Title: The extended language network: Language-responsive brain areas whose contributions to language remain to be discovered

Abbreviated title: The extended language network

Authors: Agata Wolna[1], Aaron Wright[1], Colton Casto[1,2,3], Samuel Hutchinson[1], Benjamin Lipkin[1], Evelina Fedorenko[1,3]

Affiliations: [1]McGovern Institute for Brain Research and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA; [2]Kempner Institute for the Study of Natural and Artificial Intelligence, Harvard University, Cambridge, MA; [3]Program in Speech and Hearing Bioscience and Technology (SHBT), Harvard University, Cambridge, MA

Corresponding authors: awolna@mit.edu, evelina9@mit.edu

Number of pages: 18 (incl. title, abstract, introduction, results, methods, and discussion)

Number of figures: 5 (not incl. supplement)

Number of tables: 2 (not incl. supplement)

Number of words:

Abstract: 243
Introduction: 656
Discussion: 1,549

Conflict of interest: The authors declare no competing financial interests.

Acknowledgements: We would like to acknowledge the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research at MIT, and its support team (Steve Shannon and Atsushi Takahashi). We would also like to thank Moshe Poliak for statistical consultations as well as former and current EvLab members for contributions to the fMRI data collection efforts over the years. This work was partially supported by a grant from the Simons Foundation to the Simons Center for the Social Brain at MIT. AW was supported by a postdoctoral fellowship from the Simons Center for the Social Brain. CC as supported by the Kempner Institute for the Study of Natural and Artificial Intelligence at Harvard University. EF was additionally supported by research funds from the McGovern Institute for Brain Research, the Department of Brain and Cognitive Sciences, and the MIT Siegel Family Quest for Intelligence.

Author Contributions: Conceptualization: AgW, EF; Methodology: all authors; Software: AgW, AaW; Validation: AaW; Formal analysis: AgW, AaW; Investigation: all authors; Data

curation: all authors; Writing – original draft: AgW, EF; Writing – review & editing: AaW, CC, SH, BL; Visualization: AgW, AaW, SH; Supervision: EF; Project administration: AgW, EF.

Abstract

Although language neuroscience has largely focused on 'core' left frontal and temporal brain areas and their right-hemisphere homotopes, numerous other areas—cortical and subcortical—have been implicated in linguistic processing. However, these areas' contributions to language remain unclear given that the evidence for their recruitment comes from diverse paradigms, many of which conflate language processing with perceptual, motor, or task-related cognitive processes. Using fMRI data from 772 participants (438 females, 334 males) performing an extensively-validated language 'localizer' paradigm that isolates language processing from other processes, we a) delineate a comprehensive set of areas that respond reliably to language across written and auditory modalities, and b) evaluate these areas' selectivity for language relative to a demanding non-linguistic task. In line with prior claims, many areas outside the core fronto-temporal network respond during language processing, and most of them show selectivity for language relative to general task demands. These language-selective areas of the extended language network include areas around the temporal poles, in the medial frontal cortex, in the hippocampus, and in the cerebellum, among others. Although distributed across many parts of the brain, the extended language-selective network still only comprises ~3.5% of the grey matter volume, challenging the view that the entire brain processes language. These newly identified language-selective areas can now be systematically characterized to decipher their contributions to language processing, including testing whether these contributions differ from those of the core language areas.

Significance statement

Language processing consistently recruits a left-lateralized fronto-temporal brain network, but language tasks often additionally engage areas outside this core system. In an fMRI dataset of 772 participants performing a validated language localizer task, we identified 17 brain areas outside the core fronto-temporal network that respond to both auditory and written language. Most of these areas show selectivity for language, including regions in the temporal poles, medial frontal cortex, hippocampus, and cerebellum. Despite its large number of components, this extended language network still only takes up ~3.5% of the grey matter volume, challenging the view that the entire brain processes language. These findings lay the foundation for systematic characterization of these newly identified non-canonical language areas.

Introduction

Language processing engages a left-lateralized fronto-temporal brain network—the "language network"—which selectively supports comprehension and production across modalities, including computations related to word retrieval and combinatorial syntactic and semantic processing (Vigneau et al., 2006; Turker et al., 2023; Fedorenko et al., 2024). These areas have been consistently identified across diverse neuroimaging approaches (PET, fMRI, MEG, and intracranial recordings) and paradigms, and also implicated in studies of patients with aphasia (see **Table OSF1** for select reviews, meta-analyses, and empirical papers). In addition to this 'core' network and its right-hemispheric homotopic regions (Lindell, 2006; Ferstl et al., 2008; Mahowald & Fedorenko, 2016; Martin et al., 2022), several other brain areas have been implicated in aspects of language. In the cortex, these include medial frontal areas (Ardila, 2020; Bohsali & Crosson, 2016; Hertrich et al., 2016), medial and lateral parietal areas (Hasson et al., 2007; Peschke et al., 2012; Papagno et al., 2017), temporal poles (Ferstl et al., 2008; Holland & Lambon Ralph, 2010; Lambon Ralph et al., 2017), ventral temporal areas (Krauss et al., 1996; Bédos Ulvin et al., 2017; Snyder et al., 2023), and even occipital areas in the blind (Bedny et al., 2011; Reich et al., 2011; Kim et al., 2017) but also sighted populations (Dikker et al., 2009, 2010; Seydell-Greenwald et al., 2023). Subcortically, language-implicated areas include the hippocampus (Covington & Duff, 2016; Dijksterhuis et al., 2024; reviews: Duff & Brown-Schmidt, 2012, 2017; Piai et al., 2016), the thalamus (Wahl et al., 2008; review: Klostermann & Ehlen, 2013), and basal ganglia structures, including the caudate, the pallidum, and the putamen (Booth et al., 2007; Crosson et al., 2003; Kotz, 2009; Oberhuber et al., 2013; Thibault et al., 2021, 2025). Finally, parts of the cerebellum have long been implicated in speech and language (reviews: Murdoch, 2010; Mariën et al., 2014; LeBel & D'Mello, 2023).

However, the precise contributions of these non-canonical brain areas to language remain unclear. Much of the evidence comes from studies relying on single experimental paradigms, which often fail to isolate language processing from perceptual, motor, or general cognitive demands (for discussion, see Fedorenko et al., 2024). Further, the group-averaging approach—adopted by the majority of past neuroimaging studies—complicates across-study comparisons. In this approach, individual maps are averaged voxel-wise, and group-level activation peaks are reported in a standardized coordinate system (e.g., MNI; Mazziotta et al., 2001). However, inter-individual neuroanatomical variability leads to differences in structural and functional topographies (Frost & Goebel, 2012; Tahmasebi et al., 2012; Fedorenko, 2021; Michon et al., 2022), which may shift the group peaks across studies even for the same paradigm. Besides, any macro-anatomical structure encompasses multiple functionally distinct areas, making it difficult to unambiguously determine whether nearby peaks reported in different studies correspond to the same functional area (earlier discussions: Brett et al., 2002; Saxe et al., 2006).

One approach that circumvents these challenges is functional localization (Saxe et al., 2006; Kanwisher, 2010; Fedorenko et al., 2010). Instead of averaging brains together, this approach identifies functional areas within individuals using 'localizer' tasks—robust contrasts that target a particular perceptual or cognitive process. This approach ensures that functionally equivalent area(s) are referred to across individuals and studies, supporting the systematic

characterization of each area's responses across paradigms, as needed to understand its computations. At the same time, identifying these areas and testing their properties in each brain individually—effectively, working in each individual's functional native space (regardless of the anatomical, stereotactic space that the data are registered to)—allows to preserve high functional resolution (Nieto-Castañón & Fedorenko, 2012): i.e., the ability to differentiate the relevant areas from nearby, functionally distinct areas.

Here, we use an established language localizer and a large fMRI dataset to search for language-selective areas outside the core language network. In line with past claims, we find several such areas, and many of them show selectivity for language relative to a demanding non-linguistic task, which makes them exciting targets for future investigations.

Methods

Overview

We leverage a dataset of 772 participants each of whom performed an extensively validated reading-based language localizer (Fedorenko et al., 2010) and one or two additional tasks (**Figure 1A**). In particular, 489 of the 772 participants performed a demanding spatial working memory task, which allowed us to begin evaluating the selectivity of the non-canonical language-responsive areas. 86 of these 489 participants additionally performed an auditory version of the language localizer (Malik-Moraleda, Ayyash et al., 2022), which allowed us to evaluate whether the observed language responses are robust to modality (reading vs. listening). If a brain region, akin to the core frontal and temporal language areas, responds to both written and auditory linguistic inputs (Fedorenko et al., 2010; Vagharchakian et al., 2012; Regev et al., 2013; Deniz et al., 2019), its computations plausibly have to do with the content of the linguistic signals rather than their surface (visual or auditory) properties.

In the critical analysis, we define regions of interest functionally within individual participants. To constrain the selection of these individual-level functional regions of interest (fROIs), we use functional parcels derived from a group-level representation of the language localizer activity (Fedorenko et al., 2010). Because the fMRI signal is weaker in subcortical areas, which are further away from the receiver coil, and to ensure that we don't miss any language-responsive areas, we complement these parcels with subcortical anatomical parcels, which we use in the same way (to constrain individual-level fROIs). Finally, because standard cortical atlases are commonly used for ROI definition, we comprehensively examine language responsiveness and selectivity in three such atlases, in order to specify the relevant subsets of areas within which reliable and selective responses to language can be observed, and to further highlight the importance of functional localization within individuals instead of using group-level anatomical masks.

Participants

The structure of the dataset is shown in **Figure 1A**. We used three tasks (**Figure 1B**): a standard reading-based language localizer (Fedorenko et al., 2010), an auditory language localizer based

on excerpts from *Alice in Wonderland* (Malik-Moraleda, Ayyash et al., 2022), and a spatial working memory task (Fedorenko et al., 2013; Assem et al., 2020).

772 neurotypical adults (age: mean=27, range=18-81; 438 females) each completed the standard reading-based language localizer (Fedorenko et al., 2010). A subset of 706 participants (age: mean=27, range=18-81; 404 females; *subset A*), all native speakers of English, were included in the GSS (Group-constrained Subject-Specific) analysis (Fedorenko et al., 2010), to create a set of extended language parcels (Group-level parcels used to constrain the definition of individual fROIs). 489 of the 772 participants (age: mean=29, range=17-80; 225 females; *subset B*) performed a spatial working memory task (Fedorenko et al., 2013; Assem et al., 2020), and 86 of these 489 participants (age: mean=28, range=19-45; 44 females; *subset C*) additionally performed an auditory language localizer task. This subset consisted of 20 native English speakers and 66 native speakers of diverse languages (all proficient in English), who performed the auditory localizer in their native language (data partially reported in (Malik-Moraleda, Ayyash et al., 2022). All participants gave written informed consent, in accordance with the requirements of MIT's Committee on the Use of Humans as Experimental Subjects (COUHES) and were paid for their participation.

Tasks

Reading-based language localizer task

Participants passively read English sentences and lists of pronounceable nonwords, which appeared on the screen one word/nonword at a time (**Figure 1B, top**). Each trial started with a 100ms fixation, followed by a 12-element-long sequence of words/nonwords presented at the rate of 450ms. At the end of each trial, a drawing of a hand appeared for 400ms, and participants had to press a button every time they saw this drawing. The button press task was included to help the participants stay alert and focused during the task. Each trial ended with a blank screen presented for 100ms. The task followed a blocked design. The total duration of each trial was 6 seconds, and 3 trials were included in each block. Each run included 16 18-second blocks (8 per condition). Additionally, 5 fixation blocks of 14 seconds were included in each run, for a total run duration of 358 seconds. Each participant completed two runs. The *Sentences > Nonwords* contrast targets brain areas that support language processing (including mental operations related to word retrieval, syntactic structure building, and semantic composition), while excluding perceptual processing (see **Figure OSF1**—available at https://osf.io/7594t/—for evidence that this contrast includes areas sensitive to sub-lexical/phonological processing; also: Bozic et al., 2015; Regev et al., 2022).

It is important to note that although this particular version of the language localizer uses written stimuli, the core frontal and temporal areas activated by the written *Sentences>Nonwords* contrast overlap almost perfectly with areas identified with auditory contrasts between language processing and a perceptually similar condition, such as backwards speech or acoustically degraded speech (e.g., Fedorenko et al., 2010; Scott et al., 2017; Malik-Moraleda, Ayyash et al., 2022). Moreover, the network of areas that emerges for the *Sentences>Nonwords* contrast (and similar contrasts) closely matches one of the intrinsic brain networks recoverable from individual patterns of functional connectivity, as measured during naturalistic conditions

such as resting state (e.g., Braga et al., 2020; Du et al., 2024) or even during task paradigms (e.g., Du et al., 2025; Shain & Fedorenko, 2025).

Auditory language localizer task

Participants passively listened to intact and acoustically degraded passages from *Alice in Wonderland* (Carroll, 1865; **Figure 1B, middle**). The degraded versions of each passage were created using a procedure introduced in Scott et al. (2017): intact passages were low-pass filtered (passband 500 Hz) and mixed with a time-varying noise track derived from the same clip (generated by randomizing consecutive 0.02-s segments, amplitude-modulated by the intact signal, and low-pass filtered to attenuate very high frequencies). The noise level was adjusted until the resulting clips were unintelligible. Although the degraded clips largely preserve prosodic patterns, the key computations carried out by the language areas (related to word recognition, syntactic structure building, and semantic composition; e.g., Shain, Kean et al., 2024) cannot be performed because the content is not comprehensible, leading to a much weaker response compared to the response to the intact clips.

