Functional and structural connectivity of auditory areas that process talker variability in speech

Ja Young Choi[1,3], Rebecca Belisle[2], & Tyler Perrachione[1]

[1]Dept. of Speech, Language, & Hearing Sciences, and [2]Graduate Program for Neuroscience, Boston University;
[3]Program in Speech & Hearing Bioscience and Technology, Harvard University

Summary

Speech acoustics vary substantially across talkers, affecting speech recognition efficiency. Functional neuroimaging[1], brain stimulation[2], and intracranial electrophysiology[3] implicate bilateral STG in processing talker variability. **Here, we measured functional and structural connectivity of talker variability-sensitive brain areas.**

- Sparse-sampling fMRI measured BOLD activation during a word recognition task with speech from single or mixed talkers.
- **Talker adaptation** (less activation for single vs. mixed talkers) was found in four areas: left and right, middle and posterior STG.
- Group-constrained subject-specific functional regions of interest (fROIs) in each area served as seeds for connectivity analyses.
- Generalized psychophysiological interactions (gPPI) revealed greater functional connectivity to domain-general brain areas when listening to a single talker vs. multiple talkers.
- Probabilistic tractography highlighted prominent structural connections to parietal and occipital cortices.
- Broadest functional and structural connectivity from right pSTG.

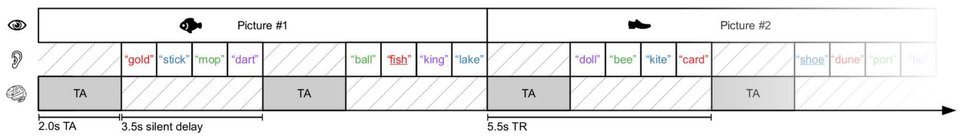
Overall, auditory areas that are sensitive to talker variability are functionally and structurally integrated in broader, whole-brain salience and attentional networks. These neural results support behavioral models of talker adaptation[4] that implicate domain-general auditory attention in explaining differences in speecdh processing efficiency when listening to single vs. mixed talkers.

Procedures

Participants: N=19 native English-speaking young adults (11F / 8M) with normal speech, language, hearing, and reading abilities.

Stimuli: 288 high-familiarity monosyllabic nouns read in isolation by 5 female young adult native English speakers. 36 words were "targets" which matched black and white line drawings of objects.

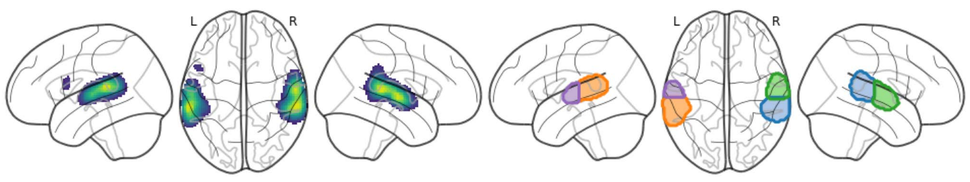
Task: Four word recordings were played during the silent delay between sparse acquisitions in blocks of 22s (4 TRs). In *single-talker* blocks, all the recordings came from one consistent talker. In the *mixed-talker* blocks, the 4 words in each TR were spoken by 4 different talkers. Participants pressed a button to indicate when the word they heard matched the picture they saw (2 per block). Responses to targets were 10% faster in *single* blocks ($p < 0.0001$).



Task design. Auditory words were presented during silent TR delays. Participants indicated when words matched the pictures they saw. A *mixed-talker* block is shown, with colors representing different talkers.

MRI data acquisition: Anatomy: Whole-brain, T1-weighted, $1mm^3$ voxels. Functional: Sparse-sampled T2*-weighted BOLD EPI (TR = 5.5s, TA = 2s, $3.125mm^2$ voxels, 32 4.0mm-slices) with 97 volumes per each of 4 runs. Diffusion: T2-weighted, TR=9.3s, $2mm^3$ voxels, 74 slices, 10 b_0 images, b = 0 & 700 s/mm^2, 30 diffusion directions.

