

Measurement of tongue movement during vowels production with computer-assisted B-mode and M-mode ultrasonography

YUH-CHYUN CHIANG, MD, FEI-PENG LEE, MD, CHIEN-LUN PENG, PHD, and CHE-TONG LIN, PHD, Taipei, Taiwan

OBJECTIVES: We sought to apply B- + M-mode ultrasonography in combination with the cushion-scanning technique (CST) in quantitative analysis of tongue movement during vowel articulation.

STUDY DESIGN AND SETTING: Twenty normal persons (10 women and 10 men) were enrolled in this study. Each individual was asked to enunciate 5 vowels: /e/, /i/, /a/, /o/, and /ju/. A noninvasive diagnostic technique, computer-assisted B-mode plus M-mode ultrasonography, was used in combination with the CST to assess their tongue movement. The sonographic signals were recorded on a video recorder and then transferred to a personal computer via a frame grabber for digital assessment.

RESULTS: M-mode images show an amplitude-time diagram, whereas B-mode images reveal midsagittal tongue configuration. The ranges of tongue movement in midsagittal plane in male speakers during production of the vowels /e/, /i/, /a/, /o/, and /ju/ were 8.29 ± 1.76 , 4.00 ± 0.78 , 13.82 ± 2.86 , 14.05 ± 1.63 , and 6.72 ± 1.66 mm, respectively; for female speakers, the averages were 7.19 ± 0.92 , 3.36 ± 1.31 , 12.74 ± 2.16 , 12.86 ± 2.18 , and 7.11 ± 2.09 mm, respectively.

CONCLUSIONS: The computer-assisted B-mode plus M-mode ultrasonography in combination with the CST provides a standardized and quantifiable ultrasonographic examination for the clinical investigation of tongue movement during vowel articulation. In our study group, there was no difference in the thickness of tongue or the range of tongue movement in midsagittal plane during articulation

of the selected vowels between males and females. Further exploration can be extended in the field of speech research by this valuable tool. (Otolaryngol Head Neck Surg 2003;128:805-14.)

The tongue is the most important articulator in speech production, yet it is difficult to obtain a direct view of the tongue during speech. Previously, cineradiography was used almost exclusively for evaluating tongue movement during speech.¹ Owing to considerable radiation exposure, the use of this technique has become limited for ethical reasons. Other measurement devices also provide data on the tongue and the vocal tract. Palatography, using a talc or powdery substance on the lingual surface to observe tongue contact, is a well-known traditional method of obtaining data on articulation.² In recent times, electropalatography (EPG) uses an artificial palate with built-in electrosensors to record tongue-palate contact during an utterance.³ In the case of electromagnetic articulography (EMA) or electromagnetic midsagittal articulometry (EMMA), transducers mounted in the helmet induce current in small coils attached on the articulators.^{4,5} However, due to insufficient display of tongue shape and physiologic disturbance of the artificial palate during tongue movements, as well as considerable cost for fabrication of individual experimental material, these techniques have limited use. Recently, magnetic resonance imaging (MRI) was used for speech production research.⁶ MRI permits observation of vocal tract features from any angle and makes possible 3-dimensional reconstruction of the entire vocal tract. However, the trade-off between imaging time and resulting image quality effectively limits application of the technique to studies of sustainable articulations. Image acquisition rates are currently still too slow by an order of magnitude to capture dynamic speech.

Ultrasonography has the advantage of allowing visualization of the actual soft tissues of the

From the Departments of Otolaryngology (Drs Chiang and Lee), Orthodontics (Dr Peng), and Prosthodontics (Dr Lin), Taipei Medical University Hospital.

This study was supported by National Scientific Council of Taiwan (grant: NSC89-2314-B-038-028).

Reprint requests: Chien-Lun Peng, PhD, Department of Orthodontics, College of Oral Medicine, Taipei Medical University, No. 250, Wu-Hsing Street, Taipei 110, Taiwan; e-mail, einstein@tmu.edu.tw.

Copyright © 2003 by the American Academy of Otolaryngology-Head and Neck Surgery Foundation, Inc.

0194-5998/2003/\$30.00 + 0

doi:10.1016/S0194-5998(03)00360-7

tongue and floor of the mouth. Over the past years, an increasing number of ultrasonographic examinations of tongue tumor have been carried out in an effort to improve preoperative staging and to help the surgeon to define adequate resection margins.^{7,8} The application of real-time scanning instrumentation, including a linear array transducer and a mechanical sector transducer, in the studies of tongue functions emerged in 1980.^{9,10} With improvement in the instrument, extended amount of tongue surface was visualized and intrinsic tongue musculature was detailed. Throughout the 1980s, many researchers used this technique to observe tongue functions, such as speaking or swallowing.¹⁰⁻¹² However, these previous ultrasonographic studies are subject to potential artifacts caused by varying transducer-skin contact surfaces of the submental area due to movements during function. Additional movements also occurred readily because the ultrasonographic transducers in previous studies were handheld by either the examinee or the examiner. Consequently, the images were not obtained in the exact same position from study to study.

A dynamic tongue imaging technique, the cushion-scanning technique (CST), was developed to ensure positional stability of the transducer during repeated trials.¹³ This sonographic technique has been used to investigate and measure tongue movement during swallowing in normal persons.¹⁴ The purposes of this study are to introduce this technique in the field of speech research and to determine if ultrasonography can be used to measure tongue movement during vowel production.

