Temporal regions are the epicenter of language processing in the human brain

Greta Tuckute (gretatu@mit.edu)
Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology,
43 Vassar Street Cambridge, MA 02139 USA

Aalok Sathe (asathe@mit.edu)
Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology,
43 Vassar Street Cambridge, MA 02139 USA

Jingyuan Selena She (jshe@mit.edu)
Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology,
43 Vassar Street Cambridge, MA 02139 USA

Evelina Fedorenko (evelina9@mit.edu)
Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology,
43 Vassar Street Cambridge, MA 02139 USA

Abstract:

Human language processing is seemingly effortless and is supported by an interconnected set of frontal and temporal areas in the left-hemisphere. These areas respond in similar ways to diverse linguistic manipulations, suggesting some degree of redundancy. Here, we ask: what are the most essential parts of the human language system? We define a region as *essential* if it shows ubiquitous engagement during the processing of any linguistic input, written or spoken. We find that the temporal areas are essential to language processing, and that frontal areas do not appear necessary when linguistic input is sufficiently easy to process. Our results suggest that although in most scenarios language areas work together to efficiently process linguistic input, the temporal component of the language network is indispensable whereas other language areas play an auxiliary role and perhaps help interface linguistic representations with other networks across the brain.

Keywords: language; redundancy; amodal processing; fMRI

Introduction

Language processing recruits an interconnected network of frontal and temporal areas in the left hemisphere (Binder et al., 1997, Lipkin et al., 2022). These brain areas are functionally similar: they are all highly selective for language relative to diverse inputs and tasks and they show similar responses to diverse linguistic manipulations (Fedorenko et al., 2024). This similarity among different parts of the language network may suggest that the language system is characterized by some redundancy. Relatedly, in many cases, focal damage to language regions leads to relatively quick recovery (Kertesz, 1977; Wilson et al., 2023). Here we use fMRI to ask: which part(s) of the language network are essential to language processing and which parts may be redundant? Specifically, we aim to identify the essential component(s), *the epicenter,* of language processing in the brain which is needed to process any linguistic input.

Methods

Listening/reading fMRI dataset: 9 native English speakers (4 female; 2 left-handed) read and listened to n=400 corpus-extracted sentences in an event-related design while their brain activity was recorded using 3T fMRI (TR=2s, 2mm^3 voxels). Half of the sentences were presented visually (full sentence on the screen), and the other half—auditorily (2s stimulus presentation, 4s inter-stimulus interval). Critically, because surprisal has been shown to strongly affect the level of response in the language network (e.g., Shain et al., 2020; Heilbron et al., 2022), sentences were designed to fall into a low or high surprisal condition (200 in each) based on 5-gram surprisal estimates and were sampled from both transcribed spoken and written text corpora. Sentences were all 6 words long and closely matched on number of characters and—for auditory stimuli—speech rate. During the fMRI session, the sentences were randomized, and presentation modality was counterbalanced across participants. In addition to the critical experiment, participants completed a reading-based language 'localizer' task contrasting sentences and strings of nonwords in a blocked fMRI design (Fedorenko et al., 2010), as well as an auditory version of this localizer with identical design but consisting of intact and degraded speech (Scott et al., 2017).

Reading fMRI dataset: For replication, we used the dataset from Tuckute et al. (2024), where 5 participants each read n=1,000 semantically diverse corpus-extracted sentences (also 6 word-long) in an event-related design.

Results

Isolating the essential part(s) of the language network: Our goal is to identify brain regions that are ubiquitously engaged in processing *any kind of linguistic input* irrespective of both i) linguistic complexity/content, and ii) modality (reading or listening). To do so, we propose a new metric that targets these ubiquitous amodal language regions ("UbiquitousAmodal metric") and leverages both the critical event-related data and the language localizer data. In particular, for each participant, we created 3 brain maps:

1) **Ubiquitous engagement:** The number of sentences (ranging from 0 to 400) that elicit a positive (above-fixation) response in a given voxel.
2) **Language-selectivity (written):** The *t*-statistic comparing responses to visually presented sentences and strings of nonwords.
3) **Language-selectivity (spoken):** The *t*-statistic comparing responses to auditorily presented intact and degraded speech.

