# HHS Public Access
Author manuscript
*Nat Neurosci*. Author manuscript; available in PMC 2023 August 10.

# An investigation across 45 languages and 12 language families reveals a universal language network

**Saima Malik-Moraleda**[1,2,3,*,^], **Dima Ayyash**[1,2,*], **Jeanne Gallée**[3], **Josef Affourtit**[1,2], **Malte Hoffmann**[4,5], **Zachary Mineroff**[1,2,6], **Olessia Jouravlev**[1,2,7], **Evelina Fedorenko**[1,2,3,^]

[1]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139

[2]McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA 02139

[3]Program in Speech and Hearing Bioscience and Technology, Harvard University, Boston, MA 02114

[4]Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, 02129

[5]Department of Radiology, Harvard Medical School, Boston, MA, 02115

[6]Eberly Center, Carnegie Mellon University, Pittsburgh, PA 15289

[7]Department of Cognitive Science, Carleton University, Ottawa, Canada

## Abstract

To understand the architecture of human language, it is critical to examine diverse languages; yet most cognitive neuroscience research has focused on a handful of primarily Indo-European languages. Here, we report an investigation of the fronto-temporo-parietal language network across 45 languages and establish the robustness to cross-linguistic variation of its topography and key functional properties, including left-lateralization, strong functional integration among its brain regions, and functional selectivity for language processing.

Approximately 7,000 languages are currently spoken and signed across the globe[1]. These are distributed across more than 100 language families—groups of languages that have descended from a common ancestral language, called the proto-language—which vary in size from 2 to over 1,500 languages. Certain properties of human languages have been argued to be universal, including their capacity for productivity [2] and communicative

efficiency[3]. However, language is the only animal communication system that manifests in so many different forms[4]. The world's languages exhibit striking diversity[4], with differences spanning the sound inventories, the complexity of derivational and functional morphology, the ways in which the conceptual space is carved up into lexical categories, and the rules for how words can combine into phrases and sentences. To truly understand the nature of the cognitive and neural mechanisms that can handle the learning and processing of such diverse languages, we have to go beyond the limited set of languages used in most psycho- and neuro-linguistic studies[5,6]. This much needed step will also foster inclusion and representation in language research[7].

Here, in a large-scale fMRI investigation, we evaluate the claim of language universality with respect to core features of its neural architecture. In the largest to date effort to sample many diverse languages, we tested native speakers of 45 languages across 12 language families (Afro-Asiatic, Austro-Asiatic, Austronesian, Dravidian, Indo-European, Japonic, Koreanic, Atlantic-Congo, Sino-Tibetan, Turkic, Uralic, and an isolate—Basque, which is effectively a one-language family). To our knowledge, about a third of these languages have never been investigated with functional brain imaging (or only probed in clinical contexts), no experimental paradigm has been tested with more than four languages at a time[8], and no attempts have been made to standardize tasks / language network definitions across languages, as needed to enable meaningful comparisons across studies (Supp. Table 1).

Using a powerful individual-subject analytic approach[9], we examined the cross-linguistic generality of the following properties of the language network: i) topography (robust responses to language in the frontal, temporal, and parietal brain areas), ii) lateralization to the left hemisphere, iii) strong functional integration among the different regions of the network as assessed with inter-region functional correlations during naturalistic cognition, and iv) functional selectivity for language processing. All these properties have been previously shown to hold for English speakers. Because of their robustness at the individual-subject level[10], and in order to test speakers of as many languages as possible, we adopted a 'shallow' sampling approach—testing a small number (n=2) of speakers for each language. The goal was not to evaluate any particular hypothesis/-es about cross-linguistic differences in the neural architecture of language processing (see discussion toward the end of the paper for examples), but rather to ask whether the core properties that have been attributed to the 'language network' based on data from English and a few other dominant languages extend to typologically diverse languages. Although we expected this to be the case, this demonstration is an essential foundation for future systematic, in-depth, and finer-grained cross-linguistic comparisons. Another important goal was to develop robust tools for probing diverse languages in future neuroscientific investigations.

Each participant performed several tasks during the scanning session. First, they performed two language 'localizer' tasks: the English localizer based on the contrast between reading sentences and nonword sequences[9] (all participants were fluent in English; Supp. Table 3), and a critical localizer task, where they listened to short passages from *Alice in Wonderland* in their native language, along with two control conditions (acoustically degraded versions of the native language passages where the linguistic content was not discernible and passages in an unfamiliar language). Second, they performed one or two non-linguistic

tasks that were included to assess the functional selectivity of the language regions[11] (a spatial working memory task, which everyone performed, and an arithmetic addition task, performed by 67 of the 86 participants). Finally, they performed two naturalistic cognition paradigms that were included to examine correlations in neural activity among the language regions, and between the language regions and regions of another network supporting high-level cognition: a ~5 min naturalistic story listening task in the participant's native language, and a 5 min resting state scan.