MATERIAL AND METHODS

Twenty medical students (10 women, 23.4 ± 0.5 years; 10 men, 23.5 ± 0.7 years) were enrolled in this study. All were screened by one certified speech pathologist and were found to be free of oral-motor, structural, or rhythmic impairment. Our study was performed in compliance with guidelines for research involving humans and with the informed consent for all participants. A non-invasive diagnostic technique, computer-assisted B-mode plus M-mode ultrasonography, was used in combination with the CST to assess tongue movement. The CST is based on 6 components,

which are illustrated in Fig 1. A 500-mL polyvinyl chloride (PVC) bag (18 cm \times 11.5 cm) for intravenous fluid injection (Y. F. Chemical Corp, Taipei, Taiwan) filled with water as transmedium was used as a cushion. The PVC bag was connected to 2 drainage pipes, ensuring constant pressure in the cushion and an even distribution of local pressure arising from movements of the submental area or the entire mandible. The cushion device was supported by a plastic plate with a semicircular cutout (5.5 cm in diameter) for the entry of the ultrasound transducer (Fig 2).

The ultrasound transducer was fixed by a holder with adjustable hinges, which had scales to allow for reproducible registration of different scanning directions. The head supporter provided a firm and stable support for the forehead. This was supplemented with a head position recording device, which consisted of frontal and lateral transparent acrylic plates. These plates allowed the head to be oriented repeatedly in relation to the Frankfort horizontal plane, extending from the upper rim of the external auditory canal to the lower rim of the orbit.

Tongue movement during the production of 5 vowels was investigated with B-mode and M-mode sonographic technique (Panasonic Panavista-LSC I with 3.5 MHz, 13 mm in diameter, 100-degree mechanical sector transducer; Matsushita Corp, Tokyo, Japan). The ultrasound transducer was placed midway between the posterior border of the symphysis and the anterior margin of the hyoid bone in the midsagittal plane. At the same time, the transducer was oriented with its long axis perpendicular to the Frankfort horizontal plane. Ultrasound gel was applied to all coupling surfaces. The scan line (M-position) was placed through the middle of the B-mode sector image (Fig 3). Image contrast was deliberately enhanced to emphasize the tongue surface. The sonographic signals were recorded on a digital video recorder (Sony DCR-TRV 110; Sony Corporation; Tokyo, Japan) and then transferred into an IBM-compatible personal computer via a frame grabber (Upmost DV FUN; Upmost Technology Corp; Taipei, Taiwan) for digital assessment. The uppermost and the lowest positions on the surface of the tongue on the M-mode image of each phonation were digitized

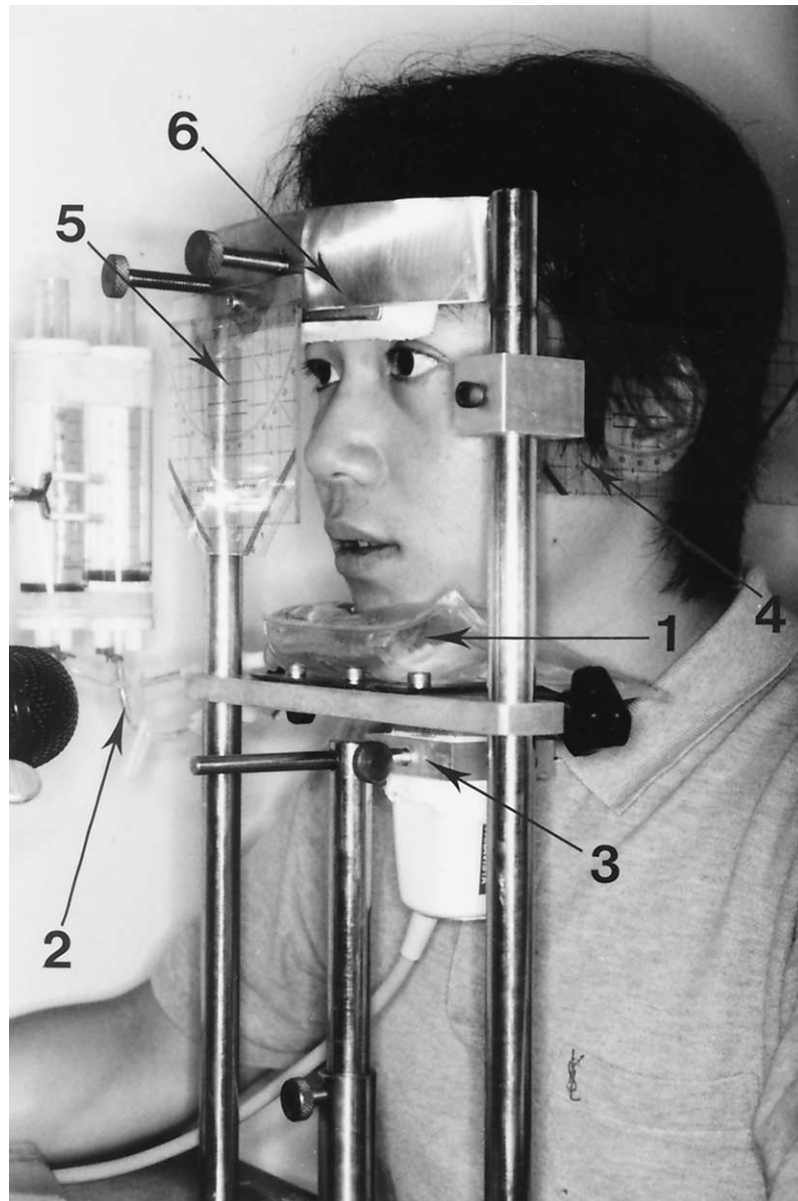


Fig 1. Illustration of the components of the CST system used in this study. (1, cushion; 2, drainage pipe; 3, ultrasound transducer holder; 4, lateral head position recording device; 5, frontal head position recording device; 6, head support).

and measured with the help of 2 graphic programs (Video Studio 3.0 SE; Ulead Systems Inc, Taipei, Taiwan; and Photoshop 5.0; Adobe Systems Inc, San Jose, CA). The distance between the 2 positions (the amplitude) was defined as the range of tongue movement.

