	2950.	3670.	245.	2	464.	1794.	2698.	79.
1	278.	2780.	3780.	231.	2	476.	1807.	2707.	78.
1	338.	2753.	3150.	189.	2	449.	1746.	2694.	96.
1	362.	2895.	3800.	259.	2	389.	1963.	2638.	124.
1	300.	3000.	3300.	200.	2	386.	1915.	2591.	120.
1	380.	2715.	3080.	232.	2	391.	1921.	2715.	132.
1	280.	2715.	3010.	248.	2	382.	1899.	2621.	125.
1	324.	2685.	3200.	241.	2	371.	1942.	2634.	127.
1	380.	2850.	3330.	245.	2	417.	2124.	2947.	141.
1	403.	2600.	3023.	202.	2	427.	2018.	2788.	147.
1	260.	2850.	3400.	217.	2	400.	2145.	2847.	144.
1	350.	2880.	3350.	229.	2	407.	2027.	2945.	138.
1	230.	2805.	3550.	210.	2	408.	2135.	2837.	142.
1	286.	2710.	3185.	260.	2	558.	2389.	3120.	188.
1	349.	2830.	3355.	278.	2	552.	2234.	3324.	181.
1	288.	2800.	3410.	231.	2	602.	2396.	3109.	194.
1	282.	2840.	3130.	216.	2	566.	2344.	3132.	184.
1	430.	3350.	3900.	214.	2	586.	2371.	3355.	195.
1	335.	3450.	4180.	298.	2	490.	2242.	2967.	207.
1	345.	3245.	3675.	249.	2	474.	2197.	2993.	197.
1	415.	3300.	3850.	278.	2	443.	2272.	3234.	208.
1	475.	2765.	3550.	332.	2	454.	2217.	3130.	208.
1	465.	3505.	4270.	219.	2	510.	2085.	2965.	211.
1	335.	3175.	3560.	258.	2	403.	2220.	2877.	195.
1	310.	3080.	3600.	260.	2	421.	2113.	2798.	207.
1	299.	2973.	3450.	240.	2	428.	2122.	2874.	213.
1	290.	3270.	3890.	240.	2	409.	2170.	3089.	203.
1	335.	2955.	3705.	281.	2	371.	2191.	2795.	170.
1	365.	3225.	3850.	281.	2	437.	2043.	2775.	201.
1	288.	3265.	3825.	288.	2	414.	2108.	2853.	194.
1	370.	2930.	3440.	278.	2	398.	2117.	2742.	200.
1	347.	3180.	3700.	332.	2	407.	2091.	2768.	197.
1	325.	2270.	2780.	135.	2	465.	2059.	2724.	184.
1	240.	2125.	2880.	122.	2	530.	2730.	3600.	269.
1	284.	2475.	2790.	154.	2	430.	2480.	3070.	232.
					2	443.	1899.	2738.	104.
2	433.	1944.	2738.	102.	2	350.	1890.	2490.	113.
2	432.	1944.	2740.	99.	2	370.	2050.	2730.	100.
2	340.	2020.	2600.	112.	2	464.	2673.	3145.	232.
2	340.	1940.	2460.	118.	2	481.	2640.	3100.	241.
2	380.	2220.	2700.	125.	2	430.	2538.	3150.	234.

2	440.	2310.	2940.	212.
2	410.	2030.	2860.	225.
2	430.	2420.	3150.	228.
2	440.	2550.	3280.	234.
2	480.	3000.	3600.	300.
2	359.	1857.	2516.	132.
2	399.	1874.	2592.	132.
2	295.	1895.	2858.	132.
2	384.	1920.	2711.	132.
2	444.	2079.	2770.	132.
2	471.	1920.	2789.	132.
2	442.	1804.	2650.	132.
2	431.	1881.	2519.	132.
2	444.	2089.	2716.	132.
2	408.	1960.	2655.	102.
2	363.	1905.	2415.	114.
2	415.	1870.	2590.	140.
2	333.	1970.	2470.	125.
2	393.	2130.	2850.	146.
2	435.	1970.	2545.	136.
2	375.	1940.	2460.	122.
2	396.	2246.	2793.	155.
2	345.	1985.	2660.	135.
2	490.	1980.	2610.	140.
2	350.	2005.	2595.	193.
2	370.	1775.	2725.	131.
2	402.	2205.	2710.	130.
2	440.	2020.	2688.	169.
2	403.	2000.	2705.	129.
2	426.	2030.	2430.	155.
2	308.	2165.	2625.	107.
2	395.	2020.	2463.	145.
2	420.	2025.	2670.	136.
2	411.	1933.	2575.	144.
2	458.	1765.	2450.	120.
2	414.	1985.	2475.	111.
2	466.	1979.	2639.	148.
2	300.	2150.	2600.	100.
2	412.	1948.	2513.	152.
2	353.	2210.	2705.	138.
2	295.	1975.	2690.	129.
2	445.	2005.	2380.	151.
2	388.	2169.	2715.	129.
2	430.	2055.	2675.	146.
2	425.	2140.	2600.	137.
2	405.	1790.	2400.	125.
2	392.	2095.	2515.	157.
2	453.	2460.	3160.	227.
2	470.	2320.	2960.	255.
3	547.	1819.	2626.	113.
3	581.	1401.	2288.	106.
3	531.	1437.	2301.	106.
3	534.	1452.	2276.	102.

2	500.	2420.	2975.	181.
2	410.	2610.	3100.	205.
2	475.	2385.	3093.	237.
2	475.	2645.	3320.	259.
2	531.	2520.	3220.	266.
2	432.	2155.	2760.	211.
2	488.	2375.	3105.	244.
2	432.	2628.	3115.	232.
2	407.	2541.	3045.	194.
2	518.	2665.	3310.	267.
2	434.	2605.	3040.	217.
2	439.	2355.	2940.	220.
2	455.	2490.	3265.	261.
2	414.	2250.	2955.	235.
2	494.	2630.	2950.	235.
2	440.	2303.	2818.	205.
2	370.	2565.	3090.	232.
2	507.	2625.	3165.	248.
2	322.	2500.	3015.	222.
2	461.	2290.	2985.	263.
2	473.	2599.	3060.	291.
2	384.	2410.	3065.	233.
2	420.	2400.	2980.	207.
2	605.	2525.	4050.	205.
2	645.	3100.	3730.	324.
2	540.	2710.	3450.	270.
2	570.	2870.	3860.	286.
2	570.	2500.	3450.	310.
2	560.	3125.	4085.	228.
2	515.	2565.	3640.	264.
2	545.	2590.	3580.	288.
2	440.	2500.	3220.	221.
2	425.	2770.	3675.	238.
2	550.	2735.	3455.	255.
2	450.	2930.	3700.	297.
2	528.	2740.	3425.	277.
2	500.	2670.	3240.	281.
2	565.	2830.	3500.	314.
3	609.	1809.	2667.	78.
3	633.	1818.	2646.	104.
3	622.	1835.	2616.	97.
3	597.	1858.	2667.	99.
3	654.	1782.	2663.	103.
3	576.	1868.	2737.	118.
3	589.	1831.	2723.	103.
3	605.	1811.	2668.	110.
3	626.	1858.	2692.	110.
3	626.	1858.	2692.	110.
3	540.	2415.	2900.	217.
3	558.	2570.	3200.	214.
3	710.	1900.	2700.	244.
3	740.	2210.	2950.	246.

