Prosodic and Phonetic Coordination of Speech: fMRI Evidence for Parallel Circuits

Minkyu Kim[1], Oren Poliva[2], Chris Naber[3], Bradley Buchsbaum[4], Connor Mayer[1], Caroline Niziolek[3], Jonathan Venezia[2], Gregory Hickok[1]

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

UCI University of California, Irvine

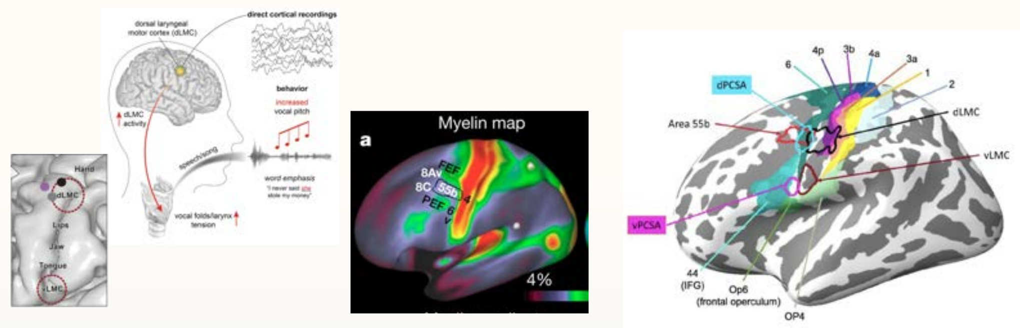
***Human speech** relies on precisely coordinated control of **prosodic** and **segmental features,** yet whether these functions are supported by distinct neural pathways remains unresolved. This study used **functional MRI** with **two complementary approaches**. **Expt 1** investigated **syllable sequence repetition** with stimuli varying in prosodic and phonetic demands to examine their neural representation. **Expt 2** used **Altered Auditory Feedback** to test whether **pitch** and **formant perturbations** engage separable error-monitoring and motor pathways. Results indicate a functional segregation within premotor cortex, with **partly distinct networks** for prosodic and segmental control that **differentially interface with auditory cortex**. These findings refine models of speech motor control and advance understanding of the **neural basis of fluent speech and its disorders**.*

Introduction

Speech can be decomposed into two broad acoustic dimensions

- **Prosody** – pitch/intensity, extend across syllables, larger timescale, intonation/rhythm, largely **laryngeal**
- **Segmental structure** – consonants/vowels, shorter timescale (faster), largely **supralaryngeal**

Hypothesis for two functionally distinct speech coordination areas in premotor cortex



Dichter et al. (2018) showed that bilateral **dorsal laryngeal motor cortex** (dLMC) selectively encode **produced pitch** but not non-laryngeal articulatory movements by high-density intracranial recordings.

Glasser et al. (2016) identified a new region, **area 55b**, in the **posterior middle frontal gyrus**, characterized by **relatively low myelination** and engagement during **language tasks** such as story listening.

Hickok et al. (2022) proposed **two distinct regions** in the **precentral gyrus** (main motivation of this research):

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

RQ: Does the human speech production system comprise two anatomically separable neural pathways that differentially govern prosodic (largely laryngeal) and segmental (largely supralaryngeal) control?

Experiment 1

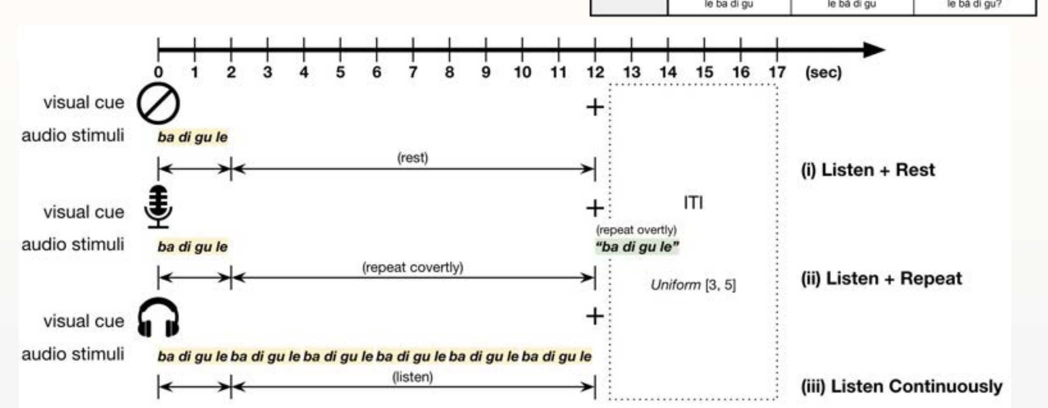
Repetition of Syllable Sequences

TL;DR You repeat syllable sequences that differ in **prosodic demands** (*ba-ba-ba-ba* vs. *ba-bá-ba-ba?*) and **segmental demands** (*ba-ba-ba-ba* vs. *ba-di-gu-le*). Neural activation that **correlates specifically** with one of these dimensions is taken as evidence for the representation of its control. This will help us break down the **auditory-motor network** into distinct sub-networks for **prosodic control** and **supralaryngeal articulatory movements**.