All participants were told to place their tongue on the floor of the mouth holding the tongue still for 5 seconds with the mouth closed. During this period, continuous imaging was performed and

recorded. This provided the resting configuration of the tongue on the B-mode image. Any fasciculations or tremors, if ever occurred, would be easily revealed on the M-mode image. Thickness of the tongue was measured midsagittally along the scan line from the surface of the tongue to the floor of the intramuscular septum, the junction of the tongue with the geniohyoid-myohyoid muscles (Fig 3). Thereafter, each individual was asked to enunciate each vowel according to the phonetic

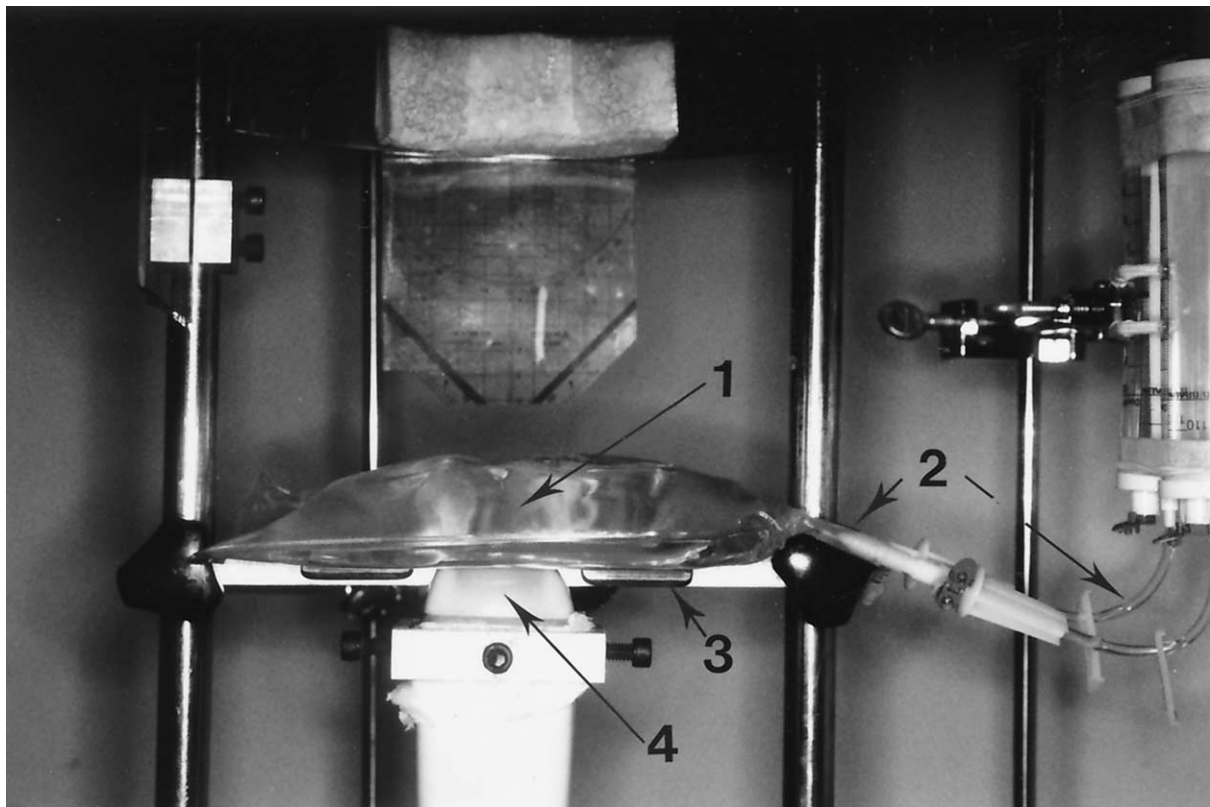


Fig 2. Illustration of the relationship between the cushion device and the ultrasound transducer (1, cushion; 2, drainage pipe; 3, cushion device supporter; 4, ultrasound transducer).

symbols on the paper shown and to read each [e], [aI], [o], and [ju] 3 times at the rate of one vowel per second. No production cues or information regarding correctness of an utterance were provided.

The range of tongue movement of each phonation from these 20 examined individuals was calculated and analyzed. All measurements were performed by the same investigator to diminish interexaminer variability.

To evaluate the intraindividual reproducibility, one randomly selected participant was asked to pronounce each of the 5 vowels 10 times. Coefficient of variation (CV) was calculated for these phonations.

Data analysis was computed using SPSS for Windows (SPSS Inc, Chicago, IL). The Student's *t* test was used to determine whether there were differences in tongue thickness or range of tongue movement between the female and the male participants.

RESULTS

Range of Tongue Movement

Normal resting nonphonation thickness of tongue, measured from the tongue surface to the floor intramuscular septum, averaged 40.91 ± 3.36 mm in male speakers and 39.55 ± 3.81 mm in female speakers. As shown in Table 1, average range of tongue movement in male speakers during production of the vowels /e/, /i/, /aI/, /o/, and /ju/ was 8.29 ± 1.76 , 4.00 ± 0.78 , 13.82 ± 2.86 , 14.05 ± 1.63 , and 6.72 ± 1.66 mm, respectively. The averages for female speakers were 7.19 ± 0.92 , 3.36 ± 1.31 , 12.74 ± 2.16 , 12.86 ± 2.18 , and 7.11 ± 2.09 mm, respectively. There was no difference in thickness of tongue or range of tongue movement during phonation of these vowels of the two study groups ($P > 0.05$).