Each trial consisted of an 18-second recording. Each run included 12 18-second trials (4 per condition; in addition to the intact and degraded conditions, the experiment included a third condition not included in any analyses here—recordings in an unfamiliar foreign language). Additionally, 3 fixation blocks of 12 seconds were included in each run, for a total run duration of 252 seconds. Each participant completed three runs. Similar to the reading-based localizer, the *Intact* > *Degraded* contrast targets brain areas that support language processing.

A demanding non-linguistic (spatial working memory) task

Participants performed a spatial working memory task consisting of easier and harder trials (e.g., Fedorenko et al., 2013; Assem et al., 2020). Each trial started with a 500ms fixation. Then, a 3 x 4 grid appeared in the center of the screen, and a series of locations were flashed in blue at the rate of 1 second per flash (**Figure 1B, bottom**). In the easy condition, one location flashed at a time (four locations in total); in the hard condition, two locations flashed at a time (eight in total). Participants were instructed to keep track of the locations. Each trial ended with a two-choice forced choice task: two grids with different sets of locations were shown side by side, and participants had to choose the grid that shows the locations just presented. They were told whether they chose correctly (green checkmark) or not (red cross). The task followed a blocked design. The total duration of each trial was 8 seconds, and 4 trials were included in each block. Each run included 10 32-second blocks (5 per condition). Additionally, 6 fixation blocks of 16 seconds were included in each run, for a total run duration of 448 seconds. Each participant completed 2 runs. The *Hard* > *Easy* contrast targets brain areas that support demanding cognitive tasks, including mental operations associated with working memory, attention, and cognitive control.

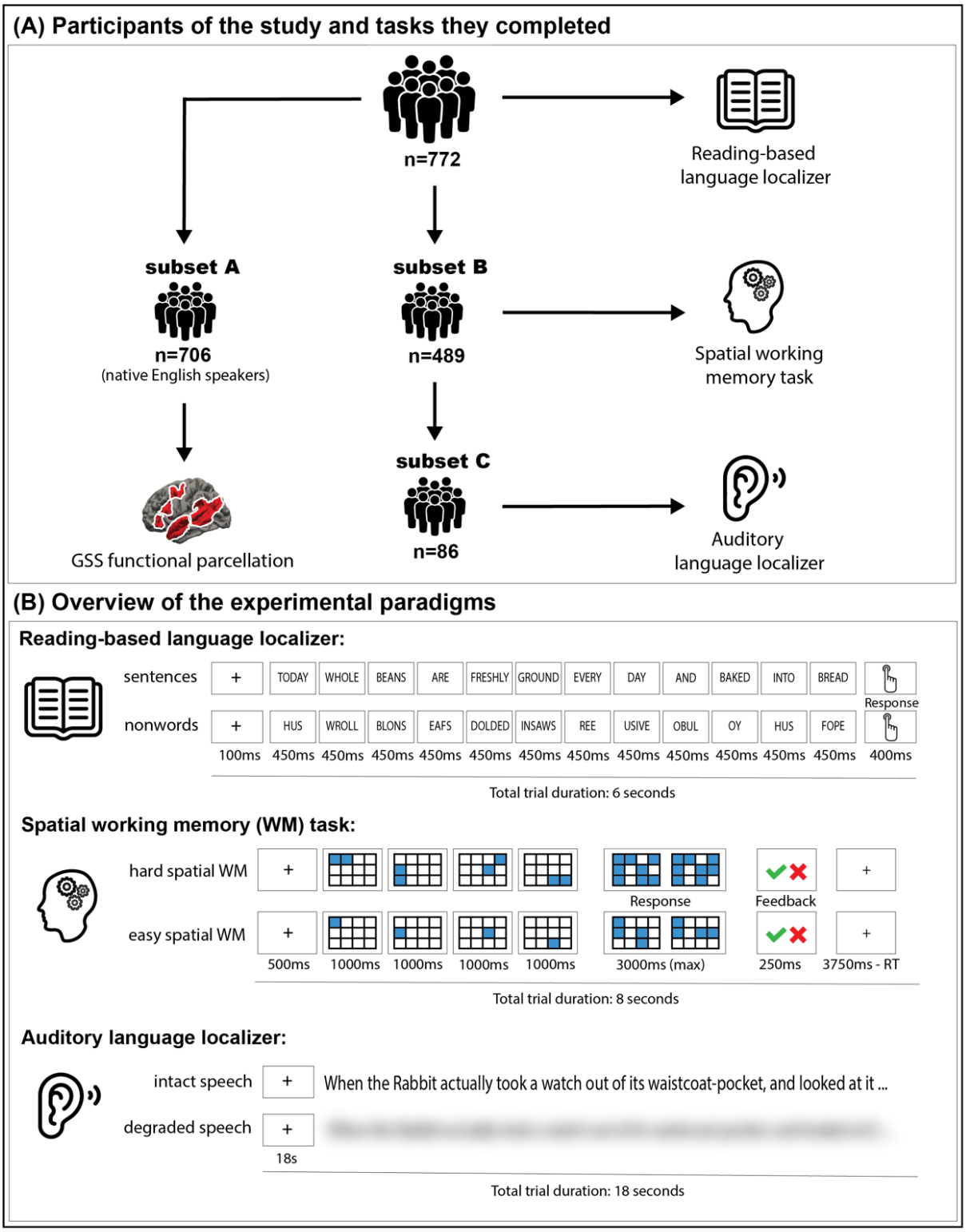

Figure 1. Overview of the dataset and paradigms. A) The structure of the dataset. A total of n=772 participants were included. All 772 completed the reading-based language localizer. A subset of n=706 (*subset A*) were native speakers of English and were used in the GSS analysis for creating the extended set of parcels. A subset of n=489 participants (*subset B*) completed the spatial working memory task, and a subset of n=86 participants (*subset C*) completed the auditory language localizer. B) The paradigms. For each task, a sample trial in each condition is shown (see Tasks for a detailed description).

fMRI data acquisition

Structural and functional data collected in the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research at MIT. Structural and functional data were acquired using a Siemens Trio 3 Tesla scanner before or after a Prisma upgrade (n=508 and n=264, respectively) and a 32-channel head coil. High-resolution, whole-brain anatomical images were acquired using two different T1-weighted MPRAGE sequence (**Sequence 1**: n=508, 176 sagittal slices; voxel size 1 x 1 x 1 mm^3; TR = 2530 ms, TE = 3.48 ms, flip angle = 9°; **Sequence 2**: n=264, 208 sagittal slices; voxel size 0.9 x 0.9 x 0.9 mm^3; TR = 1800 ms, TE = 2.40 ms, flip angle = 9°). Functional T2*-weighted images were acquired using three different whole-brain echo-planar (EPI) pulse sequences with phase encoding direction A >> P (**Sequence 1, Siemens Trio**: n = 508, 31 near-axial slices, 2.1 × 2.1 × 4 mm voxels; TR = 2,000 ms; TE = 30 ms; flip angle = 90°; matrix size 96 x 96 mm; MB acceleration factor 0; **Sequence 2, Siemens Prisma**: n = 73, 52 near-axial slices, 2 × 2 × 2 mm isotropic voxels; TR = 2,000 ms; TE = 30 ms; flip angle = 90°; matrix size 104 x 104 mm; MB acceleration factor 2; **Sequence 3, Siemens Prisma**: n = 191, 72 near-axial slices, 2 × 2 × 2 mm isotropic voxels; TR = 2,000 ms; TE = 30 ms; flip angle = 90°; matrix size 104 x 104 mm; MB acceleration factor 3). The first 10 s of each run were excluded to allow for steady state magnetization.

fMRI data preprocessing

fMRI data were analyzed using SPM12 (release 7487), CONN EvLab module (release 19b), and other custom MATLAB scripts. Each participant's functional and structural data were converted from DICOM to NIFTI format. All functional scans were coregistered and resampled using 4th degree B-spline interpolation to the first scan of the first session (Friston et al., 1995). Potential outlier scans were identified from the resulting subject-motion estimates as well as from BOLD signal indicators using default thresholds in CONN preprocessing pipeline (5 standard deviations above the mean in global BOLD signal change, or framewise displacement values above 0.9 mm; Nieto-Castañón, 2020). Functional and structural data were independently normalized into a common space (the Montreal Neurological Institute [MNI] template; IXI549Space) using SPM12 unified segmentation and normalization procedure (Ashburner & Friston, 2005) with a reference functional image computed as the mean functional data after realignment across all timepoints omitting outlier scans. The output data were resampled to a common bounding box between MNI-space coordinates (-90, -126, -72) and (90, 90, 108), using 2mm isotropic voxels and 4th degree B-spline interpolation for the functional data, and 1mm isotropic voxels and trilinear interpolation for the structural data. Last, the functional data were smoothed spatially using spatial convolution with a 4 mm FWHM Gaussian kernel (following a reviewer's request, we additionally provide unsmoothed activation maps on OSF (https://osf.io/7594t/); see also **Figure OSF2** for evidence that individual-level activation maps are highly similar regardless of the preprocessing approach: with vs. without normalization to the common space, and between our volumetric SPM-based pipeline and a surface-based Freesurfer pipeline).

fMRI data first-level modeling

Responses in individual voxels were estimated using a General Linear Model (GLM) in which each experimental condition was modeled with a boxcar function convolved with the canonical hemodynamic response function (HRF) (fixation was modeled implicitly). Temporal autocorrelations in the BOLD signal timeseries were accounted for by a combination of high-pass filtering with a 128 seconds cutoff, and whitening using an AR(0.2) model (first-order autoregressive model linearized around the coefficient a=0.2) to approximate the observed covariance of the functional data in the context of Restricted Maximum Likelihood estimation (ReML). In addition to experimental condition effects, the GLM design included first-order temporal derivatives for each condition (included to model variability in the HRF delays), as well as nuisance regressors to control for the effect of slow linear drifts, subject-specific motion parameters (6 parameters), and potential outlier scans (identified during preprocessing as described above) on the BOLD signal.

Group-level parcels used to constrain the definition of individual fROIs

The group-level functional parcels were created using the group-constrained subject-specific approach (GSS; Fedorenko et al., 2010) as implemented in the spm_ss toolbox (available for download from: https://www.evlab.mit.edu/resources). This approach helps determine which parts of the activation landscape are consistent across participants and establish correspondences across individuals. To do so, thresholded binarized individual maps for a particular contrast of interest are overlaid in a common brain space to create a probabilistic activation overlap map (or 'atlas'). Subsequently, this map is divided into parcels using a watershed image parcellation algorithm (see Julian et al., 2012 for the application of this approach to the ventral visual areas).

Here, we used a set of 706 individual maps (the subset of the native English speakers; **Figure 1A**) and selected the 20% most language-responsive voxels across the brain in each individual. This relatively liberal individual-level threshold increases the likelihood of capturing all language-responsive areas, including those with weaker and less spatially consistent responses. It is worth noting that the average *t*-value across participants for these voxels was still quite high: 2.02 (SD=0.6), with the average minimum *t*-value of 0.92, and the average maximum *t*-value of 14.17). In the resulting overlap map, the value of each voxel, divided by the total number of participants, represents the proportion of participants for whom that voxel belongs to the top 20% of most language-responsive voxels (the highest value in this map is ~0.78, in line with Lipkin et al., 2022). The overlap map was thresholded to remove voxels with values below 0.1 (i.e., voxels present in fewer than 10%, or 71 of the 706 participants) and smoothed with a 6 mm full-width half-maximum Gaussian kernel, to avoid over-parcellation. Finally, an image parcellation (watershed) algorithm was run to identify the main "hills" in the activation landscape (Meyer, 1991).

In our dataset, this algorithm identified 27 parcels (of size 150 voxels or larger): 11 in the left hemisphere, 8 in the right hemisphere, 3 midline bilateral parcels, and 5 parcels in the

cerebellum (**Figures 2** and **3** and **Table 1**; see **Figure OSF3** for evidence that the probabilistic overlap map and the parcels are similar if the individual maps are thresholded more conservatively—at top 10% of most language-responsive voxels instead of top 20%). Six additional parcels (smaller than 150 voxels in size) were excluded from the main analysis, but we report their profiles for completeness in **Figure S1**.

Because the BOLD signal is weaker in brain areas that are further away from the receiver coil, we complemented our functional parcels with a set of 7 bilateral subcortical parcels from the Harvard-Oxford Subcortical atlas (Desikan et al., 2006), to the exclusion of the brain stem, ventricles, and white matter.

Estimation of response to the three tasks in the individual language fROIs

Within each parcel, we defined a subject-specific functional ROI (fROI) by selecting the 10% of voxels showing the strongest response to the language localizer contrast (*Sentences > Nonwords*) (see **Figures 2** and **3** for sample fROIs; see **Figure OSF4** for evidence that the results are near-identical if the fROIs are defined as voxels that pass a fixed whole-brain significance threshold of $p<.001$). Subsequently, we estimated the responses in these fROIs to the conditions of the auditory language localizer (*Intact speech* and *Degraded speech*) and the spatial working memory task (*Hard spatial WM* and *Easy spatial WM*). To estimate the responses to the *Sentences* and *Nonwords* conditions, we relied on an across-runs cross-validation procedure, which ensures that different subsets of the data are used for fROI definition and response estimation (e.g., Kriegeskorte et al., 2010).

