COMPRESSIBILITY ANALYSIS OF THE TONGUE DURING SPEECH

Devrim Unay[1], Cengizhan Ozturk[1], Maureen Stone[2]

[1] Institute of Biomedical Engineering, Bogazici University, 80815 Bebek, Istanbul, Turkey.

[2] University of Maryland School of Dentistry, Baltimore MD, USA

Abstract – **The motivation of this study is to observe the compressibility of the tongue during speech. The tongue has a complicated muscular structure. Real-time MRI (16 frames/s) with tagging has been used for imaging during the utterance of four short syllables, "sha", "ga", "ta" and "ba." Four-dimensional parametric motion field analysis has been used which allows point tracking everywhere on the tongue. In this paper, 3D compression and expansion analysis of the tongue will be presented. Patterns of expansion and compression have been compared for different syllables and various repetitions of each syllable. The long-term objective of this research is to provide important information about the motion of the internal tongue muscles and shed light on the intricate relationship between these muscles and the final shape of the tongue during regular and abnormal speech patterns.**

Keywords –tongue motion, speech analysis, 4D B-splines

I. INTRODUCTION

This paper analyzes and describes the compressibility of the tongue during speech using real-time MR images with tagging at a 62 msec temporal resolution. We compared the maximum compression and expansion of the tongue during the utterance of four short syllables, "sha", "ga", "ta" and "ba". The acquisition of images was done several times for each syllable and the compression/expansion of the tongue was measured.

Previous work on this topic relied on finite element methods [1] or ultrasound imaging, the latter being restricted to surface analysis [2]. Early MRI studies in this field have examined non-speech movements [3] and static speech [4]. For dynamic MR imaging; first, several oblique slices that are orthogonal to the surface were proposed [5]. Also, a synchronized sampling technique was borrowed from cardiac imaging for tongue imaging [6]. Using this technique, last year for the first time, MR tagging for tongue motion analysis was introduced [7] and the principle strains of the tongue were obtained [8]. At that time, the subject had to repeat each syllable 13 times for each slice. In this study, sample images, compression and expansion analysis of real time MRI (16 images/s) with tagging are presented based on a single repetition per slice.

II. METHODS

A 26-year-old male, native speaker of English with no dental fillings was used for the study. Four consonant-to-vowel syllables ("sha," "ga," "ta," "ba") were chosen as the speech material for their large tongue motions and short durations. Only one repetition per slice was needed to acquire all 16 phases of a given slice. However, the subject repeated each syllable several times at a rate of one per second, from an auditory cue, to provide repeatability data. At each slice, three syllable repetitions were acquired (only two for "ga"). A gating input was provided to the scanner in sync with the cue. During one second, 16 phases were obtained for a given slice (Figure 1). The acquisition was repeated for all the syllables and slices (3 sagittal and 8 axial slices), with grid tagging. We used a modified gated multi-echo SPGR sequence in a 1.5 T GE Signa Cardiac Scanner. (TR 11.3 ms, TE 1.5 ms, FA 10, ETL 8, NVP 40, BW 125, 128 x 96 matrix, NEX 0.8, FOV 30x15 cm, ST 7mm, tag spacing 5). Following contouring and tag segmentation (using XBS and Findtags) a four-dimensional B-spline based parametric motion field was computed for a volume of interest encompassing the whole tongue [9]. Using this motion field, 3D displacements, velocities and strains can be calculated at any material point inside the tongue.
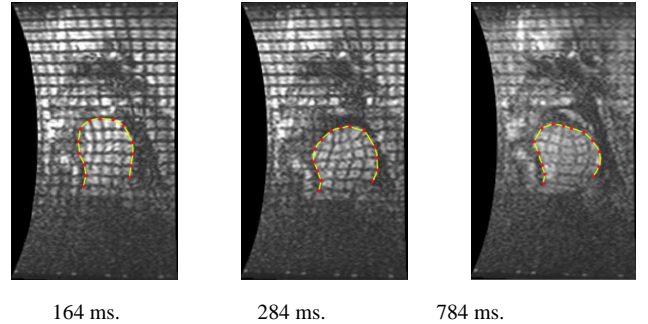

164 ms.   284 ms.   784 ms.

Figure 1: Real-time, tagged MRI images of the tongue during the syllable "ga" at three selected time frames.

