

# Vocal tract physiology and its MRI evaluation

## Fisiología del tracto vocal y su evaluación por resonancia magnética

Bruno Murmura<sup>1</sup>  , Filippo Barbiera<sup>1</sup>  , Francesco Mecorio<sup>2</sup>  , Giovanni Bortoluzzi<sup>3</sup>  , Ilaria Orefice<sup>4</sup>  , Elena Vetrano<sup>1</sup>  , Alfonso Gianluca Gucciardo<sup>5</sup>  

<sup>1</sup> ASP Agrigento - Dipartimento di Scienze Radiologiche; U.O. Radiologia - Ospedale di Sciacca; Sciacca; Italy.

<sup>2</sup> Associazione E.U.R.E.C.A.; Marta (VT); Italy.

<sup>3</sup> Sherden Overtone Singing School; Treviso (TV); Italy.

<sup>4</sup> Singing school; Sherden Overtone Singing School; Mogoro (Or); Italy.

<sup>5</sup> Chair, Performing Arts Medicine Dept.; CEIMARS – Italian interdisciplinary Center for Performing Arts Medicine; Agrigento; Italy.

### Correspondence

Bruno Murmura.  
 email: [brunomurmura@hotmail.com](mailto:brunomurmura@hotmail.com)

### How to cite

Murmura, Bruno; Barbiera, Filippo; Mecorio, Francesco; Bortoluzzi, Giovanni; Orefice, Ilaria; Vetrano, Elena; Gucciardo, Alfonso Gianluca. (2021). Vocal tract physiology and its MRI evaluation. *Revista de Investigación e Innovación en Ciencias de la Salud*. 3(2), 47-56. <https://doi.org/10.46634/riics.84>

**Received:** 28/09/2021

**Revised:** 04/11/2021

**Accepted:** 06/11/2021

### Invited editor

Carlos Manzano Aquihuatl, MD, MSc. 

### Editor en jefe

Jorge Mauricio Cuartas Arias, Ph.D. 

### Coeditor

Fraidy-Alonso Alzate-Pamplona, MSc. 

**Copyright** © 2021. Fundación Universitaria María Cano. The *Revista de Investigación e Innovación en Ciencias de la Salud* provides open access to all its content under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International \(CC BY-NC-ND 4.0\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

### Conflicts of Interest

The authors have declared that no competing interests exist.

## Abstract

**Introduction.** The rapid technological evolution in Magnetic Resonance Imaging (MRI) has recently offered a great opportunity for the analysis of voice production.

**Objectives.** This article is aimed to describe main physiological principles at the base of voice production (in particular of vocal tract), and an overview about literature on MRI of the vocal tract. This is presented in order to analyze both present results and future perspectives.

**Method.** A narrative review was performed by searching the MeSH terms “vocal tract” and “MRI” in Pub Med database. Then, the obtained studies were subsequently selected by relevancy.

**Results.** Main fields described in literature concern technical feasibility and optimization of MRI sequences, modifications of vocal tract in vowel or articulatory phonetics, modifications of vocal tract in singing, 3D reproduction of vocal tract and segmentation, and describing vocal tract in pathological conditions.

**Conclusions.** MRI is potentially the best method to study the vocal tract physiology during voice production. Most recent studies have achieved good results in representation of changes in the vocal tract during emission of vowels and singing. Further developments in MR technique are necessary to allow an equally detailed study of faster movements that participate in the articulation of speaking, which will allow fascinating perspectives in clinical use.

## Key words

Real time MRI; dynamic MRI; sequence; vocal tract; speech research; vowels; singing; overtones; phonetics; voice physiology; speech science.

#### Data Availability Statement

All relevant data is in the article and in the appendices. For more detailed information, write to the Corresponding Author.

#### Funding

None. This research did not receive any specific grants from funding agencies in the public, commercial, or non-profit sectors.

#### Disclaimer

The content of this article is the sole responsibility of the authors and does not represent an official opinion of their institutions or the *Revista de Investigación e Innovación en Ciencias de la Salud*.

#### Author Contributions

**Bruno Murmura:** conceptualization, data curation, formal analysis, investigation, methodology, resources, validation, visualization, writing – original draft, writing – review and editing.

