Coordination of pitch versus phonetic features in speech motor control rely on distinct sensorimotor circuits

Gregory Hickok(greg.hickok@uci.edu)[1], **Minkyu Kim**[1], **Chris Naber**[2], **Bradley Buchsbaum**[3], **Caroline Niziolek**[2], **Kouroshi Saberi**[1], **Jonathan Venezia**[4]

[1]University of California, Irvine, CA, USA; [2]University of Wisconsin–Madison, Madison, WI, USA; [3]Rotman Research Institute, Toronto, ON, Canada; [4]VA Loma Linda Healthcare System, Loma Linda, CA, USA

UCI University of California, Irvine

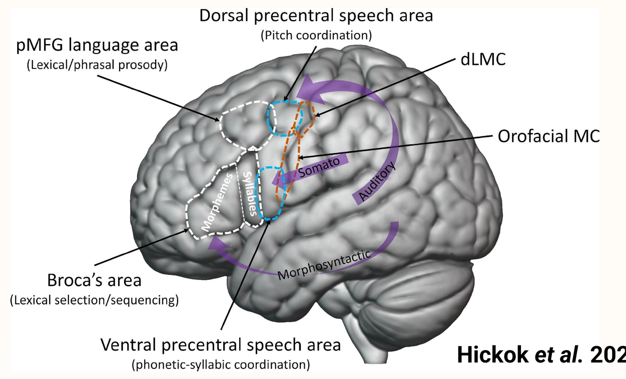
Abstract

Human speech *requires tightly coordinated control of* ***prosodic*** *and* ***segmental features****, but whether these functions rely on* ***shared or distinct neural pathways*** *remains unclear. Using* ***Altered Auditory Feedback*** *(AAF) of* ***fundamental frequency*** *(F0) and* ***formants*** *(F1, F2) during speech production, we identified both* ***common network components*** *and* ***dissociable pathways****. These findings support partially overlapping and functionally specialized auditory–motor control systems.*

Introduction

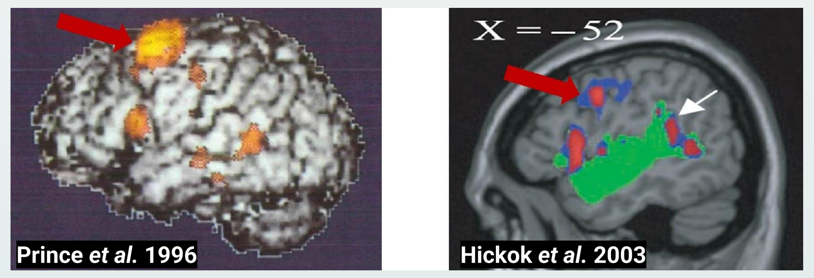
Recent evidence supports two parallel hierarchies of speech motor control

- **Ventral frontal:** segmental articulation & syllabic sequencing ***~ supralaryngeal control***
- **Dorsal frontal:** pitch & prosodic processing ***~ laryngeal control***
- Hickok et al. (2023) proposes two distinct regions in the precentral gyrus, namely dPCSA and vPCSA:



dorsal precentral speech coordination area (**dPCSA**)	**ventral** precentral speech coordination area (**vPCSA**)
auditory-weighted sensorimotor control	**somatosensory**-weighted sensorimotor control
prosodic, **pitch**-related, laryngeal effector	**syllabic**, **phonetic**-related, supralaryngeal effector

- Previous fMRI studies of the speech motor control circuit using a listen and repeat paradigm conflated prosodic and segmental features and identified both dorsal and ventral frontal regions



In this study, we aimed to separate prosodic and segmental features, predicting neural dissociations

- This design enables us to directly test whether these parallel pathways are functionally dissociable during speech production.

