Luuk Voskuilen

Advancing MRI of the tongue

Imaging muscular architecture and swallowing functionality

Advancing MRI of the tongue:
imaging muscular architecture and swallowing functionality

Luuk Voskuilen

A grant for the purchase of the flexible receiver coil for tongue imaging was provided by the Maurits en Anna de Kock Stichting.

ISBN: 978-94-6423-641-5
Cover design & lay-out: L. Voskuilen & M. Solleveld
Printed by: Proefschriftmaken.nl

Advancing MRI of the tongue:
imaging muscular architecture and swallowing functionality

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor

aan de Universiteit van Amsterdam

op gezag van de Rector Magnificus

Prof. dr. ir. K.I.J. Maex

ten overstaan van een door het College voor Promoties ingestelde commissie,

in het openbaar te verdedigen in de Agnietenkapel

op donderdag 17 februari 2022, te 16:00 uur

door **Luuk Voskuilen**

geboren te Amersfoort

Promotiecommissie

Promotores:	Prof. dr. L.E. Smeele	Universiteit van Amsterdam
	Prof. dr. ir. A.J. Nederveen	Universiteit van Amsterdam
Copromotores:	Prof. dr. A.J.M. Balm	Universiteit van Amsterdam
	Dr. ir. F. van der Heijden	Universiteit Twente
Overige leden:	Dr. M.M.L. de Win	Universiteit van Amsterdam
	Prof. dr. M.P. Schijven	Universiteit van Amsterdam
	Prof. dr. A.G. Becking	Universiteit van Amsterdam
	Prof. dr. ir. B ten Haken	Universiteit Twente
	Prof. dr. A.G. Webb	Universiteit Leiden

Faculteit der Tandheelkunde, Universiteit van Amsterdam

Table of contents

1

Chapter 1

General introduction

General introduction

Notice how, at this moment, your tongue is resting in your mouth. Is it resting against your palate, or is there some air in between? Notice the build-up of saliva in your mouth. When you are swallowing it, consider the series of muscle contractions required for this most basic function of the human body. I apologise for the next few minutes, during which you will be conscious of your tongue, but it does indicate that, most of the time, we take basic functionalities such as swallowing for granted. Imagine how a patient suffering from a disease of the tongue such as cancer may be affected in his or her swallowing, chewing, speaking, kissing, or laughing. Considering the severe impact on a patient's life, this thesis is dedicated to advance our understanding of tongue morphology and its functions by optimizing the medical imaging of this group of muscles. First, however, I will briefly summarise our current understanding of the musculature, functions, and cancer of the tongue.

Tongue anatomy

As the tongue is involved in a wide variety of vital functions such as speaking, chewing, and swallowing, it has to perform a wide range of different motions. These motions are based on a complex architecture of eight tongue muscles. In contrast to other skeletal muscles that originate from bone and insert onto bone, these tongue muscles are only connected to one bony structure (the extrinsic muscles) or no bones at all (the intrinsic muscles). Therefore, the tongue can only partly rely upon the skeleton to provide support.

Similar to trunks and tentacles, the tongue is a muscular hydrostat[1], which means that it maintains its form due to the incompressibility of water. The intrinsic muscles are aligned along three roughly perpendicular directions. In the centre of the tongue, vertical and transverse muscle fibres appear to cross. Microscopically, however, the tongue's centre is constructed from thin sheets or laminae alternating between the vertical and transverse muscles.[2] When these muscles contract, the tongue becomes thinner and narrower, and as a result this lengthens the tongue due to the incompressibility of water. Inversely, by contracting the longitudinal muscles that are positioned around the transverse and vertical muscles, the tongue becomes not only shorter, but also thicker. By contracting only certain parts of tongue muscles, an even greater series of motion can be executed, e.g. by only activating the superior longitudinal fibres the tongue is bent upwards.

In addition to the intrinsic muscles, the extrinsic muscles add further flexibility to the tongue by not only reshaping the tongue, but also displacing it with respect to the skull, jaw or hyoid. The genioglossus protrudes the tongue, the hyoglossus depresses the tongue, and the styloglossus and palatoglossus elevate the tongue.[3] By combining the actions of these extrinsic and the intrinsic muscles fibres, the

wide array of motions required for the tongue's functions can be achieved.

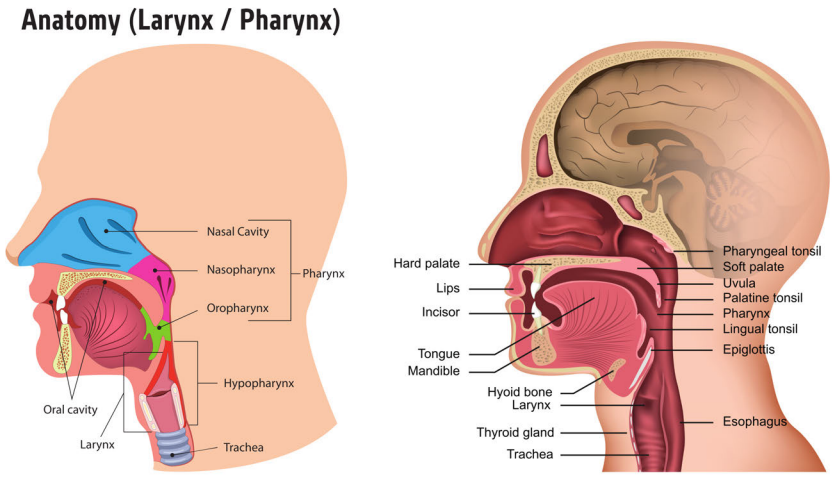


FIGURE 1.1 — The anterior two-thirds of the tongue are located in the oral cavity, which ranges from the lips to circumvallate papillae and to the transition from the hard to the soft palate. The posterior third of the tongue is situated in the oropharynx. Inferiorly, the oropharynx continues as the hypopharynx. As the hypopharynx is involved in the respiratory and the digestive tract, a valve (the epiglottis) is present that prevents food or fluids from entering the larynx, guiding it into the oesophagus, thus preventing aspiration.

Functions of the tongue

One of these functions, in which we differ from the rest of the animal kingdom, is speech.[4] Speech allows us to transmit information to others at a higher rate than other sounds we can produce. To accommodate this, we are able to produce a larger amount of phonemes, speech elements, than our close relatives. These phonemes can be divided into two groups, vowels and consonants.

Vowels are produced when there is 'no major stricture in the vocal tract', which consists of the oral cavity, pharynx, and larynx.[5] As the shape of the tongue largely determines tract narrowing in the oral cavity and the pharynx, the tongue's shape determines the produced vowel. Elevation of the front of the tongue produces sounds like the [i], while elevation of the back of the tongue produces [o]-like vowels. In contrast, depression of the tongue creates sounds like the [a].

The other group of phonemes, consonants, can be divided in to labial, coronal, dorsal, radical, and laryngeal articulations.[5] Although labial articulations, produced by the lips, and laryngeal articulations do not involve the tongue, it is involved in the other three categories of articulations. Coronal articulations such as the [t], require the tongue to temporarily obstruct the air flow by pressing the tip of the tongue against the upper alveolar ridge. The [k], an example of an dorsal articulation, is formed by the tongue by temporarily closing the vocal tract at the

level of the soft palate and the tongue. Finally radical articulations, which are common in Middle-Eastern language, can be produced by deforming of the vocal tract between the base tongue and the pharyngeal wall.

Another basic function in which the tongue is heavily involved is swallowing. Swallowing is defined as the act of transporting fluids, food and saliva from the oral cavity into the stomach.[6] Although swallowing can be performed consciously, the majority of the swallows are involuntary in response to the continuous saliva production in the mouth. A complete swallowing motion is commonly described by the following four phases: the oral preparation, the oral transit, the pharyngeal and the oesophageal phases.

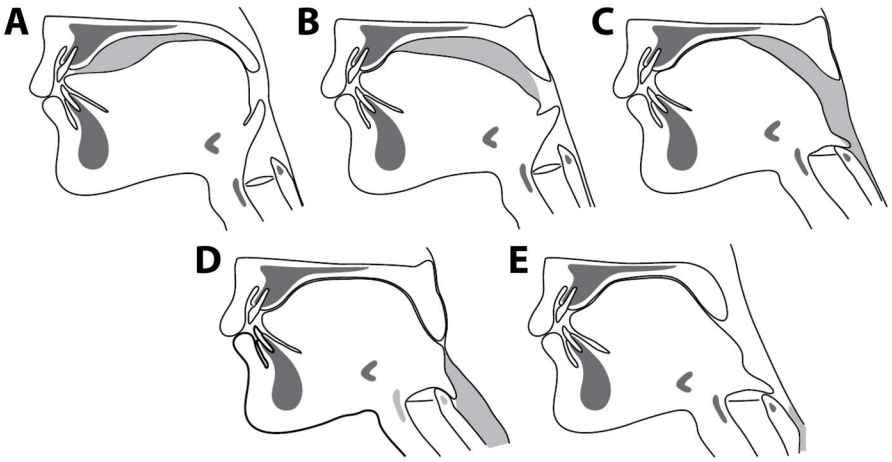


FIGURE 1.2 — The four phases of swallowing: oral preparation phase (a); the oral transit phase (b); the pharyngeal phase (c & d); and the oesophageal phase (e). This figure was adapted from Matsuo et Palmer, 2008[7].

In the oral preparation phase, the tongue and the cheeks direct solid food onto the molars and premolars. With a cyclical jaw motion, chewing or mastication, the food is mixed with saliva and its particle size is reduced to form a bolus, a ball of food ready to be swallowed.[7] At the beginning of the oral transit phase, the bolus is placed on the dorsum of the tongue. The tip of the tongue then pushes against the upper alveolar ridge and gradually extends the area of contact between the tongue and palate posteriorly, thereby propelling the bolus into the oropharynx. During the pharyngeal phase, the soft palate elevates and closes off the nasopharynx. Simultaneously, the hyoid and larynx are pulled upward and anteriorly, which presses the epiglottis over the laryngeal opening sealing off the air way. While the upper oesophageal sphincter opens, the base of the tongue and the muscles in the pharyngeal wall contract to propel the bolus further into the oesophagus. In the oesophageal phase, the bolus is transported by peristalsis, as it is in the rest of the digestive tract. During a peristaltic wave, the muscles in

the oesophageal wall ahead of the bolus relax, and the muscles behind the bolus contract, directing the bolus into the stomach.

The last function, where the tongue is involved in, is breathing. Although the tongue is only minimally active during the actual breathing motion, a constant muscle tone in the base of the tongue keeps the oropharynx open during breathing, thus facilitating the flow of air.[8] This constant muscle tone is accommodated by specialised muscle fibres present in higher numbers in the posterior part of the tongue. In contrast, muscle fibres that are better suited to rapid repositioning of the tongue, are located in the anterior part of the tongue. Although, at first, this seems inconsequential to breathing, a smaller proportion of the slow-acting muscle fibres in the genioglossus has been identified in patients with obstructive sleep apnoea.[9] Although benign diseases such as obstructive sleep apnoea may have a severe impact on the quality-of-life of patients, the remainder of this chapter will focus on cancer of the oral cavity and the oropharynx.

Oral & oropharyngeal cancer

In the United States the incidence rate of oral or oropharyngeal cancer was 8.4 cases per 100,000 in the period from 2000 until 2010.[10] However, the incidence of oral cancer in the US had decreased, which was attributed to less smoking. In the Netherlands, the incidence of oral cancer had increased between 1989 and 2012, which some attribute to increased smoking in women.[11] This explanation is, however, debated as no increase in other smoking-related cancers is observed in women. For oral cancer in the Netherlands, the 5-year survival was a mere 52% in the same period.[11] For oropharyngeal cancer, which includes the base of the tongue, the incidence had increased, which has been attributed to more human papilloma virus (HPV) infections.[10]

Recognising HPV as an increasingly important prognostic factor in oropharyngeal cancer, the TNM classification has been updated with a separate grading system for HPV-mediated oropharyngeal cancer in the eighth edition (table 1.1).[12] The tumour's clinical TNM stage is determined by clinical and radiological factors such as size, depth of invasion, and extension into bone, skin, and major vessels, and regional lymph node or distant metastasis. Pathological findings such as surgical margins and extranodal extension of lymph node metastasis determine the pathologiacal TNM stage. As the TNM classification mainly grades cancer on its anatomical extent, this classification can be converted into four stages that better approximate the patient's prognosis.

TABLE 1.1 — Eighth-edition TNM classification for cancer of the oral cavity and oropharynx.[13] The 'T' indicates the extent of the primary tumour, the 'N' the presence of lymph node metastases, and the 'M' the presence of distant metastases.

	Oral cavity	*Oropharynx non-HPV-related*	*Oropharynx HPV-related*
T1	≤ 2 cm in greatest dimension or ≤ 5 mm depth of invasion	≤ 2 cm	≤ 2 cm
T2	≤ 2 cm and > 5 mm depth, or > 2 cm, ≤ 4 cm and ≤ 10 mm depth	> 2 cm, and ≤ 4 cm	> 2 cm, and ≤ 4 cm
T3	> 2 cm, ≤ 4 cm, and > 10 mm depth, or > 4 cm and ≤ 10 mm depth	> 4 cm, or extension to lingual surface of epiglottis	> 4 cm, or extension to lingual surface of epiglottis
T4a	> 4 cm and > 10 mm depth, or invading through cortical bone of the mandible or maxilla, involving the maxillary sinus or invading the skin of the face	Tumour invades larynx, extrinsic tongue muscle, medial pterygoid, hard palate, or mandible	Tumour invades larynx, extrinsic tongue muscle, medial or lateral pterygoid muscle, pterygoid plates, hard palate, mandible, lateral nasopharynx, or skull base, or encases carotid artery
T4b	Invading masticator space, pterygoid plates, or skull base, or encases internal carotid artery	Tumour invades lateral pterygoid, pterygoid plates, lateral nasopharynx, or skull base, or encases carotid artery	
N1	Single ipsilateral node, ≤ 3 cm	Single ipsilateral node, ≤ 3 cm	Unilateral, all ≤ 6 cm
N2a	Single ipsilateral lymph node, > 3 cm, and ≤ 6 cm	Single ipsilateral node, ≤ 3 cm	Contralateral or bilateral, all ≤ 6 cm
N2b	Multiple ipsilateral nodes, all ≤ 6 cm	Multiple ipsilateral nodes, all ≤ 6 cm	
N2c	Bilateral or contralateral nodes, all ≤ 6 cm	Bilateral or contralateral nodes, all ≤ 6 cm	
N3a	> 6 cm	> 6 cm	> 6 cm
N3b	Clinical extra-nodal extension	Clinical extra-nodal extension	
M1	Distant metastasis	Distant metastasis	Distant metastasis

The TNM prognostic stage in combination with the tumour localization determine, to a large extent, the treatment choices offered to a patient. Surgery is often performed in cancer stages I–III[11], as these tumours generally are well-accessible surgically without causing much functional impairment. In locally-advanced cancers (stage IV), the tumour is surgically resected if it is anatomically and functionally operable. Although anatomical and functional inoperability are not clearly defined in the literature, a tumour is considered anatomically inoperable if the surgery is too great of a risk to the patient, e.g. if the tumour encases the internal carotid artery or invades the skull base. Functional inoperability refers to an unacceptable loss of functionality, but what encompasses unacceptable loss differs

between surgeons[14]. For nearly all physicians, the total resection of the tongue results in an unacceptable loss of functionality.

In locally-advanced cancers that are functionally inoperable, radiotherapy with or without chemotherapy or radiation therapy can be an organ-sparing alternative to surgery.[15] In locoregionally advanced cancers that are treated with surgery, radiotherapy may be added to improve the treatment outcome.[16] For lower stage cancers of the oropharynx, radiotherapy is the preferred treatment because of better functional outcomes in terms of speech and swallowing than the functional outcome following surgery. Moreover, HPV-related tumours have been shown to be more sensitive to irradiation than their non-HPV-related counterparts.[17]

If there is a nodal involvement or if unfavourable pathological findings are present such as extranodal spread or positive margins after surgery, the addition of chemotherapy to the treatment regimen has been shown to improve the survival of patients with locally-advanced head and neck cancer.[18] In addition, a recent and highly-active field of clinical investigation aims to research the first immunotherapeutic agents that may further improve the survival of patients with advanced head-and-neck cancer.

TABLE 1.2 — Conversion table from the TNM classification to the prognostic stages.[13] A higher prognostic stage indicates a worse prognosis.

	Oral cavity or oropharynx non-HPV-related	*Oropharynx HPV-related*
Stage I	T1 N0 M0	T1–2 N0 M0
Stage II	T2 N0 M0	T0–2 N2 M0, T3 N0–2 M0
Stage III	T3 N0 M0, T1–3 N1 M0	T1–4 N3 M0, T4 N0–3 M0
Stage IVa	T4a N0–1 M0, T1-4a N2 M0	T1–4 N0–3 M1
Stage IVb	T1–4 N3 M0, T4b N0–3 M0	
Stage IVc	T1–4 N0–3 M1	

Treatment outcome

Although the five-year survival of patients with oral cancer has improved[11], the treatment of oral cancer may severely reduce the functionality of the tongue. For surgery, the amount of speech loss has been correlated with the volume of the resected tissue.[19] For chewing and swallowing, a similar correlation to this volume was found.[20] The location of the tumour also determines the loss of functionality. A resection of the floor of the mouth mostly impacts chewing, while a resection of the tongue base affects swallowing to a major extent.

Although radiotherapy can be described as an organ-sparing treatment modality, it does induce several side effects that impact the patient's functionality. One common side effect of radiotherapy is trismus, or limited mouth opening.[21] Xerostomia, or a persistent dry mouth, is another side effect caused by reduced salivary production after irradiation of the salivary glands.[22] Both trismus and xerostomia

can directly influence a patient's ability to swallow (dysphagia) or to speak fluently by limiting jaw opening. In the case of dysphagia, this may even lead to aspiration and a potentially lethal pneumonia in an already compromised patient.[23]

Imaging in cancer

Oral and oropharyngeal cancer are primarily evaluated by clinical examination and, if required, endoscopy. During this clinical examination, a biopsy is usually taken to confirm a cancer diagnosis. In order to determine the extent of the cancer, i.e. tumour invasion in surrounding tissues and spread through the lymph nodes, additional imaging is required. In the past, computed tomography (CT) was the preferred imaging modality as it offers high-resolution images. Moreover, a CT exam is performed relatively quickly and easily.[24] However, CT requires ionising radiation, and it is subject to a poor soft-tissue contrast and artefacts (image distortions) by heavy metals in dental fillings.

Currently, magnetic resonance imaging (MRI) is supplanting CT for imaging of oral and oropharyngeal cancer. MRI does not require ionising radiation and the artefacts from dental fillings are less severe.[24] In addition, due to the better soft-tissue contrast of MRI, it better detects 'early bone marrow involvement, pterygopalatine fossa infiltration, prevertebral muscle involvement and perineural tumour spread'.[24] The examination time of MRI is however quite long, increasing its cost and the risk of motion artefacts.

MRI of tongue functionality

As MRI has proven its usefulness in diagnostic imaging, first steps are being taken to investigate its use in the imaging of functionality. Although MRI generally requires a long acquisition time, the incorporation of new acquisition and reconstruction techniques have accelerated MRI in such a way that one slice can be imaged in 20 ms.[25] Due to this excellent temporal resolution, this approach has been coined real-time MRI. Real-time MRI is a promising technique that has been able to visualise normal speech[26] and swallowing[27] in a research setting. However, real-time MRI has not yet entered clinical practice, for which the following reasons may be given: only a single 2D slice is imaged; scanning with a custom-built MRI receiver coil is recommended[28]; and the complex reconstruction techniques are not readily available on clinical scanners.

Although tongue functionality can be imaged directly using real-time MRI, MRI may also visualise changes in the musculature of the tongue.[29] From these changes in musculature, alterations in tongue functionality might be derived. Currently, the musculature of the tongue can be reconstructed using an MRI technique called diffusion tensor imaging (DTI). In DTI, the diffusion of water molecules is measured along multiple directions, which is then modelled by a diffusion tensor. As we

know that the diffusion of water along muscle fibres is larger than that perpendicular to the fibres, the direction of the muscle fibre, and globally, the musculature can be reconstructed.[30] Although DTI of the tongue is promising, it is susceptible to distortions caused by an inhomogeneous magnetic field. Moreover, crossing muscle fibres, which are abundant in the tongue, cannot be modelled correctly by the diffusion tensor.

If these hurdles were to be overcome, there currently is no way to relate changes in tongue musculature to treatment-induced functional loss. Finite-element modelling (FEM) has been shown promising in translating anatomical knowledge of the tongue into a model that can reproduce the wide range tongue motions.[31] However, assuming that differences exist between the musculature of the tongue between two individuals, a separate finite-element model should be constructed for each individual to correctly estimate that individual's functionality.

In the present work, we attempted to answer the question: How can a patient's tongue functionality after treatment for oral or oropharyngeal cancer be estimated more accurately? In order to address this question, we improved upon the currently available MRI methods of imaging the musculature of the tongue (chapter two & three). These images of tongue musculature were incorporated in a personalised FEM that simulates tongue motions of a specific individual (chapter four). Finally, a technique for real-time 3D MRI was implemented, which allows imaging of the entire tongue during swallowing (chapters five & six). In the future, such real-time images may be valuable in the validation of personalised FEM models (chapter four). These validated models hopefully will assist physicians and patients with choosing better treatment options, minimising the loss of functionality after treatment. Furthermore, these models may be helpful in rehabilitation of speech and swallowing after cancer treatment.

References

1. 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, Apr. 1985. DOI: 10.1111/j.1096-3642.1985.tb01178.x.
2. H. Takemoto, "Morphological Analyses of the Human Tongue Musculature for Three-Dimensional Modeling," *Journal of Speech, Language, and Hearing Research*, vol. 44, no. 1, pp. 95–107, 2001. DOI: 10.1044/1092-4388(2001/009).
3. I. Sanders, L. Mu, A. Amirali, H. Su, and S. Sobotka, "The human tongue slows down to speak: Muscle fibers of the human tongue," *Anatomical Record*, vol. 296, no. 10, pp. 1615–1627, 2013. DOI: 10.1002/ar.22755. arXiv: NIHMS150003.
4. W. T. Fitch, "The evolution of speech: A comparative review," *Trends in Cognitive Sciences*, vol. 4, no. 7, pp. 258–267, 2000. DOI: 10.1016/S1364-6613(00)01494-7.
5. P. Ladefoged and I. Maddieson, *The Sounds of the World's Languages*. Blackwell Publishers Ltd, Mar. 1996.
6. W. J. Dodds, E. T. Steward, and J. A. Logemann, "Physiology and Radiology of the Normal Oral and Pharyngeal Phases of Swallowing," *American journal of roentgenology*, vol. 154, pp. 953–963, 1990.
7. K. Matsuo and J. B. Palmer, "Anatomy and Physiology of Feeding and Swallowing: Normal and Abnormal," *Physical Medicine and Rehabilitation Clinics of North America*, vol. 19, no. 4, pp. 691–707, 2008. DOI: 10.1016/j.pmr.2008.06.001.
8. P. Stål, S. Marklund, L. E. Thornell, R. De Paul, and P. O. Eriksson, "Fibre composition of human intrinsic tongue muscles," *Cells Tissues Organs*, vol. 173, no. 3, pp. 147–161, 2003. DOI: 10.1159/000069470.
9. J. M. Cori, F. J. O'donoghue, and A. S. Jordan, "Sleeping tongue: Current perspectives of genioglossus control in healthy individuals and patients with obstructive sleep apnea," *Nature and Science of Sleep*, vol. 10, pp. 169–179, 2018. DOI: 10.2147/NSS.S143296.
10. D. J. Weatherspoon, A. Chattopadhyay, S. Boroumand, and I. Garcia, "Oral cavity and oropharyngeal cancer incidence trends and disparities in the United States: 2000-2010," *Cancer Epidemiology*, vol. 39, no. 4, pp. 497–504, 2015. DOI: 10.1016/j.canep.2015.04.007.
11. B. A. Van Dijk, M. T. Brands, S. M. Geurts, M. A. Merkx, and J. L. Roodenburg, "Trends in oral cavity cancer incidence, mortality, survival and treatment in the Netherlands," *International Journal of Cancer*, vol. 139, no. 3, pp. 574–583, 2016. DOI: 10.1002/ijc.30107.
12. S. H. Huang and B. O'Sullivan, "Overview of the 8th Edition TNM Classification for Head and Neck Cancer," *Current Treatment Options in Oncology*, vol. 18, no. 7, 2017. DOI: 10.1007/s11864-017-0484-y.
13. J. D. Brierley, M. K. Gospodarowicz, and C. Wittekind, *TNM Classification of Malignant Tumours*, 8th edition. 2016.
14. M. A. Kreeft, L. Van der Molen, F. J. Hilgers, and A. J. Balm, "Speech and swallowing after surgical treatment of advanced oral and oropharyngeal carcinoma: a systematic review of the literature," *Eur Arch Otorhinolaryngol*, vol. 266, pp. 1687–1698, 2009. DOI: 10.1007/s00405-009-1089-2.
15. C. Nutting, "Radiotherapy in head and neck cancer management: United Kingdom National Multidisciplinary Guidelines," *The Journal of laryngology and otology*, vol. 130, no. S2, S66–S67, 2016. DOI: 10.1017/S0022215116000463.
16. A. Argiris, M. V. Karamouzis, D. Raben, and R. L. Ferris, "Head and neck cancer.," *Lancet (London, England)*, vol. 371, no. 9625, pp. 1695–1709, 2008. DOI: 10.1016/S0140-6736(08)60728-X.
17. K. Lindel, K. T. Beer, J. Laissue, R. H. Greiner, and D. M. Aebersold, "Human papillomavirus positive squamous cell carcinoma of the oropharynx: A radiosensitive subgroup of head and neck carcinoma," *Cancer*, vol. 92, no. 4, pp. 805–813, 2001. DOI: 10.1002/1097-0142(20010815)92:4<805::AID-CNCR1386>3.0.CO;2-9.
18. P. Savvides, "The Role of Chemotherapy in the Management of Patients with Head and Neck Cancer," *Seminars in Plastic Surgery*, vol. 24, no. 02, pp. 137–147, May 2010. DOI: 10.1055/s-0030-1255331.
19. G. Nicoletti, D. S. Soutar, M. S. Jackson, A. a. Wrench, G. Robertson, and C. Robertson, "Objective assessment of speech after surgical treatment for oral cancer: experience from 196 selected cases.," *Plastic and reconstructive surgery*, vol. 113, no. 1, pp. 114–125, 2004. DOI: 10.1097/01.PRS.0000095937.45812.84.
20. G. Nicoletti, D. S. Soutar, M. S. Jackson, A. a. Wrench, and G. Robertson, "Chewing and swallowing after surgical treatment for oral cancer: functional evaluation in 196 selected cases.," *Plastic and reconstructive surgery*, vol. 114, no. 2, pp. 329–338, 2004. DOI: 10.1097/01.PRS.0000131872.90767.50.
21. C. Weber, S. Dommerich, H. W. Pau, and B. Kramp, "Limited mouth opening after primary therapy of head and neck cancer," *Oral and Maxillofacial Surgery*, vol. 14, no. 3, pp. 169–173, 2010. DOI: 10.1007/s10006-010-0220-2.
22. A. P. Jellema, B. J. Slotman, P. Doornaert, C. R. Leemans, and J. A. Langendijk, "Impact of Radiation-Induced Xerostomia on Quality of Life After Primary Radiotherapy Among Patients With Head and Neck Cancer," *International Journal of Radiation Oncology Biology Physics*, vol. 69, no. 3, pp. 751–760, 2007. DOI: 10.1016/j.ijrobp.2007.04.021.
23. H. M. Starmer, H. Quon, R. Kumar, S. Alcorn, E. Murano, B. Jones, and I. Humbert, "The Effect of Radiation Dose on Swallowing: Evaluation of Aspiration and Kinematics," *Dysphagia*, vol. 30, no. 4, pp. 430–437, 2015. DOI: 10.1007/s00455-015-9618-1.
24. D. W. T. Vogel, T. Vogel, P. Zbaeren, and H. C. Thoeny, "Cancer of the oral cavity and oropharynx," *Cancer Imaging*, vol. 10, no. 1, pp. 62–72, 2010. DOI: 10.1102/1470-7330.2010.0008.
25. M. Uecker, S. Zhang, D. Voit, A. Karaus, K. D. Merboldt, and J. Frahm, "Real-time MRI at a resolution of 20 ms," *NMR in Biomedicine*, vol. 23, no. 8, pp. 986–994, 2010. DOI: 10.1002/nbm.1585.
26. S. G. Lingala, Y. Zhu, Y.-C. Kim, A. Toutios, S. Narayanan, and K. S. Nayak, "A fast and flexible MRI system for the study of dynamic vocal tract shaping," *Magnetic Resonance in Medicine*, vol. 77, no. 1, pp. 112–125, Jan. 2017. DOI: 10.1002/mrm.26090.
27. S. Zhang, A. Olthoff, and J. Frahm, "Real-time magnetic resonance imaging of normal swallowing," *Journal of Magnetic Resonance Imaging*, vol. 35, no. 6, pp. 1372–1379, 2012. DOI: 10.1002/jmri.23591.
28. S. G. Lingala, B. P. Sutton, M. E. Miquel, and K. S. Nayak, "Recommendations for real-time speech MRI," *Journal of Magnetic Resonance Imaging*, vol. 43, no. 1, pp. 28–44, 2016. DOI: 10.1002/jmri.24997.

29. E. Z. Murano, H. Shinagawa, J. Zhuo, R. P. Gullapalli, R. A. Ord, J. L. Prince, and M. Stone, "Application of diffusion tensor imaging after glossectomy," *Otolaryngol Head Neck Surg*, vol. 143, no. 2, pp. 304–306, 2010. DOI: 10.1017/CBO9781107415324.004. arXiv: arXiv:1011.1669v3.

30. T. a. Gaige, T. Benner, R. Wang, V. J. Wedeen, and R. J. Gilbert, "Three dimensional myoarchitecture of the human tongue determined in vivo by diffusion tensor imaging with tractography," *Journal of Magnetic Resonance Imaging*, vol. 26, no. 3, pp. 654–661, 2007. DOI: 10.1002/jmri.21022.

31. S. Buchaillard, "Activations musculaires et mouvements linguaux: modélisation en parole naturelle et pathologique," Ph.D. dissertation, Université Joseph-Fourier - Grenoble, 2007.

2

Chapter 2

Crossing muscle fibers of the human tongue resolved in vivo using constrained spherical deconvolution

Luuk Voskuilen, Valentina Mazzoli, Jos Oudeman, Alfons J.M. Balm, Ferdinand van der Heijden, Martijn Froeling, Maartje M.L. de Win, Gustav J. Strijkers, Ludi E. Smeele, and Aart J. Nederveen

Journal of Magnetic Resonance Imaging

2019

Abstract

Background: Surgical resection of tongue cancer may impair swallowing and speech. Knowledge of tongue muscle architecture affected by the resection could aid in patient counseling. Diffusion tensor imaging (DTI) enables reconstructions of muscle architecture in vivo. Reconstructing crossing fibers in the tongue requires a higher-order diffusion model.

Purpose: To develop a clinically feasible diffusion imaging protocol, which facilitates both DTI and constrained spherical deconvolution (CSD) reconstructions of tongue muscle architecture in vivo.

Study Type: Cross-sectional study.

Subjects/Specimen: One ex vivo bovine tongue resected en bloc from mandible to hyoid bone. Ten healthy volunteers (mean age 25.5 years; range 21-–34 years; four female).

Field Strength/Sequence: Diffusion-weighted echo planar imaging at 3 T using a high-angular resolution diffusion imaging scheme acquired twice with opposing phase-encoding for B_0-field inhomogeneity correction. The scan of the healthy volunteers was divided into four parts, in between which the volunteers were allowed to swallow, resulting in a total acquisition time of 10 minutes.

Assessment: The ability of resolving crossing muscle fibers using CSD was determined on the bovine tongue specimen. A reproducible response function was estimated and the optimal peak threshold was determined for the in vivo tongue. The quality of tractography of the in vivo tongue was graded by three experts.

Statistical Tests: The within-subject coefficient of variance was calculated for the response function. The qualitative results of the grading of DTI and CSD tractography were analyzed using a multilevel proportional odds model.

Results: Fiber orientation distributions in the bovine tongue specimen showed that CSD was able to resolve crossing muscle fibers. The response function could be determined reproducibly in vivo. CSD tractography displayed significantly improved tractography compared with DTI tractography ($P = 0.015$).

Data Conclusion: The 10-minute diffusion imaging protocol facilitates CSD fiber tracking with improved reconstructions of crossing tongue muscle fibers compared with DTI.

Level of Evidence: 2

Technical Efficacy: Stage 1

Introduction

Carcinomas involving the tongue are preferably removed surgically.[1] The amount of resected tissue, which can be substantial, is associated with loss in tongue functionality affecting speech,[2] mastication, and swallowing.[3] The location and type of resection play a prominent role in the expected extent of loss of functionality.[4] Detailed insights into how individual tongue muscles are involved in complex tongue functionality such as speech and swallowing, however, are lacking, especially in the presence of a tumor. Therefore, some might want to study the complex tongue muscle anatomy in vivo.

This complex tongue muscle anatomy is comprised of intrinsic and extrinsic muscle groups. The intrinsic musculature of the tongue belongs to the category of muscular hydrostats,[5] consisting of transverse (TRA) and vertical (VER) muscles, surrounded by longitudinally oriented fibers (superior longitudinal [SL] and inferior longitudinal [IL] muscles).[5,6] Furthermore, the tongue contains four extrinsic muscles (genioglossus [GG], hyoglossus [HG], styloglossus [SG], and palatoglossus [PG] muscles), which originate from the mandible, hyoid bone, styloid process of the skull, and the palate, respectively.[7] To predict the functional outcome after surgical resection in such complex musculature, one cannot simply rely on experience or common imaging modalities.[8] To better model and understand residual functionality after surgery, it is essential to obtain the muscular architecture on an individual patient basis, especially the musculature that is to be resected.

Diffusion weighted imaging (DWI) has been used as a non-invasive method to image tongue muscular architecture in vivo.[9] DWI is capable of quantifying the self-diffusion of water in biological tissues by measuring the attenuation of the magnetic resonance imaging (MRI) signal in the presence of diffusion-encoding gradients. Diffusion of water in skeletal muscle is anisotropic[10]: Perpendicular to muscle fibers, water diffusion is restricted by cell membranes as well as by structured intracellular and extracellular proteins, whereas along the muscle fibers water molecules diffuse more freely.

In diffusion-tensor imaging (DTI), this orientation dependence of diffusion is geometrically described by a rank-2 semi-definite positive tensor. The first eigenvector of the diffusion tensor, which corresponds to the largest eigenvalue, has been shown to align with the local muscle fiber direction in skeletal muscle.[11] In a technique called tractography, the principal diffusion vectors of individual voxels are connected via streamlines, which results in 3D reconstructions of the underlying muscle architecture.[12]

A drawback of DTI is that the diffusion tensor can only describe a single fiber orientation per voxel. However, due to the relatively large voxel size in DWI compared with the diameter of the myocytes, multiple muscle fiber population appear to cross or merge within a voxel. This is particularly relevant for the tongue, due to its complex structure with many crossing muscle fibers. DTI is

therefore generally not capable of resolving the true tongue muscle architecture.

Crossing fibers can be resolved, though, by higher-order diffusion models, such as by constrained spherical deconvolution (CSD).[13] CSD is based on the assumption that a fiber orientation distribution (FOD) can be calculated by deconvolution of the measured DWI signal and a response function (RF), which corresponds to the DWI signal of a single fiber population.[14] The FOD, DWI signal, and RF are described using spherical harmonics. Therefore, a maximum harmonic order (l_{max}) has to be defined, where a higher l_{max} allows a higher maximal angular resolution to be obtained. Additionally, a peak threshold is chosen to remove small spurious peaks in the calculated FOD.

While for DTI the measurement of six noncollinear diffusion directions is theoretically sufficient, CSD requires sampling of the diffusion along many more directions. The minimal number of gradient directions required is defined by

$$n_{min} = \frac{1}{2}(l_{max} + 1)(l_{max} + 2)$$

thus, for $l_{max} = 8$ at least 45 unique gradient directions are necessary. Therefore, the diffusion sampling strategy required for CSD is referred to as high-angular resolution diffusion imaging (HARDI).[15] Application of this approach to the human tongue is not straightforward, since measurements of many diffusion directions necessitates a long scan time. Such a long scan time for tongue imaging is undesirable, due to signal dropouts from motions such as involuntary swallowing. In DWI, the scan time is commonly reduced by using an echo-planar imaging (EPI) readout, and because other acceleration techniques such as multishot imaging introduce motion-related phase errors.

A major disadvantage of spin-echo (SE)-EPI is the small bandwidth per voxel in the phase-encoding direction, which increases the susceptibility to B_0-inhomogeneity artifacts. Therefore, air–tissue interfaces and metal-based dental prostheses, which distort the B0-field, will result in deformations in the phase-encoding direction. These distortions can be corrected for by using an algorithm from the FSL library, Topup.16 Topup estimates the inhomogeneity field by estimating the distortions, which are equal in magnitude but opposite in direction, between two DWI datasets acquired with opposite phase-encoding directions. Subsequently, the two images are combined, partially restoring spatial encoding information lost in the distorted images due to signal pile-up, using information from the opposite image, where the signal is dispersed.[16]

The aim of this study was to develop a clinically feasible scan protocol and postprocessing pipeline for reconstructing the complex muscular architecture of the human tongue in vivo allowing both DTI and CSD fiber tractography.

Methods

Bovine Tongue Specimen

One ex vivo bovine tongue was acquired from a local slaughterhouse. The specimen was resected en bloc from mandible and hyoid bone. The imaging of the tongue was performed within 24 hours after harvest. The bovine tongue (~45 cm in length) was cast into alginate polymer before scanning. The alginate cast reduced B_0-inhomogeneity, and provided support for the tongue to minimize deformation due to the weight of the coil.

Imaging of Bovine Tongue

Standard torso (16 channels) and posterior coils (12 channels) were used for acquisition on a 3 T Ingenia scanner (Philips Medical Systems, Best, The Netherlands). Diffusion-weighted images were acquired using a HARDI[15] gradient scheme, in which 64 diffusion directions were evenly spaced over a sphere, optimized for gradient load.[17] Although 45 gradient directions are theoretically sufficient for l_{max} = 8, we acquired more diffusion directions to reduce the susceptibility of the calculated FODs to noise. A b-value of 1000 s/mm^2 was used, which is higher than the b-value for the in vivo acquisitions to compensate for the loss of diffusivity postmortem. After every eight diffusion-weighted images one b0-image was acquired. All images were acquired with two opposing phase-encoding directions, RL and LR. Other imaging parameters were: SE-EPI; echo train length (ETL): 31; field of view (FOV): 192 (AP) x 154 (RL) x 420 (FH) mm^3; voxel size: 2.4 x 2.4 x 2.4 mm^3; matrix size: 80 x 62; echo time (TE): 68 msec; repetition time (TR) = 40.5 sec; SENSE: 2; no partial Fourier; bandwidth per pixel in the phase-encoding direction: 46.4 Hz; number of signal acquired (NSA): 1; SPIR and SSGR fat suppression; scan time: 2 hours.

Volunteers

Ten healthy volunteers (mean age 25.5 years; range 21–34 years; 4 female) were scanned. All volunteers provided written informed consent. Volunteers were excluded if a contraindication for MRI or orthodontic braces was present. This study was approved by the institutional Ethics Committee Review Board.

In Vivo Imaging

Two flexible surface coils, 20 cm in diameter, were gently strapped to the cheeks of the volunteers for acquisition on a 3 T Ingenia scanner (Philips Medical Systems). The scan protocol was repeated with an interval of no more than 10 minutes between scans to assess reproducibility. In between the two scans, the volunteer

left the table and was repositioned. During scanning, volunteers were instructed to let their tongue rest against their palate, stabilizing the tongue and removing as much excessive air from the oral cavity as possible.

The same diffusion gradient scheme as for the ex vivo experiment with 64 diffusion-encoding directions was divided into two sets of 32 directions in order to maintain an equally spaced distribution. Other imaging parameters were: SE-EPI; ETL: 25; FOV: 192 (AP) x 156 (RL) x 84 (FH) mm^3; voxel size: 3 x 3 x 3 mm; matrix size: 64 x 49; TE: 60 msec; TR: 3.4 sec; SENSE: 2; no partial Fourier; bandwidth per pixel in the phase-encoding direction: 56.4 Hz; NSA: 1; SPIR and SSGR fat suppression; b-value: 700 s/mm^2; scan time: 10:03 minutes. The total scan protocol consisted of four parts (phase-encoding: RL & LR and 2 x 32 gradient directions) of 2:30 minutes each. In between these parts, we allowed the volunteer to swallow. This approach minimized tongue motion due to the swallowing reflex initiated by the build-up of saliva in the oral cavity.