**Filippo Barbiera:** conceptualization, investigation, project administration, resources, supervision, validation, visualization, writing – review and editing.

**Francesco Mecorio:** resources, validation, visualization, writing – review and editing.

**Giovanni Bortoluzzi:** investigation, resources, visualization, writing – review and editing.

**Ilaria Orefice:** investigation, resources, visualization, writing – review and editing.

**Elena Vetrano:** visualization, writing – review and editing.

**Alfonso Gianluca Gucciardo:** investigation, project administration, resources, supervision, validation, visualization, writing – review and editing.

## Resumen

**Introducción.** La rápida evolución tecnológica en la resonancia magnética (MRI) ha ofrecido recientemente una gran oportunidad para el análisis de la producción de voz.

**Objetivos.** Este artículo tiene como objetivo describir los principales principios fisiológicos en la base de la producción de la voz (en particular, del tracto vocal) y una descripción general de la literatura sobre resonancia magnética del tracto vocal. Esto se presenta con el fin de analizar tanto los resultados actuales como las perspectivas futuras.

**Método.** Se realizó una revisión narrativa mediante la búsqueda de los términos MeSH “tracto vocal” y “MRI” en la base de datos PubMed. Los estudios obtenidos se seleccionaron posteriormente por relevancia.

**Resultados.** Los campos principales descritos en la literatura se refieren a la viabilidad técnica y optimización de secuencias de resonancia magnética, modificaciones del tracto vocal en una vocal o fonética articulatoria, modificaciones del tracto vocal en el canto, reproducción 3D del tracto vocal y segmentación y descripción del tracto vocal en condiciones patológicas.

**Conclusiones.** La resonancia magnética es potencialmente el mejor método para estudiar la fisiología del tracto vocal en el momento de la producción de la voz. Los estudios más recientes han obtenido buenos resultados en la representación de cambios en el tracto vocal durante la emisión de vocales y el canto. Se necesitan más desarrollos en la técnica de RM para permitir un estudio igualmente detallado de los movimientos más rápidos que participan en la articulación del habla, lo que permitirá perspectivas fascinantes en el uso clínico.

## Palabras clave

Resonancia magnética en tiempo real; resonancia magnética dinámica; secuencia; tracto vocal; investigación del habla; vocales; canto; armónicos; fonética; fisiología de la voz; ciencias del habla.

## Introduction

MRI is a powerful imaging technique used in radiology to generate anatomical and functional image of the human body.

MRI scans offer numerous advantages in observing anatomy such as its absence of ionizing radiation, non-invasive methods, and excellent contrast in the representation of soft tissue.

Historically, an important limitation was, and still is in certain instances, the time-consuming nature of examination. Often, a fair amount of time is needed to obtain areas of various regions of the body.

Continuous technological evolution of MRI, which consist in magnetic fields and radiofrequencies (RF) used to obtain anatomical images based on some parameters —e.g. T1 and T2, which are different and intrinsic to each tissue—, led to birth of dynamic and real time MRI by use of fast MRI sequences [1].

This has allowed representation of moving organs and new applications in studies related to intestinal and cardiac imaging, and has opened interesting and promising perspectives in other fields, like speech research.

MRI is increasingly used in studies of speech as it enables noninvasive visualization of the vocal tract and articulators, thus providing information about their shape, size, location, motion and position.

However, the use of MRI is limited by some intrinsic characteristics caused by its high technological complexity, such as the elevated cost of use and possibility of causing claustrophobia in patients. Additionally, the use of magnetic fields and gradients prevents it from being safe to use on patients with implants or ferromagnetic material/devices present in their bodies.

The production of voice, sound, and spoken language is the result of the dynamic and fast-moving interaction of organs and anatomical structures. This function needs the interaction and the coordination of the respiratory muscles that control the lungs and the breath flow, which initiates vocalization and provides the energy to produce sound; the vocal folds (in voiced sounds), which convert the airflow in an audible vibration that produces the pitch and the tone; and upper airways, including the mouth and nasal cavities, which act as a resonating chamber or more precisely as an acoustic filter.

Particular attention with MRI is paid to the vocal tract, and the rapid movement and articulation of its anatomical structures.