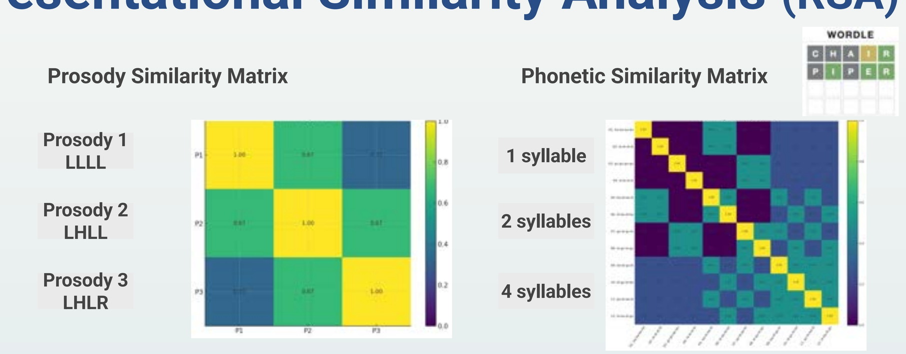
Paradigm & Stimuli

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

Prosody	P1 (L L L L)	P2 (L H L L)	P3 (L H L R?)
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?

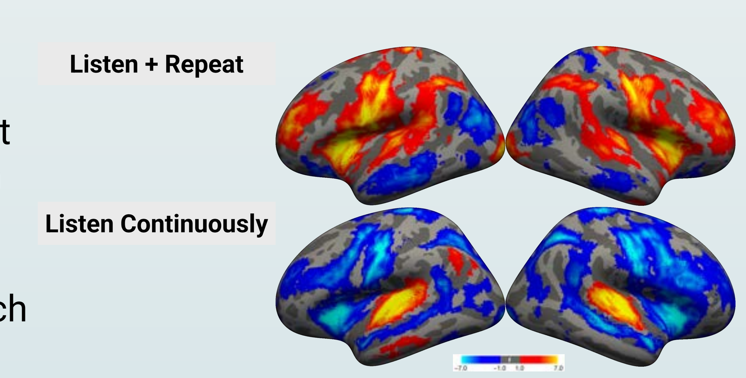


Representational Similarity Analysis (RSA)



Simple GLM Results

Contrasting Listen + Repeat over Listen + Rest gives localization of the broad auditory–motor pathway of speech



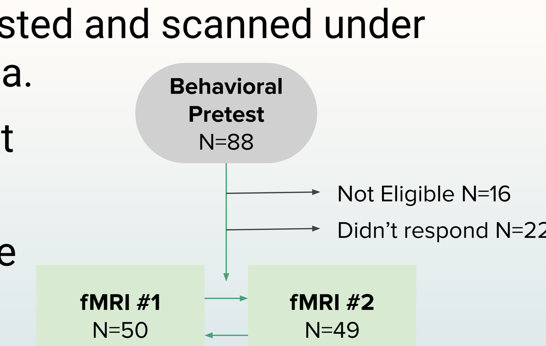
Notes on Project Structure

Same cohort took part in both fMRI paradigms (final sample: N ≈ 50).

Recruited, screened, pre-tested and scanned under identical inclusion criteria.

Eliminates between-subject variability

Within-subject convergence



Experiment 2

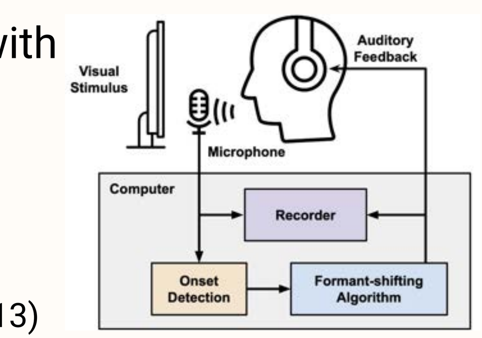
Altered Auditory Feedback: F0 vs F1

TL;DR While reading words on a screen, your voice is **altered in real time**, with either **pitch** or **vowel quality** manipulated and **fed back to your ears**. This will help us find the regions responsible for **pitch control** and **vowel quality control**.