Data Processing

Diffusion-weighted images were smoothed with a Rician noise filter in DTITools for Mathe-matica.[12] The B_0-inhomogeneity field was estimated using the previously described Topup. Subsequently, the DWIs were corrected for inhomogeneity, rigid motion, and eddy currents using FSL.[18] A mask of the whole ex vivo tongue was created by applying a threshold including the top 95% voxels regarding intensity of the corrected b0-image. Masks of the in vivo tongues were defined by manually delineation in ITK-Snap[19] on the corrected b0-images. These masks reduced computation time and served as an outer boundary for tractography.

For both ex vivo and in vivo data, diffusion tensors were fitted using RE-STORE[20], and fractional anisotropy (FA) and mean diffusivity (MD) were calculated in ExploreDTI.[21] The average FA and MD were calculated for a manually delineated region of interest (ROI) in the GG muscle in a known noncrossing area. The diffusion signal in spherical harmonic representation of voxels exceeding these average FA and MD values were averaged to obtain the CSD response function (RF). FODs were subsequently calculated in the entire tongue using $l_{max} = 8$, which was chosen to obtain the largest angular resolution in the FODs possible.

For the in vivo data, whole tongue deterministic DTI tractography was performed using the following parameters chosen empirically: FA range: 0.1–0.6; seed point resolution: 3 mm; step size: 1 mm; angular threshold: 15° tract length: 10–100 mm.

We performed a Bland–Altman analysis on the two repeated measurements of the FA and MD that are used to define the RF. Additionally, the interscan variability was determined by calculating the within-subject coefficient of variation (wsCV) by dividing the standard deviation of the paired difference by the mean of the paired average.

In muscular tissues, the diffusion anisotropy is generally much lower than in brain tissue.[22] Therefore, the RF in muscles is more spherical, increasing the susceptibility to overfitting and thus spurious peaks in FODs. In order to minimize the effects on in vivo tractography of these spurious peaks, true peaks, an FOD peak threshold is employed. This peak threshold removes peaks shorter than the peak threshold times the length of the largest peak. We determine the optimal FOD peak threshold[23] by calculating the number of peaks per voxel for 10 different peak thresholds ranging from 0.05–0.50. The number of voxels containing three peaks, which may be considered spurious, false-positive peaks, has to be minimized. However, a peak threshold that is too high results in an increasing number of false-negative peaks, mainly in voxels with crossing fibers that should contain two peaks. A tradeoff has to be made between false-positive and false-negative peaks.

Finally, whole tongue deterministic CSD tractography was performed in ExploreDTI with an FOD peak threshold of 0.10. The other tracking parameters were the same as for deterministic DTI tractography, however, without the FA constraint. Representative images of DTI and CSD tractography for both repetitions and all 10 subjects were graded by three reviewers on a three-point ordinal scale, where three is the best and one the worst grade. Tractography of the tongue is rare and therefore the reviewers had limited experience with these images. Therefore, two reviewers were chosen with extensive knowledge of tongue muscular anatomy (head and neck surgeons) and one reviewer experienced in brain DTI (radiologist). Five questions were asked: "Can you grade the overall quality of the data", "How well can you distinguish the transverse & vertical muscles?", "How well can you distinguish the hyoglossus and styloglossus?", "How well can you appreciate the superior longitudinal muscle?", "How would you grade the influence of gaps in tractography on the image quality?"

The grades provided by the expert reviewers were modeled using a multilevel proportional odds model in R using the ordinal package.[24] We modeled the grades given by the expert reviewers using the method (DTI or CSD) and repetition as fixed effects, and subjects, expert reviewers, and the five asked questions as random effects with random intercepts. Random slopes were used for variables reviewers and questions regarding the method (DTI or CSD). P-values of the effect of using DTI or CSD and the effect of repetition were calculated by likelihood ratio tests of the full model compared with models without either the method or repetition as fixed effects.

Results

In the bovine tongue specimen (figure 2.1), DTI primary eigenvectors and FOD display similar fiber directionality in noncrossing muscles such as the HG, SG, and GH muscles. However, the enlarged sections of the intrinsic tongue musculature show that DTI (figure 2.1a) is unable to distinguish between TRA and VER muscles. The CSD FODs (figure 2.1b) show two peaks, which indicates that CSD appears to be able to distinguish the TRA and VER muscles. On the interface between the TRA, VER, and SL, where a partial volume effect is present in three muscle directions, three peaks corresponding to these three muscles can be appreciated.

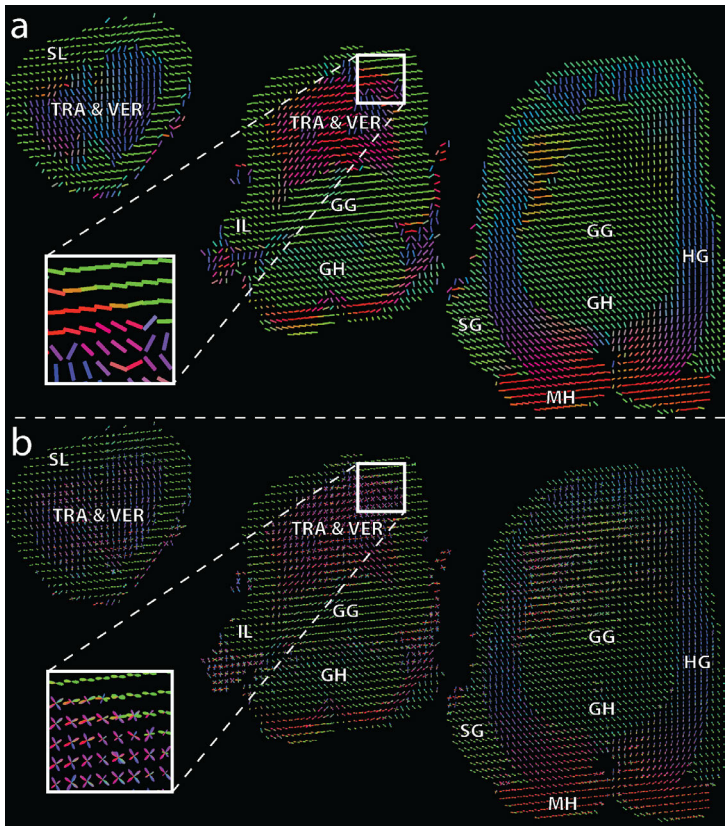


FIGURE 2.1 — Fiber directionality of the bovine tongue determined by the primary eigenvector of DTI (a) and the fiber orientation distribution of CSD (b), color-coded similarly to the human tongue (red: right-left, green: cranio-caudal, blue: anteriorposterior). The following muscles are annotated in the slices: superior longitudinal (SL); transverse (TRA); vertical (VER); genioglossus (GG); inferior longitudinal (IL); geniohyoid (GH); hyoglossus (HG); styloglossus (SG); mylohyoid (MH). The white boxes indicate the location of the inlays, which highlight to ability of CSD to resolve crossing fibers of TRA and VER, while maintaining the correct reconstruction of single fibers of the SL.

FIGURE 2.2 — B_0-inhomogeneity correction using Topup. a: Transverse b0-image with phase encoding along the left-right direction. b: Similar b0-image with phase encoding opposite to a. Image distortions are present along the phase-encoding direction but opposite between a and b; for example, lateral to the tongue (white arrows). c: B_0-inhomogeneity map estimated by Topup. Strong inhomogeneities are present at air–tissue interfaces and near teeth lateral to the tongue (white arrow). d: Resulting combined b0-image with corrected for B_0-inhomogeneity. e: T2-weighted Dixon image serving as a high-resolution undistorted gold standard for the distortion correction.

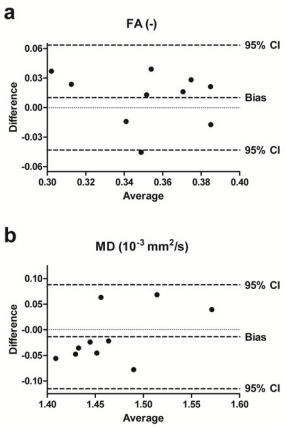


FIGURE 2.3 — Bland–Altman plots of FA and MD used to estimate the CSD response function (RF) using two repeated measurements for 10 healthy volunteers. The RF was determined by drawing an ROI in the GG muscle, where only one fiber direction is present. The two metrics, FA (a) and MD (b) that define the RF were calculated by fitting a diffusion tensor.

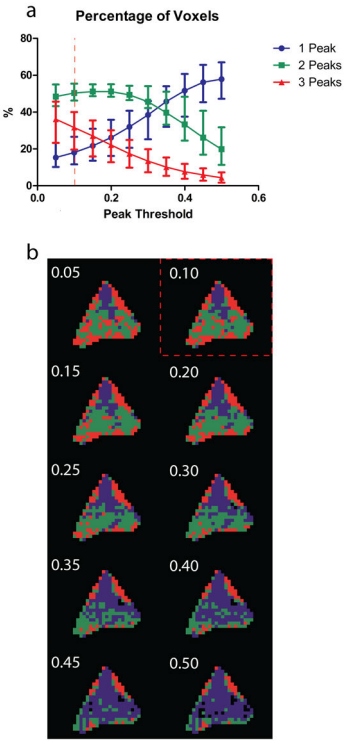


FIGURE 2.4 — The effect of changing the peak threshold on the percentage of voxels with one, two, or three peaks on all volunteers (a); brackets represent the range of percentages. The spatial distribution of the number of peaks per voxel in a single volunteer (b) colored similarly to a. Note that increasing the peak threshold decreases the number of 3-peak-voxels, which generally contain spurious peaks. However, increasing the peak threshold also decreases the number of two-peak-voxels in the center of the tongue, therefore reducing the ability of resolving crossing muscle fibers

In all 10 volunteers, both repetitions of the DWI acquisition were acquired successfully without any subject dropouts due to, for example, tongue motion. Figure 2.2 shows the effect of inhomogeneity correction using Topup on transverse diffusion-weighted images of a single in vivo tongue, where inhomogeneity artifacts are clearly present. Distinct but opposite distortions of b0-images are present along the phaseencoding direction (RL & LR, figure 2.2a,b). The estimated B_0-inhomogeneity map (figure 2.2c) displays field inhomogeneity at expected locations due to differences in magnetic field permeability, i.e., at tissue–tooth interfaces, laterally to the tongue, and at air–tissue interfaces, near the oropharynx and skin. The inhomogeneity map was used to combine and correct the images with opposite phase-encoding, resulting in the corrected images (figure 2.2d), which correspond better to the actual morphology of the tongue (figure 2.2e).

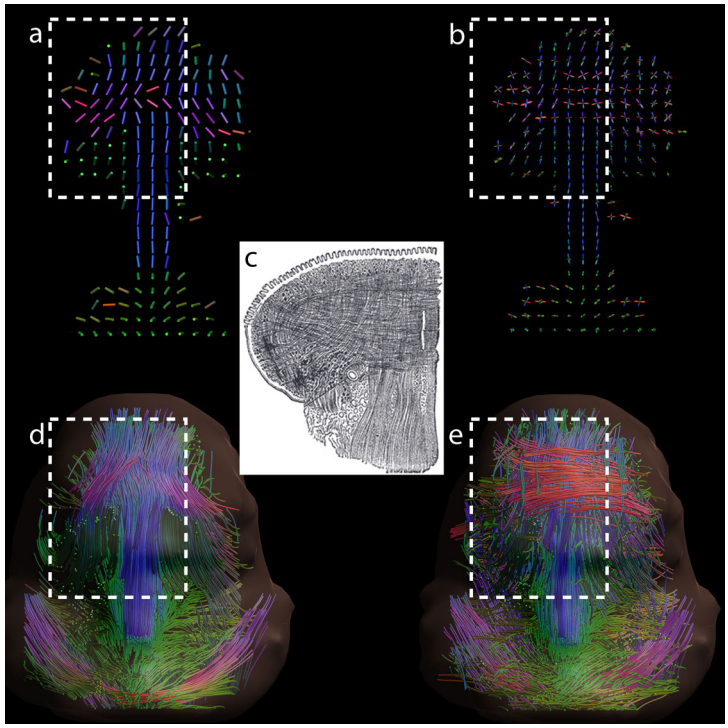


FIGURE 2.5 — Coronal view of in vivo DTI and CSD fiber directionality and tractography in a healthy volunteer compared with textbook anatomy.[25] Images are conventionally color-coded (red: right–left, green: anterior–posterior, blue: inferior–superior). a: Coronal slice of the in vivo human tongue displaying the primary eigenvector of the diffusion tensor as a color-coded cylinder. b: Coronal slice identical to (a) but displaying FODs calculated by CSD. c: Illustration (adapted from Gray's Anatomy) of a coronal slice of the human tongue displaying crossing transverse and vertical muscle fibers. d: Coronal view of DTI fiber tractography. e: Coronal view of CSD fiber tractography. The areas corresponding to the slice from the anatomical atlas have been marked. In (d,e), fibers originating from the apex of the tongue have been removed to improve the visibility of the crossing fibers in the tongue core.

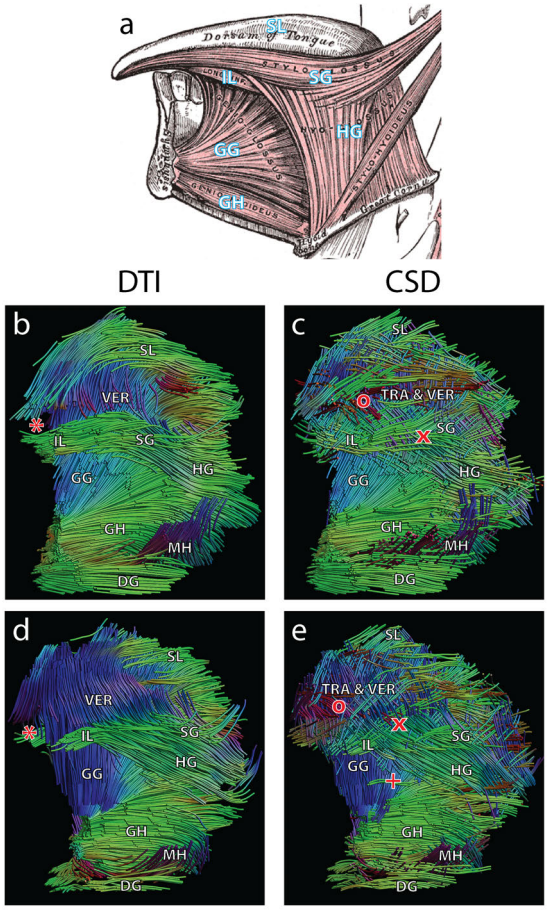


FIGURE 2.6 — Textbook anatomy[25] (a) compared with DTI and CSD tractography of volunteers 9 (b,c) and 5 (d,e) in a frontosagittal view. The following muscles of the tongue and floor of the mouth have been annotated: superior longitudinal (SL); transverse (TRA); vertical (VER); inferior longitudinal (IL); styloglossus (SG); genioglossus (GG); hyoglossus (HG); geniohyoid (GH); mylohyoid (MH); digastric (DG). DTI tractography is unable to reconstruct the tip of the tongue in these cases (*). CSD tractography is able to resolve transverse fibers (o) in the tongue; however, DTI tractography cannot. In DTI tractography the SG and HG are merged into a single fiber bundle, while CS is able to resolve both muscles (x). An example of spurious tracts is present in CSD tractography (+).

The Bland–Altman analysis of two repeated measurements of the FA and MD used for RF estimation is displayed in figure 2.3. For the two DTI metrics defining the RF, the bias is close to zero, suggesting that the systematic error in determining the RF is low. The wsCV was 7.3% for FA and 3.3% for MD.

Figure 2.4a shows the percentage of voxels with one, two, or three peaks as a function of FOD peak threshold. The percentage of voxels with two peaks reaches a maximum around 0.15, suggesting that a peak threshold of 0.15 is ideal for resolving two crossing muscle fibers, which most commonly occurs in the tongue. An example of the spatial distribution of the number of peaks per voxel in a single transverse slice is displayed in figure 2.4b. The three-peak-voxels, which may be considered to contain spurious, false-positive peaks, are mainly located in the periphery and are therefore of little influence on tractography. However, a peak threshold of 0.15 or higher results in an increasing number false-negative peaks, namely, the number of voxels in the posterior part of the tongue containing only one peak. The latter does impact tractography; thus, a peak threshold of 0.10 was chosen as a reasonable compromise between false-positive and false-negative

peaks.

The coronal view of one in vivo human tongue (figure 2.5) illustrates the ability of CSD to resolve crossing fibers in the tongue core in vivo contrary to DTI. The images of FODs and CSD tractography both more closely resemble the anatomical atlas compared with their DTI counterparts. The DTI principal eigenvectors in the TRA & VER muscles appear to be an oblique average of the two FOD peaks. Note that fiber tracts belonging to the tip of the tongue for both DTI and CSD have been removed to improve visibility of the crossing fibers of the tongue core.

DTI and CSD tractography of two typical volunteers is compared with textbook anatomy[25] in figure 2.6. The TRA muscle fibers and fibers in the tip of the tongue are missing in DTI tractography. Additionally, the extrinsic muscles SG and HG appear to have merged in the DTI tractography, while these muscles cross in CSD tractography, the latter of which is in accordance with textbook anatomy. Finally, CSD tractography displays spurious tracts, which are not visible in DTI tractography.

The proportion of the grades given by the expert reviewers for both DTI and CSD are summarized in figure 2.7. Images of CSD tractography have significantly higher odds of being graded higher than images of DTI tractography ($P = 0.015$). These odds were not significantly different between the two repetitions ($P = 0.161$). Images of CSD and DTI tractography that were graded by expert reviewers can be found in figure 2.8.

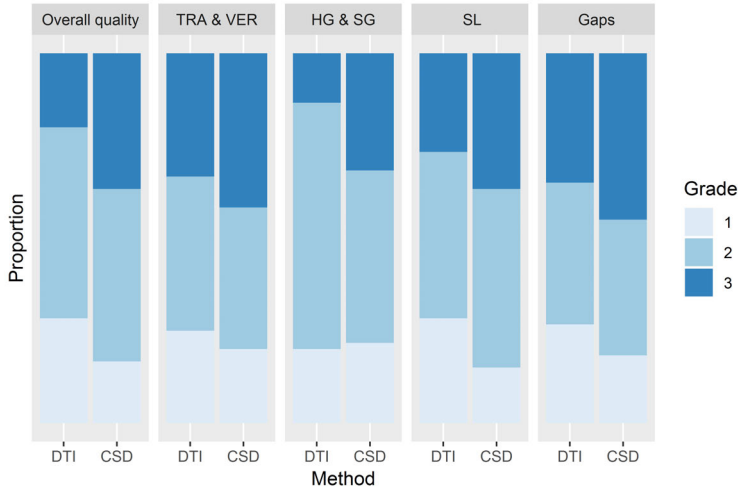


FIGURE 2.7 — Visualization of the distribution of the selected grades for evaluating the quality of CSD and DTI, where three is the best and one the worst grade. Note that the data from all expert reviewers and both repetitions are included here. Five questions were asked focused on overall quality of the data and the visualization of crossing muscles. The odds of receiving a higher grade is significantly higher ($P = 0.015$) for CSD compared with DTI, while no such difference was detected between the two repetitions ($P = 0.161$).

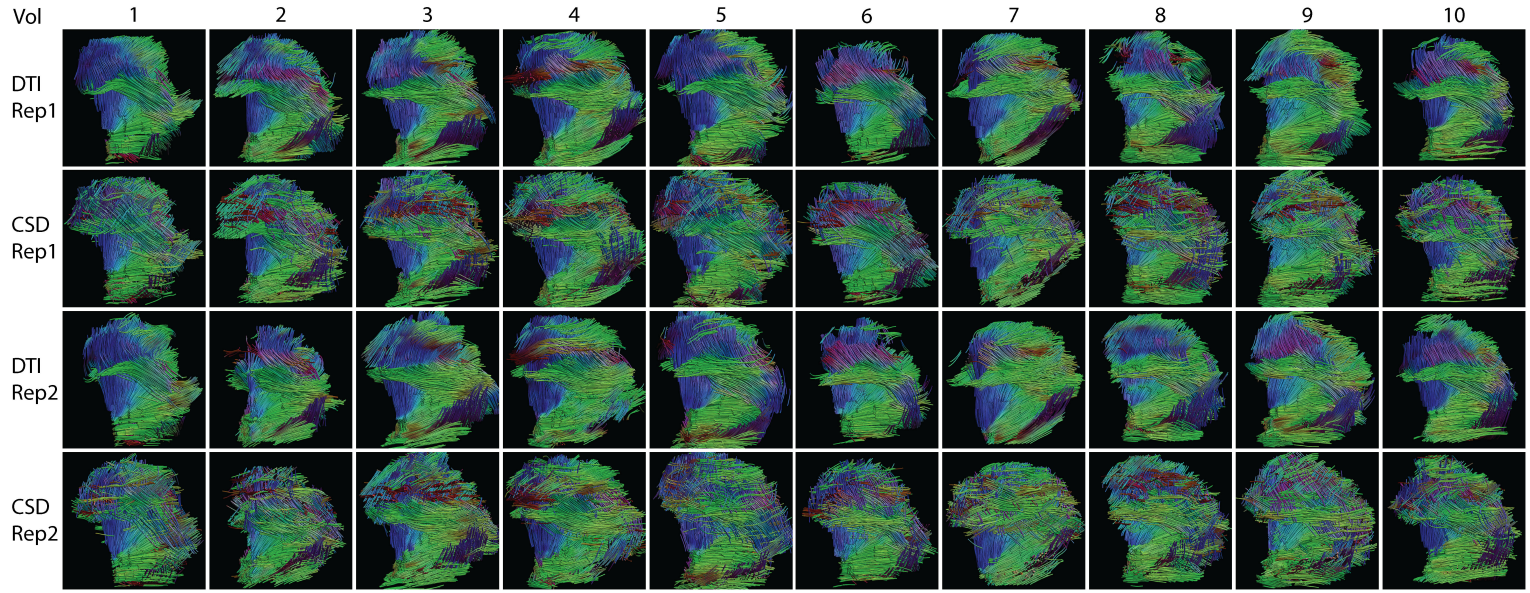


FIGURE 2.8 — Sagittal/oblique images of DTI and CSD tractography of the 10 volunteers (Vol) for the first (Rep1) and second (Rep2) acquisitions. These images were used for the qualitative grading by three expert reviewers. Similar to Fig. 5, transverse muscle fibers are better reconstructed in CSD tractography compared with DTI tractography. The SL is also represented better in CSD tractography; however, CSD tractography is more prone to spurious tracts than DTI tractography. Note that in some cases of DTI tractography, the tip of the tongue is missing (vol0, vol5, vol6, vol7, vol9).

Discussion

In this study we created a scan protocol and processing pipeline for human in vivo CSD tractography within a clinically acceptable scan time of 10 minutes. CSD allow more accurate reconstructions of tongue musculature in regions containing crossing fibers compared with DTI tractography in vivo and ex vivo. However, CSD tractography is more susceptible to erroneous fibers due to the presence of spurious peaks.

The ex vivo mammalian tongue has been a frequently used model to test higher-order diffusion models in the past,[26–28] arguably due to the better conservation of crossing fibers postmortem compared with brain tissue. Diffusion spectrum imaging (DSI) has been successfully used to resolve the crossing fibers of the transverse and vertical muscles, and the surrounding sheet of longitudinal muscles in the bovine[26] and murine[27] ex vivo tongue. In the latter example, Gaige et al were also able to relate DSI tractography and microscopy.[27] We were able to reconstruct a similar muscular architecture (containing transverse, vertical, and longitudinal fibers) in the bovine tongue as these previous DSI experiments; thus, we suppose that CSD is also able to correctly capture the complex tongue architecture ex vivo.

For human measurements, we could not find such a "gold standard". Therefore, we could not validate our results, except for the comparison with textbook anatomy. Although DSI has been validated ex vivo, the relatively high number of gradient directions and high maximum b-values require a long scan time; for example, ∼25 minutes for a single slice with a b-value of 8000 s/mm^2.[29] This excessively long scan time makes DSI unsuitable for a clinical setting. Another diffusion model, generalized q-space imaging (GQI), also has the ability to discern between crossing muscle fibers in the murine tongue[28] using a similar HARDI scheme and scan time as CSD. However, the typical array of transverse, vertical, and longitudinal fibers could only be detected in the murine model and not in the in vivo human case. This absence may be caused by B_0-inhomogeneities, which were not accounted for, or by the lower angular resolution of GQI compared with CSD, as shown in computer simulations.[30] Therefore, we deem CSD to be a reasonable compromise between accuracy in resolving crossing fibers and scan time.

Gaige et al previously used DTI to determine the human tongue muscular architecture in vivo.[9] Due to the inherent limitation of the tensor model, namely, only one fiber direction per voxel, crossing fibers from the transverse muscle fibers could not be detected. The vertical fibers appear slightly oblique, similar to our findings with in vivo DTI tractography. These similarities suggest that the acquired data are similar to previous work, but that the diffusion tensor is an insufficient model to describe tongue muscular architecture. However, the present imaging protocol requires a substantial increase in scan time compared with an imaging protocol designed only for DTI. Future studies need to prove whether the additional information gained by CSD are worth the additional scan time required.

DTI of patients following partial glossectomy has been performed;[31,32] however, the tongue musculature superior to the IL muscle was not reported. Due to the crossing nature of the muscle fibers and the influence of B_0-inhomogeneities, we suspect that these muscles could not be reconstructed. In another case, prior information from a manually segmented high-resolution anatomical image was used to fit a multitensor model to a DTI acquisition scheme.[33] Although this method allows for the reconstruction of the crossing transverse and vertical fibers in a limited scan time, this method is highly dependent on prior knowledge. This prior is created manually and might not always be correct, for example with variations between subjects.

In the present study the RF was determined by calculating the average FA and MD in a single fiber ROI of the GG muscle. Comparing the present reproducibility to the human forearm,[34] the wsCV of the FA appears to be similar, while the wsCV of the MD appears to be better for the ROI in the GG. Although the RF is determined in a reproducible manner by manual delineation in the GG, this RF might not be the most suitable RF for other tongue muscles due to differences in muscle fiber diameter and fiber type composition. Methods of automatically determining the RF have been developed for brain CSD.[35,36] These methods are designed to circumvent the fact that in the brain few voxels contain a single fiber direction, while these voxels are more prevalent in muscle diffusion imaging. Nonetheless, we recommend adapting such an automatic method to muscle CSD, because of the less user input required and possibly increased reproducibility of RF estimation due to variation between tongue muscles.

To minimize the number of spurious peaks caused by the more spherical RF due to the low FA in muscle CSD compared with brain CSD, the FOD peak threshold has to be chosen carefully. We chose a relatively low peak threshold so the number of false-negative peaks will be low. We assume that the false-positive peaks will have a limited effect on the final tractography results, because spurious peaks tend to have a large angular difference with surrounding voxels, and therefore have a higher chance of being ignored during tractography. Nevertheless, spurious tracts do appear in the final tractography results. To further reduce the number of spurious tracts, we suggest using a higher b-value, which results in a sharper RF taking into account that sufficient signal-to-noise ratio (SNR) is maintained,[36] or improving the SNR.[35]

The SNR may be increased by increasing the number of signal averages, which will lead to a longer scan time, and therefore a higher chance of motion artifacts. We optimized the SNR using our current hardware by using two flexible surface coils strapped to the cheeks of the volunteers instead of the standard neurovascular coil. However, custom-built receive coils have been shown to improve SNR,[37] which may reduce the number of spurious peaks, or which may increase spatial resolution to reduce the partial volume effect.

We used a high b-value of 700 s/mm^2 compared with other DTI studies of

muscle, but low compared with CSD studies of the brain. Our chosen b-value is a trade-off between the high b-value necessary for a high angular resolution for CSD,[30] and a short TE to obtain sufficient SNR in muscular tissue, which has a comparatively short T2 compared with gray and white matter. Because the diffusivity of brain[38] tissue is approximately half the magnitude of muscular tissue,12 the b-value used in this study (700 s/mm^2) compares approximately to a b-value of 1400 s/mm^2 in the brain, which is in the range of typically used b-values.

In the present study we were able to reconstruct most tongue muscles except for the smallest extrinsic muscle: the PG. Most probably, the PG muscle merges with the styloglossus muscle nearly immediately after insertion in the body of the tongue,[6] hindering the differentiation between each other. The PG muscle could also be missed due to the fact that tractography was constrained to the whole tongue masks. Additionally, the origins of the PG and SG muscles (the styloid process and soft palate, respectively) were not imaged to reduced scan time. As reported previously,[32] we also could not distinguish the IL from the SG muscles in vivo. Anatomical dissection has shown that the parallel fiber bundles of these two muscles are even difficult to distinguish on a microscopic level, especially near the apex of the tongue.[6,39] Finally, we had difficulties in tracking the SL. The relatively large voxel size causes a partial volume effect that may terminate SL fiber tracking prematurely. Also, normal breathing of the subjects, causing slight deviations of the base of tongue over time, which contributed to impairment of reconstruction of the SL muscle in the posterior tongue.

Three expert reviewers graded images for both DTI and CSD tractography in five categories. Additionally, a limited three-point scale was chosen, because the ability of discerning between more than three points comes with experience. Unfortunately, this limited scale also reduces the power of the statistical analyses. It was therefore not feasible to obtain meaningful qualitative results on questions individually with only 10 subjects.

During acquisition, B_0-inhomogeneity distortions were minimized by exclusion of volunteers with braces and dental splints, careful instruction of the healthy volunteers, and by correcting the DWIs using Topup. However, dental work involving metals remains an issue, especially regarding the clinical use of this scan protocol. Fortunately, orthodontic braces and wires are generally not present in the intended patient population for this scan protocol, which is generally older than 50 years. Most dental materials used in these patients, such as amalgam and zirconia, do not induce severe inhomogeneity artifacts.

The main limitation of our study is the lack of a gold standard to which ex vivo and in vivo muscle CSD can validated, although the presented tractography appears to agree with results previously reported in the literature. Anatomical dissection including microscopy[6,27] has been able to resolve crossing muscle fibers; however, this technique is time-consuming, cannot be performed in vivo, and issues arise with the reconstruction into a 3D volume. Another limitation to the

study is the participation of relatively young healthy volunteers without any dental prosthetics, which cause inhomogeneity artifacts and are often present in tongue cancer patients. Because these patients are generally in their sixth decade of life, and because the tongue composition changes with age due to muscular atrophy and lipomatosis,[40] care has to be taken to extrapolate the present results to the patient population. Finally, the scan protocol has not been validated for tongue cancer patients, where the diffusion signal may be affected by tumor infiltration and inflammation.

In conclusion, we have shown that the muscular architecture of the human tongue could be reconstructed in vivo in a clinically acceptable scan time. The present tractography results agreed with textbook anatomy. In the future, the proposed scan protocol may improve our understanding of how the complex muscular architecture of the tongue relates to tongue functionality after surgery for tongue cancer.

References

1. J. C. Sowder, R. B. Cannon, L. O. Buchmann, J. P. Hunt, Y. Hitchcock, S. Lloyd, K. F. Grossmann, and M. M. Monroe, "Treatment-related determinants of survival in early-stage (T1-2N0M0) oral cavity cancer: A population-based study," *Head & Neck*, vol. 39, no. 5, pp. 876–880, May 2017. DOI: 10.1002/hed.24679. arXiv: NIHMS150003.

2. G. Nicoletti, D. S. Soutar, M. S. Jackson, A. a. Wrench, G. Robertson, and C. Robertson, "Objective assessment of speech after surgical treatment for oral cancer: experience from 196 selected cases.," *Plastic and reconstructive surgery*, vol. 113, no. 1, pp. 114–125, 2004. DOI: 10.1097/01.PRS.0000095937.45812.84.

3. G. Nicoletti, D. S. Soutar, M. S. Jackson, A. a. Wrench, and G. Robertson, "Chewing and swallowing after surgical treatment for oral cancer: functional evaluation in 196 selected cases.," *Plastic and reconstructive surgery*, vol. 114, no. 2, pp. 329–338, 2004. DOI: 10.1097/01.PRS.0000131872.90767.50.

4. B. R. Pauloski, J. A. Logemann, L. A. Colangelo, A. W. Rademaker, F. M. McConnel, M. A. Heiser, S. Cardinale, D. Shedd, D. Stein, Q. Beery, E. Myers, J. Lewin, M. Haxer, and R. Esclamado, "Surgical variables affecting speech in treated patients with oral and oropharyngeal cancer.," *The Laryngoscope*, vol. 108, no. 6, pp. 908–916, 1998. DOI: 10.1097/00005537-199806000-00022.

5. 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, Apr. 1985. DOI: 10.1111/j.1096-3642.1985.tb01178.x.

6. H. Takemoto, "Morphological Analyses of the Human Tongue Musculature for Three-Dimensional Modeling," *Journal of Speech, Language, and Hearing Research*, vol. 44, no. 1, pp. 95–107, 2001. DOI: 10.1044/1092-4388(2001/009).

7. I. Sanders and L. Mu, "A three-dimensional atlas of human tongue muscles," *Anatomical Record*, vol. 296, no. 7, pp. 1102–1114, 2013. DOI: 10.1002/ar.22711. arXiv: NIHMS150003.

8. A. Kreeft, I. B. Tan, M. W. M. van den Brekel, F. J. Hilgers, and A. J. M. Balm, "The surgical dilemma of 'functional inoperability' in oral and oropharyngeal cancer: current consensus on operability with regard to functional results," *Clinical Otolaryngology*, vol. 34, no. 2, pp. 140–146, 2009. DOI: 10.1111/j.1749-4486.2009.01884.x.

9. T. a. Gaige, T. Benner, R. Wang, V. J. Wedeen, and R. J. Gilbert, "Three dimensional myoarchitecture of the human tongue determined in vivo by diffusion tensor imaging with tractography," *Journal of Magnetic Resonance Imaging*, vol. 26, no. 3, pp. 654–661, 2007. DOI: 10.1002/jmri.21022.

10. G. G. G. Cleveland, D. C. C. Chang, C. F. F. Hazelwood, and H. E. E. Rorschach, "Nuclear Magnetic Resonance Measurement of Skeletal Muscle," *Biophysical journal*, vol. 16, pp. 1043–1053, 1976. DOI: 10.1016/S0006-3495(76)85754-2.

11. C. C. Van Donkelaar, L. J. Kretzers, P. H. Bovendeerd, L. M. Lataster, K. Nicolay, J. D. Janssen, M. R. Drost, and V. Leeuwen, "Diffusion tensor imaging in biomechanical studies of skeletal muscle function," *J Anat*, vol. 194, no. 1, pp. 79–88, 1999.

12. M. Froeling, A. J. Nederveen, D. F. R. Heijtel, A. Lataster, C. Bos, K. Nicolay, M. Maas, M. R. Drost, and G. J. Strijkers, "Diffusion-tensor MRI reveals the complex muscle architecture of the human forearm.," *Journal of magnetic resonance imaging : JMRI*, vol. 36, no. 1, pp. 237–248, 2012. DOI: 10.1002/jmri.23608.

13. J.-D. Tournier, F. Calamante, and A. Connelly, "Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution.," *NeuroImage*, vol. 35, pp. 1459–1472, May 2007. DOI: 10.1016/j.neuroimage.2007.02.016.

14. J. D. Tournier, F. Calamante, D. G. Gadian, and A. Connelly, "Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution," *NeuroImage*, vol. 23, no. 3, pp. 1176–1185, 2004. DOI: 10.1016/j.neuroimage.2004.07.037.

15. D. S. Tuch, T. G. Reese, M. R. Wiegell, N. Makris, J. W. Belliveau, and J. Van Wedeen, "High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity," *Magnetic Resonance in Medicine*, vol. 48, no. 4, pp. 577–582, 2002. DOI: 10.1002/mrm.10268.

16. J. L. R. Andersson, S. Skare, and J. Ashburner, "How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging.," *NeuroImage*, vol. 20, no. 2, pp. 870–888, Oct. 2003. DOI: 10.1016/S1053-8119(03)00336-7.

17. D. K. Jones, M. a. Horsfield, and a. Simmons, "Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging.," *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, vol. 42, no. 3, pp. 515–525, Sep. 1999.

18. J. L. R. Andersson and S. N. Sotiropoulos, "An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging," *NeuroImage*, vol. 125, pp. 1063–1078, 2016. DOI: 10.1016/j.neuroimage.2015.10.019.

19. P. a. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability.," *NeuroImage*, vol. 31, no. 3, pp. 1116–28, Jul. 2006. DOI: 10.1016/j.neuroimage.2006.01.015. [Online]. Available: www.itksnap.org.

20. L.-C. Chang, D. K. Jones, and C. Pierpaoli, "RESTORE: robust estimation of tensors by outlier rejection.," *Magnetic Resonance in Medicine*, vol. 53, pp. 1088–1095, 2005. DOI: 10.1002/mrm.20426.

21. A. Leemans and B. Jeurissen, "ExploreDTI: a graphical toolbox for processing, analyzing, and visualizing diffusion MR data," *Proceedings of the International Society of Magnetic Resonance in Medicine*, vol. 245, p. 3300, 2009.

22. M. Froeling, A. J. Nederveen, K. Nicolay, and G. J. Strijkers, "DTI of human skeletal muscle: the effects of diffusion encoding parameters, signal-to-noise ratio and T2 on tensor indices and fiber tracts," *NMR in Biomedicine*, vol. 26, no. 11, pp. 1339–1352, 2013. DOI: 10.1002/nbm.2959.

23. R. E. Smith, J. D. Tournier, F. Calamante, and A. Connelly, "SIFT: Spherical-deconvolution informed filtering of tractograms," *NeuroImage*, vol. 67, pp. 298–312, 2013. DOI: 10.1016/j.neuroimage.2012.11.049.

24. R. H. B. Christensen. (2018). "ordinal - Regression Models for Ordinal Data," [Online]. Available: http://www.cran.r-project.org/package=ordinal/.

25. H. Gray, "Muscles of the Tongue," in *Anatomy of the Human Body*, W. H. Lewis, Ed., 20th, Philadelphia: Lea & Febiger, 1918, pp. 1128–1131.

26. R. J. Gilbert, V. J. Wedeen, L. H. Magnusson, T. Benner, R. Wang, G. Dai, V. J. Napadow, and K. K. Roche, "Three-dimensional myoarchitecture of the bovine tongue demonstrated by diffusion spectrum magnetic resonance imaging with tractography.," *The anatomical record. Part A, Discoveries in molecular, cellular, and evolutionary biology*, vol. 288, no. 11, pp. 1173–1182, 2006. DOI: 10.1002/ar.a.20387.

27. T. A. Gaige, H. S. Kwon, G. Dai, V. C. Cabral, R. Wang, Y. S. Nam, B. P. Engelward, V. J. Wedeen, P. T. C. So, and R. J. Gilbert, "Multiscale structural analysis of mouse lingual myoarchitecture employing diffusion spectrum magnetic resonance imaging and multiphoton microscopy," *Journal of Biomedical Optics*, vol. 13, no. 6, p. 064 005, 2008. DOI: 10.1117/1.3046724.

28. E. N. Taylor, M. P. Hoffman, G. E. 2. Aninwene, and R. J. Gilbert, "Patterns of Intersecting Fiber Arrays Revealed in Whole Muscle with Generalized Q-Space Imaging," *Biophysical Journal*, vol. 108, no. 11, pp. 2740–2749, 2015. DOI: 10.1016/j.bpj.2015.03.061.

29. V. J. Wedeen, P. Hagmann, W.-Y. I. Tseng, T. G. Reese, and R. M. Weisskoff, "Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging," *Magnetic Resonance in Medicine*, vol. 54, no. 6, pp. 1377–1386, 2005. DOI: 10.1002/mrm.20642.

30. J. D. Tournier, C. H. Yeh, F. Calamante, K. H. Cho, A. Connelly, and C. P. Lin, "Resolving crossing fibres using constrained spherical deconvolution: Validation using diffusion-weighted imaging phantom data," *NeuroImage*, vol. 42, no. 2, pp. 617–625, 2008. DOI: 10.1016/j.neuroimage.2008.05.002.

31. E. Z. Murano, H. Shinagawa, J. Zhuo, R. P. Gullapalli, R. A. Ord, J. L. Prince, and M. Stone, "Application of diffusion tensor imaging after glossectomy," *Otolaryngol Head Neck Surg*, vol. 143, no. 2, pp. 304–306, 2010. DOI: 10.1017/CBO9781107415324.004. arXiv: arXiv:1011.1669v3.

