Article published online: 2024-07-18

Estimating Palatal and Pharyngeal Muscle Contraction in Hindi Syllable Pronunciation using Computational Modeling

Madhubari Vathulya[1] Subrato Sarkar[2] Indra Vir Singh[3] Tripta Prajapati[4] Pankaj Sharma[5]

[1] Department of Burns and Plastic Surgery, All India Institute of Medical Sciences, Rishikesh, Uttarakhand, India
[2] Rensselear Polytechnic Institute, Troy, New York, United States
[3] Department of Mechanical and Industrial Engineering, Indian Institute of Technology, Rourkee, Uttarakhand, India
[4] Department of Radiology, All India Institute of Medical Sciences, Rishikesh, Uttarakhand, India
[5] Department of Radiodiagnosis, All India Institute of Medical Sciences, Rishikesh, Uttarakhand, India

Address for correspondence Madhubari Vathulya, MS, MCH, Department of Burns and Plastic Surgery, All India Institute of Medical Sciences, Rishikesh, Uttarakhand 249203, India (e-mail: madhubari@yahoo.co.in).

Indian J Plast Surg 2024;57(Suppl S1):S24–S29.

Abstract

Introduction Speech, one of the main functions affected by cleft palate, involves a complex orchestra of sound deformation by various organs including the larynx, pharynx, epiglottis, palate, tongue, lip, and other structures complementing them. Although the effects of palatoplasty are traditionally evaluated subjectively, objective parameters have seldom been described to compare the pre- and postrepair functions. The study tries to extract the palatal and pharyngeal muscles while uttering various Hindi syllables and tries to find the percentage contraction of these structures as an objective assessment.

Materials and Methods Digital Imaging and Communications in Medicine (DICOM) images while uttering each syllable of the Hindi syllable are obtained by subjecting a healthy volunteer to a dynamic magnetic resonance imaging (MRI). Using freeform geomagic software, 3D models of the structure of the pharynx and palate were created, which were then used to create a finite element model. Different anatomical constraints were applied to the muscles obtained. The finite element model was tested for convergence and a suitably fine mesh was used to obtain the results. The percentage of contraction of the palate and pharynx for uttering different syllables was thus evaluated.

Results The palate and the posterior pharyngeal walls yielded different contractions for different syllables independent of each other. The highest contraction for the palate and posterior pharyngeal wall was for the syllable /k/ and the lowest regarding the palate was for /h/ and /e:/ for the pharyngeal wall.

Conclusion Using computational modeling, quantification of speech in terms of percentage contraction of the palate and pharynx has been attempted for the Hindi

Keywords
- Devnagri script
- computational modeling
- speech
- syllables
- palate
- pharyngeal
- contraction

article published online
July 18, 2024

DOI https://doi.org/10.1055/s-0044-1788591.
ISSN 0970-0358.

Thieme Medical and Scientific Publishers Pvt. Ltd., A-12, 2nd Floor, Sector 2, Noida-201301 UP, India

language. Once validated with a larger population, the database may be used to quantify speech deformities due to structural pathologies in terms of palatal and pharyngeal contractions and help us assess the effectiveness of corrective surgeries for them.

Introduction

The mechanism of speech is complex and determined by dynamic deformations of the structures involved in speech including the larynx, epiglottis, pharynx, tongue, palate, and lip.[1]

The speech apparatus is affected at the level of the palate and posterior pharyngeal wall in cases of the cleft palate, which is one of the common congenital deformities.[2] The idea of cleft palate repair is to address the same by reorienting the abnormally oriented muscle fibers, especially of the levator and palatopharyngeus muscle. While exhaustive research has been done to assess the result of such repairs subjectively by evaluating the quality of speech, very little work has been done to assess speech objectively. In the recent publication by Rizzo et al,[3] subjective and objective analyses of repaired palates were compared in terms of speech improvement, which is the effect of the repair. Since even a well-repaired palate might sometimes not be able to produce the satisfactory speech desired due to the need for speech therapy, it is imperative to understand if the problem in speech lies in the structural anatomical reconstruction or in the dearth of adequate speech therapy. To facilitate the same, it would be important to understand the normal contraction rates or deformation of the various soft-tissue structures involved in speech while uttering these syllables in normal individuals so that the repaired palates can then be compared from the preoperative status.