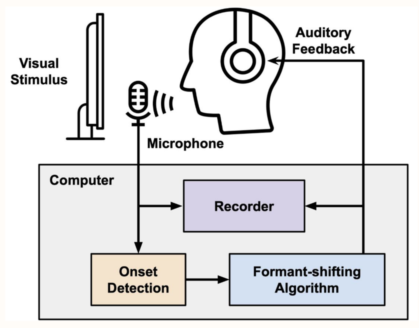
Materials & Methods

"Altered Auditory Feedback" Paradigm

Experimental Setup

MR-safe audio equipment with active **noise-cancellation** system (Optoacoustics)

Pitch/Formant-shifting algorithm ***Audapter*** (Cai et al., 2008; Tourville et al., 2013)

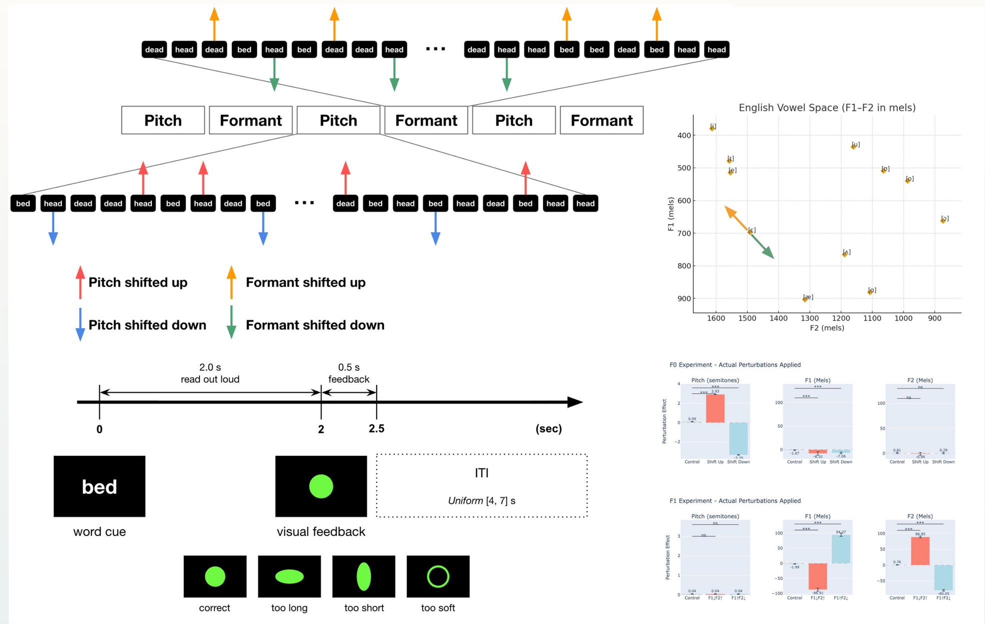


Paradigm

Visual stimuli: 3 English words -ɛd: *bed, dead, head*

F0 shift ±3 st / **F1+F2 shift** 125 mel (45° in F1–F2 plane)
⅙ up-shift + ⅙ down-shift + ⅔ control (no-shift)
total 36 trials per run (≈ 5 min) / total 8 runs

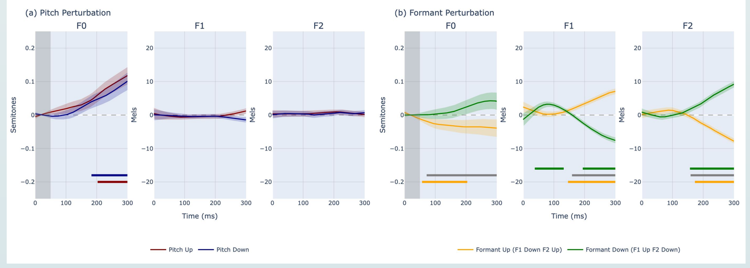
Following Niziolek (2021) F1 AAF Paradigm



A **formant** is a peak in the speech frequency spectrum, representing **resonant frequencies** in the vocal tract. Shifting F1 can change **vowel perception**; for example, shifting F1 up in ***bed*** /bɛd/ makes it sound like ***bad*** /bæd/, and shifting it down makes it sound like ***bid*** /bɪd/.