Midline Sagittal Tongue Configuration

The tongue configurations along the midline sagittal section for the 5 tested vowels are shown

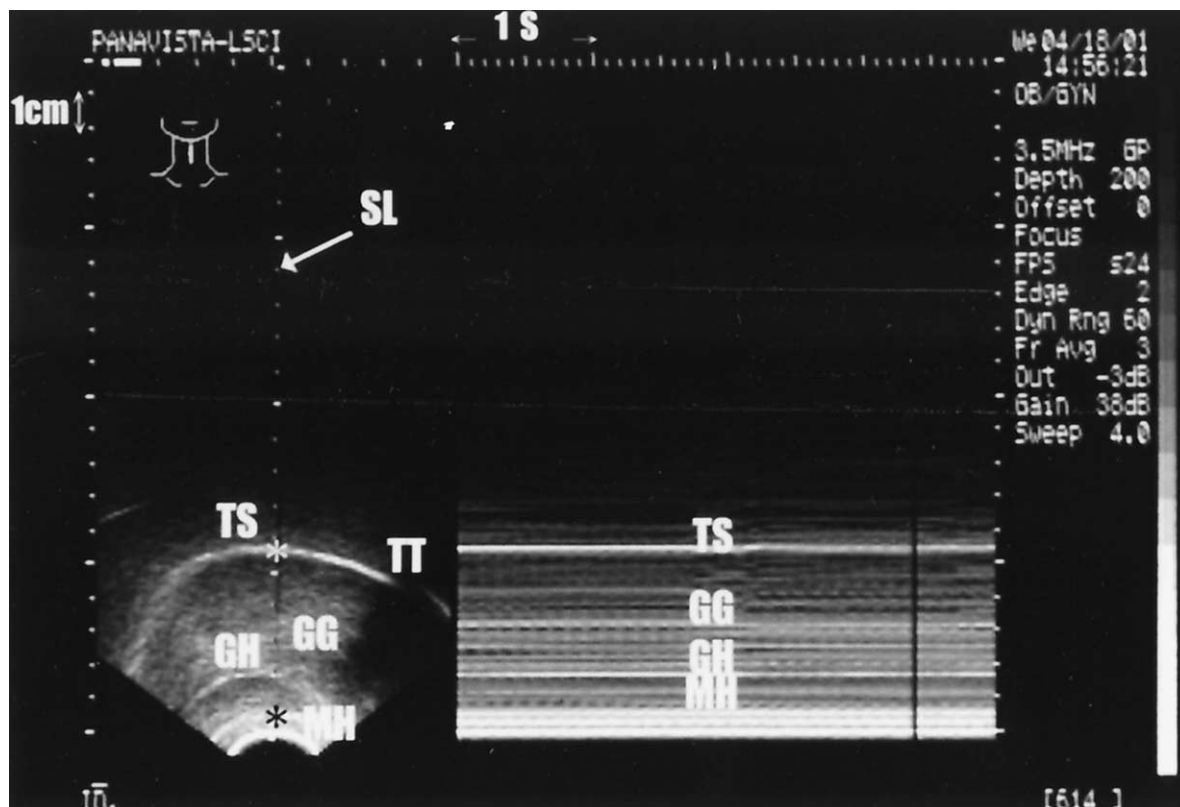


Fig 3. M-mode ultrasonogram of the resting, nonphonating tongue. Both B- and M-mode images illustrate anatomic structures of tongue along the chosen scan line (SL). The distance between the 2 asterisks was measured as the thickness of the tongue: tongue surface (TS), geniohyoid muscle (GH), mylohyoid muscle (MH), genioglossus muscle (GG).

Table 1. Range of tongue movement in mid-sagittal plane during five vowels articulation in the male and the female speakers

Range of tongue movement (mm)	/e/	/i/	/aI/	/o/	/ju/
	Male/ Female	Male/ Female	Male/ Female	Male/ Female	Male/ Female
Mean	8.29/7.19	4.00/3.36	13.82/12.74	14.05/12.86	6.72/7.11
Maximum	11.04/8.06	5.22/4.78	22.39/15.82	17.01/16.12	9.85/9.40
Minimum	5.22/5.52	3.13/1.49	10.75/10.60	11.49/9.40	4.03/3.73
Standard Deviation	1.76/0.92	0.78/1.31	2.86/2.16	1.63/2.18	1.66/2.09

in Fig 4A-E. Observation of these B-mode images revealed that each of these 5 vowels showed regular configurations. In general, the tongue root moved forward with the anterior part of tongue shoveled for the vowel /e/ and /i/. On the contrary, for the vowels /aI/ and /o/, the tongue root moved backward with the anterior tongue flattened. When phonating the diphthongal vowel /ju/, biphasic movement of tongue was noted; there was initial forward and later

backward motion of the posterior tongue. The anterior tongue depressed slightly inferiorly and was retracted posteriorly.

Intraindividual Reproducibility

The CV obtained from a single randomly selected participant during 10-time production of the vowels /e/, /i/, /aI/, /o/, and /ju/ was 7%, 5%, 4%, 5%, and 6%, respectively.