Consistent with prior investigations of a subset of these languages (e.g., Supp. Table 1), the activation landscape for the *Native-language>Degraded-language* contrast, which targets high-level language processing and activates the same set of brain areas as those activated by a more commonly used language localizer based on reading sentences versus nonword sequences (see [12] for a direct comparison; also Extended Data Figure 1), is remarkably consistent across languages and language families. The activations cover extensive portions of the lateral surfaces of left frontal, temporal, and parietal cortex (Figures 1, 2; see Extended Data Figure 2 and Supp. Figure 1 for right hemisphere (RH) maps, and Extended Data Figure 3 for volume-based maps). In the left-hemisphere language network (defined by the English localizer; see Extended Data Figure 4 for evidence that similar results obtain in functional regions of interest (fROIs) defined by the Alice localizer), across languages, the *Native-language* condition elicits a reliably greater response than both the *Degraded-language* condition (2.13 vs. 0.84 % BOLD signal change relative to the fixation baseline; t(44)=21.0, p<0.001) and the *Unfamiliar-language* condition (2.13 vs. 0.76; t(44)=21.0, p<0.001) (Figure 3a; see Extended Data Figure 5, Supp. Figures 2,3 for data broken down by language, language family, and fROI, respectively; see Supp. Table 2 for analyses with linear mixed effects models). Across languages, the effect sizes for the *Native-language>Degraded-language* and the *Native-Language>Unfamiliar-language* contrasts range from 0.49 to 2.49, and from 0.54 to 2.53, respectively; importantly, for these and all other measures, the inter-language variability is comparable to, or lower than, inter-individual variability (Extended Data Figure 6, Supp. Figures 4,5).

The *Native-language>Degraded-language* effect is stronger in the left hemisphere fROIs than the right hemisphere ones (2.13 vs. 1.47; t(44)=7.00, p<0.001), and more spatially extensive (318.2 vs. 203.5 voxels; t(44)=6.97, p<0.001; Figure 3b). Additionally, in line with prior data from English[13], the regions of the language network (here and elsewhere, by 'regions' we mean individually defined fROIs) exhibit strong correlations in their activity during naturalistic cognition, with the average LH within-network correlation of r=0.52 during story comprehension and r=0.41 during rest, both reliably higher than zero (ts(44)>31.0, ps<0.001) and phase-shuffled baselines (ts(44)>10.0, ps<0.001; Figure 3c; see Extended Data Figure 7 and Supp. Figure 6 for data broken down by language). The correlations are stronger during story comprehension than rest (t(44)=−6.34, p<0.01). Further, as in prior work in English[13], and mirroring lateralization effects in the strength and extent of activation, the inter-region correlations in the LH language network are reliably stronger than those in the RH during both story comprehension (0.52 vs. 0.35; t(44)=8.00, p<0.001) and rest (0.41 vs. 0.28; t(44)=8.00, p<0.001; Figure 3c).

Finally, brain regions that support language processing have been shown to exhibit strong selectivity for language over many non-linguistic tasks, including executive function tasks, arithmetic processing, music perception, and action observation[11,14]. This selectivity appears to be robustly present across speakers of diverse languages. Responses to the *Native-language* condition are significantly higher than those to the spatial working memory (WM) task (2.13 vs. −0.01; t(44)=20.7, p<0.001), and the math task (2.13 vs. 0.03; t(40)=21.5, p<0.001; Figure 3a, Extended Data Figures 4,5, Supp. Figures 2,3). Furthermore, as in English[13], the language regions are robustly dissociated in their intrinsic fluctuation patterns from the regions of the bilateral domain-general multiple demand (MD) network implicated in executive functions[15]: within-network correlations are reliably greater than between-network correlations both during story comprehension (0.43 (language network, across the left and right hemisphere), 0.40 (MD network) vs. −0.01 (language-MD); ts(44)>23, ps<0.001), and during rest (0.34 (language, across hemispheres), 0.43 (MD) vs. −0.03 (language-MD), ts(44)>20, ps<0.001; Figure 3c, Extended Data Figure 7, Supp. Figure 6).

