Magnetic Resonance in Medicine 69:477–485 (2013)

Real-Time MRI of Speaking at a Resolution of 33 ms: Undersampled Radial FLASH with Nonlinear Inverse Reconstruction

Aaron Niebergall,[1] Shuo Zhang,[1]* Esther Kunay,[2] Götz Keydana,[2] Michael Job,[2] Martin Uecker,[1] and Jens Frahm[1]

Dynamic MRI studies of the upper airway during speaking, singing or swallowing are complicated by the need for high temporal resolution and the presence of air-tissue interfaces that may give rise to image artifacts such as signal void and geometric distortions. This work exploits a recently developed real-time MRI technique to address these challenges for monitoring speech production at 3 T. The method combines a short-echo time radial FLASH MRI sequence (pulse repetition time/echo time = 2.22/1.44 ms; flip angle 5°) with pronounced undersampling (15 radial spokes per image) and image reconstruction by regularized nonlinear inversion. The resulting serial images at 1.5 mm in-plane resolution and 33.3 ms acquisition time are free of motion or susceptibility artifacts. This application focuses on a dynamic visualization of the main articulators during natural speech production (Standard Modern German). Respective real-time MRI movies at 30 frames per second clearly demonstrate the spatiotemporal coordination of lips, tongue, velum, and larynx for generating vowels, consonants, and coarticulations. The quantitative results for individual phonetic events are in agreement with previous non-MRI findings. **Magn Reson Med 69:477–485, 2013. © 2012 Wiley Periodicals, Inc.**

Key words: real-time MRI; speaking; articulation; tongue dynamics; speech production; phonetics

Since the earliest studies of human articulation in the late 18th century, methods and applications have continuously been expanded. Nowadays, articulatory processes during speaking may be analyzed by a wide range of techniques including laryngoscopy and laryngography, electropalatography, electromyography, electromagnetic articulography, or imaging based on X-ray videofluoroscopy or sonography (1). However, while former approaches require an invasive procedure, which eventually disturbs the natural articulation of vowels and consonants, videofluoroscopy involves the exposure to radiation and sonography is restricted to specific image orientations.

[1]Biomedizinische NMR Forschungs GmbH am Max-Planck-Institut für biophysikalische Chemie, Göttingen, Germany.

[2]Sprachwissenschaftliches Seminar, Georg-August-Universität, Göttingen, Germany.

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Shuo Zhang, Biomedizinische NMR Forschungs GmbH am Max-Planck-Institut für biophysikalische Chemie, 37070 Göttingen, Germany. E-mail: szhang1@gwdg.de

Received 22 December 2011; revised 14 February 2012; accepted 9 March 2012.

DOI 10.1002/mrm.24276

Published online 12 April 2012 in Wiley Online Library (wileyonlinelibrary.com).

Although MRI promises noninvasive examinations as well as more flexible and detailed insights into the vocal tract, insufficient temporal resolution, and limited image quality still emerge as two major obstacles for studying speech production. For example, because long image acquisition times require either repetitively generated (2,3) or continuant (4) speech sounds, such strategies preclude a proper visualization of the lips and tongue for plosive consonants that are characterized by rapid movements within 50–100 ms (5). In addition, motion-induced image blurring may result from data sharing when using long acquisitions with sliding-window reconstructions to achieve high frame rates (6,7). It has therefore been concluded that the true temporal resolution for MRI should be comparable to that of standard videofluoroscopy, ideally no longer than 40 ms per image (8). Other problems are the occurrence of focal signal losses and geometric distortions that depend on the sensitivity of a chosen MRI acquisition technique to the unavoidable susceptibility differences near air-tissue interfaces in the upper airway (9,10).

Most recently, we have described a real-time MRI technique using highly undersampled radial FLASH with image reconstruction by regularized nonlinear inversion (11–13). Preliminary applications to cardiovascular function (14), quantitative blood flow (15), and normal swallowing (16) allow for serial real-time images (i.e., movies) with 1.5 to 2.0 mm in-plane resolution and 20–30 ms acquisition time. Accordingly, the purpose of this study was to exploit this real-time MRI method to address the challenges posed by dynamic MRI of speaking and to investigate the key articulatory configurations during natural speech production including vowels, consonants, and coarticulation effects in vowels preceded by consonants.

METHODS

Subjects

Twelve subjects (6 men and 6 women; age 28 ± 8 years [mean ± standard deviation]; range 23–54 years) with no known illness were recruited from the local University. All subjects gave written informed consent before each MRI examination and were paid for their participation in this study.

Magnetic Resonance Imaging

All studies were conducted at 3 T using a commercial MRI system (TIM Trio, Siemens Healthcare, Erlangen,

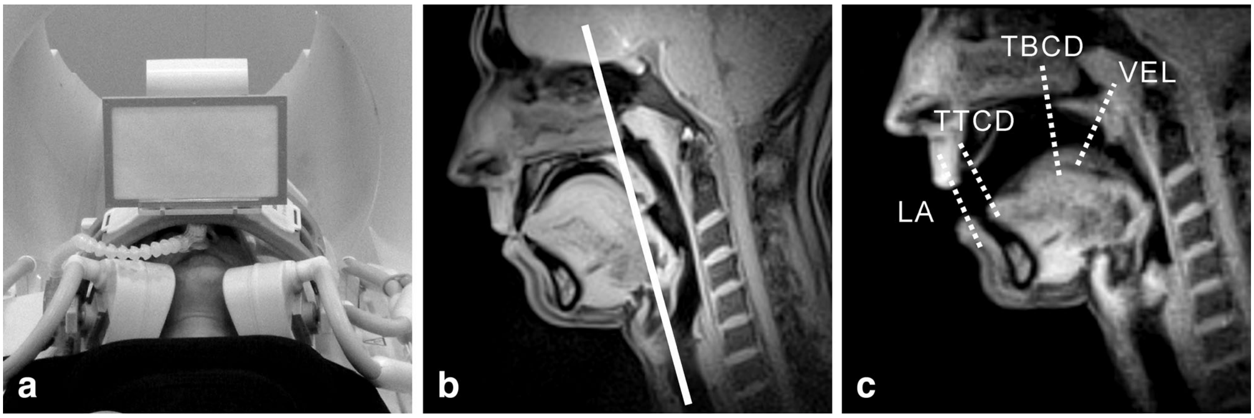

FIG. 1. a: Setup for real-time MRI of speaking. The arrangement combines a flexible 4-channel receiver coil covering the lower face with a bilateral 2 × 4 array coil centered to the thyroid prominence and further involves an MR-compatible optical microphone and projection system (screen and mirror). b: T1-weighted scout image (radial FLASH, pulse repetition time/echo time = 2.22/1.44 ms, flip angle 5°, and 1.5 × 1.5 × 10 mm^3) used for defining a midsagittal and coronal (solid line) plane for real-time MRI. c: Reference lines for deriving spatiotemporal patterns for the lip aperture (LA), tongue tip constrict degree (TTCD), tongue body constrict degree (TBCD), and velum aperture (VEL) during speaking (compare Figs. 6, 8, and 9).