3	552.	1401.	2280.	102.	3	720.	2220.	3400.	250.
3	527.	1440.	2325.	106.	3	630.	2230.	3490.	253.
3	517.	1790.	2690.	80.	3	770.	2420.	3780.	295.
3	549.	1741.	2621.	80.	3	516.	1848.	2525.	132.
3	542.	1744.	2770.	79.	3	453.	1807.	2596.	132.
3	563.	1797.	2770.	77.	3	565.	1954.	2912.	132.
3	553.	1777.	2728.	84.	3	513.	1795.	2620.	132.
3	511.	1848.	2605.	119.	3	524.	1964.	2687.	132.
3	441.	1875.	2580.	115.	3	659.	1889.	2789.	132.
3	473.	1752.	2552.	125.	3	500.	1827.	2635.	132.
3	498.	1830.	2618.	123.	3	500.	1777.	2465.	132.
3	487.	1799.	2537.	122.	3	567.	1984.	2654.	132.
3	547.	2074.	2876.	136.	3	545.	1760.	2440.	98.
3	517.	2011.	2850.	139.	3	515.	1755.	2288.	104.
3	527.	2045.	2841.	136.	3	505.	1735.	2445.	136.
3	463.	2017.	2947.	132.	3	425.	2015.	2405.	116.
3	519.	1987.	2898.	134.	3	520.	1853.	2510.	145.
3	765.	2085.	3090.	177.	3	565.	1840.	2450.	139.
3	703.	1957.	2900.	175.	3	545.	1685.	2345.	120.
3	739.	1876.	3000.	173.	3	524.	2035.	2680.	143.
3	767.	2030.	2946.	172.	3	412.	1800.	2470.	124.
3	833.	1888.	2895.	163.	3	598.	1935.	2580.	118.
3	688.	1837.	3153.	188.	3	530.	1785.	2510.	158.
3	677.	1888.	2907.	184.	3	373.	1690.	2530.	124.
3	688.	2060.	2905.	184.	3	516.	1854.	2495.	116.
3	723.	1777.	2927.	171.	3	622.	1870.	2815.	157.
3	719.	2243.	2903.	206.	3	585.	1845.	2535.	117.
3	597.	2089.	2905.	178.	3	553.	1975.	2475.	165.
3	402.	2102.	2974.	188.	3	473.	1935.	2550.	106.
3	519.	2110.	2897.	188.	3	500.	1823.	2475.	126.
3	655.	2169.	2943.	177.	3	569.	1870.	2585.	130.
3	599.	2035.	3056.	102.	3	587.	1758.	2575.	140.
3	570.	2120.	2767.	191.	3	573.	1800.	2480.	121.
3	552.	2088.	2764.	192.	3	433.	2010.	2460.	109.
3	556.	2038.	2759.	199.	3	539.	1962.	2570.	135.
3	580.	1996.	2757.	203.	3	504.	1995.	2780.	105.
3	562.	2098.	2733.	200.	3	559.	1773.	2378.	140.
3	690.	2610.	3570.	260.	3	569.	1900.	2355.	133.
3	610.	2330.	2990.	223.	3	480.	1920.	2615.	125.
3	530.	1780.	2410.	127.	3	590.	1805.	2380.	138.
3	560.	1680.	2500.	96.	3	540.	1965.	2410.	138.
3	580.	1840.	2500.	125.	3	540.	1910.	2525.	135.
3	580.	1940.	2560.	125.	3	550.	1885.	2360.	130.
3	560.	2300.	2970.	129.	3	575.	1760.	2450.	118.
3	500.	1830.	2620.	168.	3	587.	1870.	2225.	146.
3	630.	1930.	2800.	233.	3	690.	2310.	3140.	230.
3	610.	1930.	2440.	235.	3	655.	2235.	2930.	248.
3	570.	1710.	2370.	237.	3	635.	2465.	2945.	231.
3	620.	2480.	2900.	221.	4	828.	1812.	2582.	132.
3	559.	2435.	3175.	190.	4	886.	1772.	2620.	132.
3	553.	2360.	3120.	231.	4	676.	1465.	2262.	98.
3	535.	2355.	3090.	217.	4	692.	1478.	2238.	103.
3	725.	2430.	3125.	244.	4	697.	1714.	2423.	85.

3	550.	2275.	2765.	200.
3	753.	2335.	3118.	236.
3	621.	2400.	3020.	226.
3	564.	2266.	2818.	189.
3	683.	2480.	3300.	252.
3	516.	2530.	3185.	210.
3	560.	2150.	2960.	191.
3	740.	2320.	3110.	251.
3	657.	2275.	2940.	223.
3	544.	2460.	3021.	224.
3	554.	2138.	2800.	199.
3	645.	2395.	3030.	217.
3	650.	2438.	3060.	227.
3	616.	2340.	2850.	216.
3	630.	2045.	2930.	240.
3	719.	2338.	3040.	245.
3	439.	2365.	3080.	220.
3	560.	2170.	2770.	209.
3	680.	2625.	3725.	218.
3	775.	3010.	3715.	278.
3	726.	2500.	3400.	245.
3	795.	2680.	4150.	281.
3	818.	2360.	3745.	298.
3	635.	2830.	3630.	214.
3	730.	2515.	3515.	253.
3	650.	2430.	3490.	272.
3	665.	2288.	3300.	234.
3	600.	2655.	3660.	218.
3	758.	2505.	3300.	252.
3	645.	2770.	3550.	278.
3	610.	2890.	3550.	274.
3	598.	2480.	3135.	267.
3	815.	2695.	3595.	308.
4	716.	1932.	2670.	99.
4	713.	1933.	2601.	101.
4	703.	1914.	2540.	95.
4	762.	1905.	2627.	94.
4	760.	1925.	2587.	101.
4	733.	1617.	2597.	102.
4	799.	1747.	2616.	94.
4	766.	1685.	2631.	97.
4	778.	1671.	2619.	105.
4	727.	1748.	2556.	104.
4	707.	1480.	2208.	108.
4	686.	1552.	2166.	105.
4	699.	1559.	2284.	98.
4	524.	1925.	2599.	132.
4	774.	1774.	2751.	132.
4	758.	1785.	2539.	132.
4	727.	1731.	2404.	132.
4	723.	1934.	2651.	132.
4	644.	1733.	2353.	94.
4	691.	1758.	2573.	81.
4	714.	1744.	2652.	78.
4	671.	1736.	2594.	79.
4	619.	1717.	2431.	78.
4	654.	1830.	2758.	108.
4	692.	1846.	2717.	110.
4	636.	1844.	2738.	112.
4	712.	1831.	2739.	115.
4	614.	1842.	2707.	110.
4	624.	2033.	2936.	137.
4	663.	2057.	2873.	132.
4	682.	2097.	2790.	135.
4	718.	2032.	2890.	127.
4	637.	1981.	2849.	132.
4	833.	1888.	2895.	163.
4	802.	1880.	2826.	153.
4	741.	1931.	2868.	160.
4	782.	1883.	2906.	157.
4	803.	2000.	2955.	160.
4	906.	2011.	3003.	200.
4	759.	1758.	3063.	188.
4	800.	1701.	3063.	189.
4	1012.	1895.	2903.	197.
4	830.	1693.	3072.	184.
4	862.	1954.	2804.	165.
4	803.	1847.	2725.	160.
4	787.	1944.	2831.	158.
4	672.	1894.	2920.	154.
4	750.	2003.	2793.	182.
4	697.	1955.	2647.	156.
4	787.	2165.	2638.	172.
4	741.	2020.	2734.	177.
4	756.	2020.	2834.	175.
4	631.	2160.	2793.	168.
4	1010.	2320.	3320.	251.
4	860.	2050.	2850.	210.
4	710.	1560.	2470.	124.
4	700.	1590.	2400.	125.
4	760.	1760.	2560.	128.
4	760.	1600.	2430.	128.
4	990.	1850.	2920.	215.
4	990.	2140.	3350.	254.
4	960.	1970.	2720.	259.
4	910.	2060.	2880.	260.
4	1120.	2380.	3610.	280.
4	677.	1753.	2552.	132.
4	623.	1691.	2489.	132.
4	859.	2390.	3060.	250.
4	880.	2225.	2840.	194.
4	830.	1950.	2830.	196.
4	875.	1995.	2795.	237.
4	802.	1775.	2820.	229.
4	743.	2245.	2980.	211.