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)

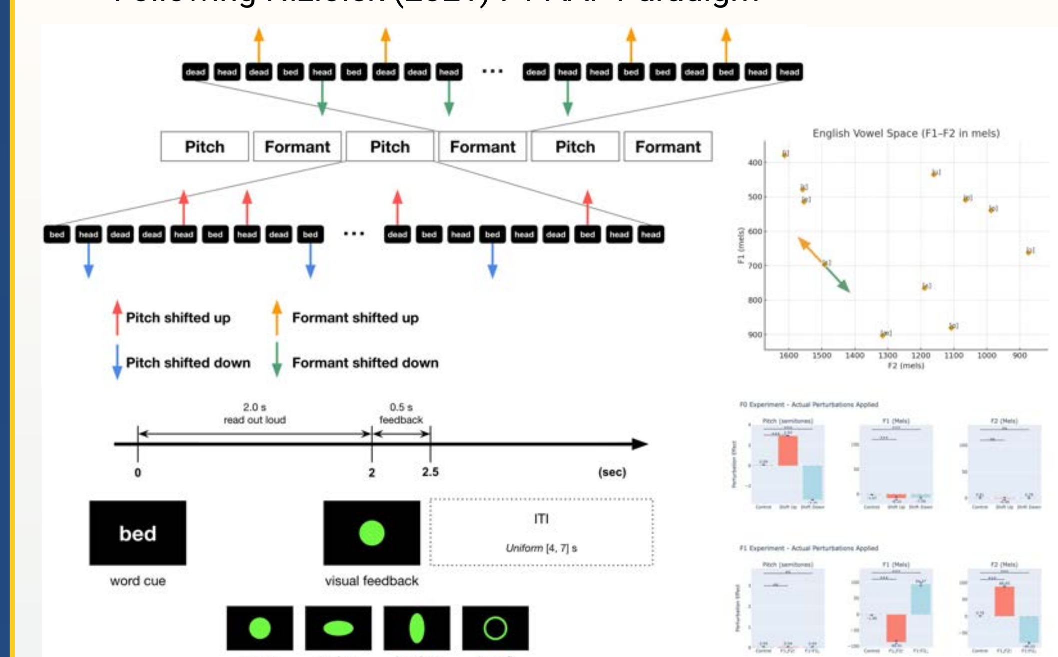


Paradigm

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

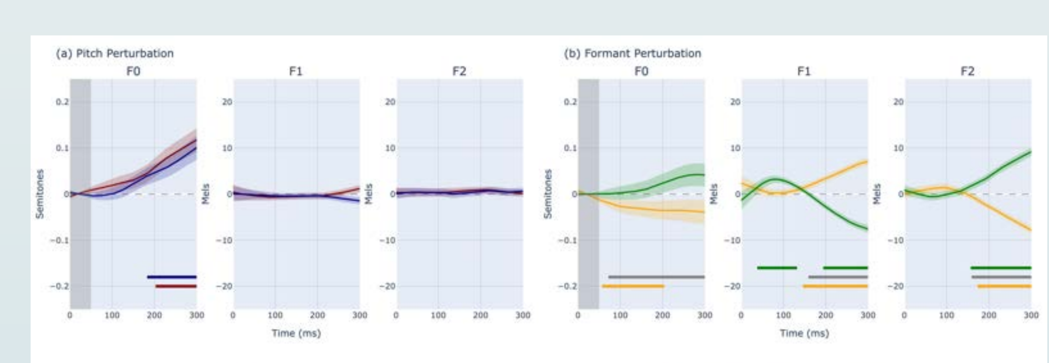
Following Niziolek (2021) F1 AAF Paradigm



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/.