Germany) and a body coil for radiofrequency (RF) excitation. Subjects were examined in a supine position and MRI signals were acquired by combining a small flexible 4-channel receiver coil covering the lower face (Siemens Healthcare, Erlangen, Germany) with a bilateral 2 × 4 array coil centered to the thyroid prominence on both sides of the neck (NORAS MRI products, Hoechberg, Germany) as shown in Fig. 1a.

Successive image acquisitions during speech production relied on a highly undersampled RF-spoiled radial FLASH MRI sequence with image reconstruction by regularized nonlinear inversion as described (13). The imaging parameters were: repetition time pulse repetition time = 2.22 ms, echo time echo time = 1.44 ms, flip angle 5°, field of view FOV 192 × 192 mm^2, in-plane resolution 1.5 × 1.5 mm^2, and section thickness 10 mm. Individual images were obtained from a single set of 15 spokes, which resulted in a temporal resolution of 33.3 ms or 30 fps.

Prior to dynamic MRI, scout images (FOV 256 × 256 mm^2) were obtained in the mid-sagittal plane using the same FLASH sequence but with full radial sampling and conventional gridding reconstruction. An example is shown in Fig. 1b. Speech production was then studied by multiple real-time MRI movies in mid-sagittal and coronal orientations. The midsagittal plane covered the entire vocal tract from the lips, tongue, and nasopharynx to larynx as well as the upper airway including vocal folds. The coronal plane was located parallel and central to the trachea as indicated in Fig. 1b.

During data acquisition, online image control was ensured by sliding-window gridding reconstructions of 75 spokes obtained by combining five consecutive data sets each comprising 15 spokes at complementary positions (17). At the same time, the incoming data are automatically exported for immediate offline reconstruction by regularized nonlinear inversion to a computer equipped with 8 GTX580 graphical processing units each providing 512 processing cores (Nvidia, CA). Once the offline calculation is completed, the images are reimported into the MRI system's database. Typically, most reconstructions of a speaking study were already available at the end of a 20 min in-room time.

Speech Tasks and Acoustic Recording

Table 1 summarizes four sets of speech tasks. The first set of seven vowels in German was employed to study tongue gestures at vowel production under different linguistic conditions, i.e., isolated segments as well as segments in words and sentences, which vary with respect to coarticulation, intonation, and speech rate. The second set served to visualize consonant production using meaningless logatom words with a consonant-vowel-consonant-vowel structure. To characterize the main articulators without putative coarticulation effects, the first vowel was a long [a:], while the second was a short [a] in all cases. The third experiment involved real words with or without a nasal consonant ([n], [m], or [n]) between vowels (VCV structure) to identify coarticulation effects, i.e., the overlap of articulatory gestures such as the timing of velum activity. The final set of logatom words was acquired in a coronal plane as the optimal orientation for studying larynx motion (18). In general, individual MRI recordings lasted for up to 88 s, while the total examination time was about 15 min per subject.

All subjects were native speakers of Standard Modern German (Neuhochdeutsch) and instructed to speak at a natural rate. For producing isolated segments, they were asked to repeat the vowel of the preceding word. The speech tasks were generated on a computer and projected into the MRI magnet using a setup (video projector, lens, screen, and mirror) developed for MRI studies of human brain function (Fig. 1a). To allow for acoustic recordings, subjects were equipped with an MR-compatible optical microphone with integrated software for adaptive cancellation of the gradient noise (Dual Channel-FOMRI, Optoacoustics, or Yehuda, Israel). Its two channels were positioned central to the lips in a unidirectional and omnidirectional orientation. The speech

Table 1
Speech Tasks

Vowels									
Word	bist	Beet	Bett	Bus	Boot	Post		hat	
Segment	i	e	e	u	o	o		a	
Sentence		Zwölf große Boxkämpfer jagen den schwachen Viktor über den Sylter Deich							
		Vier hungrige Angler verputzen fröhlich in der Garage leckeren Käsekuchen							
IPA	[i]	[e:]	[ɛ]	[ʊ]	[o:]	[ɔ]		[a]	
Consonants									
Logatom	*bata*	*bada*	*bapa*	*baba*	*bafa*	*bava*	*baka*	*baga*	*bapfa*
IPA	[t]	[d]	[p]	[b]	[f]	[v]	[k]	[g]	[pf]
Logatom	*bassa*	*basa*	*baja*	*bacha*	*bala*	*bascha*	*bama*	*bana*	*bange*
IPA	[s]	[z]	[j]	[χ]	[l]	[ʃ]	[m]	[n]	[ŋ]
Coarticulation									
Word	Pate		Sahne						
Structure	V-C-V		V-C-V						
Larynx									
Logatom	*bafa*	*bava*	*basa*	*bassa*					
IPA	[f]	[v]	[z]	[s]					

IPA, international phonetic alphabet; V-C-V, vowel-consonant-vowel.

recording was triggered by the radial FLASH sequence and thus synchronized to the data acquisition.

Image Analysis

A quantitative image analysis was accomplished by a Speech Analysis Toolbox developed in-house and written in MATLAB (MathWorks, Natick, MA). In general, the image series and synchronous audio files were divided into small fractions each containing only one vowel (segment), word, or sentence. The articulatory configuration of a particular phonetic event was derived from the image best matching the time point of production of the corresponding vowel or consonant. The selection of images was also verified by a spectrogram analysis (www.praat.org). The tongue gesture during vowel production was characterized by determining the positions of the dorsum. Respective vowel locations were then mapped to a simple chart, the so-called vowel diagram, for each of the three linguistic conditions.