In summary, we have here established that key properties of the neural architecture of language hold across speakers of 45 diverse languages spanning 12 language families; and the variability observed across languages is comparable to, or lower than, the inter-individual variability among speakers of the same language[10] (Extended Data Figures 6,8, Supp. Figures 4,5). Presumably, these features of the language network, including a) its *location* with respect to other—perceptual, cognitive, and motor—systems, b) *lateralization* to the left hemisphere (in most individuals), c) strong *functional integration* among the different components, and d) *selectivity* for linguistic processing, make it well-suited to support the broadly common features of languages, shaped by biological and cultural evolution.

In spite of their shared features, languages do exhibit remarkable variation[4]. How this variation relates to the neural implementation of linguistic computations remains a largely open question. By establishing broad cross-linguistic similarity in the language network's properties and making publicly available the 'localizer' tasks (https://evlab.mit.edu/aliceloc) for 46 languages (to be continuously expanded over time), this work lays a critical foundation for future in-depth cross-linguistic comparisons along various dimensions of interest. In contrast to the shallow sampling approach adopted here (testing a small number of speakers across many languages), such investigations will require testing large numbers of speakers for each language / language family in question, while matching the groups carefully on all the factors that may affect neural responses to language. Such 'deep' sampling of each language / language family is necessary because cross-linguistic differences in the neural implementation of language processing are likely to be relatively subtle and they would need to exceed the (substantial) variability that characterizes speakers of the same language in order to be detected[10,16]. The language localizer tasks enable narrowing in on the system of interest—the fronto-temporo-parietal network that selectively supports linguistic processing—thus yielding greater statistical power[17], critical for detecting small effects, and interpretability, and leading to a robust and cumulative research enterprise. Future investigations of cross-linguistic similarities and differences may also call for i) more fine-tuned/targeted paradigms (cf. the broad language contrast examined here), ii) multivariate analytic approaches, and iii) methods with high temporal resolution,

like MEG or intracranial recordings (e.g., see [18–20] for past reports of cross-linguistic differences as measured with EEG).

What might hypotheses about cross-linguistic differences in neural implementation of language look like? Some examples include the following: i) languages with relatively strict word orders, compared to free-word-order languages, may exhibit a higher degree of left lateralization, given the purportedly greater role of the left hemisphere in auditory and motor sequencing abilities[21,22], or stronger reliance on the dorsal stream, for similar reasons[23]; ii) tonal languages may exhibit stronger anatomical and functional connections between auditory areas that process pitch[24] and the higher-level language areas given the need to incorporate pitch information in interpreting word meanings (see [25] for evidence of a cross-linguistic difference in the lower-level speech perception cortex between speakers of a tonal vs. a non-tonal language); and iii) languages where utterances tend to underdetermine the meaning, like Riau Indonesian[26], may place greater demands on inferential processing to determine speaker intent and thus exhibit stronger reliance on brain areas that support such processes, like the right hemisphere language areas[27] and/or the system that supports mental state attribution[28].

Another class of hypotheses might come from the field of natural language processing (NLP). Recent advances in artificial intelligence have given rise to artificial neural network (ANN) models that achieve impressive performance on diverse language tasks[29] and capture neural responses during language processing in the human brain[30]. Although, like cognitive neuroscience, NLP has been dominated by investigations of English, there is growing awareness of the need to increase linguistic diversity in the training and evaluation of language models[31], and some work has begun to probe cross-linguistic similarities and differences in the models' learned representations[32]. A promising future direction is to relate these cross-linguistic differences to neural differences observed during language processing across languages in an effort to illuminate how language implementation—in silico or in biological tissue—may depend on the properties of a particular language. More generally, because searching for cross-linguistic neural differences is a relatively new direction for language research[cf. 5,6], it will likely require a combination of top-down theorizing and bottom-up discovery. But no matter what discoveries about cross-linguistic differences in neural implementation lie ahead, the ability to reliably identify the language network in speakers of diverse languages opens the door to investigations of linguistic phenomena that are present in a small subset of the world's languages, to paint a richer picture of the human language system.