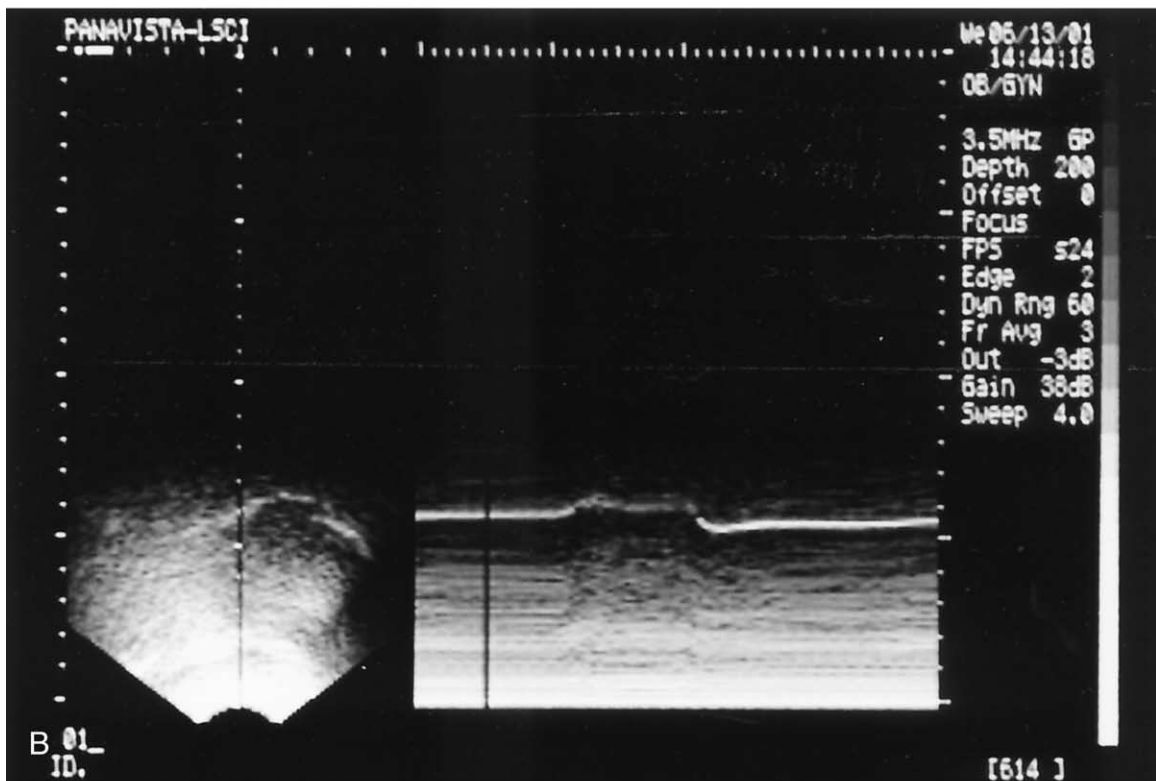
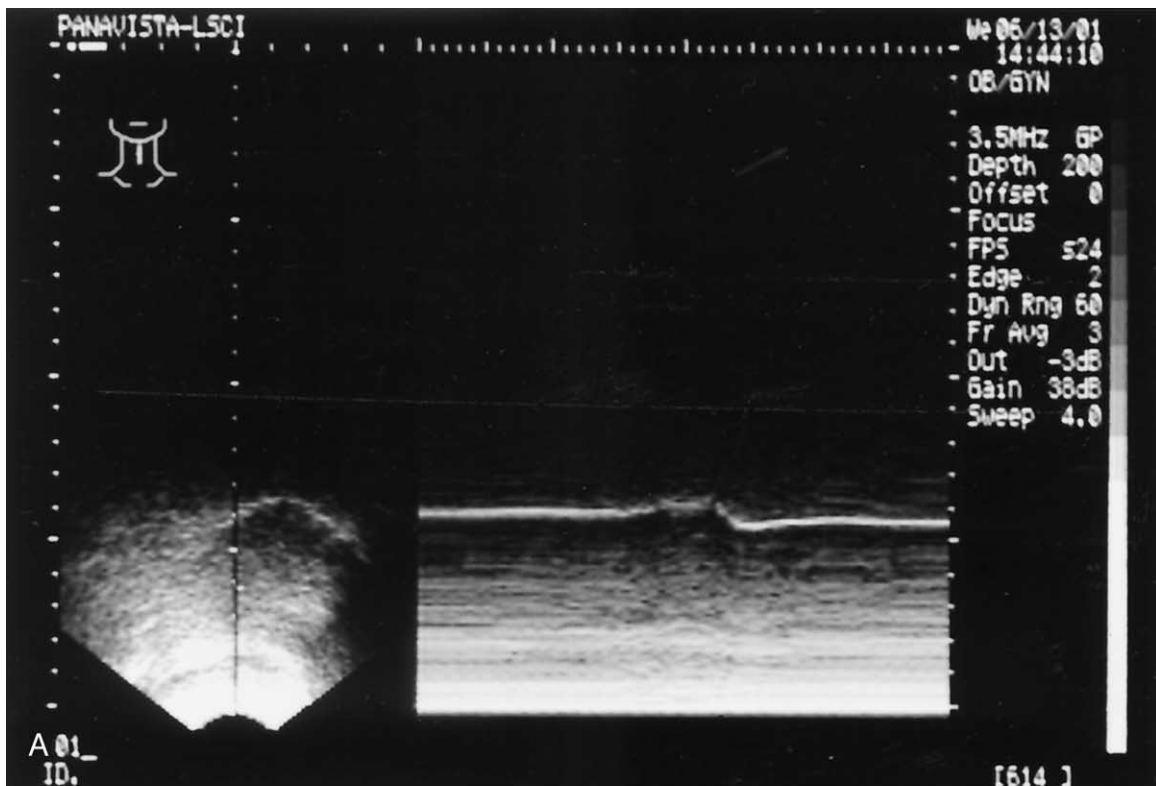


Fig 4. A-E, Five B- + M-mode ultrasonogram of one participant during production of /e/ (A), /i/ (B), /a/ (C), /o/ (D), and /ju/ (E). On the M-mode image (right side), the uppermost and the lowest positions on the surface of the tongue of each phonation were measured as the range of tongue movement.

