

# Real-time MRI of Speaking: Preliminary Experience at a Temporal Resolution of 33ms

Aaron Niebergall<sup>1</sup>, Shuo Zhang<sup>1</sup>, Martin Uecker<sup>1,2</sup>, and Jens Frahm<sup>1</sup>

<sup>1</sup>Biomedizinische NMR Forschungs GmbH am Max-Planck-Institut fuer biophysikalische Chemie, Goettingen, Germany, <sup>2</sup>Electrical Engineering and Computer Science, University of California, Berkeley, California, United States

## Introduction

To understand human speech production is of great interest from phonetic, linguistic, engineering, and several other research aspects, as it reflects human-specific characteristics in anatomy and physiology. Compared to the other measurement approaches, MRI has the advantage of non-invasiveness and large spatial coverage of the entire vocal tract. However, several existing problems include: i) susceptibility artifacts due to tissue-air-boundaries in the oro-pharyngo-laryngeal area, ii) insufficient temporal resolution to resolve the rapid and coordinated movements, and iii) lack of synchronous audio recording for more detailed evaluation. In this work, a recently introduced real-time MRI technique [1, 2] was applied to solve these problems.

## Materials and Methods

The approach combines an undersampled radial fast low-angle shot (FLASH) MRI acquisition technique with image reconstruction by regularized nonlinear inversion [1]. Studies of healthy subjects ( $n = 12$ , 6 female, 6 male, Age 22 - 54) were performed on a conventional 3 T MRI system (Siemens Tim Trio) using a combination of a multi-element flexible (Siemens AG, Erlangen, Germany) and a bilateral  $2 \times 4$  array coil (Noras MRI Products, Hoechberg, Germany). T1-weighted data sets were obtained using a spoiled radial FLASH sequence (TR/TE = 2.22/1.44 ms, FA = 5°, FOV = 192 x 192 mm<sup>2</sup>, spatial resolution 1.5 x 1.5 x 10 mm<sup>3</sup>, bandwidth = 1560 Hz/Px). Speech tasks include i) vowels under different linguistic conditions, namely standalone letters or segments, words and sentences, ii) consonants contained in the meaningless logatom words with the same linguistic structure as carriers, and iii) words with different structure. The subjects were asked to speak in Standard Modern German in a natural rate with no intentional sustaining. The synchronous acoustic recording was accomplished by using an optical microphone (Optoacoustics, Or Yehuda, Israel) with online noise cancellation.

## Results

Preliminary applications demonstrate a high quality of real-time MRI movies that monitor the speaking process. Typical images have acquisition times of 33 ms corresponding to a temporal resolution of 30 frames per second. The images are free from susceptibility and motion artifacts. Various synergic movements of the articulators such as lip protrusion and opening, dorsum position, velum opening, larynx height, can be visualized during speech production. Particularly, articulatory configurations of vowels (Fig. 1) and consonants (Fig. 2) were well captured and characterized.

## Conclusion

The proposed real-time MRI method demonstrates unique potential for anatomical and functional studies of the speaking process at high spatiotemporal resolution, and thus promises a useful tool in studying the coordinated behaviors of the articulators.

## References

- [1] Uecker M, et al. NMR Biomed 2010;23:986-994.
- [2] Zhang S, et al. J Cardiovasc Magn Reson 2010;12:39-46.

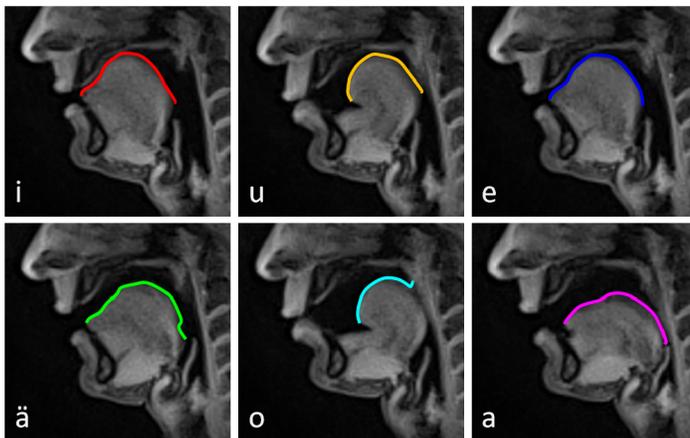


Figure 1: Real-time MRI of speaking at vowel production.

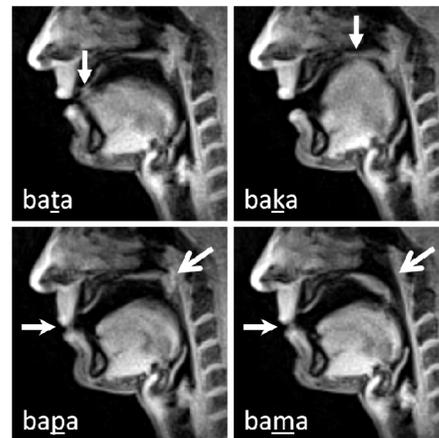


Figure 2: Real-time MRI of speaking at consonant production.