32. H. Shinagawa, E. Z. Murano, J. Zhuo, B. Landman, R. P. Gullapalli, J. L. Prince, and M. Stone, "Tongue muscle fiber tracking during rest and tongue protrusion with oral appliances: A preliminary study with diffusion tensor imaging," *Acoustical Science and Technology*, vol. 29, no. 4, pp. 291–294, 2008. DOI: 10.1250/ast.29.291.

33. C. Ye, E. Murano, M. Stone, and J. L. Prince, "A Bayesian approach to distinguishing interdigitated tongue muscles from limited diffusion magnetic resonance imaging," *Computerized Medical Imaging and Graphics*, vol. 45, pp. 63–74, 2015. DOI: 10.1016/j.compmedimag.2015.07.005.

34. M. Froeling, J. Oudeman, S. van den Berg, K. Nicolay, M. Maas, G. J. Strijkers, M. R. Drost, and A. J. Nederveen, "Reproducibility of diffusion tensor imaging in human forearm muscles at 3.0 T in a clinical setting.," *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, vol. 64, no. 4, pp. 1182–1190, 2010. DOI: 10.1002/mrm.22477.

35. C. M. W. Tax, B. Jeurissen, S. B. Vos, M. A. Viergever, and A. Leemans, "Recursive calibration of the fiber response function for spherical deconvolution of diffusion MRI data," *NeuroImage*, vol. 86, pp. 67–80, 2014. DOI: 10.1016/j.neuroimage.2013.07.067.

36. J. D. Tournier, F. Calamante, and A. Connelly, "Determination of the appropriate b value and number of gradient directions for high-angular-resolution diffusion-weighted imaging," *NMR in Biomedicine*, vol. 26, no. 12, pp. 1775–1786, 2013. DOI: 10.1002/nbm.3017.

37. S. G. Lingala, Y. Zhu, Y.-c. Kim, A. Toutios, S. Narayanan, and K. S. Nayak, "A fast and flexible MRI system for the study of dynamic vocal tract shaping," *Magnetic Resonance in Medicine*, vol. 00, n/a–n/a, 2016. DOI: 10.1002/mrm.26090.

38. A. L. Alexander, J. E. Lee, M. Lazar, and A. S. Field, "Diffusion Tensor Imaging of the Brain," *Neurotherapeutics : the journal of the American Society for Experimental NeuroTherapeutics*, vol. 4, no. 3, p. 316, 2007.

39. H. Takemoto, "Morphological analyses and 3d modeling of the tongue musculature of the chimpanzee {Pan troglodytes)}," *American Journal of Primatology*, vol. 70, no. 10, pp. 966–975, 2008. DOI: 10.1002/ajp.20589.

40. R. Bässler, "Histopathology of different types of atrophy of the human tongue," *Pathology Research and Practice*, vol. 182, no. 1, pp. 87–97, 1987. DOI: 10.1016/S0344-0338(87)80147-4.

3

Chapter 3

Diffusion MRI-based tongue muscle fibre atlases of healthy volunteers and tongue-cancer patients

Luuk Voskuilen, Valentina Mazzoli, Alfons J.M. Balm, Gustav J. Strijkers, Ludi E. Smeele, Aart J. Nederveen

Manuscript in preparation

Abstract

Background: In cancer of the tongue, changes in the diffusion properties of the tongue's musculature may reflect the extent of the cancer and assess the tissue damage caused by it. Diffusion-weighted MRI, in combination with constrained spherical deconvolution (CSD), is able to resolve this musculature, even in the presence of crossing muscle fibres, which diffusion tensor imaging (DTI) cannot. However, the tongue musculature reconstructed by CSD can currently not be examined automatically and quantitatively.
Purpose: To create tongue muscle atlases to automatically segment the musculature and objectively compare diffusion metrics between study populations.
Study type: Prospective cohort study.
Populations: Ten young healthy volunteers (mean age: 25.5 years), ten elderly healthy volunteers (61.4 years), and eight patients with cancer of the tongue scheduled for a partial glossectomy (57.6 years).
Sequence: High-angular resolution diffusion imaging at 3T.
Assessment: CSD-based population averages were generated for the three populations. The population average of the young population was segmented to create an atlas for tongue muscles. Within these muscles, two DTI metrics: fractional anisotropy (FA) and mean diffusivity (MD), and one CSD metric: apparent fibre density (AFD) were calculated.
Statistical tests: One-way ANOVA with post-hoc two-sample t-tests on the diffusion metrics.
Results: The DTI-based metrics FA and MD were not able to identify differences between the three populations. CSD-based metric AFD could identify differences between the populations.
Data conclusion: Using population averaging, AFD differences could be identified between populations, which may be explained by microstructural changes in the muscle fibres of the tongue. However, further histological study is necessary to study the cause and clinical significance of these changes due to aging or in the presence of cancer.

Introduction

The tongue is the largest muscle group in the oral cavity and pharynx. It is vital to basic human functions such as speech, swallowing, and mastication. These functions may become impaired as carcinomas invade, cause inflammation, and destroy muscular tissue of the tongue. In addition to the direct effect of tumour growth, the functionality of the tongue will be further reduced by the preferred treatment, surgical removal, as this includes resection of the tumour including a margin of normal tissue. To assist in choosing the optimal treatment strategy that leads to a complete removal of the tumour, while minimizing the post-operative loss of functionality, it is important to know which tongue muscles will be affected by the tumour or its resection. This knowledge might also be used to improve the rehabilitation of these patients, who often suffer from impaired tongue functionality such as compromised speech and swallowing.

The normal muscular architecture of the tongue has been previously characterised by anatomical dissection.[1] The tongue is classified as a muscular hydrostat, which means that it relies on hydrostatic pressure for support rather than skeletal bones.[2] The tongue is comprised of eight muscles of which four are extrinsic muscles that are connected to bony structures (genioglossus; hyoglossus; styloglossus; and palatoglossus), and four are intrinsic muscles without any connection to bones at all (transversus; verticalis; and longitudinalis superior and inferior). A successful prediction surgical outcome depends on the knowledge of the anatomy of these individual muscles and how these muscles will be changed by the surgery. However, determining these changes in the tongue musculature of patients non-invasively by current imaging methods is challenging.

Diffusion tensor imaging (DTI) is an non-invasive MRI technique that is able to reconstruct the muscular architecture of a wide variety of muscles ranging from relatively simple muscular structures such as the forearm muscles[3] to complex muscular organs such as the heart[4] and the tongue[5]. By exploiting self-diffusion of water molecules, and their anisotropic behaviour in muscle, DTI allows the reconstruction of muscle fibres at the voxel level, which in turn can be used to create 3D models of the muscular architecture. Besides 3D visualization, DTI can characterise the microstructure of the tongue musculature by metrics such as the mean diffusivity (MD) and the fractional anisotropy (FA).

However, DTI only assumes one single muscle fibre direction in every voxel, while in large parts of the tongue, muscle fibres interlace at a microscopic level[6], meaning that multiple crossing muscle fibre populations are present within one voxel. Constrained Spherical Deconvolution (CSD) has been developed in order to deal with the "crossing fibres" problem and correctly reconstructs multiple fibre population within a voxel.[7] In short, in CSD a response function (RF) is defined that corresponds to the diffusion signal of a single fibre population. By deconvolving the measured diffusion signal in each voxel with this response function, a

fibre orientation distribution (FOD) is obtained. Usually, these calculations are performed in spherical harmonics to reduce their complexity.

CSD has been successfully applied to the in vivo tongue, and was proven to provide more anatomically accurate reconstructions of the muscle architecture than DTI reconstructions.[8] Despite these encouraging results, its clinical implementation is currently hampered by lack of automatic and objective ways to segment the individual muscles. If these segmentations would be known, changes in muscle architecture could be related, via the segmented muscles, to changes in tongue functionality. Therefore, an automatic and objective method of comparing CSD-based tongue reconstructions is required.

In the field of neuroimaging, study populations are often compared using an atlas of the 'average' brain, such as the MNI atlas.[9] These brain atlases are usually created by combining the imaging data of several brains that were annotated or segmented manually resulting in a template that contains a segmentation layer and an imaging layer. Once the atlas is created, every new image can automatically be segmented by registering the atlas to that image and applying that same transformation to the segmentation layer. These segmentations can then be utilized to perform volumetric and morphological measurements, and to extract quantitative MRI metrics, such as the FA and MD, for each individual muscle.

In a similar way as how an atlas is created, the images of a population can be combined into a population average (PA), a generalisation of the atlas that may lack the segmentation layer. As the names of these PAs imply, they aggregate the information of a population, and as such define morphology and quantitative MRI metrics of the 'average' tongue within that population. Once PAs containing quantitative MRI metrics are defined, they can be used to compare different populations[10], e.g. healthy volunteers and tongue cancer patients. A tongue PA or atlas for automatic segmentation has previously been created using high resolution scalar images[11], but this method fails to correctly identify all the different muscle group of the tongue. In addition, given the ability of CSD to provide microstructural information about the separate muscles, a tongue atlas based on the CSD fibre orientation distribution might be more sensitive in detecting microstructural changes in the tongue.

Based on these premises, our goal was to implement a method for generating CSD-based PAs of the tongue, and using these PAs to automatically segment the musculature and to objectively compare diffusion metrics between different populations. In order to achieve that, we first determined which type of image registration should be used to generate PAs. Subsequently, we used CSD to reconstruct the 3D architecture of the tongue in the healthy PA and created an atlas of the in vivo tongue. Finally, after generating a PA for each of our three data sets (young volunteers, elderly volunteers, and tongue cancer patients), we compared the three populations in a CSD-based diffusion metric for every muscle individually.

Methods

Participants

Twenty-eight ($n = 28$) subjects took part in the study. Before scanning, all volunteers provided written informed consent. This study was approved by the institutional medical ethics review committees at our institutions.

Ten ($n = 10$) healthy young volunteers (mean age 25.5 years; range 21–34 years, four women) with no orthodontic appliances were scanned (*young group*). The scan protocol was repeated within one hour after repositioning the subject. These images were previously used in the publication on the feasibility of CSD for MRI of the tongue.[8]

Eight ($n = 8$) tongue carcinoma patients (mean age 57.6 years; age range 33–71 years; two women) were included (*patient group*). All subjects had a T1–3 carcinoma of the tongue scheduled for a partial glossectomy. Subjects were prospectively excluded if they had braces or any contraindications to MRI, and retrospectively if severe motion artefacts were present in the images.

Ten ($n = 10$) older healthy volunteers (mean age 61.4 years; range 56–71 years; three women) were scanned (*elderly group*). The older volunteers were age matched to the patient group ($P = 0.383$). Metal orthodontic appliances (excluding braces) were allowed.

MRI acquisition

Participants from the 3 groups were scanned using the same acquisition protocol with a 3T scanner (Philips Healthcare, Best, Netherlands). Before the scan, participants were instructed to place their tongue against the palate to minimise tongue motion and reduce the amount of air in the oral cavity. The previously described diffusion-weighted imaging (DWI) protocol[8] was used with the following scan parameters: SE single-shot EPI; shot length: 25; TE/TR: 60/3400 ms; two repetitions with opposing phase-encoding directions (RL & LR); NSA=1; SPIR and SSGR fat suppression; FOV: 192 x 156 x 84 mm^3; voxel size: 3 x 3 x 3 mm^3; b-value: 700 s/mm^2 along 64 directions evenly spaced over a hemisphere and optimised for gradient load; after every eight volumes with a b-value of 700 s/mm^2, a b0-volume with a b-value of 0 s/mm^2 was acquired, totalling eight b0-volumes; total scan time: 10 minutes.

DWI processing

All datasets were denoised according to the method proposed by Veraart et al.[12], and corrected for B_0-inhomogeneity, eddy current distortions, and rigid motion in

FSL[13]. To reduce computation time, masks of the whole tongue were created by manual delineation (figure 3.1a).[8]

In MRtrix3[14], a response function (RF) was estimated for each participant (figure 3.1b) according to Tournier et al.[15], which determines the RF from the voxels with the highest FA. Within each data set, these response functions were averaged. Using a b-spline interpolation, the corrected DWIs were resampled to a resolution of 1.5 x 1.5 x 1.5 mm isotropic. Using CSD, the resampled and corrected DWIs were deconvolved with the averaged response functions to obtain the FOD volumes to a maximal spherical harmonic degree (l_{max}) of 8.[16]

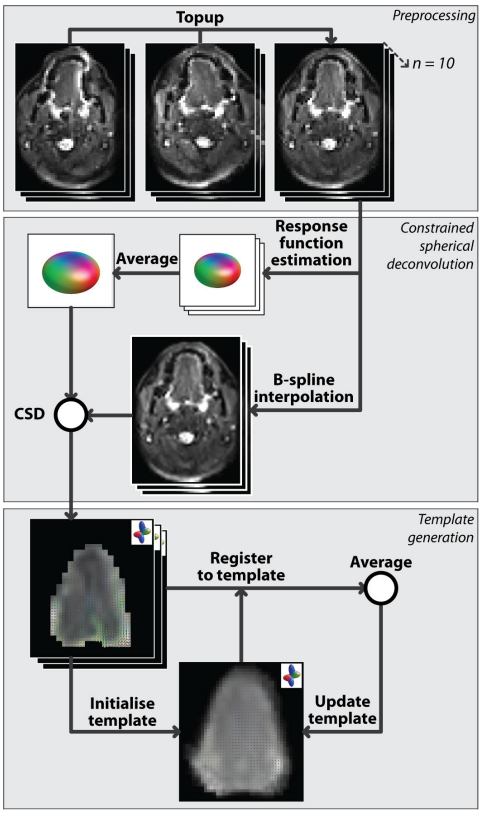


FIGURE 3.1 — An overview of the processing pipeline for generating an unbiased FOD-based population average (PA) from uncorrected diffusion-weighted images (DWIs). The stacks of images indicate a data sets of multiple individuals, an FOD in the upper right corner of an image indicates that the underlying data is an FOD volume. In the first block (*preprocessing*), the DWIs were corrected for B_0-inhomogeneity, eddy current distortions, and rigid motion. In the second block (*constrained spherical deconvolution*), the response functions corresponding to a single fibre population were estimated and then averaged, and DWIs were scaled to a higher resolution. Then, the FOD volumes were calculated using CSD. In the final block (*template generation*), the population average generation was initialised by averaging the FOD volumes aligned on their centres of gravity. Subsequently, the FOD volumes were registered to the current template and averaged to create the template for the next iteration. This last step was repeated for six iterations with rigid registrations, six with affine registrations, and sixteen with non-linear registrations.

PA generation

All PAs were generated by iteratively aligning and averaging all volumes within a subject group[17], in order to avoid biases toward individual subjects (figure 3.1c). An initial template was generated by aligning all FOD volumes in a population based on their centres of gravity and averaging the FOD volumes on a voxel level. Subsequently, using an CSD-based registration with the appropriate FOD reorientation[18], the FOD volumes were registered to the current template, which was

updated by the average of registered FOD volumes. This registration and averaging step was repeated 28 times, while the leniency of the registration step was increased with every step, i.e. lowering the scale, increasing the l_{max}, and changing the registration type from rigid to affine to a non-linear symmetric registration. The final template was considered the PA for the subject group.

To assess the reproducibility of this method for the generation of PAs, one PA was generated for the first repetition and one for the second repetition of data set young. To assess the error in the PA generation, we calculated the average L_2-error (L2) and the average angular correlation coefficient (ACC)[18] between the PA and each individual's FOD volume transformed to the PA space. By also generating PAs for the young group using a conventional registration based on the image intensity of the corrected DWIs instead of the fibre orientation, we evaluated whether a fibre-orientation-based registration is beneficial to PA generation. For the PAs based on registering DWIs, L2 and ACC were also calculated. Using two-tailed t-tests and a significance level of 5%, we tested for differences in the error between the first and second repetition and between the DWI-based and CSD-based registration.

Fibre tracking and segmentation

Whole tongue fibre tractography[19] was performed for one individual in the young group, and for the population average of young (PAy). Fibre tractography was performed in MRtrix3 with the following parameters: step size: 3 mm; angular threshold 15; minimal length: 10 mm; maximal length: 100 mm; FOD lobe cut-off value: 0.05; and 10,000 seed point randomly placed within the mask of the tongue. In TrackVis[20], the tractography of PAy was subsequently segmented into the four intrinsic tongue muscles (transverse, vertical, superior longitudinal, and inferior longitudinal muscle), two extrinsic tongue muscles (genioglossus and hyoglossus) and three muscles of the floor of the mouth (geniohyoid, mylohyoid, and the anterior belly of the digastric muscle).

Comparison of data sets

To compare the three groups with one another, we calculated the following three quantitative diffusion metrics for each muscle: FA, MD, and AFD, which is defined as the integral under an FOD lobe.[21] As multiple fibre populations and therefore multiple AFDs may be present per voxel, a CSD-based PA is necessary for this metric. The population average of the elderly group (PAe) and that of the patient group (PAp) were subsequently registered to the young group (PAy) using a non-linear CSD-based registration. A single segmentation could thus be applied to all three groups.

FA and MD maps were computed for each participant and morphed to PAy space using the CSD-based registration. For each tongue muscle, FA and MD were averaged over the voxels in which the muscle was located. Similar as for the FA and MD maps, the FOD volumes were transformed to the PAy template for each participant. In each voxel, the FOD was divided into fixels, which contain an orientation vector of each FOD lobe and the corresponding integral, the AFD.[22] Subsequently, we assigned the fixels to a tongue muscle if a fibre track was present in their voxel and if the difference in orientation between the fixel and the track was smaller than 30°. Over each muscle tract, the AFD was averaged. Finally, the differences between the three populations in FA, MD and AFD were evaluated using one-way ANOVAs using a significance level of 0.05. In the case of significantly differences, post-hoc two-tailed t-tests were performed.

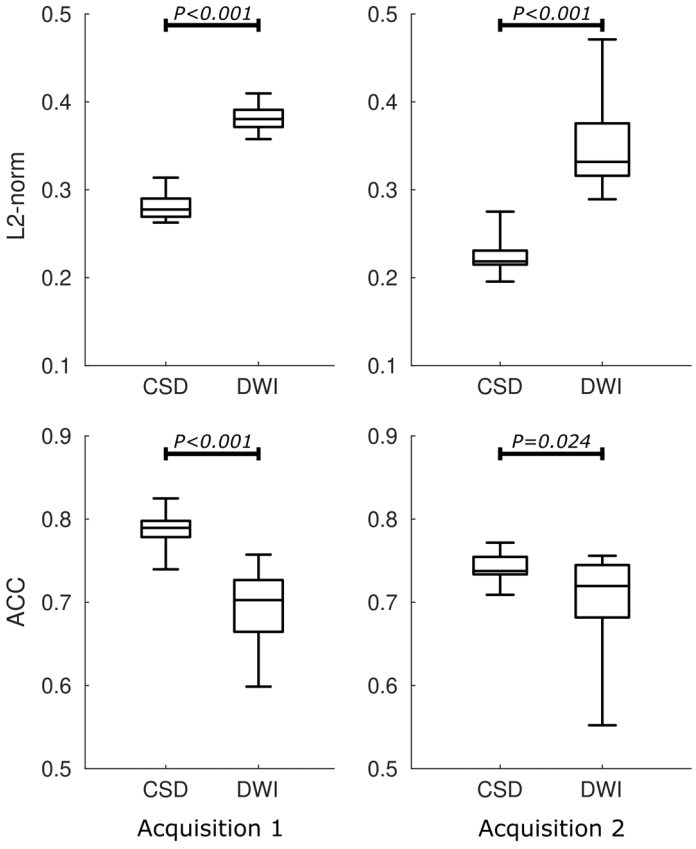


FIGURE 3.2 — Box plots of the L2-norm and the angular correlation coefficient (ACC) between the population average (PA) and the volumes from which it was generated. For each of the two acquisitions, two PAs were generated using either the CSD-based registration or the DWI-based. The median, 25^{th} and 75^{th} percentile are displayed as a box, and the minimum and maximum are displayed as the whiskers. The L2-norm and ACC were both significantly poorer for all PAs based on the DWI registration than for those based on the FOD registration.

Results

Population averages (PAs) could be successfully generated for all three groups, i.e. the younger population (both first and second repetition), the elderly population, and the patients, even though the FOD maps of the patients contained some artefacts (ESM 1). For the image-intensity-driven registration (figure 3.2) the L2-norm was significantly higher (first repetition $P < 0.001$, second repetition $P < 0.001$) and the ACC significantly lower (first repetition $P < 0.001$, second repetition $P < 0.024$) as compared to FOD-based registration. This indicates worse repeatability for image-intensity-driven registration than for the FOD-based registration.

In figure 3.3, the whole-tongue tractography of the PAy is displayed alongside the tractography of a young individual and an anatomical reference. As the PA is an average and therefore inherently less noisy, the tractography of the atlas appears to contain fewer spurious tracts upon visual inspection compared to the tractography of the single young subject. In addition, muscles that are inherently difficult to track such as the transverse and superior longitudinal muscles, may be better appreciated visually. The tractography of PAy was segmented in the individual muscles of the tongue and the floor of the mouth mentioned in the methods section.

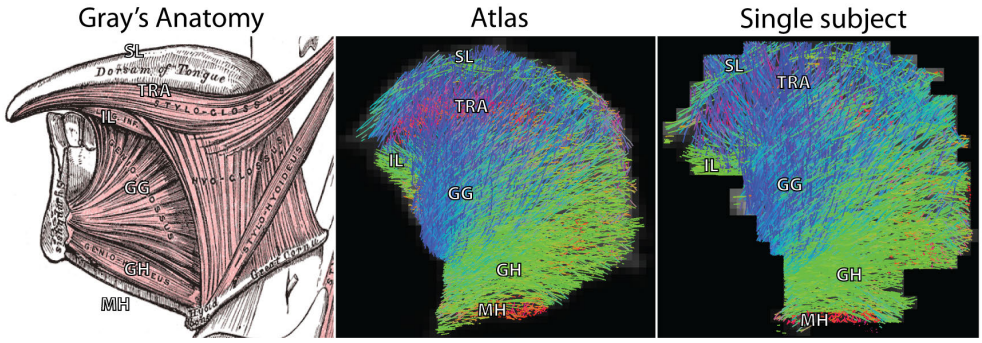


FIGURE 3.3 — Midsagittal view of the CSD-based tractography of the atlas or annotated population average of the young population (PAy), and one individual of that dataset, compared to an image from Gray's Anatomy. Fibres are colour-coded based on their direction (red: RL, green AP, blue IS). The following tongue muscles can be observed: the superior longitudinal (SL, green), transverse (TRA, red), inferior longitudinal (IL, green), genioglossus (GG, blue/green), geniohyoid (GH, green), and mylohyoid (MH, red) muscles. Visually, the tractography of the atlas compared to that the single subject is overall similar except for the following two cases: The SL muscle can be tracked along the entire dorsum of tongue in the population average, compared to a premature termination approximately halfway through its course in the individual. The TRA muscle can only be tracked sparsely in the individual compared to more densely tracking in the population average.

By averaging the three metrics (FA, MD, and AFD) over the segmented muscles, the three populations (young, elderly, and patients) were compared in the space of PAy. For the two DTI metrics FA and MD, no significant differences were detected between the young, elderly, and patient populations (table 3.1 and 3.2). A significant decrease in AFD was detected in five muscles between the young and elderly, and in nine muscles between the young and patient population (table 3.3). Compared to the elderly population, the AFD was significantly lower in four muscles of the patients (table 3.3). Overall, the AFD decreased in the elderly compared to the young subjects, and decreased even further in patients.

In figure 3.4, an example is given of the effect a tumour has on the AFD of the surrounding musculature. Although an overall decrease in AFD was detected in the patients compared to either the young or elderly data sets (table 3.3), the AFD was locally increased near the tumour.

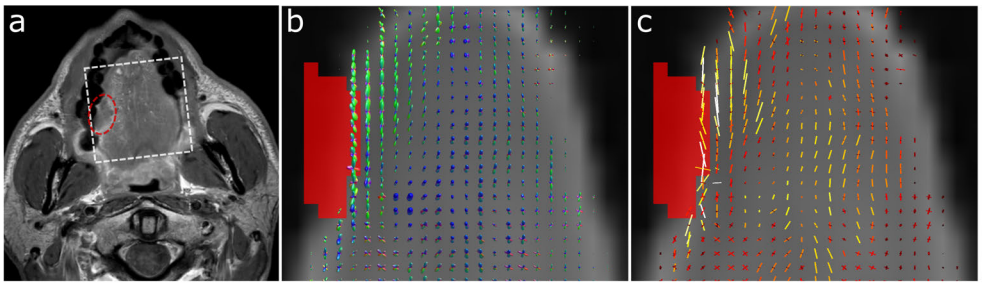


FIGURE 3.4 — An example of the change in AFD in a patient with tongue cancer. A tumour on the right side of the tongue can be identified on a contrast-enhanced T1-weighted scan. The tumour is highlighted by a red ellipse in a, and by a red region-of-interest in b and c. Using CSD as the diffusion model, the fibre-orientation distribution (FOD) in each voxel of the tongue was calculated (b), and the FOD volume was transformed to the space of the population average of the young (PAy). The approximate location of this transformed FOD volume is displayed by the white box in a. The FODs were segmented into separate FOD lobes (or fixels), visualised as vectors with their colour and magnitude proportional to the apparent fibre density (AFD) of each lobe (c). Comparing the AFD of the right side of the tongue that contains the tumour, to that of the left healthy side, the AFD of the tongue musculature around the tumour appeared to be increased.

TABLE 3.1 — Mean and standard deviation (std) of the fractional anisotropy (FA) in the population average of the three data sets: healthy young (PAy), healthy elderly (PAe), and patients (PAp). Using t-tests, P-values were calculated between the following segmented muscles (and where applicable the left-sided (l) and right-sided (r) muscle): the anterior belly of the digastric (DG), genioglossus (GG), geniohyoid (GH), hyoglossus (HG), inferior longitudinal (IL), mylohyoid (MH), superior longitudinal (SL), transverse (TRA), and vertical (VER) muscle.

Muscle	Young		Elderly		Patients		P-value
	Mean	*Std*	*Mean*	*Std*	*Mean*	*Std*	
DGl	0.196	0.018	0.190	0.021	0.192	0.017	0.806
DGr	0.200	0.015	0.189	0.009	0.187	0.026	0.189
GG	0.215	0.009	0.207	0.018	0.200	0.016	0.814
GH	0.208	0.008	0.198	0.016	0.191	0.023	0.089
HGl	0.184	0.008	0.182	0.024	0.170	0.016	0.111
HGr	0.180	0.008	0.177	0.020	0.174	0.009	0.197
ILl	0.194	0.009	0.186	0.031	0.175	0.014	0.654
ILr	0.190	0.012	0.183	0.020	0.181	0.016	0.190
MH	0.198	0.006	0.199	0.014	0.194	0.025	0.427
SL	0.153	0.006	0.154	0.022	0.145	0.011	0.735
TRA	0.173	0.005	0.178	0.019	0.174	0.015	0.191
VER	0.175	0.008	0.177	0.022	0.165	0.009	0.401

TABLE 3.2 — Mean and standard deviation (std) of the mean diffusivity (MD) in the population average of the three data sets: healthy young (PAy), healthy elderly (PAe), and patients (PAp). Using t-tests, P-values were calculated between the following segmented muscles (and where applicable the left-sided (l) and right-sided (r) muscle): the anterior belly of the digastric (DG), genioglossus (GG), geniohyoid (GH), hyoglossus (HG), inferior longitudinal (IL), mylohyoid (MH), superior longitudinal (SL), transverse (TRA), and vertical (VER) muscle.

Muscle	Young		Elderly		Patients		P-value
	Mean	*Std*	*Mean*	*Std*	*Mean*	*Std*	
DGl	1.43	0.20	1.38	0.14	1.36	0.13	0.661
DGr	1.42	0.21	1.36	0.13	1.35	0.16	0.680
GG	1.46	0.05	1.52	0.11	1.50	0.11	0.479
GH	1.51	0.08	1.63	0.14	1.59	0.15	0.124
HGl	1.49	0.05	1.57	0.10	1.54	0.12	0.370
HGr	1.50	0.08	1.50	0.14	1.52	0.12	0.169
ILl	1.43	0.06	1.51	0.11	1.46	0.12	0.936
ILr	1.50	0.09	1.47	0.17	1.48	0.11	0.147
MH	1.47	0.08	1.53	0.13	1.49	0.14	0.863
SL	1.34	0.06	1.25	0.12	1.33	0.14	0.916
TRA	1.38	0.05	1.39	0.11	1.40	0.12	0.539
VER	1.41	0.07	1.36	0.10	1.38	0.13	0.131

TABLE 3.3 — Mean and standard deviation of the apparent fibre density (AFD) in the population average of the three data sets: healthy young (PAy), healthy elderly (PAe), and patients (PAp). Using one-way ANOVAs, P-values were calculated between the following segmented muscles (and where applicable the left-sided (l) and right-sided (r) muscle): the anterior belly of the digastric (DG), genioglossus (GG), geniohyoid (GH), hyoglossus (HG), inferior longitudinal (IL), mylohyoid (MH), superior longitudinal (SL), transverse (TRA), and vertical (VER) muscle. Significantly different means are indicated by a bold P-value, for which post-hoc t-tests were performed. The following letters in superscript indicate a significant difference between two groups: a, between young and elderly; b, between young and patients; and c, between elderly and patients.

Muscle	Young		Elderly		Patients		P-value
	Mean	*Std*	*Mean*	*Std*	*Mean*	*Std*	
DGl	0.504	0.065	0.542	0.060	0.456	0.035	**0.012**^{c}
DGr	0.511	0.057	0.510	0.068	0.430	0.020	**0.006**^{b,c}
GG	0.634	0.064	0.615	0.078	0.547	0.090	**0.007**^{b}
GH	0.575	0.048	0.550	0.050	0.491	0.058	**0.008**^{b,c}
HGl	0.592	0.073	0.523	0.068	0.468	0.091	0.065
HGr	0.558	0.068	0.476	0.099	0.457	0.109	**0.008**^{a,b}
ILl	0.733	0.082	0.634	0.112	0.542	0.101	0.063
ILr	0.636	0.087	0.545	0.101	0.525	0.126	**0.002**^{a,b}
MH	0.438	0.043	0.424	0.063	0.361	0.034	0.068
SL	0.558	0.066	0.483	0.078	0.469	0.118	**0.036**^{a}
TRA	0.446	0.041	0.441	0.077	0.371	0.066	0.634
VER	0.483	0.051	0.496	0.061	0.457	0.137	0.079

Discussion

We were able to successfully generate unbiased population averages (PAs) based on the muscular architecture of the tongue for three different groups: young and elderly control subjects and a group of patients suffering for tongue cancer.

Compared to PAs generation with the regular image-intensity or DWI-based registration, we showed that the error in PAs generation using an FOD-based registration was significantly lower, as demonstrated by the lower L2-norm and the higher ACC. Based on a manual segmentation of the population average of the young population (PAy), tongue muscles could be automatically segmented in the elderly and patient groups. Finally, significant changes in apparent fibre density (AFD) were detected between the young, elderly, and patient groups, which could net be detected in the DTI-based metrics.

Upon review of the MD maps, we noticed a hyperintense halo around the tongues. This increase in MD may have originated from a partial volume effect of muscles and the sublingual salivary glands. This partial volume effect was also present at other locations, such as the interface of tongue and saliva. The effect of free water surrounding the tongue muscles (either in the saliva or the salivary glands) could be minimized in future studies. Instead of using reference images with a b-value of 0 mm^2/s, a low b-value (20 mm^2/s for example) could be used to strongly reduce the intensity of free water in the DWI, without affecting the intensity of muscles to a large extent. In addition, a CSD variant that uses multiple response functions for different tissue types within a voxel, multi-tissue CSD[23], could overcome the partial volume effect.[23] However, as multi-tissue CSD generally works best with multiple high-angular resolution diffusion imaging (HARDI)[24] shells, i.e. multiple b-values, this methodology could not be applied to our single-shell data sets.

Due to its highly complicated anatomy, the interpretation of diffusion parameters such as MD and FA in the tongue is more challenging than for skeletal muscles. In fact, as the tongue is known for its crossing muscle fibres, i.e. multiple fibre populations exist within one voxel, the diffusion tensor is unable to properly describe the tongue muscle architecture. Although a higher-order diffusion model such as the CSD better represents the complex muscular architecture of the tongue[8], no other studies have yet calculated CSD-derived metrics such as the AFD in muscle. As there were significantly different muscles for the AFD and no significant differences in either the FA or MD, CSD may be a more sensitive method for detecting changes in the tongue muscular architecture than DTI.

However, the exact physiological basis of the AFD in muscle is yet unknown. In the brain, the AFD has been attributed to the intra-axonal water volume along a certain direction.[21] In muscles, we might therefore similarly interpret the AFD as representing the intracellular water content. As the AFD was decreased in the elderly compared to the patients, this may indicate a smaller intracellular water

volume, which may be expected in the older subjects and patients due to muscular atrophy and potentially lipomatosis (the replacement of muscle with fatty tissue).[25]

In contrast to the decrease of the AFD with ageing, for one patient, the AFD of the musculature near a tumour was increased, which could indicate an increase of the intracellular water fraction. This could originate from increased pressure from the tumour, forcing the extracellular fluid out of the surrounding muscular tissue. Another possibility may be that the inflammatory response to the tumour has resulted in damage and swelling of the surrounding muscle fibres.

Although the generation of population averages based on CSD MRI has been previously reported for the brain, to our knowledge, this work is the first to apply this methodology to musculature. In principle, this methodology should be applicable to any muscle or group of muscles, especially in other anatomies where crossing fibres occur, such as the perineal body of the pelvic floor.[26] If no crossing fibres are present in the musculature of interest, this approach may also be adapted to muscle DTI, given the similarities between DTI and CSD with an $l_{max} = 2$. For DTI however, it is recommended to use registration algorithms that are driven by DTI and therefore would not require the estimation of a response function.[27]

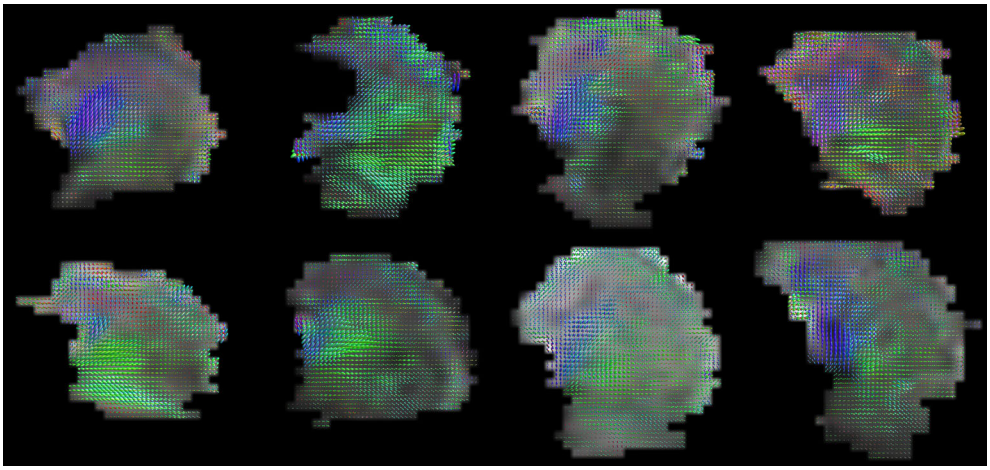
The main limitation of this study is the limited group size, especially in the patient group. Although we initially included more patients than the eight reported in this work, patients had to be excluded due to the presence of motion artifact in the images. These motion artefact may be mitigated during the image acquisition by acceleration techniques, such as multiband imaging.[28,29]

In conclusion, we provided a methodology for the analysis of the muscular architecture of the tongue by generating population averages. As we were unable to determine differences between our three groups in FA and MD, the CSD-based metric AFD appears to be more sensitive to changes in tongue musculature than a DTI-based approach. Using population averaging, AFD differences could be identified between groups, which may be explained by microstructural changes in the muscle fibres of the tongue. Further histological validation studies are necessary to study the cause and clinical significance of these changes due to aging or in the presence of cancer.

Acknowledgements

We would like to thank Kilian Kappert for helping with patient inclusion, Anita Paape and Bas Jasperse for their help with the scheduling of the patient scans, and Ferdinand van der Heijden for our critical discussion about the methodology. We are grateful to the developers of the excellent software package MRtrix3.

Supplementary materials



SUPPLEMENTARY MATERIAL 3.1 — For the eight included patients, an overview is presented of the midsagittal slice of the fibre-orientation distribution (FOD) maps. The FOD maps have not been registered to a common space. The FOD maps are coloured based on their direction: left-right in red; anterior-posterior in green; feet-head in blue. In the second tongue, note that the tip is missing due to an artefact caused by a metal orthodontic appliance.

References

1. S. Abd-el-Malek, "Observations on the morphology of the human tongue," *Journal of anatomy*, vol. 73, pp. 201–210.3, 1939. DOI: 10.1099/0022-1317-37-2-435.

2. 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, Apr. 1985. DOI: 10.1111/j.1096-3642.1985.tb01178.x.

3. M. Froeling, J. Oudeman, S. van den Berg, K. Nicolay, M. Maas, G. J. Strijkers, M. R. Drost, and A. J. Nederveen, "Reproducibility of diffusion tensor imaging in human forearm muscles at 3.0 T in a clinical setting.," *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, vol. 64, no. 4, pp. 1182–1190, 2010. DOI: 10.1002/mrm.22477.

4. M. Froeling, G. J. Strijkers, A. J. Nederveen, S. A. Chamuleau, and P. R. Luijten, "Diffusion Tensor MRI of the Heart – In Vivo Imaging of Myocardial Fiber Architecture," *Current Cardiovascular Imaging Reports*, vol. 7, no. 7, p. 9276, 2014. DOI: 10.1007/s12410-014-9276-y.

5. R. J. Gilbert, T. a. Gaige, R. Wang, T. Benner, G. Dai, J. N. Glickman, and V. J. Wedeen, "Resolving the three-dimensional myoarchitecture of bovine esophageal wall with diffusion spectrum imaging and tractography.," *Cell and tissue research*, vol. 332, no. 3, pp. 461–468, 2008. DOI: 10.1007/s00441-008-0601-0.

6. I. Sanders, L. Mu, A. Amirali, H. Su, and S. Sobotka, "The human tongue slows down to speak: Muscle fibers of the human tongue," *Anatomical Record*, vol. 296, no. 10, pp. 1615–1627, 2013. DOI: 10.1002/ar.22755. arXiv: NIHMS150003.

7. J.-D. Tournier, F. Calamante, and A. Connelly, "Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution.," *NeuroImage*, vol. 35, pp. 1459–1472, May 2007. DOI: 10.1016/j.neuroimage.2007.02.016.

8. L. Voskuilen, M. Italiaander, P. de Heer, A. J. Balm, F. van der Heijden, G. J. Strijkers, L. E. Smeele, and A. J. Nederveen, "A 12-channel flexible receive coil for accelerated tongue imaging," in *Proceedings of the International Society for Magnetic Resonance in Medicine*, 2019.

9. J. Mazziotta, A. Toga, A. Evans, P. Fox, J. Lancaster, K. Zilles, R. Woods, T. Paus, G. Simpson, B. Pike, C. Holmes, L. Collins, P. Thompson, D. MacDonald, M. Iacoboni, T. Schormann, K. Amunts, N. Palomero-Gallagher, S. Geyer, L. Parsons, K. Narr, N. Kabani, G. Le Goualher, D. Boomsma, T. Cannon, R. Kawashima, and B. Mazoyer, "A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM)," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 356, no. 1412, pp. 1293–1322, 2001. DOI: 10.1098/rstb.2001.0915.

10. V. Pedoia, X. Li, F. Su, N. Calixto, and S. Majumdar, "Fully automatic analysis of the knee articular cartilage T1rho relaxation time using voxel-based relaxometry," *Journal of Magnetic Resonance Imaging*, vol. 43, no. 4, pp. 970–980, 2016. DOI: 10.1002/jmri.25065.

11. J. Woo, J. Lee, E. Z. Murano, F. Xing, M. Al-Talib, M. Stone, and J. L. Prince, "A High-resolution Atlas and Statistical Model of the Vocal Tract from Structural MRI," *Comput Methods Biomech Biomed Eng Imaging Vis.*, vol. 3, no. 1, pp. 47–60, 2015. DOI: 10.1017/CBO9781107415324.004. arXiv: arXiv:1011.1669v3.

12. J. Veraart, D. S. Novikov, D. Christiaens, B. Ades-aron, J. Sijbers, and E. Fieremans, "Denoising of diffusion MRI using random matrix theory," *NeuroImage*, vol. 142, pp. 1–28, 2016. DOI: 10.1016/j.neuroimage.2016.08.016.

13. J. L. R. Andersson and S. N. Sotiropoulos, "An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging," *NeuroImage*, vol. 125, pp. 1063–1078, 2016. DOI: 10.1016/j.neuroimage.2015.10.019.