Acoustic analysis of recorded responses

Confirms selective responses to AAF



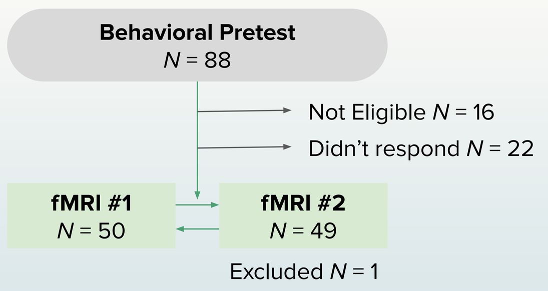
Notes on Project Structure

Same cohort took part in two distinct fMRI paradigms.

Recruited, screened, pre-tested and scanned under identical inclusion criteria (healthy R-handed adults)

Final sample size for this analysis: *N* = 48

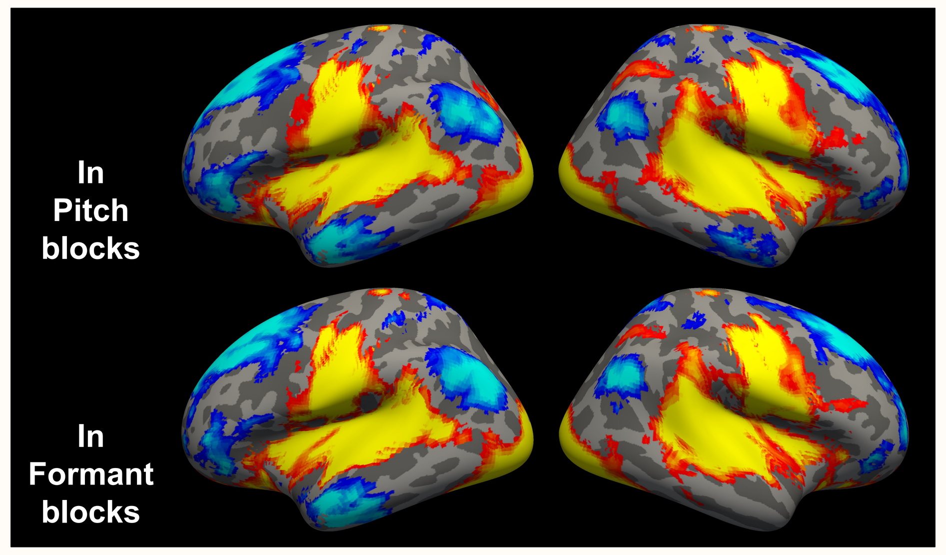
One subject (*N* = 1) excluded due to AAF malfunction



Results

Baseline Production Network Activation

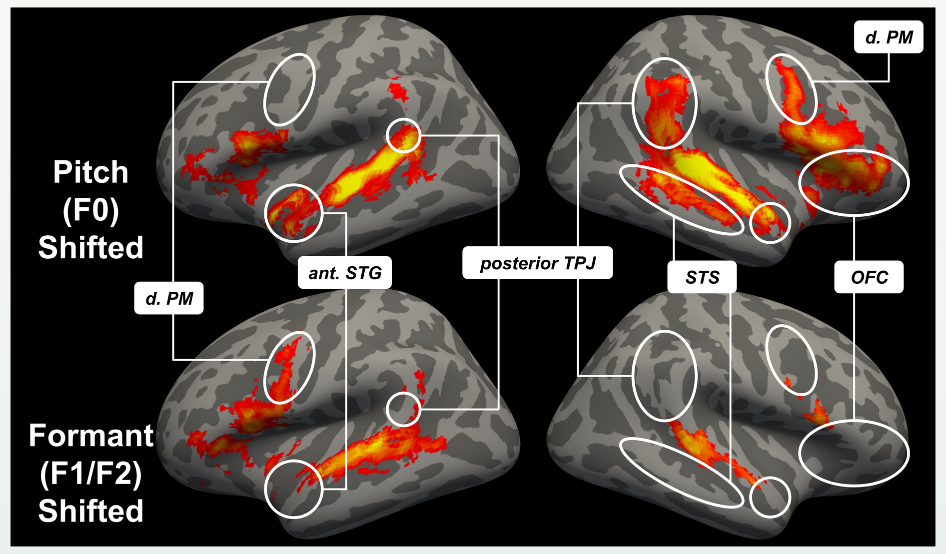
Group-level GLM [non-altered > baseline]