MRI data analysis: Anatomy: Default surface reconstruction in FreeSurfer 6.0.0. Functional: Preprocessing (motion correction, spatial smoothing, coregistration) and mixed-effects model estimation were done in Lyman v1.0.0 using FSL, FreeSurfer, and custom algorithms; GCSS parcellation[5] was used to obtain subject-specific fROIs as seeds for connectivity analyses; We contrasted the *single* and *mixed* talker gPPI effects. Diffusion: Preprocessing via FSL (*eddy*, *topup*, *bedpostx*); connectivity from fROI seeds to cortical targets was estimated using *probtrackx*.

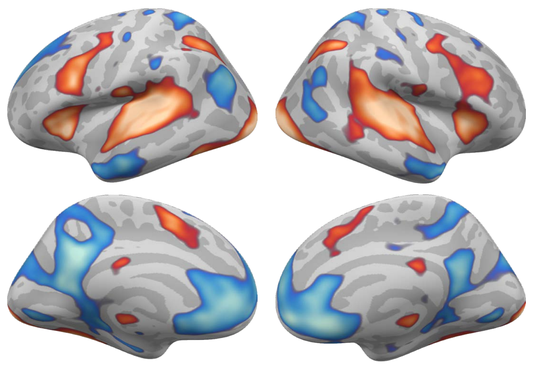


Left: **Probability** of significant within-subject talker-variability effects (*mixed* > *single*). ***Right:*** **GCSS parcellation** of variability-sensitive areas.

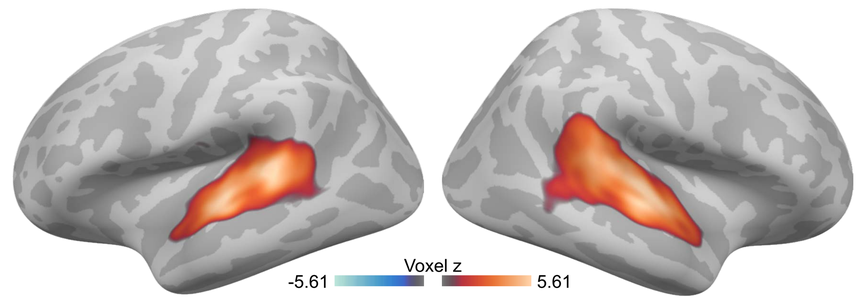
References & Acknowledgments

References: [1] Wong et al. (2004) *J. Cog. Neuro.* [2] Choi & Perrachione (2019) *Brain & Lang.* [3] Sjerps et al. (2019) *Nat. Comm.* [4] Choi & Perrachione (2019) *Cognition;* [5] Fedorenko et al. (2010) *J. Neurophys.*
Supported by NIH grants : T32 DC01301, R03 HD096098, R03 DC014045.
Contact: tkp@bu.edu http://sites.bu.edu/cnrlab/

Task Activation & Talker Adaptation

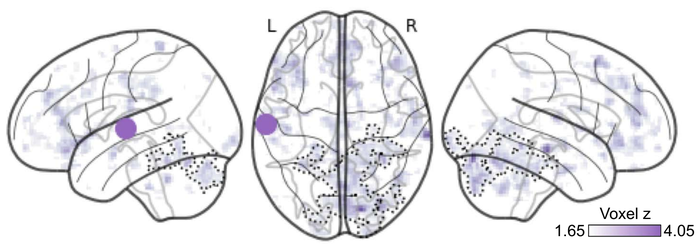


Task activation shown here for the contrast *mixed-talkers* > rest at the group level (voxel $p < 0.001$; cluster FDR $p < 0.05$).