14. J. D. Tournier, F. Calamante, and A. Connelly, "MRtrix: Diffusion tractography in crossing fiber regions," *International Journal of Imaging Systems and Technology*, vol. 22, no. 1, pp. 53–66, 2012. DOI: 10.1002/ima.22005.

15. J. D. Tournier, F. Calamante, and A. Connelly, "Determination of the appropriate b value and number of gradient directions for high-angular-resolution diffusion-weighted imaging," *NMR in Biomedicine*, vol. 26, no. 12, pp. 1775–1786, 2013. DOI: 10.1002/nbm.3017.

16. J. D. Tournier, C. H. Yeh, F. Calamante, K. H. Cho, A. Connelly, and C. P. Lin, "Resolving crossing fibres using constrained spherical deconvolution: Validation using diffusion-weighted imaging phantom data," *NeuroImage*, vol. 42, no. 2, pp. 617–625, 2008. DOI: 10.1016/j.neuroimage.2008.05.002.

17. S. Joshi, B. Davis, M. Jomier, and G. Gerig, "Unbiased diffeomorphic atlas construction for computational anatomy," *NeuroImage*, vol. 23, no. SUPPL. 1, 2004. DOI: 10.1016/j.neuroimage.2004.07.068.

18. D. Raffelt, J. D. Tournier, J. Fripp, S. Crozier, A. Connelly, and O. Salvado, "Symmetric diffeomorphic registration of fibre orientation distributions," *NeuroImage*, vol. 56, no. 3, pp. 1171–1180, 2011. DOI: 10.1016/j.neuroimage.2011.02.014.

19. J.-D. Tournier, F. Calamante, and A. Connelly, "Improved probabilistic streamlines tractography by 2nd order integration over fibre orientation distributions," *Proceedings of the International Society for Magnetic Resonance in Medicine*, vol. 88, no. 2003, p. 2010, 2010.

20. R. Wang and V. J. Wedeen, *TrackVis*, 2015. [Online]. Available: http://trackvis.org/.

21. D. Raffelt, J. D. Tournier, S. Rose, G. R. Ridgway, R. Henderson, S. Crozier, O. Salvado, and A. Connelly, "Apparent Fibre Density: A novel measure for the analysis of diffusion-weighted magnetic resonance images," *NeuroImage*, vol. 59, no. 4, pp. 3976–3994, 2012. DOI: 10.1016/j.neuroimage.2011.10.045.

22. D. A. Raffelt, J. D. Tournier, R. E. Smith, D. N. Vaughan, G. Jackson, G. R. Ridgway, and A. Connelly, "Investigating white matter fibre density and morphology using fixel-based analysis," *NeuroImage*, vol. 144, pp. 58–73, 2016. DOI: 10.1016/j.neuroimage.2016.09.029.

23. B. Jeurissen, J. D. Tournier, T. Dhollander, A. Connelly, and J. Sijbers, "Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data," *NeuroImage*, vol. 103, pp. 411–426, 2014. DOI: 10.1016/j.neuroimage.2014.07.061.

24. D. S. Tuch, T. G. Reese, M. R. Wiegell, N. Makris, J. W. Belliveau, and J. Van Wedeen, "High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity," *Magnetic Resonance in Medicine*, vol. 48, no. 4, pp. 577–582, 2002. DOI: 10.1002/mrm.10268.

25. R. Bässler, "Histopathology of different types of atrophy of the human tongue," *Pathology Research and Practice*, vol. 182, no. 1, pp. 87–97, 1987. DOI: 10.1016/S0344-0338(87)80147-4.

26. F. M. Zijta, M. Froeling, M. P. Van Der Paardt, M. M. E. Lakeman, S. Bipat, a. D. Montauban Van Swijndregt, G. J. Strijkers, A. J. Nederveen, and J. Stoker, "Feasibility of diffusion tensor imaging (DTI) with fibre tractography of the normal female pelvic floor," *European Radiology*, vol. 21, no. 6, pp. 1243–1249, 2011. DOI: 10.1007/s00330-010-2044-8.

27. H. Zhang, P. A. Yushkevich, D. C. Alexander, and J. C. Gee, "Deformable registration of diffusion tensor MR images with explicit orientation optimization," *Medical Image Analysis*, vol. 10, no. 5, pp. 764–785, 2006. DOI: 10.1016/j.media.2006.06.004. arXiv: 1008.5070.

28. L. Voskuilen, P. de Heer, L. van der Molen, A. J. M. Balm, F. van der Heijden, G. J. Strijkers, L. E. Smeele, and A. J. Nederveen, "A 12-channel flexible receiver coil for accelerated tongue imaging," *Magnetic Resonance Materials in Physics, Biology and Medicine*, 2020. DOI: 10.1007/s10334-019-00824-5.

29. K. Setsompop, B. A. Gagoski, J. R. Polimeni, T. Witzel, V. J. Wedeen, and L. L. Wald, "Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty," *Magnetic Resonance in Medicine*, vol. 67, no. 5, pp. 1210–1224, 2012. DOI: 10.1002/mrm.23097.

4

Chapter 4

Personalized biomechanical tongue models based on diffusion-weighted MRI and validated using optical tracking of range of motion

Kilian D.R. Kappert, Luuk Voskuilen, Ludi E. Smeele, Alfons J.M. Balm, Bas Jasperse, Aart J. Nederveen, Ferdinand van der Heijden

Biomechanics and Modeling in Mechanobiology

2021

Abstract

For advanced tongue cancer, the choice between surgery or organ-sparing treatment is often dependent on the expected loss of tongue functionality after treatment. Biomechanical models might assist in this choice by simulating the post-treatment function loss. However, this function loss varies between patients and should, therefore, be predicted for each patient individually. In the present study, the goal was to better predict the postoperative range-of-motion (ROM) of the tongue by personalizing biomechanical models using diffusion-weighted MRI and constrained spherical deconvolution (CSD) reconstructions of tongue muscle architecture.

Diffusion-weighted MRI scans of ten healthy volunteers were obtained to reconstruct their tongue musculatures, which were subsequently registered to a previously described population average or atlas. Using the displacement fields obtained from the registration, the segmented muscle fiber tracks from the atlas were morphed back to create personalized muscle fiber tracks. Finite element models were created from the fiber tracks of the atlas and those of the individual tongues. Via inverse simulation of a protruding, downward, left and right movement, the ROM of the tongue was predicted. This prediction was compared to the ROM measured with a 3D camera.

It was demonstrated that biomechanical models with personalized muscles bundles are better in approaching the measured ROM than a generic model. However, to achieve this result a correction factor was needed to compensate for the small magnitude of motion of the model. Future versions of these models may have the potential to improve the estimation of function loss after treatment for advanced tongue carcinoma.

Introduction

The incidence of tongue cancer is rising worldwide, accounting for almost 20% of all head and neck cancers.[1,2] Locally advanced tongue cancer is usually treated by surgery and/or chemoradiation, which may have a serious impact on the mobility of the tongue due to surgical defects and/or radiation-induced fibrosis. This often leads to difficulties with speech, mastication, and swallowing.[3,4] The choice between surgical or organ-sparing treatment is dependent on expected function loss after treatment, which is difficult to predict.[5] The prediction of the expected function loss would be of great benefit for the decision-making process shared between physician and patient. Biomechanical modeling of the tongue would be a logical next step in the process of the prediction of functional loss.

The biomechanics of the tongue however are complex.[5–7] The tongue consists of four extrinsic and four intrinsic muscles, which interdigitate and seem to follow a strict pattern.[8] Although we know that all muscles, except for the palatoglossus muscle, are innervated by the hypoglossal nerve[9,10], the complex neural strategies that are required for shaping the tongue during speech and mastication, are currently unknown.[11,12] Moreover, the tongue shape varies between individuals, and knowledge about anatomical variations in muscle structure does not yet exist.[13] Because of the difficulties in obtaining the right mechanical properties and simulating the viscoelastic nature of the tongue, mechanical properties are often approached by a hyperplastic model using ex-vivo data.[14–17]

Despite the challenges biomechanical finite-element (FE) models have shown to be a promising method to predict functional loss after treatment in their current form.[14–16] However, these FE models are generally generic and are therefore unable to predict functional loss on an individual level and should be personalized.

One way of creating personalized FE models is by morphing of a generic FE model to a subject-specific situation.[18,19] Previous work has shown that this morphing can be driven by imaging data such as anatomical slices[20], computed tomography[21,22], and MRI[23,24]. Alternatively, personalized models can also be constructed by embedding mesh and muscle structures in a coarse FE model that is generated according to the shape of the mesh.[25,26] If muscles are however included in personalized models, the morphing should not only be driven by the outline of anatomical structures or meshes, but also the internal structure of the muscle, such as the muscle fiber directionality.

This muscle fiber directionality can be measured by exploiting the possibilities of diffusion-weighted MRI.[27] Using diffusion-sensitizing gradients, it is possible to encode MR images with diffusion information along a certain direction. As the diffusivity of water is higher along muscle fibers than perpendicular to them, it is possible to reconstruct the fiber orientation using the diffusion tensor. In fiber tracking or tractography, fiber tracks are computed from these fiber orientations[28], easing the visualization of the tongue musculature.[29–32] These tracks have even

been used as an input for biomechanical models of the tongue.[33] Despite this potential of DTI, it is unable to resolve crossing or merging muscle fibers on the tongue. Recently, a diffusion-weighted MRI technique called constrained spherical deconvolution (CSD), which can resolve the interdigitating muscle fibers of the tongue in vivo, was applied to the tongue.[34] This technique enables us to reconstruct the tongue muscle architecture of the individual more accurately.

The goal of the current work was to create personalized biomechanical models of the tongue by using CSD MRI. As manual embedding of all the fibers of this muscle architecture in the FE model would be very laborious, it is hardly feasible. This motivated us to use automated methods to embed these fibers. While CSD MRI is superior in resolving crossing and merging fibers, the high noise level resulted in relatively low reconstructions of the tongue musculature of a single subject. We, therefore, proposed to use a population average or atlas of the tongue muscle architecture[35], which is more resistant to noise and artifacts. By mapping the tongue of an individual to the atlas tongue, we hypothesized that the segmented fiber tracks of the atlas can be morphed back to an individual's space and, subsequently, from these segmented fiber tracks a personalized biomechanical model can be created. The effect of this personalization step was evaluated by comparing models with both personalized and generic muscles bundles to the predicted range-of-motion (ROM) of the tongue measured in vivo using 3D optical tracking.[36]

Methods

The following section covers the characteristics of volunteers and the measurement of their ROM. Next, the creation of the personalized biomechanical models is described, which is summarized in figure 4.1. Methods are divided into those applied to the atlas (figure 4.1A1–9) and those applied to personalized models (figure 4.1P1–7). Finally, the ROM predicted by these biomechanical models and the atlas were compared to the measured ROM.

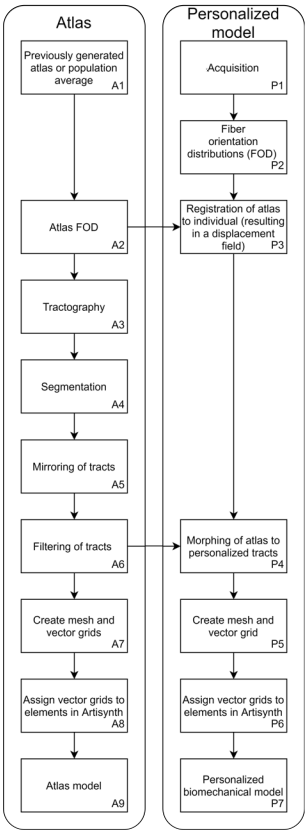


FIGURE 4.1 — A flow chart of the steps required to create an atlas-based (A1–9) and a personalized model (P1–7)).

Volunteers & ROM measurement

A total of ten healthy volunteers were included with a mean age of 61 years (range: 56 to 71; seven men) to match the same age group of tongue cancer patients. Volunteers with steel braces or any contra-indication to an MRI scan were excluded. The ROM of the tongue was obtained by optical tracking of a marker on the tip of the tongue using a 3D camera. The volunteers were asked to perform four different tongue movements: left, right, down, and protrusion as described in the paper by Kappert et al.[36] The up-movement was left out since it was proven to be unreliable. Written informed consent was obtained from all volunteers before inclusion. This study was approved by the medical ethical committee of the Netherlands Cancer Institute (ref: N17BTM).

CSD MRI acquisition and processing

The volunteers were scanned in a 3 T Philips MRI scanner (Philips Healthcare, Best, The Netherlands) using a neurovascular coil according to the CSD scan protocol by Voskuilen et al. (figure 4.1P1).[34] The raw diffusion-weighted images were acquired using the following parameters: single-shot spin-echo echo-planar imaging; echo-train length 25; repetition time: 3.4 s; echo time: 60 ms; two repetitions with opposing phase-encoding directions; number of signal averages: 1; fat suppression: spectral presaturation with inversion recovery and slice-selection gradient reversal; field-of-view: 192 by 156 by 84 mm; voxel size: 3 mm isotropic; b-value: 700s/mm^2 along 64 directions evenly spaced over a hemisphere and optimized for gradient load; total scan time: 10 minutes.

Subsequently, the noise of the diffusion-weighted images was reduced using the method of Veraart et al.[37] Using FSL, a software library for diffusion MRI[38], the diffusion-weighted images were corrected for distortions caused by B_0-inhomogeneity, eddy currents from the diffusion-encoding gradients, and rigid motion[39]. For all subjects, masks of the tongue were created by manual delineation in ITK-Snap[40]. In MRtrix3[41], the corrected diffusion-weighted images were upsampled to a resolution of 1.5 mm isotropic using a b-spline interpolation. For each volunteer, we estimated a CSD response function, which corresponds to the diffusion signal of a single fiber population.[42] By deconvolving, using CSD, the corrected diffusion-weighted images with this response function[43], fiber-orientation distribution (FOD) maps were calculated up to a maximum spherical harmonic degree of 8 (figure 4.1P2 , figure 4.2).

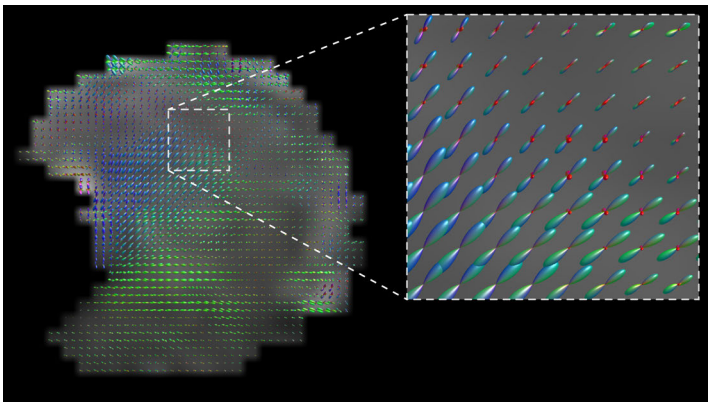


FIGURE 4.2 — Mid-sagittal view of an fibre-orientation distribution (FOD) map of volunteer 2. In the inlay, the map is enlarged so the individual FODs can better be appreciated. The FODs are coloured according to their direction: red for right-left; green for anterior-posterior; and blue for feet-head.

Finally, using symmetric diffeomorphic registration based on the FOD maps[44], displacement fields were calculated from each volunteer to the tongue muscle atlas described in Voskuilen et al.[35] (figure 4.1P3). This atlas is a population average of ten volunteers different than those included in this study (mean age of 25.5 years; four female). The atlas has a voxel size of 1.5 mm isotropic after up sampling. The registration error between the FOD maps of the individuals and that of the atlas was quantified by the L_2-norm and the angular correlation coefficient.[45] The calculated displacement fields were used later in this work to morph the generical biomechanical model of the atlas.

Fiber tracking and filtering

Although it would be possible to compute a 3D vector field (required to build a biomechanical model) from the FOD maps directly, CSD-based fiber tracking was first performed on the atlas (figure 4.1A3).[46] Fiber tracking ignored many spurious vectors, and the segmentation of streamlines was less time-consuming than segmentation of vectors. For this global fiber tracking, the following parameters were used: step size: 1.5 mm; angular threshold: 15°; FOD peak threshold: 10% of the largest peak; maximal length: 100 mm; minimal length: 10 mm; number of seed points: 10,000 randomly placed within the mask. In TrackVis[47], the fiber tracking was manually segmented into the following muscle tracts: genioglossus, geniohyoid, hyoglossus, inferior longitudinal, superior longitudinal, transverse, and vertical muscles (figure 4.1A4 and 4.3). These segmentations were subsequently checked by a head-and-neck surgeon. The styloglossus muscle could not be distinguished from the inferior longitudinal and was therefore not included.[34]

In Matlab R2019a (Mathworks, Natick, MA), these atlas tracks were mirrored in the midsagittal plane to ensure the symmetry of the atlas (figure 4.1A5). To remove faulty tracks, while preserving the muscle shape, the muscle tracts were filtered using the criteria shown in appendix I (figure 4.1A6). These criteria were chosen empirically based on the reduction in outliers and the expected curvature obtained from earlier anatomical research.[9]

For each volunteer of the study group, the displacement fields obtained from the registration earlier were used to morph the filtered tracts from the atlas into personalized tracts (figure 4.1P4). The personalized tracts may have been rotated during morphing and were therefore reoriented, based on the orientation of the muscles before morphing.

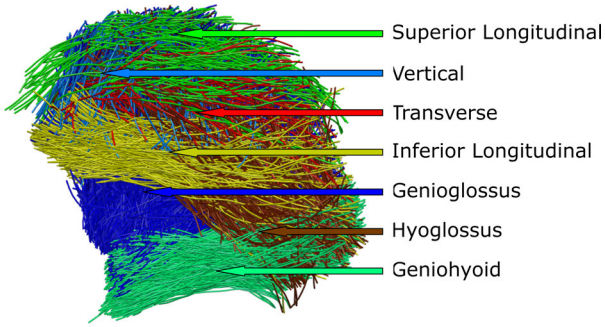


FIGURE 4.3 — Side view of the segmented fiber tracts of the atlas.

FE model construction

The tracks of both the atlas model and personalized models were converted into vector fields of muscle fiber direction (figure 4.1P5 and 4.1A7), using the following steps. For each muscle, a convex hull was calculated that enclosed the filtered tract (figure 4.4A–B). These convex hulls were filled with a uniformly-distributed grid of vectors, where the direction of these vectors was determined by an inverse distance interpolation of nearby tracks (figure 4.4C; figure 4.5; equation 4.1). This was done for both left and right muscles independently if applicable (figure 4D).

$$\mathbf{n}_{\text{grid}}(\mathbf{x}) = \sum_i \left(1 - \sqrt{\frac{d(\mathbf{x}, \text{track}_i)}{\max_{\mathbf{x}}(d(\mathbf{x}, \text{track}_i))}}\right) \mathbf{n}_{\text{track}_i}$$

$$\text{with } \mathbf{n}_{\text{grid}}(\mathbf{x}) = \frac{\mathbf{n}_{\text{grid}}(\mathbf{x})}{\|\mathbf{n}_{\text{grid}}(\mathbf{x})\|}$$

EQUATION 4.1 — $\mathbf{n}_{\text{grid}}(\mathbf{x})$ is the uniformly distributed grid of vectors within the convex hull, Σ_i the sum over all vectors and $\mathbf{n}_{\text{track}_i}$, the original tracks. $d(\mathbf{x},\text{track}_i)$ is the distance of vector $\mathbf{x}$ to track_i.

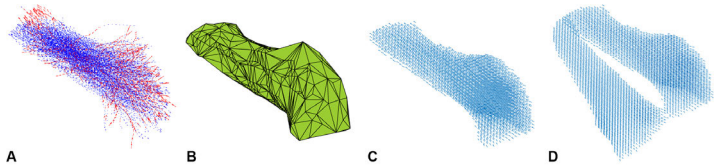


FIGURE 4.4 — The tracks (blue) and filtered tracks (red) from the inferior longitudinal muscle (a); the convex hull enclosing these tracks (b); a uniformly-distributed vector field based on the direction of the tracks within the convex hull (c); and vector fields of both left and right muscles (d).

The genioglossus muscle was divided into an oblique and horizontal part based on the estimated position of the short tendon.[8] With all tracks combined, a convex hull was generated to create a mesh of the tongue. In Meshlab[48], the HC Laplacian filter[49] was used to smoothen the mesh. Attachment points for the mandible and hyoid bones were determined based on the endpoints of the extrinsic muscle tracts from the atlas model.

Using ArtiSynth[50] —a platform for combined multibody and FE modeling— the muscle vector fields and tongue meshes of the atlas and the ten volunteers were integrated into a biomechanical model (figure 4.1P6, 1A8 and figure 4.6) using the following steps, which are similar to those described by Kappert et al.[16] The surface mesh of the tongue, obtained from the previous step, is embedded into a FE model consisting of 16 mm^3 cubic hexahedral elements, generated to match the shape of the surface mesh. The benefit of only using cubic elements is that they can easily be removed or added, which is an essential feature to

simulate surgical resections on personalized biomechanical models in the future.[16] The attachment points to the hyoid and mandible were simulated by making FE nodes non-dynamic.

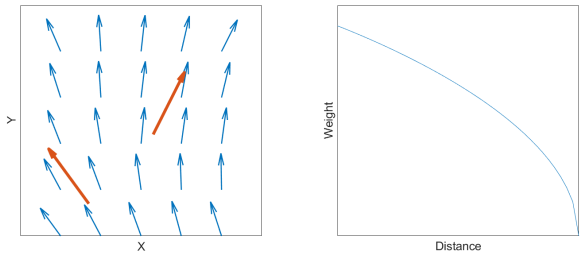


FIGURE 4.5 — The direction of $\mathbf{n}_{\text{grid}}(\mathbf{x})$ (small blue arrows) is determined by nearby $\mathbf{n}_{\text{track}_i}$ (large red arrows) by means of inverse distance interpolation (equation 4.1) visualized in the right figure.

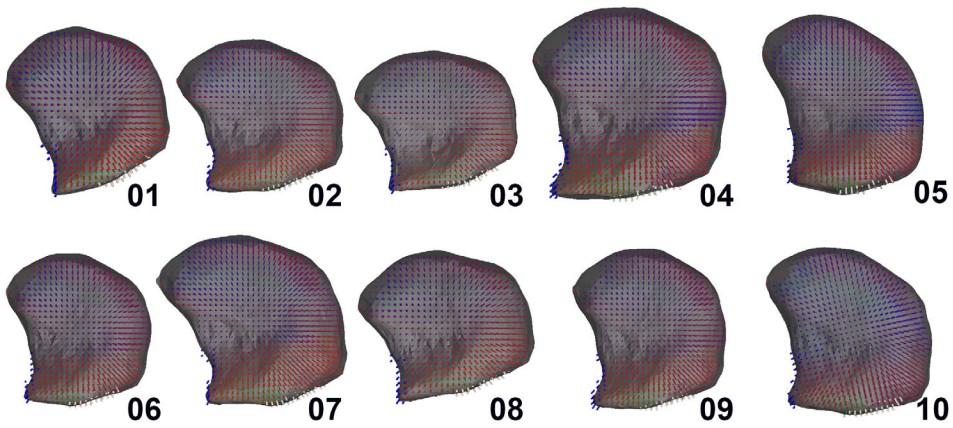


FIGURE 4.6 — Sagittal section view of personalized FE tongue models of the ten healthy volunteers. The direction of force of the muscle elements has been color-coded: anterior-posterior in red; right-left in green; and feet-head in blue. Bone attachment points are visualized as floating point outside the mesh. The mandible attachment points are visualized in blue, and those of the hyoid bone in white.

Based on previous work[51] an incompressible Moony-Rivlin material was chosen for mechanical properties of the tissue:

$$W = C_{10}(I_1 - 3) + C_{20}(I_1 - 3)^2 + \kappa(\ln J)^2$$

EQUATION 4.2 — Where I_1 is the first invariant of the left Cauchy-Green deformation tensor, C_{10} and C_{20} stiffness parameters equal to 1037 Pa and 486 Pa, $\kappa = 100{*}C_{10}$ the bulk modulus to obtain a Poisson's ratio close to 0.5, and J the determinant of the deformation gradient.

The stiffness parameters are obtained from a fresh cadaver in[52] and later scaled by a factor of 5.4 in[51] to match in-vivo measurements. C_{10} is equal to 1037 Pa, C_{20} equal is to 486 Pa and the other parameters all zero. Rayleigh damping coefficients of $\alpha = 40$ s^{-1} and $\beta = 0.03$, and a density of 1040 kg/m^3 were used, comparable to those used by Buchaillard et al.[51], Stavness et al.[53] and Kappert et al.[16].

Muscle contraction was simulated using "muscle material" in ArtiSynth. With this technique, when a muscle bundle is activated, it applies external stresses on the elements associated with the muscle bundles, in addition to the regular tissue material.[50] The transversely isotropic properties of the muscles where included using ArtiSynth's "Muscle Material" interpretation of the method by Blemker et al.[54]:

$$\sigma(\lambda) = \sigma_{\max}(\alpha f_{\text{act}}(\lambda) + f_{\text{pass}}(\lambda)) \left(\frac{\lambda}{\lambda_{\text{opt}}}\right)$$

EQUATION 4.3 — With $\sigma_{\max}$ the max isometric stress in the muscle, α the normalized activation level, λ the stretch along fiber, λ_{opt} the optimal fiber stretch, f_{act} the active force-length relationship, and f_{pass} the passive.

Simulation and analysis

Inverse simulation, provided by ArtiSynth[50,53], was used to instruct the tongue tip of the personalized FE models to consecutively move to a point anterior, inferior, left and right of its initial location. The predicted ROM was defined as the distance from the initial location to maximal deflection in one of the instructed directions.

The tongue can reach strain values of 200% for elongation and 160% for contraction.[55] Using the current biomechanical model, it is not possible to simulate these magnitudes of deformation. As all model simulations use the same constitutive models, tissue properties and FE generation technique, the effect thereof is considered to be constant for all models. To evaluate the differences between the

FE models with personalized muscle bundles and those using generic muscle bundles (the atlas model), we will therefore focus on the relative differences between individuals and not on the absolute values. In order to compare the simulations to the measured ROM, a scaling factor is needed to compensate for the reduced magnitude of motion of the model. The reduced magnitude of motion will be different depending on the movement direction and therefore four scaling factors were calculated and applied to all simulations equally. To make sure outliers would not affect the scaling factors, these factors were determined by an iterative process to achieve the maximum number of predicted ROMs ($ROM_{\text{pred}}(i,j)$) that were within the CI ($ROM_{2\sigma}$) of the measured ROM ($ROM_{\text{meas}}(i,j)$) (Equation 4.4).

$$S_i = \arg\max_{S_i} \sum_{j=1}^{10} [|S_i \cdot ROM_{\text{pred}}(i,j) - ROM_{\text{meas}}(i,j)| < ROM_{2\sigma}]$$

$$\text{where } [x] = \begin{cases} 0 & \text{if } x = false \\ 1 & \text{if } x = true \end{cases}$$

In which,

$i = 1,\cdots,4$	Index for the 4 different movements
$j = 1,\cdots,10$	Index for the 10 volunteers from the study group
$ROM_{\text{pred}}(i,j)$	Predicted ROM
$ROM_{\text{meas}}(i,j)$	Measured ROM
S_i	The scaling factor applied to predicted $ROM_{\text{pred}}(i,j)$
$ROM_{2\sigma}$	Twice the standard deviation of the measured ROM (i.e. 6 mm)

EQUATION 4.4 — Optimization of the scaling factor such that the number of predictions $ROM_{\text{pred}}(i,j)$ within the bounds $ROM_{2\sigma}$ of the measured ROM ($ROM_{\text{meas}}(i,j)$) is maximized.

The predicted ROM was compared to the *in vivo* measured ROM of the individual on which the personalized model was based. To show the benefit of personalization, also the Atlas model (essentially a generic model) will be compared with the measured ROM. Only when the personalized models perform better than the atlas, we can conclude that personalization improves the ROM prediction.

Previously, the precision of the ROM measurements, quantified by the standard deviation, was determined to a range from 2.3 mm to 3.2 mm.[36] We, therefore, assumed a precision of 3 mm (3.2 mm rounded off) for all ROM measurements. If a predicted ROM fell within the 95% confidence interval (CI), i.e., within two times the standard deviation, we judged the measurement to be correct.

Results

Visually, the FOD maps were well aligned to the atlas. The error in alignment or registration error was quantified by the L_2-norm and the angular cross correlation. The mean L_2-norm between the FOD maps and the atlas was 0.302 (SD 0.030). The mean angular correlation coefficient was 0.634 (SD 0.057).

In figure 4.7, the distances for specific tongue movements of both the measured ROM and predicted ROM are shown for all ten subjects. For the predicted ROM, both the scaled and non-scaled movements are shown. The scaling factors are 2.6 for protrusion, 2.2 for down, 2.4 for left, and 2.6 for right. Protrusion and down movements show the best agreement between predictions and measurements, as nine out of ten (90%) predicted ROM's are within the CI. Eight movements (80%) to the right were predicted to be within the CI, but for the movement to the left, only six (60%) were predicted to be within the interval. In total, 32 out of the 40 predictions (4 movements, 10 volunteers) from the personalized models were within the CI. The largest disagreements between the prediction and in vivo measurement were found in subject 08. The atlas model is the same for every subject which is depicted more clearly in the following table.

TABLE 4.1 — The percentual difference between the personalized model or Atlas (generic model) and the measured ROM per subject. The last column shows the mean percentual difference.

Movement	Model	01	02	03	04	05	06	07	08	09	10	Mean
Out	Personalized	8%	4%	1%	16%	3%	0%	12%	21%	8%	36%	11%
	Atlas	30%	12%	47%	4%	17%	17%	5%	17%	22%	28%	20%
Down	Personalized	11%	7%	7%	8%	1%	6%	12%	35%	5%	4%	10%
	Atlas	6%	23%	23%	4%	0%	21%	6%	25%	12%	37%	16%
Left	Personalized	5%	28%	17%	8%	1%	26%	4%	28%	6%	11%	13%
	Atlas	34%	36%	27%	8%	10%	26%	19%	35%	8%	16%	22%
Right	Personalized	8%	0%	2%	27%	11%	12%	3%	40%	9%	2%	11%
	Atlas	11%	18%	14%	9%	17%	9%	21%	27%	0%	6%	13%

In table 4.1 the percentual differences of the models with the measured ROM are shown for every subject. The model that approaches the measured ROM better differs between subject and movement, but a majority of the measurements are approached better using the personalized model. The mean percentual difference per movement shows that for all movements the difference with the measured ROM is lower for the personalized models.

In figure 4.8 the atlas model and the personalized model of subject 3 are shown within the ArtiSynth environment. For the four simulated movements, the maximal extension is shown. Subject 3 demonstrated a ROM that that in 3 out of 4 movements could not be predicted using the atlas, but could be using the personalized model. The movement of the atlas, relative to its rest-state, looks larger in most directions than the personalized model as confirmed by the bar charts

in figure 4.7. Also, the tongue moves more upwards during the right movement. The magnitude of the movements of both models is smaller than what would be expected from a real tongue.

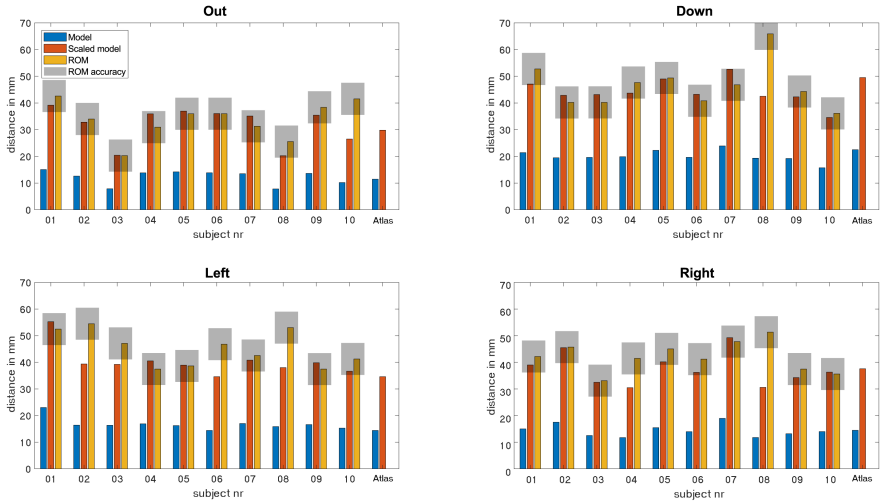


FIGURE 4.7 — Range-of-motion (ROM) in mm for the ten healthy volunteers (01–10) and the Atlas (Generic model), for protrusion, and the down, left, and right movements. The predicted ROM of the personalized and atlas (generic model) is given in blue, the scaled predicted ROM in orange, and the measured ROM in yellow. The grey box depicts the interval of two times the standard deviation of the measured ROM within which the predicted ROM values of both atlas and personalized models are assumed to be accurate.

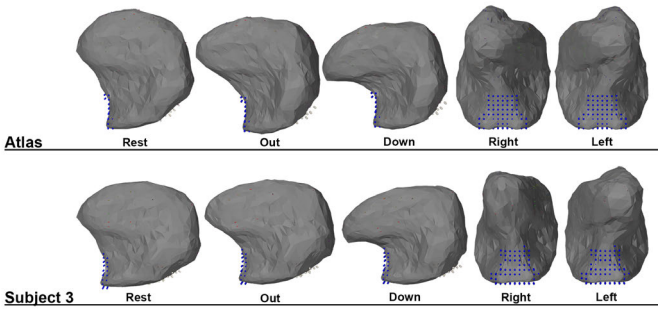


FIGURE 4.8 — An example of the maximum range in the ROM prediction for protrusion, down, left, and right using the personalized model of subject 1, 10 and the atlas.

Discussion

This study was a first approach to combine CSD MRI and FE modeling to create personalized biomechanical models of ten healthy volunteers. The results show that, after applying a correction factor to the simulations, the personalized models where comparable to the measured ROM in 80% of the cases, whereas the atlas model was only comparable in 50% of the cases. For every individual movement the personalized models also performed better than the atlas model predicting up to 90% of the down movements correctly. This confirms that using muscle bundles based on CSD in an FE model of the tongue contributes significantly to the personalization of a biomechanical tongue model.

Although the downward movement of subject 8 with a relatively small tongue was exceptionally high, this measurement was confirmed to be correct upon reviewing the images from the 3D camera. The model was not able to reproduce this large ROM, which may indicate that a large ROM is not only a result of differences in tongue muscle morphology. The ROM may also have been affected by other quantities that were not accounted for, such as the number of motor units or the stiffness, anisotropy, and density of the tissue.

In the posterior part or base of the tongue, breathing motion impaired the tracking of the superior longitudinal and transverse muscle.[34] Although these muscle tracts were filtered less vigorously, this could not resolve the absence of muscle tracts. Fortunately, the effect of the absence of these tracks is expected to be minimal, as the simulation of the ROM is less dependent on the musculature of the posterior tongue. Other artifacts, such as those caused by ferromagnetic crowns, resulted in signal voids in the diffusion-weighted images, and therefore gaps in the tractography of the tongue. As the biomechanical models were based on the atlas, where such gaps are not present, we assume these artifacts that occurred in individual data sets would cause minimal errors in the personalized models.

The genioglossus and geniohyoid muscles form one large continuous fan of tracks. As described in the literature some FE models divide the genioglossus into an anterior, middle, and posterior part[15,56–58], while others separate the genioglossus into a horizontal and an oblique subdivision[8,10,59]. We chose the latter because the location of the short tendon could be inferred from our atlas and could, therefore, be used as an anatomical marker to split the genioglossus into two. As the styloglossus could not be distinguished from the inferior longitudinal muscle in the fiber tracking, the styloglossus was omitted from the model. The effect on our simulations was expected to be limited since the styloglossus is mainly involved in retracting the tongue and swallowing.

Similar to CSD in the brain, the apparent fiber density could be derived from CSD in muscles, which should in principle relate to muscle strength.[60,61] Therefore, incorporating this apparent fiber density into our biomechanical models might improve the ROM predictions. However, since CSD MRI in the tongue is subject to

higher noise levels and more motion artifacts than for example in the brain, in our opinion, the apparent fiber density can currently not be quantified accurately. We, therefore, assumed that the vectors describing the muscle direction were equally distributed within the area of specific muscles.

While large parts of the methods were automated, some key elements were still done manually. For the atlas, the segmentation of the fiber tracts and the subsequent filtering were done manually. Techniques to automate these segmentation steps are not matured yet and therefore manual input is still needed. For the personalized models, only the initial masks were manually delineated. After this step, the models could be processed without manual interference.

In this study, we mirrored the atlas to make it symmetric. However, by applying the displacement fields to create personalized biomechanical models, asymmetry was reintroduced. In this study, the orientation of the personalized biomechanical models was based on the former position of the tracks within the atlas. This leads to small lateral asymmetry in the distribution of muscles. An alternative method would be to label based on its new midline. However, determining the exact midline remains challenging, and without a gold standard, there is no way to determine which method is best.

Similar to previous work[16], we used hexahedral cubic elements with embedded muscles and mesh for the FE model, which do not optimally represent the shape of a surface. As stated in the previous work, the effect of this method on the mobility of the model is minimal. The choice for this embedded design was made so that in the future the virtual surgery method introduced in the aforementioned study can be used in combination with the personalization proposed in the current study.

In this study, optical tracking of the tongue tip was used to determine the ROM. Not only the tongue but also the mandible and hyoid bone assist the tongue tip in reaching the desired position. How much influence these structures have on the tongue ROM depends on the anatomy, innervation, and brain-muscle control. This influence had not been measured and, instead, a marker on the mandible was used to compensate for the movement of the mandible.[36] This marker may however not always reliably compensate for all complex movements, and an error should be expected in the measured ROM. Predicted ROMs were therefore judged on whether they fell within the CI of this error. This CI was relatively wide, and might, therefore, have hampered the correct judgment of small variations between the predicted ROM and measured ROM.

In the biomechanical models, the magnitude of the predicted ROM was much smaller than that of the measured ROM which, therefore, had to be scaled in order to be compared with the measured ROM. Incorporating movement of adjacent connected structured such as the hyoid bone might improve the range. Also, the mechanical properties were based on the model of Buchaillard et al. (2009)[51], which uses stiffer material properties than those originally measured in a cadaver

study, to simulate an active state of the tongue[52]. A more recent publication showed that the stiffness of the tongue in rest might be four times less.[17] Moreover, hyperelastic material models used in most FE tongue models cannot cover all the complex properties of the tongue.[15] In this study, the FE model became unstable in extreme positions using lower stiffness values. Because the same mechanical and muscle properties are used for all personalized models, we assumed that the relative difference between models could still be used to analyze the effect of personalizing the muscle bundles of the FE model. However, mechanical properties were not the only limiting factor. Also, the specific muscle morphology obtained from CSD MRI contributed to the small magnitude of motion that is smaller than other non-personalized models in the literature that use the same mechanical properties.[15,16,51] Manual editing of the muscle morphology might improve the magnitude of motion, but it was not considered as it would compromise the goal of this research, which was to automate FE modeling based on CSD data. Finally, the scaling of the predicted ROMs was, although very close, not the same in every movement direction. In part this can also be contributed to the material properties which can have a different impact on the deformation in for example a down and a left movement. The difference between left and right however corresponded to an asymmetry in the measured ROM, specifically, a deviation to the left. As explained previously[36], this may have been caused by the order of instructions given by the investigator.

In conclusion, we demonstrated that biomechanical models based on CSD MRI contributes significant to the personalization of biomechanical models. To our knowledge, we are the first to report this personalization step for improving the prediction of tongue mobility. Additional research is needed to improve the performance of biomechanical models to match the same magnitude of motion as a real tongue. In the future, personalization may improve other biomechanical models such as those of speech and swallowing, potentially leading to better simulations of actual tongue functionality. In rehabilitation after tongue cancer surgery, models can potentially be used to simulate the tongue function that could be regained by practice. In the preoperative setting, we would expect an even larger potential for the prediction of tongue function, as alterations in tongue shape and musculature due to tumor growth would also be accounted for. Therefore, these results harbor a promising perspective for the development of biomechanical models that would better predict function loss of oral cancer patients and thus improve the choice of treatment in these patients.

Acknowledgements

This study was financially supported by private funding from the Virtutis Opis Foundation (www.anbiportal.nl/web/svo), Mr B. Verwelius† and, Mr W. de Graaf. The Department of Head and Neck Oncology and Surgery of the Netherlands Cancer Institute received a research grant from Atos Medical AB (Malmö, Sweden, www.atosmedical.nl). The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

Appendix

To avoid manual manipulation of the tracts in the atlas model, four filtering steps were used. A filter was only applied if it contributed to removal of faulty tracks. The first filter checked for deviations in the tracts from a global angle. This only worked well for muscle bundles with one global direction. The next filter calculated the angle within a single track. This did not work well with curved muscles like the superior longitudinal muscle. The third filter checked for the number of neighbors and at a certain distance, thus removing tracks that are too far from the rest. The alpha shape controlled the curvature of the convex hull that enclosed a muscle bundle and is explained in the Matlab documentation (https://nl.mathworks.com/help/matlab/ref/alphashape.html).