The finite element analysis is an effective tool for understanding contraction rates in the palatal muscles to assess speech. A full-fidelity model of the speech apparatus (accounting for deformation, acoustics, and surrounding anatomy) is a complex and challenging task that might not be feasible with the currently available tools. However, a simpler model with minimal assumptions is a viable approach to this challenging problem. Hence, in this work, the contraction of palatal muscles is analyzed using a finite element model validated using the actual contractions observed. This approach is similar to the works of Takenaka et al[4] and Arnela et al[5] in which computed tomography (CT) scans and magnetic resonance imaging (MRI) were used to generate the finite element models, which were later simulated to obtain responses during speech. Besides biomechanical response, the finite element analysis is used to analyze other aspects of speech such as acoustic resonance,[6] vibration and sound propagation,[7] and articulation.[8]

Therefore, the main objectives of this study are to acquire Digital Imaging and Communications in Medicine (DICOM) data for the various syllables of the Hindi language using dynamic MRI and use computational modeling to analyze the maximum contraction rates of pharyngeal and palatal muscles while uttering them. Once these data are compiled, the authors wish to further extrapolate this study to establish a database of palatal and pharyngeal muscle contraction rates while uttering these syllables for future references and research by performing a large-volume study.

Materials and Methods

After obtaining a written and explained consent, a healthy volunteer well versed in the Hindi language was asked to phonate each Hindi syllable at a habitual pitch and loudness. After a baseline MRI of his head and neck at rest, adopting the protocol used by Vampola et al[9] with mild modifications, he was asked to phonate each syllable until the soft palate and pharyngeal movements completed the contraction–relaxation cycle for an average of 10 times for each syllable using a dynamic MRI of 3-T GE 3T Discovery 50 W. Sagittal fiesta sequences were used with repetition time (TR) of 4.2, echo time (TE), minimum frequency, of 224, number of excitations (NEX) of 1, bandwidth of 125, and slice thickness of 6 mm. The 3D reconstructed images from the DICOM images for each model were then extracted at the maximum level of distortion using geomagic freeform software. The central palate and posterior pharynx at the level of the superior constrictor were considered for the study and compared with their resting positions. Using the freeform geomagic software, various 3D images of muscles contributing to the palate and pharyngeal wall of the speech apparatus were extracted.

The 3D images were used to create a finite element model of the palate and the posterior pharyngeal wall, as shown in ►**Figs. 1** and **2**. It is seen that the palate comprises several muscles, namely the palatopharyngeus, levator palati, and tensor palati, and structures like the uvula, while the posterior pharyngeal wall principally consists of the superior constrictor muscle contributing to speech. The muscle groups (palate and posterior pharyngeal wall) contract in the direction of their respective muscle fibers to simulate speech. Different anatomical constraints are applied to the muscles to capture the muscle movements during speech accurately using the finite element analysis software (COMSOL). The speech muscles are assumed to be passive, and an isotropic hyperelastic model is adopted with parameters taken from studies of previous literature. The contraction loads are applied such that the maximum contraction is obtained in the direction of muscle fibers. The contraction is defined as the percentage change in the largest dimension in the direction of muscle fibers. This parameter, which is used to measure the muscle contraction, is independent of the muscle and its features. Hence, this

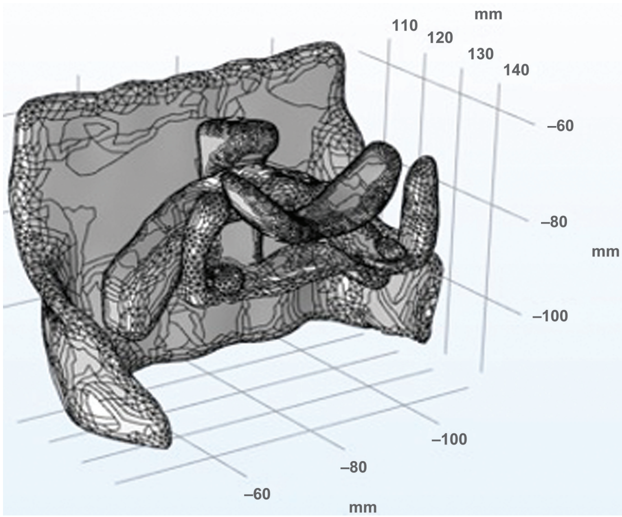

Fig. 1 The entire model with overall dimensions.