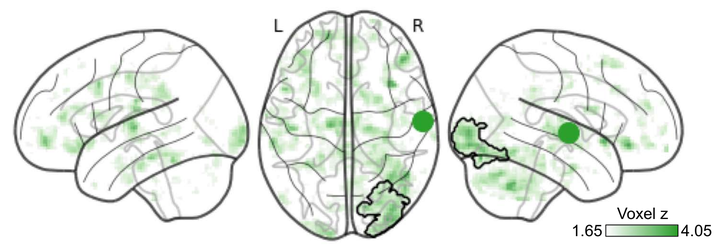


Brain areas sensitive to talker variability (based on the *mixed* > *single* talkers contrast) comprised bilateral middle and posterior STG (voxel $p < 0.001$; cluster FDR $p < 0.05$). No areas responded to reverse contrast (single > mixed talkers) at this threshold, but lower thresholds identified task-deactivated regions: medial prefrontal & anterior cingulate.

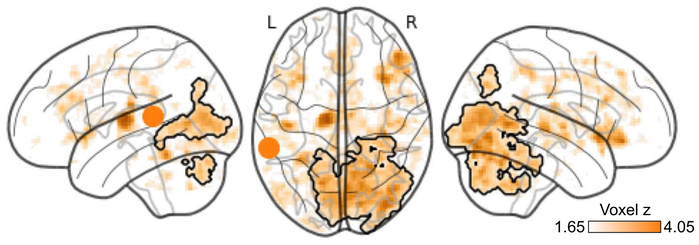
Task-Dependent Functional Connectivity



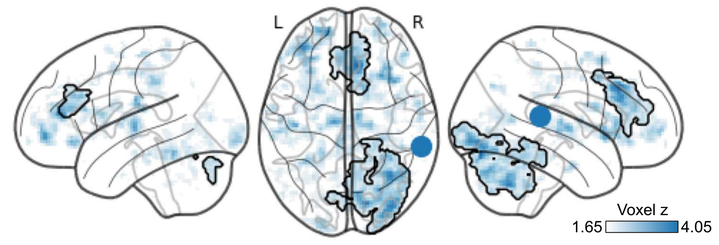
Left mSTG connectivity differences (single > mixed): visual cortex and cerebellum (vox. $p < 0.05$; clust. FDR $p < 0.05$).



Right mSTG connectivity differences (single > mixed): visual cortex ($p < 0.025$; clust. FDR $p < 0.05$).