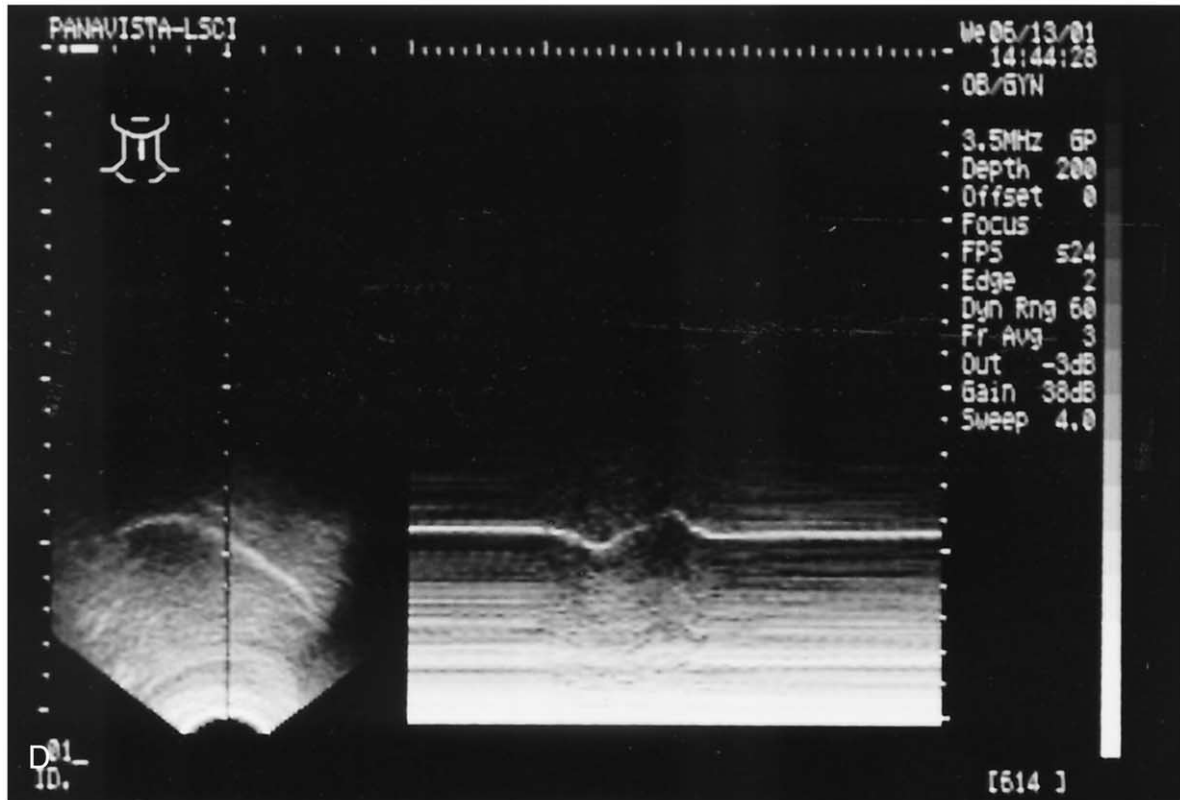
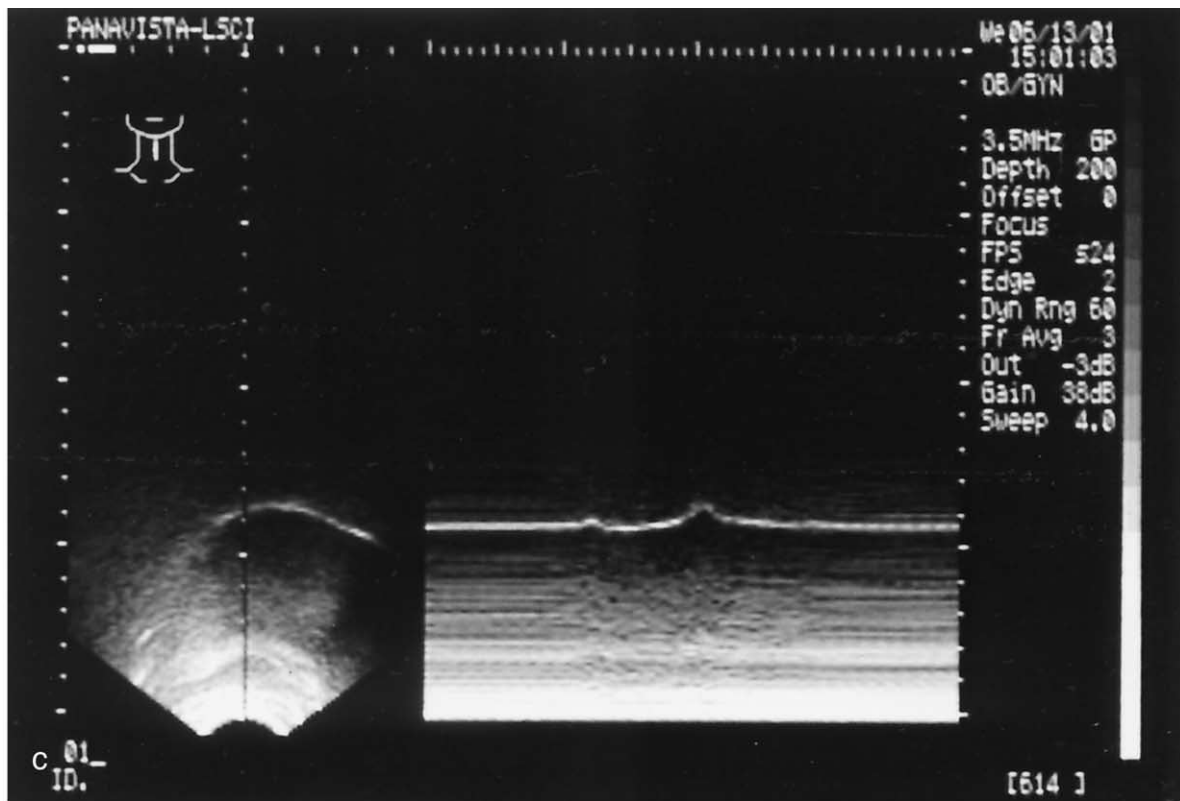


Fig 4. (Continued).