- Word production under non-altered conditions yielded **widespread activation** in the **superior temporal and sensorimotor cortex** bilaterally compared to resting baseline.
- This constitutes the **speech production network** at the word level.
- Visual cortex was also activated due to the visual stimulus presentation.

Neural Effects of *AAF*

Group-level GLM [altered > non-altered]



- Comparing altered versus non-altered trials modulated **subparts of the speech production** network, largely outside of sensorimotor cortex, and did so **differently for F0 vs. F1.**
- **F0 shift** had more activation in **posterior temporal-parietal** *bilaterally*, **anterior STG** *bilaterally*, **right STS**, **right premotor cortex**, and **right orbitofrontal cortex**.
- **F1 shift**: more in **left premotor cortex.**
- **Ventral premotor areas** were active during both shift types in both hemispheres.
- **Dorsal premotor areas** showed opposite patterns in the two hemispheres for the two shift types.

Discussions

Shared and Distinct Neural Networks for F0 vs. F1 Perturbation

- **Both F0 and F1 perturbations** engaged overlapping portions of the speech production network, including bilateral ventral premotor area.
- Despite this overlap, each shift type also recruited distinct pathways, supporting the idea that **laryngeal** and **supralaryngeal** control draw on **partially separable neural subsystems.**

Hemispheric Differences in Dorsal Premotor Cortex

- **Premotor cortex** showed opposite patterns across hemispheres for the two shift types.
- **F0 shift** showed greater dorsal premotor activation, as expected, but **only in the right hemisphere**, indicating right dominance.
- However, **F1 shift** was associated with greater dorsal premotor activation **in the left hemisphere,** in contrast to prediction.
- Ventral premotor activation was associated with both types of shift, but showed differences in the degree of lateralization.

Temporal–Parietal Engagement

- F0 perturbations recruited **posterior temporal–parietal regions** more strongly than F1.
- This pattern may reflect reliance on distinct auditory–motor interface networks.

Remaining Questions & Future Directions

- Unexpected activation in **anterior temporal cortex** (especially from pitch perturbation)
- Future analyses will test individual differences, including individual compensation magnitude.
- Connectivity analyses will probe network-level interactions between premotor, temporal, and parietal regions.

Acknowledgement

National Science Foundation Grant #2242080
Thanks to the following research assistants involved in data collection: Lydia Cheng, Chen Chi, Wandy Kongmebhol, Trisha Le, Andre Musaderyan, Janaki Nair, Bianca (Tati) Rios, and Emily Perez Wong.

Bibliography

Cai S et al., *Proceedings of the 8th ISSP.* **2008**; 65-68.
Hickok G et al. *J Cogn Neurosci.* **2003**;15:673–682.
Hickok G et al., *Brain.* **2023**;146(5):1775-1790.
Niziolek CA et al., *J Speech Lang Hear Res.* **2021**;64(6S):2169-2181.
Price CJ et al. *Brain.* **1996**;119:919–931.
Tourville JA et al., *Lang Cogn Process.* **2011**;26(7):952-981.

