Selective neural tuning to spoken words in the dorsal auditory stream

Lillian Chang, Srikanth Damera, Xiong Jiang, Josef Rauschecker, Maximilian Riesenhuber
Georgetown University Medical Center — Department of Neuroscience

Objective

- Cortical models of auditory processing[1, 2] propose a dual-stream architecture consisting of ventral and dorsal streams, respectively.
- In the ventral stream, we have previously found lexical representations for auditory words in the anterior superior temporal cortex (aSTC) — the so-called auditory word form area (AWFA)[3-5].
- In addition to the ventral recognition stream, two-stream models of speech processing in the brain postulate that the dorsal stream is involved in transforming sensory (auditory speech) representations into motor (articulatory) representations, implementing a so-called *inverse* internal model that mediates speech production learning[1, 3].
- Our prior effective connectivity analyses between the posterior superior temporal cortex (pSTC) and the premotor cortex (PMC) support the existence of this "inverse model" of speech processing[6]. Furthermore, our prior studies elucidated a sublexical representation for *written words* in the pSTC[7].
- We here hypothesized a **hierarchical representation for auditory words in the dorsal stream**, with a sublexical representation in the pSTC (paralleling our earlier finding for written words[7]) and a lexical representation in the PMC.
- To test our hypothesis, we used an fMRI Rapid Adaptation experiment to test for selectivity for spoken words in the pSTC and PMC regions-of-interest (ROIs) defined individually for each subject.

Probing Lexical Selectivity with fMRI Rapid Adaptation (fMRI-RA)

- A single fMRI voxel captures activity from many neurons with possibly different selectivities. Changes in the BOLD signal may reflect activation of different subgroups of neurons within a voxel, and the average activation level cannot reveal selectivity of the underlying population (Fig. 2A)[11].
- Rapidly repeating the same stimulus will activate the same population of neurons, resulting in a decrease of the BOLD signal. This is known as an "adaptation response" (Fig. 2B)[11].
- Areas posited to contain an auditory lexicon will only show adaptation to rapid repeated presentations of the same spoken word (SAME). Importantly, **presenting two words that differ by only a single phoneme (1PH) should activate distinct populations of neurons resulting in a full release from adaptation** (Fig. 2C), with similar response as to two different words that share no phonemes. Alternatively, an area selective for **sublexical representations will show a graded release from adaptation with increasing phonetic distance** (Fig. 2C).
- Our lab has previously used fMRI-RA to study neural selectivity in both the visual[12-14] and auditory[5, 6, 15] domains.

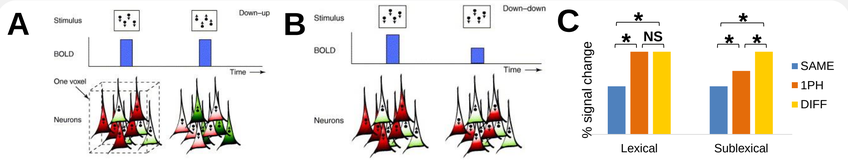

Figure 2: Illustration of the Rapid Adaptation effect. Krekelberg et al. (2006)[11] illustrates neural responses when **A.** a different stimulus or **B.** the same stimulus is rapidly presented **C.** Hypothesized adaptation patterns that reflect neuronal populations with a lexical or sublexical representation for spoken words. Asterisks indicate significance. NS = not significantly different..

Defining Individual-subject ROIs

Regions-of-interest (ROIs) were defined through contrast masking done at the first-level analyses (SPM12).

Contrasts

- AWFA: RW vs. Silence (p<0.001, uncorrected) masked by the contrast of RW vs. Scrambled Words (p-FDR<0.05).
- PMC & pSTC: RW & PW vs. Silence (p<0.001, uncorrected) masked by the contrast of RW & PW vs. Scrambled Words (p<0.05, uncorrected).

ROI clusters and sizes: Thresholds were adjusted to obtain clusters closest to the literature coordinates that were between 10 and 50 voxels.

- AWFA reference (MNI: -61, -15, -5)[3]
- PMC reference (MNI: -50, -6, 47)[8]
- pSTC reference (MNI: -56, -46, 10)[7]

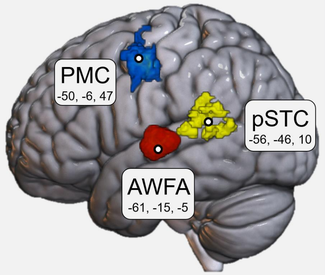

Fig 4. Reference locations to create individual-subject ROIs using the auditory word localizers. AWFA: auditory word form area, pSTC: posterior superior temporal cortex, PMC: premotor cortex. Literature coordinates used as reference points are marked with a white dot.