To assess the deformation of the vocal tract during speaking, quantitative spatiotemporal patterns were measured with a system of reference lines. As shown in Fig. 1c, such metrics included the lip aperture (LA), tongue tip constrict degree (TTCD), tongue body constrict degree (TBCD), and velum aperture (VEL). In contrast to a fixed reference system (19,20), the lines were positioned along the movement of the articulators, i.e., the opening and closing direction of the lips, the extending direction of the tongue tip, the rising direction of the tongue body, and the rising direction of the velum against the posterior pharyngeal wall, respectively.

RESULTS AND DISCUSSION

This real-time MRI study of speaking successfully visualized the dynamics of human articulation in all subjects: no scan needed to be repeated and no volunteer exhibited any problems due to the arrangement of RF coils and microphone. Moreover, the configuration and spatiotemporal coordination of the main articulators were well demonstrated without image artifacts. Details included the lip protrusion and opening, tongue gestures, velum opening, and larynx motion.

Image Quality

The proposed real-time MRI method reaches a temporal resolution that corresponds to 33 ms acquisitions for individual images or movies at 30 fps. A similar speed has recently been demonstrated to allow for a detailed visualization of all physiologic events during natural swallowing (16). These results not only benefit from the use of an advanced acquisition and reconstruction technique, but also from a combination with optimized RF coils that help to ensure adequate SNR. The sensitivity to susceptibility artifacts or off-resonance effects is drastically reduced by a high receiver bandwidth and very short echo time. The achievement of a truly short acquisition time not only avoids visible motion artifacts, but also minimizes any blurring of anatomic structures due to rapid movements. Figure 2 depicts selected movie frames of a subject pronouncing the word *Kanupolo*. Together, they outline the major phonetic events for this word of eight segments—a sequence even better visualized in Supporting Information Movie 1 (including audio). In addition, Supporting Information Movie 2 shows an MRI video (including audio) of a subject speaking a long sentence at natural speed. In this case, part of the movie demonstrates a signal below the velum (close to the larynx), which represents a partial volume effect with the uvular structure due to movement during speaking. It should also be noted that some frames exhibit residual streaking artifacts, which are caused by the high degree of data undersampling. They mainly emerge from anatomic regions that are characterized by a combination of high signal intensities and high spatial frequencies.

Vowel Production

So far, the tongue gestures during vowel production have exclusively been studied using static MRI of

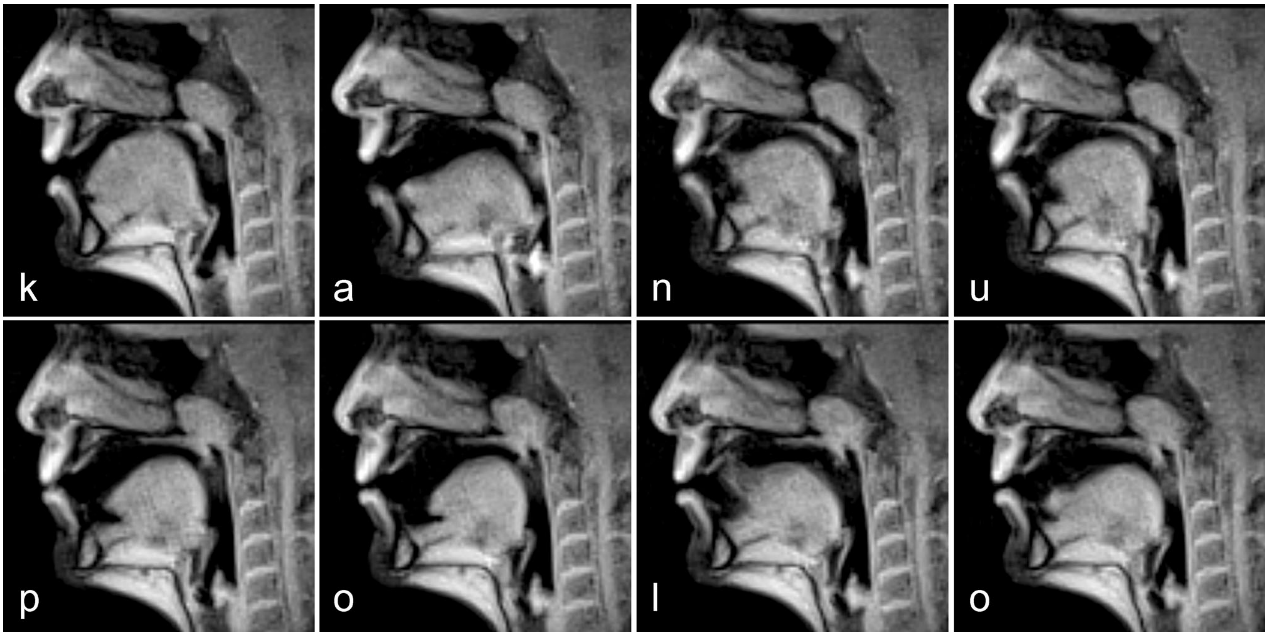

FIG. 2. Real-time MRI of uttering *Kanupolo*. Selected frames refer to the pronunciation of the eight phonetic events indicated. Individual images were obtained at 33 ms temporal resolution (15 spokes, radial FLASH, pulse repetition time/echo time = 2.22/1.44 ms, and flip angle 5°) and 1.5 mm in-plane resolution (10 mm section thickness). For a video and audio representation, see Supporting Information Movie 1.

isolated segments with sustained pronunciation. Here, vowel production was successfully captured for segments, words and sentences at a natural speaking rate (see Supporting Information Movie 2). For a single subject, Fig. 3 summarizes the contours of the tongue from tip to root for seven vowels (different colors) and three linguistic conditions. The reference images were chosen from a corresponding resting phase prior to speaking. The resulting highest tongue positions are then mapped in Fig. 4 (same subject) in comparison to an established vowel chart for Modern Standard German (21,22). It turns out that the different articulatory features of these vowels are clearly discriminated in agreement with literature findings (21).