Evidence of distinct speech control pathways for prosody and syllabic articulations using speech repetition under fMRI

Minkyu Kim(m.kim@uci.edu)[1], **Chris Naber**[2], **Caroline Niziolek**[2], **Jonathan Venezia**[3], **Bradley Buchsbaum**[4], **Gregory Hickok**[1]

[1]University of California, Irvine, CA, USA; [2]University of Wisconsin–Madison, Madison, WI, USA; [3]VA Loma Linda Healthcare System, Loma Linda, CA, USA; [4]Rotman Research Institute, Toronto, ON, Canada

UCI University of California, Irvine

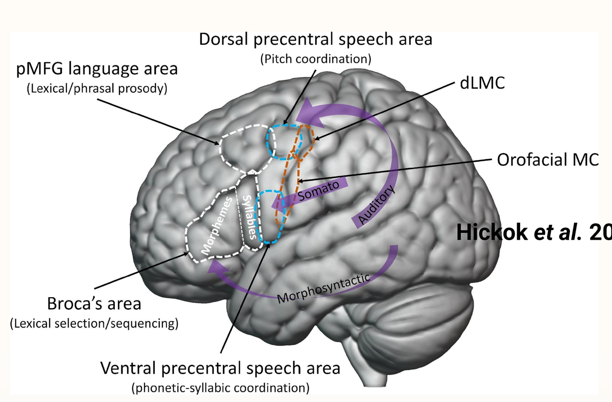
Abstract

***Human speech** requires tightly coordinated control of **prosodic** (pitch) and **segmental features** (consonants/vowels), but whether these functions rely on **shared** or **distinct neural pathways** remains unclear. We manipulated the two dimensions of difficulty during **syllable-sequence repetition under fMRI** and used Representational Similarity Analysis (RSA) to discover both **overlapping** and **dissociable** activity within the **premotor–temporal auditory–motor network**.*

Introduction

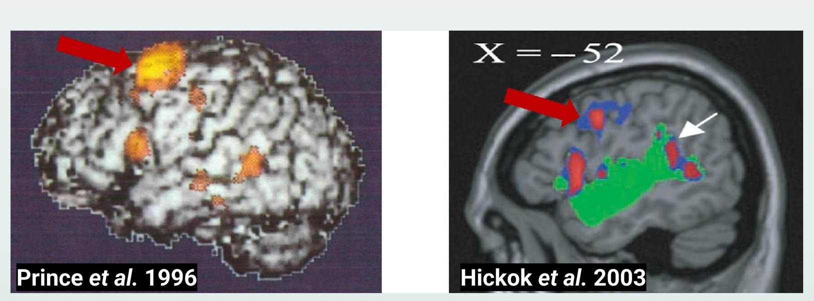
Recent evidence supports two parallel hierarchies of speech motor control

- **Ventral frontal:** segmental articulation & syllabic sequencing *~ **supralaryngeal control***
- **Dorsal frontal:** pitch & prosodic processing *~ **laryngeal control***
- Hickok et al. (2023) proposes two distinct regions in the precentral gyrus, namely dPCSA and vPCSA:



dorsal precentral speech coordination area (**dPCSA**)	**ventral** precentral speech coordination area (**vPCSA**)
auditory-weighted sensorimotor control	**somatosensory**-weighted sensorimotor control
prosodic, **pitch**-related, laryngeal effector	**syllabic**, **phonetic**-related, supralaryngeal effector

- Previous fMRI studies of the speech motor control circuit using a listen and repeat paradigm conflated prosodic and segmental features and identified both dorsal and ventral frontal regions



In this study, we aimed to separate prosodic and segmental features, predicting neural dissociations

- This design enables us to test whether these parallel pathways are functionally dissociable.