Each of the 3 maps were normalized to values between 0 and 1, and summed (other aggregation approaches yield similar maps), resulting in an UbiquitousAmodal map, where a value of each voxel is between 0 and 3. **Fig. 1A** shows the 3 separate brain maps and the resulting UbiquitousAmodal maps for sample participants. The UbiquitousAmodal maps were

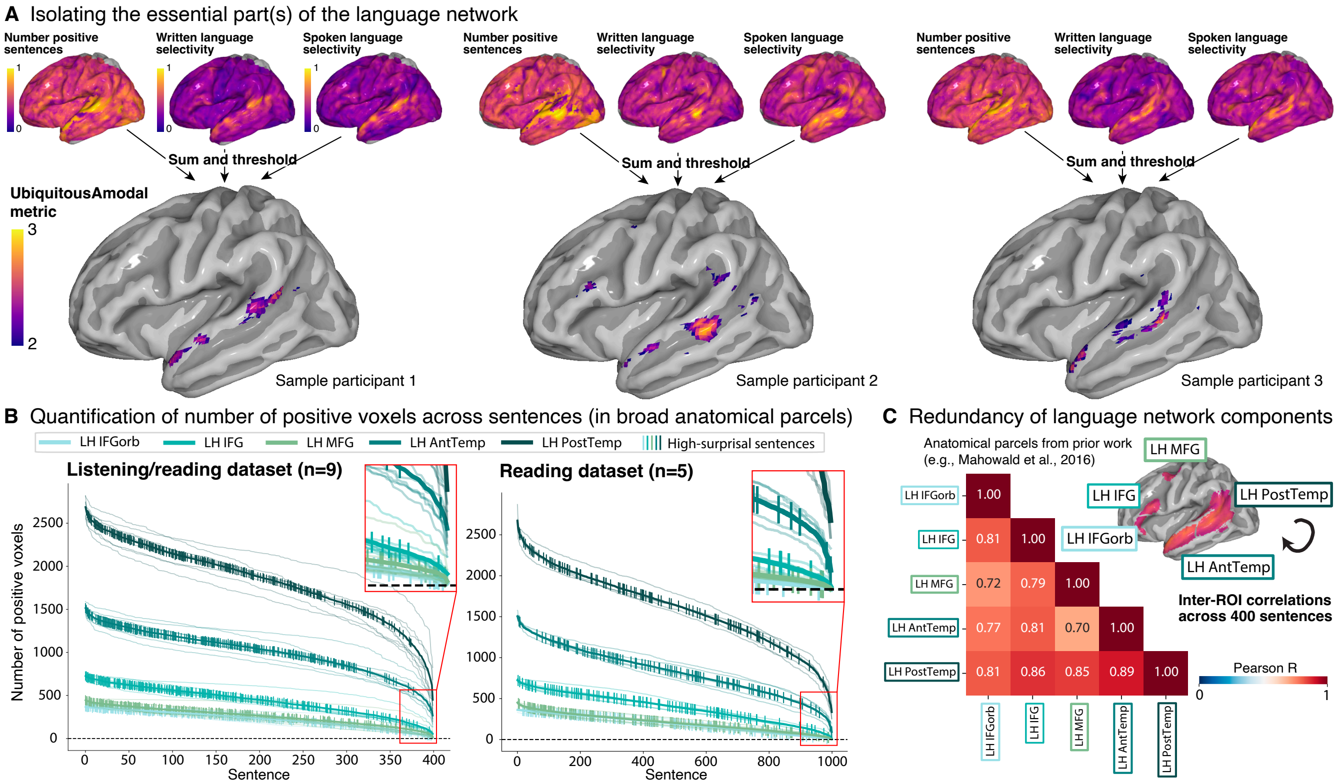

Fig. 1: **A.** For each of 3 sample participants, the upper row shows the three maps used to compute the UbiquitousAmodal maps (second row). **B.** The number of positive voxels in 5 broad anatomical parcels (visualized in C) across all sentences. Thin lines show values for each participant, thick line shows the average. Vertical markers denote high-surprisal sentences. **C.** Correlation of the mean response in the top 10% UbiquitousAmodal voxels within the 5 anatomical parcels across participants.

thresholded at a value of 2, meaning that only voxels that are engaged in processing a majority of the 400 sentences (listening/reading) as well as linguistic input from the two language localizers (listening- and reading-based) are visible on these maps. The maps qualitatively show a clear temporal component (quantified via Bonferroni-corrected, independent t-tests for eight frontal-temporal parcel pairs; p<<.001).