4D B-spline Motion Field

We describe the motion of the tongue as a four-dimensional B-spline forward motion field:

$$\phi^s\left(\vec{P},t\right)=\sum_u\sum_v\sum_w\sum_t B_u(x)B_v(y)B_w(z)B_t(t)\vec{C}^{(s)}_{u,v,w,t*}$$

By this motion field we have a parametric representation of the material coordinate displacement field. The displacement gradient can be computed at each point

$$g_{i,j}(x,y,z)=\frac{\partial U_i}{\partial X_i}$$

where U_i is the *i*th component of the material displacement field and X_i is the material coordinate position vector. The

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
25OCT2001	N/A	-

Title and Subtitle	Contract Number
Compressibility Analysis of the Tongue During Speech	
	Grant Number
	Program Element Number

Author(s)	Project Number
	Task Number
	Work Unit Number

Performing Organization Name(s) and Address(es)	Performing Organization Report Number
Institute of Biomedical Engineering, Bogazici University, 80815 Bebek, Istanbul, Turkey.	

Sponsoring/Monitoring Agency Name(s) and Address(es)	Sponsor/Monitor's Acronym(s)
US Army Research, Development & Standardization Group (UK) PSC 802 Box 15 FPO AE 09499-1500	
	Sponsor/Monitor's Report Number(s)

Distribution/Availability Statement
Approved for public release, distribution unlimited

Supplementary Notes
Papers from the 23rd Annual International Conference of the IEEE ENgineering in Medicine and Biology Society, October 25-28, 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom., The original document contains color images.

Abstract

Subject Terms

Report Classification	Classification of this page
unclassified	unclassified

Classification of Abstract	Limitation of Abstract
unclassified	UU

Number of Pages
4

displacement gradient at any point can then be computed by differentiating the 4D B-spline motion field as:

$$g_{x,x} = \sum_u \sum_v \sum_w \sum_t B'_u(x) B_v(y) B_w(z) B_t(t) C'_{u,v,w,t}$$

where B' is the derivative of the B-spline basis function. The determinant of this point displacement gradient gives us the local compression/expansion value at that point, i.e. alpha.

III. RESULTS & DISCUSSION

In our previous study we analyzed the 2D displacement of the surface points of the tongue. These points were found by intersecting the tag lines with surface contours at each time frame. It was shown that tongue motion was consistent with respect to subject repeatability [10].

In this study we completed the full 4D analysis of the tongue motion for all syllables. For the repetition analysis, we display here the results of one syllable "sha," in which the tongue moves down and back. In Figure 2, if we observe the first image group (sagittal), we see that for the first repetition, there is compression (red) in the upper-central and lower-front regions of the tongue, and expansion under the tip and in the root. Repetitions 2 and 3 show expansion either below the tip (2) or in the root (3), indicating a trade-off between these two regions. Similarly, repetition 2 does not have lower-front compression. Muscle trade-offs allow different muscle combinations to produce the same surface shape, facilitating coordination of complex speech gestures. In the axial image group, the center of the tongue compresses maximally. The outer regions show expansion. The compression region differs a little for the second repetition. It shifts to the left-bottom. In all repetitions, the tongue performs about 22 per cent compression and 20 per cent expansion. The range of alpha values are; for the first repetition 1,218 & 0,785, for the second repetition 1,234 & 0,788, and for the third repetition 1,208 & 0,818. These values indicate that tongue compression values of different repetitions for syllable "sha" are consistent.

Compression and expansion percentages of repetitions of the other syllables are; 18 per cent compression and 26 per cent expansion for "ta"; 23 per cent compression and 33 per cent expansion for "ga"; 23 per cent for both compression and expansion for "ba". The range of alpha values are; for "ta" 1,149 & 0,802 (first repetition), 1,259 & 0,700 (second rep.), 1,176 & 0,691 (third rep.); for "ga" 1,291 & 0,580 (first), 1,118 & 0,730 (second), 1,230 & 0,669 (third); and for "ba" 1,263 & 0,776 (first), 1,214 & 0,769 (second). The maximum and minimum compressions are observed in the first and second repetitions of syllable "ga" with 29 and 12 per cent, respectively. Maximum expansion is in the first repetition of "ga" with 42 per cent, whereas the minimum expansion is observed in the third repetition of "sha" with 18 per cent.

