Using a multi-task brain imaging battery to relate speech production to phonological working memory, emotion and prosody

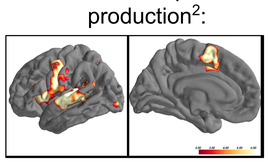
Kevin Sitek[1,2], Gregory Ciccarelli[2,3], Carlo de los Angeles[2], Thomas Quatieri[2,3], Satrajit Ghosh[1,2]

[1]Harvard Medical School, [2]MIT, [3]MIT Lincoln Laboratory

Background

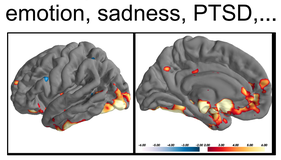
We use state-of-the-art fMRI acquisition techniques[1] to investigate multiple elements of the speech production network in a single session. We use three different tasks to provide a multifaced view of speech production that considers the influence of phonological working memory and emotion, and the neural basis of prosody.

Canonical speech production[2]:



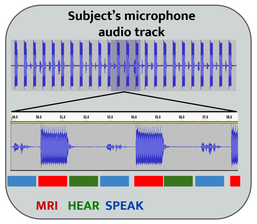
How are these linked?

Non-speech conditions: emotion, sadness, PTSD,...

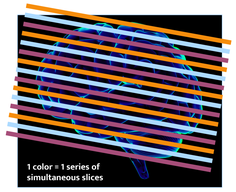


Methods

We acquired speech production in fMRI with sparse sampling, which includes silent periods for auditory stimuli and speech production. To maximize the number of trials, our silent periods were either 1 or 3 seconds (vs. 4-10 s typically).



We used simultaneous multislice imaging (SMS) to acquire 5 slices at a time. This provides **increased spatial resolution** (2mm voxels) while **reducing acquisition time** (1.1 seconds, typically 2 seconds).

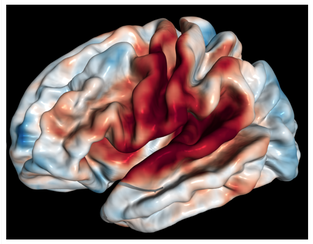


Twenty adults (mean age = 28.7 yrs, 2 left-handed, 7 female) performed a battery of speech- and emotion-related tasks.

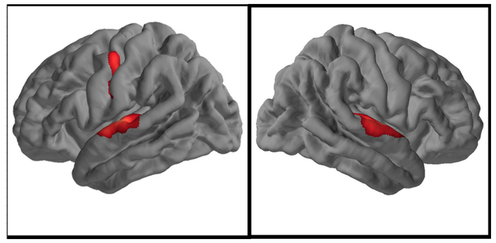
	Speech manipulation	Implicit emotion processing	Phonological load
Vocalization	Rate, pitch manipulation	Overtly read emotional sentences[3] Emotional face matching (modified Herari task)	Repetition of aurally presented non-words of 2-5 syllables
Example	"pa-ta-ka" (slow or fast) "aah" (high or low)	"I'm so sorry for hurting you" "Don't forget your jacket" "This chicken is excellent"	"bicket" "ameronable"
Target brain regions	ventral primary motor cortex, basal ganglia, cerebellum	Limbic system (amygdala, anterior & posterior cingulate)	posterior superior temporal gyrus, inferior frontal gyrus

Results

Non-word repetition

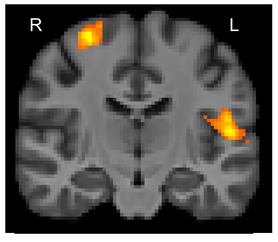


Clear speech network activation (all speech vs. baseline)

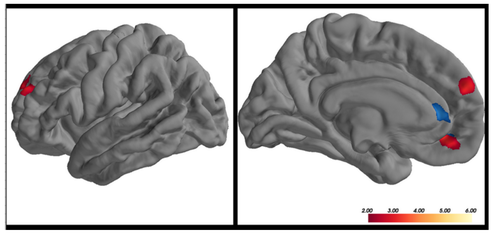


Increased phonological load by increasing the number of syllables activates left sensorimotor and bilateral auditory cortex

Emotional sentences

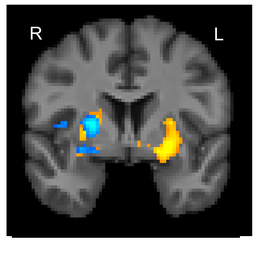


Sad (vs. happy) sentences activate L somatosensory and A1, R motor

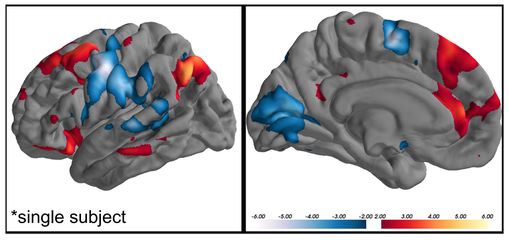


Emotional sentences involve anterior cingulate and lateral and medial prefrontal cortex (red = sad > neutral, blue = happy > neutral)

Speech manipulation



Pitch modulation activates bilateral putamen, L caudate, and R insula (blue = low > high, red = low > normal)



Increased **rate** (blue) activates motor cortex; decreased rate (red) increases cingulate/PFC

Discussion

Technical advances

Sparse SMS reduces the time per trial - increasing the number of tasks possible - and improves spatial resolution and statistical power.

Many tasks enable multifaceted view of speech production

Increasing phonological demands increases auditory and motor cortical activity due to additional auditory feedback and greater speech sequencing requirements.

Emotional sentence production involve limbic cortical regions, even with no task instruction to imitate the emotion of the sentence. Additionally, sad sentences in particular modulate auditory and sensorimotor cortex.

Explicit speech manipulation suppresses the default mode network and strongly recruits motor cortex. Basal ganglia are involved in pitch control, perhaps asymmetrically. Clear speech is "intentional" as it recruits frontal executive regions (lateral/medial PFC, cingulate) which compliment the canonical speech network.

Conclusion

Through a diverse array of speech and emotional processing tasks, we see neural activity outside of the canonical speech network. This may provide insight into understanding the neural basis of vocal modulation in neurological disorders such as depression.

References

1. Perrachione TK & Ghosh SS (2013). "Optimized design and analysis of sparse-sampling FMRI experiments." *Front Neurosci* 7:04.
2. Neurosynth.org
3. Russ JB, Gur RC, Bilker WB (2008). "Validation of affective and neutral sentence content for prosodic testing." *Behav Res Methods;*40(4):935-9.