Figure 3 already indicates that most of the contour lines are spread out and rather distinct for segments, but partially overlap for sentences. This typical observation is even better demonstrated in the diagrams of Fig. 4. From segment (solid line) to word (dashed) and sentence (dotted), the quantitative differences of the tongue gestures between vowels become reduced, whereas the general pattern of each diagram remains similar. Despite a certain variation between the 12 subjects, this finding reflects the fact that the shape of the tongue during vowel production to a certain degree depends on the articulatory condition (23).

Consonant Production

The configurations of the lips, different parts of the tongue and the velum as the main articulators for consonant production were well characterized by real-time MRI for all subjects. An example is shown in Fig. 5, where images at the time of consonant production were selected from the corresponding logatom word. Most of the articulatory gestures, i.e., the places of articulation, can unambiguously be distinguished by the position and contact (obstruction) of the articulators in the vocal tract (arrows in Fig. 5). Pertinent elements refer to the contact between the tongue tip and teeth ridge for alveolar consonants, the tongue tip and back of the teeth ridge for postalveolar consonants, the tongue body and hard palate for palatal consonants, the tongue body and soft palate for velar consonants, and the tongue body and uvula for uvular consonants.

Although alveolar and lateral alveolar consonants share the same place of articulation, a flatter contour, and lower position of the tongue lead to a different tongue shape for the latter. For bilabial consonants such as [p] and [m], the closing and contact of the upper and lower lips is well depicted. However, for labiodental consonants such as [f] and [v] a direct visualization of the contact between the upper lip, and the lower teeth is limited due to the poor visibility of the teeth. However, the nonstretched upper lip and the deformation of the lower lip due to pressure from the upper teeth differ from the pattern found for bilabial consonants. This is more easily demonstrated by the spatiotemporal pattern of the LA shown in Fig. 6 (first and second row). The detailed characterization of the affricate [pf] by the rapidly changing positions of the lips clearly benefits from the short image acquisition time. In particular, the brief contact of upper and lower lips at the initial phase of [p] and the maintaining of the lower lips shortly afterwards for contact with the upper teeth for [f] is well presented. This finding agrees with the fact that affricates begin as stops and release as fricatives (see Supporting Information Movie 3).

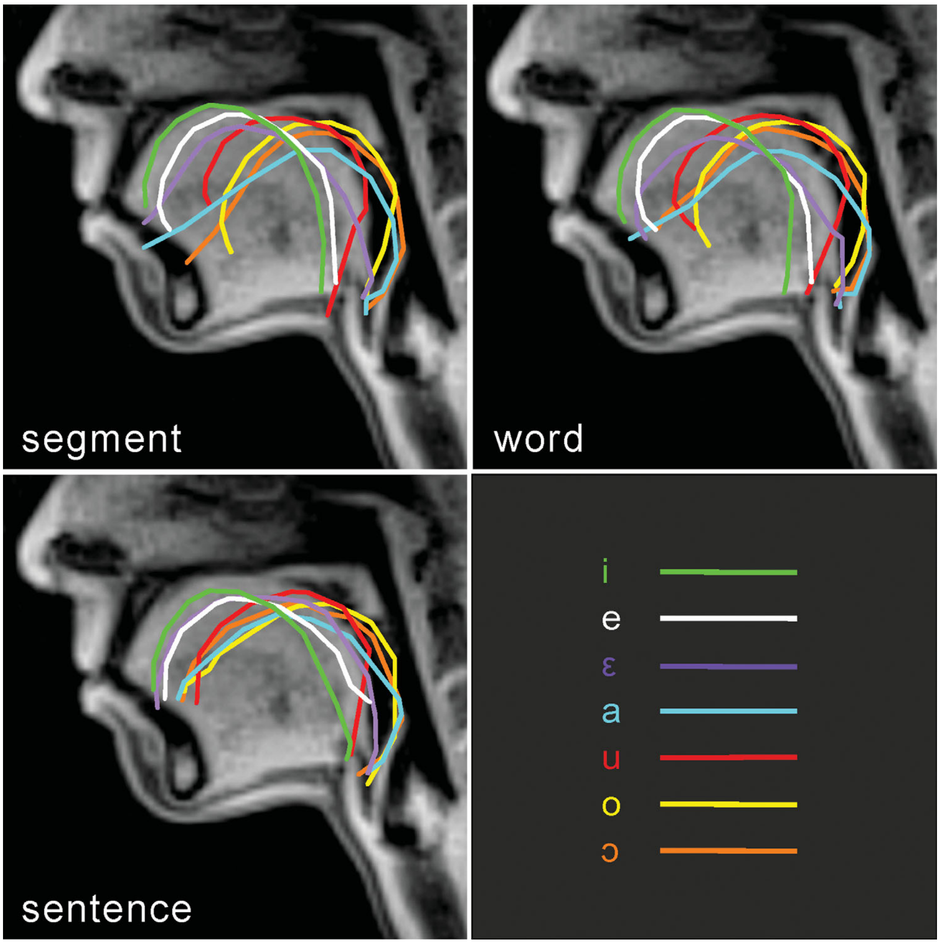

FIG. 3. Real-time MRI of vowel production. Color-coded contours represent the position of the tongue from tip to root for seven vowels and three linguistic conditions (segment, word, and sentence). Reference images were taken from the resting phase prior to speaking (same parameters as in Fig. 2).

Figure 7 focuses on the velum gesture for the plosive and nasal articulation of bilabial ([b] and [m]), alveolar ([d] and [n]), and velar consonants ([g] and [n]). For plosive consonants, the velum rises in contact with the posterior pharyngeal wall to close the nasal cavity, so that air from the lung passes through the mouth. For nasals, on the contrary, the velum is lowered to allow the air to pass through the nasal cavity.