4	630.	1706.	2260.	103.
4	695.	1650.	2570.	130.
4	640.	1755.	2130.	129.
4	685.	1780.	2725.	139.
4	820.	1695.	2240.	139.
4	650.	1525.	2300.	120.
4	694.	1899.	2539.	137.
4	674.	1585.	2340.	124.
4	691.	1622.	2525.	113.
4	570.	1745.	2600.	160.
4	540.	1590.	2625.	121.
4	625.	1864.	2565.	113.
4	758.	1805.	2325.	153.
4	701.	1679.	2520.	122.
4	595.	1975.	2515.	141.
4	533.	1970.	2550.	107.
4	642.	1675.	2488.	115.
4	680.	1695.	2555.	132.
4	718.	1675.	2323.	136.
4	690.	1640.	2415.	116.
4	736.	1695.	2420.	106.
4	640.	1807.	2525.	133.
4	670.	1860.	2500.	100.
4	617.	1773.	2413.	135.
4	657.	1779.	2400.	128.
4	645.	1600.	2280.	117.
4	760.	1595.	2345.	131.
4	616.	1722.	2465.	122.
4	655.	1810.	2460.	139.
4	715.	1885.	2520.	131.
4	738.	1540.	2370.	113.
4	683.	1815.	2280.	143.
4	1025.	1970.	3000.	223.
4	965.	1960.	2905.	244.
4	688.	2535.	3148.	229.
4	816.	1915.	2575.	161.
4	812.	2160.	2885.	203.
4	1090.	2040.	2725.	218.
4	824.	2050.	2920.	173.
4	675.	2193.	3005.	225.
4	815.	2170.	3020.	224.
4	1013.	1975.	2915.	235.
4	853.	1712.	2415.	190.
4	1024.	1995.	2835.	203.
4	995.	2110.	2910.	205.
4	700.	2085.	2694.	175.
5	640.	970.	2870.	112.
5	670.	980.	2900.	122.
5	690.	1072.	2660.	117.
5	680.	1000.	2530.	103.
5	650.	970.	2580.	162.
5	650.	980.	2350.	163.
5	762.	1150.	2680.	125.

4	840.	1827.	2653.	191.
4	880.	2125.	2935.	203.
4	1010.	2115.	2735.	225.
4	700.	2185.	2640.	211.
4	850.	1830.	2660.	214.
4	891.	2150.	3130.	209.
4	879.	2070.	2900.	206.
4	830.	1985.	2660.	205.
4	900.	2400.	3975.	200.
4	1135.	2280.	3530.	271.
4	710.	2875.	3335.	228.
4	1065.	2450.	3975.	258.
4	975.	2150.	3675.	296.
4	1145.	2625.	3595.	215.
4	1125.	2245.	3510.	239.
4	960.	2335.	3200.	255.
4	839.	1975.	3250.	228.
4	920.	2360.	3305.	237.
4	1005.	2170.	2740.	234.
4	1100.	2350.	3105.	269.
4	1300.	2220.	3215.	260.
4	943.	2080.	2890.	269.
4	1135.	2500.	3190.	267.
5	640.	1080.	2140.	91.
5	720.	1090.	2230.	94.
5	750.	1150.	2440.	100.
5	703.	1092.	2320.	95.
5	650.	1040.	2450.	140.
5	670.	1100.	2430.	136.
5	750.	980.	2460.	122.
5	550.	760.	2500.	119.
5	840.	1330.	2200.	129.
5	770.	1330.	2310.	137.
5	760.	1220.	2140.	131.
5	720.	1260.	2020.	126.
5	630.	980.	2330.	115.
5	670.	940.	2380.	121.
5	825.	1168.	2750.	137.
5	840.	1210.	2680.	135.
5	712.	1024.	2250.	125.
5	670.	1080.	2830.	125.
5	730.	1203.	2700.	112.
5	752.	1125.	2620.	107.
5	740.	1070.	2490.	148.
5	800.	1060.	2640.	170.
5	760.	1160.	3100.	200.
5	1030.	1340.	3100.	206.
5	840.	1110.	2930.	150.
5	850.	1120.	2850.	170.
5	730.	1210.	2740.	178.
5	735.	1220.	2850.	175.
5	845.	1334.	2890.	222.