### The Acoustic Theory of Language production

In speech production, we identify a sound source, which is either a reaction to the vibrating vocal folds (in voiced sounds), or a turbulent flow of air due to some constriction of some parts of the vocal tract (in voiceless sounds), or a combination of the two (Image 1). The periodicity in the acoustic waveform is a consequence of the recurring vibration of the vocal folds. This periodicity characterizes the voiced sounds and determines sinusoidal sound waves. The production of most voiceless sounds is the result of turbulent airflow localized in some part of the vocal tract [2].

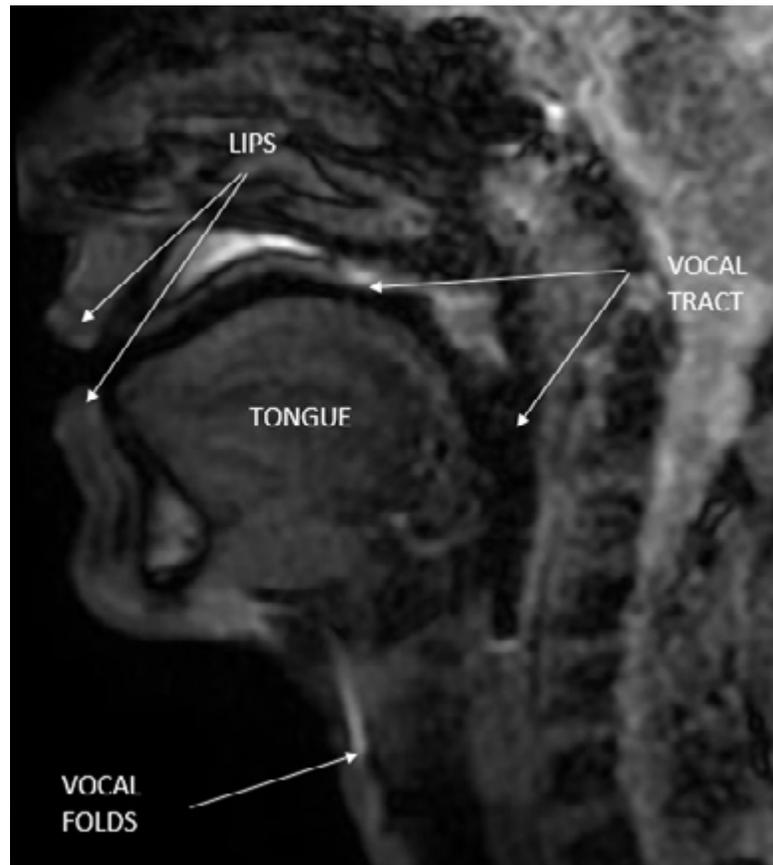
The shape of the vocal tract, which is modeled as an acoustic filter (filter applied to the sound produced by the source through the property of resonance), can be analyzed to a large extent independently of the source itself.

In speech production, formants are the spectral peaks of the sound spectrum that result from acoustic resonance in the vocal tract. The idea that the source and the filter make an (almost) independent contribution to the spectrum of the combined sound output in speech production [3-5] is a critical part of the acoustic theory of language production.

If the characteristics of the source change but the filter stays the same, the two spectra of the sound output remain very similar.

It is possible to change the pitch by producing a sound of the same phonetic quality and varying the fundamental frequency, as vocal folds vibrate with the same shape of the vocal tract.

The position of the noise source, which participates in the production of consonants, in contrast to that of the glottal source, varies with the point of articulation. Points of articulation can be found anywhere from the labiodental region to the glottal region.



**Image 1.** Magnetic Resonance Image (MRI) done by the authors with annotated anatomy of the vocal tract.

For instance, in fricative consonants, the noise source is represented by a turbulent air flow, which is the result of a jet of air being conveyed at high speed through a narrow passage by the teeth. This causes fricative production of [s] and [z] (alveolar fricatives) and [ʃ] and [ʒ] (postalveolar fricatives), or palatal and velar fricative by pressure on the palate [6].