measure can be applied to models taken from different subjects and can yield reliable measurements. The vectors of muscle action are designated based on these studies and also demonstrated by finite element analysis.[10] The use of dynamic MRI to assess the different sites of the speech apparatus has been attempted in the past too.[11,12] These dynamic MRI data are used to validate the finite element model. A flowchart outlining the iterative simulation and validation process is shown in ►**Fig. 3**. The finite element model is tested for convergence and a suitably fine mesh is used to obtain the results. The reference for the contribution of each palatal and pharyngeal muscle taken into consideration for computation was from the study by Anderson et al.[13]

Since the syllables are the simplest and basic building blocks of speech, this effort to record the functional DICOM images while uttering them and later extracting models helps us to analyze the muscle effort for each such syllable. Since the number of words that are present in the language far exceeds the researchable possibility at the moment, this endeavor can at least help us to understand the basic effort for uttering the units making each part of speech, and later may be extrapolated for words using artificial intelligence (AI) or future technologies.

Studies using finite element analysis for speech have been performed in the past, but their fidelity and endpoints have differed from our present study. For instance, the initial studies considered 2D planar models with a simplistic representation of soft palate muscles.[14] Later, slightly more complex models using simple 3D representations were used.[15] Although these were 3D models, they did not include all the palatal muscles involved in speech. Recently, models with higher fidelity have been used that were created from scanned images.[16,17]

Results

The results obtained from the finite element simulations are shown as a bar plot in ►**Fig. 4**. The bar plot indicates contractions in the soft palate and the superior constrictor needed for each syllable. The palate and the posterior pharyngeal walls yielded different contractions for different syllables independent of each other. The highest contraction for the palate and posterior pharyngeal wall was for the syllable /k/ and the lowest with regards to the palate was for /h/and /e:/ for the pharyngeal wall (►**Fig. 5**).

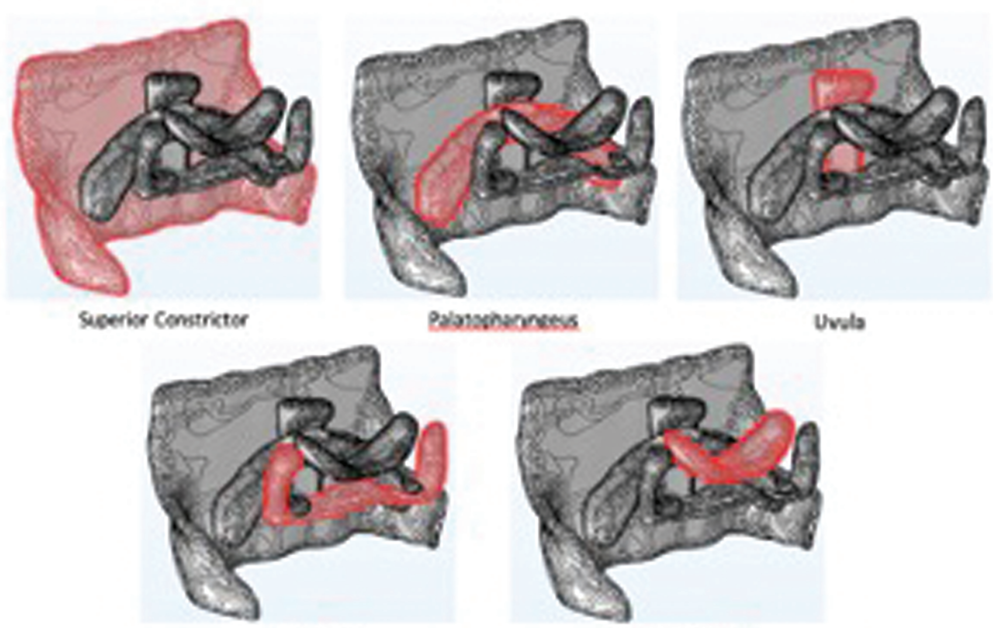

Fig. 2 Schematic showing important parts of the model.