5	652.	960.	2650.	113.	5	888.	1290.	2800.	233.
5	730.	1048.	2450.	146.	5	900.	1290.	2750.	252.
5	730.	1130.	2320.	155.	5	900.	1240.	3110.	260.
5	714.	1170.	2420.	111.	5	978.	1290.	2840.	258.
5	650.	1150.	2350.	97.	5	935.	1230.	2730.	246.
5	750.	1080.	2680.	114.	5	700.	1080.	2420.	200.
5	777.	1026.	2625.	114.	5	767.	1150.	2590.	192.
5	725.	1160.	2860.	129.	5	970.	1343.	3018.	206.
5	680.	1170.	2930.	136.	5	592.	1230.	2600.	236.
5	725.	1046.	2325.	145.	5	770.	1189.	2640.	220.
5	746.	1018.	2775.	131.	5	1000.	1260.	2900.	210.
5	730.	1160.	2340.	111.	5	766.	1180.	2340.	163.
5	758.	1280.	2470.	176.	5	750.	1065.	2640.	167.
5	686.	1078.	2570.	98.	5	820.	1050.	2800.	234.
5	700.	1050.	2680.	103.	5	855.	1170.	2940.	243.
5	780.	1170.	2640.	130.	5	960.	1280.	3000.	200.
5	788.	1115.	2645.	131.	5	810.	1120.	2400.	180.
5	740.	1115.	2330.	96.	5	830.	1200.	2880.	188.
5	740.	1115.	2330.	96.	5	880.	1240.	2870.	210.
5	721.	1080.	2280.	150.	5	790.	1250.	3080.	240.
5	700.	1048.	2335.	116.	5	820.	1210.	2960.	241.
5	735.	1070.	2100.	133.	5	834.	1282.	2800.	225.
5	713.	1180.	2200.	117.	5	850.	1270.	2760.	212.
5	700.	1100.	2240.	118.	5	843.	1190.	2860.	183.
5	670.	1100.	2220.	120.	5	740.	1160.	2780.	205.
5	750.	1100.	2550.	125.	5	915.	1280.	2530.	194.
5	800.	1120.	2500.	138.	5	723.	1196.	2600.	206.
5	777.	1170.	2600.	111.	5	920.	1350.	2550.	195.
5	750.	1175.	2820.	114.	5	920.	1470.	2690.	210.
5	740.	1110.	2500.	145.	5	987.	1330.	2830.	214.
5	700.	1060.	2720.	143.	5	1009.	1415.	3080.	214.
5	700.	1040.	2120.	141.	5	995.	1392.	2290.	199.
5	750.	1160.	2080.	125.	5	1000.	1400.	2440.	200.
5	710.	950.	2520.	119.	5	830.	1020.	2650.	225.
5	690.	960.	2520.	120.	5	830.	1095.	2610.	219.
5	760.	1260.	2120.	140.	5	990.	1237.	2360.	275.
5	770.	1140.	2020.	135.	5	987.	1172.	3180.	267.
5	985.	1260.	2740.	210.	5	860.	1103.	2700.	208.
5	900.	1350.	2990.	225.	5	860.	1103.	2700.	208.
5	950.	1130.	3160.	250.	5	800.	1200.	2920.	200.
5	850.	1150.	2940.	256.	5	760.	1140.	2850.	190.
5	978.	1362.	2724.	227.	5	1220.	1560.	3650.	205.
5	933.	1350.	2610.	233.	5	1300.	1800.	3450.	200.
5	800.	1200.	2800.	200.	5	1170.	1500.	3440.	265.
5	714.	1154.	2850.	205.	5	980.	1300.	3100.	283.
5	940.	1380.	2400.	250.	6	660.	980.	2220.	158.
5	1200.	1500.	3160.	276.	6	560.	860.	2900.	113.
5	1040.	1350.	3850.	275.	6	570.	820.	2820.	121.
5	1100.	1460.	4250.	250.	6	510.	700.	2650.	117.
5	990.	1410.	3750.	299.	6	504.	756.	2540.	120.
5	1050.	1320.	3730.	280.	6	430.	720.	2450.	145.
5	1090.	1380.	3050.	218.	6	510.	800.	2500.	171.
5	860.	1250.	2800.	212.	6	660.	1030.	2690.	120.

5	940.	1400.	3400.	240.
5	930.	1370.	3120.	245.
5	950.	1200.	2950.	278.
5	920.	1080.	2770.	250.
5	950.	1350.	3100.	250.
5	910.	1360.	2950.	227.
5	970.	1450.	3260.	242.
5	1010.	1650.	3150.	225.
5	780.	1250.	3180.	245.
5	970.	970.	3120.	236.
5	1110.	1630.	2780.	278.
5	1130.	1400.	3000.	280.
5	1230.	1300.	3200.	250.
5	1090.	1230.	2980.	286.
5	810.	1350.	2940.	270.
5	1000.	1360.	3000.	275.
5	1190.	1470.	3150.	350.
5	1070.	1460.	2950.	314.
5	750.	1020.	2620.	125.
5	730.	1040.	2300.	100.
5	710.	1080.	2530.	125.
5	850.	1120.	2750.	125.
6	550.	870.	2300.	92.
6	540.	840.	2280.	120.
6	565.	780.	2350.	106.
6	584.	849.	2460.	106.
6	580.	580.	2470.	149.
6	560.	560.	2410.	140.
6	550.	550.	2250.	117.
6	640.	830.	2550.	106.
6	590.	1010.	2100.	140.
6	620.	1050.	2180.	144.
6	540.	970.	1980.	136.
6	550.	880.	1950.	124.
6	560.	790.	2480.	112.
6	610.	840.	2420.	120.
6	671.	1000.	2670.	143.
6	690.	968.	2660.	147.
6	550.	913.	2360.	125.
6	550.	890.	2280.	126.
6	507.	755.	2420.	108.
6	538.	816.	2450.	116.
6	600.	970.	2280.	161.
6	579.	856.	2790.	214.
6	545.	905.	2750.	205.
6	530.	800.	2870.	190.
6	380.	720.	2700.	212.
6	560.	960.	2850.	160.
6	536.	850.	2850.	192.
6	631.	923.	2250.	225.
6	543.	980.	2300.	233.
6	700.	1050.	2750.	175.
6	720.	960.	2700.	125.
6	600.	900.	2400.	150.
6	640.	890.	2280.	178.
6	590.	965.	2500.	107.
6	578.	970.	2460.	109.
6	580.	800.	2650.	115.
6	585.	819.	2625.	117.
6	530.	800.	3400.	133.
6	512.	805.	3480.	128.
6	560.	840.	2500.	140.
6	560.	924.	2350.	140.
6	560.	810.	2290.	114.
6	584.	840.	2280.	116.
6	560.	665.	2620.	102.
6	550.	650.	2700.	106.
6	633.	891.	2500.	138.
6	600.	935.	2550.	150.
6	494.	789.	2420.	105.
6	628.	900.	2575.	157.
6	625.	875.	2375.	125.
6	625.	875.	2180.	125.
6	700.	1000.	2250.	115.
6	460.	720.	2180.	122.
6	470.	690.	2200.	118.
6	540.	850.	2320.	143.
6	555.	890.	2370.	150.
6	630.	891.	2519.	105.
6	572.	924.	2660.	114.
6	600.	970.	2570.	146.
6	650.	880.	2660.	138.
6	670.	920.	2240.	133.
6	570.	850.	2250.	142.
6	460.	610.	2500.	125.
6	470.	710.	2500.	120.
6	500.	800.	1850.	145.
6	600.	1000.	2000.	132.
6	650.	900.	2960.	225.
6	640.	915.	2970.	228.
6	530.	870.	2680.	242.
6	600.	900.	2770.	250.
6	700.	1080.	2810.	240.
6	720.	1090.	2840.	240.
6	530.	800.	2780.	267.
6	485.	810.	2750.	180.
6	750.	1250.	3450.	250.
6	675.	950.	3240.	225.
6	770.	1150.	3950.	286.
6	710.	1200.	3900.	273.
6	770.	940.	3750.	285.
6	680.	1020.	3700.	333.
6	800.	1220.	3700.	211.
6	640.	1070.	3000.	214.
6	530.	860.	3400.	240.

