

Experimental Assessment of the Tongue Incompressibility Hypothesis During Speech Production

Zisis Iason Skordilis¹, Vikram Ramanarayanan², Louis Goldstein³, Shrikanth S. Narayanan^{1,3}

¹Signal Analysis & Interpretation Lab, University of Southern California, Los Angeles, CA, USA

²Educational Testing Service R&D, 90 New Montgomery St, #1500, San Francisco, CA, USA

³Department of Linguistics, University of Southern California, Los Angeles, CA, USA

skordili@usc.edu, vramanarayanan@ets.org, louisgol@usc.edu, shri@sipi.usc.edu

Abstract

The human tongue is an important organ for speech production. Its deformation and motion control the shape of the vocal tract significantly and thereby the acoustic properties of the speech signal produced. Thus, much effort in the speech research community has been directed towards its biomechanical modeling. A common assumption incorporated into many models of the human tongue is the tissue incompressibility hypothesis: the tongue is considered a muscular hydrostat and therefore its volume should remain constant regardless of its posture. To the best of our knowledge, experimental assessment of the constant volume hypothesis during actual speech production is limited. In this work, the aim is to experimentally assess the incompressibility hypothesis during actual speech production using a dataset of volumetric Magnetic Resonance (MR) images of 17 subjects sustaining contextualized continuants (27 continuants per subject). A seeded region growing based algorithm is used to segment the tongue and calculate its volume. Then the intra-subject variability of the tongue volume along the different tongue postures is examined. Within the accuracy of our tongue volume measurements, our empirical results seem consistent with the incompressibility hypothesis.

Index Terms: speech production, tongue volume, muscular hydrostat, tissue incompressibility, volumetric MRI

1. Introduction

Speech production is a complex process demanding the synergy of a multitude of organs and muscular structures. Airflow from the lungs, possibly modulated by the vocal folds, passes through the vocal tract to produce an acoustic wave. The acoustic properties of the resulting speech signal depend, therefore, on the shape of the vocal tract, which is in turn determined by the position and movement of the various articulators, such as the tongue, the lips, and the velum. In the speech research community, significant effort has been devoted to uncovering and understanding the properties of this mapping between the articulatory space and the acoustic space [1]. Arguably, among the various articulators, the tongue plays a decisive role in controlling the shape of the vocal tract and thereby the acoustic properties of the produced sound [2]. Thus, modeling its motion and dynamics has received particular attention and various models have been proposed [2–8]. An assumption commonly incorporated in tongue models [2–7] is tissue incompressibility: the tongue is considered a muscular hydrostat [6,9] and as such incompressible. However, to the best of our knowledge, there has been limited experimental evaluation of this volume conservation hypothesis during actual speech production. Badin and Serrurier [10] also mention this lack of experimental assessment

and conduct a brief examination of the incompressibility hypothesis by computing tongue volumes from volumetric Magnetic Resonance Imaging (MRI) of 46 French allophone articulations for 1 subject in that particular study [10]. They find the maximum departure from the mean tongue volume to be $\pm 5\%$ [10]. This experimentation, however, was not the main focus of that study and it is also limited to only 1 subject.

In this work, the focus is on experimentally evaluating the constant tongue volume hypothesis during continuant production for multiple speakers and tongue postures. To this end, we utilize a dataset of volumetric MRI of 17 subjects sustaining contextualized continuants (27 per subject) [11]. The MR images were collected with an accelerated MR protocol [12] which requires 8s to scan the 3D vocal tract. This enables, with few exceptions, collection of the data while the subject is actively producing speech and eliminates the need for artificially holding tongue postures. For each speaker and tongue posture, the tongue is segmented using an adapted version of the semi-automatic tongue segmentation algorithm presented by Bone et al. [13], which is based on seeded region growing [14]. The intra-subject variability in tongue volume measurements is statistically examined. Within the accuracy of our measurements, the empirical results are consistent with the incompressibility hypothesis as relatively small standard deviations in tongue volumes are observed. A subset of the results presented here was presented in a preliminary study by Skordilis et al. [15].