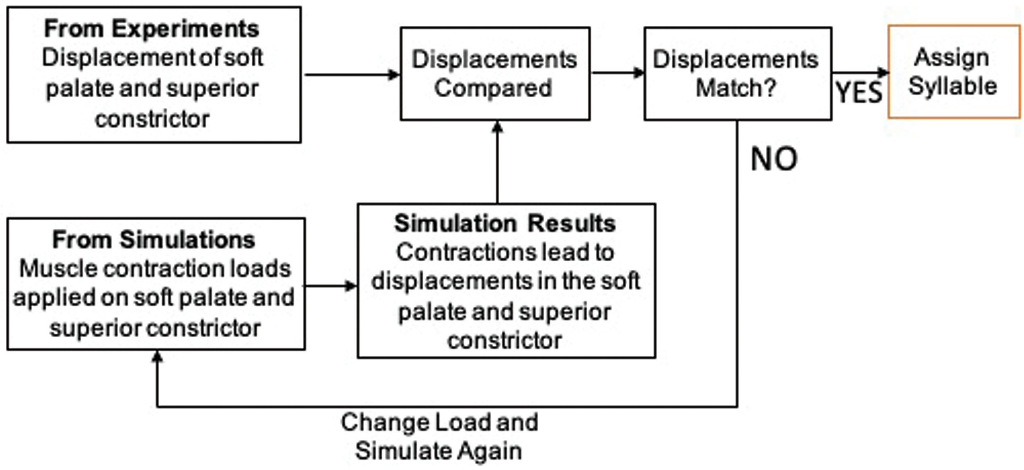

Fig. 3 Flowchart showing the steps performed to obtain results.

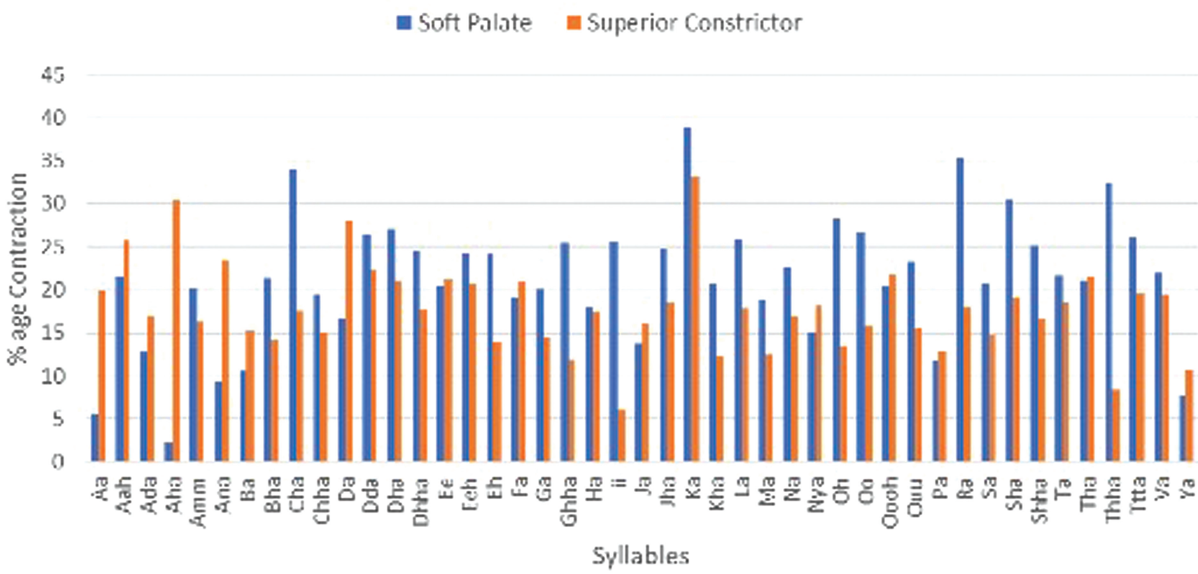

Fig. 4 Individual contractions for each of the 45 Hindi syllables. The *bars* represent the contractions for the soft palate and the superior constrictor for each syllable. For example, for uttering Aa, the soft palate contracted ~5.1% and the superior constrictor contracted 20%.