If we are to observe compression of tongue for different syllables, let's take a look at the Figure 4. In the first column all the syllables start with the first image at about 40 msec. For "sha", we observe two compression regions; one in the upper-center and the other in the bottom-left part of the tongue. The latter can be a trade-off or due to artifacts sometimes seen at the edges of segmented volumes in our motion field analysis. The syllable "ta" compresses in the lower tongue tip, consistent with rapid tip lowering from the "t" to the "a.". "Ba" compresses in the upper-back part of the tongue. Since the "b" does not engage the tongue, motion into the "a" is from a neutral position. In that case, unlike the others the tongue moves up and back, compressing as it reverses direction. On the other hand, as a surprising result, we do not observe that much compression in "ga". We would have expected a downward and backward compression. Let's take a look at the axial compressions of syllables (Figure 4). The tongue compresses in the middle for all the syllables, except "ta". "Sha" compresses in the middle-left region of the tongue, whereas "ta" compresses in the front. "Ga" exhibits compression beginning in the bottom-center and spreading to the right of the tip, however "ba" has compression limited to the center region. We observe expansion in the front-right regions of the tongue in "sha", but "ta" mainly expands in the back. Expansion in "ga" is observed in the left-outer side of the tongue, however "ba" expands around the edges.

Maximum compression and expansion values of tongue motion for the first repetitions of all the syllables are; for "sha" 1,218 & 0,785; for "ta" 1,149 & 0,802; for "ga" 1,291 & 0,580; and for "ba" 1,263 & 0,776. "Sha", "ta", "ga", "ba" perform about 21, 17, 30, 25 per cent compression and expansion, respectively. Within the first repetitions of the syllables we observe maximum compression and expansion in "ga", minimum compression and expansion in "ta".

IV. CONCLUSION

We show for the first time detailed compression and expansion analysis during the uttering of certain syllables everywhere in the tongue. These results came as the first by-products of the four-dimensional motion field analysis of tagged-MR imaging of the tongue. Detailed findings of this initial study provide us the framework for planning detailed experiments, where we aim for detailed motion analysis during normal and abnormal speech.

V. ACKNOWLEDGMENT

VI. REFERENCE

[1] Wilhelms-Tricarico R., J. Acoust. Soc. Am. 97(5):3085-98, 1995.
[2] Akgül Y. S., et al., IEEE Trans. on Med. Imag. 18. 1035-45, 1999.
[3] Napadow V.J., et al J. of Biomechanics 32:1-12, 1999.
[4] Baer T. et al., J. Acoust. Soc. Am. 90(2): 779-828, 1991.
[5] Demolin D., et al., Proc. ICSLP-96, 272-275, 1998.
[6] Masaki S., et al., J. Acous. Soc. Jap. (E), 20, 5. 1999.
[7] Dick D. et al, Proc. of ISMRM, 8: 553, 2000.
[8] Stone M. et al. Proc. of 5th Speech Prod. Sem., Kloster-Seeon, Germany, p. 133, 5/ 2000.
[9] Ozturk C., et al., Physics in Med.&Biol., 45(6):1683-1702, 2000.
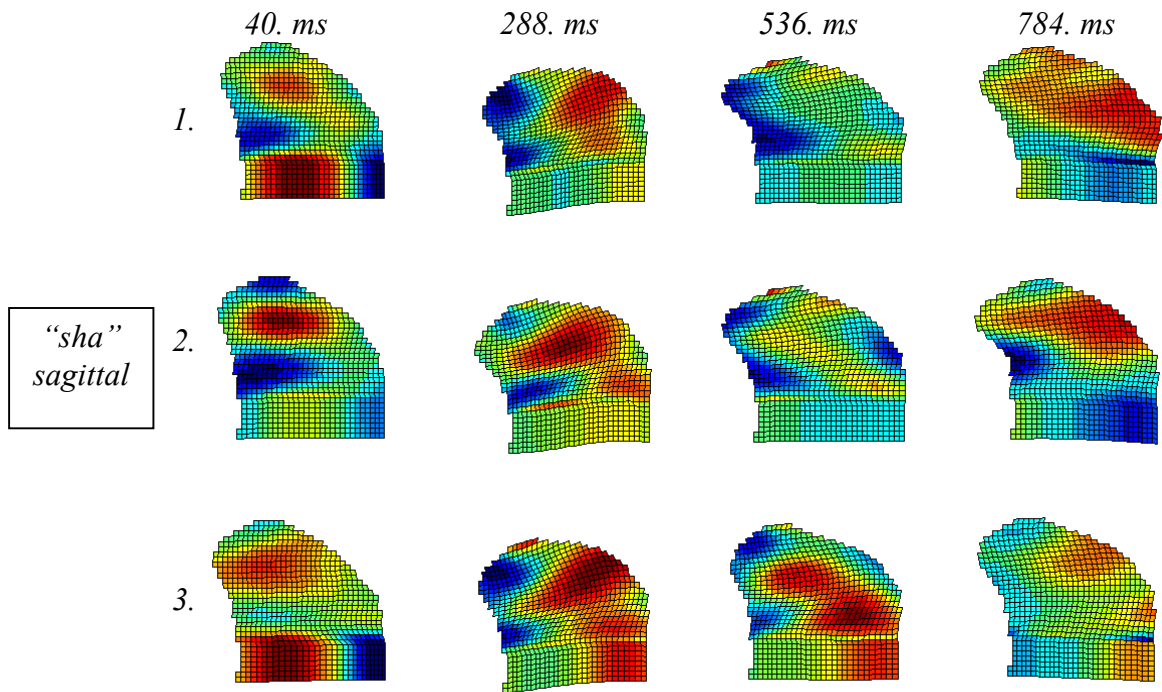[10] Ozturk C., et al., poster #1609, ISMRM Joint Annual Meeting, April 21-27, 2001, Glaskow, Scotland.