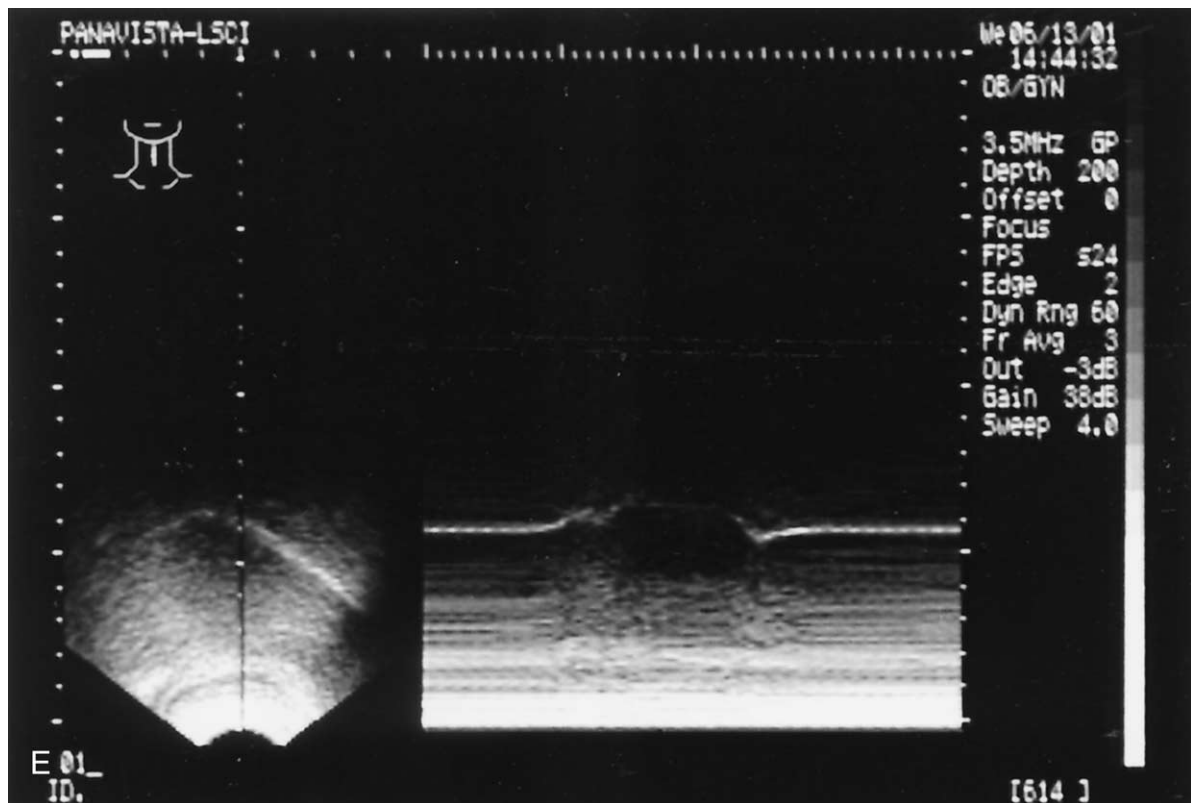


Fig 4. (Continued).

DISCUSSION

The main principle of the CST is to create a buffer area between the transducer and the skin so as to avoid accidental transducer movements during tongue function. In addition, the cushion device provides more comfortable and stable experimental status than those of the conventional techniques.¹³

An advantage of using M-mode ultrasonography in observation of tongue movement is the determination of the position of the palate. Previously, the palate was not visible unless the tongue had been filled with water. With the CST, the position of the palate can be determined easily by asking the participant to swallow. The uppermost level in the M-mode image during swallowing, indicating the contact of tongue to palate, can serve as a reference to the position of the palate.

In this study, we set the M position at the center of tongue in midsagittal plane because this area is easier to define identically on persons with various tongue sizes. This provides a

premise for statistical comparison between experimental individuals. Furthermore, the midsagittal view of M-mode ultrasonography gives the observer a brief review of the whole tongue movement during function. However, such brief review possessed some limitations: The tip of the tongue and the posterior part of tongue were not visualized in the M-mode sonogram. Nevertheless, because the tongue is a musculature organ, the movements in the tongue tip or posterior part of tongue may more or less be manifested in the central part of tongue and may be detected with a midsagittal scan line.

We applied CST in combination with B-mode and M-mode ultrasonography to investigate the tongue movement during production of five vowels, including 2 diphthongal vowels ([aI] and [ju]). The sonogram obtained during the experiment was clear and stable. On B-mode image, midsagittal tongue configuration was visualized and much of the intrinsic and extrinsic soft tissue anatomy of tongue was identified. These observed tongue configurations and posi-

tions are consistent with those obtained by previous ultrasonic observation.^{9,10} Also, continuous movement of tongue, as well as changes in soft tissue organization, during the production of these vowels was perceived on real-time basis. Tongue movement and configuration were found to be consistent for these normal speakers. On M-mode image, a time-amplitude diagram was provided. This offered quantitative data for digital assessment of tongue movement. The CV values obtained from the randomly selected participant were low. The results suggested high intraindividual reproducibility.