Movements of the larynx are shown in Fig. 8 in a coronal plane (top part) comparing the voiceless labiodental consonant [f] with its voiced counterpart [v]. While the former requires opening of the vocal folds, so that air can flow freely through the larynx and generate the sound, the latter is produced by vibrating the vocal folds closely against each other, causing a tone. Two reference lines in the images at the level of the ventricular (upper

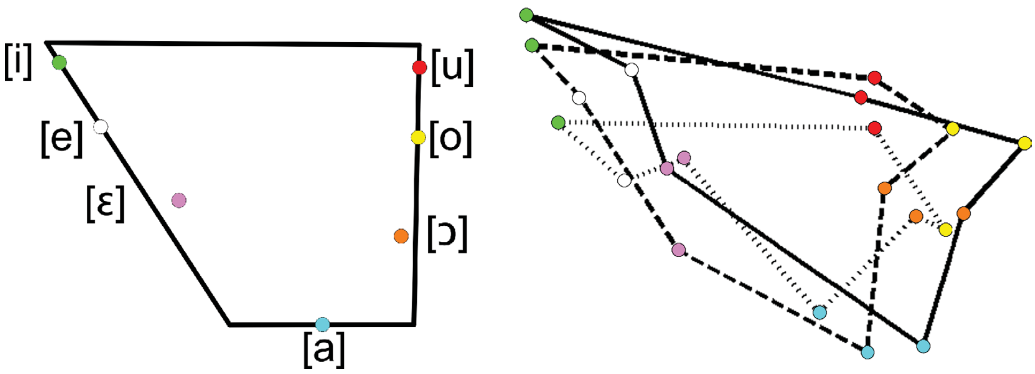

FIG. 4. Vowel diagrams for Standard Modern German (left) according to Ref. (21) and (right) as obtained by real-time MRI for segments (solid line), words (broken), and sentences (dotted). The diagrams represent the relative top positions of the tongue (compare Fig. 3, same color code).

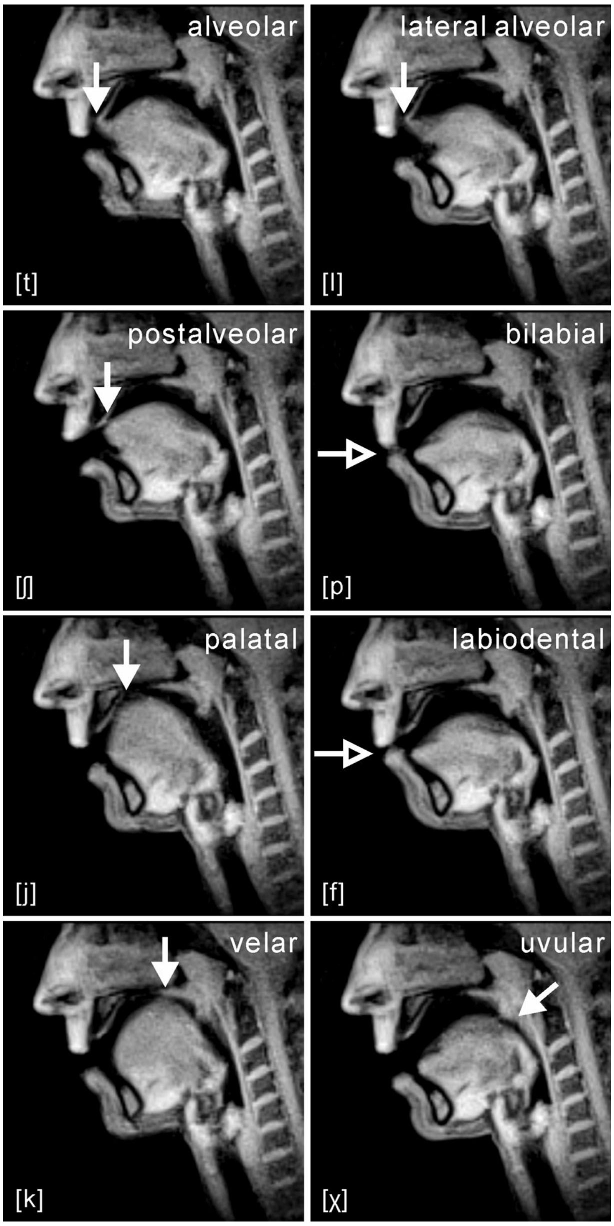

FIG. 5. Real-time MRI of consonant production. For eight consonants the selected images demonstrate the place of articulation and the configuration of the articulators (arrows for tongue and lips) at the time of sound production. Experimental parameters as in Fig. 2.

position) and vocal folds (lower position) served to derive corresponding spatiotemporal profiles as shown in the bottom part of Fig. 8. Whereas the ventricular folds open for a short period of about 130–200 ms (4–6 images) for the production of [f] and remain almost closed for [v], the vocal folds apparently stay in a similarly open position in both cases. However, this latter finding may be due to the limited MRI visibility of their inner structures, which is caused by its short T2* as a ligament tissue and the extremely fast movement of up to 230 Hz (24). Nevertheless, the overall motion of the larynx was well depicted for all subjects (see Supporting Information Movie 4).

Coarticulation

The prenasalization of vowels followed by a nasal consonant is a frequent phenomenon, e.g., see (25). Figure 9 details the gestures of a typical coarticulation effect, i.e., the velum movement in a nasal context, by the simultaneously recorded spatiotemporal profiles for LA, tongue body constrict degree, TTCD, and VEL (compare Fig. 1c). The data demonstrate that it is possible to identify prenasalization effects on the first vowel by comparing the results for *Pate* with those for *Sahne*. The former contains the same vowels as in *Sahne*, but lacks an intervocalic nasal consonant. In either case the articulation of the vowel [a:] in the spatiotemporal profiles of Fig. 9 extends from the solid to the broken vertical line where it ends right before the tongue tip reaches the teeth ridge (arrows for TTCD). In contrast, the velum behaves quite differently for both words. Whereas *Sahne* is characterized by a lowering of the velum already during production of the vowel (open arrow before the broken line), in *Pate* no such movement is observable. These imaging results confirm previous electromyographic findings (25) by adding a clear visualization of anticipatory coarticulation.