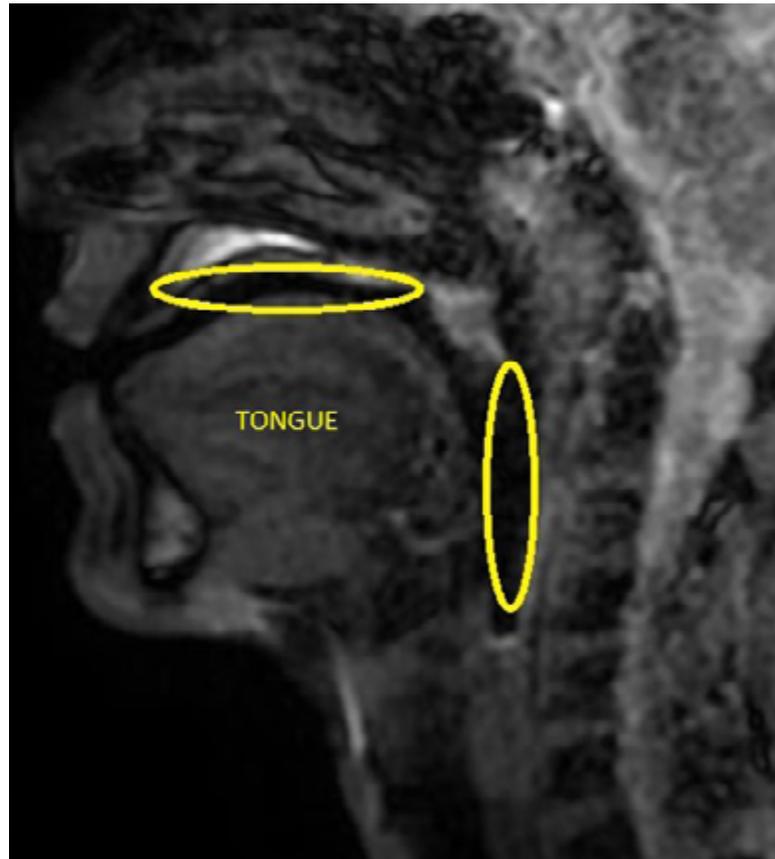
In the production of a vowel, the vocal tract can be considered as a tube from vocal folds to the lip with a constantly varying cross-sectional area between.

The right-angled curve at the junction of the pharynx and oral cavity is not important in regards to spectrum estimation, as the air flow propagates perpendicular to the vocal tract itself.

Resonance in the straight tube occurs when standing waves are produced from air pressure waves that cancel and reinforce each other at various points along the tube itself (Image 2).

One of the most influential models is the three-parameter model of speech production, in which the vocal tract is represented by four interconnected cylinders [7].

The first tube to consider models the constriction position of the vowel. Radiological analysis of vocal production has shown that it is accompanied by narrowing or constriction in the vocal tract, in a similar manner to how the vocal tract is restricted at a point of articulation in the production of consonants.



**Image 2.** Magnetic Resonance Image (MRI) done by the authors with annotated major resonance cavities of the vocal tract: oral cavity (over the tongue) and pharyngeal cavity (posterior to the tongue).

After identifying the constriction, a cavity will appear behind it, extending from the glottis to the constriction, and a cavity will be in front of it, extending from the front of the constriction to the lips, so the vocal tract can be divided into three cylinders representing the posterior cavity, constriction, and anterior cavity. A fourth tube represents the lip configuration, which is an important articulatory parameter in vocal production [2,4].

With vowels [i], the constricted tube is closer to the anterior part of the vocal tract; with vowels [u], the constriction is close to the soft palate; and with open vowels [a], the constriction is in the pharynx.

Additionally, the vowels differ in diameter, in the area of constriction, and in the extent of the opening of the lips.

In the case of nasal consonants, two acoustic pathways are formed from the glottis to the nasal cavity and from the glottis to the oral cavity following the opening of the velar door.

One method for modeling nasal consonants is to input three tubes: one for the nasal cavity, other for the oral cavity (closed at the articulation point of the nasal consonant), and another one for the pharyngeal cavity [4,7,8].

The spectra of nasal consonants are characterized by nasal formants, which are due to the combined nasal-pharyngeal tube.

Knowledge of the physiology of the vocal tract is important both for the MRI study of the vocal tract itself and for the understanding of pathologies and rehabilitation therapies [9].