6	720.	1080.	3030.	190.
6	500.	750.	2750.	250.
6	632.	850.	2850.	243.
6	600.	860.	2410.	200.
6	600.	900.	2400.	200.
6	650.	900.	2920.	233.
6	687.	1060.	2780.	229.
6	440.	749.	2640.	220.
6	567.	752.	2600.	210.
6	595.	918.	2600.	170.
6	630.	985.	2630.	176.
6	470.	895.	2870.	236.
6	480.	880.	2880.	240.
6	520.	880.	2500.	220.
6	574.	890.	2510.	217.
6	700.	1000.	3130.	200.
6	470.	830.	3300.	207.
6	408.	695.	3040.	234.
6	420.	590.	3100.	246.
6	688.	1029.	2750.	229.
6	670.	1040.	2640.	222.
6	623.	1022.	2700.	222.
6	594.	990.	2640.	220.
6	575.	1073.	2490.	192.
6	600.	1100.	2600.	200.
6	720.	1110.	2420.	194.
6	700.	1100.	2780.	200.
6	672.	1084.	2495.	226.
6	627.	1045.	2504.	209.
6	656.	944.	2250.	205.
6	720.	960.	2380.	200.
6	650.	960.	2760.	240.
6	455.	810.	2750.	253.
6	587.	1068.	3270.	267.
6	560.	990.	3150.	293.
6	606.	910.	2900.	202.
6	583.	860.	2840.	201.
6	560.	760.	2800.	200.
6	560.	770.	3000.	207.
6	660.	1100.	3850.	219.
6	690.	1090.	3900.	217.
6	530.	1060.	3450.	265.
6	540.	1080.	3000.	272.
7	570.	1050.	2500.	118.
7	590.	1100.	3000.	125.
7	628.	1254.	2470.	114.
7	617.	1255.	2480.	114.
7	620.	1240.	2410.	167.
7	640.	1250.	2400.	160.
7	680.	1150.	2560.	118.
7	726.	1270.	2560.	125.
7	630.	1140.	2200.	157.
7	630.	1170.	2280.	186.
6	520.	910.	3420.	240.
6	790.	1050.	2900.	262.
6	750.	1000.	2500.	250.
6	700.	1120.	3070.	203.
6	690.	920.	2760.	230.
6	670.	1160.	3550.	232.
6	720.	1260.	3400.	225.
6	825.	1210.	3100.	258.
6	930.	930.	2900.	300.
6	580.	930.	2950.	292.
6	540.	1070.	3000.	270.
6	535.	970.	2960.	270.
6	550.	1080.	2850.	275.
6	910.	1200.	3180.	300.
6	830.	1250.	3250.	330.
6	620.	870.	2630.	124.
6	540.	840.	2150.	91.
6	610.	850.	2680.	122.
6	610.	890.	2620.	125.
6	660.	880.	2620.	125.
7	600.	1200.	2320.	100.
7	612.	1160.	2350.	105.
7	601.	1273.	2130.	103.
7	590.	1283.	2150.	105.
7	650.	1080.	2420.	140.
7	625.	1060.	2490.	125.
7	660.	1000.	2380.	120.
7	660.	960.	2450.	110.
7	590.	1450.	2330.	140.
7	560.	1410.	2140.	140.
7	630.	1300.	1950.	136.
7	650.	1170.	2000.	130.
7	620.	1100.	2390.	120.
7	640.	1110.	2370.	125.
7	675.	1320.	2550.	150.
7	704.	1393.	2550.	150.
7	565.	1157.	2310.	128.
7	550.	1150.	2250.	130.
7	633.	1260.	2530.	115.
7	660.	1213.	2460.	120.
7	590.	1250.	2620.	144.
7	620.	1300.	2530.	148.
7	750.	1500.	2750.	250.
7	738.	1300.	2820.	217.
7	820.	1240.	2600.	217.
7	860.	1300.	2670.	216.
7	733.	1468.	2700.	183.
7	740.	1280.	2900.	200.
7	675.	1551.	2923.	225.
7	690.	1630.	2900.	233.
7	840.	1400.	2750.	280.
7	760.	1330.	2950.	270.

7	640.	1300.	2300.	108.	7	750.	1500.	2850.	263.
7	624.	1350.	2410.	104.	7	850.	1400.	2750.	250.
7	650.	1220.	2550.	122.	7	720.	1440.	2380.	200.
7	672.	1260.	2500.	120.	7	707.	1470.	2440.	191.
7	645.	1190.	2520.	140.	7	830.	1540.	2860.	220.
7	640.	1240.	2580.	128.	7	900.	1510.	2840.	237.
7	586.	1078.	2300.	136.	7	818.	1515.	2500.	217.
7	627.	1038.	2360.	136.	7	700.	1240.	2650.	200.
7	600.	1250.	2300.	117.	7	618.	1518.	2700.	187.
7	575.	1170.	2240.	125.	7	624.	1430.	2660.	183.
7	552.	1122.	2500.	104.	7	765.	1270.	3060.	254.
7	580.	1150.	2600.	115.	7	775.	1385.	2830.	262.
7	672.	1272.	2640.	143.	7	816.	1450.	2700.	204.
7	658.	1241.	2560.	146.	7	858.	1500.	2700.	214.
7	630.	1127.	2420.	105.	7	700.	1200.	3100.	200.
7	630.	1127.	2420.	105.	7	635.	1200.	3250.	227.
7	600.	1050.	2100.	150.	7	770.	1540.	2840.	257.
7	641.	1162.	2275.	125.	7	800.	1410.	2860.	257.
7	740.	1370.	2180.	137.	7	765.	1300.	2700.	225.
7	625.	1312.	2250.	125.	7	730.	1390.	2790.	221.
7	610.	1100.	2230.	121.	7	800.	1340.	2700.	200.
7	620.	1120.	2330.	126.	7	772.	1280.	2660.	214.
7	660.	1200.	2330.	132.	7	720.	1500.	2560.	200.
7	675.	1140.	2380.	150.	7	800.	1400.	2420.	200.
7	583.	1110.	2360.	116.	7	750.	1540.	2800.	214.
7	608.	1120.	2700.	117.	7	770.	1530.	2780.	214.
7	630.	1200.	2600.	150.	7	788.	1462.	2920.	225.
7	680.	1290.	2600.	140.	7	736.	1500.	2900.	217.
7	810.	1110.	2100.	139.	7	800.	1520.	2380.	221.
7	770.	1150.	2100.	131.	7	780.	1470.	2400.	210.
7	620.	880 .	2500.	124.	7	710.	1340.	2780.	230.
7	650.	1000.	2520.	124.	7	740.	1470.	2430.	245.
7	660.	1370.	2110.	143.	7	920.	1512.	2950.	275.
7	680.	1300.	2100.	145.	7	910.	1688.	3050.	260.
7	770.	1365.	3060.	220.	7	654.	1160.	2800.	218.
7	715.	1340.	3010.	223.	7	654.	1160.	2800.	218.
7	750.	1280.	2760.	251.	7	690.	1200.	2900.	208.
7	770.	1340.	2800.	258.	7	666.	1206.	2900.	201.
7	768.	1440.	2855.	240.	7	1000.	1750.	3550.	200.
7	794.	1447.	2920.	234.	7	1110.	1690.	4040.	223.
7	708.	1485.	2760.	186.	7	1000.	1800.	3450.	300.
7	676.	1500.	2590.	188.	7	880.	1500.	3200.	250.
7	850.	1700.	3250.	243.	8	490.	990.	1920.	132.
7	970.	1600.	3950.	286.	8	360.	860.	2200.	121.
7	800.	1680.	3800.	250.	8	400.	840.	2200.	120.
7	880.	1700.	3750.	293.	8	443.	1273.	2430.	143.
7	900.	1600.	3650.	340.	8	459.	1286.	2410.	153.
7	820.	1470.	3500.	216.	8	360.	1028.	2160.	128.
7	970.	1410.	3200.	211.	8	390.	1060.	2150.	140.
7	770.	1540.	3500.	256.	8	456.	1040.	2300.	114.
7	800.	1490.	3300.	257.	8	480.	1120.	2160.	120.
7	780.	1650.	3350.	270.	8	440.	1120.	2210.	163.
7	720.	1500.	3240.	250.	8	400.	1070.	2280.	190.