2. Dataset

The dataset used for this study consists of volumetric MR images of subjects sustaining contextualized continuants. The data collection process, including the MRI acquisition and reconstruction process, is described in Kim et al. [11]. The GE 3.0 Tesla HDxt scanner system at the Healthcare Consultation Center II, University of Southern California, was used. For radiofrequency (RF) transmission and reception, a body coil and a commercial 8-channel neurovascular receiver coil were utilized, respectively. The spatial resolution and size of the resulting MR image were $1.25\text{mm} \times 1.25\text{mm} \times 1.25\text{mm}$ and 160 (axial) \times 160 (coronal) \times 80 (sagittal), respectively. Using the accelerated MRI protocol proposed by Kim et al. [12] and performing an accelerated acquisition [11], the scan time was reduced to 8s per volume. Data were collected from 17 subjects, all native speakers of American English. Each subject lay in the scanner in supine position, read the stimuli on a mirror-projector setup, and sustained a contextualized continuant for the duration of the scan (8s). A total of 27 contextualized continuants of American English were recorded: 13 vowels (abbot, bat, pot, but, bird, bait, bet, bit, beet, boat, bought, boot, put), 9 fricatives (afa, ava, aha, asha, aga as in beige, asa, aza, atha as in thing,

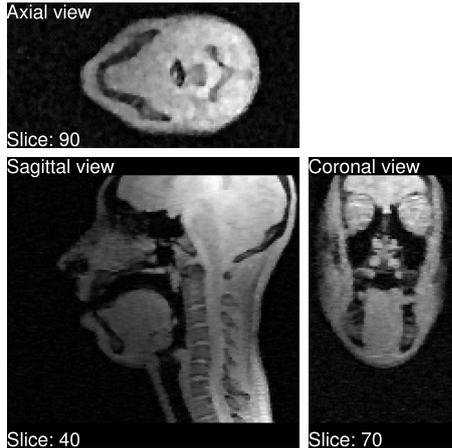


Figure 1: Example of collected volumetric MRI data: an axial, a sagittal, and a coronal slice of the volumetric MR image are shown for the female subject W3 for the stimulus *abbot*. The volumetric image consists of stacked slices.

ath as in this), 3 nasals (*ama*, *ana*, *anga*), and 2 liquids (*ala*, *ara*). Aside from the the speech stimuli, 6 non-speech tongue postures were also collected (e.g. normal breathing, clenched teeth, breath hold, and others), however these are not used in this study. An example of the MRI data collected is shown in Fig. 1 where an axial, a sagittal and a coronal view of a female subject sustaining *abbot* are depicted.

The advantage of using an accelerated MRI protocol with short scan duration was that it enabled the collection of an entire volume to complete while the subject was still able to actively produce speech (without requiring any repeated scans to acquire the same volume). Moreover, there was no need for the subjects to artificially hold a tongue posture (hold a posture without producing sound). This is important, because the goal of this study is to evaluate the constant tongue volume hypothesis during continuant productions that are as ecologically valid as possible. The only exception was the stimulus *aha*, which most subjects were unable to sustain for the entire duration of the scan. In those cases, the subjects were instructed to artificially hold the tongue posture without producing sound after they ran out of air and until scan completion.

The drawback of using the accelerated MRI protocol was lack of tissue contrast: low density regions (airways) have flat low intensity, while high density regions have flat high intensity without any distinction between different types of tissue. Only contrast between air and tissue is present in the reconstructed image (Fig. 1). This severely limits the accuracy of any tissue segmentation attempt. Specifically for the task of segmenting the tongue, it is difficult to differentiate between lingual and surrounding tissue (the jaw, musculature, the epiglottis, etc.).

3. Tongue segmentation algorithm

To segment the tongue, an adapted version of the semi-automatic algorithm presented by Bone et. al. [13] was used. The algorithm consists of the following stages: image preprocessing and denoising; manually guided tongue region bounding; and finally, seeded region growing [14] to segment tongue tissue. We used a modified region growing algorithm, which we first describe. Then the stages of the tongue segmentation algorithm are described.