Two limitations of the current investigation are worth noting. First, all participants were bilingual (fluent in English, in addition to their native language), which was difficult to avoid given that the research was carried out in the U.S. Some have argued that knowledge of two or more languages affects the neural architecture of one's native language processing[33], but this question remains controversial[34]. More generally, finding 'pure' monolingual speakers with no knowledge of other languages is challenging, especially in globalized societies, and is nearly impossible for some languages (e.g., Dutch, Galician, Kashmiri). The approach advocated here—where the language network is defined in each individual participant and individual-level neural markers are examined—allows taking into account and explicitly

modeling inter-individual variability in participants' linguistic profiles (and along other dimensions), as will be important when evaluating specific hypotheses about cross-linguistic differences in future work, as discussed above. Another limitation is the over-representation of Indo-European languages (31 of the 45 languages). The analysis in Supp. Figure 2, which shows that the key statistics hold across language families, ameliorates this concern to some extent. Nevertheless, development of language localizers and collection of data for non-Indo-European languages remains a priority for the field. Our group will continue to develop and release the localizers for additional languages (https://evlab.mit.edu/aliceloc), and we hope other labs across the world will join this effort.

In conclusion, probing human language in all its diverse manifestations is critical for uncovering additional shared features, understanding the cognitive and neural basis of different solutions to similar communicative demands, characterizing the processing of unique/rare linguistic properties, and fostering diversity and inclusion in language sciences.

## Methods

### Participants.

Ninety-one participants were recruited from MIT and the surrounding Boston community. Participants were recruited on the basis of their native language (the language acquired during the first few years of life; Supp. Table 3). All participants were proficient in English (Supp. Table 3). Data from 5 participants were excluded from the analyses due to excessive in-scanner motion or sleepiness. The final set included 86 participants (43 males) between the ages of 19 and 45 (M=27.52, SD=5.49; Supp. Table 4). All participants were right-handed, as determined by the Edinburgh Handedness Inventory[35] (n=83) or self-report (n=3), and had normal or corrected-to-normal vision. All participants gave informed written 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.

Participants' native languages spanned 12 language families (Afro-Asiatic, Austro-Asiatic, Austronesian, Dravidian, Indo-European, Japonic, Koreanic, Atlantic-Congo, Sino-Tibetan, Turkic, Uralic, Isolate (Basque)) and 45 languages (Supp. Table 3). We tested 2 native speakers per language (one male, one female) when possible; for 4 of the 45 languages (Tagalog, Telugu, Slovene, and Swahili), we were only able to test one native speaker.

### Experimental Design.

Each participant completed i) a standard language localizer task in English[9], ii) the critical language localizer in their native language, iii) one or two non-linguistic tasks that were included to assess the degree of functional selectivity of the language regions (a spatial working memory task, which everyone performed, and an arithmetic addition task, performed by 67 of the 86 participants), and iv) two naturalistic cognition paradigms that were included to examine correlations in neural activity among the language regions, and between the language regions and regions of another network supporting high-level cognition—the domain-general multiple demand (MD) network[15] (a ~5 min naturalistic story listening task in the participant's native language, and a 5 min resting state scan).

With the exception of two participants, everyone performed all the tasks in a single scanning session, which lasted approximately two hours. One participant performed the English localizer in a separate session, and another performed the spatial working memory task in a separate session. (We have previously established that individual activations are highly stable across scanning sessions[10, see also 36].)

**Standard (English-based) language localizer**—Participants passively read English sentences and lists of pronounceable nonwords in a blocked design. The *Sentences>Nonwords* contrast targets brain regions that support high-level linguistic processing, including lexico-semantic and combinatorial syntactic/semantic processes[37–39]. Each trial started with 100 ms pre-trial fixation, followed by a 12-word-long sentence or a list of 12 nonwords presented on the screen one word/nonword at a time at the rate of 450 ms per word/nonword. Then, a line drawing of a finger pressing a button appeared for 400 ms, and participants were instructed to press a button whenever they saw this icon, and finally a blank screen was shown for 100 ms, for a total trial duration of 6 s. The simple button-pressing task was included to help participants stay awake and focused. Each block consisted of 3 trials and lasted 18 s. Each run consisted of 16 experimental blocks (8 per condition), and five fixation blocks (14 s each), for a total duration of 358 s (5 min 58 s). Each participant performed two runs. Condition order was counterbalanced across runs. (We have previously established the robustness of the language localizer contrast to modality (written/auditory), materials, task, and variation in the experimental procedure.[9,12,40])