Quantifying the brain areas that are responsive to all linguistic input: Next, we quantified which parts of the brain are engaged during processing of any sentence, irrespective of its surprisal. Our main experiment consisted of fine-grained, sentence-level responses to 400 short (6-word-long) sentences: 200 low-surprisal sentences (e.g., *"I was just a little nervous."*) and 200— high-surprisal *("A chill skated down Cassie's spine."*). **Fig. 1B** shows the number of positive (above-fixation) voxels in 5 broad anatomical parcels (see **Fig. 1C**) within which individuals typically show activity for the language localizer contrast (e.g., Lipkin et al., 2022). For the listening/reading dataset (left panel in **1B**), the vertical markers denote if a sentence at that index was categorized as high-surprisal for more than half of the participants (the reading dataset was not constructed based on surprisal, but instead we categorized high-surprisal as surpassing the 75th percentile of surprisal in the set).

These plots show that certain sentences, largely low-surprisal sentences, can be processed focally in the temporal areas (especially the posterior one), without the engagement of the frontal language areas (~0 positive voxels to several sentences).

Finally, having established that the temporal areas appear more essential than the frontal ones, we show that even though frontal areas do not "come online" for certain linguistic inputs, the frontal and temporal areas show similar fine-grained preferences across the 400 sentences (**Fig. 1C**). This high degree of inter-region correlation suggests that the frontal regions may be functionally redundant with the temporal ones, opening the door to questions regarding when and why the frontal areas are recruited during language processing.

Discussion

This work represents a step towards distilling the core parts of the language processing system in the brain. Thus, although different language areas appear to be functionally similar, some areas appear to play a more critical role in language processing. This work reconciles seemingly disparate evidence from i) prior fMRI studies that typically observe the entire fronto-temporal language network active during language tasks and ii) aphasia research, which has suggested that the posterior temporal component of the language network is particularly important given damage to it leads to longer-lasting linguistic deficits.

Acknowledgments

G.T. was supported by the Amazon Fellowship from the Science Hub, the International Doctoral Fellowship from the AAUW, and the K. Lisa Yang ICoN Graduate Fellowship. E.F. was supported by National Institutes of Health award U01-NS121471 and by research funds from the McGovern Institute for Brain Research, the Department of Brain and Cognitive Sciences, the Simons Center for the Social Brain, and the MIT Quest for Intelligence.

References

Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. Journal of Neuroscience, 17(1), 353-362.

Fedorenko, E., Hsieh, P. J., Nieto-Castañón, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New method for fMRI investigations of language: defining ROIs functionally in individual subjects. Journal of neurophysiology, 104(2), 1177-1194.

Fedorenko, E., Ivanova, A. A., & Regev, T. I. (2024). The language network as a natural kind within the broader landscape of the human brain. Nature Reviews Neuroscience, 1-24.

Heilbron, M., Armeni, K., Schoffelen, J. M., Hagoort, P., & De Lange, F. P. (2022). A hierarchy of linguistic predictions during natural language comprehension. Proceedings of the National Academy of Sciences, 119(32), e2201968119.

Kertesz, A., & McCabe, P. (1977). Recovery patterns and prognosis in aphasia. Brain: a journal of neurology, 100, 1-18.

Lipkin, B., Tuckute, G., Affourtit, J., Small, H., Mineroff, Z., Kean, H., ... & Fedorenko, E. (2022). Probabilistic atlas for the language network based on precision fMRI data from> 800 individuals. Scientific Data, 9(1), 529.

Scott, T. L., Gallée, J., & Fedorenko, E. (2017). A new fun and robust version of an fMRI localizer for the frontotemporal language system. Cognitive neuroscience, 8(3), 167-176.

Shain, C., Blank, I. A., van Schijndel, M., Schuler, W., & Fedorenko, E. (2020). fMRI reveals language-specific predictive coding during naturalistic sentence comprehension. Neuropsychologia, 138, 107307.

Tuckute, G., Sathe, A., Srikant, S., Taliaferro, M., Wang, M., Schrimpf, M., ... & Fedorenko, E. (2024). Driving and suppressing the human language network using large language models. Nature Human Behaviour, 1-18.

Wilson, S. M., Entrup, J. L., Schneck, S. M., Onuscheck, C. F., Levy, D. F., Rahman, M., ... & Kirshner, H. S. (2023). Recovery from aphasia in the first year after stroke. Brain, 146(3), 1021-1039.