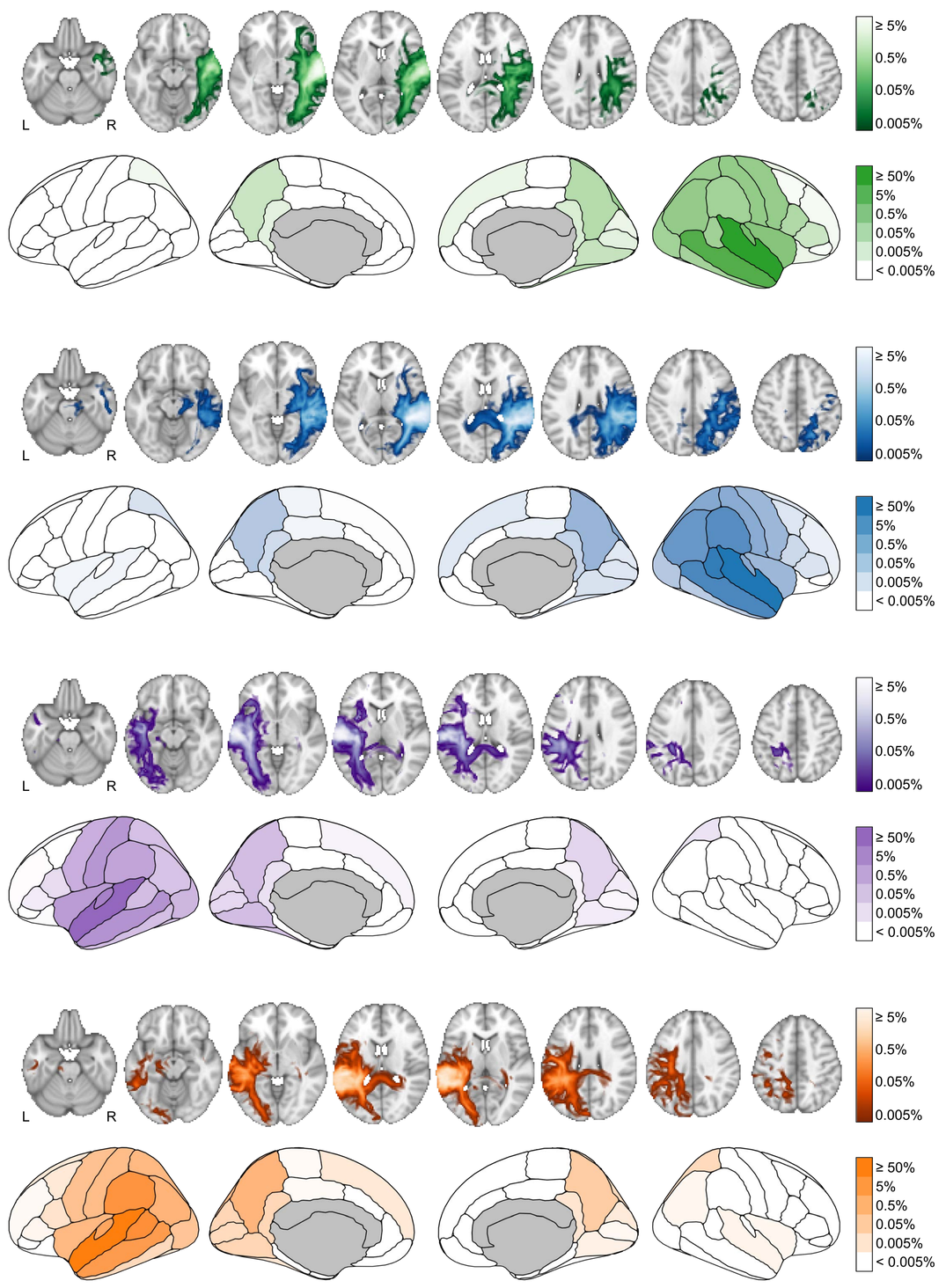


Left pSTG connectivity differences (single > mixed): visual cortex and cerebellum (vox. $p < 0.025$; clust. FDR $p < 0.05$).



Right pSTG connectivity differences (single > mixed): also to anterior cingulate (vox. $p < 0.025$; clust. FDR $p < 0.05$).

Structural Connectivity



Probability of t**racks** seeded from individual **right mSTG** fROIs; projections run primarily dorsally into parietal lobe and ventrally into temporal lobe.

Proportion of **terminations** for tracks from **right mSTG** fROIs, prominently arriving locally in STG; ipsilaterally in MTG, IFG, IPL, visual cortex; bilaterally in cuneus and SPL.

Probability of t**racks** seeded from individual **right pSTG** fROIs; projections run primarily dorsally into parietal lobe; note also contralateral projections.

Proportion of **terminations** for tracks from **right pSTG** fROIs, terminating locally in STG; ipsilaterally in MTG, SMG, IPL, cuneus; contralaterally in SPL, cuneus, STG.

Probability of t**racks** seeded from individual **left mSTG** fROIs; projections run dorsally to parietal lobe and ventrally to occipital and temporal lobes.

Proportion of **terminations** for tracks from **left mSTG** fROIs. This region had the most local terminations, plus ipsilateral parietal and MTG, contralateral SPL and cuneus.

Probability of t**racks** seeded from individual **left pSTG** fROIs; projections run primarily dorsally into parietal lobe; note also contralateral projections.

Proportion of **terminations** for tracks from **left pSTG** fROIs: locally in STG; ipsilateral MTG, SMG, cuneus; contralateral SPL, cuneus, STG.