	The angle of the vector cannot deviate more than . . . ° from the total mean vector direction of all tracks.	The angle of the vector cannot deviate more than . . . ° from the total mean vector direction of one muscle track.	Tracks with less than . . . neighbors at a distance of . . . mm will be removed	Alpha shape
Vertical	45 (z-axis)	45	3 / 0.03	0.011
Transverse			5 / 0.20	0.03
Superior longitudinal			6 / 0.20	0.014
Mylohyoid			5 / 0.05	0.015
Inferior longitudinal	45 (y-axis)	45	5 / 0.04	0.011
hyoglossus		45	3 / 0.04	0.011
Geniohyoid			3 / 0.02	
Genioglossus		45	10 / 0.10	0.011
Digastricus		45	3 / 0.03	0.011

References

1. J. E. Tota, W. F. Anderson, C. Coffey, J. Califano, W. Cozen, R. L. Ferris, M. St. John, E. E. Cohen, and A. K. Chaturvedi, "Rising incidence of oral tongue cancer among white men and women in the United States, 1973–2012," *Oral Oncology*, vol. 67, pp. 146–152, 2017. DOI: 10.1016/j.oraloncology.2017.02.019.

2. UK Cancer Research, *Head and neck cancers incidence statistics*, 2019. [Online]. Available: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/head-and-neck-cancers/incidence (visited on 04/26/2019).

3. M. A. Kreeft, L. Van der Molen, F. J. Hilgers, and A. J. Balm, "Speech and swallowing after surgical treatment of advanced oral and oropharyngeal carcinoma: a systematic review of the literature," *Eur Arch Otorhinolaryngol*, vol. 266, pp. 1687–1698, 2009. DOI: 10.1007/s00405-009-1089-2.

4. V. S. Konstantinović and N. D. Dimić, "Articulatory function and tongue mobility after surgery followed by radiotherapy for tongue and floor of the mouth cancer patients," *British Journal of Plastic Surgery*, vol. 51, no. 8, pp. 589–593, 1998. DOI: 10.1054/bjps.1998.0202.

5. A. Kreeft, I. B. Tan, M. W. M. van den Brekel, F. J. Hilgers, and A. J. M. Balm, "The surgical dilemma of 'functional inoperability' in oral and oropharyngeal cancer: current consensus on operability with regard to functional results," *Clinical Otolaryngology*, vol. 34, no. 2, pp. 140–146, 2009. DOI: 10.1111/j.1749-4486.2009.01884.x.

6. T. Bressmann, R. Sader, T. L. Whitehill, and N. Samman, "Consonant intelligibility and tongue motility in patients with partial glossectomy," *Journal of Oral and Maxillofacial Surgery*, vol. 62, no. 3, pp. 298–303, Mar. 2004. DOI: 10.1016/j.joms.2003.04.017.

7. Y. Matsui, K. Ohno, Y. Yamashita, and K. Takahashi, "Factors influencing postoperative speech function of tongue cancer patients following reconstruction with fasciocutaneous/myocutaneous flaps–a multicenter study.," *International journal of oral and maxillofacial surgery*, vol. 36, no. 7, pp. 601–9, Jul. 2007. DOI: 10.1016/j.ijom.2007.01.014.

8. I. Sanders and L. Mu, "A three-dimensional atlas of human tongue muscles," *Anatomical Record*, vol. 296, no. 7, pp. 1102–1114, 2013. DOI: 10.1002/ar.22711. arXiv: NIHMS150003.

9. H. Takemoto, "Morphological Analyses of the Human Tongue Musculature for Three-Dimensional Modeling," *Journal of Speech, Language, and Hearing Research*, vol. 44, no. 1, pp. 95–107, 2001. DOI: 10.1044/1092-4388(2001/009).

10. L. Mu and I. Sanders, "Human tongue neuroanatomy: nerve supply and motor endplates," *Clinical Anatomy*, no. 201, 2010.

11. K. Slaughter, H. Li, and A. J. Sokoloff, "Neuromuscular Organization of the Superior Longitudinalis Muscle in the Human Tongue," *Cells Tissues Organs*, vol. 181, no. 1, pp. 51–64, 2005. DOI: DOI/10.1159/000089968.

12. M. van Alphen, M. Eskes, L. Smeele, A. Balm, and F. van der Heijden, "In vivo intraoperative hypoglossal nerve stimulation for quantitative tongue motion analysis," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 1163, no. October, pp. 1–7, 2015. DOI: 10.1080/21681163.2015.1072056.

13. M. Stone, J. Woo, J. Lee, T. Poole, A. Seagraves, M. Chung, E. Kim, E. Z. Murano, J. L. Prince, and S. S. Blemker, "Structure and variability in human tongue muscle anatomy," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization*, vol. 6, no. 5, pp. 499–507, 2016. DOI: 10.1080/21681163.2016.1162752.

14. S. Buchaillard, M. Brix, P. Perrier, and Y. Payan, "Simulations of the consequences of tongue surgery on tongue mobility: Implications for speech production in post-surgery conditions," *International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 3, no. 3, pp. 252–261, 2007. DOI: 10.1002/rcs.142.

15. N. Hermant, P. Perrier, and Y. Payan, "Human Tongue Biomechanical Modeling," in *Biomechanics of Living Organs: Hyperelastic Constitutive Laws for Finite Element Modeling*, Y. Payan and J. B. T. B. o. L. O. Ohayon, Eds., Oxford: Academic Press, 2017, ch. Chapter 19, pp. 395–411. DOI: 10.1016/B978-0-12-804009-6.00019-5.

16. K. D. Kappert, M. J. van Alphen, S. van Dijk, L. E. Smeele, A. J. Balm, and F. van der Heijden, "An interactive surgical simulation tool to assess the consequences of a partial glossectomy on a biomechanical model of the tongue," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 22, no. 8, pp. 827–839, Jun. 2019. DOI: 10.1080/10255842.2019.1599362.

17. K. D. Kappert, N. Connesson, S. A. Elahi, S. Boonstra, A. J. Balm, F. van der Heijden, and Y. Payan, "In-vivo tongue stiffness measured by aspiration: Resting vs general anesthesia," *Journal of Biomechanics*, vol. 114, p. 110 147, 2021. DOI: 10.1016/j.jbiomech.2020.110147.

18. B. Couteau, Y. Payan, and S. Lavallée, "The mesh-matching algorithm: An automatic 3D mesh generator for finite element structures," *Journal of Biomechanics*, 2000. DOI: 10.1016/S0021-9290(00)00055-5.

19. I. A. Sigal, M. R. Hardisty, and C. M. Whyne, "Mesh-morphing algorithms for specimen-specific finite element modeling," *Journal of Biomechanics*, 2008. DOI: 10.1016/j.jbiomech.2008.02.019.

20. J. W. Fernandez, P. Mithraratne, S. F. Thrupp, M. H. Tawhai, and P. J. Hunter, "Anatomically based geometric modelling of the musculo-skeletal system and other organs," *Biomechanics and Modeling in Mechanobiology*, 2004. DOI: 10.1007/s10237-003-0036-1.

21. M. Bucki, C. Lobos, and Y. Payan, "A fast and robust patient specific Finite Element mesh registration technique: Application to 60 clinical cases," *Medical Image Analysis*, vol. 14, no. 3, pp. 303–317, 2010. DOI: 10.1016/j.media.2010.02.003.

22. L. Grassi, N. Hraiech, E. Schileo, M. Ansaloni, M. Rochette, and M. Viceconti, "Evaluation of the generality and accuracy of a new mesh morphing procedure for the human femur," *Medical Engineering and Physics*, 2011. DOI: 10.1016/j.medengphy.2010.09.014.

23. D. C. Barber, E. Oubel, A. F. Frangi, and D. R. Hose, "Efficient computational fluid dynamics mesh generation by image registration," *Medical Image Analysis*, 2007. DOI: 10.1016/j.media.2007.06.011.

24. A. Bijar, P. Y. Rohan, P. Perrier, and Y. Payan, "Atlas-Based Automatic Generation of Subject-Specific Finite Element Tongue Meshes," *Annals of Biomedical Engineering*, vol. 44, no. 1, pp. 16–34, 2016. DOI: 10.1007/s10439-015-1497-y.

25. M. Nesme, P. G. Kry, L. Jeřábková, and F. Faure, "Preserving topology and elasticity for embedded deformable models," in *ACM Transactions on Graphics*, 2009. DOI: 10.1145/1531326.1531358.

26. C. A. Sánchez, Z. Li, A. G. Hannam, P. Abolmaesumi, A. Agur, and S. Fels, "Constructing Detailed Subject-Specific Models of the Human Masseter BT - Imaging for Patient-Customized Simulations and Systems for Point-of-Care Ultrasound," M. J. Cardoso, T. Arbel, J. M. R. S. Tavares, S. Aylward, S. Li, E. Boctor, G. Fichtinger, K. Cleary, B. Freeman, L. Kohli, D. Shipley Kane, M. Oetgen, and S. Pujol, Eds., Cham: Springer International Publishing, 2017, pp. 52–60.

27. C. C. Van Donkelaar, L. J. Kretzers, P. H. Bovendeerd, L. M. Lataster, K. Nicolay, J. D. Janssen, M. R. Drost, and V. Leeuwen, "Diffusion tensor imaging in biomechanical studies of skeletal muscle function," *J Anat*, vol. 194, no. 1, pp. 79–88, 1999.

28. P. J. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and a. Aldroubi, "In vivo fiber tractography using DT-MRI data.," *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, vol. 44, no. 4, pp. 625–632, Oct. 2000.

29. V. J. Napadow, Q. Chen, V. Mai, P. T. So, and R. J. Gilbert, "Quantitative analysis of three-dimensional-resolved fiber architecture in heterogeneous skeletal muscle tissue using nmr and optical imaging methods.," *Biophysical journal*, vol. 80, no. 6, pp. 2968–2975, 2001. DOI: 10.1016/S0006-3495(01)76262-5.

30. H. Shinagawa, E. Z. Murano, J. Zhuo, R. P. Gullapalli, B. Landman, J. L. Prince, and M. Stone, "Tongue muscle fiber tracking during tongue protrusion and rest," *The Journal of the Acoustical Society of America*, vol. 120, no. 5, pp. 3354–3354, Nov. 2006. DOI: 10.1121/1.4781440.

31. A. M. Heemskerk, T. K. Sinha, K. J. Wilson, Z. Ding, and B. M. Damon, "Repeatability of DTI-based skeletal muscle fiber tracking," *NMR in Biomedicine*, 2010. DOI: 10.1002/nbm.1463.

32. C. Ye, A. Carass, E. Murano, M. Stone, and J. L. Prince, "A bayesian approach to distinguishing interdigitated muscles in the tongue from limited diffusion weighted imaging," *Bayesian Graph Models Biomed Imaging*, vol. 8677, no. around 30, pp. 13–24, 2014. DOI: 10.1007/978-3-319-12289-2.

33. S. M. Mijailovich, B. Stojanovic, M. Kojic, A. Liang, V. J. Wedeen, and R. J. Gilbert, "Derivation of a finite-element model of lingual deformation during swallowing from the mechanics of mesoscale myofiber tracts obtained by MRI.," *Journal of applied physiology (Bethesda, Md. : 1985)*, vol. 109, no. 1985, pp. 1500–1514, 2010. DOI: 10.1152/japplphysiol.00493.2010.

34. L. Voskuilen, M. Italiaander, P. de Heer, A. J. Balm, F. van der Heijden, G. J. Strijkers, L. E. Smeele, and A. J. Nederveen, "A 12-channel flexible receive coil for accelerated tongue imaging," in *Proceedings of the International Society for Magnetic Resonance in Medicine*, 2019.

35. L. Voskuilen, L. E. Smeele, A. J. M. Balm, F. V. D. Heijden, G. J. Strijkers, and A. J. Nederveen, "Generation of a muscle fibre orientation atlas of the in vivo tongue," in *Proceedings of the International Society for Magnetic Resonance in Medicine*, Paris, 2018.

36. K. D. Kappert, M. J. van Alphen, L. E. Smeele, A. J. Balm, and F. van der Heijden, "Quantification of tongue mobility impairment using optical tracking in patients after receiving primary surgery or chemoradiation," *PLoS ONE*, 2019. DOI: 10.1371/journal.pone.0221593.

37. J. Veraart, D. S. Novikov, D. Christiaens, B. Ades-aron, J. Sijbers, and E. Fieremans, "Denoising of diffusion MRI using random matrix theory," *NeuroImage*, vol. 142, pp. 1–28, 2016. DOI: 10.1016/j.neuroimage.2016.08.016.

38. S. M. Smith, M. Jenkinson, M. W. Woolrich, C. F. Beckmann, T. E. J. Behrens, H. Johansen-Berg, P. R. Bannister, M. De Luca, I. Drobnjak, D. E. Flitney, R. K. Niazy, J. Saunders, J. Vickers, Y. Zhang, N. De Stefano, J. M. Brady, and P. M. Matthews, "Advances in functional and structural MR image analysis and implementation as FSL," *NeuroImage*, vol. 23, no. SUPPL. 1, pp. 208–219, 2004. DOI: 10.1016/j.neuroimage.2004.07.051.

39. J. L. R. Andersson and S. N. Sotiropoulos, "An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging," *NeuroImage*, vol. 125, pp. 1063–1078, 2016. DOI: 10.1016/j.neuroimage.2015.10.019.

40. P. a. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability.," *NeuroImage*, vol. 31, no. 3, pp. 1116–28, Jul. 2006. DOI: 10.1016/j.neuroimage.2006.01.015. [Online]. Available: www.itksnap.org.

41. J. D. Tournier, F. Calamante, and A. Connelly, "MRtrix: Diffusion tractography in crossing fiber regions," *International Journal of Imaging Systems and Technology*, vol. 22, no. 1, pp. 53–66, 2012. DOI: 10.1002/ima.22005.

42. J. D. Tournier, F. Calamante, and A. Connelly, "Determination of the appropriate b value and number of gradient directions for high-angular-resolution diffusion-weighted imaging," *NMR in Biomedicine*, vol. 26, no. 12, pp. 1775–1786, 2013. DOI: 10.1002/nbm.3017.

43. J.-D. Tournier, F. Calamante, and A. Connelly, "Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution.," *NeuroImage*, vol. 35, pp. 1459–1472, May 2007. DOI: 10.1016/j.neuroimage.2007.02.016.

44. D. Raffelt, J. D. Tournier, J. Fripp, S. Crozier, A. Connelly, and O. Salvado, "Symmetric diffeomorphic registration of fibre orientation distributions," *NeuroImage*, vol. 56, no. 3, pp. 1171–1180, 2011. DOI: 10.1016/j.neuroimage.2011.02.014.

45. D. Raffelt, J. D. Tournier, S. Crozier, A. Connelly, and O. Salvado, "Reorientation of fiber orientation distributions using apodized point spread functions," *Magnetic Resonance in Medicine*, vol. 67, no. 3, pp. 844–855, 2012. DOI: 10.1002/mrm.23058.

46. J.-D. Tournier, F. Calamante, and A. Connelly, "Improved probabilistic streamlines tractography by 2nd order integration over fibre orientation distributions," *Proceedings of the International Society for Magnetic Resonance in Medicine*, vol. 88, no. 2003, p. 2010, 2010.

47. R. Wang and V. J. Wedeen, *TrackVis*, 2015. [Online]. Available: http://trackvis.org/.

48. P. Cignoni, P. Cignoni, M. Callieri, M. Callieri, M. Corsini, M. Corsini, M. Dellepiane, M. Dellepiane, F. Ganovelli, F. Ganovelli, G. Ranzuglia, and G. Ranzuglia, "MeshLab: an Open-Source Mesh Processing Tool," *Sixth Eurographics Italian Chapter Conference*, pp. 129–136, 2008. DOI: 10.2312/LocalChapterEvents/ItalChap/ItalianChapConf2008/129-136.

49. J. Vollmer, R. Mencl, and H. Muller, "Improved Laplacian Smoothing of Noisy Surface Meshes," *Computer Graphics Forum*, vol. 18, no. 3, pp. 131–138, Sep. 1999. DOI: 10.1111/1467-8659.00334.

50. J. E. Lloyd, I. Stavness, and S. Fels, "ArtiSynth: A Fast Interactive Biomechanical Modeling Toolkit Combining Multibody and Finite Element Simulation," *Soft Tissue Biomechanical Modeling for Computer Assisted Surgery*, pp. 355–394, 2012. DOI: 10.1007/8415_2012_126.

51. S. Buchaillard, P. Perrier, and Y. Payan, "A biomechanical model of cardinal vowel production: muscle activations and the impact of gravity on tongue positioning.," *The Journal of the Acoustical Society of America*, vol. 126, no. 4, pp. 2033–2051, 2009. DOI: 10.1121/1.3204306.

52. J. M. Gerard, J. Ohayon, V. Luboz, P. Perrier, and Y. Payan, "Non-linear elastic properties of the lingual and facial tissues assessed by indentation technique: Application to the biomechanics of speech production," *Medical Engineering and Physics*, vol. 27, no. 10, pp. 884–892, 2005. DOI: 10.1016/j.medengphy.2005.08.001.

53. I. Stavness, J. E. Lloyd, and S. Fels, "Automatic prediction of tongue muscle activations using a finite element model," *Journal of Biomechanics*, vol. 45, no. 16, pp. 2841–2848, 2012. DOI: 10.1016/j.jbiomech.2012.08.031.

54. S. S. Blemker, P. M. Pinsky, and S. L. Delp, "A 3D model of muscle reveals the causes of nonuniform strains in the biceps brachii," *Journal of Biomechanics*, vol. 38, no. 4, pp. 657–665, 2005. DOI: 10.1016/j.jbiomech.2004.04.009.

55. V. J. Napadow, Q. Chen, V. J. Wedeen, and R. J. Gilbert, "Intramural mechanics of the human tongue in association with physiological deformations," *Journal of Biomechanics*, vol. 32, no. 1, pp. 1–12, 1999. DOI: 10.1016/S0021-9290(98)00109-2.

56. S. Dabbaghchian, M. Arnela, O. Engwall, O. Guasch, I. Stavness, and P. Badin, "Using a Biomechanical Model and Articulatory Data for the Numerical Production of Vowels," in *INTERSPEECH*, vol. m, 2016, pp. 3569–3573. DOI: 10.21437/Interspeech.2016-1500.

57. N. M. Harandi, J. Woo, M. Stone, R. Abugharbieh, and S. Fels, "Subject-specific biomechanical modelling of the tongue: analysis of muscle activations during speech," *Proc Int Semin Speech Prod (ISSP)*, no. 2009, pp. 174–177, 2014.

58. X. Wu, J. Dang, and I. Stavness, "Iterative method to estimate muscle activation with a physiological articulatory model," *Acoustical Science and Technology*, vol. 35, no. 4, pp. 201–212, 2014. DOI: 10.1250/ast.35.201.

59. K. Honda, E. Z. E. Murano, S. Takano, S. Masaki, and J. Dang, "Anatomical considerations on the extrinsic tongue muscles for articulatory modeling," *Proceedings of Meetings on Acoustics*, vol. 19, no. 1, p. 60 270, 2013. DOI: 10.1121/1.4800262.

60. D. Raffelt, J. D. Tournier, S. Rose, G. R. Ridgway, R. Henderson, S. Crozier, O. Salvado, and A. Connelly, "Apparent Fibre Density: A novel measure for the analysis of diffusion-weighted magnetic resonance images," *NeuroImage*, vol. 59, no. 4, pp. 3976–3994, 2012. DOI: 10.1016/j.neuroimage.2011.10.045.

61. J. L. Miller, K. L. Watkin, and M. F. Chen, "Muscle, adipose, and connective tissue variations in intrinsic musculature of the adult human tongue.," *Journal of Speech, Language and Hearing Research*, vol. 45, no. 1, pp. 51–65, 2002. DOI: 10.1044/1092-4388(2002/004).

5

Chapter 5

A 12-channel flexible receiver coil for accelerated tongue imaging

Luuk Voskuilen, Paul de Heer, Lisette van der Molen, Alfons J. M. Balm, Ferdinand van der Heijden, Gustav J. Strijkers, Ludi E. Smeele and Aart J. Nederveen

Magnetic Resonance Materials in Physics, Biology and Medicine

2020

Abstract

Objective: MRI of the tongue requires acceleration to minimise motion artefacts and to facilitate real-time imaging of swallowing. To accelerate tongue MRI, we designed a dedicated flexible receiver coil.

Materials and Methods: We designed a flexible 12-channel receiver coil for tongue MRI at 3 T and compared it to a conventional head-and-neck coil regarding SNR and g-factor. Furthermore, two accelerated imaging techniques were evaluated using both coils: multiband (MB) diffusion-tensor imaging (DTI), and real-time MRI of swallowing.

Results: The flexible coil had significantly higher SNR in the anterior (2.1 times higher, $P = 0.002$) and posterior (2.0 times higher, $P < 0.001$) parts of the tongue, while the g-factor was lower at higher acceleration. Unlike for the flexible coil, the apparent diffusion coefficient ($P = 0.001$) and fractional anisotropy ($P = 0.008$) deteriorated significantly while using the conventional coil after accelerating DTI with MB. The image quality of real-time MRI of swallowing was significantly better for hyoid elevation ($P = 0.029$) using the flexible coil.

Conclusion: Facilitated by higher SNR and lower g-factor values, our flexible tongue coil allows faster imaging, which was successfully demonstrated in MB DTI and real-time MRI of swallowing.

Introduction

Accelerated imaging is an intensively researched field in MRI over the past two decades, and resulted in techniques such as parallel imaging[1], multiband imaging[2], and compressed sensing[3]. As acceleration reduces the total scan time, the incidence of motion artefacts decreases. In approximately 20% of all MRI examinations, an imaging sequence has to be repeated due to motion artefacts[4]. Not only have these artefacts been estimated to cost the hospital $115,000 per scanner per year[4], they may also prevent a radiologist from determining the correct diagnosis.

MRI of the tongue may especially benefit from acceleration, as it is prone to motion artefacts due to breathing and swallowing. Besides motion artefact reduction, acceleration techniques have recently been used to develop new imaging protocols, such as real-time dynamic imaging of speech[5] and swallowing[6]. Additionally, acceleration may also improve the clinical feasibility of imaging techniques that currently take too much time for routine clinical examinations, such as high-angular resolution diffusion imaging of the tongue[7].

In general, acceleration reduces the signal-to-noise ratio (SNR). In parallel imaging, SNR loss is related to the acceleration factor and a geometry-specific noise-amplification factor (g-factor). In the recently developed multiband (MB) imaging approach, several slices are acquired simultaneously. The SNR loss due to acceleration is therefore no longer dependent on the acceleration factor, but only on the g-factor[2]. In another acceleration technique, compressed sensing, the relationship between acceleration and SNR loss is not as straight-forward, as compressed sensing basically acts as a denoiser[3]. Nevertheless, improving SNR has been attributed to milder artefacts and thus better image quality in real-time imaging[8]. In summary, in order to maintain image quality after acceleration in tongue MRI, we hypothesize that the SNR should be increased and the g-factor reduced.

The SNR and g-factor can both be improved by designing a receiver coil specifically for the tongue, as the SNR can be increased by using small surface-coil elements that have an intrinsically higher SNR, and the g-factor can be reduced by increasing the density of coil elements. Consequently, such a custom coil allows for higher acceleration than conventional coils, as has been proven for the breast[9], heart[10], and upper airway at 1.5 T[8] and 3 T[11].

These coil designs include either a rigid housing[8–10] or a rigid support structure[11] to improve the robustness of the coil. A rigid coil however prevents the coil elements from being placed closer to the subject, which decreases the maximal SNR that could be achieved. Additionally, these rigid coils may restrict a subject's mobility or the delivery of a contrast agent, which makes the study of swallowing of MRI more difficult. A flexible array of small coil elements may therefore not only allow the subject to move more freely, but it also improves SNR.

In this study, we designed a flexible receiver coil with 12 coil elements to accelerate tongue MRI at 3 T. We compared the flexible coil to a conventional 16-channel neurovascular coil for SNR and g-factor characteristics. Furthermore, the two coils were compared on two applications: Diffusion-weighted multiband imaging of the tongue, and real-time MRI of swallowing using compressed sensing.

Methods

Volunteers

We included five healthy volunteers (2 female; mean age 27 years, range 23–29 years), who gave written informed consent, according to the regulations of our institution. Volunteers were excluded if they had orthodontic appliances or any contraindications to MRI. The volunteers were scanned in a 3 T Philips Ingenia scanner (Philips Healthcare, Best, Netherlands). Using identical scanning protocols, we scanned the volunteers first with the conventional neurovascular coil, and afterwards with the flexible coil.

Flexible coil design

Twelve copper coil elements (53 by 32 mm) were built according to the design displayed in figure 5.1a (MRcoils, Zaltbommel, Netherlands). For each element, a housing was 3D printed to robustly attach the coaxial cable to the coil element, and also to prevent these cables from coiling. Subsequently, the elements were arranged in phased arrays of three coil elements each (figure 5.1b). Two phased arrays were sewn into a synthetic leather sheet, which improved the robustness of the coil, while maintaining flexibility. The coil was fixed on the head by a Velcro strap around the chin and two flaps, which were placed between the headphones and the subject's head (figure 5.1c). While loaded with a male volunteer, the coil elements were tuned and matched with a network analyser.

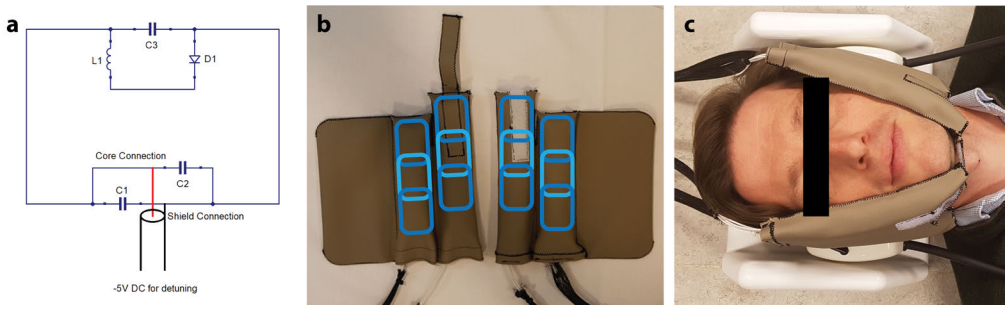


FIGURE 5.1 — Design and lay-out of the flexible coil: a: The schematic of a single coil element with three capacitors (C1-3), one inductor (L1), and one diode (D1). b: Three of these elements were assembled in a phased array indicated by the blue rounded rectangles. c: The coil was placed directly on the cheeks of a subject. The synthetic leather strap and flaps, which were placed between the headphones and the subject's head, kept the coil in position. The white 3D-printed headrest minimised head rotation during scanning.

Quantitative assessment of coil performance

To quantitatively compare the flexible to the conventional coil, a 2D gradient echo sequence was used to acquire images for all five volunteers. The images were also reconstructed for each coil element separately. The imaging parameters were: TR/TE: 6.3/3.2 ms; voxel size: 1.25 x 1.25 x 1.25 mm^3; flip angle: 8°; no parallel imaging; scan time: 52 s. To obtain noise-only images, this sequence was repeated with the gradients and RF transmitter switched off. In order to obtain g-factor maps, we repeated the sequence another five times for the last volunteer only with increasing SENSE acceleration factor along the left-right direction (1.2, 1.5, 2, 3, and 4).

In MATLAB (R2018a, MathWorks, Natick, MA), we calculated the noise correlation matrix from data of a single volunteer, in which noise-only images are correlated between all possible combinations of coil elements[12]. To obtain the mean noise correlation of this matrix, we averaged the non-diagonal elements of this noise correlation matrix.

SNR maps were calculated by dividing the signal in a voxel by the standard deviation of the noise within a radius of three voxels surrounding that voxel, compensating for the Rician distribution of this noise. In 3D Slicer[13], we manually segmented four regions: the masseter muscles, the parotid glands, and the anterior and posterior parts of the tongue, which were created by dividing the segmentation of the entire tongue into two roughly equally large parts using a coronal plane. For every region, the SNR gain was determined by dividing the average SNR of the flexible coil with the average SNR of the conventional coil. The average SNR gain was subsequently calculated by averaging over the five volunteers.

Using the vendor's reconstruction software, we created g-factor maps for various acceleration factors for a single volunteer. To calculate the average g-factor in the head, we created a mask by applying a threshold on the magnitude images and subsequently smoothening the mask with morphological operations. By combining the regions of the anterior and posterior parts of the tongue, which were also used for the SNR gain calculation, we determined the average g-factor in the tongue.

Diffusion-tensor Multiband Imaging

For all five volunteers, diffusion-tensor imaging (DTI) of the tongue was accelerated using a SE-EPI sequence with multiband (MB) SENSE[2]. For MB factor 1 (no MB acceleration), the imaging parameters were TR/TE: 2622/67 ms; scan time: 23.6 s. For MB factor 2, these imaging parameters were changed to TR/TE: 1405/71 ms; scan time: 12.6 s. Other imaging parameters that were the same for both acquisitions were: ETL: 35; matrix size: 64 x 64; voxel size: 3 x 3 x 3 mm^3; in-slice SENSE factor: 1.5; no partial Fourier imaging; NSA: 1; SPIR and SSGR fat suppression; b-value: 700 s/mm^2 along 6 different directions. An

region-of-interest was manually drawn in the genioglossus muscle, in which the average apparent diffusion coefficient (ADC) and fractional anisotropy (FA) were calculated using MRtrix3[14].

Real-time MRI of swallowing

Single-slice midsagittal MRI scans of swallowing were acquired using a single-slice golden angle radial GRE sequence[15] with the following parameters: TR/TE: 2.9/1.12 ms; flip angle: 10°; TFE factor: 22; matrix size: 128 x 128; voxel size: 2 x 2 x 6 mm^3; golden angle: 111.25°; scan time: 10 s. Volunteers were asked to swallow 10 mL of pineapple juice, which is a natural T1 contrast agent due to the presence of manganese[16].

The images were reconstructed off-line using the Berkeley Advanced Reconstruction Toolbox[17] and MATLAB. Because we assumed that the sensitivity maps were constant over time, we estimated these sensitivity maps using eSPIRIT[18] from low resolution images created from all spokes. Subsequently, we binned eight spokes in each frame (without using a sliding window approach), resulting in 43 frames per second. The images were reconstructed with compressed sensing using a locally low-rank constraint in space[19] (with regularisation parameter $\lambda = 0.005$) and a total variation constraint over time[15] (with $\lambda = 0.01$). The regularisation parameters were determined empirically based on the images from the flexible coil. A median filter over time was used to remove residual radial streaking artefacts[20], and a non-local means filter to further suppress noise[21].

Image grading

The movies of the real-time MRI of swallowing were graded by four speech therapists, who had between seven and twenty years of experience in reviewing videofluoroscopic swallowing studies. These ten movies, five for each coil, were presented to the reviewers in a random order blinded to these reviewers. We assessed three aspects of the swallowing movies, namely the visualisation of swallowing, the presence of motion artefacts, and the overall quality of the movies. For the visualisation of swallowing, we composed five questions distributed over the oral and pharyngeal phases of normal swallowing[22]. These five questions were: 'How well is the labial seal visualised?'; 'How well is the contact of the tongue and alveolar ridge visualised?'; 'How well is the velopharyngeal closure visualised?'; 'How well is the hyoid elevation visualised?'; 'How well is the contraction of tongue base and pharyngeal wall visualised?'. For the effect of motion artefacts on the image quality, we composed the following three questions: 'To what extent does radial streaking affect the quality of the movie?'; 'To what extent does inhomogeneity of intensity affect the quality of the movie?'; 'To what extent does motion blurring affect the quality of the movie?'. For the overall image quality, we composed only

one question: 'How do you grade the overall quality of the movie?'. The reviewers answered each question using a five-point Likert scale, in which 1 represented the lowest image quality (or most artefacts) and 5 the highest image quality (or fewest artefacts).

Statistical analysis

For the four regions (masseter muscles, parotid glands, and anterior and posterior parts of the tongue), the difference in average SNR between both coils over all five volunteers was tested using a paired t-test. For both coils separately, we tested the difference in average ADC and FA values between MB factor 1 and 2 over all five volunteers with a paired t-test. For every question individually, we modelled the grades provided by the speech therapists by a multilevel proportional-odds model using the ordinal package in R[23]. In this model, we included one fixed effect, the used coil (conventional or flexible), and two random effects, the subject and the speech therapist. Both random effects were modelled with random intercepts and random slopes with the fixed effect (the used coil). The P-values of the effect that the used coil had on the grades were calculated by likelihood-ratio tests of the full model compared to the model without the fixed effect (the coil used).

Results

The correlations in the noise-only images between all possible combinations of coil elements are visualised as a correlation matrix for the conventional coil (figure 5.2a) and the flexible coil (figure 5.2b). The mean noise correlation excluding the diagonal elements of the matrix was 5.9% for the conventional coil and 2.3% for the flexible coil. In the magnitude images of the coil elements (figure 5.2c), localised areas with high signal intensity can be appreciated, which would be beneficial to the parallel imaging performance of the coil. In this slice, the image intensity of four coil elements was lower, because these elements are sensitive to a region outside this slice. In the image of all elements combined (figure 5.2d), focal hyperintensities were present near the location of the coil elements.

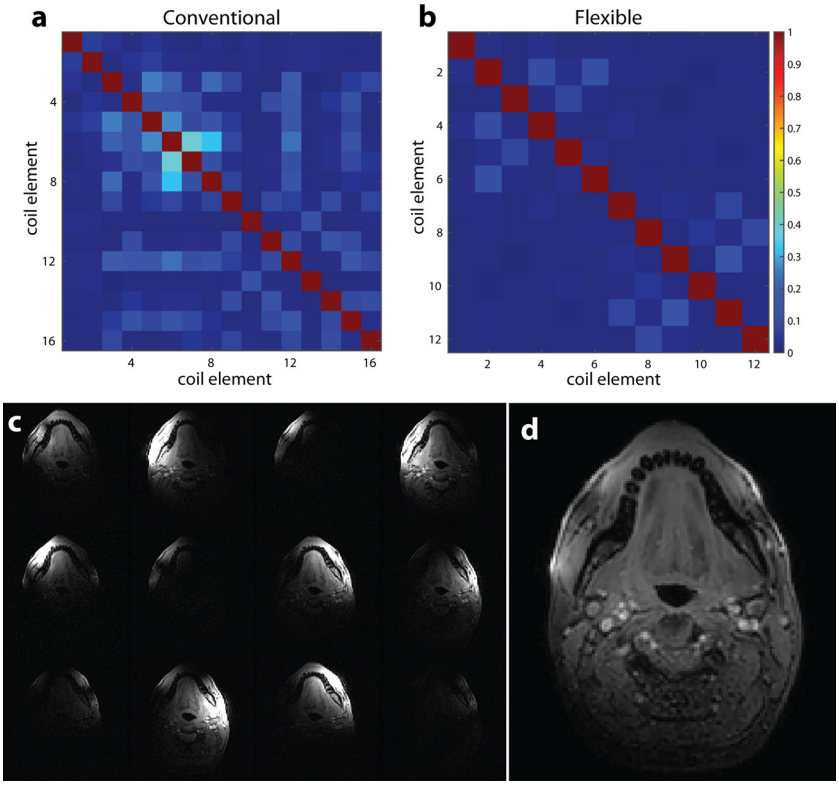


FIGURE 5.2 — Signal and noise characteristics of the flexible coil: The noise correlation matrix of (a) the conventional and (b) the flexible coil for the first volunteer. For both coils, little coupling was present between coil elements, as the average non-diagonal elements of the noise correlation matrices was 2.3% for the flexible coil and 5.9% for the conventional coil. c: The images of the same slice acquired by different coil elements confirmed that the small coil elements are sensitive to a small volume. A few images, such as the lower right-most image in c, had a low average intensity because the corresponding coil element was located further from the imaged slice. d: Combined magnitude image from all coil elements. Although the signal is generally fairly homogeneous, the image shows a few focal hyperintensities located close to the coil elements.

In figure 5.3a–b the SNR maps are shown of both coils for a single volunteer. In the cheeks, where the coil elements are positioned, the SNR improvement over the conventional coil was the largest. Also in the tongue, the SNR was higher in the flexible coil. In figure 5.3c, a representative slice is presented containing the four regions that we manually segmented. Over all volunteers, the flexible coil array had a significant SNR gain in the following regions: an SNR gain of 2.1 in the anterior part of the tongue ($P = 0.002$), 2.0 in the posterior part of the tongue ($P < 0.001$) and 6.1 in the masseter muscles ($P < 0.001$). The SNR gain of 1.5 in the parotid glands was not significant ($P = 0.171$).

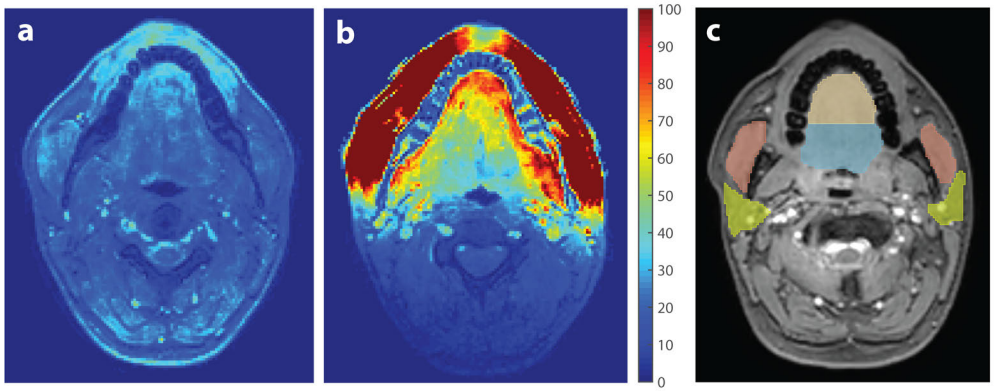


FIGURE 5.3 — Signal-to-noise ratio evaluation. SNR maps of a similar slice acquired by (a) the conventional neurovascular coil and (b) the flexible tongue coil of one volunteer. The SNR is higher in the images from the flexible coil, especially in the cheeks. Also in our primary area of interest, the tongue, the SNR is higher. c: The images were manually segmented into four regions, i.e. the anterior part of the tongue (brown), the posterior part of the tongue (blue), the masseter muscles (red), and the parotid glands (yellow).

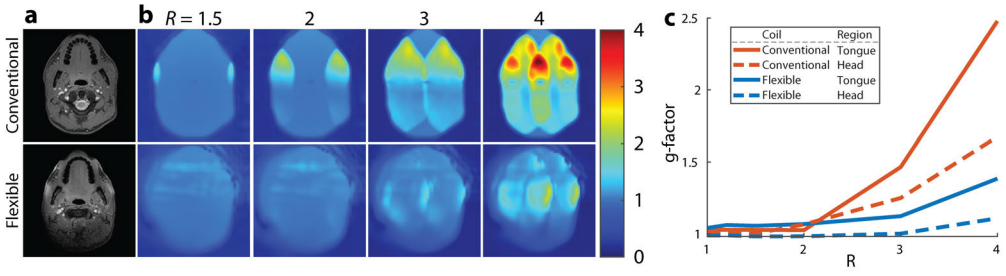


FIGURE 5.4 — G-factor characteristics of the coil: a: A magnitude image without acceleration is given as anatomical reference. b: For one volunteer, g-factor maps of the conventional coil and the flexible coil were created by increasing SENSE acceleration (R) in the left-right direction. The g-factor maps at $R = 1$ and $R = 1.2$ were not shown in this figure as these differed little from the g-factor map at $R = 1.5$. c: Averaged g-factors over the head or the tongue are displayed for both coils. At low acceleration factors ($R < 2$), the conventional and flexible coils differed only slightly. At higher acceleration however, the g-factor increased more for the conventional coil than for the flexible coil.