In 1984, Shawker and Sonies¹⁵ first used real-time ultrasonography scanning as biofeedback therapy for speech training. The ultrasound image of tongue, scanned by the patient with the transducer held by under her chin, functioned as a discriminatory sensory feedback stimulus to retrain the patient. Their preliminary results on a 9 year-old girl with a single persistent articulatory defect seemed promising. However, images produced with the hand-held transducer were essentially unstable. Borrowing from this concept, we are currently conducting pilot studies with the CST on children with hearing impairment to determine whether this technique has therapeutic potential.

We acknowledge various limitations to this study. First, we recognize that our study group was small ($n = 20$). The results of this study are preliminary. In fact, because of the way in which measurements were taken in a single spot on the tongue surface and only in a single plane (midsagittal), the whole tongue activity is not fully represented. To characterize vowel production more correctly and comprehensively, observation made on coronal ultrasound view should also be included in the future studies. This presentation, nevertheless, does point out many of the advantages of CST over previous techniques in observation of tongue movement during vowel production. Our results have revealed that digital analysis of tongue movement during vowel production is readily available by this simple method. The study should be repeated on a larger sample with more special designed scan line to provide more representative statistical data, which may serve as the

normal databank for future application in the diagnostic protocol of speech researches.

CONCLUSION

The computer-assisted B-mode plus M-mode ultrasonography in combination with the cushion-scanning technique provides a standardized and objective ultrasonographic examination for the clinical investigation of tongue movement during vowel articulation. Real-time observation, as well as quantitative measurement of tongue movement during articulation, is readily available. In our study group, there was no difference in thickness of tongue or range of tongue movement in midsagittal plane during production of the vowels /e/, /i/, /aI/, /o/, and /ju/ between the males and the females. It is our aim to continue applying this valuable tool on future patients to reach a new realm in the field of speech research.

We are indebted to those students who spent numerous hours to make our research successful. In addition, we would like to thank Dr Hung-Yi Chiou and Jayne Tchen-Ru Chen for their valuable suggestions and corrections.

REFERENCES

1. Fujimura O, Ishida Y, Kiritani S. Computer-controlled radiography for observation of movements of articulatory and other human organs. *Comp Biol Med* 1973;3:371-84.
2. Harley WT. Dynamic palatography—a study of linguopalatal contacts during the production of selected consonant sounds. *J Prosthet Dent* 1972;27:364-76.
3. Hardcastle WJ, Gibbon FE, Jones W. Visual display of tongue-palate contact: electropalatography in the assessment and remediation of speech disorders. *Br J Disord Commun* 1991;26:41-74.
4. Schönle PW, Grabe K, Wenig P, et al. Electromagnetic articulography: use of an alternating magnetic field for tracking movements of multiple points inside and outside the vocal tract. *Brain Lang* 1987;31:26-35.
5. Perkell JS, Cohen MH, Svirsky MA, et al. Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements. *J Acoust Soc Am* 1992; 92:3078-96.
6. Narayanan SS, Alwan AA, Haker K. An articulatory study of fricative consonants using magnetic resonance imaging. *J Acoust Soc Am* 1995;98:1325-47.
7. Ikezoe J, Nakanishi K, Morimoto S. Sonographic staging of cancer of the mobile tongue. *Acta Radiol* 1991;32:6-8.
8. Helbig M, Flechtenmacher C, Hansmann J, et al. Intraoperative B-mode endosonography of tongue carcinoma. *Head Neck* 2001;23:233-7.
9. Niimi S, Simada Z. Ultrasonic investigation of tongue shape—a preliminary report. *Jpn J Logop Phoniat* 1980; 21:121-5.

10. Shawker TH, Sonies BC. Tongue movement during speech: a real-time ultrasound evaluation. *J Clin Ultrasound* 1984;12:125-33.
11. Sonies BC, Shawker TH, Hall TE, et al. Ultrasonic visualization of tongue motion during speech. *J Acoust Soc Am* 1981;70:683-6.
12. Shawker TH, Sonies BC, Stone M, et al. Real time ultrasound visualization of tongue movement during swallowing. *J Clin Ultrasound* 1983;11:485-90.
13. Peng CL, Jost-Brinkmann PG, Miethke RR. The cushion scanning technique: a method of dynamic tongue sonography and its comparison with the transducer-skin coupling scanning technique during swallowing. *Acad Radiol* 1996;3:239-44.
14. Peng CL, Jost-Brinkmann PG, Miethke RR, et al. Ultrasonographic measurement of tongue movement during swallowing. *J Ultrasound Med* 2000;19:15-20.
15. Shawker TH, Sonies BC. Ultrasound biofeedback for speech training: instrumentation and preliminary results. *Invest Radiol* 1985;20:90-3.

BOUND VOLUMES AVAILABLE TO SUBSCRIBERS

Bound volumes of *Otolaryngology-Head and Neck Surgery* are available to subscribers (only) for the 2005 issues from the Publisher, at an individual cost of \$118.00 (\$146.59 for Canadian, \$137.00 for international subscribers) for Vols. 128 (January-June) and 129 (July-December). Shipping charges are included. Each bound volume contains subject and author indexes, and all advertising is removed. Copies are shipped within 60 days after publication of the last issue in the volume. The binding is durable blue buckram with the Journal name, volume number, and year stamped in gold on the spine. *Payment must accompany all orders.*

Contact Mosby, Inc, Subscription Customer Service, 6277 Sea Harbor Dr, Orlando, FL 32887; phone 800-654-2452 or 407-345-4000.

Subscriptions must be in force to qualify. Bound volumes are not available in place of a regular Journal subscription.