Figure 2: 4D B-spline motion field analysis of sagittal tongue motion for all three repetitions of syllable "sha". Images of each column belong to a unique time interval. Each row of images belongs to the same repetition.
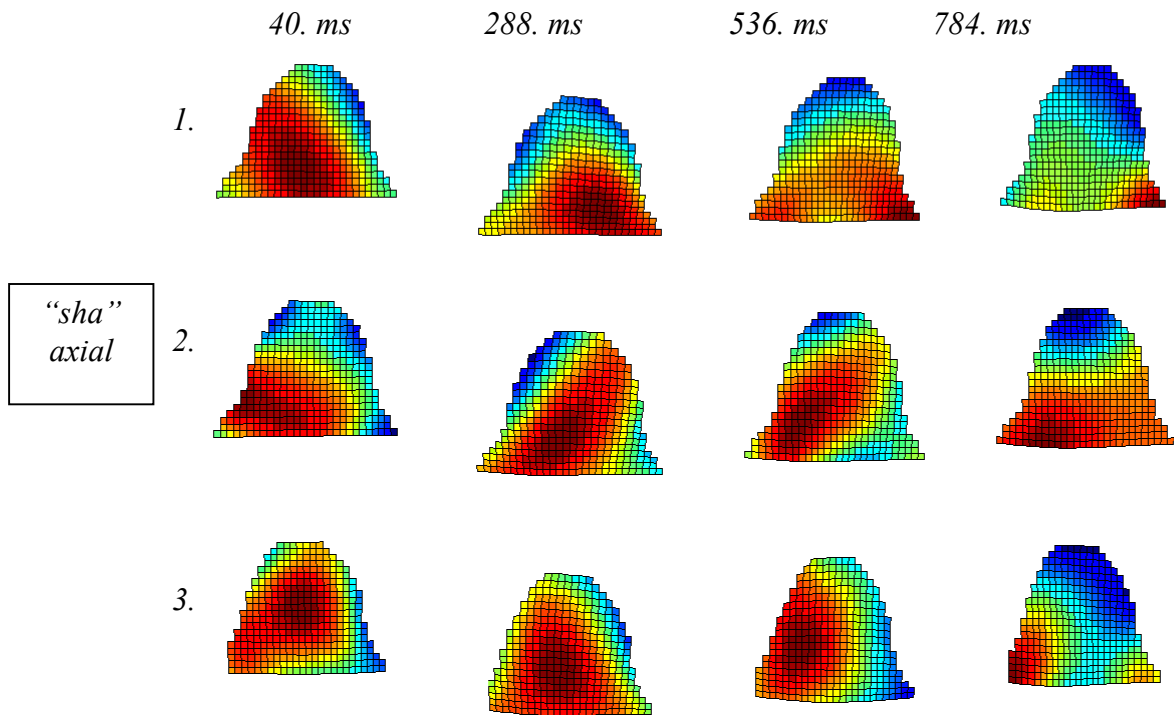