7	706.	1410.	3200.	236.	8	350.	1000.	2500.	125.
7	720.	1480.	2880.	211.	8	380.	920.	2370.	130.
7	770.	1650.	3420.	250.	8	465.	990.	2440.	122.
7	690.	1600.	3350.	230.	8	462.	976.	2450.	125.
7	860.	1530.	3100.	268.	8	460.	1120.	2150.	170.
7	970.	1500.	3050.	256.	8	493.	1120.	2300.	170.
7	700.	1730.	2960.	290.	8	456.	1080.	2520.	120.
7	725.	1570.	2900.	270.	8	450.	1140.	2600.	120.
7	850.	1540.	3020.	275.	8	448.	960.	2200.	160.
7	840.	1580.	2880.	262.	8	450.	1000.	2180.	196.
7	810.	1600.	3230.	270.	8	467.	1110.	2400.	111.
7	760.	1530.	3180.	280.	8	475.	1220.	2310.	105.
7	930.	1540.	3120.	310.	8	480.	950.	2500.	140.
7	950.	1670.	3150.	315.	8	461.	993.	2350.	127.
7	650.	1270.	2600.	132.	8	420.	975.	3450.	140.
7	690.	1260.	2650.	125.	8	450.	1030.	3500.	150.
7	680.	1240.	2610.	128.	8	495.	1080.	2275.	150.
7	640.	1260.	2550.	130.	8	500.	1100.	2275.	143.
7	660.	1290.	2720.	132.	8	455.	970.	2140.	130.
7	440.	1140.	2580.	163.	8	456.	1642.	2038.	120.
7	810.	1400.	2740.	211.	8	448.	980.	2370.	112.
7	640.	1200.	2260.	215.	8	410.	940.	2370.	104.
7	650.	1310.	2640.	218.	8	490.	1102.	2420.	175.
7	800.	1300.	2820.	229.	8	492.	1077.	2306.	154.
7	700.	1220.	2550.	260.	8	450.	1028.	2160.	128.
7	780.	1430.	3040.	260.	8	562.	1119.	2340.	167.
7	780.	1780.	3590.	260.	8	507.	1050.	2325.	133.
7	840.	1570.	3650.	281.	8	420.	1100.	2000.	150.
7	1130.	1740.	3670.	310.	8	420.	1120.	2100.	140.
					8	320.	770.	1860.	129.
8	460.	1150.	2290.	114.	8	310.	790.	1920.	130.
8	456.	1030.	2300.	114.	8	460.	960.	2210.	136.
8	420.	1100.	2140.	105.	8	460.	1000.	2350.	156.
8	422.	1200.	2175.	111.	8	438.	975.	2300.	125.
8	450.	940.	1910.	145.	8	420.	938.	2300.	140.
8	410.	830.	2190.	141.	8	430.	1130.	2440.	142.
8	390.	730.	2770.	140.	8	430.	1150.	2420.	143.
8	360.	740.	2200.	130.	8	550.	970.	2200.	140.
8	450.	1200.	2250.	150.	8	490.	870.	2240.	141.
8	415.	1140.	2210.	143.	8	390.	900.	2100.	125.
8	470.	1040.	1990.	133.	8	460.	920.	2140.	125.
8	380.	1060.	1950.	157.	8	520.	1350.	3190.	275.
8	470.	1220.	2150.	150.	8	510.	1415.	3130.	280.
8	496.	1120.	3000.	225.	8	315.	900.	2650.	225.
8	690.	1070.	3240.	230.	8	430.	1075.	2580.	215.
8	600.	1225.	2500.	250.	8	430.	1000.	2700.	213.
8	630.	1320.	2560.	264.	8	620.	1420.	3700.	206.
8	500.	1215.	2870.	243.	8	620.	1410.	3520.	220.
8	500.	1240.	2860.	239.	8	560.	1440.	3500.	285.
8	450.	1460.	2550.	214.	8	570.	1450.	3500.	294.
8	467.	1400.	2450.	233.	8	610.	1500.	3300.	300.
8	490.	1160.	2610.	233.	8	500.	1370.	3500.	275.
8	513.	1500.	2650.	250.	8	680.	1420.	3800.	285.

