Accelerated High Field CINE MR Imaging of Tongue Motion During Speech Production

S. Takano[1], J. F. Utting[1], T. Frauenrath[2], M. Kob[2], and T. Niendorf[1]

[1]Division of Experimental Magnetic Resonance Imaging, RWTH Aachen, Aachen, Germany, [2]Division of Phoniatry, Pedaudiology and Communication Disorders, RWTH Aachen, Aachen, Germany

Introduction

Imaging studies hold great potential for elucidating the mechanisms of speech production. Movement of the tongue is an essential component of human phonation, which presents the major challenge of imaging its highly dynamic motion. Motion compensation of the speech organs has been elusive hitherto. To approach this aim, this study focuses on highly accelerated CINE imaging, synchronized with tongue motion to accomplish frame rates as high as 240 fps. In order to balance the competing constraints of signal-to-noise ratio (SNR) and imaging speed the study makes use of the SNR advantage inherent to high magnetic field strengths in synergy with many-element coil arrays. In order to achieve high accelerations, supporting targeted temporal and spatial resolutions without the prohibitive noise amplification associated with coil sensitivity encoding, spatio-temporal correlations in dynamic imaging are exploited using *k-t* BLAST (1,2) with the ultimate goal of visualizing the deformation of the tongue during speech.

Methods

Healthy volunteers repeated the vowel sequence '/ei/' in a 3.0 T MR scanner (Achieva, Philips Medical Systems) prompted by 1s (3) cyclic rhythm tones synchronized with MR imaging (Figure 1). Synchronization was accomplished via sound presentation through head-phones to the subject together with external-triggering (3). The trigger waveform was delivered to the internal physiological signal controller circuitry of the clinical scanner without modifying the scanner's hardware. The imaging protocol consisted of two series: (i) segmented, 2D CINE gradient-echo imaging and (ii) segmented, 2D CINE gradient-echo imaging in conjunction with tagging (lattice pattern, 8 mm intervals) to make use of T_1-prolongation, tag mean life time and contrast-to-noise enhancement at 3.0T for motion tracking (4). Imaging parameters were: TR=4 ms, TE=1.2 ms, flip angle=10 degrees, slice thickness=10 mm, FOV=(256 x 256) mm^2, matrix was 256 x 256 (reconstructed as 448 x 448), number of phases=120, trigger delay=500 ms and net frame rate of 240 fps. Data acquisition was accelerated using undersampling rates of kt=1, kt=8, kt=12 and kt=16 together with k-t BLAST reconstruction.

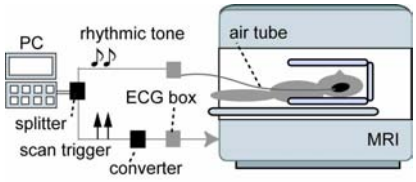

Fig. 1: Synchronization of the data acquisition with the speech paradigm by external triggering.

Results

The feasibility of highly accelerated 2D CINE imaging of tongue motion at 3.0 T using spatio-temporal correlations in the dynamic data is demonstrated in Fig. 2. The *k-t* BLAST approach enabled acceleration factors and image quality that are unattainable with one-dimensional SENSE. The imaging speed advantage of 8-fold, 12-fold and 16-fold undersampled acquisitions over the conventional approach translated in to a substantial scan time reduction from 176 s to 44 s (k-t=8), 32 s (k-t=12) and 28 s (k-t=16), while preserving image quality and millimeter in-plane spatial resolution. The effective temporal resolution of 4 ms facilitated visualisation of rapidly moving structures, a capability that was successfully exploited for the evaluation of tongue mechanics and motion patterns across the entire "/ei/"-cycle as illustrated for selected phases in Fig. 2.

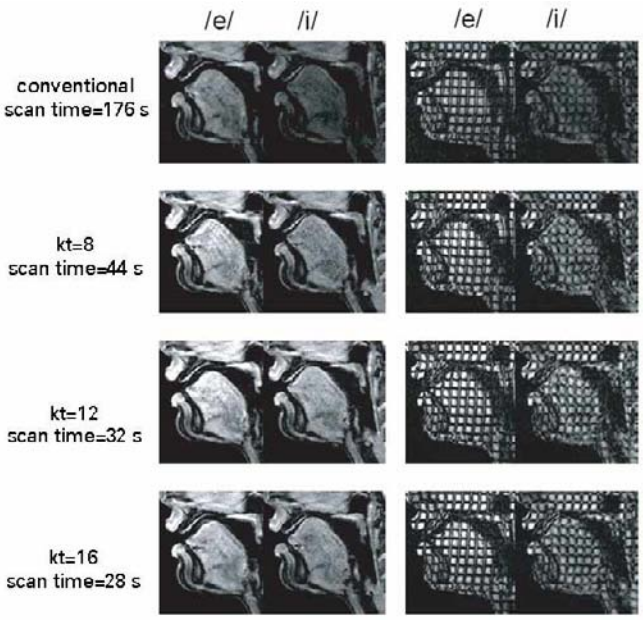

Fig. 2 Tongue motion phases /e/ and /i/ during an /ei/ cycle. Images were acquired using 2D CINE gradient-echo techniques without (left) and with (right) a tagging preparation module. For acceleration purposes 8-, 12- and 16-fold undersampling was applied. Images were reconstructed using the k-t approach.

Discussion and Conclusions

Highly accelerated MR imaging in conjunction with synchronization techniques serves to remedy many of the existing limitations associated with highly dynamic tongue motion. The speed advantages presented here offer the potential of extending the capabilities of speech production studies using 2D CINE techniques to whole speech organ coverage 3D acquisitions within acceptable scan times. In conclusion, speech organ imaging stands to benefit substantially from the improvements in effective temporal resolution and imaging speed enabled by highly accelerated parallel imaging. Consequently, we anticipate an extension of this work to evolve towards resolving vocal fold vibration with stroboscopic MRI, which has been unattainable hitherto so that overcoming the temporal resolution obstacle would be legion for imaging and understanding of highly dynamic 3D motion of speech organs.

References 1) Tsao J. et. al., Magn Reson Med 2003; 50:1031-1042. **2)** Hansen M.S. et. al. Magn Reson Med 2004; 52:1175-1183. **3)** Masaki et. al, 1999 J. Acoust. Soc. Jpn. . 375-379. **4)** Gutberlet, M. et. al. Invest. Radiology 2006. 41 154-167.