## Method

A narrative review was performed by searching the MeSH terms “vocal tract” and “MRI” in the PubMed database. Only researches in English were considered. Studies that use functional MRI with activation of areas of the brain associated with the vocal tract, studies on swallowing, and studies involving other imaging methods (e.g. CT) were excluded. Only experimental studies were included (excluding reviews). Studies were selected by relevancy, including studies that explore MRI morphological analysis of the entire vocal tract.

## Results

### Magnetic Resonance Imaging of Vocal Tract

Several articles have been published in literature regarding the application of MRI in the study of the vocal tract and in “speech research” (Table 1), a relatively new field that was first studied the mid-1990s.

**Table 1. types of main vocal tract MRI studies in literature**

Aim of main vocal tract MRI studies	N° of articles
Describe modifications of vocal tract in singing	11
Analyze technical feasibility and optimization of MRI sequences	14
Describing vocal tract in pathological conditions	5
3D reproduction of vocal tract and segmentation	8
Vocal tract in vowel or articulatory phonetics	17
Others	11

**Note.** Types of main vocal tract MRI studies in literature (searched in PubMed with key terms “vocal tract” and “MRI” and selected by relevancy), divided by aim/objective of the study. N.B.: some articles are counted in more than one field.

## Discussion

In addition to the study and representation of the vocal tract through MRI, aimed at confirming the acoustic theory of language [10], one of the first fields of interest was the technical aspect regarding the feasibility of MR study and the optimization of sequences. There are known studies with magnets from 0.5 T up to 3 T, EPI, GRE, SSFSE sequences [1], with temporal resolutions reached in the order of tens of milliseconds. However, effort to reach best temporal resolution came at the expense of the spatial resolution with a consequent search for the best trade-off and optimization of the MR technique [11].

MR images can also be obtained and analyzed on different planes. The most useful and widely-utilized is the sagittal plane.

The research also led to the development of dedicated coils and microphones.

MRI also allows you to obtain 2D images, which are especially useful for evaluation on a single plane, usually at mid-sagittal plane. Sometimes, MRI creates 3D images, allowing a better definition, and the possibility of reproduction of the conformations of the vocal tract through 3D printers. Reproduction using 3D printers has made it possible to compare the original sound with those emitted by instruments that use 3D reproduction [12].

Together with the development and optimization of MR sequences, several studies have evaluated the different applications in post-imaging such as identification of “anatomical landmarks” and analysis of segmentation in different portions of the vocal tract [13]. This is done by comparing different segmentation programs and developing a protocol for the application of segmentation programs, and its correlation with the emitted sound.

His direction provides an interesting prospect of use in future clinical practice and speech rehabilitation.

A computational analysis of the vocal tract associated with automatic analysis of the segmentation can work to identify the necessary compensatory movements of the articulators during speech production [14].

Separately, some studies have evaluated variations in the vocal tract in professional singers while they perform certain singing techniques.

In the singing voice, vocal tract adjustments are needed to change sound quality or increase carrying power [15]. Using MRI, such adjustments have been observed in professional sopranos [16], opera tenors [17], and male altos [18]. In the last ones, some observed that falsetto stage register is associated with a narrowing of the pharynx and larynx, with increases of lip and jaw opening, using different vocal strategies than tenors.

In 2020, our group analyzed the conformation of the vocal tract in a professional diphonic male singer. Using a 1,5 T MRI scanner and commercially available FIESTA sequence [1], we described in detail the conformation of the vocal tract in different overtone singing techniques (L-technique, J-technique, and NG technique), and one effect (Ezengileer) applied to L-technique. For each overtone technique we evaluated MRI movement of the lips, tongue, and velopharyngeal closure and the relationship among the tongue and pharyngeal posterior wall/soft palate. In particular, in L-technique (or “double cavity technique”), that consists in a scale of pitches with a sound similar to a lingual consonant, we demonstrated the division of the oral cavity in two chambers, divided by lingual tip attached to hard palate and modulation of pitches mostly by tongue movement and changing of conformation of size and shape of oral and pharyngeal cavity, according to fundamental frequencies and formants [19]. We also observed velum movements, noticing a specific use of the velum in a Tuvan traditional style called “Ezengileer”, in which the velum movements, which open and close rhythmically the velo-pharyngeal port, add a percussive effect to the sound produced; we have observed also an important role of the velum in the NG technique, in which the velum movement, in cooperation with the tongue, produces a sound that is similar to what happens in a sequence composed of a nasal velar consonant, followed by a velar plosive consonant (as the name of the technique, “NG” suggests).