3.1. Modified seeded region growing

Seeded region growing [14] can be used to segment connected areas of similar pixel intensity in an image. It starts with the

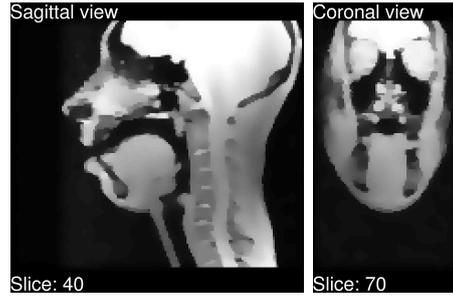


Figure 2: Example of denoised MRI data: a sagittal and a coronal slice for the female subject W3 for the stimulus *abbot*.

manual specification of a small seed inside a region of interest in the image. This is a first estimate S_0 of the region of interest. The region is iteratively updated. At each iteration the neighbors of the current region S_i are examined. The following “distance measure” is computed for each neighbor:

$$\delta(i) = |f(i) - \mathbb{E}_{j \in S_i}[f(j)]|, \quad (1)$$

where i is a neighbor pixel index, $f(\cdot)$ denotes pixel intensity and $\mathbb{E}_{j \in S_i}[f(j)]$ the mean intensity of the current region. The $\delta(i)$ measures are enqueued in a priority queue Q . The region is expanded by adding the enqueued pixel with the smallest $\delta(i)$:

$$S_{i+1} = S_i \cup \{\arg \min_{i \in Q} \delta(i)\} \quad (2)$$

This process continues until the distance $\min_{i \in Q} \delta(i)$ of the next candidate pixel exceeds a predefined threshold T . The resulting region is connected and contains pixels with similar intensity. For use on our MRI data, we modified this region growing algorithm as follows: when segmenting tissue, candidate pixels with higher intensity than the current region mean are always added to the region regardless of the intensity distance $\delta(i)$. The value of $\delta(i)$ is only taken into account if the candidate pixel to be added to the region has a lower intensity than the region mean. This modification is motivated by the fact that tissue in our MRI images has high intensity, therefore if the current region contains tissue, pixels with intensity higher than the region mean are most likely tissue regardless of their exact intensity value. The respective modification was made when segmenting airways (always add to the region neighbors with lower intensity than the region mean). In the following, seeded region growing will refer to this modified version.

3.2. Image preprocessing

Intensity correction was applied to the volumetric image, because the 8-channel neurovascular receiver used for MRI acquisition exhibits a drop in sensitivity in the anterior-posterior direction (Fig. 1) [11]. Each coronal slice was individually normalized by the average tissue intensity in the same slice [11] (tissue detected by thresholding with Otsu’s method [16]).

Then, denoising was applied to reduce noisy spikes and drops in areas of uniform intensity. Anisotropic diffusion [17] was chosen as the denoising method, since it preserves edges. For each intensity corrected volume, two denoised volumes were produced: one by applying anisotropic diffusion to each sagittal slice individually, and one by applying anisotropic diffusion to each coronal slice individually (starting again from the intensity corrected volume). The reason will become apparent in the following. A denoised image example is shown in Fig. 2.

3.3. Manually guided tongue region bounding

Since there is no contrast between different types of tissue in the volumetric image, it is necessary to provide bounds to any

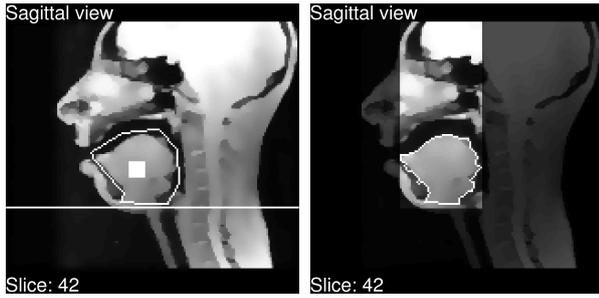


Figure 3: Tongue segmentation on the midsagittal slice. Left: Manually provided polygonal tongue boundary (polygonal line), tongue seed (solid polygon), and maximal axial slice (horizontal line) where tongue tissue is present. Right: Result of tongue segmentation. Boundary of detected tongue region. Regions to be removed are shaded.