Statistical analyses

The data were analyzed using linear mixed-effect models as implemented in the lmerTest package in R (Kuznetsova et al., 2017). We ran a separate analysis for each fROI comparing the critical and baseline conditions in each task: the reading-based language localizer, the auditory language localizer, and the spatial working memory task using the following formula:

$$\text{EffectSize} \sim 1 + \text{Condition} * \text{Task} + (1 \mid \text{Participant})$$

The categorical predictor of *Condition* was deviation-coded using the treatment contrast which specified the control condition of each task as the baseline (the reading-based language localizer: *Sentences* = 1, *Nonwords* = 0; the auditory language localizer: *Intact speech* = 1, *Degraded speech* = 0; the spatial WM task: *Hard WM* = 1, *Easy WM* = 0). The categorical predictor of *Task* was deviation-coded using the treatment contrast and specifying the reading language localizer as the baseline.

Subsequently, we ran three post-hoc pairwise comparisons to test (i) differences between the critical and control conditions for each task (i.e., *Sentences > Nonwords*, *Intact speech > Degraded speech*, and *Hard WM > Easy WM*); (ii) whether in each task the responses to the critical and control conditions significantly differed from 0 (e.g., *Sentences > 0* and *Nonwords > 0* in the reading language task); and (iii) whether the responses to the critical condition significantly differed between a) each of the two language conditions (i.e., *Sentences > Intact*

speech) and b) the non-linguistic Hard WM condition (i.e., *Sentences* > *Hard WM*, and *Intact speech* > *Hard WM*). We do not apply a correction for multiple comparisons (regions) here because we are asking an independent question for each region (rather than asking whether the effect holds in *any* of the regions). Besides, the approach already has internal replicability "checks" built into it, including an across-runs cross-validation procedure for the reading-based contrast and generalization to the auditory language contrast. That said, most of the critical effects would survive a correction for the number of regions given the size of our sample (all GSS and subcortical regions and majority of regions evaluated within standard atlases).

It is important to note that although different subsets of participants completed the three tasks (all n=772 completed the reading-based language localizer; n=489 of the 772 additionally completed the spatial WM task; and n=86 of the 489 additionally completed the auditory language localizer; **Figure 1A**), linear mixed-effects models are well-suited to handle such data imbalances (Little & Rubin, 2019). In particular, this approach efficiently estimates the effects of interest while controlling for data sparsity (assuming greater uncertainty where data are sparse by implementing partial pooling strategy, which "shrinks" the intercepts of participants with less data points towards the population mean more than those of participants who completed all tasks) and is robust under the assumption that the missing data are missing at random, as is reasonable here. In addition, this approach has the advantage of using all available data compared to analyzing each task separately (Snijders & Bosker, 2011; Brown, 2021). Nevertheless, to establish the statistical robustness of our results, we additionally report the results for the subset of the n=86 participants who completed all three tasks and thus where the effects could be estimated with a full within-participant design (**Figure OSF5**).

Estimation of the language system size

To estimate the size of the language system, we first used individual-level *t*-maps to extract the number of voxels that pass a significance threshold of p<.001, uncorrected, for *Sentences* > *Nonwords* contrast within the union of all core and extended parcels. To convert this value into cubic centimeters, we multiplied the number of voxels by their size (i.e., $8mm^3$, given that all functional data were resampled to the 2x2x2mm resolution). It is important to note that MNI template tends to be larger than a typical brain. For example, in our data the amount of grey matter in the MNI-normalized probability map was estimated to be, on average, $942cm^3$ (SD = $76cm^3$, min = $548cm^3$, max = $1112cm^3$), whereas in the native space, it was, on average, $748cm^3$ (SD = $74cm^3$, min = $483cm^3$, max = $998cm^3$). As a result, the absolute volume estimates (in cm^3) in the analysis space that we report may be overestimates of the native-space tissue volumes.

To estimate the proportion that the language system, including the cortical and cerebellar components but excluding the subcortical ones, occupies relative to all grey matter (cortical, subcortical, cerebellar, and brain stem), we used the anatomical segmentation images produced by SPM during preprocessing. In particular, we used the individual grey matter probability maps normalized to the MNI space (wc1* image), which matches the analysis space of our functional data. For each participant, we estimated the grey matter amount as the sum of voxelwise grey matter probabilities multiplied by voxel volume.

Comparison with Standard Cortical Atlases

To facilitate comparisons with other studies, where standard anatomical atlases are often used, we evaluated language responses in three cortical atlases: 1) the Desikan-Killiany-Tourville atlas (Klein & Tourville, 2012), 2) the Harvard-Oxford Cortical atlas (Desikan et al., 2006), and 3) the multi-modal Glasser atlas (Glasser et al., 2016) (see **Figure S5A** for information on the overlap between the functional and anatomical cortical parcels). The aim of these analyses is to evaluate the usefulness of standard atlases as group-constraints for subject-specific fROIs and not making any claims about the relationship of the functional language areas and macroanatomical landmarks.

To create a volume-based DKT parcellation, we ran the cortical reconstruction (using the recon-all Freesurfer standard pipeline) on the MNI template (IXI549Space), which was used in this study to normalize each participant's data. The resulting parcels (the aparc.DKTatlas+aseg.mgz file) were converted to the NIFTI format and smoothed with a 2 mm FWHM Gaussian kernel. The smoothing step is necessary to slightly increase the spatial extent of the parcels (because they are aligned to the cortical surface, they are relatively thin) as well as accommodate potential small differences between the surface space and the volumetric space. The Harvard-Oxford atlas is volume-based, so no transformation was needed. Finally, for the Glasser atlas, the MNI-based volumetric version was downloaded from https://figshare.com/articles/dataset/HCP-MMP1_0_projected_on_MNI2009a_GM_volumetric_in_NIfTI_format/3501911 and resliced to the 2mm IXI549Space in SPM.

For each atlas, we performed two analyses: an analysis where the whole parcel is used as the ROI (an identical set of voxels is used across participants), and an analysis where within each parcel, we define a subject-specific fROI, as we did in the main analysis (by selecting the 10% of voxels showing the strongest response to the language localizer contrast). Statistical analyses were performed on the responses extracted from these two kinds of ROIs (anatomical group-level ROIs and individual fROIs) using the same approach of linear mixed-effects models, as described above.

Code accessibility

All data, including 3D brain contrast maps (both smoothed and unsmoothed), parcels, and data produced by the functional localization analysis toolbox (spm_ss), as well as the code necessary to replicate all analyses and plots presented in this paper are available at: https://osf.io/7594t/.

Results

We considered a region to contain a fROI **responsive to language** if the following conditions were met: 1) the response to both of the language contrasts (*Sentences* > *Nonwords* and *Intact* > *Degraded speech*) is positive and significant ($p<.05$), and 2) the response to both of the critical conditions (i.e., *Sentences* for the reading language localizer and *Intact speech* for the auditory language localizer) falls significantly above the fixation baseline. We further

considered a region to contain a fROI **selective for language** if, in addition to the abovementioned conditions, the critical language conditions (*Sentences* or *Intact speech*) elicited a reliably greater response than the critical condition for the non-linguistic task (i.e., *Hard WM*). Note that here, we only evaluate selectivity for language relative to one task: a non-linguistic working memory task. This approach allows us to determine whether a brain region responds to any cognitive task, or whether the response reflects processing required for reading and listening to language but not when performing the working memory task. This limited definition of selectivity already rules out some of the language-responsive regions as supporting computations that are not specific to language (see Results), but it is of course possible that some of the regions found here to be selective will later be found to not be selective relative to other non-linguistic tasks (such as processing music; or processing non-verbal communicative signals; or processing non-linguistic meanings).

The extended set of language-selective regions

Twenty-two of the 27 parcels identified by the whole-brain GSS analysis (Methods) contained fROIs that responded to language across modalities, with the exception of the three parcels in the occipital cortex, the parcel in the left precuneus / posterior ACC, and the parcel in the left parahippocampal gyrus (the fROIs in these parcels either did not show a significant response to one of the language contrasts or responded at or below baseline to one or both of the critical language conditions). Of the 22 language-responsive regions, 20 (including 5 core LH regions and their RH homotopic regions) showed selective responses to language with little or no response to the demanding non-linguistic task; both of the exceptions were located in the cerebellum (**Figures 2** and **3**, **Table 1**; see **Table S1** for the details of the statistical tests; for completeness, see **Figure S1** for the profiles of the regions excluded due to small parcel size). Note that in addition to these language-selective areas, most of these parcels contain areas that respond strongly to the non-linguistic task and not to the language tasks (see **Figure S3**), in line with past evidence that distinct functional areas often lie near each other within the same anatomical structure (e.g., Fedorenko et al., 2012; Braga et al., 2020; Du et al., 2024)—a finding that reinforces the importance of functionally defining the relevant region within each individual.

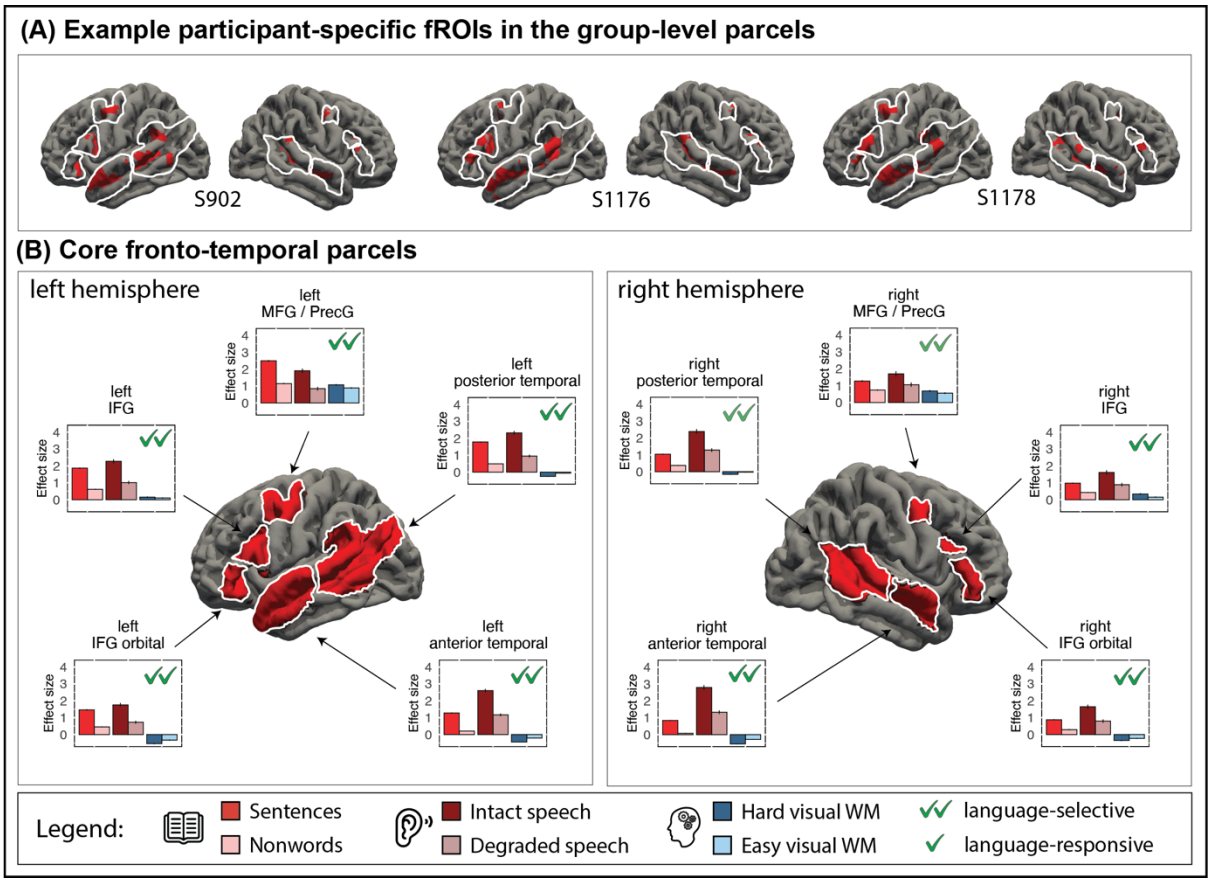

Figure 2. Language-responsive regions in the core language network and their functional profiles. A) Example participant-specific fROIs created using top 10% of most active voxels (shown in red) within the group-level parcels (white outlines) for the core language regions. B) The five core left-hemisphere parcels (left sub-panel) and their right-hemisphere homotopes (right sub-panel). The figure shows the surface projections of the parcels (all the analyses were performed in the volume space, as described in Methods; the surface projections were created using FreeSurfer). For each region (defined based on individual activation maps, as described in Methods and illustrated in panel A), we show responses to the reading-based and auditory language localizers (red bars; darker bars correspond to the critical conditions) and the non-linguistic spatial working memory task (blue bars; the darker bar corresponds to the harder condition). For the reading-based localizer, the responses were estimated using independent runs of the data. Parcels that contain language-responsive fROIs have a green checkmark in the upper right-hand corner; the ones where the fROIs are additionally language-selective have two green checkmarks (all fROIs in this set are language-selective). In all plots, the bars correspond to the mean of raw data and whiskers represent SEM of raw means.