Discussion

A literature search for objective assessment of speech revealed decreased importance in the past. Speech has so far been evaluated in a subjective manner comparing its quality with that of a normal healthy person. For speech to be legible and comprehensible, the larynx has to be rightly contorted by the various structures so that sound travels through the pharynx, soft palate, lips, and tongue. Various structures deform themselves and get into complex alignments with the other structures, which are responsible for the uniqueness of the sound produced for each syllable.

To understand the deformation of various soft-tissue structures involved in speech production, we need to first capture or image them. Various investigations and software like ultrasonography,[18] X-ray microbeam technique,[19,20] electromagnetic midsagittal articulometer,[21] static,[22] dynamic electropalatography,[23,24] glossometry,[25] and functional MRI[26] have been attempted in the past.

Similar models in the past have tried to simulate velopharyngeal opening and closure, but models using real-time data to mimic the utterance of every syllable have been attempted for the first time as far as the authors are concerned in languages including Hindi.[14] Certain studies[27] have tried to evaluate the levator palati shortening along with their angle of inclination while uttering short words using MRI and have given their observations.

Srodon et al[28] studied velopharyngeal closure using finite element analysis but did not add further information related to deformations during syllable production in any of the languages.

Ha et al[29] studied the anatomical and physiological variations of the levator palati muscle of the palate in their size, muscle fiber inclination, etc., during vocalization. They found that muscle activity decreased progressively from rest to

	Palate	Wall
अ	0.05	0.199
आ	0.21	0.25
इ	0.20	0.21
ई	0.24	0.20
उ	0.26	0.15
ऊ	0.20	0.21
ए	0.24	0.13
ऐ	0.25	0.06
ओ	0.28	0.13
औ	0.23	0.15
अं	0.20	0.16
अः	0.021	0.30

	क	ख	ग	घ	ङ
Palate	0.38	0.205	0.20	0.25	0.16
Wall	0.33	0.12	0.14	0.11	0.28
	च	छ	ज	झ	ञ
Palate	0.33	0.19	0.13	0.24	0.15
Wall	0.17	0.15	0.16	0.18	0.18
	ट	ठ	ड	ढ	ण
Palate	0.21	0.26	0.16	0.26	0.09
Wall	0.18	0.19	0.28	0.22	0.23
	त	थ	द	ध	न
Palate	0.21	0.32	0.27	0.24	0.22
Wall	0.21	0.08	0.21	0.17	0.16
	प	फ	ब	भ	म
Palate	0.11	0.19	0.10	0.21	0.18
Wall	0.12	0.20	0.15	0.14	0.12
	य	र	ल	व	श
Palate	0.07	0.35	0.25	0.21	0.30
Wall	0.1	0.18	0.17	0.19	0.18
	ष	स	ह		
Palate	0.25	0.207	0.17		
Wall	0.18	0.14	0.17		

- ▶ Lowest contraction of palate for Aha and lowest contraction of pharyngeal wall for eh
- ▶ Highest contraction of palate and wall for ka

Fig. 5 Percentage contraction of the palate and posterior pharyngeal wall for each syllable. The lowest contraction of the palate was for /ɦ/ and the lowest contraction of the pharyngeal wall was for /eː/. The highest contraction of the palate and wall was for /k/.

pronouncing nasal consonants, low vowels, high vowels, fricative, and consonants in both repaired and normal palates.

In our study, however, we could not find such observations against the sonority scale. Since the sonority scale is influenced by the airflow through the voice box, diaphragm muscles and the lungs rather than the contractions of palatal and pharyngeal structures the scale is not correlating with the percentage of contractions of these structures studied.

Harper et al[30] have studied the rhotic pharyngeal gestures in English-speaking and non-English-speaking individuals and have shown that it is difficult to acquire the production of pharyngeal gestures in the English rhotic approximant.

