University of Utah
UNDERGRADUATE RESEARCH JOURNAL

USING REAL TIME MAGNETIC RESONANCE TO STUDY VOCAL TREMOR

Bryce Larson (Julie Barkmeier-Kraemer PhD, CCC-SLP, Ganesh Adluru, PhD)

Division of Otolaryngology

Background

Vocal tremor is a neurogenic voice disorder characterized by nearly periodic changes of pitch and loudness. To date, vocal tremor has been primarily characterized using acoustic measures of fundamental frequency (f_o) and amplitude/sound pressure level (SPL). A listener can hear these acoustic measures as a the fluctuation of pitch and loudness, ultimately leading to a shaky voice.[1-4] However, literature describing limb tremor characterizes etiologic differences by measures of rate and extent of tremor movements.[5] To date, little is known about the kinematics of speech structure tremor associated with the acoustic patterns of vocal tremor. This research gap is largely due to limitations of current clinical tools for imaging speech structures. Real-time magnetic resonance imaging (rtMRI) offers a method for recording and measuring vocal tract structures during speech.[6,7] However, rtMRI spatial and temporal resolution levels for measuring speech structure oscillation rate and extent during voicing is unknown. Spatial resolution refers to the number of pixels utilized in the image while temporal resolution is how often data is collected in the same area. Due to Heisenbergs uncertainty principle, we trade spatial resolution for temporal, making it difficult to maximize both simultaneously. Since there is a trade off between the two we see smaller pixel size leads to a greater spatial resolution. Another factor that effects Spatial resolution is our voxel size. The voxel is both the pixel size and slice thickness measured in mm^3. In-plane resolution is the base unit of the voxel, with its pixel sizes being measured in mm. An increased voxel size lowers the overall resolution of the image. Therefore, we wanted to produce the smallest voxel size while maintaining a proper spatial and temporal resolution. The purpose of this study was to determine the feasibility and specifications for using rtMRI to measure rate and extent of speech structure oscillations associated with vocal tremor.

Methods

Institutional Review Board: This pilot study was completed as part of an approved University of Utah IRB protocol #0094980.

Participants: One normal speaker (1 female) and 4 speakers diagnosed with vocal tremor (1 male, 3 female) were recruited and consented to participate.

Equipment: A Siemens MAGENTOM Prisma 3 Tesl machine was used to acquire mid-sagittal and axial imaging of participants. Imaging methods varied by in-plane resolution and slice thickness across subjects until measures of tremor oscillation were achieved. Nasoendoscopic recordings using PentaxMedical HD Video Laryngostroboscope (EPK-1000) equipment were reviewed to obtain kinematic estimates of structural rates of oscillation in affected structures under similar phonation conditions to verify dynamic real-time MRI findings.

Speech Tasks: All participants completed 1 minute of quiet breathing to obtain referent recordings of non-moving upper airway structures during the mid-sagittal and axial imaging planes of view. Thereafter, all participants completed three trials each of sustained phonation of /a/ and then /i/ for as long as possible at comfortable pitch at comfortable, soft, and loud phonation levels, and then during high pitch and low pitch voicing. All trials of voicing were recorded during imaging in the mid-sagittal plane whereas only trials of sustained phonation of /i/ were recorded during recordings in the axial plane of view.

MR data acquisition: Continuous real time MR images were acquired during sustained phonation using a radial GRE pulse sequence with golden ratio. This acquisition scheme enabled retrospective reconstruction of images at a temporal resolution showing all of the oscillatory cycles. Spatial resolution parameters were adjusted across participants until speech structures observed to oscillate during endoscopy showed similar oscillatory cycles on Reconstructed MR images (FIG 1).

Image reconstruction: Reconstruction of under-sampled data was done offline using a multi-coil spatiotemporal constrained reconstruction method with total of variation as well as a patched based low rank constraints.