**Critical (native-language-based) language localizer**

**Materials.:** Translations of *Alice in Wonderland*[41] were used to create the materials. We chose this text because it is one of the most translated works of fiction, with translations existing for at least 170 languages[42], and is suitable for both adults and children. Using the original (English) version, we first selected a set of 28 short passages (each passage took between 12 and 30 sec to read out loud). We also selected 3 longer passages (each passage took ~5 min to read out loud) to be used in the naturalistic story listening task (see below). For each target language, we then recruited a native female speaker, who was asked to a) identify the corresponding passages in the relevant translation (to ensure that the content is similar across languages), b) familiarize themselves with the passages, and c) record the passages. In some languages, due to the liberal nature of the translations, the corresponding passages differed substantially in length from the original versions; in such cases, we adjusted the length by including or omitting sentences at the beginning and/or end of the passage so that the length roughly matched the original. We used female speakers because we wanted to ensure that the stimuli would be child-friendly (for future studies), and children tend to pay better attention to female voices[43]. Most speakers were paid for their help, aside from a few volunteers from the lab. Most of the recordings were conducted in a double-walled sound-attenuating booth (Industrial Acoustics). Materials for 3 of the languages (Hindi, Tamil, and Catalan) were recorded outside the U.S.; in such cases, recordings were done in a quiet room using a laptop's internal microphone. We ensured that all recordings were fluent; if a speaker made a speech error, the relevant portion/passage were re-recorded. For each language, we selected 24 of the 28 short passages to be used in the experiment, based on length so that the target passages were as close to 18 s as

possible. Finally, we created acoustically degraded versions of the target short passages following the procedure introduced in Scott et al.[12]. In particular, for each language, the intact files were low-pass filtered at a pass-band frequency of 500 Hz. In addition, a noise track was created from each intact clip by randomizing 0.02-second-long periods. In order to produce variations in the volume of the noise, the noise track was multiplied by the amplitude of the intact clip's signal over time. The noise track was then low-pass filtered at a pass-band frequency of 8,000 Hz and a stop frequency of 10,000 Hz in order to soften the highest frequencies. The noise track and the low-pass filtered copies of the intact files were then combined, and the level of noise was adjusted to a point that rendered the clips unintelligible. The resulting degraded clips sound like poor radio reception of speech, where the linguistic content is not discernible. In addition to the intact and degraded clips in their native language, we included a third condition: clips in an unfamiliar language (Tamil was used for 75 participants and Basque for the remaining 11 participants who had some exposure to Tamil during their lifetime). All the materials are available from the Fedorenko lab website: https://evlab.mit.edu/aliceloc.

**Procedure.:** For each language, the 24 items (intact-degraded pairs) were divided across two experimental lists so that each list contained only one version of an item, with 12 intact and 12 degraded trials. Any given participant was presented with the materials in one of these lists. Each list additionally contained 12 unfamiliar foreign language clips (as described above) chosen randomly from the set of 24. Participants passively listened to the materials in a long-event-related design, with the sound delivered through Sensimetrics earphones (model S14). The *Native-language* condition was expected to elicit stronger responses compared to both the *Degraded-language* condition[12] and the *Unfamiliar-language* condition[40] in the high-level language processing brain regions[9]. These language regions appear to support the processing of word meanings and combinatorial semantic/syntactic processes[37–39], and these processes are not possible for the degraded or unfamiliar conditions. Each event consisted of a single passage and lasted 18 s (passages that were a little shorter than 18 s were padded with silence at the end, and passages that were a little longer than 18 s were trimmed down). We included a gradual volume fade-out at the end of each clip during the last 2 s, and the volume levels were normalized across the 36 clips (3 conditions * 12 clips each) in each set. The materials were divided across three runs, and each run consisted of 12 experimental events (4 per condition), and three fixation periods (12 s each), for a total duration of 252 s (4 min 12 s). Each participant performed three runs. Condition order was counterbalanced across runs.

**Non-linguistic tasks**—Both tasks were chosen based on prior studies of linguistic selectivity[11]. In the *spatial working memory* task, participants had to keep track of four (easy condition) or eight (hard condition) locations in a $3 \times 4$ grid[11]. In both conditions, participants performed a two-alternative forced-choice task at the end of each trial to indicate the set of locations that they just saw. Each trial lasted 8 s (see [11] for the timing details). Each block consisted of 4 trials and lasted 32 s. Each run consisted of 12 experimental blocks (6 per condition), and 4 fixation blocks (16 s in duration each), for a total duration of 448 s (7 min 28 s). Each participant performed 2 runs. Condition order was

counterbalanced across runs. Note that in the main analyses of this task and the math task, we averaged across the hard and easy conditions (but see Extended Data Figure 9).