In addition to the core left lateral frontal and lateral temporal regions and their homotopic regions—which replicate prior findings (Mahowald & Fedorenko, 2016; see Fedorenko et al., 2024 for a review) and which we report here for completeness—the extended language-selective network (**Figure 3**) included regions in the bilateral temporal pole, in the left 'basal temporal language area' (BTLA; Lüders et al., 1986; see Li et al., 2024 and Salvo et al., 2025 for concordant evidence using a similar functional localization approach to the one used here), in the left precuneus, in three regions in the medial frontal cortex (spanning different sections of the SFG and the frontal pole), and in three cerebellar regions (left Crus II / VIIb, right Crus I / II / VIIb, and right VIIb / VIIIa; see Casto et al., 2025a for concordant evidence).

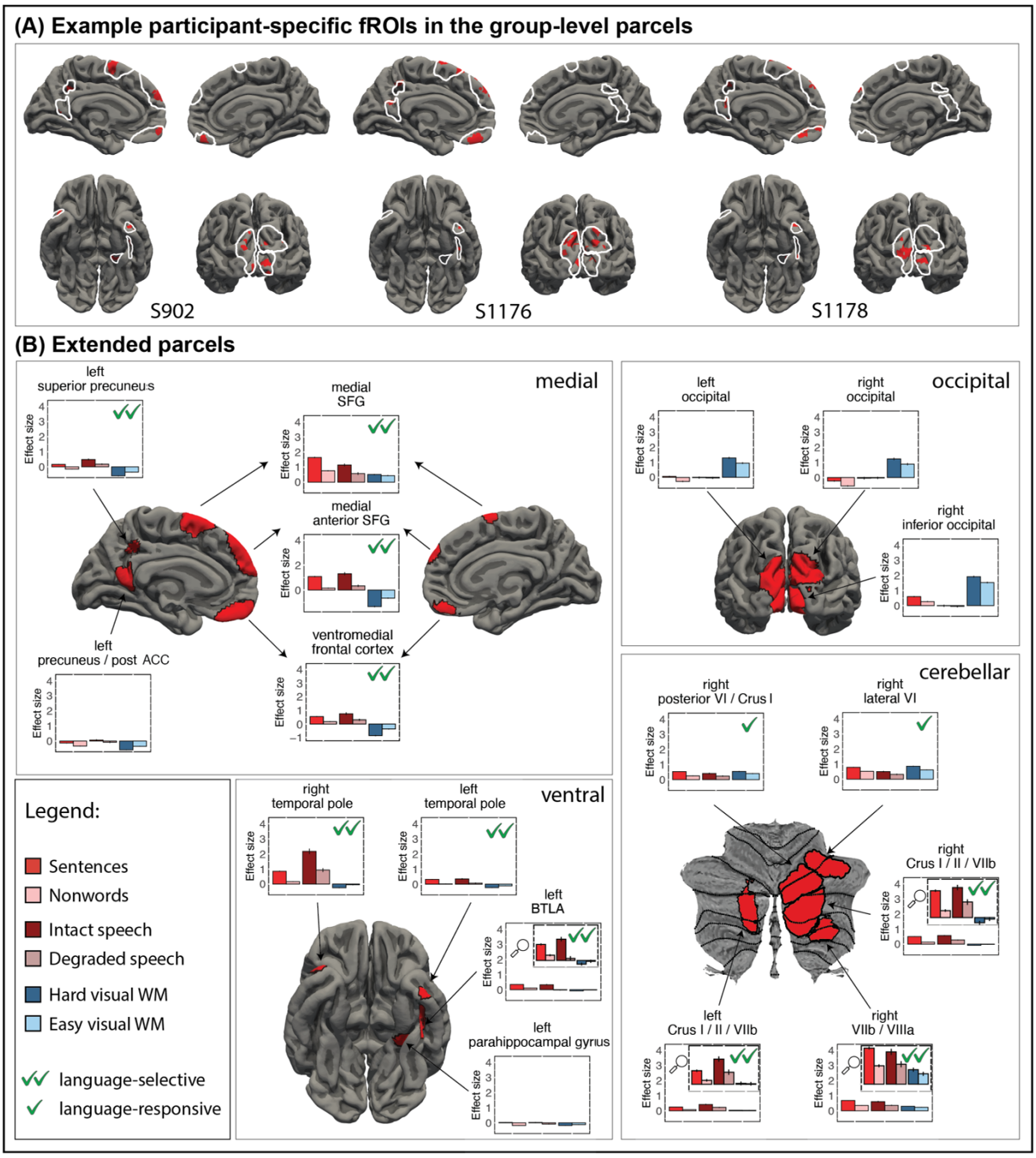

Figure 3. Language-responsive regions in the extended language network and their functional profiles. A) Example participant-specific fROIs created using top 10% of most active voxels (shown in red) within the group-level parcels (white outlines) for the non-canonical language regions. Note that due to image orientation in some cases, the fROIs are not visible in the surface projection (statistical maps for all subjects are available on OSF: https://osf.io/7594t/). B) The extended set of parcels (as identified by the GSS procedure; Methods). Separate sub-panels show the medial, occipital, ventral, and cerebellar parcels. The figure shows the surface projections of the cortical and cerebellar parcels (all the analyses were performed in the volume space, as described in Methods; the surface projections were created using FreeSurfer for the cortical and SUIT for the cerebellar parcels and are only used for visualization). For each region (defined based on individual activation maps, as described in Methods and illustrated in panel A), we show responses to the reading-based and auditory language localizers (red bars; darker bars correspond to the critical conditions) and the non-linguistic spatial working memory task (blue bars; the darker bar corresponds to the harder condition). For the reading-based localizer, the responses were estimated using independent runs of the data. Parcels that contain language-responsive fROIs have a green checkmark in

the upper right-hand corner (n=12 of the 17 fROIs satisfy this criterion); the ones where the fROIs are additionally language-selective have two green checkmarks (n=10 fROIs). We used the same range of values for the y-axis across all regions to highlight differences in response magnitudes, but for the language-selective regions with low responses, we include insets, marked with a magnifying glass, where we restrict the y-axis to make the profiles easier to see.

It is important to note that even with these additional regions, language-selective cortex takes up a relatively small amount of grey matter volume. Although the parcels are, by design, large (to capture inter-individual variability in the precise locations of functional areas), the individual fROIs constitute only a small fraction of each parcel. For example, voxels that show a significant ($p<.001$) *Sentences>Nonwords* effect across all our 20 parcels (10 core and 10 non-canonical) add up to 22.1 cm^3, on average (SD=12.8 cm^3), which corresponds to only ~3.5% of the grey matter volume (the volume estimate is similar when taking top 10% of language-responsive voxels in each parcel and language selectivity declines steeply beyond this threshold, see **OSF6**). An average of 15.2 cm^3 (SD=8.2 cm^3) is taken up by the five core left-hemisphere fROIs (defined within the following parcels: AntTemp, PostTemp, IFGorb, IFG, and MFG), and 3.6 cm^3 (SD=3.7 cm^3)—by their right-hemisphere homotopes.

In addition to the cortical language areas, 4 of the 14 parcels that we examined in the Harvard-Oxford Subcortical atlas contained language-selective fROIs: bilateral hippocampi and bilateral amygdalae (**Figure 4, Table 1**). One additional parcel (the left thalamus) contained a language-responsive, but non-selective fROI (see **Table S1** for the details of the statistical tests).

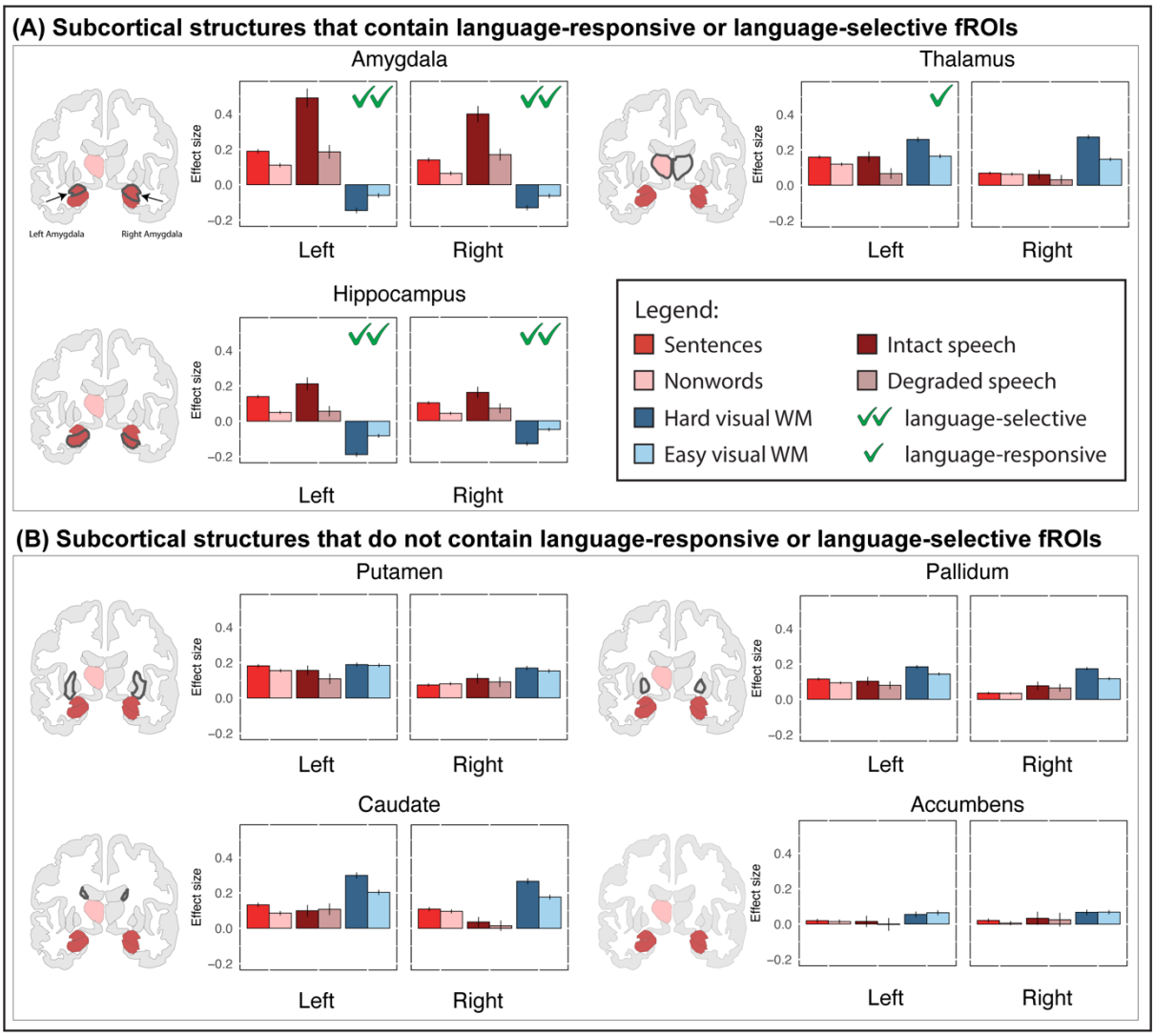

Figure 4. Language responsive regions within the parcels in the Harvard-Oxford Subcortical atlas. For each region (defined based on individual activation maps; Methods), we show responses to the reading-based and auditory language localizers (red bars; darker bars correspond to the critical conditions) and the non-linguistic spatial working memory task (blue bars; the darker bar corresponds to the harder condition). For the reading-based localizer, the responses were estimated using independent runs of the data. A) Parcels that contain language-responsive fROIs have a green checkmark; the ones where the fROIs are additionally language-selective have two green checkmarks. B) The parcels that do not contain language-responsive fROIs.

The full set of 27 cortical parcels, in both the MNI and the fsaverage spaces, is available on OSF (https://osf.io/7594t/; note that the fsaverage version was created by converting the volumetric set to the fsaverage space, not based on statistical maps generated in a surface space). Additionally, for the 20 parcels containing language-selective fROIs, we created a left-right symmetrical version, where the left-hemisphere cortical parcels were mirrored onto the right hemisphere (the medial parcels were first split along the brain midline (x = 45)), and the right-hemisphere cerebellar parcels were mirrored onto the left hemisphere. This symmetrical set (also available on OSF at https://osf.io/7594t/ and at https://evlab.mit.edu/parcels) may be useful for defining the language fROIs in atypically lateralized (right-lateralized or bilateral) participants, or for directly comparing the profiles of the left- and right-hemisphere language regions. We also provide a left-right symmetrical version that uses—for the core areas—the

parcels derived from a smaller set of participants and used in much past work (e.g., Diachek, Blank, Siegelman et al., 2020; Ivanova et al., 2020; Shain et al., 2020; Schrimpf et al., 2021; Malik-Moraleda, Ayyash, et al., 2022).