generic region growing scheme, which merely uses voxel intensity as a feature. Otherwise, any tissue region will grow into surrounding tissue and finally encompass all connected tissue in the head. To provide bounds for the tongue region, prior knowledge of anatomical landmarks must be used. Since there are no tissue boundaries in the image, these cannot be fully automatically detected and manual input is necessary.

First, a preliminary segmentation of the tongue on the midsagittal slice was performed using seeded region growing. The midsagittal slice index is manually specified. On the midsagittal slice, a rough polygonal contour that encompasses the tongue is manually drawn (Fig. 3). This can be easily done as it need not be precise, but only provide a rough boundary around the tongue to avoid potential leakage of region growing. A seed inside the tongue is also manually specified (Fig. 3). Finally, the maximal axial slice where tongue tissue exists is manually specified (Fig. 3). Then, the midsagittal tongue is segmented by region growing constrained inside the provided rough tongue delineation (Fig. 3). This midsagittal segmentation provides the tip and the back of the tongue. Outside these bounds other tissue is removed (shaded regions in Fig. 3).

Next, the palatal contour is tracked on the midsagittal slice. Tongue tissue can only lie below this contour (since this is the midsagittal palatal contour it is the highest possible), so the region above is removed. The axial slice where the top of the hard palate lies is manually provided (Fig. 4). A seed inside the midsagittal airway is also provided (Fig. 4). Then, using seeded region growing (constrained under the top of the palate), the airway is segmented and the palatal contour extracted (Fig. 4).

Finally, the upper teeth outer contour is estimated. The search region for lingual tissue is bound within it. An axial slice



Figure 4: Palatal contour tracking. Left: Manually specified airway seed (solid polygon) and axial slice at top of hard palate (horizontal line). Right: Boundary of detected airway region. Regions to be removed are shaded.

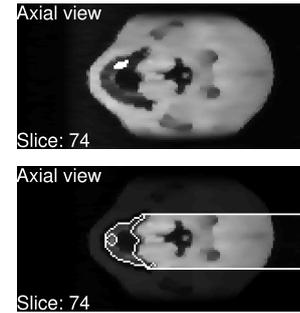


Figure 5: Upper teeth contour detection. Top: Manually specified upper teeth seed. Bottom: Boundary of detected upper teeth region (expanded towards the back of the head). Regions to be removed are shaded.

where the upper teeth are visible is manually selected and a seed within the upper teeth region is manually specified. Then, using seeded region growing the teeth region is segmented and its outer contour is extracted.

For all the above, the sagittally denoised volume was used, in order to provide uniform intensity plateaus for region growing on the midsagittal slice.

3.4. Seeded region growing for final tongue segmentation

After removal of surrounding tissue as discussed, final tongue segmentation is accomplished by seeded region growing applied iteratively over the coronal slices from the detected tip of the tongue and up to the detected back of the tongue. In each coronal slice, the intersection of the midsagittally segmented tongue with the respective coronal slice is used as a seed. For this final region growing step, to provide uniform intensity plateaus, the coronally denoised volume is used. Following tongue segmentation, tongue volume calculation is straightforward (1.25^3mm^3 per voxel of tongue tissue).

4. Results and discussion

Tongue volumes were determined for all speakers and tongue postures. For a particular female subject, example segmented tongue postures are depicted in Fig. 6. In terms of tongue volume, the results are summarized in Table 1, which shows the mean, the standard deviation, and the upper bound for a 95% confidence interval for the standard deviation of the tongue volume for all postures within each subject. For the subjects W6 and M6 (marked with *) the pot stimulus was missing from the data and results are based on the remaining 26 stimuli. For the subject W9 (marked with +), the data was ill-conditioned with intensity variations causing a poor segmentation result, so those results should be viewed with caution. This subject was not included in any statistical analyses mentioned in the following.