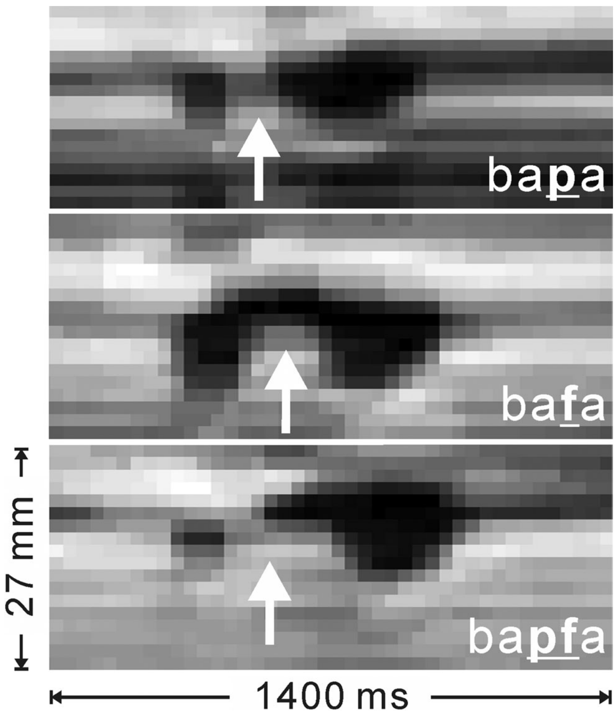

FIG. 6. Real-time spatiotemporal profiles (1400 ms duration and 33.3 ms resolution) of the LA (27 mm line LA in Fig. 1c) for (top) a bilabial consonant [p], (middle) a labiodental consonant [f], and (bottom) an affricate consonant [pf]. Arrows indicate the relative positions of the lips and teeth at the time of the underlined consonant.

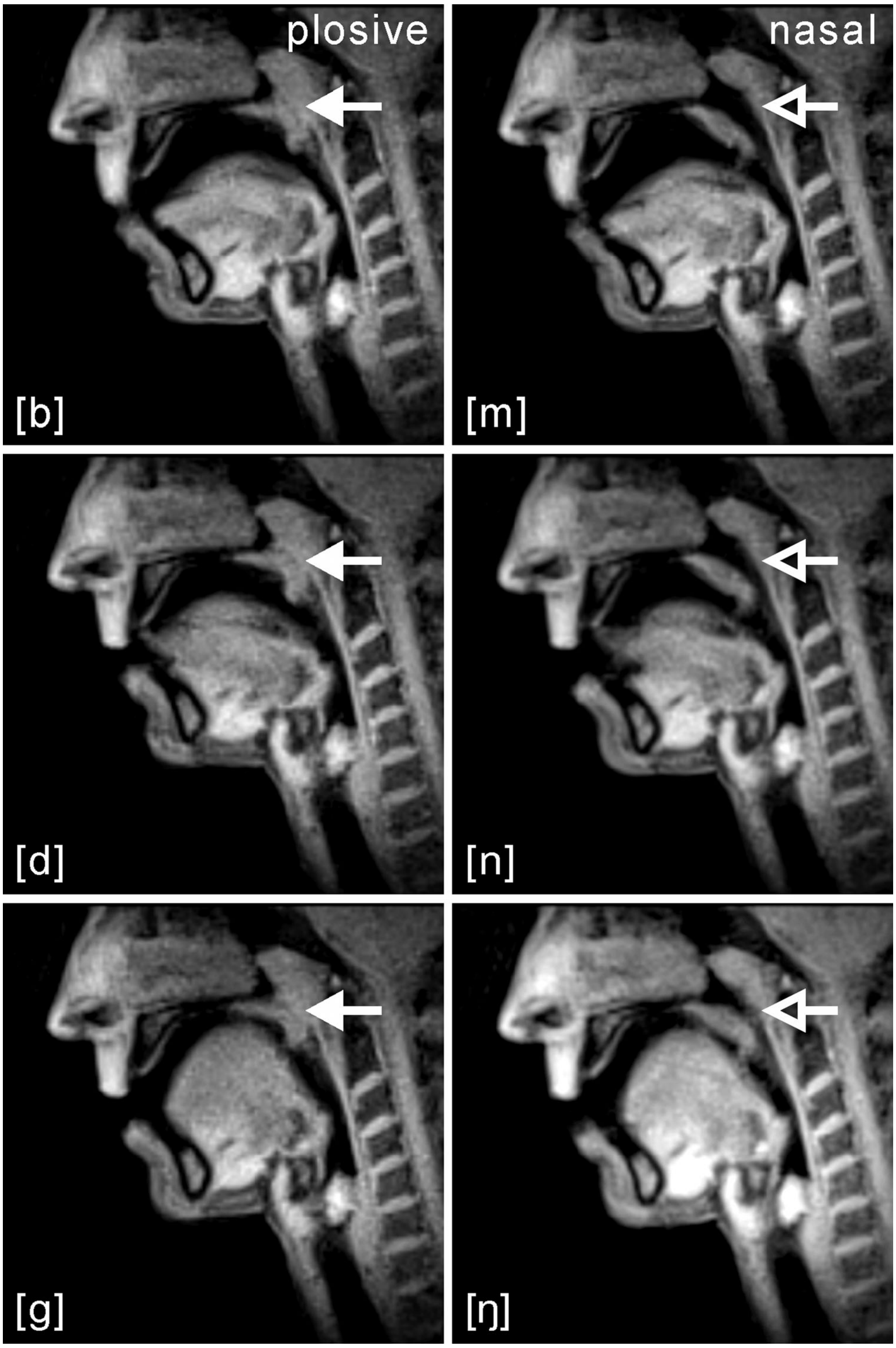

FIG. 7. Real-time MRI of (left) plosive and (right) nasal consonant production. While plosive consonants are characterized by a rise of the velum (arrows) with closure of the nasal cavity, nasal consonants involve a lowering of the velum (open arrows), and opening of the nasal cavity. Experimental parameters as in Fig. 2.

CONCLUDING REMARKS

This work demonstrates the use of a recently developed real-time MRI method based on highly undersampled radial FLASH to study articulatory dynamics during natural speech production. The method successfully addresses hitherto unmet challenges such as adequate speed, i.e., image acquisition times as short as for videofluoroscopy, and high image quality, i.e., good spatial resolution and SNR as well as minimized sensitivity to susceptibility artifacts. A quantitative description of the production of vowels, consonants, and coarticulation effects in Standard Modern German was achieved for the main articulators such as lips, tongue, velum, and larynx. The analysis employed spatiotemporal profiles along motion-adapted reference lines and revealed general agreement with previous phonetic findings. The choice of an articulation-specific reference system served to adapt changes of the vocal tract during speaking to the individual subject and motion. Nevertheless, because the real-time MRI method generates vast amounts of data, more efficient post-processing tools that allow for