In 2017, Hagedorn et al. [20] applied rtMRI (i.e., real time MRI) in a study of the vocal tract of an apraxic subject to demonstrate the production of covert, intrusive speech gestures in repetitive and non-repetitive speech, and the multiple hidden initiation gestures when attempting to produce complex change and articulation of multiple parts of vocal tract. A silent articulation of consonants suggests that rtMRI is able to capture alteration of speech, such as apraxic speech, as previously described in literature.

Yamasaki et al [21,22] suggest that reduced anterior-posterior dimensions of the larynx may be a morphological characteristic of patients with vocal nodules. They found that habitual VT adjustment of dysphonic patients are different at rest and during phonation. Furthermore, some therapeutic exercise can promote positive VT change and reduce differences.

## Conclusions

Magnetic Resonance Imaging is potentially the best method to study the vocal tract physiology during voice production, thanks to its non-invasiveness, absence of ionizing radiation, and possibility to analyze the entire vocal tract during movement of each of its parts. This, together with the co-recording of emitted sound, can provide detailed information on the physiology of speech.

Most recent studies have achieved good results in representation of changes in the vocal tract during emission of vowels and singing, which require slower movements, while further developments in MR technique are necessary to allow an equally detailed and commercially available study of faster movements that participate in the articulation of speaking.

MRI of vocal tract, although promising, is subject to the typical limitations of MRI, i.e., the impossibility of performing the examination in claustrophobic subjects or with metal objects or medical devices —not compatible with MRI and requiring the absolute collaboration of the patient in absence of dental prostheses— that can alter the images obtained or prevent the examination.

In the future, detailed analysis of the movement of anatomical structures and the segmentation of the vocal tract could perhaps offer some good prospects for its clinical use even in all conditions that require speech therapy rehabilitation.

## References

1. Brown MA, Semelka RC. MR imaging abbreviations, definitions, and descriptions: a review. *Radiology*. 1999;213:647–62. doi: <https://doi.org/10.1148/radiology.213.3.r99dc18647>.
2. Harrington J, Cassidy S. *Techniques in speech acoustics*. Text, speech and language technology. Dordrecht: Springer; 1999. doi: <https://doi.org/10.1007/978-94-011-4657-9>
3. Stevens KN, Kasowski S, Fant CGM. An electric analog of the vocal tract. *J Acoust Soc Am*. 1953;25(4):734. doi: <https://doi.org/10.1121/1.1907169>
4. Fant, G. *The acoustic theory of speech production*. The Hague: Mouton; 1960.
5. Stevens KN, House AS. Perturbation of vowel articulations by consonantal context: an acoustical study. *J. Speech Hear. Res*. 1963;6:111–28. doi: <https://doi.org/10.1044/jshr.0602.111>