Conclusions

- **Activation in the PMC follows similar adaptation patterns to the AWFA**—adaptation to the same word, and a complete release from adaptation when the two words differ by just a single phoneme, only for RW and not for novel PW. These findings support the existence of a spoken word lexicon in both the AWFA in the aSTC and the PMC.
- In contrast, we find a **graded release from adaptation in the pSTC**, which is compatible with a sublexical representation.
- In conjunction with our cross-modal functional connectivity analyses between the reading and auditory speech systems[16] and the effective connectivity from the pSTC to the PMC from our prior SEM analyses[6], **our findings support the proposed "inverse model" in the dorsal auditory stream along a hierarchy**—starting with **sublexical representation in the pSTC** and culminating in an **auditory lexicon in the PMC**.

Functional Architecture of Auditory Speech Processing

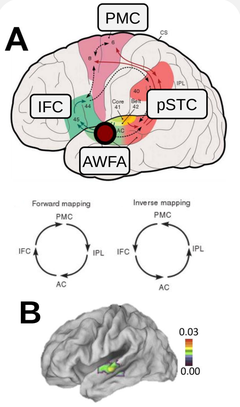

Fig 1A. Model of auditory speech processing from Rauschecker & Scott (2009)[1] with the AWFA included (dark red circle). AWFA location based on **B.** DeWitt & Rauschecker (2012)[3].

- The ventral, "what" pathway is specialized for spoken word recognition, with the AWFA in the aSTC selective for spoken whole words[3, 4].
- The dorsal, "how" stream is thought to subserve speech production and phonological processing[1-3]. Furthermore, studies have shown that motor regions in the dorsal stream, such as the PMC, are activated by both speech production and perception[8, 9].
- In the dorsal stream, we previously reported effective connectivity from the pSTC to the PMC when listening to speech sounds, thought to reflect the mapping of these sounds to their motor articulations following the "inverse model"[6].
- Based on our prior findings, we here investigated the hierarchical representation for spoken words in the pSTC and the PMC.

Experimental Design

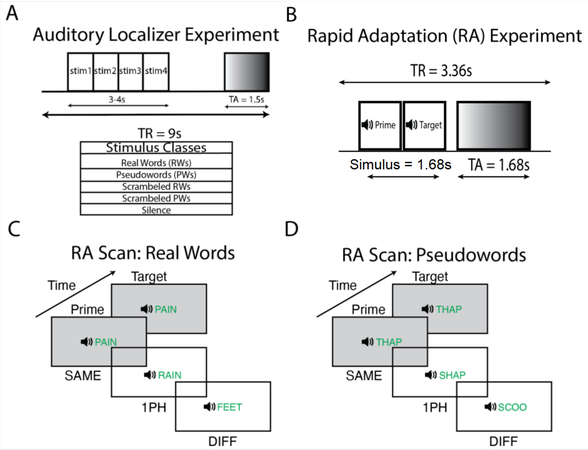

Figure 3: Rapid adaptation and Auditory Localizer Experimental Paradigms. A. Slow clustered acquisition paradigm used in the auditory localizer scan. **B.** Rapid clustered acquisition paradigm used for the RA scans. Each trial was 3.36s long and consisted of a 1.68s silent period, preceded by 1.68s stimulus presentation period during which the word pairs were presented. The first word acted as a prime and the second word as the target, which corresponded to one of the three conditions (i.e., SAME, 1PH, or DIFF). **C-D.** Example trials for Real Words (RW) and Pseudowords (PW), respectively. Figure adapted from Damera et al. (submitted)[5].

Subjects: 26 right-handed English speakers (12 females). Ages 18-34.

Auditory word localizer (Fig. 3A)

- Task: One-back. Participants responded with a button press if any of the 4 stimuli in a short block was a repeat.

fMRI-RA Experiment (Fig. 3B-D)

- Task: Oddball detection. Subjects were asked to attend to all the words and respond when they heard the oddball stimulus (RW or PW containing the rime '-ox' e.g., "socks")
- Conditions:
  - SAME – Subject hears the same word twice.
  - 1PH – Two words that differ by one phoneme.
  - DIFF – Two words that share no phonemes.

Acquisition Parameters:

- 3T Siemens Prisma-fit scanner with a 64-channel head coil.
- Auditory word localizer: TR = 9s, TA = 1.68s
- fMRI-RA: TR = 3.36s, TA = 1.68s

Lexical Representations in the PMC and AWFA

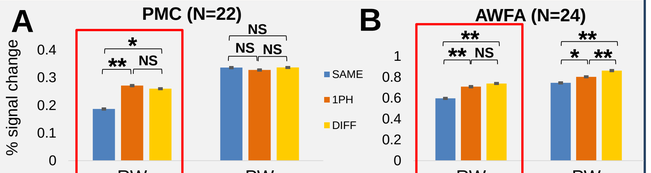

Fig 5. Neural selectivity for spoken words in the **A.** PMC and the **B.** AWFA. RW = Real Words, PW = Pseudowords. p* < 0.05, p** < 0.001, NS = not significant (p ≥ 0.05). Error bars indicate withing-subject SEM.