Table 1. Language regions. The 27 parcels that emerged in the whole-brain GSS analysis (Methods) and the 14 subcortical parcels from Harvard-Oxford Subcortical atlas (Desikan et al., 2006). The parcel sizes are reported in cm^3 (referring to volumes in the *analysis space*, see Methods) and the number of voxels (2x2x2 mm resolution); individual fROIs are 10% of the parcel size (Methods). The remaining two columns mark the parcels that contain language-responsive and language-selective fROIs (see Results for definitions).

Hemisphere	Name	Size (cm^3)	Size (n voxels)	Language-responsive	Language-selective
Core frontal and temporal parcels in the left and right hemispheres identified using GSS parcellation (included for completeness, as evidence of replicating past work)					
left	IFG orbital	12,40	1550	✓	✓
left	IFG	16,71	2089	✓	✓
left	MFG / PrecG	18,79	2349	✓	✓
left	anterior temporal	32,52	4065	✓	✓
left	posterior temporal	52,99	6624	✓	✓
right	IFG orbital	7,13	891	✓	✓
right	IFG	5,89	736	✓	✓
right	MFG / PrecG	5,90	737	✓	✓
right	anterior temporal	12,30	1538	✓	✓
right	posterior temporal	25,01	3126	✓	✓
Extended parcels identified using GSS parcellation					
left	temporal pole	1,64	205	✓	✓
left	BTLA	1,37	171	✓	✓
left	parahippocampal gyrus	1,27	159		
left	superior precuneus	2,42	302	✓	✓
left	precuneus / post ACC	4,30	538		
left	occipital	21,83	2729		
right	temporal pole	6,30	788	✓	✓
right	occipital	7,50	937		
right	inferior occipital	13,29	1661		
bilateral	ventromedial frontal cortex	6,01	751	✓	✓
bilateral	medial anterior SFG	13,96	1745	✓	✓
bilateral	medial SFG	13,85	1731	✓	✓
left	Crus I / II / VIIb	3,00	375	✓	✓
right	Crus I / II / VIIb	9,20	1150	✓	✓
right	VIIb / VIIIa	3,60	450	✓	✓
right	posterior VI / Crus I	3,92	490	✓	
right	lateral VI	3,05	381	✓	

Subcortical regions from Harvard-Oxford Subcortical Atlas					
left	accumbens	0,72	90		
left	amygdala	2,46	308	✓	✓
left	caudate	3,94	493		
left	hippocampus	5,51	689	✓	✓
left	pallidum	2,17	271		
left	putamen	6,44	805		
left	thalamus	10,47	1309	✓	
right	accumbens	0,68	85		
right	amygdala	2,90	363	✓	✓
right	caudate	4,14	517		
right	hippocampus	5,69	711	✓	✓
right	pallidum	2,14	267		
right	putamen	6,39	799		
right	thalamus	10,18	1272		

Language responses in three standard brain atlases

To facilitate comparisons with other work, we comprehensively evaluated language responses in three cortical atlases that vary in the granularity of their parcellation. These analyses serve two goals. First, they help identify anatomical areas within which reliable and selective responses to language can be observed within individuals. And second, they highlight the critical importance of individual-level functional localization.

When using the functional localization approach (i.e., selecting—for each individual separately—the top 10% of most active voxels based on their language localizer map), a number of areas emerge as containing language-selective fROIs (**Figure 5**). Full functional profiles of the atlas parcels are presented on **Figure S6** (see **Table OSF2** for details of statistical tests in all regions in the three atlases). In the DKT atlas, 20 of the 72 parcels contain language-selective fROIs: 13 in the left hemisphere and 7 in the right hemisphere. In the Harvard-Oxford Cortical atlas, 42 of the 96 parcels contain language-selective fROIs: 26 in the left hemisphere and 16 in the right hemisphere. Finally, in the Glasser multimodal atlas, 95 of the 380 parcels contain language-selective fROIs: 64 in the left hemisphere and 31 in the right hemisphere (see **Table 2** for the full list of language-responsive and language-selective regions). Across all three atlases, the topography of the parcels containing language-selective fROIs resembles the topography of the functional parcels from the GSS (see **Figure S5A** for information on the overlap between standard atlas parcels and functional language parcels), spanning bilateral frontal and temporal areas, as well as pre-central frontal and inferior parietal areas in the left hemisphere. However, in the most granular, Glasser atlas, 9 of the parcels containing language-selective fROIs showed no or minimal spatial overlap with any of the GSS parcels (see **Figure S2** for the functional profiles of these regions).

Table 2. List of parcels that contain language-responsive and language-selective fROIs in the three standard cortical atlases: the DKT, Harvard-Oxford Cortical, and Glasser atlases. In the Glasser atlas, parcels marked

with a degree sign (º) lie at the intersection of two anatomical lobes and parcels marked with a cross lie at the intersection of three anatomical lobes and are thus listed in more than one lobe.

Parcels in the DKT atlas that contain language-selective* fROIs

Lobe	Left hemisphere	Right hemisphere
Frontal	Caudal Middle Frontal, Medial Orbitofrontal, Pars Opercularis, Pars Orbitalis, Pars Triangularis, Precentral, Superior Frontal, Lateral Orbitofrontal	Medial Orbitofrontal, Pars Opercularis, Pars Triangularis, Precentral, Superior Frontal
Parietal	Supramarginal, Postcentral	
Temporal	Inferior Temporal, Middle Temporal, Superior Temporal	Middle Temporal, Superior Temporal

Parcels in the Harvard-Oxford Cortical atlas that contain language-selective fROIs

Lobe	Left hemisphere	Right hemisphere
Frontal	Frontal Pole, Sup. Frontal Gyrus, Mid. Frontal Gyrus, Inf. Frontal Gyrus Pars Triangularis, Inf. Frontal Gyrus Pars Opercularis, Precentral Gyrus, Frontal Medial Cortex, Juxtapositional Lobule Cortex, Frontal Orbital Cortex, Frontal Operculum Cortex	Inf. Frontal Gyrus Pars Triangularis, Inf. Frontal Gyrus Pars Opercularis, Frontal Medial Cortex, Frontal Orbital Cortex, Juxtapositional Lobule Cortex
Parietal	Supramarginal Gyrus Post., Angular Gyrus, Central Opercular Cortex, Parietal Operculum Cortex	Supramarginal Gyrus Post., Angular Gyrus
Temporal	Temporal Pole, Sup. Temporal Gyrus Ant., Sup. Temporal Gyrus Post., Mid. Temporal Gyrus Ant., Mid. Temporal Gyrus Post., Mid. Temporal Gyrus Temporooccipital, Inf. Temporal Gyrus Ant., Inf. Temporal Gyrus Post., Temporal Fusiform Cortex Ant., Temporal Fusiform Cortex Post., Planum Polare, Planum Temporale	Temporal Pole, Sup. Temporal Gyrus Ant., Sup. Temporal Gyrus Post., Mid. Temporal Gyrus Ant., Mid. Temporal Gyrus Post., Mid. Temporal Gyrus Temporooccipital, Inf. Temporal Gyrus Ant., Planum Temporale, Temporal Fusiform Cortex Ant.
	Additional parcels that contain language-responsive (but not language-selective) fROIs	
Frontal		Precentral Gyrus
Parietal	Supramarginal Gyrus Ant.	
Temporal	Inf. Temporal Gyrus Temporooccipital	
Occipital	Lat. Occipital Cortex Inf.	

Parcels in the Glasser atlas that contain language-selective fROIs

Lobe	Left hemisphere	Right hemisphere
Frontal	FEF, PEF, 4, 55B, SFL, 8BM, 10R, 47M, 9M, 9P, 44, 45, 47L, 6R, 8AV, 8BL, IFJA, IFSP, IFSA, 9A, 10V, 10d,	55B, SFL, 47M, 9M, 8BL, 44, 45, IFJA, IFSP, IFSA, 10V, OFC

	47S, FOP5, S32°, FOP4°, OP4°, OP2-3†, FOP1	
Parietal	3b, PFCM, PF, OP4°, 31PD°, PGI†, PSL°, OP2-3†	3B
Occipital	PGI†, PHT°, TPOJ3	PHT°, TOPJ3
Temporal	FOP4°, PEEC°, PGI†, PSL°, STV, RI, TA2, STGA, PBELT, A5, STSDA, STSDP, STSVP, TGD, TE1A, TE1P, TF, TE2P, TPOJ1, TPOJ2, TGV, A4, STSVA, TE1M, PI, PIR°, PHT°, OP2-3†	STV, TA2, STGA, A5, STSDA, STSDP, STSVP, TGD, TE1A, TE1P, TPOJ1, TPOJ2, TGV, A4, STSVA, TE1M, PHT°
Limbic**	S32°, 31PV, PEEC°, 31PD°, PIR°, H, pOFC	
	Additional parcels that contain language-responsive (but not language-selective) fROIs	
Frontal	SCEF, 6D, 6V, IFJP, P9-46V, I6-8, 8C	6r, 6V, FEF, IFJP, PEF, SCEF
Parietal	1	1
Temporal	PHA3	TF
Occipital	MST, FST	

* All DKT language-responsive regions are also language-selective.

** Lobe segmentation in the Glasser parcellation follows PALS_B12 atlas (Van Essen, 2005; https://surfer.nmr.mgh.harvard.edu/fswiki/PALS_B12) template (using FreeSurfer);

° denotes regions straddling two lobes;

† denotes regions straddling three lobes.

In stark contrast, when using the whole parcels as ROIs (taking all voxels within an ROI, with no participant-specific functional masking), many fewer areas emerge as language-selective, and in the ones that do, the degree of selectivity is substantially lower compared to the corresponding fROIs (see the histograms in **Figure 5B** for a quantitative comparison). For example, in the DKT atlas, which is the least granular (with ROIs corresponding to entire cortical gyri in many cases), only 11 regions satisfy the selectivity criterion when the entire regions are used as ROIs (cf. 20 regions when using individual-level fROIs). The selectivity of these regions (the difference between the response to *Sentences* vs. to the *Hard spatial WM* condition) is low (2-3 times lower than when using fROIs). Further, some regions, especially in the frontal cortex, show mixed selectivity (responses to both the language contrasts and the spatial WM contrast), when in reality, little or no overlap exists between these contrasts within any given individual (**Figure 5B**).

The picture is similar for the other two atlases: fewer regions emerge as language-selective, and in the ones that do, selectivity is severely underestimated. For example, in the Harvard-Oxford atlas, which has an intermediate level of granularity, in the two most selective group-level ROIs, the difference between the *Sentences* condition and the *Hard spatial WM* condition is 1.39 (left posterior superior temporal gyrus) and 1.03 (left anterior superior temporal gyrus);

the corresponding effect sizes in the individually-defined fROIs within these regions are 2.80 and 1.97. Similarly, even in the most granular atlas (Glasser et al., 2016), the selectivity of the individual fROIs is substantially higher (average across all selective fROIs: 1.02, SD = 0.65, range 0.05-2.67; compared to the average of the group-level ROIs (average: 0.50, SD = 0.39, range = -0.37 – 1.33; **Figure 5A**).

These differences in selectivity can be straightforwardly explained by i) inter-individual variability in the precise locations of language fROIs within the larger, anatomical regions (see **Figure S5B** for examples), and ii) the presence of domain-general areas that respond robustly to demanding cognitive tasks in the vicinity of the language areas (**Figures S3** and **S4**). In particular, a set of voxels that is the same across participants in a common template space is bound to include—for any given individual—some non-localizer-responsive voxels, leading to a lower response to the *Sentences* condition, and it may further include voxels that belong to the nearby Multiple Demand network (Fedorenko & Blank, 2020; Braga et al., 2020), leading to a higher response to the working memory task conditions (see e.g., Nieto-Castañón & Fedorenko, 2012 for discussion).