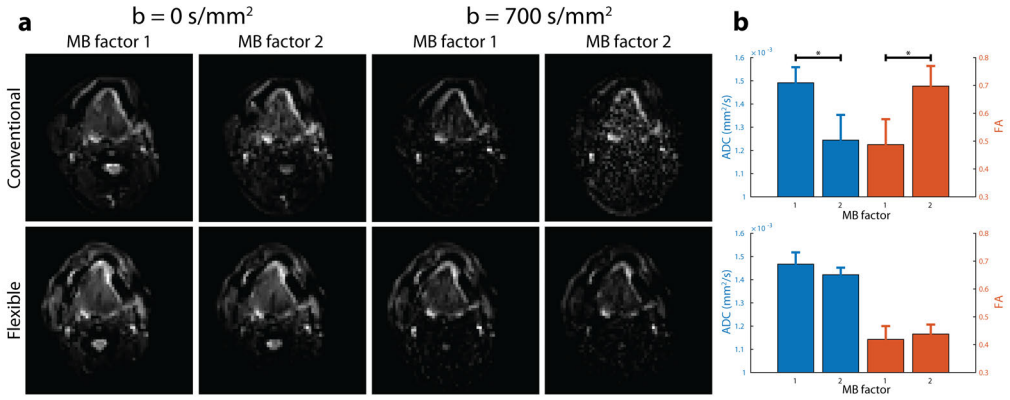


FIGURE 5.5 — Diffusion-tensor imaging accelerated with multiband-SENSE. a: For one volunteer, a similar slice is displayed for the conventional and flexible coil, with and without diffusion weighting, and using a MB factor of 1 or 2 (a). Although the noise amplification caused by MB-SENSE was low in the diffusion-weighted images from the flexible coil, this noise obscured the diffusion-weighted images acquired with the conventional coil. b: The average and standard deviation of two DTI metrics, ADC and FA over the five volunteers were calculated within a region-of-interest in the genioglossus muscle. For the conventional coil, significant differences were found between MB factor 1 and 2 for both the ADC ($P = 0.001$) and the FA ($P = 0.008$). No such significant differences were found for the flexible coil.

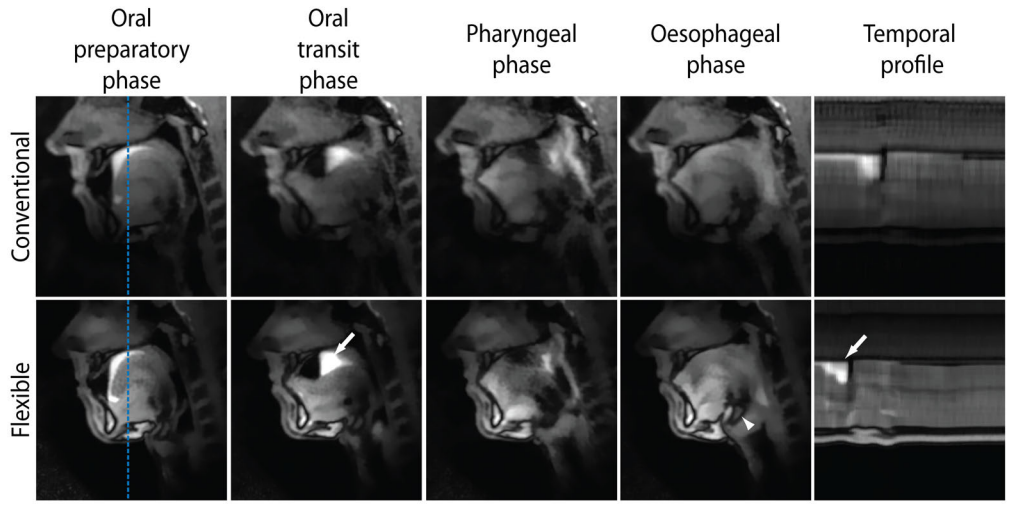


FIGURE 5.6 — For the conventional and flexible coil, real-time MRI of swallowing at 43 frames per second is displayed for one volunteer. Four swallowing phases were distinguished: The oral preparatory phase, i.e. the oral transit phase, the pharyngeal phase, and the oesophageal phase. In the rightmost panels, the temporal profile of a single line, indicated by the blue dashed line, is shown. The white arrows, in the panels of the oral transit phase and the temporal profile for the flexible coil, indicate the swallowing contrast agent (pineapple juice). The white arrowhead, in the panel for the flexible coil in the oesophageal phase, indicates the hyoid, which could be significantly better visualised with the flexible coil ($P = 0.029$). The full movies of this volunteer are available for the conventional coil (supplementary material 5.1) and the flexible coil (supplementary material 5.2).

The g-factor evolution by increasing the SENSE acceleration is displayed in figure 5.4. Although the g-factor maps at low acceleration factors ($R < 2$) are similar, at higher acceleration factors ($R \geq 2$), the g-factor increased more for the conventional coil than for the flexible coil. By averaging the g-factor over either the head or the tongue (figure 5.4c), a similar pattern could be appreciated, in which the g-factor increased more for the conventional coil.

In figure 5.5, we show that diffusion-weighted imaging could be accelerated with MB-SENSE (MB factor 2) without unfolding artefacts. Without diffusion weighting, images accelerated with MB-SENSE appeared to be similar to images without MB-SENSE for both coils. With diffusion weighting however, the noise obscured the accelerated images from the conventional coil, while the accelerated images from the flexible coil were still similar to the non-accelerated images. For the flexible coil, no statistically significant differences were found between MB factor 1 and two for ADC ($P = 0.196$) and FA ($P = 0.419$). For the conventional coil however, using MB acceleration significantly changed both ADC ($P = 0.001$) and FA ($P = 0.008$).

For the flexible and conventional coils, four frames from the real-time MRI of swallowing are displayed in figure 5.6. In these frames, four swallowing phases could be distinguished for both coils. The epiglottis and hyoid appeared to be visualised more sharply for the flexible coil. In the temporal profile, the contrast agent also appears to be visualised more sharply with the flexible coil. The differences between the conventional coil (supplementary material 5.1) and the flexible coil (supplementary material 5.2) may be better appreciated in the full movies of the real-time MRI of swallowing.

The proportion of grades given by the speech therapists for either the conventional coil or the flexible coil are presented in figure 5.7. For all questions asked except for the question about hyoid elevation, there were no statistically significant differences in odds between the coils. For the visualisation of the hyoid elevation, the odds of being graded better were significantly higher for the flexible coil ($P = 0.029$).

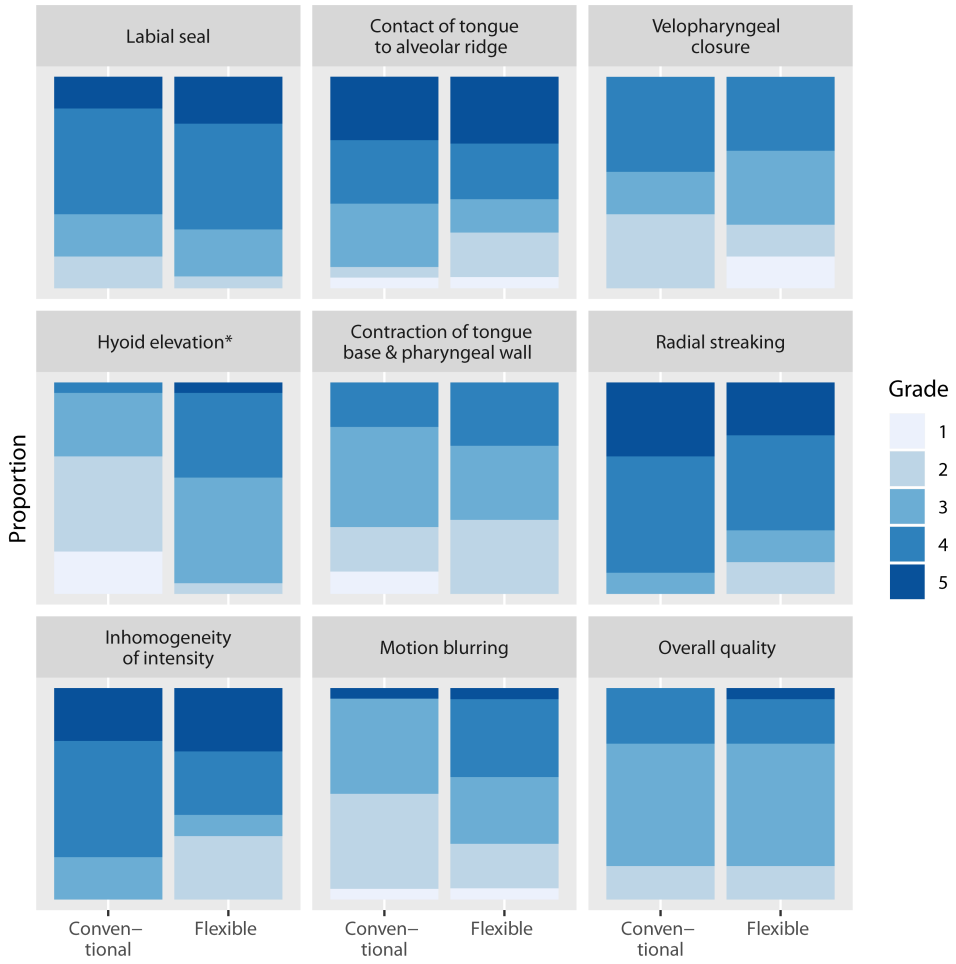


FIGURE 5.7 — The proportions of grades given by the reviewers for the nine questions about the real-time MRI of swallowing. For this figure only, the grades for the five different subjects and by the four reviewers have been pooled. The grade 5 related to the best image quality or least artefacts, while 1 related to the worst image quality or most artefacts. For the visualisation of hyoid elevation, the odds of receiving a better grade was significantly higher for the flexible coil ($P = 0.029$) than for the conventional coil. For the other aspects of real-time MRI of swallowing, no differences in odds were found.

Discussion

In this study, we designed a flexible receiver coil for tongue MRI at 3 T and evaluated its performance. Compared to a conventional coil, our flexible coil exhibits higher SNR in the tongue and lower g-factor values at high acceleration factors. For DTI of the tongue, the flexible coil facilitates faster scanning using MB-SENSE. For real-time MRI of swallowing using compressed sensing, similar image quality was found for both coils, except for the visualisation of hyoid elevation where the flexible coil performed better.

Although we found higher SNR values in the tongue with the flexible coil, the SNR rapidly decreases further away from the coil. In the parotid glands, the mean SNR no longer differed significantly between the two coils. In the back of the neck, the SNR appeared to be even lower for the flexible coil. Thus, imaging with the flexible coil is restricted to a smaller field-of-view than with the conventional coil. However, this field-of-view can easily be extended by using additional coil elements from a torso coil or the coils in the patient bed depending on the application. Another option would be to redesign the flexible coil with a higher number of coil elements. However, care should to taken to maintain the flexibility of the coil.

The flexibility of the presented tongue coil leads to the following three advantages: The coil should be more comfortable for the subject, the coil elements are closer to the volume of interest, which increases SNR, and the flexibility of the coil allows us to use auxiliary equipment during scanning, which means that we could easily administer the swallowing contrast agent. The access to auxiliary equipment also facilitates the implementation of other study protocols, such as concurrent manometry and MRI for patients with dysphagia[6], or the evaluation of obstructive sleep apnoea with MRI while patients are wearing a facial mask[24].

However, the flexibility also allows the subject to move more during the scans, which may result in more motion artefacts than with conventional coils. To prevent head rotation during scans, we therefore made a 3D printed headrest. More subtle motion originating from breathing and swallowing may still result in motion artefacts, but these artefacts are expected to occur less often due to acceleration with techniques such as MB imaging.

Other custom receiver coils for tongue or upper airway imaging have been reported for both 1.5[8] and 3 T[11]. Although these coils appear to be more rigid than the flexible coil presented in this paper, their reported SNR gains are higher, i.e. SNR gains in the tongue of between 2.6 and 5.5 for the coil at 3T. These differences in SNR gain may be explained by the different conventional coils used as a reference (an 8-channel coil compared to a 16-channel coil in this study).

Using the flexible coil, we were the first to show the feasibility of MB acceleration in the head and neck area, as we were able to accelerate DTI by a MB factor of 2, without significantly affecting the ADC or FA. With the conventional coil however, the ADC was significantly lower and the FA was significantly higher

between MB factor 1 and 2, which is consistent with the effect that a low SNR level has on these metrics[25]. Using MB imaging, we can now reduce the scan time high-angular resolution diffusion-weighted imaging of the tongue[7], which would otherwise require too much scan time in a clinical setting. Other sequences that rely on an EPI read-out such as arterial spin labelling in the parotid glands[26], could also benefit from MB imaging. Multiband imaging can even be combined with a radial acquisition[27], which we used for the real-time MRI of swallowing.

We showed that the SNR gain from the flexible coil only modestly improves the image quality of real-time imaging of swallowing. Although we measured a two-times higher SNR and a lower g-factor at high acceleration factors for the flexible coil, no difference in image quality was found in the real-time MRI of swallowing, except for hyoid elevation. For the flexible coil, the temporal profile is less blurred than for the conventional coil, but no significant difference in motion blurring was detected. As the reviewers had not graded real-time MRI of swallowing before, we might have detected smaller differences in image quality by training the reviewers more extensively. Nevertheless, a radial acquisition in combination with compressed sensing reconstruction is apparently effective in restoring image quality in undersampled data. To benefit more from the improved SNR of the flexible coil, we believe that the in-plane resolution should be increased or that multiple slices should be acquired. The latter is especially useful in cases where aspiration may be missed, if it is located laterally from a single midsagittal slice.

Real-time MRI of swallowing was previously demonstrated to be feasible for a frame rate of 24.3 frames per second[16]. Although we attained a higher frame rate, these previous studies were able to reconstruct the movies in real time using the regularised non-linear inversion method[28]. The advantage of such a real-time reconstruction method is that images can be made available for the examiner at the scanner immediately. In this way, the acquisition may be adjusted or repeated as necessary, which may help the integration of real-time MRI of swallowing into clinical practice.

The main limitation of the flexible coil are the focal hyperintensities near the coil elements. These are caused by the signal profile of the small receive elements that has a steep signal drop-off the further you move from the element. Although these hyperintensities did not affect the image quality in the sagittal images of swallowing graded by speech therapists, they were more clearly present in transverse images. These inhomogeneities were also reported in other flexible receiver coils[29]. We tested several correction methods, such as normalisation using the coil sensitivity maps, but we could not find a method that sufficiently homogenised the images. Until such a method has been developed, it is possible to reduce the inhomogeneity by placing the coil slightly further from the skin, but this will reduce the SNR.

In conclusion, the 12-channel flexible coil provides a higher average SNR and lower g-factor in the tongue, which enable higher image acceleration than a conventional neurovascular coil. This acceleration may not only reduce the incidence

of motion artefacts, but also allows the application highly-accelerated MRI protocols such as multiband diffusion tensor imaging and real-time MRI of swallowing at 43 frames per second.

Acknowledgements

We would like to thank the speech therapists at the Antoni van Leeuwenhoek hospital for reviewing the real-time MRI of swallowing, and Maurits en Anna de Kock Stichting for kindly providing a grant for the purchase of the flexible receiver coil.

Supplementary materials

SUPPLEMENTARY MATERIAL 5.1 — A movie of real-time MRI of swallowing at 43 frames per second acquired with the conventional coil. Pineapple juice was used as a contrast agent. Flickering is clearly noticeable in this movie.

SUPPLEMENTARY MATERIAL 5.2 — A movie of real-time MRI of swallowing at 43 frames per second acquired with the flexible coil. Pineapple juice was used as a contrast agent. Flickering is not as noticeable in this movie. However, the signal intensity is more inhomogeneous across the image.

References

1. K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: Sensitivity Encoding for Fast MRI.," *Magnetic resonance in medicine*, vol. 42, pp. 952–962, 1999.
2. K. Setsompop, B. A. Gagoski, J. R. Polimeni, T. Witzel, V. J. Wedeen, and L. L. Wald, "Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty," *Magnetic Resonance in Medicine*, vol. 67, no. 5, pp. 1210–1224, 2012. DOI: 10.1002/mrm.23097.
3. M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007. DOI: 10.1002/mrm.21391.
4. J. B. Andre, B. W. Bresnahan, M. Mossa-Basha, M. N. Hoff, C. Patrick Smith, Y. Anzai, and W. A. Cohen, "Toward quantifying the prevalence, severity, and cost associated with patient motion during clinical MR examinations," *Journal of the American College of Radiology*, vol. 12, no. 7, pp. 689–695, 2015. DOI: 10.1016/j.jacr.2015.03.007.
5. A. Niebergall, S. Zhang, E. Kunay, G. Keydana, M. Job, M. Uecker, and J. Frahm, "Real-time MRI of speaking at a resolution of 33 ms: Undersampled radial FLASH with nonlinear inverse reconstruction," *Magnetic Resonance in Medicine*, vol. 69, no. 2, pp. 477–485, 2013. DOI: 10.1002/mrm.24276.
6. A. Olthoff, A. A. Joseph, M. Weidenmüller, B. Riley, and J. Frahm, "Real-time MRI of swallowing: intraoral pressure reduction supports larynx elevation," *NMR in Biomedicine*, vol. 29, no. 11, pp. 1–6, 2016. DOI: 10.1002/nbm.3621.
7. L. Voskuilen, M. Italiaander, P. de Heer, A. J. Balm, F. van der Heijden, G. J. Strijkers, L. E. Smeele, and A. J. Nederveen, "A 12-channel flexible receive coil for accelerated tongue imaging," in *Proceedings of the International Society for Magnetic Resonance in Medicine*, 2019.
8. S. G. Lingala, Y. Zhu, Y. Lim, A. Toutios, Y. Ji, W. C. Lo, N. Seiberlich, S. Narayanan, and K. S. Nayak, "Feasibility of through-time spiral generalized autocalibrating partial parallel acquisition for low latency accelerated real-time MRI of speech," *Magnetic Resonance in Medicine*, vol. 00, pp. 1–8, 2017. DOI: 10.1002/mrm.26611.
9. A. N. Nnewihe, T. Grafendorfer, B. L. Daniel, P. Calderon, M. T. Alley, F. Robb, and B. A. Hargreaves, "Custom-fitted 16-channel bilateral breast coil for bidirectional parallel imaging," *Magnetic Resonance in Medicine*, vol. 66, no. 1, pp. 281–289, 2011. DOI: 10.1002/mrm.22771.
10. M. Schmitt, A. Potthast, D. E. Sosnovik, J. R. Polimeni, G. C. Wiggins, C. Triantafyllou, and L. L. Wald, "A 128-channel receive-only cardiac coil for highly accelerated cardiac MRI at 3 tesla," *Magnetic Resonance in Medicine*, vol. 59, no. 6, pp. 1431–1439, 2008. DOI: 10.1002/mrm.21598.
11. Y.-C. C. Kim, C. E. Hayes, S. S. Narayanan, and K. S. Nayak, "A Novel 16-Channel Receive Coil Array for Accelerated Upper Airway MRI at 3 Tesla," *Magn Reson Med*, vol. 65, no. 6, pp. 1711–1717, 2011. DOI: 10.1002/mrm.22742.
12. Q. Zhang, B. F. Coolen, S. van den Berg, G. Kotek, D. S. Rivera, D. W. J. Klomp, G. J. Strijkers, A. J. Nederveen, Q. Z. Id, B. F. Coolen, S. V. D. Berg, G. Kotek, S. Rivera, D. W. J. Klomp, G. J. Strijkers, and A. J. Nederveen, "Comparison of four MR carotid surface coils at 3T," *PLoS ONE*, vol. 14, no. 3, pp. 1–16, 2019. DOI: 10.1371/journal.pone.0213107.
13. A. Fedorov, R. Beichel, J. Kalpathy-cramer, J. Finet, J.-c. Fillion-robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka, J. Buatti, S. Aylward, J. V. Miller, S. Pieper, and R. Kikinis, "3D Slicer as an image computing platform for the Quantitative Imaging Network," *Magnetic Resonance Imaging*, vol. 30, no. 9, pp. 1323–1341, 2012. DOI: 10.1016/j.mri.2012.05.001.
14. J. D. Tournier, F. Calamante, and A. Connelly, "MRtrix: Diffusion tractography in crossing fiber regions," *International Journal of Imaging Systems and Technology*, vol. 22, no. 1, pp. 53–66, 2012. DOI: 10.1002/ima.22005.
15. L. Feng, R. Grimm, K. T. Block, H. Chandarana, S. Kim, J. Xu, L. Axel, D. K. Sodickson, and R. Otazo, "Golden-angle radial sparse parallel MRI: combination of compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric MRI," *Magnetic resonance in medicine*, vol. 72, no. 3, pp. 707–717, 2014. DOI: 10.1002/mrm.24980. arXiv: NIHMS150003.
16. S. Zhang, A. Olthoff, and J. Frahm, "Real-time magnetic resonance imaging of normal swallowing," *Journal of Magnetic Resonance Imaging*, vol. 35, no. 6, pp. 1372–1379, 2012. DOI: 10.1002/jmri.23591.
17. M. Uecker, F. Ong, J. I. Tamir, D. Bahri, P. Virtue, J. Y. Cheng, T. Zhang, and M. Lustig, "Berkeley Advanced Reconstruction Toolbox," in *Proc. Intl. Soc. Mag. Reson. Med*, Toronto, 2015. DOI: 10.5281/zenodo.592960.
18. M. Uecker, P. Lai, M. J. Murphy, P. Virtue, M. Elad, J. M. Pauly, S. S. Vasanawala, and M. Lustig, "ESPIRiT - An eigenvalue approach to autocalibrating parallel MRI: Where SENSE meets GRAPPA," *Magnetic Resonance in Medicine*, vol. 71, no. 3, pp. 990–1001, 2014. DOI: 10.1002/mrm.24751. arXiv: 15334406.
19. R. Otazo, E. Candès, and D. K. Sodickson, "Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components," *Magnetic Resonance in Medicine*, vol. 73, no. 3, pp. 1125–1136, 2015. DOI: 10.1002/mrm.25240. arXiv: NIHMS150003.
20. M. Uecker, S. Zhang, D. Voit, A. Karaus, K. D. Merboldt, and J. Frahm, "Real-time MRI at a resolution of 20 ms," *NMR in Biomedicine*, vol. 23, no. 8, pp. 986–994, 2010. DOI: 10.1002/nbm.1585.
21. A. Buades, B. Coll, and J. M. Morel, "A non-local algorithm for image denoising," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, 2005, pp. 60–65. DOI: 10.1109/CVPR.2005.38.
22. W. J. Dodds, E. T. Steward, and J. A. Logemann, "Physiology and Radiology of the Normal Oral and Pharyngeal Phases of Swallowing," *American journal of roentgenology*, vol. 154, pp. 953–963, 1990.
23. R. H. B. Christensen. (2018). "ordinal - Regression Models for Ordinal Data," [Online]. Available: http://www.cran.r-project.org/package=ordinal/.
24. Y.-C. Kim, R. M. Lebel, Z. Wu, S. L. D. Ward, M. C. K. K. Khoo, and K. S. Nayak, "Real-time 3D magnetic resonance imaging of the pharyngeal airway in sleep apnea," *Magnetic Resonance in Medicine*, vol. 71, no. 4, pp. 1501–1510, 2014. DOI: 10.1002/mrm.24808. arXiv: NIHMS150003.

25. M. Froeling, A. J. Nederveen, K. Nicolay, and G. J. Strijkers, "DTI of human skeletal muscle: the effects of diffusion encoding parameters, signal-to-noise ratio and T2 on tensor indices and fiber tracts," *NMR in Biomedicine*, vol. 26, no. 11, pp. 1339–1352, 2013. DOI: 10.1002/nbm.2959.

26. Y. N. Kami, M. Sumi, Y. Takagi, M. Sasaki, M. Uetani, and T. Nakamura, "Arterial spin labeling imaging for the parotid glands of patients with Sjögren's syndrome," *PLoS ONE*, vol. 11, no. 3, pp. 1–14, 2016. DOI: 10.1371/journal.pone.0150680.

27. S. Rosenzweig, H. C. M. Holme, R. N. Wilke, D. Voit, J. Frahm, and M. Uecker, "Simultaneous multi-slice MRI using cartesian and radial FLASH and regularized nonlinear inversion: SMS-NLINV," *Magnetic Resonance in Medicine*, vol. 79, no. 4, pp. 2057–2066, 2018. DOI: 10.1002/mrm.26878.

28. M. Uecker, T. Hohage, K. T. Block, and J. Frahm, "Image reconstruction by regularized nonlinear inversion - Joint estimation of coil sensitivities and image content," *Magnetic Resonance in Medicine*, vol. 60, no. 3, pp. 674–682, 2008. DOI: 10.1002/mrm.21691.

29. K. P. McGee, R. S. Stormont, S. A. Lindsay, V. Taracila, D. Savitskij, F. Robb, R. J. Witte, T. J. Kaufmann, J. Huston, S. J. Riederer, E. A. Borisch, and P. J. Rossman, "Characterization and evaluation of a flexible MRI receive coil array for radiation therapy MR treatment planning using highly decoupled RF circuits," *Physics in Medicine and Biology*, vol. 63, no. 8, 08NT02, 2018. DOI: 10.1088/1361-6560/aab691.

Chapter 6

6

Dynamic MRI of swallowing: real-time volumetric imaging at 12 frames per second at 3 T

Luuk Voskuilen, Jasper Schoormans, Oliver J. Gurney-Champion, Alfons J.M. Balm, Gustav J. Strijkers, Ludi E. Smeele, Aart J. Nederveen

Magnetic Resonance Materials in Physics, Biology, and Medicine

2021

Abstract

Objective: Dysphagia or difficulty in swallowing is a potentially hazardous clinical problem that needs regular monitoring. Real-time 2D MRI of swallowing is a promising radiation-free alternative to the current clinical standard: videofluoroscopy. However, aspiration may be missed if it occurs outside this single imaged slice. We therefore aimed to image swallowing in 3D real-time at 12 frames per second (fps).
Materials and methods: At 3 T, three 3D real-time MRI acquisition approaches were compared to the 2D acquisition: an aligned stack-of-stars (SOS), and a rotated SOS with a golden-angle increment and with a tiny golden-angle increment. The optimal 3D acquisition was determined by computer simulations and phantom scans. Subsequently, five healthy volunteers were scanned and swallowing parameters were measured.
Results: Although the rotated SOS approaches resulted in better image quality in simulations, in practice the aligned SOS performed best due to the limited number of slices. The four swallowing phases could be distinguished in 3D real-time MRI, even though the spatial blurring was stronger than in 2D. The swallowing parameters were similar between 2D and 3D.
Conclusion: At a spatial resolution of 2-by-2-by-6 mm with 7 slices, swallowing can be imaged in 3D real-time at a frame rate of 12 fps.

Introduction

Dysphagia or difficulty in swallowing is a potentially hazardous complication of diseases ranging from neurological disorders[1] to head-and-neck cancer[2]. Dysphagia can be classified either as a mechanical obstruction, for example the compression of the pharyngeal tract by a tumour mass, or as a motility disorder, often indicating a neuromuscular disease.[3] Whatever the cause, oropharyngeal dysphagia may have a serious negative impact on the food intake[4], reducing the quality of life[5], and increases the chance of aspiration and subsequent pneumonia[6], which may be life threatening. The swallowing functionality of patients suffering from dysphagia should therefore be monitored regularly.

The most commonly used method for evaluating swallowing functionality is the videofluoroscopic swallowing study (VFSS)[7], also called modified barium swallowing study. While the patient swallows a radiopaque contrast agent, a fluoroscope visualises the oral and pharyngeal phases of swallowing, and any aspiration of the contrast agent. Although the VFSS is the diagnostic test of choice for oropharyngeal dysphagia[3], it is subject to several disadvantages: it has poor soft-tissue contrast and it superimposes anatomical structures making 3D localisation difficult. Moreover, it exposes the patient to ionising radiation, which is especially harmful in persistent dysphagia patients as multiple follow-up consultations are required.[8]

Dynamic MRI of swallowing has been proposed as an alternative to the VFSS[9] that is able to visualise the mobility of soft tissues.[10] Until recently, the main disadvantage of dynamic MRI of swallowing had been the low temporal resolution of only several frames per second (fps). However, advances in MRI acceleration, such as compressed sensing[11], have shown that it is feasible to study swallowing with MRI in real time in 2D at 24.3 fps[12], which is a similar frame rate to the VFSS.

Unlike the VFSS, real-time MRI of swallowing is not based on ionising radiation, and provides better soft-tissue contrast. However, as this real-time MRI is mostly acquired only in a single 2D slice, if aspiration occurs outside this imaged midsagittal slice, it will be missed. Such aspiration could be detected on the projection images generated in VFSS, although without a 3D localisation due to the superimposition of the anatomical structures. We can overcome the limitations of the VFSS and conventional 2D real-time MRI by developing real-time 3D MRI for swallowing.

However, with a naïve extension of 2D MRI to 3D real-time MRI the frame rate would decrease linearly with the number of slices, which may hinder a correct assessment of swallowing. Recently, a frame rate of 15 fps had been shown to be sufficiently high to correctly image swallowing using the VFSS.[13] 3D real-time MRI of swallowing should therefore aim to approach this frame rate of 15 fps as closely as possible.

In this study, we aim to image swallowing in 3D real-time at 15 fps. We evaluated three acquisition patterns based on the stack-of-stars (SOS) by computer simulations, and by comparing the image quality provided by these patterns in a static phantom and in healthy volunteers. Finally, we demonstrated the swallowing features that can be visualised with this approach and compare swallowing metrics from this 3D approach to those from the 2D alternative.

Methods

Acquisition strategies

For 2D real-time MRI at 3 T, the radial acquisition pattern is often used[12,14,15], as it is more resistant to motion artefacts than Cartesian imaging. Furthermore, if subsequent radial *k*-lines or spokes are rotated by the golden angle (approximately 111.24°), a near-uniform distribution of spokes is ensured, especially if the number of spokes belongs to the Fibonacci sequence.[16] Additionally, it allows an arbitrary time window length and position, which means that a sliding window reconstruction can be used. We therefore aimed to preserve these properties of the 2D radial golden angle acquisition in a 3D stack-of-stars (SOS) acquisition.

The most straightforward way of extending 2D radial to 3D is the aligned SOS (ASOS), in which the k_z-direction is fully sampled for each angle.[17] For static SOS imaging, Zhou et al.[18] recently showed that the image quality can be improved by rotating the individual stars over the slice direction (k_z-direction) in a SOS acquisition using the golden ratio (RSOS-GR). In this implementation, however, the rotation between the stars is dependent on the number of spokes and a sliding window can no longer be used.

In this study, we rotate the stars along the k_z-direction by the same golden angle that is used to rotated the spokes in the k_x-k_y-plane (RSOS-ψ_1). The angular increment over the k_z-direction (inner scan loop) should be continued by the increment in-plane (outer scan loop). By using this continuous golden angle, we hypothesise that the image quality may be improved similar to the work by Zhou et al.[18], while a sliding window reconstruction can also be used.

As the angle increment in the radial golden angle acquisition is large (111.25°), eddy currents may be induced that distort the images.[19] Eddy currents could therefore negate the image-quality improvement by rotating the stars. Similar to the golden angle, tiny golden angles allow a sliding window reconstruction, while the spokes are distributed evenly.[20] Although the first tiny golden angle (ψ_1) is equal to the golden angle, the subsequent tiny golden angles are smaller and thus should induce fewer eddy-currents. To ensure an even distribution of spokes in the tiny-golden-angle rotated stack-of-stars (RSOS-ψ_9), we chose the tiny golden angle number equal to the number of samples in the k_z-direction ($\psi_9 \approx 18.71°$).

Simulations

In order to compare four sampling strategies (ASOS, RSOS-GR, RSOS-ψ_1, and RSOS-ψ_9), we determined the incoherence of the PSF similar to the simulations by Zhou et al.[18]. The number of read-out points (256) and the number of slices (nine including oversampling) were the same as those of the phantom and in vivo scans. To obtain the PSF in image space, the spokes were transformed

using a non-uniform fast Fourier transform without Toeplitz embedding in BART (version 5.0.0)[21] and MATLAB (R2019a, Mathworks, Natick, MA). Finally, the incoherence of the PSF was quantified by dividing the magnitude of the central peak by the standard deviation of the side lobes.

Pineapple phantom

Three acquisition patterns (ASOS, RSOS-ψ_1, and RSOS-ψ_9) were implemented on a 3T Philips Ingenia scanner (Best, Netherlands) and evaluated by scanning a pineapple. The RSOS-GR acquisition strategy, for which the spoke angle is dependent on the number of spokes, is not compatible with a sliding window approach, in which the number of spokes may vary, and was therefore not implemented. For each acquisition pattern, a scan was performed for 402 spokes per slice (fully sampled), resulting in a scan time of 1.2 s per slice. First, the number of slices was set to 7, which would be sufficient to cover the oral cavity and oropharynx given the slice thickness of 6 mm. Subsequently, these three acquisition patterns were repeated with 21 slices. The other scan parameters were identical for the three acquisitions: 3D FFE; TR/TE = 3.0/1.26 ms; flip angle: 10°; read-out samples: 256; FOV: matrix size: 128x128; voxel size 2x2x6 mm^3 (APxFHxRL); slice oversampling: 1.28 times the number of slices; receiver coils: torso and table coils.

The images of the pineapple phantom were reconstructed using Matlab (R2019a, Mathworks, Natick, MA) and BART (version 5.0.0)[21]. A radial phase shift correction[22] and noise pre-whitening were performed on the spokes before reconstruction. Assuming that sensitivity maps are constant over time, these maps were estimated from low resolution images reconstructed from all spokes using ESPIRiT[23]. From all 402 spokes per slice, a fully sampled reference image was reconstructed using SENSE[24]. Subsequently, 6 spokes per slice were binned for each frame, resulting in a frame rate of 6.2 fps for the 7-slice acquisition and 2.1 fps for 21-slice acquisition. The real-time images were reconstructed using compressed sensing with locally-low rank regularisation in plane (with regularisation parameter λ=0.001)[25] and total variation regularisation over time (with λ=0.005), which were chosen empirically based on the image quality of the reconstructed images.

To determine the quality of the compressed-sensing reconstruction, we calculated the structural similarity index (SSIM)[26] between the fully-sampled ground truth image and the central frame of the accelerated images.

Healthy volunteers

Five healthy volunteers were included (1 female, mean age 28 years, range 26–29 years), from whom we obtained written informed consent. For this study, we received the appropriate approval from the institutional medical ethical committee.

Exclusion criteria were metal braces, dental splints, or any general MRI contraindication. Based on the results from the phantom scans, only the 7-sliced ASOS acquisition was acquired and compared to conventional 2D golden-angle imaging. During these acquisitions, the volunteers were asked to swallow 20 mL of pineapple juice in one swallow, which they administered to themselves with a syringe.

Only the 7-slice acquisition was able to achieve a frame rate of 12 fps, and hence, unlike for the phantom, no acquisitions with 21 slices were made for the volunteers. Additionally, a custom 12-channel flexible surface coil for tongue imaging[14] was used instead of the torso coil. As the two halves of this coil are strapped directly to the cheeks of the volunteers and as it is specifically tuned for tongue imaging, the SNR in the tongue is approximately doubled compared to a conventional neurovascular coil. In addition, the coil features a higher density of coil elements which resulted in better image quality for parallel imaging and compressed sensing reconstructions.[14] For phantom scans however, this coil is not suitable as the coil elements have been tuned to the human head. The remaining acquisition parameters were the same as for the pineapple phantom.

The image reconstruction was equal to that of the pineapple phantom, except for two additional steps: a sliding window approach and a flat-field filter. The sliding window was implemented by repeating the reconstruction of the MRI series twice, independently, with half a frame rate shift. Hence, each individual reconstruction was identical to that of the pineapple and simulations and only after reconstruction, frames were combined to achieve a doubling of the frame rate. The flat-field filter was necessary in vivo to adjust for the local sensitivity changes of the small coil elements used in the dedicated coil array. By dividing an 3D image by that same image convolved with a wide Gaussian kernel, this filter was able to correct for the inhomogeneous image intensity caused by the flexible receiver coil. The flat-field filter also made the comparison possible between the in vivo and phantom scans, which inhomogeneous due to the larger acquisition coils used.

From the 2D and 3D real-time imaging, quantitative parameters of swallowing were measured according to Olthoff et al.[27]: the duration of swallowing, the area of the bolus (in the middle slice for 3D imaging), laryngeal elevation, and contraction of the submental muscles such as the anterior belly of the digastric muscle. For the 3D imaging of swallowing, the bolus volume measurement was added, which was the only measurement based on multiple slices. These parameters were measured by a single rater with four years of experience in real-time MRI. The measurements are described in more detail in supplementary material 6.1.

Results

The PSF simulations indicated that the incoherence was higher for all three RSOS approaches (RSOS-GR, RSOS-ψ_1, and RSOS-ψ_9) than for the aligned SOS, regardless of the number of spokes per slice used (figure 6.1). The difference in incoherence between the three RSOS approaches was marginal, where RSOS-ψ_1 and RSOS-ψ_9 performed slightly better at fewer than 18 spokes per slice.

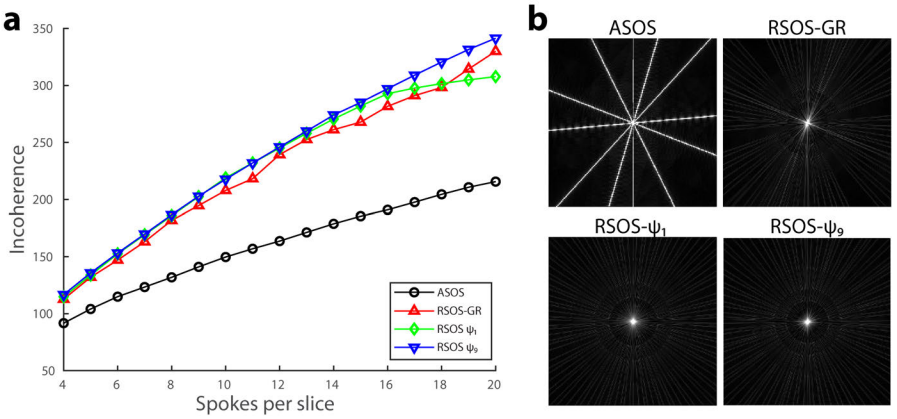


FIGURE 6.1 — Point-spread function (PSF) simulations for four 3D radial stack-of-stars acquisition patterns. The incoherence was quantified by the ratio of the central peak relative and the standard deviation of the side lobes. In PSFs with higher incoherence, the streaking artefacts were more noise-like, which should benefit a compressed sensing reconstruction. For the three rotated stack-of-stars (RSOS) acquisition patterns, the incoherence was better than that of the aligned stack-of-stars (ASOS) acquisition (a). The difference between the RSOS-GR (Zhou et al.[18]), and the continuous RSOS with the golden angle (RSOS-ψ_1) and tiny golden angle (RSOS-ψ_9) were small. For 6 spokes per slice, the central slices of the PSFs for the four acquisition strategies are displayed with the display range from 0 to 10% of the central peak intensity (b).

For the pineapple phantom, the images reconstructed using compressed sensing displayed a reduction in detail (blurring), which may have been caused by the broader PSF or the compressed sensing regularisation (figure 6.2). For the acquisitions with 7 slices (figure 6.2a), the SSIM was 0.737 for the ASOS, 0.692 for RSOS-ψ_1, and 0.693 for RSOS-ψ_9. Contrary to the computer simulations, the image quality of the ASOS acquisition (quantified by the SSIM) was the best of the three, as the RSOS-ψ_1 suffered mainly from more radial streaking and the RSOS-ψ_9 displayed a stronger halo, which may be best appreciated in the difference images (figure 6.2a).

For the acquisitions of the pineapple phantom with 21 slices (figure 6.2b), the SSIM was 0.708 for the ASOS, 0.589 for RSOS-ψ_1, and 0.731 for RSOS-ψ_9. The image quality was quantitatively the best for RSOS-ψ_9, which was in line with the results from the simulations.

All volunteers were able to swallow the contrast agent (pineapple juice) in a supine position without problems (figure 6.3). The temporal resolution was sufficient such to distinguish the oral, pharyngeal, and the early oesophageal phases of swallowing in all volunteers (figure 6.4, supplementary material 6.2). In the 3D real-time imaging, the duration of swallowing and the laryngeal elevation were similar to those in 2D (table 6.1). The contraction of the submental muscles however was smaller for the 3D imaging than for 2D imaging. The bolus volume (only measured in 3D) was larger than the administered volume, most probably due the mixing of the contrast agent with saliva. For 3D imaging, the contraction of the submental muscles is less than for 2D imaging in both studies.