The average tongue volumes estimated for each subject are in the order of magnitude reported in other studies. Lauder and Muhl [18] report average tongue volumes of 71.2cm^3 up to 79.3cm^3 . Lowe et al. [19] report tongue volumes ranging from 44.03cm^3 to 99.56cm^3 with a mean and standard deviation of $(71.96 \pm 13.41) \text{cm}^3$. Badin and Serrurier [10] report an average (across postures) of 115cm^3 for their 1 male subject.

Our empirical results indicate that the standard deviation of the tongue volume for all subjects is less than 13cm^3 with 95% confidence. This is less than 11.5% of the tongue volume mean (although the percentage should be viewed with caution since there is overestimation of tongue volume due to the limitations of the data, see Section 5). Furthermore, for all subjects except one (M6), the tongue volume standard deviation is

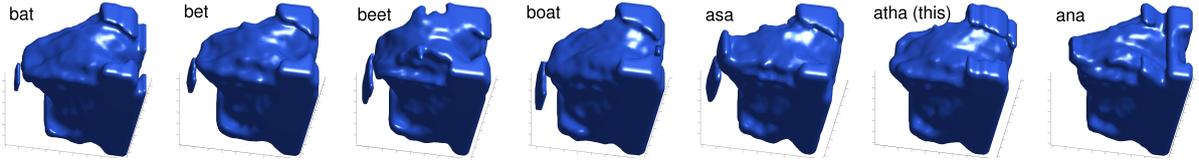


Figure 6: Examples of segmented tongue for some tongue postures for the female subject W3.

less than 10cm^3 (less than 10% of the respective mean) with 95% confidence. These results show that tongue volume varies within a relatively small margin around its mean. So our results are not inconsistent with the incompressibility hypothesis within the accuracy of our measurements. Further, we conducted some limited Friedman tests [20] to check for differences in tongue volume between postures. Unfortunately, conducting a Friedman test for all postures does not yield a meaningful result, because associations between tongue postures and volume are introduced by segmentation artifacts (details in Section 5). Friedman tests were conducted for the following sets of postures: {ama, ana, ala, afa, ava, asa}, {abbot, pot, but}, and {aga (as in beige), aza, asha}, which were selected so that within each set the tongue limits (tip and back) do not vary significantly and thus segmentation artifacts are minimized. The tests indicated no significant difference between postures within each considered set ($p = 0.91, 0.11, 0.27$, respectively). These Friedman tests provide some insight but they cannot be meaningfully extended to more postures due to segmentation artifacts (Section 5) and should be viewed with caution.

The gender variability of tongue volume can also be examined. For each subject, tongue volume is estimated as the mean volume across postures. The mean and standard deviation of tongue volume within gender groups are $(111.44 \pm 11.31)\text{cm}^3$ for male and $(83.36 \pm 11.79)\text{cm}^3$ for female. A Mann-Whitney test [21] showed statistically significant distribution difference between genders at the 0.01 level ($p = 0.002$). Within gender, distribution of tongue volume does not significantly deviate from normal (Shapiro-Wilk test [22] $p > 0.4$). The variances of the two gender groups are not significantly different (Levene test [23] $p > 0.6$). A one-way Analysis of Variance (ANOVA)

Subject	Mean (cm^3)	Std (cm^3)	UB of 95% bCI for Std (cm^3)
W1	92.14	3.88	4.75
W2	71.48	3.49	4.49
W3	73.69	3.77	4.75
W4	95.77	6.26	7.34
W5	65.81	3.40	4.11
W6*	96.92	4.16	5.30
W7	83.89	4.87	6.84
W8	87.20	5.02	5.90
M1	127.59	5.93	7.12
M2	113.88	4.26	5.42
M3	113.11	7.99	9.25
M4	109.26	7.09	8.86
M5	125.27	5.95	6.79
M6*	105.25	10.01	12.10
M7	93.08	4.55	5.49
M8	104.10	3.78	4.43
W9 ⁺	105.15	4.40	5.17