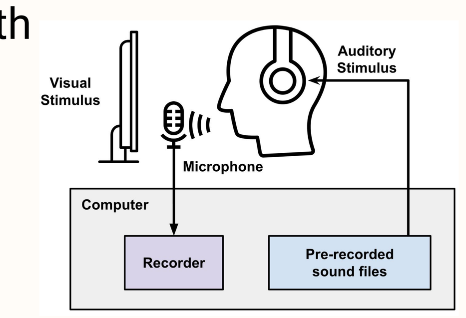
Materials & Methods

Speech Repetition Paradigm

Fifty healthy right-handed adults (N = 50) listened and *covertly* rehearsed **four-syllable sequences** designed to manipulate **two orthogonal dimensions** of articulatory difficulty: (1) **prosodic demands** (e.g. monotone vs. question) and (2) **segmental demands** (e.g. *ba-ba-ba-ba* vs. *ba-di-gu-le*).

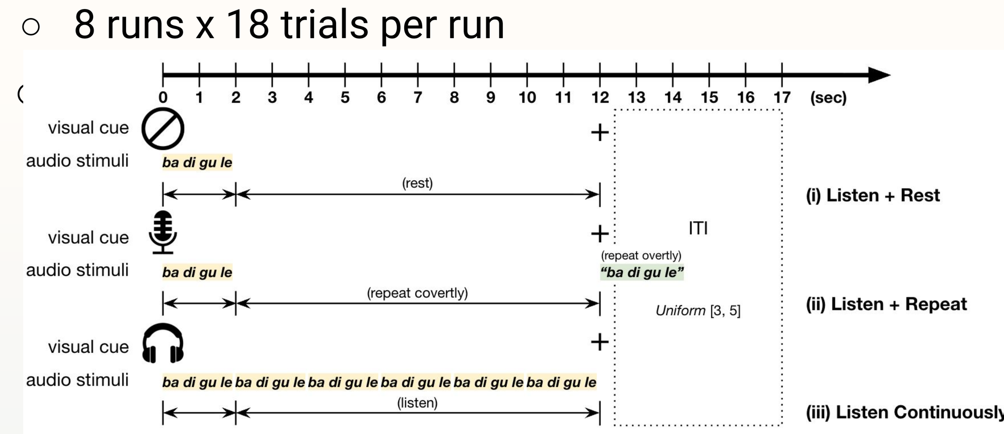
Experimental Setup

MR-safe audio equipment with active **noise-cancellation** system (Optoacoustics)



Paradigm

- Block design, 12 sec/trial
- 3 conditions, 48 stimuli
- 8 runs x 18 trials per run



Auditory Stimuli

Prosodic patterns

- Monotone
- Stressed
- Question intonation

Segmental patterns

- 1 distinct syllables
- 2 distinct syllables
- 4 distinct syllables ∈ {*ba, di, gu, le*}
- hummed melody (= 0 distinct syllables)

Prosody	P1 (L L L L)	P2 (L H L L)	P3 (L H L R?)
Humming	mm–mm–mm–mm (record 4 tokens)	mm–***mm***–mm–mm (record 4 tokens)	mm–***mm***–mm–***mm?*** (record 4 tokens)
S1 (1 syllable)	ba ba ba ba	ba bá ba ba	ba bá ba ba?
	di di di di	di dí di di	di dí di di?
	gu gu gu gu	gu gú gu gu	gu gú gu gu?
	le le le le	le lé le le	le lé le le?
S2 (2 syllables)	ba di ba di	ba dí ba di	ba dí ba di?
	di ba di ba	di bá di ba	di bá di ba?
	gu le gu le	gu lé gu le	gu lé gu le?
	le gu le gu	le gú le gu	le gú le gu?
S4 (4 syllables)	ba di gu le	ba dí gu le	ba dí gu le?
	di gu le ba	di gú le ba	di gú le ba?
	gu le ba di	gu lé ba di	gu lé ba di?
	le ba di gu	le bá di gu	le bá di gu?