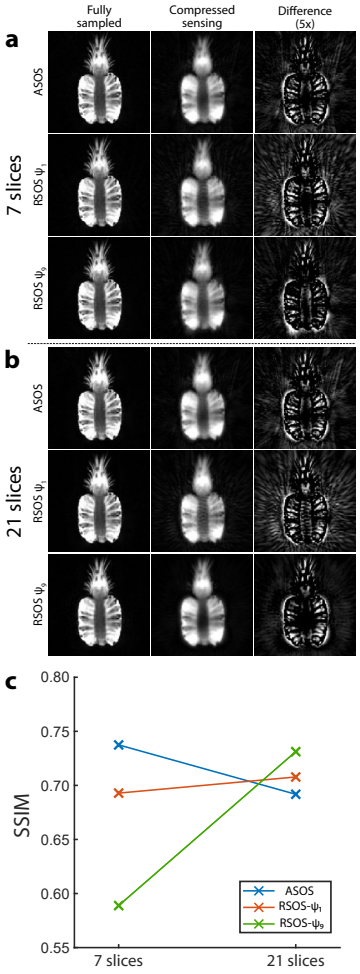


FIGURE 6.2 — In a pineapple phantom, a fully-sampled reconstruction (first column) was compared to a compressed-sensing reconstruction with six spokes per slice (second column). The difference between these reconstructions (third column) is displayed with five times higher window level settings. The three acquisition patterns were acquired with 7 slices (a) and 21 slices (b), and the structural similarity index (SSIM) between the fully-sampled and compressed-sensing reconstruction was calculated (c). For both 7 and 21 slices, the RSOS-ψ_1 pattern resulted in more radial streaking, and a lower SSIM than for the ASOS. Although the image quality of the ASOS was better than that of the RSOS-ψ_9 acquisition for the 7-slice acquisition, the image quality of the RSOS-ψ_9 acquisition was better when 21 slices were acquired.

TABLE 6.1 — Quantitative swallowing parameters from 2D and 3D real-time imaging.

	This study				*Olthoff et al., 2016*	
Scan mode	3D		2D		2D	
Population size	5		5		11	
Bolus volume (mL)	20		20		10	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
Duration (s)	2.2	0.39	1.9	0.58	4.6	2.0
Bolus (cm^2)	7.5	2.8	10.5	2.0	3.0	1.4
Bolus (cm^3)	24.4	5.2	n/a	n/a	n/a	n/a
Laryngeal elevation (mm)	20.4	5.8	20.5	4.5	27	5.0
Contraction of submental muscles (%)	-25.2	5.0	-30.6	2.9	-29	6.0

FIGURE 6.3 — For each of the five healthy volunteers, a frame is displayed just before swallowing pineapple juice. The initial position of the contrast agent (pineapple juice) differed between the volunteers: The first three volunteers have the juice on top of the tongue, while the last two volunteers have the juice in front of the tongue.

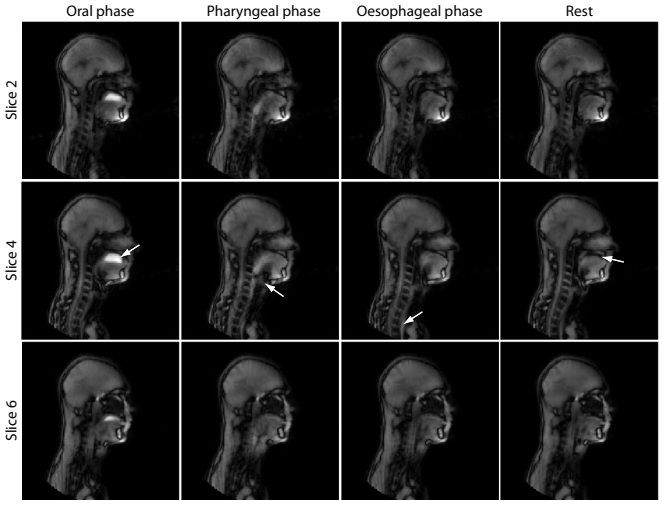


FIGURE 6.4 — Overview of real-time 3D MRI of swallowing four frames (at equal distance in time) that represent four swallowing phases. Three of the total of seven slices are shown. The contrast agent (pineapple juice) is displayed as a hyperintense fluid in the oral phase (arrow in the first column). Laryngeal elevation and closure by the epiglottis can be appreciated in the pharyngeal phase (arrow in second column). In the oesophageal phase, the contrast agent is located in the oesophagus (arrow third column). Finally in rest, tongue relaxes and creates space between the palate and itself (arrow last column).

Discussion

In this study, we aimed to image swallowing in 3D real-time at 3 T. Computer simulations showed that a rotated stack-of-stars (RSOS) approach results in a more incoherent PSF, which should benefit a compressed sensing reconstruction. For the experiments with the pineapple phantom, the image quality of the RSOS acquisition with the golden angle (RSOS-ψ_1) was lower than that of the aligned stack-of-stars (ASOS), which we attribute to increased presence of eddy currents in RSOS-ψ_1. A tiny golden angle approach (RSOS-ψ_9) reduced these distortions. However, the image quality of the RSOS-ψ_9 approach was only superior to that of the ASOS acquisitions if a sufficient number of slices were acquired. Using the ASOS approach, we were able to image swallowing in 3D real-time with a sufficient spatial and temporal resolution to distinguish several swallowing phases and to derive quantitative swallowing parameters: duration, laryngeal elevation, bolus area and volume, and contraction of the submental muscles.

Previously, Zhou et al.[18] found that an RSOS acquisition reduced streaking artefacts. In computer simulations, the incoherence produced by their RSOS-GR approach was higher than that produced by the ASOS; a result confirmed in our study. In this study, we found that a sufficiently high number of slices should be used to benefit from a RSOS approach, and that a tiny golden angle may further improve image quality by reducing eddy-current distortions.

Although we also intended to apply an RSOS approach to our volunteer scans, we observed an unexpected requirement for such an approach. Namely, the phantom scans show that a sufficiently high number of slices is required to benefit from an RSOS approach. This effect may be explained in the following way: If a small number of slices is chosen, the sparsity over slice direction is bad, i.e. there are few zero-valued elements along this dimension in the transformed domain. Therefore, the compressed-sensing reconstruction benefits little from the 3D acquisition. However, as we aimed for a frame rate of 15 fps, which was incompatible with a high number of slices, we ultimately chose for the conventional ASOS approach with seven slices.

Even though the number of slices was reduced to seven, we still only reached a temporal resolution of 12 fps. We decided not to increase the frame rate further, as it would reduce the image quality too much. In order to push the frame rate to our aim of 15 fps, more time-efficient read-outs could be used.

There are several alternative approaches to real-time MRI that we did not explore in this work. One could use a stack-of-spiral approach to achieve higher frame rate, as spiral trajectories cover more k-space per read-out than radial trajectories and therefore are more time-efficient. Although spiral read-outs can be more challenging than radial read-outs, spiral approaches have already been applied successfully for real-time 3D imaging of speech[28,29]. It would be interesting to see whether these approaches can be adapted for imaging of swallowing to achieve a

frame rate of at least 15 fps.

Although a spiral approach is already being successfully applied to real-time 3D imaging of speech at 1.5 T[28], B_0-inhomogeneities are more pronounced due to the higher magnetic field at 3 T. Simply acquiring an additional off-resonance field map before the real-time scan may be used to correct for off-resonance artefacts, but it does not contain time-resolved information. Recently, Lim et al. described a method for dynamic off-resonance correction for speech imaging[29] that could help bring spiral imaging to MRI of swallowing.

In addition to rotating in the k_x-k_y-plane, the spokes may also be rotated in the k_z-direction, a so-called koosh-ball trajectory. This acquisition pattern is a useful alternative, especially for isometric field-of-views (FOVs). However the koosh-ball trajectory requires more spokes to be acquired to meet the Nyquist criterion than a RSOS acquisition would. To account for this, Burdumy et al.[30] implemented an RSOS acquisition pattern, in which the number of spokes per slice was also reduced as the distance from the k-space centre increased. Alternatively to 3D MRI, 2D images may be acquired along three orthogonal orientations[31], which already provides a better overview of the swallowing than a single 2D acquisition. However, this imaging approach required multiple swallows, which are difficult to exactly reproduce leading to inconsistencies across slices.

Finally, Fu et al.[32] developed a low-rank method for speech imaging in which a common temporal-spatial subspace is determined from data over repeated talking. If this method could be adopted for imaging of swallowing, this could allow for a substantially higher resolution and frame rate. However, this method would require repeated swallowing, which in its current format (over 7 minutes of repeated speech) would be physiologically unsafe when applied to swallowing. How these methods perform compared to our SOS approaches should be investigated in future work.

The current reconstruction pipeline is an important hurdle for clinical implementation of the real-time 3D MRI of swallowing. As the whole 4D volume is reconstructed at once, it requires lengthy off-line reconstruction on a high-performance computer (in our case nearly three hours using four CPUs). If the inverse Fourier transform is first applied along the k_z-direction, the reconstruction can be considerably accelerated by reconstructing the slices separately in parallel, but this is only an option for the ASOS acquisition.

For 2D and 3D real-time imaging, we determined several parameters to quantify the swallowing movement, which could be compared to previous work by Olthoff et al.[27]. In their study, the duration of swallowing is much longer (mean of 4.6 s), as our definition of a swallowing movement was stricter, meaning that in our case, only a single swallows was included. Between our 2D and 3D scans, the mean laryngeal elevation was nearly equal, but lower than that of Olthoff et al., which may have been caused by inter-rater variability or the smaller fluid bolus administered in their study. Finally, the mean contraction of the submental muscles

in 2D is approximately equal to that of Olthoff et al., but for our 3D acquisitions, this contraction is slightly lower. We attribute this decrease to stronger blurring that is present in 3D images compared to the 2D acquisition.

With the exception the bolus volume, none of the measurements took full advantage of the 3D imaging, as we only intended to determine that there are only minor differences between the 2D acquisition and the 3D acquisition. To better utilise the 3D acquisition, new parameters should be developed. As we are missing patient data in this work, and as there is no similar 3D imaging modality from which metrics may be derived, we decided that the development of new metrics was out of the scope of this work.

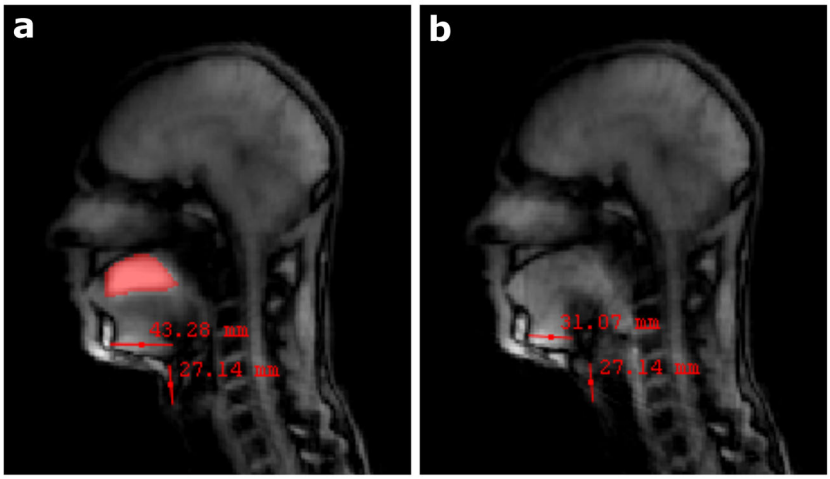
Although we were able to perform 3D real-time MRI of swallowing, the main limitation of this study is that we did not compare this technique with the current gold standard for grading dysphagia, VFSS. In the case of inclusion body myositis, 2D real-time MRI has been shown to be able to identify the cause of dysphagia as well as VFSS[15]. However, whether the spatial or temporal resolution of the 3D real-time MRI is sufficient for the detection of aspiration or bolus retention still has to be examined. Comparing VFSS to 3D real-time MRI, we would expect VFSS to provide a better temporal resolution and to be cheaper. In contrast, 3D real-time MRI does not require ionising radiation and provides better soft-tissue contrast, thereby allowing the analysis of swallowing beyond grading the aspiration. A disadvantage may be that MRI scans are generally performed in a supine or prone position, while the natural position for swallowing is upright. Further research should prove whether the position during scanning affects swallowing and thus the grading of dysphagia.

In conclusion, we were able to image and quantify swallowing in 3D real-time using MRI at 3 T. We evaluated three SOS acquisition patterns. Although computer simulations showed that the RSOS acquisitions produced a more incoherent PSF that should provide better image quality, the image quality was reduced in RSOS acquisitions probably due to eddy currents. Eddy currents were mitigated by a tiny golden angle radial *k*-space filling, which resulted in similar image quality to the aligned SOS. Using this SOS approach, we achieved the imaging of swallowing in 3D with 12 fps, visualising several swallowing phases. We demonstrated that real-time 3D MRI is a potential radiation-free alternative to the VFSS, which can also visualise soft tissues and localise the origin of swallowing problems in 3D.

Acknowledgements

We would like to thank the Maurits en Anna de Kock Stichting for kindly providing a grant for the purchase of the flexible receiver coil for tongue imaging.

Supplementary materials



SUPPLEMENTARY MATERIAL 6.1 — The definition of the quantitative swallowing metrics visualised on a sagittal slice of the 3D real-time MRI of swallowing. The bolus area and volume were measured by delineating the hyperintense pineapple juice on the image during rest (a). The length of the submental muscles was defined as the distance from the right edge of the mandible to the left edge of the hyoid bone. The contraction of the submental muscles was calculated as the decrease in length during the swallow (b) relative to the length before swallowing (a). The laryngeal elevation was measured as the largest distance between the position of the most inferior point of the thyroid cartilage before swallowing and it position during swallowing. The duration was calculated by counting the frames from the last stationary motion state until the oral tract returned to this state. Eventual additional swallows to completely clear the oral or pharyngeal cavities were not included in the duration.

SUPPLEMENTARY MATERIAL 6.2 — Visualisation of the full real-time 3D MRI movie of a single swallowing movement. Subsequent slices are placed in reading order: from left to right, up to down. As this is the same movie as visualised in figure 6.6, the same swallowing phases can be distinguished.

References

1. B. Kertscher, R. Speyer, M. Palmieri, and C. Plant, "Bedside screening to detect oropharyngeal dysphagia in patients with neurological disorders: An updated systematic review," *Dysphagia*, vol. 29, no. 2, pp. 204–212, 2014. DOI: 10.1007/s00455-013-9490-9.

2. K. Manikantan, S. Khode, S. I. Sayed, J. Roe, C. M. Nutting, P. Rhys-Evans, K. J. Harrington, and R. Kazi, "Dysphagia in head and neck cancer," *Cancer Treatment Reviews*, vol. 35, no. 8, pp. 724–732, 2009. DOI: 10.1016/j.ctrv.2009.08.008.

3. A. A. Jalil, D. A. Katzka, and D. O. Castell, "Approach to the patient with dysphagia," *American Journal of Medicine*, vol. 128, no. 10, 1138.e17–1138.e23, 2015. DOI: 10.1016/j.amjmed.2015.04.026.

4. S. Mercadante, F. Aielli, C. Adile, P. Ferrera, A. Valle, F. Fusco, A. Caruselli, C. Cartoni, P. Massimo, F. Masedu, M. Valenti, and G. Porzio, "Prevalence of oral mucositis, dry mouth, and dysphagia in advanced cancer patients," *Supportive Care in Cancer*, vol. 23, no. 11, pp. 3249–3255, 2015. DOI: 10.1007/s00520-015-2720-y.

5. N. P. Nguyen, C. Frank, C. C. Moltz, P. Vos, H. J. Smith, U. Karlsson, S. Dutta, A. Midyett, J. Barloon, and S. Sallah, "Impact of dysphagia on quality of life after treatment of head-and-neck cancer," *International Journal of Radiation Oncology Biology Physics*, vol. 61, no. 3, pp. 772–778, 2005. DOI: 10.1016/j.ijrobp.2004.06.017.

6. A. Eisbruch, T. Lyden, C. R. Bradford, L. A. Dawson, M. J. Haxer, A. E. Miller, T. N. Teknos, D. B. Chepeha, N. D. Hogikyan, J. E. Terrell, and G. T. Wolf, "Objective assessment of swallowing dysfunction and aspiration after radiation concurrent with chemotherapy for head-and-neck cancer," *International Journal of Radiation Oncology Biology Physics*, vol. 53, no. 1, pp. 23–28, 2002. DOI: 10.1016/S0360-3016(02)02712-8.

7. B. A. Murphy and J. Gilbert, "Dysphagia in Head and Neck Cancer Patients Treated With Radiation: Assessment, Sequelae, and Rehabilitation," *Seminars in Radiation Oncology*, vol. 19, no. 1, pp. 35–42, 2009. DOI: 10.1016/j.semradonc.2008.09.007.

8. M. M. Szczesniak, J. Maclean, T. Zhang, P. H. Graham, and I. J. Cook, "Persistent dysphagia after head and neck radiotherapy: A common and under-reported complication with significant effect on non-cancer-related mortality," *Clinical Oncology*, vol. 26, no. 11, pp. 697–703, 2014. DOI: 10.1016/j.clon.2014.08.009.

9. A. Olthoff, P. O. Carstens, S. Zhang, E. von Fintel, T. Friede, J. Lotz, J. Frahm, and J. Schmidt, "Evaluation of dysphagia by novel real-time magnetic resonance imaging," *Neurology*, pp. 2132–2138, 2016. DOI: 10.1212/WNL.0000000000003337.

10. A. M. Kreeft, C. R. N. Rasch, S. H. Muller, F. A. Pameijer, E. Hallo, and A. J. M. Balm, "Cine MRI of swallowing in patients with advanced oral or oropharyngeal carcinoma: A feasibility study," *European Archives of Oto-Rhino-Laryngology*, vol. 269, no. 6, pp. 1703–1711, 2012. DOI: 10.1007/s00405-011-1861-y.

11. M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing MRI," *IEEE Signal Processing Magazine*, vol. 25, no. March 2008, pp. 72–82, 2008. DOI: Doi10.1109/Tit.2006.871582. arXiv: 1204.4227v1.

12. S. Zhang, A. Olthoff, and J. Frahm, "Real-time magnetic resonance imaging of normal swallowing," *Journal of Magnetic Resonance Imaging*, vol. 35, no. 6, pp. 1372–1379, 2012. DOI: 10.1002/jmri.23591.

13. J. Layly, F. Marmouset, G. Chassagnon, P. Bertrand, D. Sirinelli, J. P. Cottier, and B. Morel, "Can We Reduce Frame Rate to 15 Images per Second in Pediatric Videofluoroscopic Swallow Studies?" *Dysphagia*, vol. 35, no. 2, pp. 296–300, 2019. DOI: 10.1007/s00455-019-10027-8.

14. L. Voskuilen, P. de Heer, L. van der Molen, A. J. M. Balm, F. van der Heijden, G. J. Strijkers, L. E. Smeele, and A. J. Nederveen, "A 12-channel flexible receiver coil for accelerated tongue imaging," *Magnetic Resonance Materials in Physics, Biology and Medicine*, 2020. DOI: 10.1007/s10334-019-00824-5.

15. P.-O. Carstens, S. Zhang, A. Olthoff, E. Bremen, J. Lotz, J. Frahm, and J. Schmidt, "Evaluation of dysphagia in inclusion body myositis by novel real-time MRI," *European Journal of Neurology*, vol. 84, no. 14, p. 110, 2015.

16. S. Winkelmann, T. Schaeffter, T. Koehler, H. Eggers, and O. Doessel, "An optimal radial profile order based on the golden ratio for time-resolved MRI," *IEEE Transactions on Medical Imaging*, vol. 26, no. 1, pp. 68–76, 2007. DOI: 10.1109/TMI.2006.885337.

17. L. Feng, R. Grimm, K. T. Block, H. Chandarana, S. Kim, J. Xu, L. Axel, D. K. Sodickson, and R. Otazo, "Golden-angle radial sparse parallel MRI: combination of compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric MRI," *Magnetic resonance in medicine*, vol. 72, no. 3, pp. 707–717, 2014. DOI: 10.1002/mrm.24980. arXiv: NIHMS150003.

18. Z. Zhou, F. Han, L. Yan, D. J. J. Wang, and P. Hu, "Golden-ratio rotated stack-of-stars acquisition for improved volumetric MRI," *Magnetic Resonance in Medicine*, vol. 78, no. 6, pp. 2290–2298, 2017. DOI: 10.1002/mrm.26625.

19. S. Wundrak, J. Paul, J. Ulrici, E. Hell, M. A. Geibel, P. Bernhardt, W. Rottbauer, and V. Rasche, "Golden ratio sparse MRI using tiny golden angles," *Magnetic Resonance in Medicine*, vol. 75, no. 6, pp. 2372–2378, 2016. DOI: 10.1002/mrm.25831.

20. S. Wundrak, J. Paul, J. Ulrici, E. Hell, and V. Rasche, "A small surrogate for the golden angle in time-resolved radial MRI based on generalized fibonacci sequences," *IEEE Transactions on Medical Imaging*, vol. 34, no. 6, pp. 1262–1269, 2015. DOI: 10.1109/TMI.2014.2382572.

21. M. Uecker, F. Ong, J. I. Tamir, D. Bahri, P. Virtue, J. Y. Cheng, T. Zhang, and M. Lustig, "Berkeley Advanced Reconstruction Toolbox," in *Proc. Intl. Soc. Mag. Reson. Med*, Toronto, 2015. DOI: 10.5281/zenodo.592960.

22. G. Buonincontri, C. Methner, T. Krieg, T. A. Carpenter, and S. J. Sawiak, "Trajectory correction for free-breathing radial cine MRI," *Magnetic Resonance Imaging*, vol. 32, no. 7, pp. 961–964, 2014. DOI: 10.1016/j.mri.2014.04.006.

23. M. Uecker, P. Lai, M. J. Murphy, P. Virtue, M. Elad, J. M. Pauly, S. S. Vasanawala, and M. Lustig, "ESPIRiT - An eigenvalue approach to autocalibrating parallel MRI: Where SENSE meets GRAPPA," *Magnetic Resonance in Medicine*, vol. 71, no. 3, pp. 990–1001, 2014. DOI: 10.1002/mrm.24751. arXiv: 15334406.

24. K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: Sensitivity Encoding for Fast MRI.," *Magnetic resonance in medicine*, vol. 42, pp. 952–962, 1999.

25. F. Ong and M. Lustig, "Beyond low rank+ sparse: Multi-scale low rank matrix decomposition," *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 10, no. 4, pp. 4663–4667, 2016. DOI: 10.1109/ICASSP.2016.7472561. arXiv: 1507.08751.

26. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

27. A. Olthoff, A. A. Joseph, M. Weidenmüller, B. Riley, and J. Frahm, "Real-time MRI of swallowing: intraoral pressure reduction supports larynx elevation," *NMR in Biomedicine*, vol. 29, no. 11, pp. 1–6, 2016. DOI: 10.1002/nbm.3621.

28. Y. Lim, Y. Zhu, S. G. Lingala, D. Byrd, S. Narayanan, and K. S. Nayak, "3D dynamic MRI of the vocal tract during natural speech," *Magnetic Resonance in Medicine*, vol. 81, no. 3, pp. 1–10, 2018. DOI: 10.1002/mrm.27570.

29. Y. Lim, S. Goud, L. Shrikanth, S. N. Krishna, S. G. Lingala, S. S. Narayanan, and K. S. Nayak, "Dynamic off-resonance correction for spiral real-time MRI of speech," *Magnetic Resonance in Medicine*, vol. 81, no. 1, pp. 234–246, 2019. DOI: 10.1002/mrm.27373.

30. M. Burdumy, L. Traser, F. Burk, B. Richter, M. Echternach, J. G. Korvink, J. Hennig, and M. Zaitsev, "One-second MRI of a three-dimensional vocal tract to measure dynamic articulator modifications," *Journal of Magnetic Resonance Imaging*, vol. 46, no. 1, pp. 1–8, 2016. DOI: 10.1002/jmri.25561.

31. A. Olthoff, S. Zhang, R. Schweizer, and J. Frahm, "On the physiology of normal swallowing as revealed by magnetic resonance imaging in real time," *Gastroenterology Research and Practice*, vol. 2014, 2014. DOI: 10.1155/2014/493174.

32. M. Fu, M. S. Barlaz, J. L. Holtrop, J. L. Perry, D. P. Kuehn, R. K. Shosted, Z. P. Liang, and B. P. Sutton, "High-frame-rate full-vocal-tract 3D dynamic speech imaging," *Magnetic Resonance in Medicine*, vol. 77, no. 4, pp. 1619–1629, 2017. DOI: 10.1002/mrm.26248.

Chapter 7

General discussion

7

General discussion

In this thesis, we have introduced a new MRI protocol that corrects for B_0-inhomogeneities and that can distinguish the crossing muscle fibres in the tongue using the constrained spherical deconvolution (CSD). Using an image registration method driven by this CSD, we were able to construct a population average or atlas of the tongue that not only allows rapid segmentation of the tongue's musculature, but it also allows us to compare separate study populations. Based on this tongue atlas, we also proposed a method that semi-automatically creates personalised finite-element models of the tongue, which better models the tongue's range-of-motion than their generic counterpart does. A major disadvantage of the CSD-MRI, and, consequentially, of the models that are derived from it, is the excessive time such a scan takes. We therefore developed a custom flexible receiver coil for tongue MRI that accelerates CSD imaging with a factor of two. In addition, the signal-to-noise ratio (SNR) gained by this coil can be employed to produce 2D real-time MRI videos of tongue functions such as swallowing. We even showed that 3D real-time MRIs of swallowing can be produced with an acceptable frame rate of 12 frames per second.

Improving imaging of the musculature of the tongue

In chapter two, relatively high-quality CSD images could be acquired in the healthy young study population compared to the images form the elderly volunteers and T1–3 tongue carcinoma patients in chapter three. The younger population was more carefully curated than the other two groups, as younger volunteers with metal dental splints or wires were excluded. These generally steel wires and splints cause a large signal void in MR images of the tongue. Although the current elderly population that now receives MRI scans for carcinoma of the tongue, does not have such fixed retainers, the high prevalence of these steel wires in the group of young volunteers warrants a problem for head-and-neck MRI in the future.

In the elderly and patient populations, another kind of imaging artefacts was often observed that could not be easily corrected for during inclusion: motion artefacts. CSD images had to be excluded if these motion artefacts were present that resulted in misalignment between or signal voids in DWI volumes with different b-values. These artefacts were partially mitigated by using the DWI preprocessing algorithms from FSL[1], in which two scans with opposing phase encodings are registered to the same coordinate system, and combined to correct for B_0-inhomogeneities. If a signal void caused by motion is present in one scan, it can therefore be filled by data from the other. However, if too many motion artefacts are present, no registration may be found, and signal voids may be present at the same location in both scans with opposing phase-encoding, causing signal void in the combined image. Fortunately, the incidence of motion artefacts that

are primarily caused by swallowing, can be reduced by accelerating the scan. In chapter five, a flexible receiver coil is described that can accelerate CSD imaging by a factor of two. By building a coil array with more coil elements, it should be possible to push the acceleration even further.

Instead of correcting for B_0-inhomogeneities, it is also possible to replace the EPI sequence that is the standard for diffusion MRI. A diffusion-prepared TSE sequence has been developed that is almost completely insensitive to these B_0-inhomogeneities.[2] As the EPI is such a fast read-out technique, this diffusion-prepared TSE sequence does take more scan time than the EPI, which increases the chance of motion artefacts in the images. However, the acceleration method used for real-time MRI, compressed sensing, could allow higher acceleration factors than multiband imaging, which we used in chapter five. In animal models, compressed sensing has been adapted for diffusion MRI, showing an up to eight-fold decrease in scan time.[3]

In addition to the imaging artefacts, inconsistencies were present in the quantification of the elderly and patient datasets. As stated in chapter three, these were attributed to the partial volume effect of free water and the tongue muscles. Due to the high diffusivity of free water, the apparent fibre density (AFD), which is dependent on the diffusivity of water in muscle fibres, was distorted. To combat this partial-volume effect, we recommend to extend the CSD model to include the free water diffusion. Such models have already been developed for brain imaging[4], and indeed show a reduced partial volume effect with free water in brain areas that border cerebrospinal fluid. We experimented with incorporating this multi-tissue CSD model in our pipe-line, but we initially could not detect improvement in the estimated fibre-orientation distributions of the tongue (figure 7.1) near the skin of the tongue, where the partial volume effect would be present. We assume that these results may be improved by scanning the tongue with additional high-angular resolution diffusion imaging (HARDI) shells, similar to brain imaging.[4]

Clinical applications of tongue CSD

Although the fibre tracts that can be reconstructed from CSD imaging of the tongue, appear to be a visualisation of macroscopic muscle fibres, they are actually derived from the hindered diffusion on a microscopic level by cellular structures such as cell membranes and proteins. Therefore, changes in the tongue musculature captured by CSD-MRI would reflect changes on a microscopic level. These changes could range from increased diffusion due to an inflammatory process, or altered diffusion anisotropy as cell membranes and protein chains are destroyed by tumour invasion. Currently, these microscopic changes are identified using invasive methods, such as histological or pathological examination of excised tissues. In the future, CSD may prove to be a method of histological examination that is non-invasive, and provides better spatial localisation of these microscopic changes.

This improved spatial localisation may also benefit treatment options. The CSD imaging could be used to better plan surgical procedures, especially in the more complex procedures such as the COMMANDO (combined mandibular operation). If the musculature of the tongue can be measured during a surgery using an ultrasound probe, and if these 2D US images could be mapped to the 3D CSD images, it would even be possible to navigate based on the tongue's musculature.

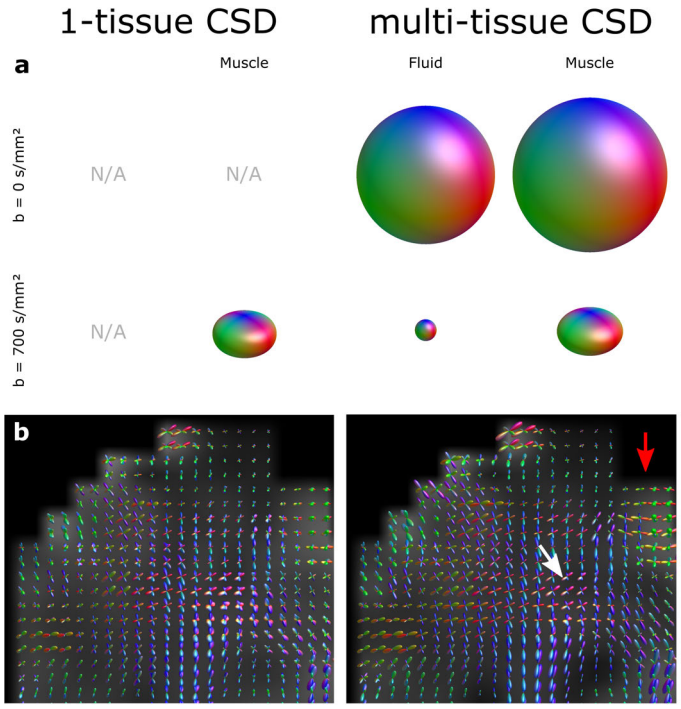


FIGURE 7.1 — The multi-tissue CSD method requires at least two HARDI shells (in this case 0 and 700 s/mm^2) from which it is able to resolve two tissue types: an isotropic free-water component and an anisotropic muscle component. A typical example of a healthy young volunteer is shown in this figure. Response functions (a) were calculated for both tissue types and HARDI shells, and for the conventional one-tissue type method. In coronal slices of the tongue, the resulting FOD images (b) show sharper FOD lobes in the centre of the tongue for the multi-tissue implementation (white arrow), which may indicate that noise from the free water component has been removed. At the edges of the tongue however, where we would expect the most partial voluming with free water, the FODs appear to have been inflated, which may sign to an exacerbation of the partial volume effect. The results may become more consistent across the tongue, if more HARDI shells are included, but this would require and even longer scan time.

As the muscles are the actuators of the tongue, they play a large role in its functionality. The musculature determined by CSD may therefore be correlated to functional outcomes to better predict and potentially adjust rehabilitation after treatment. However, care has to be taken, as the tongue's functionality is dependent on many other factors than only tongue musculature, e.g. fibrosis and innervation.

Fibrosis, which is a well-known and common side-effect of radiation therapy, stiffens the tissue of the tongue, not only after irradiation, but also after surgery. Therefore, fibrosis limits the range-of-motion of the tongue. Some attempts have been made to visualise and quantify fibrotic changes using MRI, such as T1ρ mapping[5] and MR elastography[6], which may fill in these gaps in our knowledge in the future. However, these techniques are not yet available clinically.

Another important factor in tongue functionality is the innervation of muscles and its disruption caused by surgery. As the innervation of the tongue has been shown to differ between individuals[7], imaging of the hypoglossal nerves could be used to adapt a surgical plan to avoid these nerves. To our knowledge however, current MRI methods have not been able to images the tiny distal branches present in the tongue. A promising technique is diffusion-prepared MR neurography[8], that basically suppresses all MR signals in an image, except those signals with diffusion characteristics of (or similar to) nerves. This MR neurography technique works well for the imaging of large and relatively stationary neural structures such as the brachial plexus[9]. However, in mobile organs with thin nerves ($<$1 mm), such as the tongue, we have found that a long scan time (10 minutes or more) is required to obtain enough signal of these nerves to be able to reconstruct an image. This long scan time increases the probability of motion of the tongue corrupting the images, and making this technique difficult to apply clinically.

Although it may difficult to quantify confounders such as fibrosis and innervation directly, it may be possible to determine them by inverse modelling. In forward modelling, as we showed in chapter four using finite-element models, the tongue's functionality is determined based on tongue's musculature, stiffness, and innervation. In inverse modelling, the model is inverted in such a way that, based on the tongue's functionality, the stiffness and innervation could be determined. However, this would probably require extensive information about an individual's functionality, which requires too much time to acquire if at all possible.

Although we have only focussed in CSD-MRI of the tongue in this work, other clinical fields could benefit from detailed information of muscular architectures. The lips contain a large number of muscles in a small volume and these muscles appear to cross. Contrary to the tongue, these muscles do not interdigitate on a microscopic level, but due to the partial volume effect, the muscles appear to cross within the voxels of the image. We also experimented with reconstructing the muscular architecture of the lips based on CSD-MRI (figure 7.3). This resulted in anatomically plausible reconstructions of the musculature of the lips. In the

future, CSD-MRI may also be used to better simulate the functionality in the lips, similarly to the personalised finite-element models of the tongue created in chapter four.

A similar musculature to the lips, where muscles appear to cross, is the pelvic floor. We have been able to apply CSD to the pelvic floor with promising results (figure 7.2). In the future, CSD may aid in early detection of pelvic floor weakness —a side-effect of vaginal child-birth—, hopefully preventing urinal or faecal incontinence.[10]

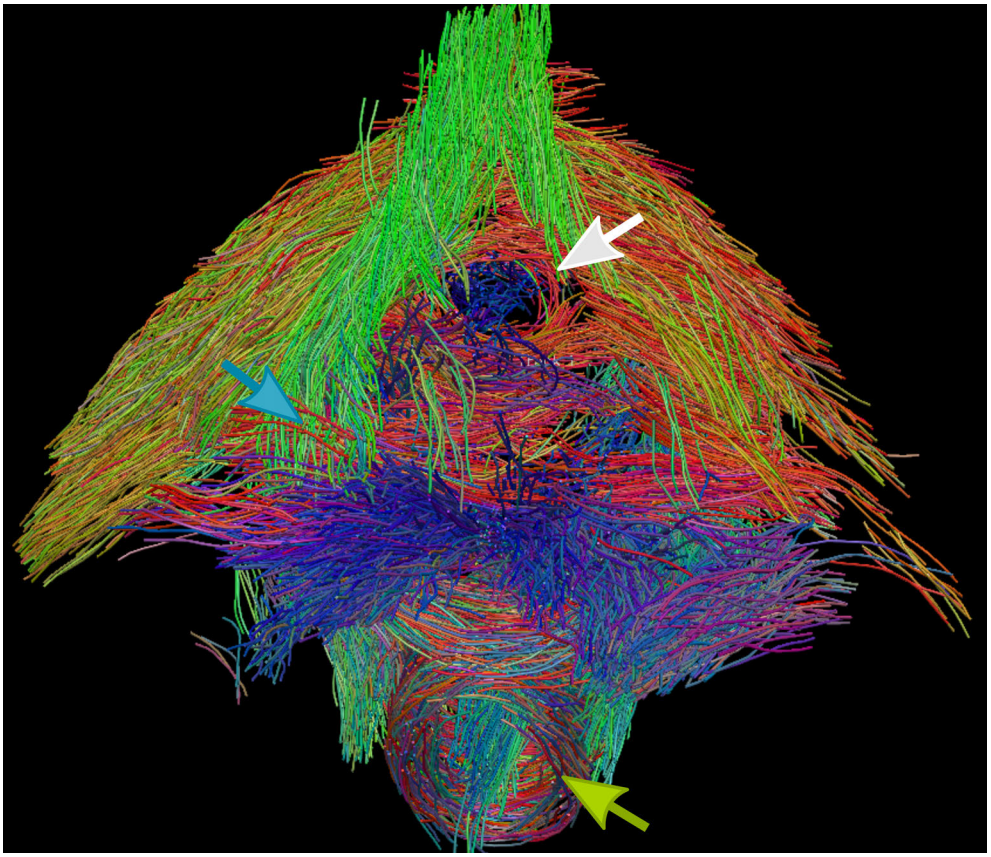


FIGURE 7.2 — Inferior view of CSD-based fibre tractography of the pelvic floor of a healthy female volunteer. The tracks are coloured based on their direction: left-right in red; anterior-posterior in green; and inferior-superior in blue. Two circular fibre population mat be appreciated: anteriorly, the urethral sphincter (white arrow); and posteriorly, the anal sphincter (green arrow). Similarly to the tongue, the pelvic floor contains interleaving fibre populations (blue arrow), that may be resolved by using CSD tractography.

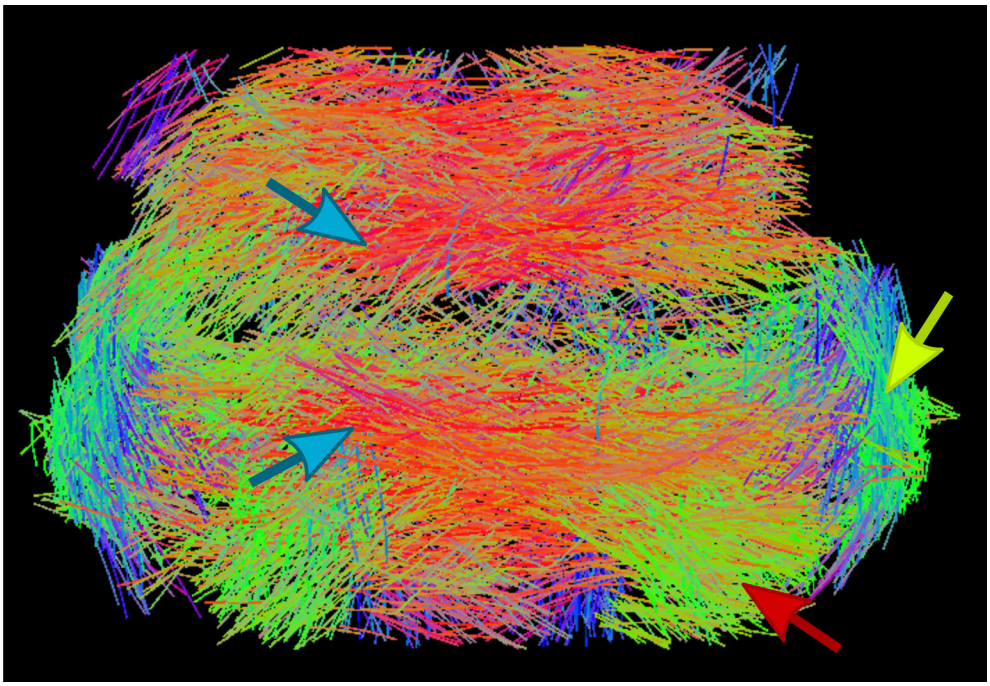


FIGURE 7.3 — Coronal view of CSD tractography of the facial muscles surrounding the mouth of a healthy male volunteer. Various muscles of the mouth may be appreciated such as the orbicularis oris (blue arrow); the depressor labii inferior (blue arrow); and the depressor anguli oris (green arrow). Similarly to tractography of the tongue, such data may be used to personalise the FEM models of the vocal tract.

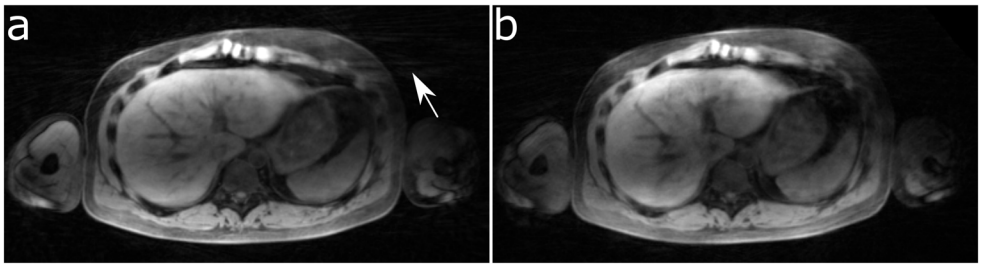


FIGURE 7.4 — An example of real-time MRI of the liver of a healthy male volunteers acquired using two different stack-of-stars acquisitions and reconstructed using compressed sensing with the same parameters. In an image generated using an aligned stack-of-stars (ASOS) acquisition (a), more streaking artefacts may be appreciated than in image from the rotated stack-of-stars (RSOS) acquisition (b). The reduction in streaking is in line with the PSF simulations in chapter six. Although we found no improvement in image quality in the 3D real-time MRI of swallowing, probably due to the limited number of slices, this number is higher in liver imaging, and RSOS is therefore expected to result in a higher image quality.