8	400.	940.	2820.	235.	8	640.	1350.	3950.	278.
8	380.	860.	2680.	214.	8	550.	1195.	3750.	322.
8	424.	1040.	2780.	212.	8	550.	1340.	3500.	350.
8	520.	1060.	2670.	200.	8	660.	1360.	3700.	219.
8	537.	1360.	2920.	233.	8	730.	1500.	3600.	214.
8	480.	1345.	2680.	240.	8	510.	1250.	3320.	255.
8	540.	1200.	2860.	286.	8	520.	1140.	3320.	260.
8	570.	1200.	2970.	205.	8	540.	1430.	3320.	275.
8	350.	1250.	2750.	250.	8	530.	1580.	3200.	263.
8	522.	1044.	2800.	266.	8	475.	1250.	3150.	250.
8	546.	1090.	2400.	210.	8	460.	1210.	2750.	212.
8	462.	1240.	2310.	210.	8	500.	1640.	3580.	216.
8	512.	1211.	2630.	233.	8	450.	1440.	3500.	250.
8	467.	1167.	2595.	233.	8	490.	1460.	2860.	260.
8	460.	1045.	2504.	204.	8	570.	1320.	2840.	286.
8	480.	1105.	2400.	240.	8	450.	1350.	3000.	300.
8	420.	1200.	2600.	200.	8	520.	1600.	3150.	320.
8	460.	1260.	2640.	200.	8	540.	1420.	3050.	283.
8	544.	1088.	3000.	272.	8	600.	1440.	2900.	300.
8	500.	1000.	2850.	250.	8	550.	1420.	3040.	275.
8	466.	1330.	2750.	233.	8	570.	1500.	3000.	295.
8	466.	1165.	2800.	233.	8	630.	1310.	3270.	327.
8	450.	970.	3190.	225.	8	610.	1550.	3400.	322.
8	410.	940.	3040.	240.	8	470.	1000.	2450.	131.
8	500.	1230.	2520.	251.	8	480.	1020.	2100.	127.
8	480.	1230.	2750.	256.	8	470.	960.	2520.	128.
8	427.	1506.	2640.	251.	8	460.	1080.	2270.	128.
8	460.	1370.	2610.	240.	8	490.	1220.	2560.	136.
8	480.	960.	2820.	240.	8	570.	1060.	2290.	220.
8	484.	900.	2640.	242.	8	470.	1150.	2570.	234.
8	520.	1210.	2420.	202.	8	500.	1260.	3180.	263.
8	468.	1275.	2550.	212.	8	500.	1290.	3480.	264.
8	470.	1200.	2900.	222.	8	560.	1320.	3780.	282.
8	470.	1190.	2800.	237.					
8	500.	1200.	2450.	250.	9	340.	950.	2240.	112.
8	460.	1150.	2880.	230.	9	326.	900.	2190.	112.
8	335.	1049.	2470.	223.	9	315.	995.	2260.	117.
8	420.	1009.	2300.	210.	9	326.	1125.	2210.	125.
8	400.	1070.	2530.	282.	9	280.	650.	3300.	140.
8	450.	1050.	2450.	250.	9	260.	660.	3300.	137.
9	260.	720.	2100.	131.	9	290.	1000.	2300.	145.
9	260.	740.	3030.	132.	9	300.	600.	2300.	150.
9	244.	1090.	2260.	122.	9	230.	570.	2100.	148.
9	225.	1000.	2300.	125.	9	250.	690.	2080.	125.
9	380.	1010.	2140.	141.	9	270.	650.	2050.	130.
9	330.	800.	2130.	133.	9	324.	800.	2220.	162.
9	280.	670.	2140.	140.	9	290.	800.	2150.	139.
9	250.	720.	2190.	126.	9	360.	870.	2480.	224.
9	395.	1300.	2160.	146.	9	360.	865.	3100.	240.
9	400.	1320.	2150.	153.	9	440.	1290.	2530.	258.
9	294.	930.	2050.	133.	9	460.	1080.	2640.	269.
9	280.	1000.	2160.	140.	9	470.	1000.	2820.	263.
9	344.	960.	2150.	123.	9	378.	950.	2990.	272.

9	350.	1000.	2250.	125.	9	450.	1080.	2350.	225.
9	240.	1040.	2150.	160.	9	400.	1000.	2400.	200.
9	270.	930.	2280.	157.	9	400.	1250.	2500.	250.
9	250.	1000.	2100.	130.	9	405.	1080.	2500.	225.
9	210.	960.	1940.	140.	9	330.	760.	2870.	196.
9	324.	708.	2440.	120.	9	350.	710.	2760.	188.
9	387.	786.	2518.	157.	9	380.	770.	2900.	190.
9	380.	1040.	2260.	175.	9	340.	750.	2780.	187.
9	400.	1000.	2350.	200.	9	400.	1180.	2760.	235.
9	313.	838.	2340.	125.	9	396.	1120.	2560.	233.
9	288.	938.	2450.	125.	9	400.	980.	2630.	328.
9	333.	835.	2170.	167.	9	440.	990.	2900.	290.
9	280.	750.	2170.	198.	9	358.	640.	2560.	256.
9	270.	910.	2200.	107.	9	300.	750.	2500.	250.
9	260.	975.	2320.	108.	9	360.	930.	2260.	257.
9	280.	950.	2300.	140.	9	440.	1100.	2300.	220.
9	266.	920.	2300.	133.	9	450.	875.	2750.	250.
9	250.	760.	2780.	146.	9	420.	935.	2710.	233.
9	310.	650.	2580.	130.	9	420.	1000.	2500.	250.
9	290.	760.	2300.	162.	9	350.	1100.	2400.	275.
9	315.	850.	2025.	157.	9	375.	1124.	2685.	187.
9	350.	820.	2130.	125.	9	375.	1143.	2700.	188.
9	366.	772.	2058.	128.	9	390.	1020.	2700.	260.
9	232.	696.	2200.	116.	9	390.	990.	2600.	261.
9	222.	665.	2200.	117.	9	300.	850.	2800.	180.
9	320.	960.	2240.	160.	9	350.	840.	2300.	175.
9	320.	960.	2290.	160.	9	395.	810.	2900.	208.
9	350.	898.	2140.	116.	9	480.	955.	2960.	238.
9	300.	1092.	2365.	182.	9	419.	1050.	2850.	263.
9	272.	1089.	2350.	136.	9	390.	1060.	2800.	278.
9	350.	980.	2200.	125.	9	378.	1416.	2580.	236.
9	320.	918.	2100.	133.	9	380.	1430.	2610.	239.
9	210.	670.	1900.	140.	9	370.	933.	2520.	233.
9	240.	730.	1850.	148.	9	325.	750.	2500.	250.
9	380.	1110.	2050.	140.	9	370.	1000.	2470.	207.
9	385.	880.	2330.	148.	9	350.	1320.	2550.	220.
9	333.	800.	2130.	133.	9	380.	980.	3100.	240.
9	320.	840.	2150.	140.	9	420.	990.	2860.	267.
9	280.	990.	2330.	142.	9	380.	893.	2920.	190.
9	329.	877.	2550.	219.	10	444.	1300.	1625.	111.
9	340.	900.	2530.	230.	10	469.	1288.	1600.	109.
9	290.	670.	2380.	260.	10	510.	1210.	1570.	145.
9	330.	630.	2460.	275.	10	510.	1130.	1510.	145.
9	420.	1045.	3060.	300.	10	450.	1230.	1600.	125.
9	390.	960.	3030.	300.	10	460.	1300.	1650.	127.
9	308.	1025.	2650.	205.	10	405.	1490.	1760.	135.
9	330.	840.	2550.	220.	10	415.	1490.	1790.	138.
9	280.	850.	2500.	213.	10	560.	1510.	1800.	143.
9	440.	900.	3900.	233.	10	460.	1380.	1650.	132.
9	400.	650.	3800.	200.	10	480.	1410.	1760.	120.
9	350.	1280.	3650.	333.	10	470.	1330.	1700.	121.
9	340.	1160.	2950.	290.	10	532.	1500.	1890.	140.
9	300.	1280.	3150.	256.	10	538.	1460.	1818.	146.