Table 1: Tongue volume estimation results (W=female, M=male, UB=Upper bound, bCI=bootstrap confidence interval). For subjects marked with * the pot stimulus was missing from the data and results are based on the remaining 26 stimuli. For the subject marked with +, the data was ill-conditioned with intensity variations causing a poor segmentation result.

showed that the difference in mean tongue volume between genders is statistically significant at the 0.01 level ($p < 10^{-3}$). This evidence supports the conclusion that on average, tongue volume for males is higher than tongue volume for females.

5. Limitations

Due to the accelerated MRI protocol required to be able to collect volumetric MRI data during actual speech production, no contrast between different types of tissue is available in the reconstructed MR image. This poses a major limitation to any tongue segmentation technique (manual or automatic). It is not possible to distinguish between the various muscles in the tongue, between the tongue and the jaw, the tongue and the gums, etc. Some manual bounding of the tongue region is possible, however the tongue-jaw boundary at the base of the tongue is impossible to detect, as are the boundaries between different muscles in the tongue and surrounding tissue. For this reason, region growing leaks into the jaw and the sides of the tongue producing overestimated tongue volumes. Furthermore, an artificial association between tongue volumes and postures is produced: the further back the tongue is positioned the more overestimation occurs, since towards the back of the head the tongue is surrounded by tissue it cannot be distinguished from. This prohibits any meaningful analysis of variance to detect association between postures and tongue volumes, as the correlation between postures is an artifact of the segmentation. Back vowels for instance consistently produce larger than the mean volumes, but this is an artifact.

6. Conclusions and future work

In this study, an attempt was made to evaluate the constant tongue volume hypothesis during actual speech production. A dataset of volumetric MRI of 17 American English speakers sustaining continuants (27 continuants per subject) was used. An accelerated MRI protocol requiring only 8s to scan the entire volume enabled, with few exceptions, collection of data during active speech production without the need of artificially sustained postures. A seeded region growing based algorithm was used to segment the tongue and calculate its volume. Intra-subject standard deviation of tongue volume was less than 13cm^3 (or 11.5% of the mean) and less than 10cm^3 for all except one subject. Within the limited accuracy of our tongue volume measurements, these empirical results are not inconsistent with the constant tongue volume hypothesis. Gender variability was also examined and statistically significant difference in tongue volume was observed: on average, for female subjects tongue volume is smaller than for male subjects.

In future work, we plan to combine the static in time volumetric MRI data with real-time MRI videos of the midsagittal slice to estimate the motion of the tongue in 3 dimensions. Based on the results of our current study, tongue incompressibility can be incorporated as a constraint into the aforementioned estimation problem.

7. Acknowledgements

This work was supported by the NIH and a USC Viterbi Graduate School PhD fellowship.