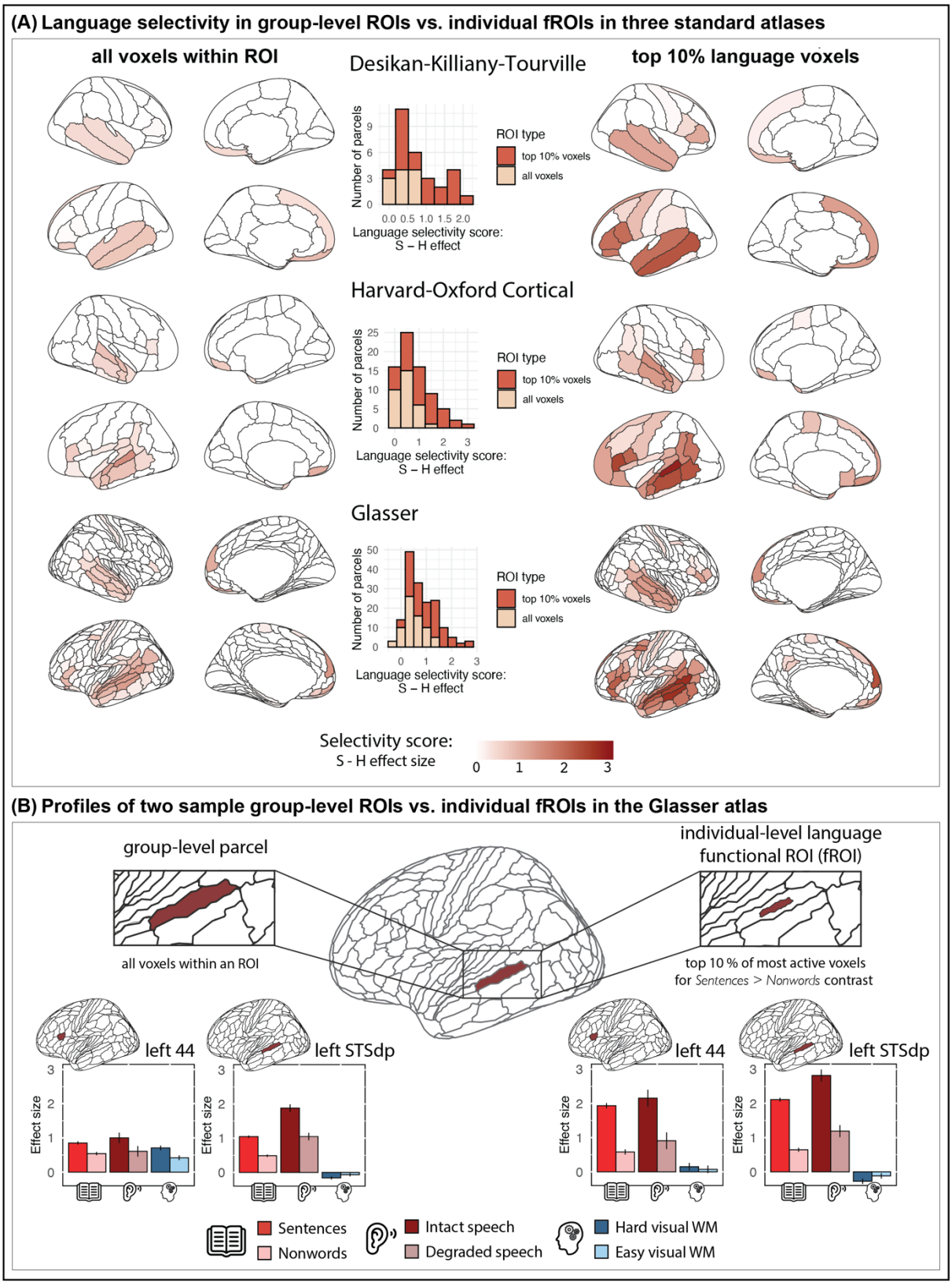

Figure 5. Consequences of using entire group-level anatomical ROIs vs. individually-defined functional ROIs for evaluating language selectivity. A) Across three standard cortical parcellations (rows; top: DKT: Klein & Tourville, 2012; middle: HOC: Desikan et al., 2006; bottom: Glasser: Glasser et al., 2016), we show the selectivity scores (difference between the response to the *Sentences* condition and the *Hard spatial WM* condition) for all regions. Higher selectivity scores are shown with darker red hues (see bottom of the panel for the legend). In the left column, the selectivity scores are based on using group-level anatomical ROIs (all voxels in a parcel

are used); in the right column, the selectivity scores are based on using individually-defined functional ROIs (top 10% of voxels are selected based on individual activation maps; the responses to the Sentence condition are estimated in a left-out run of the data; Methods). For each atlas and each ROI definition (group-level vs. individual-level), four views of the brain are shown: right lateral and medial and left lateral and medial. (For a figure showing, for each atlas, effect sizes for the three individual contrasts—*Sentences>Nonwords*, *Intact>Degraded speech*, and *Hard>Easy spatial WM*—see **Figure S6**). In the middle column, we show—for each atlas—a quantitative comparison of the distributions of selectivity scores from the group ROIs vs. individual fROIs. B) For two sample parcels in the Glasser atlas (Glasser et al., 2016)—left 44 and left STSdp (selected based on high overlap with the probabilistic atlas of the language network; Lipkin et al., 2022)—we show the response profiles for the group-level anatomical ROIs (left) and individual-level fROIs (right).

Discussion

Many brain areas outside the left fronto-temporal network have been previously implicated in language processing (**Table OSF1**). Here, we systematically searched for language-responsive areas across the brain. To do so, we a) used a paradigm that isolates language processing from perceptual, motor, and task-related cognitive demands (Fedorenko et al., 2010), b) included a large dataset (n=772 participants), c) examined language responses across modalities, and d) evaluated language selectivity relative to a demanding cognitive task. In addition to the core left hemisphere areas and their homotopes, we identified 17 areas (7 cortical and 10 subcortical, including 5 cerebellar) that respond to auditory and written language, with 14 of these areas showing selectivity for language relative to a demanding non-linguistic task.

The contributions of non-canonical language areas to language and cognition.

Establishing that a brain area responds to language across modalities and not during a non-linguistic task ensures that its activity is not driven by perceptual or general cognitive demands. Still, the precise contributions of these language-responsive and language-selective areas remain to be discovered. Past studies may hold some clues, but relating our results to prior work is challenging: we cannot be certain that the areas we found here correspond to those identified in prior work, because macroanatomy is a poor guide to function in the association cortex (Frost & Goebel, 2012; Tahmasebi et al., 2012; also, **Figure S5B**). Adopting the functional localization approach has led to several robust and replicable findings about the core language network, including its selectivity for language and its role in linguistic computations (see Fedorenko et al., 2024 for a review). The extended language areas we identified here constitute new targets for similar systematic investigations, which should i) evaluate the selectivity of these areas relative to other functions that bear similarities to language, ii) probe sensitivity of these areas to linguistic demands associated with lexical and combinatorial processing (in controlled and naturalistic paradigms), and iii) directly compare these areas' contributions to those of the core language areas.

The non-selective language-responsive areas may still meaningfully contribute to language processing. We have identified three such areas: two in the right cerebellum (Casto et al., 2025a) and one in the left thalamus. Higher spatial resolution approaches may eventually reveal that distinct neural sub-populations create the appearance of mixed selectivity. However, if the mixed selectivity is real, it may suggest that these regions play an integratory role, perhaps

combining inputs from the core language areas and other networks. Indeed, both the cerebellum and the thalamus have been hypothesized to serve as information integrators (Barbas et al., 2013; Theofanopoulou & Boeckx, 2016; Wolpert et al., 1998).

Activation during a "language task" may not always reflect linguistic processing.

Some past studies have reported neural activity during language tasks in brain areas that do not show language-responsiveness in our study. However, activation during "language tasks" (i.e., tasks that use verbal materials) may not always reflect linguistic processing. Many language paradigms tax both linguistic and general cognitive demands, so a response in such paradigms could reflect the latter. For example, Thibault et al. (2021, 2025) have argued for overlap between syntactic processing and tool use in the basal ganglia. However, both the syntactic and tool-use tasks in those studies require cognitive processes beyond syntactic and tool-use-related computations. We show that the language paradigms that are not confounded with task demands do not reliably engage any basal ganglia structures (**Figure 4B**), which suggests that the syntactic complexity manipulation in Thibault et al.'s studies may be engaging these areas due to general cognitive demands and not by virtue of syntactic computations (see **Figure S4** for evidence that basal ganglia areas respond strongly to task demands).

Similarly, Seydell-Greenwald et al. (2023) report responses to auditory language tasks in the primary visual cortex (V1) of sighted individuals. We do not replicate this finding:, the weak response in occipital areas to the reading-based contrast (mostly due to greater deactivation during nonword reading) does not generalize to the auditory modality, which rules out the possibility that V1 implements linguistic computations (cf. evidence of V1 supporting language processing in blind individuals; Bedny et al., 2011; Lane et al., 2015; Pant et al., 2020; Czarnecka et al., 2025). The effects observed in Seydell-Greenwald et al., (2023) might be explained by the fact that their contrasts are confounded with general cognitive demands: one contrast entails listening to sentences and answering questions about them vs. listening to reversed speech and responding to a beep (see Ozernov-Palchik, O'Brien et al., in press for additional discussion of difficulty confounds in this paradigm), and the other contrast compares listening to words and pressing a button for catch trials (thus requiring attention to the task beyond the lexical processing demands) vs. a low-level baseline. Attentional modulation of V1 is well-established (Somers et al., 1999), so the observed responses plausibly reflect attention-related engagement.

To argue that a brain area supports linguistic computations, one needs to ensure that it responds to language across tasks and modalities but also that the language task is not confounded by difficulty or attentional demands. Previous work has established that the core language areas are relatively insensitive to task demands, showing similarly strong responses during passive comprehension and paradigms that include tasks (Fedorenko et al., 2010; Diachek, Blank, Siegelman et al., 2020; Gao et al., 2025). Here, we extend these findings to several non-canonical language areas but also question some past findings of "language" responses that have alternative explanations.

A small and well-defined subset of the brain implements language processing.

Despite ample evidence for functional specialization in the brains of humans (Kanwisher, 2010) and non-human animals (Tsao et al., 2006), some continue to argue against the idea of stable structure in the brain, emphasizing the distributed, dynamic, and interactive nature of cognitive processes, including lnguage (e.g., Pessoa, 2022; Forkel & Hagoort, 2024; Drijvers et al., 2025). Deep engagement with this debate is beyond the scope of this article, but two points are worth clarifying. First, the fact that many areas—sometimes in distant parts of the brain—are engaged by language comprehension does not imply that the "entire brain" supports this function. Although the extended language network spans almost every major component of the brain, within each component, language regions occupy a small fraction of brain tissue. The entire language network constitutes <3.5% of the grey matter volume in individual brains. Second, linguistic inputs can unquestionably engage many brain regions beyond those that specifically support language processing: vivid descriptions of faces or scenes can engage category-selective visual areas, a story about a misunderstanding can engage the Theory of Mind network, and a horror story can engage the amygdala (see Casto, et al., 2025b for discussion). However, all these brain regions can also be engaged by non-linguistic inputs. The ability of a brain region to be engaged by language does not make it a "language region" any more than its ability to be engaged by visual inputs makes it a "visual region". Furthermore, the fact that the language network needs to interact with other brain areas does not undermine its functional distinctness from those areas and its special role in language processing. As long as different components within the language network interact more strongly with one another than with other networks—for which ample evidence exists (Blank et al., 2014; Braga et al., 2020; Du et al., 2024, 2025; Shain & Fedorenko, 2025)—the language network and other cognitive networks can be treated as meaningfully distinct objects of study (Simon, 1962).

Parcels derived from functional data have advantages over anatomical parcels for constraining individual fROIs.

For completeness and ease of comparison with past studies, we explored the possibility of using standard anatomical atlases to constrain individual fROIs (rather than the parcels derived from a group-level representation of brain activity; Fedorenko et al., 2010; Julian et al., 2012). Examining individual activation maps (or maps derived from functional connectivity patterns: Braga et al., 2020; Du et al., 2024, 2025; Shain & Fedorenko, 2025) against standardized brain parcellations reveals two issues. First, individual topographies often do not align with the boundaries of the atlas areas: a contiguous functional region may get broken up by a boundary, or—for finer-grained atlases—may get assigned to different atlas areas across individuals because of inter-individual topographic variability (see **Figure S5B** for examples). For studies focusing on a particular functional network, we therefore recommend functional parcels over anatomical/multimodal atlases. Critically, however, whichever approach is chosen, the regions of analysis should be defined functionally within individuals (using localizers or functional-connectivity-based parcellations) to avoid conflating nearby networks at the group level (**Figure 4**; see **Figure S3** for evidence that most language parcels include regions of the Multiple Demand network). And second, using large, anatomical ROIs creates an illusion that language processing engages large chunks of the cortex, while in reality these extensive regions contain (sometimes multiple) small language-responsive regions (see **Figures 2A and 3A** for

illustration of the size of individual-level fROIs relative to the parcels and **Figure 5** and **Figure S5A** for comparison of the extent of language responsive areas in coarser- vs. finer-grained parcellations).

In conclusion, we have comprehensively searched for language-selective areas across the brain and identified several new components of the extended language network. This work lays the groundwork for future investigations of these new components, including their similarities to and differences from the core language areas.