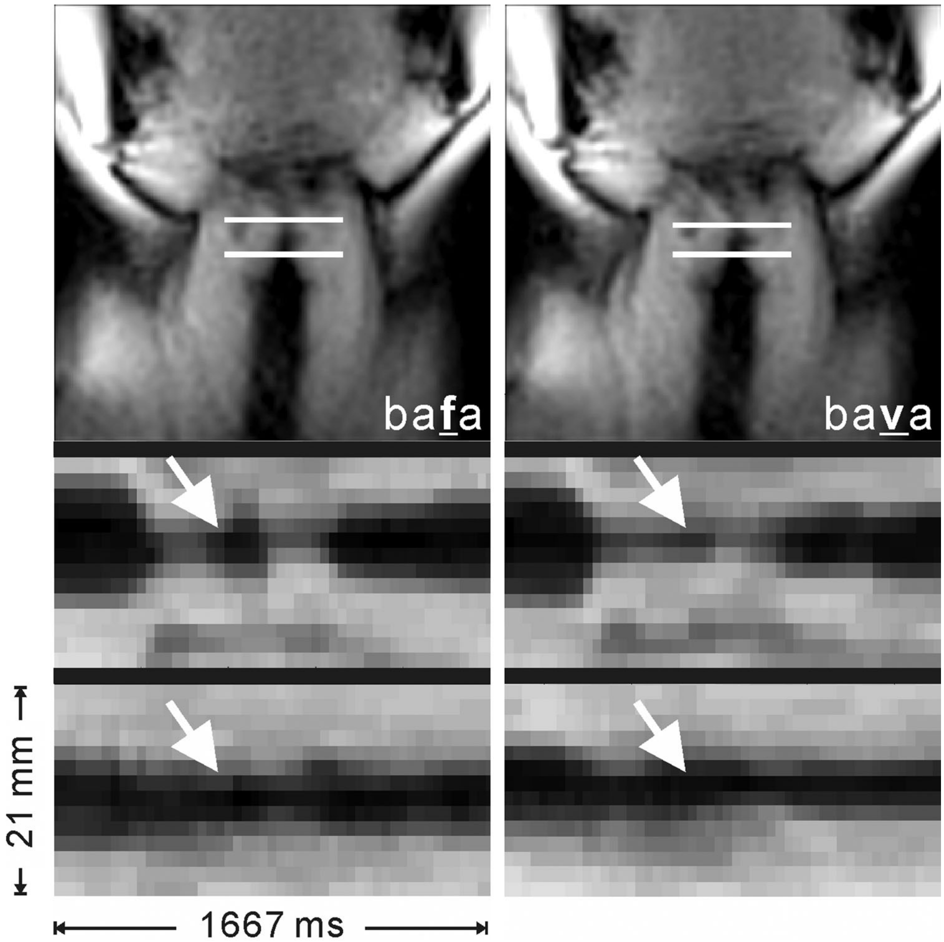

FIG. 8. (Top) Real-time MRI of the larynx for (left) the voiceless consonant [f] and (right) the voiced labiodental consonant [v] (coronal image). Two reference lines (21 mm) at the level of the ventricular (upper line) and vocal folds (lower line) served to derive spatiotemporal profiles (1667 ms duration and 33.3 ms resolution) that delineate the positions of the ventricular (upper profile) and vocal folds (lower profile) at the time of the consonant (arrows).

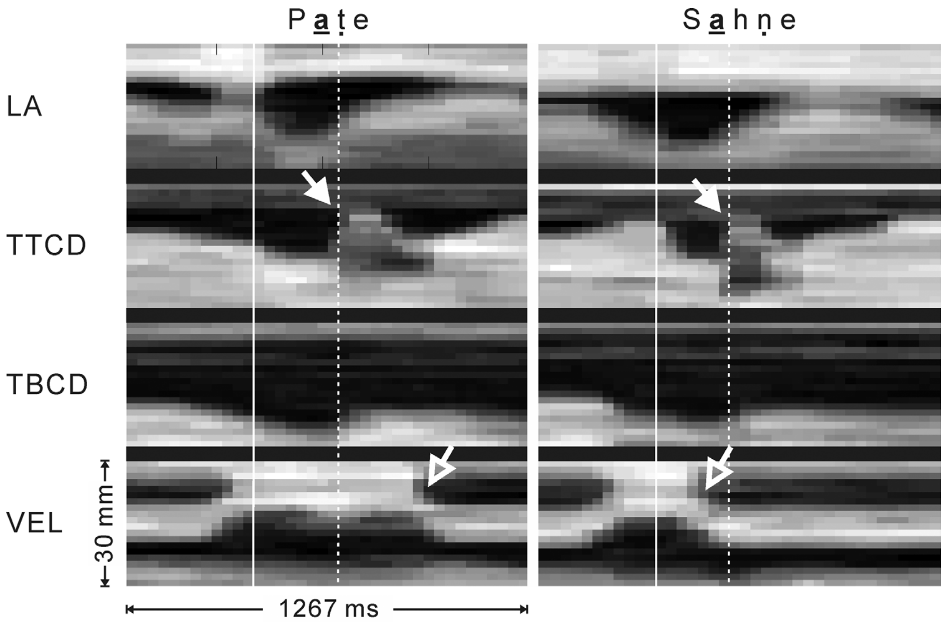

FIG. 9. Real-time MRI of coarticulation: prenasalization of the vowel [a:] followed by a nasal consonant. Spatiotemporal profiles (1267 ms duration and 33.3 ms resolution) for the LA, TTCD, tongue body constrict degree (TBCD), and VEL while speaking *Pate* in comparison to *Sahne* (compare Fig. 1c). Vertical lines refer to the beginning (solid) and end (broken) of the vowel. TTCD arrows refer to the contact of the tongue tip with the teeth ridge, VEL arrows indicate the lowering of the velum, which for *Sahne* already occurs during articulation of the vowel.

automatic edge detection, object tracking and segmentation (20) need to be combined with future developments.

A limitation of this real-time MRI study is the restriction to a single section. While a simultaneous multislice acquisition would be desirable for synchronizing articulatory configurations in different planes, a 3D MRI assessment promises even better access to vocal tract shaping and an improved characterization of articulatory processes (26). At this stage, however, real-time 3D MRI does not seem to be a realistic option. Extensions of cross-sectional real-time MRI movies to a few sections are certainly conceivable, but at the expense of compromising the temporal and/or spatial resolution. Pertinent examples have been reported for real-time MRI of heart function (14) and lingual articulation (27).