- Significant adaptation effects for **RWs in the PMC** (Fig. 5A).
  - SAME vs. DIFF: p=0.005; SAME vs. 1PH: p<0.001
  - No significant differences between 1PH and DIFF, p=0.547, indicating full release from adaptation already for two words differing by a single phoneme.
- No adaptation effects for **PWs in the PMC**.
- Our results for RWs in the PMC replicate the adaptation patterns for the AWFA (Fig. 5B) in the anteroventral stream.

Sublexical Representation in pSTC

- Gradual release from adaptation in the **pSTC for RW** (Fig. 6), compatible with a sublexical representation.
  - SAME vs. 1PH: p<0.001
  - 1PH vs. DIFF: p=0.021
  - SAME vs. DIFF: p<0.001)

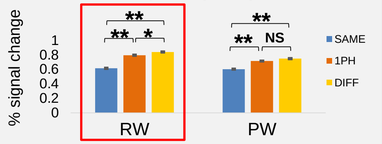

Fig 6. Neural selectivity for spoken words in the pSTC. RW = Real Words, PW = Pseudowords. p* < 0.05, p** < 0.001, NS = not significant (p ≥ 0.05). Error bars indicate within-subject SEM.

Acknowledgements

- Research supported by the National Science Foundation (BCS-1756313).
- Dr. Ian DeWitt for the auditory scrambling code used in the localizer.

References

1. Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature neuroscience*, *12*(6), 718-724.
2. Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature reviews neuroscience*, *8*(5), 393-402.
3. DeWitt, I., & Rauschecker, J. P. (2012). Phoneme and word recognition in the auditory ventral stream. *Proceedings of the National Academy of Sciences*, *109*(8), E505-E514.
4. Damera, S. R., Mattei, J., Glezer, L., Cox, P., Jiang, X., Rauschecker, J., & Riesenhuber, M. A Lexicon in the Anterior Auditory Ventral Stream: Preliminary evidence from an fMRI-RA study [abstract]. In: Eleventh Annual Meeting of the Society for the Neurobiology of Language; 2019 Aug 20-22; Helsinki, Finland. Abstract E66.
5. Damera, S. R., Chang, L., Nikolov, P., Mattei, J. A., Banerjee, S., Glezer, L. S., Cox, P. H., Jiang, X., Rauschecker, J. P., Riesenhuber, M. (2022). Evidence for a Spoken Word Lexicon in the Auditory Ventral Stream. *In submission*.
6. Chevillet, M. A., Jiang, X., Rauschecker, J. P., & Riesenhuber, M. (2013). Automatic phoneme category selectivity in the dorsal auditory stream. *Journal of Neuroscience, 33*(12), 5208-5215.
7. Glezer, L. S., Eden, G., Jiang, X., Luetje, M., Napoliello, E., Kim, J., & Riesenhuber, M. (2016). Uncovering phonological and orthographic selectivity across the reading network using fMRI-RA. *Neuroimage, 138*, 248-256.
8. Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature neuroscience, 7*(7), 701-702.
9. Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D., & Iacoboni, M. (2007). The essential role of premotor cortex in speech perception. *Current Biology, 17*(19), 1692-1696.
10. Rauschecker, J. P. (2011). An expanded role for the dorsal auditory pathway in sensorimotor control and integration. *Hearing research, 271*(1-2), 16-25.
11. Krekelberg, B., Boynton, G. M., & van Wezel, R. J. (2006). Adaptation: from single cells to BOLD signals. *Trends in neurosciences, 29*(5), 250-256.
12. Glezer, L. S., & Riesenhuber, M. (2013). Individual variability in location impacts orthographic selectivity in the "visual word form area". *Journal of Neuroscience, 33*(27), 11221-11226.
13. Glezer, L. S., Kim, J., Rule, J., Jiang, X., & Riesenhuber, M. (2015). Adding words to the brain's visual dictionary: novel word learning selectively sharpens orthographic representations in the VWFA. *Journal of Neuroscience, 35*(12), 4965-4972.
14. Jiang, X., Rosen, E., Zeffiro, T., VanMeter, J., Blanz, V., & Riesenhuber, M. (2006). Evaluation of a shape-based model of human face discrimination using FMRI and behavioral techniques. *Neuron, 50*(1), 159-172.
15. Jiang, X., Chevillet, M. A., Rauschecker, J. P., & Riesenhuber, M. (2018). Training humans to categorize monkey calls: auditory feature-and category-selective neural tuning changes. *Neuron, 98*(2), 405-416.
16. Chang, L., Banerjee, N., Cho, P., Pendri, V., Nikolov, P., Damera, S., Jiang, X., Rauschecker, J., Riesenhuber, M. Evidence for selective and modality-dependent functional connectivity and cross-modal activation between auditory and visual word form areas: A possible substrate to learn a visual lexicon? [abstract]. In: Cognitive Neuroscience Society 2022 Annual Meeting; 2022 April 23-26; San Francisco, California. Abstract B76.