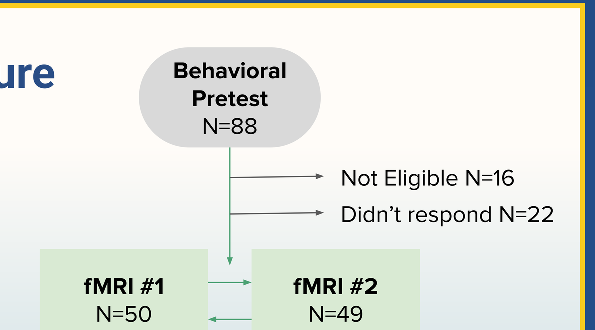
Statistical Analyses

- We used **GLM contrasts** to identify (localize) the **auditory–motor network** via conjunctions of auditory and motor trial activations.
- **Within the resulting ROIs,** we performed Representational Similarity Analysis (RSA) using binary dissimilarity matrices in two dimensions.

Notes on Project Structure

Same cohort took part in two fMRI paradigms

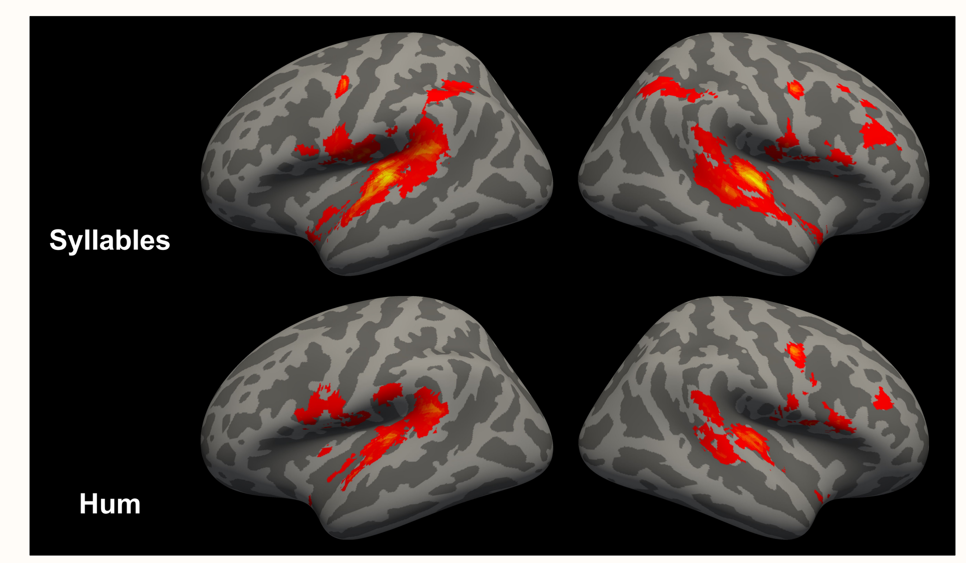
The final sample size was N = 50



Results

Auditory–Motor Network (GLM Analysis)

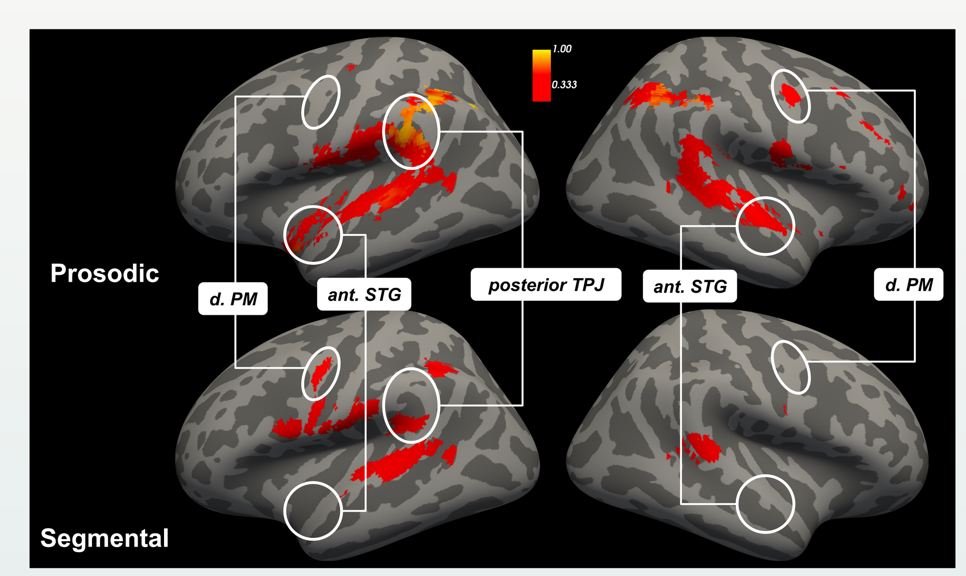
Conjunction of [Listen and Rehearse > Listen and Rest]
***AND** [Listen Continuously > Listen and Rest]*