9	400.	1300.	3700.	250.	10	440.	1250.	1625.	125.
9	420.	1110.	3640.	300.	10	480.	1160.	1520.	130.
9	505.	1050.	3400.	280.	10	539.	1370.	1800.	112.
9	600.	1200.	3600.	316.	10	549.	1353.	1728.	117.
9	550.	1100.	3470.	345.	10	370.	1520.	1670.	177.
9	620.	1100.	3250.	220.	10	460.	1330.	1590.	164.
9	600.	1280.	3650.	216.	10	360.	1300.	2800.	130.
9	360.	660.	3050.	274.	10	370.	1300.	1760.	133.
9	310.	730.	3500.	260.	10	488.	1468.	1712.	122.
9	420.	1500.	3010.	295.	10	472.	1465.	1725.	118.
9	450.	1330.	2840.	260.	10	570.	1300.	1750.	167.
9	403.	1100.	2950.	244.	10	565.	1370.	2440.	157.
9	363.	920.	2900.	242.	10	503.	1305.	1775.	120.
9	350.	1160.	3260.	290.	10	505.	1320.	1750.	120.
9	330.	1090.	3350.	273.	10	488.	1300.	1600.	163.
9	470.	1400.	2800.	275.	10	490.	1380.	1620.	163.
9	370.	1160.	2800.	286.	10	460.	1400.	1790.	107.
9	460.	1460.	3070.	307.	10	425.	1410.	1760.	103.
9	400.	1700.	3000.	300.	10	500.	1340.	1700.	128.
9	390.	1340.	2830.	280.	10	532.	1275.	1600.	133.
9	340.	1110.	3080.	284.	10	470.	1460.	1770.	126.
9	510.	1700.	3020.	283.	10	509.	1365.	1670.	144.
9	500.	1640.	3050.	278.	10	511.	1561.	1876.	150.
9	520.	1250.	3460.	345.	10	530.	1450.	1887.	138.
9	500.	1140.	3380.	334.	10	450.	1420.	1870.	111.
9	280.	980.	2240.	140.	10	472.	1430.	1840.	118.
9	270.	940.	2140.	134.	10	432.	1300.	1400.	120.
9	290.	1030.	2190.	137.	10	420.	1300.	1570.	111.
9	380.	900.	2380.	238.	10	543.	1286.	1643.	143.
9	310.	810.	2150.	239.	10	508.	1309.	1600.	145.
9	400.	850.	2190.	249.	10	512.	1280.	1570.	128.
9	380.	1010.	2770.	252.	10	512.	1280.	1570.	128.
9	400.	1080.	2610.	269.	10	559.	1324.	1618.	147.
9	390.	1080.	3080.	277.	10	532.	1330.	1600.	133.
9	410.	1070.	3190.	290.	10	554.	1480.	1800.	150.
9	340.	920.	3050.	188.	10	484.	1505.	1890.	128.
10	500.	1370.	1780.	100.	10	390.	1320.	1550.	128.
10	530.	1330.	1800.	106.	10	420.	1240.	1510.	124.
10	590.	1400.	1840.	150.	10	560.	1600.	1900.	200.
10	555.	1430.	1730.	145.	10	514.	1540.	1955.	206.
10	480.	1320.	1870.	120.	10	530.	1670.	2050.	222.
10	483.	1335.	1844.	127.	10	500.	1720.	1900.	200.
10	420.	1350.	1600.	150.	10	610.	1630.	2020.	246.
10	450.	1350.	1600.	150.	10	585.	1700.	1850.	225.
10	560.	1520.	2100.	140.	10	510.	1430.	1523.	206.
10	540.	1570.	2050.	140.	10	400.	1240.	1480.	201.
10	540.	1280.	1720.	122.	10	500.	1630.	2040.	240.
10	510.	1280.	1650.	118.	10	490.	1580.	2190.	243.
10	560.	1350.	1780.	150.	10	460.	1860.	2250.	230.
10	600.	1470.	1820.	150.	10	504.	1820.	2290.	214.
10	590.	1580.	1900.	226.	10	533.	1425.	1830.	213.
10	550.	1600.	1945.	229.	10	533.	1425.	1830.	213.
10	600.	1500.	2000.	250.	10	430.	1800.	1930.	205.

10	610.	1520.	1950.	254.
10	480.	1410.	1700.	243.
10	493.	1580.	1775.	243.
10	524.	1700.	2130.	193.
10	507.	1800.	2380.	180.
10	466.	1860.	2260.	233.
10	540.	1780.	2220.	225.
10	550.	1780.	2080.	182.
10	600.	1750.	2000.	201.
10	490.	2120.	2480.	177.
10	493.	1930.	2300.	197.
10	450.	1640.	2250.	225.
10	489.	1630.	2090.	233.
10	570.	2000.	2480.	286.
10	510.	1770.	2060.	196.
10	520.	1560.	1820.	260.
10	500.	1500.	1750.	250.
10	540.	1400.	1800.	200.
10	460.	1350.	1560.	204.
10	622.	1750.	2070.	230.
10	652.	1710.	2043.	225.
10	487.	1500.	1780.	217.
10	467.	1420.	1640.	206.
10	504.	1565.	1835.	180.
10	513.	1578.	1830.	183.
10	500.	1625.	1875.	250.
10	460.	1610.	1910.	230.
10	432.	1790.	2060.	216.
10	360.	1900.	2320.	205.
10	560.	1750.	2100.	200.
10	500.	1850.	2100.	200.
10	420.	1720.	1900.	220.
10	510.	1680.	1890.	255.
10	460.	1700.	1909.	230.
10	410.	1580.	1800.	225.
10	584.	1680.	2050.	225.
10	466.	1630.	1865.	233.
10	530.	1680.	2000.	263.
10	540.	1620.	2320.	270.
10	380.	1490.	1760.	226.
10	420.	1740.	1960.	200.
10	610.	2300.	2900.	210.
10	450.	2150.	2550.	200.
10	560.	1740.	2460.	275.
10	600.	1800.	2200.	302.
10	500.	1540.	1700.	250.
10	580.	1620.	1790.	242.
10	640.	1940.	2820.	320.
10	610.	2100.	2600.	265.
10	805.	1705.	2420.	310.
10	710.	1700.	2400.	310.
10	670.	2130.	2360.	222.
10	760.	2240.	2460.	205.
10	550.	1500.	1800.	250.
10	480.	1650.	1960.	239.
10	570.	1880.	2400.	272.
10	510.	1610.	1910.	255.
10	452.	1580.	1810.	226.
10	510.	1550.	1740.	232.
10	430.	1800.	2400.	240.
10	470.	1840.	2400.	233.
10	510.	1660.	2100.	268.
10	480.	1700.	1830.	250.
10	540.	1770.	2040.	300.
10	540.	2050.	2300.	286.
10	530.	1650.	1740.	280.
10	550.	1660.	1770.	286.
10	522.	1830.	2350.	261.
10	530.	1800.	2250.	282.
10	740.	1850.	2160.	308.
10	660.	1830.	2200.	328.
10	390.	1340.	1640.	129.
10	510.	1360.	1650.	127.
10	410.	1320.	1560.	128.
10	490.	1410.	1780.	132.
10	670.	1610.	2160.	240.
10	510.	1570.	1900.	253.

GRADUATE SCHOOL
UNIVERSITY OF ALABAMA AT BIRMINGHAM
DISSERTATION APPROVAL FORM

Name of Candidate Sung-Bok Lee

Major Subject Biomedical Engineering

Title of Dissertation A Study of Vowel Articulation in a Perceptual Space

Dissertation Committee:

Martin J. McCutcheon, Chairman

James E. Flege

Samuel G. Fletcher

Linda C. Lucas

A. V. Lakshminarayanan

Director of Graduate Program

Dean, UAB Graduate School

Date March 8, 1991

PS-1428