Figure 3: 4D B-spline motion field analysis of axial tongue motion for all three repetitions of syllable "sha". Images of each column belong to a unique time interval. Each row of images belong to the same repetition.
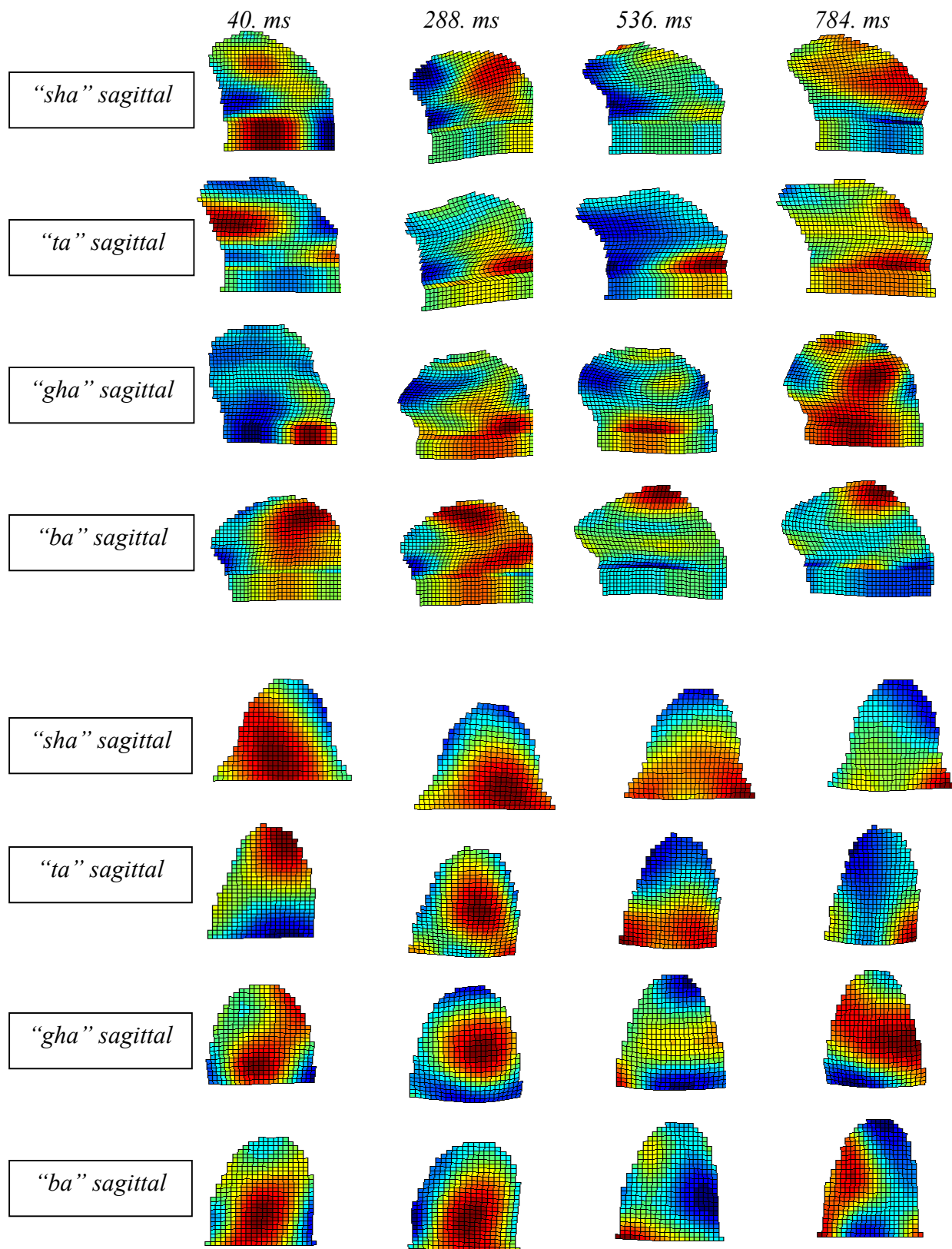
Figure 4: Different compression characteristics of different syllables. We see eight images, four sagittal and four axial, for each syllable. In the figure, the first four rows show the sagittal images and the rest rows below show the axial ones. Images in each column belong to the same time interval as indicated.