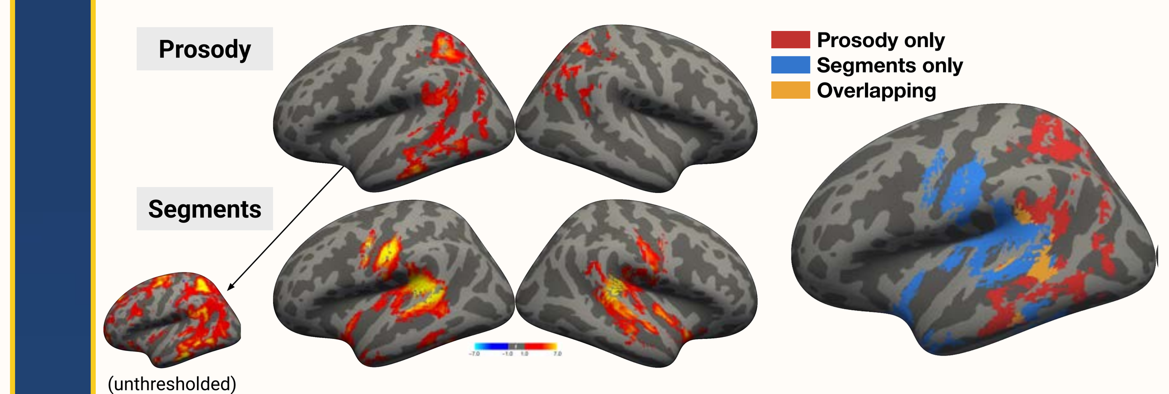
	Treatment	Behavioral Response	Neural Response
Pitch Perturbation	Selectively shift fundamental frequency (F0)	Selectively compensate F0	Prosody-related pathway activated (dPCSA + ...)
Formant Perturbation	Selectively shift first two formants (F1, F2)	Selectively compensate F1 and F2	Segment-related pathway activated (vPCSA + ...)

Behavioral Data



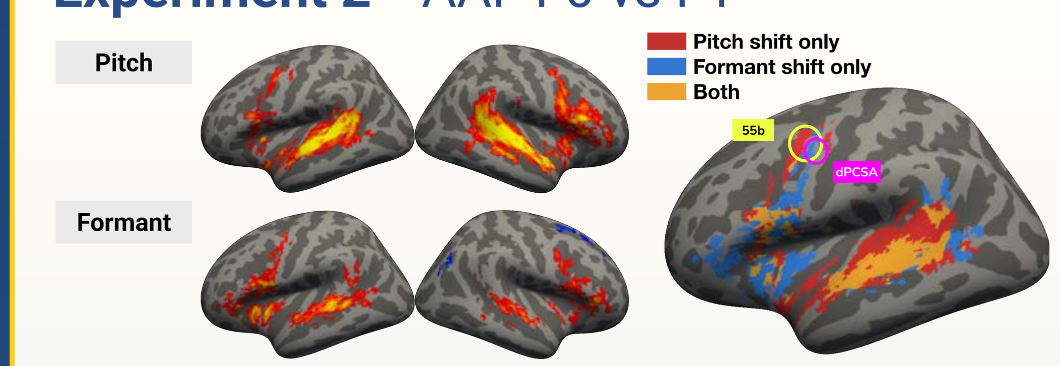
Results

Experiment 1 Speech Repetition RSA



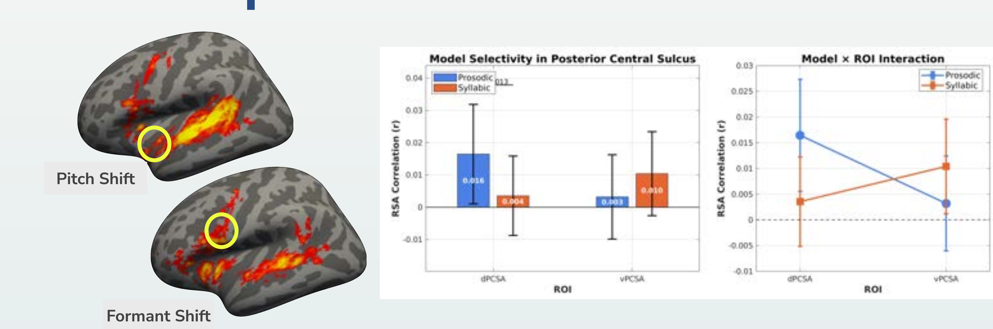
- **Whole-brain segregation** for prosodic/segmental representations achieved with significant overlapping in PT (area Spt)
- Perisylvian connection distinct

Experiment 2 AAF F0 vs F1



- Segregation in the **premotor cortex** (dorso-ventral)
- Differential activation pattern in IFG & temporal lobe

Cross-experiment validation



- Learned ROIs from Expt 2 and tested on Expt 1.
- Interaction effect (ROI:model) observed ($p < 0.05$)

Next Steps

- Individual-level topography, myelination estim.
- Surface-based analysis, ROI-based RSA
- Use same paradigm to test different features
- Combine with MEG/EEG/ECoG paradigms

Acknowledgement

National Science Foundation Grant #2242080

Undergraduate 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!