Although the results of our study are based on observations made on DICOM data obtained from an adult, not much change in the parameter is expected as compared with children and adolescents, as already described previously by other authors,[31] despite children having a smaller structure than adults. In addition, the percentage of contraction used in the study to quantify muscle contraction is independent of the muscle and is solely based on the muscle's initial and final states. Hence, the results can be easily applied to other models created from other subjects. This measure also ensures reliability due to its independence from the muscles and their particular features. Loudness, pitch, and timber-dependent variables do not affect the parameters being measured since the maximum difference in the lengths of the muscle at rest and maximum contraction while uttering the syllable is taken into consideration here.

Our next stage is directed toward more extensive research with a bigger sample size to increase the significance of the study.

Since the sonority scale depends on the airflow through the voice box, diaphragm and lungs and seldom influenced by the contractions of the palatal or pharyngeal structures, the scale seems to be not correlating with contraction rates of these structures as shown by the results. In the future, these results may enable one to quantify the speech abnormality in structural pathologies like cleft palate in terms of palatal and pharyngeal muscle contraction and help in assessing the improvement after corrective surgeries. However, this study's limitation is that the model is capable but limited only by the available data. The model will become more robust and versatile if more data are available.

Conclusion

Quantification of speech is an important aspect that is seldom researched. The authors have tried to effectively reconstruct speech models for the entire spectrum of syllables of the Hindi language. Future research should focus on a larger sample size to validate and then evaluate the possibility of using the dataset to evaluate speech after procedures like palatoplasty.

Funding
We would like to acknowledge UCOST (Uttarakhand State Council for Science and Technology, Vigyan Sadan Block) for having provided us a grant to conduct this study (UCS&T/R&D-06/19–20/18033/1).

Conflict of Interest
None declared.

Acknowledgment
We would like to acknowledge the technical assistance provided by Mr. Tarun Naugin, radiographic technician, grade 1 in the radiodiagnosis department for assisting us in this work.