8. References

- [1] Y. Zhu, Y.-C. Kim, M. Proctor, S. Narayanan, and K. Nayak, "Dynamic 3-D visualization of vocal tract shaping during speech," *IEEE Transactions on Medical Imaging*, vol. 32, no. 5, pp. 838–848, May 2013.
- [2] V. Sanguineti, R. Laboissire, and Y. Payan, "A control model of human tongue movements in speech," *Biological Cybernetics*, vol. 77, no. 1, pp. 11–22, 1997.
- [3] P. Perrier, Y. Payan, M. Zandipour, and J. Perkell, "Influences of tongue biomechanics on speech movements during the production of velar stop consonants: A modeling study," *The Journal of the Acoustical Society of America*, vol. 114, no. 3, pp. 1582–1599, 2003.
- [4] J. Dang and K. Honda, "Construction and control of a physiological articulatory model," *The Journal of the Acoustical Society of America*, vol. 115, no. 2, pp. 853–870, 2004.
- [5] C. Pelachaud, C. van Overveld, and C. Seah, "Modeling and animating the human tongue during speech production," in *Proc. Computer Animation*, May 1994, pp. 40–49.
- [6] W. S. Levine, C. Essex Torcaso, and M. Stone, "Controlling the shape of a muscular hydrostat: A tongue or tentacle," in *New Directions and Applications in Control Theory*, ser. Lect. Notes Control, W. P. Dayawansa, A. Lindquist, and Y. Zhou, Eds., 2005, vol. 321, pp. 207–222.
- [7] Y. Payan and P. Perrier, "Synthesis of V-V sequences with a 2D biomechanical tongue model controlled by the equilibrium point hypothesis," *Speech Communication*, vol. 22, no. 2–3, pp. 185–205, 1997.
- [8] P. Badin, G. Bailly, L. Revéret, M. Baciú, C. Segebarth, and C. Savariaux, "Three-dimensional linear articulatory modeling of tongue, lips and face, based on MRI and video images," *Journal of Phonetics*, vol. 30, no. 3, pp. 533–553, 2002.
- [9] W. M. Kier and K. K. Smith, "Tongues, tentacles and trunks: the biomechanics of movement in muscular-hydrostats," *Zoological Journal of the Linnean Society*, vol. 83, no. 4, pp. 307–324, 1985.
- [10] P. Badin and A. Serrurier, "Three-dimensional linear modeling of tongue: Articulatory data and models," in *Proc. International Seminar on Speech Production (ISSP)*, Brazil, Dec. 2006, pp. 395–402.
- [11] Y.-C. Kim, J. Kim, M. Proctor, A. Toutios, K. Nayak, S. Lee, and S. Narayanan, "Toward automatic vocal tract area function estimation from accelerated three-dimensional magnetic resonance imaging," in *Proc. ISCA Workshop on Speech Production in Automatic Speech Recognition (SPASR)*, France, 2013.
- [12] Y.-C. Kim, S. S. Narayanan, and K. S. Nayak, "Accelerated three-dimensional upper airway MRI using compressed sensing," *Magnetic Resonance in Medicine*, vol. 61, no. 6, pp. 1434–1440, 2009.
- [13] D. K. Bone, M. I. Proctor, Y. Kim, and S. S. Narayanan, "Semi-automatic modeling of tongue surfaces using volumetric structural MRI," *The Journal of the Acoustical Society of America*, vol. 130, no. 4, pp. 2549–2549, 2011.
- [14] R. Adams and L. Bischof, "Seeded region growing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 6, pp. 641–647, Jun 1994.
- [15] Z. I. Skordilis, V. Ramanarayanan, L. Goldstein, and S. S. Narayanan, "Experimental evaluation of the constant tongue volume hypothesis," *The Journal of the Acoustical Society of America*, vol. 136, no. 4, pp. 2143–2143, 2014.
- [16] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
- [17] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.
- [18] R. Lauder and Z. F. Muhl, "Estimation of tongue volume from magnetic resonance imaging," *The Angle Orthodontist*, vol. 61, no. 3, pp. 175–184, 1991.
- [19] A. A. Lowe, N. Gionhaku, K. Takeuchi, and J. A. Fleetham, "Three-dimensional CT reconstructions of tongue and airway in adult subjects with obstructive sleep apnea," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 90, no. 5, pp. 364–374, 1986.
- [20] M. Friedman, "The use of ranks to avoid the assumption of normality implicit in the analysis of variance," *Journal of the American Statistical Association*, vol. 32, no. 200, pp. 675–701, 1937.
- [21] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *The Annals of Mathematical Statistics*, vol. 18, no. 1, pp. 50–60, 1947.
- [22] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, no. 3-4, pp. 591–611, 1965.
- [23] H. Levene, "Robust tests for equality of variances," in *Contributions to Probability and Statistics*, I. Olkin, Ed. Stanford University Press, 1960, pp. 278–292.