6. Behrens S, Blumstein SE. On the role of the amplitude of the fricative noise in the perception of place of articulation in voiceless fricative consonants. *J Acoust Soc Am*. 1988 Sep;84(3):861-7. doi: <https://doi.org/10.1121/1.396655>.
7. Stevens KN. On the quantal nature of speech. *J Phon*. 1989;17:3-45. doi: [https://doi.org/10.1016/S0095-4470\(19\)31520-7](https://doi.org/10.1016/S0095-4470(19)31520-7)
8. Flanagan JL. *Speech analysis synthesis and perception*. 2nd Ed. New York: Springer-Verlag; 1972. doi: <https://doi.org/10.1007/978-3-662-01562-9>
9. Manzano Aquihuatl, C., Guzmán, M. Rehabilitación vocal fisiológica con ejercicios de tracto vocal semiocluído. *Revista de Investigación e Innovación en Ciencias de La Salud*. 2021;3(1):61–86. doi: <https://doi.org/10.46634/riics.68>
10. Sulter AM, Miller DG, Wolf RF, Schutte HK, Wit HP, Mooyaart EL. On the relation between the dimensions and resonance characteristics of the vocal tract: a study with MRI . *Magn Reson Imaging*. 1992;10(3):365-73. doi: [https://doi.org/10.1016/0730-725x\(92\)90507-v](https://doi.org/10.1016/0730-725x(92)90507-v)
11. Lingala SG , Sutton BP, Miquel ME, Nayak KS. Recommendations for real-time speech MRI. *J Magn Reson Imaging*. 2016 Jan;43(1):28-44. doi: <https://doi.org/10.1002/jmri.24997>
12. Birkholz P, Kürbis S, Stone S, Häsner P, Blandin R, Fleischer M. Printable 3D vocal tract shapes from MRI data and their acoustic and aerodynamic properties. *Sci Data*. 2020 Aug 5;7(1):255. doi: <https://doi.org/10.1038/s41597-020-00597-w>
13. Eslami M, Neuschaefer-Rube C, Serrurier A. Automatic vocal tract landmark localization from midsagittal MRI data. *Sci Rep*. 2020 Jan 30;10(1):1468. doi: <https://doi.org/10.1038/s41598-020-58103-6>.
14. Vasconcelos MJ, Ventura SM, Freitas DR, Tavares JM. Towards the automatic study of the vocal tract from magnetic resonance images. *J Voice*. 2011 Nov;25(6):732-42. Epub 2010 Oct 16. doi: <https://doi.org/10.1016/j.jvoice.2010.05.002>.
15. Echternach M, Markl M, Richter B. Dynamic real-time magnetic resonance imaging for the analysis of voice physiology. *Curr Opin Otolaryngol Head Neck Surg*. 2012 Dec;20(6):450-7. doi: <https://doi.org/10.1097/MOO.0b013e3283585f87>
16. Bresch E, Narayanan S. Real-time magnetic resonance imaging investigation of resonance tuning in soprano singing. *J Acoust Soc Am*. 2010 Nov;128(5):335-41. doi: <https://doi.org/10.1121/1.3499700>
17. Echternach M, Sundberg J, Arndt S, Markl M, Schumacher M, Richter B. Vocal tract in female registers--a dynamic real-time MRI study. *J Voice*. 2010 Mar;24(2):133-9. Epub 2009 Jan 29. doi: <https://doi.org/10.1016/j.jvoice.2008.06.004>.
18. Echternach M, Sundberg J, Baumann T, Markl M, Richter B. Vocal tract area functions and formant frequencies in opera tenors' modal and falsetto registers. *J Acoust Soc Am*. 2011 Jun;129(6):3955-63. doi: <https://doi.org/10.1121/1.3589249>

19. Barbiera F, Lo Casto A, Murmura B, Bortoluzzi G, Orefice I, Gucciardo AG. Dynamic Fast Imaging Employing Steady State Acquisition Magnetic Resonance Imaging of the Vocal Tract in One Overtone Male Singer: Our Preliminary Experience. *J Voice*. 2020 Jun 26;1997(20):30184-3. doi: <https://doi.org/10.1016/j.jvoice.2020.05.016>
20. Hagedorn C, Proctor M, Goldstein L, Wilson SM, Miller B, Gorno-Tempini ML, Narayanan SS. Characterizing Articulation in Apraxic Speech Using Real-Time Magnetic Resonance Imaging. *J Speech Lang Hear Res*. 2017 Apr 14;60(4):877-891. doi: [https://doi.org/10.1044/2016\\_JSLHR-S-15-0112](https://doi.org/10.1044/2016_JSLHR-S-15-0112)
21. Yamasaki R, Behlau M, do Brasil OdeO, Yamashita H. MRI anatomical and morphological differences in the vocal tract between dysphonic and normal adult women. *J Voice*. 2011 Nov;25(6):743-50. doi: <https://doi.org/10.1016/j.jvoice.2010.08.005>
22. Yamasaki R, Murano EZ, Gebrim E, Hachiya A, Montagnoli A, Behlau M, Tsuji D. Vocal Tract Adjustments of Dysphonic and Non-Dysphonic Women Pre- and Post-Flexible Resonance Tube in Water Exercise: A Quantitative MRI Study. *J Voice*. 2017 Jul;31(4):442-454. doi: <https://doi.org/10.1016/j.jvoice.2016.10.015>