References

Ardila, A. (2020). Supplementary motor area aphasia revisited. *Journal of Neurolinguistics*, *54*, 100888. https://doi.org/10.1016/j.jneuroling.2020.100888

Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *NeuroImage*, *26*(3), 839–851. https://doi.org/10.1016/j.neuroimage.2005.02.018

Assem, M., Blank, I. A., Mineroff, Z., Ademoğlu, A., & Fedorenko, E. (2020). Activity in the fronto-parietal multiple-demand network is robustly associated with individual differences in working memory and fluid intelligence. *Cortex*, *131*, 1–16. https://doi.org/10.1016/j.cortex.2020.06.013

Barbas, H., García-Cabezas, M. Á., & Zikopoulos, B. (2013). Frontal-thalamic circuits associated with language. *Brain and Language*, *126*(1), 49–61. https://doi.org/10.1016/j.bandl.2012.10.001

Bedny, M., Pascual-Leone, A., Dodell-Feder, D., Fedorenko, E., & Saxe, R. (2011). Language processing in the occipital cortex of congenitally blind adults. *Proceedings of the National Academy of Sciences*, *108*(11), 4429–4434. https://doi.org/10.1073/pnas.1014818108

Bédos Ulvin, L., Jonas, J., Brissart, H., Colnat-Coulbois, S., Thiriaux, A., Vignal, J.-P., & Maillard, L. (2017). Intracerebral stimulation of left and right ventral temporal cortex during object naming. *Brain and Language*, *175*, 71–76. https://doi.org/10.1016/j.bandl.2017.09.003

Blank, I., Kanwisher, N., & Fedorenko, E. (2014). A functional dissociation between language and multiple-demand systems revealed in patterns of BOLD signal fluctuations. *Journal of Neurophysiology*, *112*(5), 1105–1118. https://doi.org/10.1152/jn.00884.2013

Bohsali, A., & Crosson, B. (2016). The Basal Ganglia and Language: A Tale of Two Loops. In J.-J. Soghomonian (Ed.), *The Basal Ganglia: Novel Perspectives on Motor and Cognitive Functions* (pp. 217–242). Springer International Publishing. https://doi.org/10.1007/978-3-319-42743-0_10

Booth, J. R., Wood, L., Lu, D., Houk, J. C., & Bitan, T. (2007). The role of the basal ganglia and cerebellum in language processing. *Brain Research*, *1133*, 136–144. https://doi.org/10.1016/j.brainres.2006.11.074

Bozic, M., Fonteneau, E., Su, L., & Marslen-Wilson, W. D. (2015). Grammatical analysis as a distributed neurobiological function. *Human Brain Mapping*, *36*(3), 1190–1201. https://doi.org/10.1002/hbm.22696

Braga, R. M., DiNicola, L. M., Becker, H. C., & Buckner, R. L. (2020). Situating the left-lateralized language network in the broader organization of multiple specialized large-scale distributed networks. *Journal of Neurophysiology*, *124*(5), 1415–1448. https://doi.org/10.1152/jn.00753.2019

Brett, M., Johnsrude, I. S., & Owen, A. M. (2002). The problem of functional localization in the human brain. *Nature Reviews Neuroscience*, *3*(3), 243–249. https://doi.org/10.1038/nrn756

Brown, V. A. (2021). An Introduction to Linear Mixed-Effects Modeling in R. *Advances in Methods and Practices in Psychological Science*, *4*(1), 2515245920960351. https://doi.org/10.1177/2515245920960351

Carroll, L. (1865). *Alice's adventures in Wonderland.*

Casto, C., Poliak, M., Tuckute, G., Small, H., Sherlock, P., Wolna, A., Lipkin, B., D'Mello, A. M., & Fedorenko, E. (2025a). *The cerebellar components of the human language network* (p. 2025.04.14.645351). bioRxiv. https://doi.org/10.1101/2025.04.14.645351

Casto, C., Ivanova, A., Fedorenko, E., & Kanwisher, N. (2025b). *What does it mean to understand language?* (arXiv:2511.19757). arXiv. https://doi.org/10.48550/arXiv.2511.19757

Covington, N. V., & Duff, M. C. (2016). Expanding the Language Network: Direct Contributions from the Hippocampus. *Trends in Cognitive Sciences*, *20*(12), 869–870. https://doi.org/10.1016/j.tics.2016.10.006

Crosson, B., Benefield, H., Cato, M. A., Sadek, J. R., Moore, A. B., Wierenga, C. E., Gopinath, K., Soltysik, D., Bauer, R. M., Auerbach, E. J., Gökçay, D., Leonard, C. M., & Briggs, R. W. (2003). Left and right basal ganglia and frontal activity during language generation:

Contributions to lexical, semantic, and phonological processes. *Journal of the International Neuropsychological Society*, *9*(7), 1061–1077. https://doi.org/10.1017/S135561770397010X

Czarnecka, M., Hryniewiecka, K., Wolna, A., Baumbach, C., Beck, J., Vadlamudi, J., Collignon, O., Jednoróg, K., Szwed, M., & Stroh, A.-L. (2025). Association between cortical thickness and functional response to linguistic processing in the occipital cortex of early blind individuals. *Cerebral Cortex*, *35*(11), bhaf317. https://doi.org/10.1093/cercor/bhaf317

Deniz, F., Nunez-Elizalde, A. O., Huth, A. G., & Gallant, J. L. (2019). The Representation of Semantic Information Across Human Cerebral Cortex During Listening Versus Reading Is Invariant to Stimulus Modality. *Journal of Neuroscience*, *39*(39), 7722–7736. https://doi.org/10.1523/JNEUROSCI.0675-19.2019

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021

Diachek, E., Blank, I., Siegelman, M., Affourtit, J., & Fedorenko, E. (2020). The Domain-General Multiple Demand (MD) Network Does Not Support Core Aspects of Language Comprehension: A Large-Scale fMRI Investigation. *The Journal of Neuroscience*, *40*(23), 4536–4550. https://doi.org/10.1523/JNEUROSCI.2036-19.2020

Dijksterhuis, D. E., Self, M. W., Possel, J. K., Peters, J. C., van Straaten, E. C. W., Idema, S., Baaijen, J. C., van der Salm, S. M. A., Aarnoutse, E. J., van Klink, N. C. E., van Eijsden, P., Hanslmayr, S., Chelvarajah, R., Roux, F., Kolibius, L. D., Sawlani, V., Rollings, D. T., Dehaene, S., & Roelfsema, P. R. (2024). Pronouns reactivate conceptual representations in human hippocampal neurons. *Science*, *385*(6716), 1478–1484. https://doi.org/10.1126/science.adr2813

Dikker, S., Rabagliati, H., Farmer, T. A., & Pylkkänen, L. (2010). Early Occipital Sensitivity to Syntactic Category Is Based on Form Typicality. *Psychological Science*, *21*(5), 629–634. https://doi.org/10.1177/0956797610367751

Dikker, S., Rabagliati, H., & Pylkkänen, L. (2009). Sensitivity to syntax in visual cortex. *Cognition*, *110*(3), 293–321. https://doi.org/10.1016/j.cognition.2008.09.008

Drijvers, L., Small, S. L., & Skipper, J. I. (2025). Language is widely distributed throughout the brain. *Nature Reviews Neuroscience*, 1–1. https://doi.org/10.1038/s41583-024-00903-0

Du, J., DiNicola, L. M., Angeli, P. A., Saadon-Grosman, N., Sun, W., Kaiser, S., Ladopoulou, J., Xue, A., Yeo, B. T. T., Eldaief, M. C., & Buckner, R. L. (2024). Organization of the human cerebral cortex estimated within individuals: Networks, global topography, and function. *Journal of Neurophysiology*. https://doi.org/10.1152/jn.00308.2023

Du, J., Tripathi, V., Elliott, M. L., Ladopoulou, J., Sun, W., Eldaief, M. C., & Buckner, R. L. (2025). Within-individual precision mapping of brain networks exclusively using task data. *Neuron*, *0*(0). https://doi.org/10.1016/j.neuron.2025.08.029

Duff, M. C., & Brown-Schmidt, S. (2012). The hippocampus and the flexible use and processing of language. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00069

Duff, M. C., & Brown-Schmidt, S. (2017). Hippocampal Contributions to Language Use and Processing. In D. E. Hannula & M. C. Duff (Eds.), *The Hippocampus from Cells to Systems: Structure, Connectivity, and Functional Contributions to Memory and Flexible Cognition* (pp. 503–536). Springer International Publishing. https://doi.org/10.1007/978-3-319-50406-3_16

Fedorenko, E. (2021). The early origins and the growing popularity of the individual-subject analytic approach in human neuroscience. *Current Opinion in Behavioral Sciences*, *40*, 105–112. https://doi.org/10.1016/j.cobeha.2021.02.023

Fedorenko, E., & Blank, I. A. (2020). Broca's Area Is Not a Natural Kind. *Trends in Cognitive Sciences*, *24*(4), 270–284. https://doi.org/10.1016/j.tics.2020.01.001

Fedorenko, E., Duncan, J., & Kanwisher, N. (2012). Language-Selective and Domain-General Regions Lie Side by Side within Broca's Area. *Current Biology*, *22*(21), 2059–2062. https://doi.org/10.1016/j.cub.2012.09.011

Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences*, *110*(41), 16616–16621. https://doi.org/10.1073/pnas.1315235110

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. https://doi.org/10.1152/jn.00032.2010

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. https://doi.org/10.1038/s41583-024-00802-4

Fedorenko, E., & Kanwisher, N. (2009). Neuroimaging of Language: Why Hasn't a Clearer Picture Emerged? *Language and Linguistics Compass*, *3*(4), 839–865. https://doi.org/10.1111/j.1749-818X.2009.00143.x

Ferstl, E. C., Neumann, J., Bogler, C., & Von Cramon, D. Y. (2008). The extended language network: a meta-analysis of neuroimaging studies on text comprehension. *Human brain mapping*, *29*(5), 581-593. https://doi.org/10.1002/hbm.20422

Forkel, S. J., & Hagoort, P. (2024). Redefining language networks: Connectivity beyond localised regions. *Brain Structure and Function*, *229*(9), 2073–2078. https://doi.org/10.1007/s00429-024-02859-4

Friston, Karl. J., Ashburner, J., Frith, C. D., Poline, J.-B., Heather, J. D., & Frackowiak, R. S. J. (1995). Spatial registration and normalization of images. *Human Brain Mapping*, *3*(3), 165–189. https://doi.org/10.1002/hbm.460030303

Frost, M. A., & Goebel, R. (2012). Measuring structural–functional correspondence: Spatial variability of specialised brain regions after macro-anatomical alignment. *NeuroImage*, *59*(2), 1369–1381. https://doi.org/10.1016/j.neuroimage.2011.08.035

Gao, R., Cheung, C., Siegelman, M., Pongos, A. L., Kean, H. H., Tanner, A., Fedorenko, E., Ivanova, A. A. (2025) The language network responds robustly to sentences across diverse tasks. *bioRxiv.*

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C. F., Jenkinson, M., Smith, S. M., & Van Essen, D. C. (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, *536*(7615), 171–178. https://doi.org/10.1038/nature18933

Hasson, U., Nusbaum, H. C., & Small, S. L. (2007). Brain Networks Subserving the Extraction of Sentence Information and Its Encoding to Memory. *Cerebral Cortex*, *17*(12), 2899–2913. https://doi.org/10.1093/cercor/bhm016

Hertrich, I., Dietrich, S., & Ackermann, H. (2016). The role of the supplementary motor area for speech and language processing. *Neuroscience & Biobehavioral Reviews*, *68*, 602–610. https://doi.org/10.1016/j.neubiorev.2016.06.030

Holland, R., & Lambon Ralph, M. A. (2010). The Anterior Temporal Lobe Semantic Hub Is a Part of the Language Neural Network: Selective Disruption of Irregular Past Tense Verbs by rTMS. *Cerebral Cortex*, *20*(12), 2771–2775. https://doi.org/10.1093/cercor/bhq020

Ivanova, A. A., Srikant, S., Sueoka, Y., Kean, H. H., Dhamala, R., O'Reilly, U.-M., Bers, M. U., & Fedorenko, E. (2020). Comprehension of computer code relies primarily on domain-general executive brain regions. *eLife*, *9*, e58906. https://doi.org/10.7554/eLife.58906

Julian, J. B., Fedorenko, E., Webster, J., & Kanwisher, N. (2012). An algorithmic method for functionally defining regions of interest in the ventral visual pathway. *NeuroImage*, *60*(4), 2357–2364. https://doi.org/10.1016/j.neuroimage.2012.02.055

Kanwisher, N. (2010). Functional specificity in the human brain: A window into the functional architecture of the mind. *Proceedings of the National Academy of Sciences*, *107*(25), 11163–11170. https://doi.org/10.1073/pnas.1005062107

Kim, J. S., Kanjlia, S., Merabet, L. B., & Bedny, M. (2017). Development of the Visual Word Form Area Requires Visual Experience: Evidence from Blind Braille Readers. *Journal of Neuroscience*, *37*(47), 11495–11504. https://doi.org/10.1523/JNEUROSCI.0997-17.2017

Klein, A., & Tourville, J. (2012). 101 Labeled Brain Images and a Consistent Human Cortical Labeling Protocol. *Frontiers in Neuroscience*, *6*. https://doi.org/10.3389/fnins.2012.00171

Klostermann, F., & Ehlen, F. (2013). Functional roles of the thalamus for language capacities. *Frontiers in Systems Neuroscience*, *7*. https://doi.org/10.3389/fnsys.2013.00032

Kotz, S. A. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain and Language*, *109*(2), 68–74. https://doi.org/10.1016/j.bandl.2008.06.002

Krauss, G. L., Fisher, R., Plate, C., Hart, J., Uematsu, S., Gordon, B., & Lesser, R. P. (1996). Cognitive Effects of Resecting Basal Temporal Language Areas. *Epilepsia*, *37*(5), 476–483. https://doi.org/10.1111/j.1528-1157.1996.tb00594.x

Kriegeskorte, N., Lindquist, M. A., Nichols, T. E., Poldrack, R. A., & Vul, E. (2010). Everything you never wanted to know about circular analysis, but were afraid to ask. *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, *30*(9), 1551–1557. https://doi.org/10.1038/jcbfm.2010.86

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(1), Article 1. https://doi.org/10.18637/jss.v082.i13

Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, *18*(1), 42–55. https://doi.org/10.1038/nrn.2016.150

Lane, C., Kanjlia, S., Omaki, A., & Bedny, M. (2015). "Visual" Cortex of Congenitally Blind Adults Responds to Syntactic Movement. *Journal of Neuroscience*, *35*(37), 12859–12868. https://doi.org/10.1523/JNEUROSCI.1256-15.2015

LeBel, A., & D'Mello, A. M. (2023). A seat at the (language) table: Incorporating the cerebellum into frameworks for language processing. *Current Opinion in Behavioral Sciences*, *53*, 101310. https://doi.org/10.1016/j.cobeha.2023.101310

Li, J., Hiersche, K. J., & Saygin, Z. M. (2024). Demystifying visual word form area visual and nonvisual response properties with precision fMRI. *iScience*, *27*(12), 111481. https://doi.org/10.1016/j.isci.2024.111481

Lindell, A. K. (2006). In Your Right Mind: Right Hemisphere Contributions to Language Processing and Production. *Neuropsychology Review*, *16*(3), 131–148. https://doi.org/10.1007/s11065-006-9011-9

Lipkin, B., Tuckute, G., Affourtit, J., Small, H., Mineroff, Z., Kean, H., Jouravlev, O., Rakocevic, L., Pritchett, B., Siegelman, M., Hoeflin, C., Pongos, A., Blank, I. A., Struhl, M. K., Ivanova, A., Shannon, S., Sathe, A., Hoffmann, M., Nieto-Castañón, A., & Fedorenko, E. (2022). Probabilistic atlas for the language network based on precision fMRI data from >800 individuals. *Scientific Data*, *9*(1), 529. https://doi.org/10.1038/s41597-022-01645-3

Little, R. J. A., & Rubin, D. B. (2019). *Statistical Analysis with Missing Data*. John Wiley & Sons.