References

1 Zhang Z. Mechanics of human voice production and control. J Acoust Soc Am 2016;140(04):2614
2 Kummer AW. Speech and resonance disorders related to cleft palate and velopharyngeal dysfunction: a guide to evaluation and treatment. Perspect School Based Issues 2014;15(02):57–74
3 Rizzo MI, Fallico N, Beneduce N, et al. Objective and subjective evaluation of velopharyngeal dysfunction (VPD) following surgical repair of the cleft palate using the furlow palatoplasty: a new tool. J Plast Reconstr Aesthet Surg 2022;75(09):3448–3456
4 Takenaka S, Yoshinaga T, Nozaki K, Ii S, Wada S. A mechanical deformation model of the tongue during speech production considering personalized anatomical structure. J Acoust Soc Am 2018;144(3, supplement):1907–1907
5 Arnela M, Dabbaghchian S, Guasch O, Engwall O. MRI-based vocal tract representations for the three-dimensional finite element synthesis of diphthongs. IEEE/ACM Trans Audio Speech Lang Process 2019;27(12):2173–2182
6 Vampola T, Horáček J, Švec JG. FE modeling of human vocal tract acoustics. Part I: production of Czech vowels. Acta Acust United Acust 2008;94(03):433–447
7 Alipour F, Berry DA, Titze IR. A finite-element model of vocal-fold vibration. J Acoust Soc Am 2000;108(06):3003–3012
8 Gérard JM, Ohayon J, Luboz V, Perrier P, Payan Y. Non-linear elastic properties of the lingual and facial tissues assessed by indentation technique. Application to the biomechanics of speech production. Med Eng Phys 2005;27(10):884–892
9 Vampola T, Laukkanen AM, Horácek J, Svec JG. Vocal tract changes caused by phonation into a tube: a case study using computer tomography and finite-element modeling. J Acoust Soc Am 2011; 129(01):310–315
10 Dao TT, Tho MHB. A systematic review of continuum modeling of skeletal muscles: current trends, limitations, and recommendations. Appl Bionics Biomech 2018;2018:7631818
11 Lim Y, Zhu Y, Lingala SG, Byrd D, Narayanan S, Nayak KS. 3D dynamic MRI of the vocal tract during natural speech. Magn Reson Med 2019;81(03):1511–1520
12 Erattakulangara S, Kelat K, Meyer D, Priya S, Lingala SG. Automatic multiple articulator segmentation in dynamic speech MRI using a protocol adaptive stacked transfer learning U-NET model. Bioengineering (Basel) 2023;10(05):623
13 Anderson P, Fels S, Stavness I, Pearson WG Jr, Gick B. Intravelar and extravelar portions of soft palate muscles in velic constrictions: a three-dimensional modeling study. J Speech Lang Hear Res 2019; 62(04):802–814
14 Berry DA, Moon JB, Kuehn DP. A finite element model of the soft palate. Cleft Palate Craniofac J 1999;36(03):217–223
15 Inouye JM, Pelland CM, Lin KY, Borowitz KC, Blemker SS. A computational model of velopharyngeal closure for simulating cleft palate repair. J Craniofac Surg 2015;26(03):658–662
16 Anderson P, Harandi NM, Moisik S, Stavness I, Fels SS. A comprehensive 3D biomechanically-driven vocal tract model including inverse dynamics for speech research. Paper presented at: 16th Annual Conference of the International Speech Communication Association, INTERSPEECH 2015; September 6–10, 2015; Dresden, Germany
17 Ma J, Zhang M, Yang D, et al. Three-dimensional finite element analysis on stress distribution after different palatoplasty and levator veli palatini muscle reconstruction. Clin Oral Investig 2024;28(04):221
18 Keller E, Ostry DJ. Computerized measurement of tongue dorsum movements with pulsed-echo ultrasound. J Acoust Soc Am 1983; 73(04):1309–1315
19 Stone M. A three-dimensional model of tongue movement based on ultrasound and X-ray microbeam data. J Acoust Soc Am 1990; 87(05):2207–2217
20 Kiritani S, Ito K, Fujimura O. Tongue-pellet tracking by a computer-controlled X-ray microbeam system. J Acoust Soc Am 1975;57 (6, Pt 2):1516–1520
21 Perkell JS, Cohen MH, Svirsky MA, Matthies ML, Garabieta I, Jackson MT. Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements. J Acoust Soc Am 1992;92(06):3078–3096
22 Ladefoged P. Use of palatography. J Speech Hear Disord 1957;22 (05):764–774
23 Fletcher SG, McCutcheon MJ, Wolf MB. Dynamic palatometry. J Speech Hear Res 1975;18(04):812–819
24 Hardcastle WJ, Jones W, Knight C, Trudgeon A, Calder G. New developments in electropalatography: a state of the art report. Clin Linguist Phon 1989;3:1–38
25 Fletcher SG, McCutcheon MJ, Smith SC, Smith WH. Glossometric measurements in vowel production and modification. Clin Linguist Phon 1989;3:359–375
26 Shinagawa H, Ono T, Honda E, et al. Dynamic analysis of articulatory movement using magnetic resonance imaging movies: methods and implications in cleft lip and palate. Cleft Palate Craniofac J 2005;42(03):225–230
27 Perry JL, Sutton BP, Kuehn DP, Gamage JK. Using MRI for assessing velopharyngeal structures and function. Cleft Palate Craniofac J 2014;51(04):476–485
28 Srodon PD, Miquel ME, Birch MJ. Finite element analysis animated simulation of velopharyngeal closure. Cleft Palate Craniofac J 2012;49(01):44–50
29 Ha S, Kuehn DP, Cohen M, Alperin N. Magnetic resonance imaging of the levator veli palatini muscle in speakers with repaired cleft palate. Cleft Palate Craniofac J 2007;44(05):494–505
30 Harper S, Goldstein L, Narayanan S. L2 acquisition and production of the English rhotic pharyngeal gesture. Paper presented at: INTERSPEECH 2016; September 8–12, 2016; San Francisco, CA
31 Tian W, Yin H, Redett RJ, et al. Magnetic resonance imaging assessment of the velopharyngeal mechanism at rest and during speech in Chinese adults and children. J Speech Lang Hear Res 2010;53(06):1595–1615