In summary, the proposed real-time MRI method offers dynamic imaging with acquisition times of 33 ms and 1.5 mm in-plane resolution. The achieved image quality, spatiotemporal resolution, and access to quantitative speaking parameters are expected to advance further applications in linguistics, singing, clinical phonetics, and logopedics.

ACKNOWLEDGMENTS

The authors thank Xuan Chang and Johannes Roemer for their contributions at an earlier stage of this work. The authors also thank Dr. Dirk Voit for technical support and Dr. Arno Olthoff for valuable discussions.

REFERENCES

1. Stone M. Laboratory techniques for investigating speech articulation. In: Hardcastle WJ, Laver J, Gibbon FE, editors. The Handbook of Phonetic Sciences. Oxford, UK: Blackwell Publishing Ltd., 2010. p 9–38.
2. Demolin D, Hassid S, Metens T, Soquet A. Real time MRI and articulatory coordinations in speech. C R Biol 2002;325:547–556.
3. Kröger BJ, Hoole P, Sader R, Geng C, Pompino-Marschall B, Neuschaefer-Rube C. MRT sequences as a database for a visual articulatory model. HNO 2004;52:837–843.
4. Mady K, Sader S, Zimmermann A, Hoole P, Beer A, Zeilhofer H, Hannig C.Assessment of consonant articulation in glossectomee speech by dynamic MRI. In: Proceedings of the 7th International Conference on Spoken Language Processing, Denver, Colorado, USA, 2002, p. 961.
5. Shriberg LD, Kent RD, Karlsson HB, McSweeny JL, Nadler CJ, Brown RL. A diagnostic marker for speech delay associated with otitis media with effusion: backing of obstruents. Clin Linguist Phon 2003; 17:529–547.
6. Narayanan S, Nayak K, Lee S, Sethy A, Byrd D. An approach to real-time magnetic resonance imaging for speech production. J Acoust Soc Am 2004;115:1771–1776.
7. Byrd D, Tobin S, Bresch E, Narayanan S. Timing effects of syllable structure and stress on nasals: a real-time MRI examination. J Phon 2009;37:97–110.
8. Beer AJ, Hellerhoff P, Zimmermann A, Mady K, Sader R, Rummeny EJ, Hannig C. Dynamic near-real-time magnetic resonance imaging for analyzing the velopharyngeal closure in comparison with videofluoroscopy. J Magn Reson Imaging 2004;20:791–797.
9. Kim YC, Narayanan SS, Nayak KS. Flexible retrospective selection of temporal resolution in real-time speech MRI using a golden-ratio spiral view order. Magn Reson Med 2011;65:1365–1371.
10. Sutton BP, Conway CA, Bae Y, Seethamraju R, Kuehn P. Faster dynamic imaging of speech with field inhomogeneity corrected spiral fast low angle shot (FLASH) at 3 T. J Magn Reson Imaging 2010;20: 791–797.
11. Uecker M, Hohage T, Block KT, Frahm J. Image reconstruction by regularized nonlinear inversion–joint estimation of coil sensitivities and image content. Magn Reson Med 2008;60:674–682.
12. Uecker M, Zhang S, Frahm J. Nonlinear inverse reconstruction for real-time MRI of the human heart using undersampled radial FLASH. Magn Reson Med 2010;63:1456–1462.
13. Uecker M, Zhang S, Voit D, Karaus A, Merboldt KD, Frahm J. Real-time MRI at a resolution of 20 ms. NMR Biomed 2010;23:986–994.
14. Zhang S, Uecker M, Voit D, Merboldt KD, Frahm J. Real-time cardiovascular magnetic resonance at high temporal resolution: radial FLASH with nonlinear inverse reconstruction. J Cardiovasc Magn Reson 2010;12:39–46.
15. Joseph AA, Merboldt KD, Voit D, Zhang S, Uecker M, Lotz J, Frahm J. Real-time phase-contrast MRI of cardiovascular blood flow using undersampled radial FLASH and nonlinear inverse reconstruction. NMR Biomed, in press.
16. Zhang S, Olthoff A, Frahm J. Real-time magnetic resonance imaging of normal swallowing. J Magn Reson Imaging, in press.
17. Zhang S, Block KT, Frahm J. Magnetic resonance imaging in real time: Advances using radial FLASH. J Magn Reson Imaging 2010;31: 101–109.
18. Hollien H. Vocal fold thickness and fundamental frequency of phonation. J Speech Hear Res 1962;5:237–243.
19. Browman CP andGoldstein L. Articulatory phonology: an overview. Haskins Laboratories Status Report on Speech Research 1992; SR–111/112:23–42.
20. Bresch E, Kim YC, Nayak K, Byrd D, Narayanan S. Seeing speech: capturing vocal tract shaping using real-time magnetic resonance imaging. IEEE Trans Signal Proc 2008;128–132.
21. Kohler KJ.Handbook of the International Phonetic Association.Cambridge:Cambridge University Press; 1999. p 87.
22. Iivonen A.Zur regionalen Variation der betonten Vokale im gehobenen Deutsch. Leena Kahlas-Tarkka (toim.), Neophilologica Fennica. Neuphilologischer Verein 100 Jahre. Mémoires de la Société Néophilologique de Helsinki XLV; 1987. p 87–119.
23. Kohler KJ. Segmental reduction in connected speech in German: Phonological facts and phonetic explanations. In: Hardcastle WJ and Marchal A, editors.Speech Production and Speech Modelling.Dordrecht:Kluwer; 1990. p 69–92.
24. Ladefoged P. A course in phonetics. Fort Worth, TX: Harcourt College Publishers; 2001.
25. Ushijima T, Hirose H. Electromyographic study of the velum during speech. J Phonetics 1974;2:315–326.
26. Kim YC, Narayanan SS, and Nayak KS. Accelerated three-dimensional upper airway MRI using compressed sensing. Magn Reson Med 2009;61:1434–1440.
27. Kim YC, Proctor MI, Narayanan SS, Nayak KS. Improved imaging of lingual articulation using real-time multislice MRI. J Magn Reson Imaging, in press.