Real-time MRI of swallowing

With the advent of reconstruction techniques such as compressed sensing (CS)[11] and non-linear inversion (NLINV)[12], new avenues have opened for dynamic MRI. Series of images or videos are possible at a much higher temporal resolution (or frame rate) than before. In addition, acquisitions that had required multiple repetitions of a certain task, might now be acquired within one such repetition (in real-time). The nomenclature of dynamic and real-time MRI occasionally differs, but we agree with the analysis of Nayak[13]: Dynamic MRI captures "dynamic processes (e.g. cardiac function) with adequate temporal resolution to resolve the dynamics"[13]. Real-time MRI is a subdivision of dynamic MRI that requires only a single repetition to render a video. Interactive MRI is a further subdivision of real-time MRI, in which the latency of the acquisition and reconstruction is low —generally within seconds—, so the user can act based on these images, e.g. MRI-guided tumour-ablations. In our research, swallowing is a physiological process that is difficult to repeat in exactly the same way, and the imaging of swallowing does not require interaction. MR imaging of swallow would therefore fall into the category of real-time MRI.

One unresolved issue of real-time MRI of swallowing is the effect of the supine position of the participant or patient. As gravity forces the bolus to the subject's posterior side, it is simultaneously guided into the oesophagus, which is positioned posteriorly to the trachea. Aspiration may therefore be missed on real-time MRI, contrary to a videofluoroscopic swallowing study, that is performed in a more natural upright examination. A solution could be the use of an upright or rotatable MRI scanner. These scanners however generally have a weaker primary magnetic field, and thus a lower SNR and a lower temporal resolution. A more practical solution may be to scan the subject in the current high-field (1.5–3 T) machines while in a prone position. This reverses the direction of the gravitational force and also reverses its effect. For the conventional head coils, this position probably is impossible, but for flexible coils such as the one presented in chapter five, a prone position would not be an issue. Regardless, the effect of positioning whether upright, prone, or supine remains to be investigated on a population with dysphagia.

Although we have focussed on the imaging of swallowing, the methods introduced in this work might also benefit other fields. We were able to generalise the rotated stack-of-stars acquisition pattern to dynamic liver imaging with encouraging image quality (figure 7.4). As many real-time MRI techniques are already used in phonetic science, we are confident that this technique may also improve speech imaging.

We experimented with speech imaging, but we were unable to correlate our real-time images with speech data, as it would require the simultaneous acquisition of *k*-lines and the recording of speech. A regular microphone in the scanner's room or the intercom of the scanner itself can be used to record sound, but the noise

of an active MRI machine drowns out speech. We experimented with acquisitions with lower gradient slew rates, which is defined as the time in which a magnetic gradient reaches its intended value. Lowering this slew rate reduced the noise but also elongated the TR, but the quality of the recorded speech was not sufficient. The only proper solution currently is placement of a recorder next to the location where the sound is produced, i.e. inside the bore of the scanner, and active noise cancelation. These set-ups are commercially available but pricy.

In addition to lowering the slew rate to improve the acquisition of speech real-time MRI, there is a wide array of possibilities of altering the MR sequence mentioned in this work to better suit the application. The fast field echo (FFE) sequence could be exchanged for a balanced steady-state free precession (bSSFP) sequence, which should improve the SNR and delivers a different image contrast[14]. In contrast to FFE, the bSSFP sequence better visualises of certain structures, such as the temporomandibular joint (figure 7.5). Furthermore, imaging has been further accelerated by combining a multiband sequence with radial imaging[15], allowing for the real-time acquisition of multiple parallel slices.

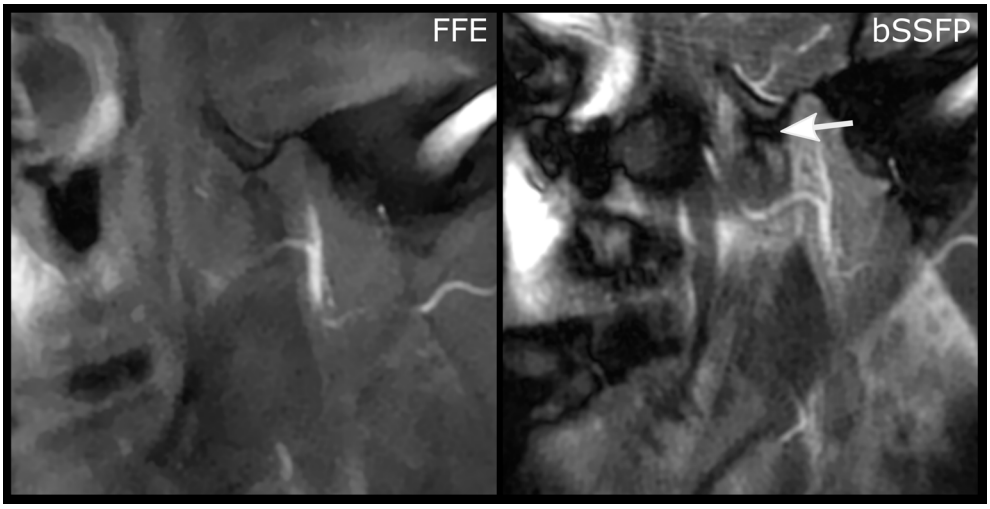


FIGURE 7.5 — 2D real-time MR images of the right temporomandibular joint (TMJ) of a healthy volunteer during maximal opening of the mouth. The left image was taken using an fast-field echo (FFE) sequence, which is nearly identical to the sequence using for the 2D real-time imaging of swallowing in chapters five and six. By changing the MR sequence to a balanced steady-state free precession sequence (right), a different contrast is obtained, in which the articular disc of the TMJ may be better identified (white arrow).

Improving upon the work by Felton et al.[16], we also experimented with the addition of phase-contrast MRI of swallowing that yields velocity information of the tongue. This information could especially benefit finite-element models of the tongue, as this velocity information can be converted to strains that may be used to drive deformation of the finite elements. However, as the phase of the images

is proportional to the velocity within the voxel, accurate phase correction is pivotal to phase-contrast MRI of swallowing, which is difficult to achieve in radial imaging, referring to the eddy-current distortions from chapter six. A cartesian acquisition pattern[17] would negate this problem, but it generally requires multiple repetitions of the same motion, which is difficult for swallowing. Our experiment with a cartesian (pseudo-spiral) undersampling pattern[18] therefore resulted in images with noticeable temporal blur, even while scanning a well-trained volunteer (figure 7.6). Note that this experiment did not include a phase-contrast gradient, which would multiply the total scan time by four, and consequentially, would introduce more blurring.

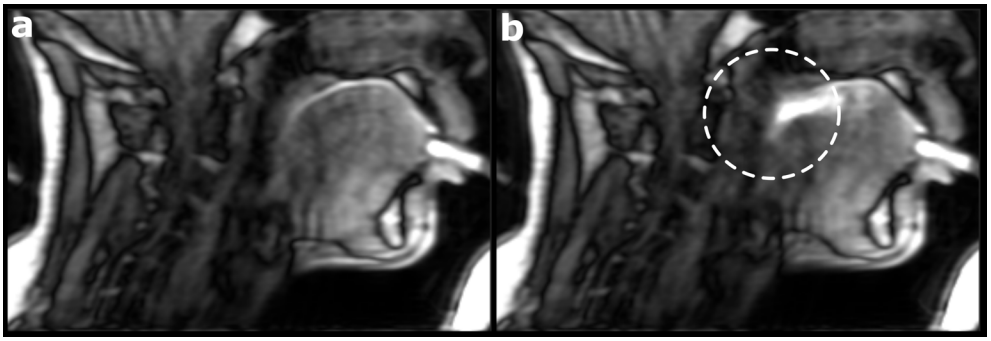


FIGURE 7.6 — A sagittal slice of dynamic MRI of swallowing using a Cartesian pseudo-spiral acquisition pattern and reconstructed from multiple swallows of pineapple juice in a well-trained healthy volunteer. Although the image quality appear reasonably well at rest (a), during swallowing (b), the moving anatomical structures, such as the palate and the back of the tongue, have become blurred. The blurring was probably caused because the timing of the swallows was slightly off, or the motion was slightly different between swallows. For MRI of swallowing, a real-time acquisition would therefore be preferred. Note that a silicon tube was used to administer the pineapple juice to the volunteer, as multiple swallows were required without motion of the head.

Image reconstruction

In addition to undersampling, compressed sensing is an inverse problem in which data fidelity and sparsity are balanced by a regularisation factor. Similarly to the many options in acquisition patterns available for compressed sensing MRI, many image reconstruction algorithms have been developed to solve this inverse problem. In our work, we employed the BART toolbox[19] in combination with two sparsity constraints: a locally-low rank constraint over the spatial dimensions[20] and a total variation constraint over time[21]. Other options that we evaluated, included the wavelet decomposition[22] and the total variation in space, but these produced more image artifacts than the finally chosen constraints.

Not only do the undersampling pattern and the sparsity constraints, but also the regularisation factors play an important role in obtaining the optimal image quality. Arguably, these regularisation factors (generally one for each constraint) are the most difficult element of the inverse problem that we have to choose, as it is dependent on many factors such as the SNR, the type of regularisation or sparsity constraint, the nature of the dimensions over which is regularised (time, space, or breathing states for example), and the contents of the image, i.e. the imaged anatomy. Therefore, the regularisation factors are generally determined by trial-and-error or empirically.

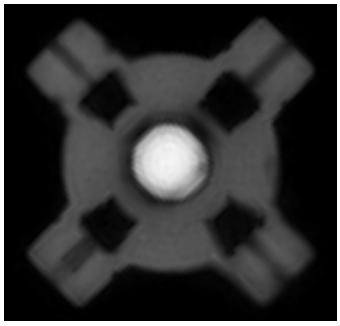


FIGURE 7.7 — The proposed basic element of an MRI phantom that simulates swallowing. This element filled with pineapple juice, is a silicon tube with four chambers that can be filled with compressed air to constrict the lumen. By combining multiple of these elements, and carefully regulating the flow of compressed air, a peristatic motion may be simulated.

A better way of determining the regularisation factor would be to first obtain a fully sampled data set, and subsequently minimising the difference between the images reconstructed from the fully-sampled data set and an undersampled data set. However, for physiological processes that are dynamic and difficult to repeat in exactly the same manner, such as swallowing, the acquisition of a fully sampled data set is practically impossible. In these instances, a dynamic phantom that mimics the physiological process may help in determining the optimal image-reconstruction parameters. Therefore, we initiated the development of an MRI-compatible dynamic swallowing phantom. The first prototype was a short tube that was able to constrict and relax on command, and it was visible on MR images (figure 7.7). If multiple of these tubes are connected, a peristatic motion may be simulated, which is swallowing in its most basal form.

Even if a compressed sensing model is constructed with the optimal regularisation parameters that result in the highest image quality, the clinical implementation of this imaging technique may be hindered by the long reconstruction time of the images. In chapter six, a typical reconstruction took approximately three hours to complete on four CPU cores. As obtaining images may take such a long time, it is not possible to redo a scan because the subject, during the initial scan, moved his or her head or swallowed the bolus prematurely. Fortunately, efforts to accelerate the reconstruction time have been made.

A promising solution, developed by Schaetz et al., is to solve the inverse compressed-sensing model using a fast optimiser such as the conjugate gradient method, which is parallelised and performed on multiple graphical processing units (GPUs)[23]. This results in a reconstruction time per frame of less than a second if only a single GPU is used. However, some approximations were made regarding the total variation constraint.

Other applications for real-time MRI

Deep learning can also be inserted at other locations in the imaging pipe-line that the reconstruction step. Arguably the most researched areas are image classification and segmentation. In a publicly-available data set of real-time speech MRI, which is similar to our 2D real-time swallowing data, we were able to automatically classify phonemes (vowels and consonants) on single images from the video feed.[24] This model may be extended to identify pathological speech, or be used as a basis from which a similar model for swallowing MRI may be constructed. Unfortunately, we were however unable to replicate these results with our own data due to the aforementioned issues with the simultaneous acquisition of MRI data and speech.

In addition to the imaging of fast physiological processes such as speech and swallowing aiming for improving the diagnosis, real-time MRI can also be applied to therapies or interventions, especially if a low-latency reconstruction is implemented. If an organ such as the liver is tracked during breathing, a beam of radiation may be guided to target a metastasis. Similarly an ablation probe or other surgical instruments may be tracked in real-time. Compared to similar techniques using X-rays or ultrasound, MRI would allow for simultaneous probe tracking and high-contrast imaging of anatomical structures.

Finally, real-time MRI of swallowing, especially the variant in 3D, may be used to validate finite-element models (FEMs) of the tongue or the head and neck in general. In addition, certain model assumptions, such as certain parameter values of the FEM may be determined through inverse modelling. Although we made a first effort in personalising a FEM based on CSD MRI data, many more aspects of the model contain assumptions that may need to be personalised, potentially by MRI data.

Finite-element modelling

Not only the muscular architecture of the tongue, which contains the location and direction of muscles, but also the density of muscle fibres is an important factor for FE modelling, as a higher density of muscle fibres translates into a higher muscle strength. In chapter three, we calculated a metric, the apparent fibre density, which is thought to represent the fibre density of a fibre population. In future FEMs, a higher AFD should therefore translate to stronger muscle fibres. However, the assumption that AFD correlates to fibre density still has to be confirmed in histological studies.

Not only the muscles in an FEM are important for simulating physiological processes such as swallowing. The elastic and/or rigid structures such as the lingual septum or bones play an important role in the dynamics of swallowing and speech, but are currently not personalised in our models. In chapter four, bony attachments were merely modelled by fixing certain nodes in the FE model.

Ideally, for a well-personalised FEM, rigid structures such as the bones should be personalised as well.

A similar approach to the implementation of the tongue musculature using a CSD-based atlas (chapter three) could be taken by creating an atlas of the bones. Based on T1-weighted MRI contrast, we made an initial effort for this bone atlas by first generating an atlas of the mandible. This approach was subsequently applied to the parotid glands and multiple MR contrasts indicating that an atlas can be created for hard and soft tissues. For FE modelling, a better approach would be an atlas of all anatomical structures in the head-and-neck, which also contains information about the position of these anatomical structures relative to one another.

However, even if all assumptions of the FEM model can be validated in a model with tiny elements, which should result in accurate simulations of swallowing and speech, then the calculation of these simulations would still take a long time. If a surgeon wants to interactively explore the outcome after different surgical approaches, a long calculation time would impair the clinical application of FE modelling. This necessitates a significant acceleration of computation speed.

Initially, the efforts in accelerating the computing time of FEM have focussed on applying the principal component analysis (PCA) to FE modelling.[25] Although the PCA does accelerate the computing time, it also aggravates the error in the model. Acceleration of FEMs has however been achieved with lower errors by using a deep-learning approach.[25] Although these machine learning techniques are promising for FE modelling, they also introduce the question: Why not replace the entire FEM by a deep-learning model? However, there is no deep learning with big data, and as there are no large (publicly-available) data sets of (MRI of) swallowing and speech, completely replacing the FEM would be difficult.

In any case, whether by finite element modelling or deep learning, personalised modelling of treatment outcomes should be encouraged. At the very least, it may support physicians in choosing the optimal treatment for their patients. Furthermore, provided that the information can be presented in a way that is easy to understand, patients may benefit from the same better informed choice of treatment. Finally, if the expected causes in loss of functionality is known, post-treatment rehabilitation may be improved by focussing on these causes. Thus, personalised modelling fits perfectly with the current trends in personalised medicine to improve overall patient care.

References

1. J. L. R. Andersson and S. N. Sotiropoulos, "An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging," *NeuroImage*, vol. 125, pp. 1063–1078, 2016. DOI: 10.1016/j.neuroimage.2015.10.019.

2. Q. Zhang, B. F. Coolen, A. J. Nederveen, and G. J. Strijkers, "Three-dimensional diffusion imaging with spiral encoded navigators from stimulated echoes (3D-DISPENSE)," *Magnetic Resonance in Medicine*, vol. 81, no. 2, pp. 1052–1065, 2019. DOI: 10.1002/mrm.27470.

3. C. Zhang, T. M. Arefin, U. Nakarmi, C. H. Lee, H. Li, D. Liang, J. Zhang, and L. Ying, "Acceleration of three-dimensional diffusion magnetic resonance imaging using a kernel low-rank compressed sensing method," *NeuroImage*, vol. 210, no. October 2019, 2020. DOI: 10.1016/j.neuroimage.2020.116584.

4. B. Jeurissen, J. D. Tournier, T. Dhollander, A. Connelly, and J. Sijbers, "Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data," *NeuroImage*, vol. 103, pp. 411–426, 2014. DOI: 10.1016/j.neuroimage.2014.07.061.

5. A. Singh, D. Reddy, M. Haris, K. Cai, K. Rajender Reddy, H. Hariharan, and R. Reddy, "T1rhoMRI of healthy and fibrotic human livers at 1.5 T," *Journal of Translational Medicine*, vol. 13, no. 1, pp. 1–7, 2015. DOI: 10.1186/s12967-015-0648-0.

6. S. Cheng, S. C. Gandevia, M. Green, R. Sinkus, and L. E. Bilston, "Viscoelastic properties of the tongue and soft palate using MR elastography," *Journal of Biomechanics*, vol. 44, no. 3, pp. 450–454, 2011. DOI: 10.1016/j.jbiomech.2010.09.027.

7. M. J. A. van Alphen, *Towards a predictive model for functional loss after oral cancer treatment*. 2015, pp. 9–28. DOI: 10.3990/1.9789036539173.

8. M. Yoneyama, T. Takahara, T. C. Kwee, M. Nakamura, and T. Tabuchi, "Rapid High Resolution MR Neurography with a Diffusion-weighted Pre-pulse," *Magnetic Resonance in Medical Sciences*, vol. 12, no. 2, pp. 111–119, 2013. DOI: 10.2463/mrms.2012-0063.

9. J. Oudeman, B. F. Coolen, V. Mazzoli, M. Maas, C. Verhamme, W. M. Brink, A. G. Webb, G. J. Strijkers, and A. J. Nederveen, "Diffusion-prepared neurography of the brachial plexus with a large field-of-view at 3T," *Journal of Magnetic Resonance Imaging*, n/a–n/a, 2015. DOI: 10.1002/jmri.25025.

10. F. M. Zijta, M. M. E. Lakeman, M. Froeling, M. P. van der Paardt, C. S. V. Borstlap, S. Bipat, a. D. Montauban van Swijndregt, G. J. Strijkers, J. P. Roovers, A. J. Nederveen, and J. Stoker, "Evaluation of the female pelvic floor in pelvic organ prolapse using 3.0-Tesla diffusion tensor imaging and fibre tractography," *European Radiology*, vol. 22, no. 12, pp. 2806–2813, 2012. DOI: 10.1007/s00330-012-2548-5.

11. M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007. DOI: 10.1002/mrm.21391.

12. M. Uecker, T. Hohage, K. T. Block, and J. Frahm, "Image reconstruction by regularized nonlinear inversion - Joint estimation of coil sensitivities and image content," *Magnetic Resonance in Medicine*, vol. 60, no. 3, pp. 674–682, 2008. DOI: 10.1002/mrm.21691.

13. K. S. Nayak, *Response to Letter to the Editor: "Nomenclature for real-time magnetic resonance imaging"*, 2019. DOI: 10.1002/mrm.27770.

14. S. Wundrak, J. Paul, J. Ulrici, E. Hell, and V. Rasche, "A small surrogate for the golden angle in time-resolved radial MRI based on generalized fibonacci sequences," *IEEE Transactions on Medical Imaging*, vol. 34, no. 6, pp. 1262–1269, 2015. DOI: 10.1109/TMI.2014.2382572.

15. S. Rosenzweig, H. C. M. Holme, R. N. Wilke, D. Voit, J. Frahm, and M. Uecker, "Simultaneous multi-slice MRI using cartesian and radial FLASH and regularized nonlinear inversion: SMS-NLINV," *Magnetic Resonance in Medicine*, vol. 79, no. 4, pp. 2057–2066, 2018. DOI: 10.1002/mrm.26878.

16. S. M. Felton, T. A. Gaige, T. G. Reese, V. J. Wedeen, and R. J. Gilbert, "Mechanical basis for lingual deformation during the propulsive phase of swallowing as determined by phase-contrast magnetic resonance imaging.," *Journal of applied physiology*, vol. 103, no. 1, pp. 255–265, 2007. DOI: 10.1152/japplphysiol.01070.2006.

17. V. Mazzoli, L. M. Gottwald, E. S. Peper, M. Froeling, B. F. Coolen, N. Verdonschot, A. M. Sprengers, P. van Ooij, G. J. Strijkers, and A. J. Nederveen, "Accelerated 4D phase contrast MRI in skeletal muscle contraction," *Magnetic Resonance in Medicine*, vol. 80, no. 5, pp. 1799–1811, 2018. DOI: 10.1002/mrm.27158.

18. L. M. Gottwald, E. S. Peper, Q. Zhang, B. F. Coolen, G. J. Strijkers, A. J. Nederveen, and P. van Ooij, "Pseudo-spiral sampling and compressed sensing reconstruction provides flexibility of temporal resolution in accelerated aortic 4D flow MRI: A comparison with k-t principal component analysis," *NMR in Biomedicine*, vol. 33, no. 4, pp. 1–13, 2020. DOI: 10.1002/nbm.4255.

19. M. Uecker, F. Ong, J. I. Tamir, D. Bahri, P. Virtue, J. Y. Cheng, T. Zhang, and M. Lustig, "Berkeley Advanced Reconstruction Toolbox," in *Proc. Intl. Soc. Mag. Reson. Med*, Toronto, 2015. DOI: 10.5281/zenodo.592960.

20. T. Zhang, J. M. Pauly, and I. R. Levesque, "Accelerating parameter mapping with a locally low rank constraint," *Magnetic Resonance in Medicine*, vol. 73, no. 2, pp. 655–661, 2015. DOI: 10.1002/mrm.25161.

21. L. Feng, R. Grimm, K. T. Block, H. Chandarana, S. Kim, J. Xu, L. Axel, D. K. Sodickson, and R. Otazo, "Golden-angle radial sparse parallel MRI: combination of compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric MRI," *Magnetic resonance in medicine*, vol. 72, no. 3, pp. 707–717, 2014. DOI: 10.1002/mrm.24980. arXiv: NIHMS150003.

22. M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing MRI," *IEEE Signal Processing Magazine*, vol. 25, no. March 2008, pp. 72–82, 2008. DOI: Doi10.1109/Tit.2006.871582. arXiv: 1204.4227v1.

23. S. Schaetz, D. Voit, J. Frahm, and M. Uecker, "Accelerated Computing in Magnetic Resonance Imaging: Real-Time Imaging Using Nonlinear Inverse Reconstruction," *Computational and Mathematical Methods in Medicine*, vol. 2017, 2017. DOI: 10.1155/2017/3527269. arXiv: 1701.08361.

24. K. G. van Leeuwen, P. Bos, S. Trebeschi, M. J. van Alphen, L. Voskuilen, L. E. Smeele, F. van der Heijden, and R. J. van Son, "CNN-based phoneme classifier from vocal tract MRI learns embedding consistent with articulatory topology," *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, pp. 909–913, 2019. DOI: 10.21437/Interspeech.2019-1173.

25. F. Roewer-despres, N. Khan, and I. Stavness, "Towards Finite-Element Simulation Using Deep Learning," *15th International Symposium on Computer Methods in Biomechanics and Biomedical Engineering*, 2018.

Chapter 8

8

Summary

Samenvatting

Author contributions

Dankwoord

Summary

The tongue is a muscle involved in a wide variety of vital functions such as speaking, chewing, and swallowing and therefore has to perform a wide range of different motions. These motions are based on a complex muscular architecture of interdigitating muscle fibres and, in contrast to other skeletal muscles, on the incompressibility of water. In patients with oral or oropharyngeal cancer the functionality of the tongue may become severely affected, resulting in difficulties with speaking chewing of swallowing. Moreover, the treatment of tongue cancer often involves a surgical resection or irradiation, causing even more damage to the functionality of the tongue.

In order plan the tongue cancer treatment with minimal loss of functionality, or to improve the counselling of the patient, the patient's tongue functionality after treatment tongue cancer should be estimated. Currently, this after-treatment functionality can be estimated by biomechanical modelling using a finite-element model that is constructed from post-mortem data of a single subject. However, in-vivo information of the tongue's muscular architecture of an individual patient should be integrated into the biomechanical model to better estimate the port-treatment functionality. In this thesis, we improved currently-available MRI techniques in order to obtain in-vivo information of the tongue's musculature and it's functionality and use this information to improve biomechanical model of the tongue.

The MRI technique diffusion tensor imaging (DTI) is able to reconstruct the muscular architecture of muscles in vivo. The direction of muscle fibres is estimated from the anisotropic diffusion of water in muscles, caused by intra- and extracellular structures such as proteins and cell membranes. However, DTI averages the directions of fibres that interdigitate microscopically, such as the muscles in the tongue, resulting in incorrect reconstructions of the tongue's musculature. In chapter two, we applied the MRI method constrained spherical deconvolution (CSD) to diffusion MRI of the tongue, which resulted in more accurate in-vivo reconstructions of the tongue's muscular architecture than with DTI.

These CSD-based reconstruction could not be directly integrated into the present biomechanical models of the tongue, as the different tongue muscles could not be distinguished. Although these muscles could be segmented manually, this process would be time-consuming and prone to user error. Therefore, in chapter three, we introduced the concept of population averaging to CSD of the tongue. By registering (or morphing) the CSD images to a common template (also called a population average or atlas), the segmentation of the tongue can be transferred from the population average to the individual's tongue. This reduces user errors and saves time, as only one population average has to be segmented. Moreover, this methodology allows the analysis of differences between different populations. We showed that the metric apparent fibre density is significantly lower in several

tongue muscles in elderly populations than in a younger population.

Using these new population averages of the tongue, we automated the generation personalised biomechanical models of the tongue in chapter four. By firstly constructing a biomechanical model of the population average, and by secondly mapping the CSD image of an individual to this population average, we were able to morph the biomechanical model into the shape of the tongue of that individual. Furthermore, we showed that the range-of-motion predictions of the personalised tongues were more accurate than that of the average model.

Although CSD MRI is a promising technique for estimating the loss of tongue functionality, it's largest draw-back is the long scan time. Such a long scan time in not only more costly, but also increases the chance of motion artefacts distorting the images. Therefore, in chapter five, we designed and built a custom 12-channel MRI receiver coil for tongue imaging. Using this coil, we were able to accelerate the CSD scan resulting in the halving of the scan time.

Furthermore, this acceleration provided by the custom coil allowed us to implement real-time MRI, in the form of single-slice videos of swallowing. Such single slice real-time MRI videos might improve the assessment of impaired swallowing functionality, and may improve the predictions made by biomechanical models. However, as only a single slice is imaged, swallowing dysfunctionality such as aspiration may be missed if it occurs outside the imaged slice. In chapter six, we therefore developed a real-time MRI sequence that covers a 3D volume instead of a slice, at a frame rate of 12 frames per second.

Samenvatting

De tongspier is betrokken in een grote verscheidenheid aan lichaamsfuncties zoals spraak, kauwen en slikken, en moet daarom veel verschillende bewegingen kunnen uitvoeren. In tegenstelling tot andere skeletspieren zijn deze bewegingen gebaseerd op de complexe spierarchitectuur van in elkaar grijpende spiervezels en de onsamendrukbaarheid van water. Bij patiënten met mond- of keelkanker kan de functionaliteit van de tong zwaar zijn aangedaan, wat kan zorgen voor spraak- of slikproblemen. Daarbovenop komt nog de behandeling van tongkanker, vaak chirurgische verwijdering en/of bestraling, wat nog meer schade aan de tongfunctionaliteit veroorzaakt.

Om een behandeling voor tongkanker te plannen met zo min mogelijk verlies van functionaliteit, of om de voorlichting van de patiënt te verbeteren, zou de tongfunctionaliteit na behandeling moeten worden bepaald. Momenteel kan deze functionaliteit na behandeling worden benaderd door het gebruik van biomechanisch modelleren met een eindig-elementenmodel opgebouwd aan de hand van post mortem gegevens. In vivo informatie van de tongmusculatuur van de patiënt zou echter moeten worden toegepast op dit biomechanische model om de tongfunctionaliteit na behandeling beter te kunnen bepalen. In dit proefschrift hebben we reeds beschikbare MRI technieken verbeterd, zodat we in vivo informatie kunnen vergaren van de musculatuur en de functionaliteit van de tong. Deze informatie kunnen we vervolgens gebruiken om het biomechanisch modelleren van de tong te verbeteren.

De MRI-techniek diffusion tensor imaging (DTI) kan de spierarchitectuur in vivo reconstrueren. Door de anisotrope diffusie van water in spieren te meten, die wordt veroorzaakt door intra- en extracellulaire structuren als proteïnen en celmembranen, kan de richting van spiervezels worden bepaald. Echter, DTI middelt de richtingen uit van de spiervezels, die op een microscopisch niveau verweven zijn, bijvoorbeeld in de tong, wat resulteert in foutieve reconstructies van de tongmusculatuur. In hoofdstuk twee hebben we de MRI-methode constrained spherical deconvolution (CSD), die verweven spiervezels wel kan onderscheiden, toegepast op MRI van de tong, waardoor er nauwkeurigere in vivo reconstructies van de tongmusculatuur konden worden gemaakt.

Deze reconstructies gebaseerd op CSD konden niet zomaar worden toegepast in het bestaand biomechanisch model van de tong, omdat de verschillende tongspieren niet konden worden onderscheden. Hoewel deze spieren met de hand zouden kunnen worden gesegmenteerd, zou dit proces tijdrovend zijn, en vatbaar zijn voor fouten van de beoordelaar. Daarom hebben we in hoofdstuk drie het concept populatiegemiddelde geïntroduceerd aan CSD van de tong. Door CSD beelden van de tong te registreren naar (of te vervormen tot) een gedeeld sjabloon (het populatiegemiddelde of de atlas), kan de segmentatie van dit sjabloon overgebracht worden naar de tongreconstructie van een individu. Dit vermindert fouten

van de beoordelaars en bespaart tijd, omdat er slechts het populatiegemiddelde gesegmenteerd hoeft te worden. Bovendien kunnen we met deze methodiek verschillen tussen verschillende populaties analyseren. Zo konden we aantonen dat de maat 'apparent diffusion density' significant lager was in verschillende tongspieren in oudere populaties dan in een jongere populatie.

Door middel van deze nieuwe populatiegemiddelden van de tong, hebben we in hoofdstuk vier het genereren van biomechanische modellen van de tong geautomatiseerd. Door eerst een biomechanisch model van het populatiegemiddelde aan te maken, en vervolgens de CSD beelden van een individu aan dat van het populatiegemiddelde te koppelen, konden we het biomechanisch model vervormen tot de tong van het individu. Bovendien toonden we aan dat de voorspellingen over de bewegelijkheid van deze gepersonaliseerde tongmodellen nauwkeuriger waren dan die van het populatiegemiddelde.

Hoewel CSD MRI een veelbelovende techniek is om het functionaliteitsverlies van de tong te bepalen, heeft deze techniek een groot nadeel: de lange scantijd. Dit zorgt niet alleen voor hoge kosten, maar vergroot ook de kans op bewegingsartefacten die de beelden verstoren. Daarom hebben we in hoofdstuk vijf een 12-kanaals MRI ontvangstspoel naar ons ontwerp laten maken specifiek voor MRI van de tong. Door middel van deze spoel konden we de CSD scan zo versnellen dat de scantijd was gehalveerd.

Daarnaast hebben we de versnelling, die deze spoel toelaat, gebruikt om een real-time MRI scan te implementeren die slikvideo's produceert van een enkele plak. Dergelijke video's van real-time MRI zouden niet alleen de beoordeling van slikproblemen, maar ook de voorspellingen van biomechanische modellen kunnen verbeteren. Omdat er slechts één enkele plak wordt gescand, kunnen slikproblemen zoals aspiratie worden gemist als deze plaatsvinden buiten de plak die is gescand. In hoofdstuk zes hebben we daarom een MRI sequentie ontwikkeld die een 3D volume beslaat in plaats van een plak met een verversingssnelheid van 12 beelden per seconde.

Author contributions

Listed below are the author contributions of the peer-reviewed chapters in alphabetical order:

Chapter	*Study conception and design*	*Acquisition of data*	*Analysis and interpretation of data*	*Drafting of manuscript*	*Critical revision*
2	AB; AN; FH; JO; LS; LV; VM	LV	AN; GS; LV; MF; VM	LV	AB; AN; FH; GS; JO; LS; LV; MF; MW; VM
3	AN; LV	LV	AJ; LV; VM	LV	AB; AN; GS; LS; LV; VM
4	FH; KK; LV	KK; LV	AB; FH; KK; LV	KK; LV	AB; AN; BJ; FH; KK; LS; LV
5	AN; LM; LV;PH	LV	AN; GS; LM; LV	LV	AB; AN; FH; GS; LM; LS; LV; PH
6	AN; LV	LV; OG	AN; GS; JS; LV; OG	LV	AB; AN; GS; LS; LV; OG

AB	Alfons J.M. Balm
AN	Aart J. Nederveen
BJ	Bas Jasperse
FH	Ferdinand van der Heijden
GS	Gustav J. Strijkers
JO	Jos Oudeman
JS	Jasper Schoormans
KK	Kilian D.R. Kappert
LM	Lisette van der Molen
LS	Ludi E. Smeele
LV	Luuk Voskuilen
MF	Martijn Froeling
MW	Maartje M.L. de Win
OG	Oliver J. Gurney-Champion
PH	Paul de Heer
VM	Valentina Mazzoli

Dankwoord

Van de colleges wetenschapsfilosofie op de Universiteit Twente, is één quote me altijd bijgebleven. Deze quote, onder meer toegeschreven aan Newton, luidt: "If I have seen further it is by standing on the shoulders of giants." Hoewel deze quote origineel de bedoeling had om aan te geven dat je in de wetenschap altijd voortbouwt op het werk van anderen, kan je dit ook zonder meer toepassen op de totstandkoming van dit boekje. Dit had nooit kunnen worden bereikt, zonder de steun van anderen, veel grotere personen dan ik.

Prof. Smeele, Ludi, allereest wil ik je bedanken voor het vertrouwen dat je in mijn hebt gesteld als promovendus. Tijdens onze lunches waardeerde ik je pragmatische blik op mijn werk en problemen die naar boven kwamen tijdens mijn promotie. Dit hielp enorm om gedachten op orde te krijgen en met een frisse blik weer verder te gaan.

Prof. Nederveen, Aart, ook jou wil ik bedanken voor het vertrouwen en de kans die je me hebt geboden om mijn promotie deels op het AMC uit te mogen voeren. Zonder de kennis en kunde die je op Z0 hebt verzameld, zou ik nooit tot dit, al zeg ik het zelf, diverse proefschrift zijn gekomen. Onze wekelijkse brainstorm heeft tot interessante discussies en veel uitstapjes naar aanverwante projecten geleid, waarvan een deel is opgenomen in de discussie van dit proefschrift.

Prof. Balm, Fons, ik wil je graag bedanken voor je idealistische blik op de wetenschap en geneeskunde, die voor mij vooral tot uiting kwam in het Virtual Therapy project en de klankbordsessies met de patiëntvertegenwoordigers. Je hebt een duidelijk doel hoe je de zorg voor je patiënten het best kon verbeteren, en hielp mij ditzelfde doel voor ogen te houden tijdens mijn promotie.

Dr. Van der Heijden, Ferdi, natuurlijk bedankt voor het aandragen van mij als promovendus voor dit project bij het Antoni van Leeuwenhoek. Ik vond het altijd fijn om mijn werk aan je kritische blik van buiten de MRI-wereld te spiegelen. Niet alleen leidde dit tot een beter begrip bij mij van de technisch complexe materie, maar ook tot nieuwe mogelijkheden waar de MRI-wereld nog niet eerder aan had gedacht.

Uiteraard wil ik ook de overige leden van mijn promotiecommissie bedanken. Dr. De Win, prof. Schijven, prof. Becking, prof. Ten Haken, en prof. Webb: Hartelijk dank voor jullie toetreden tot mijn promotiecommissie, en de kritische beoordeling van mijn proefschrift. Ik kijk uit naar de aanstaande gedachtewisseling tijdens de verdediging van dit werk.

Verder wil ik alle vrijwilliger en patiënten bedankten die hebben willen deelnemen aan de studies in dit proefschrift. Zonder hen zou dit proefschrift alleen uit interessante maar onbeproefde methodes hebben bestaan. Hetzelfde geldt voor de Maurits en Anna de Kock Stichting. Zonder haar financiële steun hadden we de tongspoel, gebruikt in hoofdstuk 5 en 6, nooit kunnen laten maken, en hadden we nooit MRI-slikvideo's kunnen maken van deze kwaliteit.

Naast de bovenstaande personen, was de wetenschappelijke inhoud van dit proefschrift nooit tot dit niveau gekomen zonder de inzet van de coauteurs. Bas Jasperse, Gustav Strijkers, Jasper Schoormans, Jos Oudeman, Kilian Kappert, Lisette van der Molen, Martijn Froeling, Oliver Gurney-Champion, Paul de Heer, en Valentina Mazzoli, bedankt voor jullie onmisbare bijdrage aan dit proefschrift. Ik waardeer jullie opbouwende kritiek, maar ook jullie hulp en steun enorm.

Bovendien ben ik dank verschuldigd aan alle nog niet genoemde collega's uit het Antoni van Leeuwenhoek: De onderzoekers uit de Virtual Therapy groep, Maarten, Merijn, en Rob voor het eindeloos aanhoren van mijn MRI-praatjes. Anita voor haar onschatbare hulp bij het inplannen van de patiëntscans. Paula, Joost, Leon, en andere collega's die met mij hebben gediscussieerd over MRI technieken en machine learning. Bedankt aan alle studenten die ik tijdens mijn promotie heb mogen begeleiden. Kilian, Bence en Rebecca, bedankt voor alle gezellige momenten op onze kamer, ook al was ik slechts een parttimer.

Voor de andere helft van de tijd ben ik natuurlijk ook dank verschuldigd aan alle nog niet genoemde collega's uit het Academisch Medisch Centrum: Bobby, Carmen, Eva, Jasper, Jithsa, Jos, Kerry, Koen, Lena, Liza, Lukas, Oliver, Paul, Sophie, en Valentina, bedankt voor alle borrels en feestjes in en rond club Z0-178. Joena, Paul, Raschel, en Sandra, jullie zijn van onschatbare waarde voor Z0 (geweest). En ten slotte alle andere onderzoekers van Z0 die mijn praatjes hebben aangehoord: Anouk, Antonia, Bram, Claudia, Esther, Geor, Jules, Laura, Mariah, Marieke, Matthan, Melissa, Ot, Pim, Sofieke, en Susi.

Bruce, Rob, Seline, Thijs, en Thijs bedankt voor alle ontspanning gedurende mijn promotie, en jullie steun en luisterend oor, met name tijdens de laatste loodjes van dit promotietraject. Ryanne, bedankt voor je ongefilterde mening, de onverwachte cadeautjes, maar bovenal onze vriendschap ondanks dat je aan de andere kant van het land woont, en ik nu bij de concurrent werk. Rob, ik vind het bijzonder dat ik je na bijna 28 jaar nog steeds mijn beste vriend mag noemen. Samen hebben we al genoeg meegemaakt, en ik blij dat je de verdediging samen met mij wilt doen.

Ten slotte kan ik natuurlijk mijn familie niet vergeten: Oma, Ronald, Rob en Thijs, Suze en Vincent, Peter en Marian, en Feline en Antoine, bedankt voor al jullie steun. Conny, bedankt voor de mental support op de momenten dat ik het het meest nodig had. Miek, mama, bedankt dat ik altijd op je heb kunnen, en kan terugvallen en me altijd een schop onder mijn kont gaf als dat nodig was.

Dit proefschrift is niet het belangrijkste resultaat van mijn promotietraject. Dat is voorbehouden aan het feit dat ik tijdens mijn promotie jou, Michelle, heb ontmoet. Ook al zeg ik het te weinig, maar ik ben nog elke dag dankbaar dat je bij me bent, en dat je me door dik en dun steunt.

"You did it. You crazy son-of-a-bitch, you did it."
- Ian Malcolm [seeing the dinosaurs for the first time]