- **Repetition of syllable sequences** activated the expected auditory–motor network: dorsal/ventral premotor cortex, area Spt, and superior temporal gyrus.
- **Repetition of hummed melodies** (no segmental content) produced **a dropout of left dorsal premotor** activation, while right dorsal premotor activity persisted, indicating **hemispheric asymmetry** in **prosodic/laryngeal control**.

Feature-Specific Representations (RSA)

TFCE-corrected RSA t-statistics Within GLM-defined ROIs



- **Prosodic difficulty** mapped more extensively onto **left tempero-parietal junction, bilateral anterior STG, right ventral sensorimotor cortex** and **right dorsal premotor cortex**.
- **Segmental difficulty** mapped more extensively onto **left ventral premotor and left dorsal premotor**.
- **Left-right asymmetry** is observed in the **dorsal premotor cortex** (prosodic → RH, segmental → LH)
- **Anterior STG** represents **prosodic difficulty** more than segmental difficulty, **bilaterally**.

Discussions

Distinct Networks for Prosodic vs. Segmental Processing

- **Prosodic** and **segmental demands** activated **overlapping** parts of the auditory–motor network **but also diverged** in key regions.
- **Prosodic processing** showed stronger engagement of the right dorsal premotor cortex, left TPJ, and left & right anterior STG, consistent with laryngeal / pitch-based control.
- **Segmental processing** was more left-lateralized, recruiting **left dorsal and left ventral premotor** regions which are hypothesized to be associated with phonetic and articulatory sequencing.
- The result supports a functional division between prosodic (laryngeal/pitch) and segmental (supralaryngeal) control pathways.

Hemispheric Asymmetry and Implications for Speech Motor Models

- The dissociation between **prosodic** and **segmental** processing interacted with hemisphere: **prosodic → RH, segmental → LH.**
- This cross-cutting organization suggests speech motor control relies on partially separable **dorsal–ventral** and **left–right** subsystems.
- Findings refine current models by highlighting that prosodic and phonetic features are represented through distinct but coordinated networks rather than a single unified pathway.

Remaining Questions and Future Works

- Future analyses will test **individual variability** (e.g. prosodic imitation ability, laterality).
- **Connectivity analyses** will help determine the interaction between these regions.
- Extending to additional features (e.g., **rhythm**, **phonations, phonotactics**) may reveal finer subdivisions within the auditory–motor network.

Acknowledgement

National Science Foundation Grant #2242080

Thanks to the following research assistants involved in data collection: Lydia Cheng, Chen Chi, Wandy Kongmebhol, Trisha Le, Andre Musaderyan, Janaki Nair, Bianca (Tati) Rios, and Emily Perez Wong.

Bibliography

Cai S et al., Proceedings of the 8th ISSP (**2008**): 65-68.
Dicther BK et al, *Cell.* **2018**;174(1):21-31.e9.
Glasser MF et al., *Nature*. **2016**; 536(7615):171–178.
Hickok G et al., *Brain*. **2023**;146(5):1775-1790.
Niziolek CA et al., *J Speech Lang Hear Res.* **2021**;64(6S):2169-2181.
Tourville JA et al., *Lang Cogn Process.* **2011**;26(7):952-981.

visit my website!