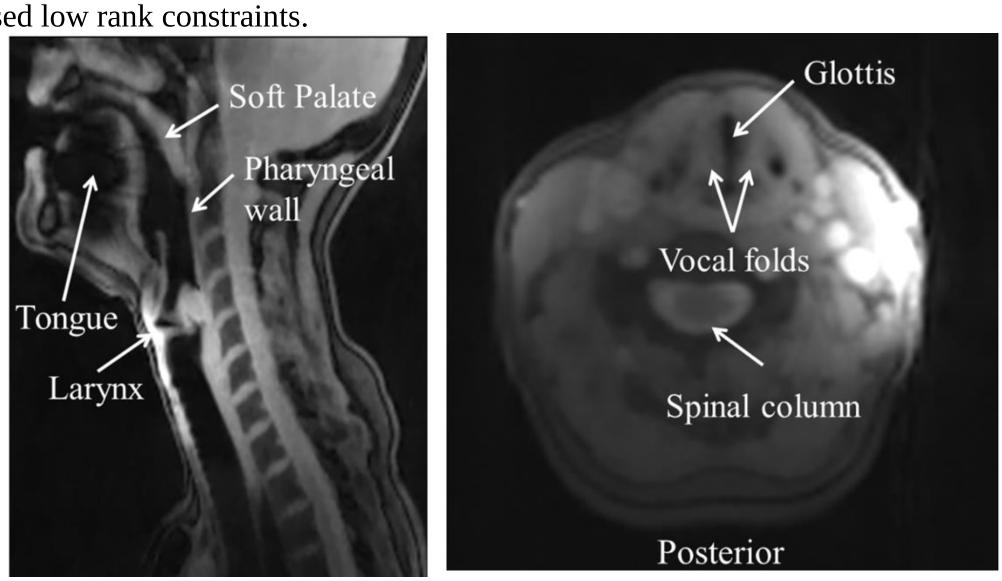

FIG 1. *Using an anchor point of the vertebrae, we measured the rate and extent of structure oscillations to measure the maximum range (i.e. extent) and rate of the tremor cycle for the soft palate, pharyngeal wall, tongue, and larynx (mid-sagittal plane) and the lengthwise and lateral oscillations of the vocal fold slow-rate oscillations (that is, ~4-8 Hz) in the axial plane. The figure to the left shows the mid-sagittal plane MR image. The figure on the right shows the axial plane MR image.*

Results/Conclusion

Findings showed that rtMRI was a feasible imaging modality for acquiring and measuring vocal tract and laryngeal oscillations during voicing in those with vocal tremor. Optimal temporal resolution was achieved at 30 fps. Image resolution for target measures was achieved with mid-sagittal in-plane resolution of 1.4 x 1.4 mm^2 with a slice thickness of 8 mm and axial in-plane resolution of 1.3 x 1.3 mm^2 with a slice thickness of 3 mm (FIG 2). Horizontal vocal fold and laryngeal vertical oscillation during vibrato in a normal subject was successfully measured and found to be comparable to acoustic measures of pitch and loudness.

Participants	Sagittal resolution	Axial resolution	Comments
1	In-plane: 1.6x1.6 mm^2 Slice thickness: 8 mm	In-plane: 1.6x1.6 mm^2 Slice thickness: 5 mm	No tremor Oscillation observed. Movements were too subtle to detect in the upper airway.
2	In-Plane: 1.6x1.6 mm^2 Slice thickness: 8 mm	In-plane: 1.6x1.6 mm^2 Slice thickness: 5 mm	Pitch and loudness tasks used to identify condition of largest tremor movements. Ideal temporal resolution is 30 FPS.
3-5	In-plane: 1.4x1.4 mm^2 Slice thickness: 8 mm	In plane: 1.3x1.3 mm^2 Slice thickness: 3 mm	Pitch and loudness tasks used to identify condition of largest tremor movements. Idea temporal resolution is 30 FPS.

FIG 2. *The optimal temporal resolution needed to capture tremor oscillations is > 30 fps. In-plane resolution is adequate at 1.4 x 1.4 mm^2 in the sagittal plane and 1.3 x 1.3 mm^2 in the axial plane when speech structure oscillation is clearly visible during nasoendoscopic examination.*

References

1.Lederle A, Barkmeier-Kraemer J, Finnegan E. Perception of vocal tremor during sustained phonation compared with sentence context. J Voice. 2012;26(5):668 e661-669.

2.Brown JR, Simonson J. Organic voice tremor. A tremor of phonation. Neurology. 1963;13:520-525.

3.Dromey C, Warrick P, Irish J. The influence of pitch and loudness changes on the acoustics of vocal tremor. J Speech Lang Hear Res. 2002;45(5):879-890.

4.Gamboa J, Jimenez-Jimenez FJ, Nieto A, et al. Acoustic voice analysis in patients with essential tremor. J Voice. 1998;12(4):444-452.

5.Louis ED. Diagnosis and Management of Tremor. Continuum (Minneap Minn). 2016;22(4 Movement Disorders):1143-1158.

6.Perry JL, Sutton BP, Kuehn DP, Gamage JK. Using MRI for assessing velopharyngeal structures and function. Cleft Palate Craniofac J. 2014;51(4):476-485.

7.Lingala SG, Zhu Y, Kim YC, Toutios A, Narayanan S, Nayak KS. A fast and flexible MRI system for the study of dynamic vocal tract shaping. Magn Reson Med. 2017;77(1):112-125.