Lüders, H., Lesser, R. P., Hahn, J., Dinner, D. S., Morris, H., Resor, S., & Harrison, M. (1986). Basal temporal language area demonstrated by electrical stimulation. *Neurology*, *36*(4), 505–510. https://doi.org/10.1212/wnl.36.4.505

Mahowald, K., & Fedorenko, E. (2016). Reliable individual-level neural markers of high-level language processing: A necessary precursor for relating neural variability to behavioral and genetic variability. *NeuroImage*, *139*, 74–93. https://doi.org/10.1016/j.neuroimage.2016.05.073

Malik-Moraleda, S., Ayyash, D., Gallée, J., Affourtit, J., Hoffmann, M., Mineroff, Z., Jouravlev, O., & Fedorenko, E. (2022). An investigation across 45 languages and 12 language families reveals a universal language network. *Nature Neuroscience*, *25*(8), 1014–1019. https://doi.org/10.1038/s41593-022-01114-5

Mariën, P., Ackermann, H., Adamaszek, M., Barwood, C. H. S., Beaton, A., Desmond, J., De Witte, E., Fawcett, A. J., Hertrich, I., Küper, M., Leggio, M., Marvel, C., Molinari, M., Murdoch, B. E., Nicolson, R. I., Schmahmann, J. D., Stoodley, C. J., Thürling, M., Timmann, D., … Ziegler, W. (2014). Consensus Paper: Language and the Cerebellum: an Ongoing Enigma. *The Cerebellum*, *13*(3), 386–410. https://doi.org/10.1007/s12311-013-0540-5

Martin, K. C., Seydell-Greenwald, A., Berl, M. M., Gaillard, W. D., Turkeltaub, P. E., & Newport, E. L. (2022). A Weak Shadow of Early Life Language Processing Persists in the Right Hemisphere of the Mature Brain. *Neurobiology of Language*, *3*(3), 364–385. https://doi.org/10.1162/nol_a_00069

Meyer, F. (1991). Un algorithme optimal de lignes de partage des eaux. *8ieme congres Reconnaissance des Formes et Intelligence Artificielle*, Lyon, France

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., Holmes, C., Collins, L., Thompson, P., MacDonald, D., Iacoboni, M., Schormann, T., Amunts, K., Palomero-Gallagher, N., Geyer, S., … Mazoyer, B. (2001). A Four-Dimensional Probabilistic Atlas of the Human Brain. *Journal of the American Medical Informatics Association*, *8*(5), 401–430. https://doi.org/10.1136/jamia.2001.0080401

Michon, K. J., Khammash, D., Simmonite, M., Hamlin, A. M., & Polk, T. A. (2022). Person-specific and precision neuroimaging: Current methods and future directions. *NeuroImage*, *263*, 119589. https://doi.org/10.1016/j.neuroimage.2022.119589

Murdoch, B. E. (2010). The cerebellum and language: Historical perspective and review. *Cortex*, *46*(7), 858–868. https://doi.org/10.1016/j.cortex.2009.07.018

Nieto-Castanon, A. (2020). *Handbook of functional connectivity Magnetic Resonance Imaging methods in CONN*. Hilbert Press.

Nieto-Castañón, A., & Fedorenko, E. (2012). Subject-specific functional localizers increase sensitivity and functional resolution of multi-subject analyses. *NeuroImage*, *63*(3), 1646–1669. https://doi.org/10.1016/j.neuroimage.2012.06.065

Oberhuber, M., Parker Jones, ‘Ōiwi, Hope, T. M. H., Prejawa, S., Seghier, M. L., Green, D. W., & Price, C. J. (2013). Functionally distinct contributions of the anterior and posterior putamen during sublexical and lexical reading. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00787

Ozernov-Palchik, O., O’Brien, A. M., Lee, E. J., Richardson, H., Romeo, R., Poliak, M., Lipkin, B., Small, H., Capella, J., Nieto-Castañón, A., Saxe, R., Gabrieli, J. D. E., & Fedorenko, E. (2025). Precision fMRI reveals that the language network exhibits adult-like left-hemispheric lateralization by 4 years of age. *bioRxiv*, 2024.05.15.594172. https://doi.org/10.1101/2024.05.15.594172

Pant, R., Kanjlia, S., & Bedny, M. (2020). A sensitive period in the neural phenotype of language in blind individuals. *Developmental Cognitive Neuroscience*, *41*, 100744. https://doi.org/10.1016/j.dcn.2019.100744

Papagno, C., Comi, A., Riva, M., Bizzi, A., Vernice, M., Casarotti, A., Fava, E., & Bello, L. (2017). Mapping the brain network of the phonological loop. *Human Brain Mapping*, *38*(6), 3011–3024. https://doi.org/10.1002/hbm.23569

Peschke, C., Ziegler, W., Eisenberger, J., & Baumgaertner, A. (2012). Phonological manipulation between speech perception and production activates a parieto-frontal circuit. *NeuroImage*, *59*(1), 788–799. https://doi.org/10.1016/j.neuroimage.2011.07.025

Pessoa, L. (2022). *The Entangled Brain: How Perception, Cognition, and Emotion Are Woven Together*. MIT Press.

Piai, V., Anderson, K. L., Lin, J. J., Dewar, C., Parvizi, J., Dronkers, N. F., & Knight, R. T. (2016). Direct brain recordings reveal hippocampal rhythm underpinnings of language processing. *Proceedings of the National Academy of Sciences*, *113*(40), 11366–11371. https://doi.org/10.1073/pnas.1603312113

Regev, M., Honey, C. J., Simony, E., & Hasson, U. (2013). Selective and Invariant Neural Responses to Spoken and Written Narratives. *Journal of Neuroscience*, *33*(40), 15978–15988. https://doi.org/10.1523/JNEUROSCI.1580-13.2013

Regev, T. I., Casto, C., Hosseini, E. A., Adamek, M., Brunner, P., & Fedorenko, E. (2022). *Intracranial recordings reveal three distinct neural response patterns in the language network* (p. 2022.12.30.522216). bioRxiv. https://doi.org/10.1101/2022.12.30.522216

Reich, L., Szwed, M., Cohen, L., & Amedi, A. (2011). A Ventral Visual Stream Reading Center Independent of Visual Experience. *Current Biology*, *21*(5), 363–368. https://doi.org/10.1016/j.cub.2011.01.040

Salvo, J. J., Anderson, N. L., & Braga, R. M. (2025). Intrinsic functional connectivity delineates transmodal language functions. *Imaging Neuroscience*, *3*, IMAG.a.25. https://doi.org/10.1162/IMAG.a.25

Saxe, R., Brett, M., & Kanwisher, N. (2006). Divide and conquer: A defense of functional localizers. *NeuroImage*, *30*(4), 1088–1096. https://doi.org/10.1016/j.neuroimage.2005.12.062

Schrimpf, M., Blank, I. A., Tuckute, G., Kauf, C., Hosseini, E. A., Kanwisher, N., Tenenbaum, J. B., & Fedorenko, E. (2021). The neural architecture of language: Integrative modeling converges on predictive processing. *Proceedings of the National Academy of Sciences*, *118*(45), e2105646118. https://doi.org/10.1073/pnas.2105646118

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. https://doi.org/10.1080/17588928.2016.1201466

Seydell-Greenwald, A., Wang, X., Newport, E. L., Bi, Y., & Striem-Amit, E. (2023). Spoken language processing activates the primary visual cortex. *PLOS ONE*, *18*(8), e0289671. https://doi.org/10.1371/journal.pone.0289671

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. https://doi.org/10.1016/j.neuropsychologia.2019.107307

Shain, C., Kean, H., Casto, C., Lipkin, B., Affourtit, J., Siegelman, M., ... & Fedorenko, E. (2024). Distributed sensitivity to syntax and semantics throughout the language network. *Journal of Cognitive Neuroscience, 36*(7), 1427-1471. https://doi.org/10.1162/jocn_a_02164

Shain, C., & Fedorenko, E. (2025). *A language network in the individualized functional connectomes of over 1,000 human brains doing arbitrary tasks*. bioRxiv. https://doi.org/10.1101/2025.03.29.646067

Simon, H. A. (1962). The Architecture of Complexity. *Proceedings of the American Philosophical Society*, *106*(6), 467–482.

Snijders, T. A. B., & Bosker, R. (2011). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. 1–368.

Snyder, K. M., Forseth, K. J., Donos, C., Rollo, P. S., Fischer-Baum, S., Breier, J., & Tandon, N. (2023). Critical role of the ventral temporal lobe in naming. *Epilepsia*, *64*(5), 1200–1213. https://doi.org/10.1111/epi.17555

Somers, D. C., Dale, A. M., Seiffert, A. E., & Tootell, R. B. H. (1999). Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proceedings of the National Academy of Sciences*, *96*(4), 1663–1668. https://doi.org/10.1073/pnas.96.4.1663

Tahmasebi, A. M., Davis, M. H., Wild, C. J., Rodd, J. M., Hakyemez, H., Abolmaesumi, P., & Johnsrude, I. S. (2012). Is the Link between Anatomical Structure and Function Equally Strong at All Cognitive Levels of Processing? *Cerebral Cortex*, *22*(7), 1593–1603. https://doi.org/10.1093/cercor/bhr205

Theofanopoulou, C., & Boeckx, C. (2016). The central role of the thalamus in language and cognition*. In *Advances in Biolinguistics*. Routledge.

Thibault, S., Py, R., Gervasi, A. M., Salemme, R., Koun, E., Lövden, M., Boulenger, V., Roy, A. C., & Brozzoli, C. (2021). Tool use and language share syntactic processes and neural patterns in the basal ganglia. *Science*, *374*(6569), eabe0874. https://doi.org/10.1126/science.abe0874

Thibault, S., Py, R., Koun, E., Boulenger, V., Roy, A. C., & Brozzoli, C. (2025). *Reaching for a domain-general syntax: Processing linguistic structures and grasping an object with a tool share similar neural codes in the basal ganglia.* bioRxiv. https://doi.org/10.1101/2025.10.17.683107

Tsao, D. Y., Freiwald, W. A., Tootell, R. B. H., & Livingstone, M. S. (2006). A Cortical Region Consisting Entirely of Face-Selective Cells. *Science*, *311*(5761), 670–674. https://doi.org/10.1126/science.1119983

Turker, S., Kuhnke, P., Eickhoff, S. B., Caspers, S., & Hartwigsen, G. (2023). Cortical, subcortical, and cerebellar contributions to language processing: A meta-analytic review of 403 neuroimaging experiments. *Psychological Bulletin*, *149*(11–12), 699–723. https://doi.org/10.1037/bul0000403

Vagharchakian, L., Dehaene-Lambertz, G., Pallier, C., & Dehaene, S. (2012). A Temporal Bottleneck in the Language Comprehension Network. *Journal of Neuroscience*, *32*(26), 9089–9102. https://doi.org/10.1523/JNEUROSCI.5685-11.2012

Van Essen, D. C. (2005). A population-average, landmark-and surface-based (PALS) atlas of human cerebral cortex. *Neuroimage, 28*(3), 635-662. https://doi.org/10.1016/j.neuroimage.2005.06.058

Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., Mazoyer, B., & Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *NeuroImage*, *30*(4), 1414–1432. https://doi.org/10.1016/j.neuroimage.2005.11.002

Wahl, M., Marzinzik, F., Friederici, A. D., Hahne, A., Kupsch, A., Schneider, G.-H., Saddy, D., Curio, G., & Klostermann, F. (2008). The Human Thalamus Processes Syntactic and Semantic Language Violations. *Neuron*, *59*(5), 695–707. https://doi.org/10.1016/j.neuron.2008.07.011

Wolpert, D.M., Miall, R.C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences, 2(*9), 338–47. https://doi.org/10.1016/S1364-6613(98)01221-2