In the *arithmetic addition* task, participants had to solve a series of addition problems with smaller (easy condition) vs. larger (hard condition) numbers. In the easy condition, participants added two single-digit numbers. In the hard condition, participants added two numbers, one of which was double-digits. In both conditions, participants performed a two-alternative forced-choice task at the end of each trial to indicate the correct sum. Each trial lasted 3 s. Each block consisted of 5 trials and lasted 15 s. Each run consisted of 16 experimental blocks (8 per condition), and 5 fixation blocks (15 s in duration each), for a total duration of 315 s (5 min 15 s). Most participants performed 2 runs; 12 participants performed 1 run; 19 participants did not perform this task due to time limitations. Condition order was counterbalanced across runs when multiple runs were performed.

**Naturalistic cognition paradigms**—In the *story listening* paradigm, participants were asked to attentively listen to one of the long passages in their native language. The selected passage was 4 min 20 s long in English. Recordings in other languages were padded with silence or trimmed at the end, to equalize scan length across languages. The same 2 sec fade-out was applied to these clips, as to the shorter clips used in the critical experiment. In addition, each run included 12 s of silence at the beginning and end, for a total duration of 284 s (4 min 44 s). In the *resting state* paradigm, following Blank et al. (2014)[13], participants were asked to close their eyes but to stay awake and let their mind wander for 5 minutes. The projector was turned off, and the lights were dimmed.

### fMRI data acquisition.

Structural and functional data were collected on the whole-body 3 Tesla Siemens Trio scanner with a 32-channel head coil at the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research at MIT. T1-weighted structural images were collected in 179 sagittal slices with 1 mm isotropic voxels (TR = 2,530 ms, TE = 3.48 ms). Functional, blood oxygenation level dependent (BOLD) data were acquired using an EPI sequence (with a 90° flip angle and using GRAPPA with an acceleration factor of 2), with the following acquisition parameters: thirty-one 4mm thick near-axial slices, acquired in an interleaved order with a 10% distance factor; 2.1 mm × 2.1 mm in-plane resolution; field of view of 200mm in the phase encoding anterior to posterior (A >> P) direction; matrix size of 96 × 96; TR of 2,000 ms; and TE of 30 ms. Prospective acquisition correction[44] was used to adjust the positions of the gradients based on the participant's motion one TR back. The first 10 s of each run were excluded to allow for steady-state magnetization.

### fMRI data preprocessing and first-level analysis.

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 B-spline interpolation to the first scan of the first session. 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). 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 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 2 mm isotropic voxels and 4th order 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. For the language localizer task and the non-linguistic tasks, effects 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 s 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 main condition effects, other model parameters in the GLM design included first-order temporal derivatives for each condition (for modeling spatial variability in the HRF delays) as well as nuisance regressors to control for the effect of slow linear drifts, subject-motion parameters, and potential outlier scans on the BOLD signal.

The naturalistic cognition paradigms (story listening and resting state) were preprocessed using the CONN toolbox[45] with default parameters, unless stated otherwise. First, in order to remove noise resulting from signal fluctuations originating from non-neuronal sources (e.g., cardiac or respiratory activity), the first five BOLD signal time points extracted from the white matter and CSF were regressed out of each voxel's time-course. White matter and CSF voxels were identified based on segmentation of the anatomical image[46]. Second, the residual signal was band-pass filtered at 0.008–0.09 Hz to preserve only low-frequency signal fluctuations[47].

To create aesthetically pleasing activation projection images for Figure 1, the data were additionally analyzed in FreeSurfer[48]. Although all the analyses were performed on the data analyzed in the volume (in SPM12), these surface-based maps are available at OSF, along with the volume-analysis-based maps: https://osf.io/cw89s/.

### fROI definition and response estimation.

For each participant, functional regions of interest (fROIs) were defined using the Group-constrained Subject-Specific (GSS) approach[9], whereby a set of parcels or "search spaces" (i.e., brain areas within which most individuals in prior studies showed activity for the localizer contrast) is combined with each individual participant's activation map for the same contrast.

To define the language fROIs, we used six parcels derived from a group-level representation of data for the *Sentences>Nonwords* contrast in 220 participants (Figure 3a). These parcels

included three regions in the left frontal cortex: one in the inferior frontal gyrus (LIFG, 740 voxels; given that each fROI is 10% of the parcel, as described below, the fROI size is a tenth of the parcel size), one in its orbital part (LIFGorb, 370 voxels), and one in the middle frontal gyrus (LMFG, 460 voxels); and three regions in the left temporal and parietal cortex spanning the entire extent of the lateral temporal lobe and extending into the angular gyrus (LAntTemp, 1,620 voxels; LPostTemp, 2,940 voxels; and LAngG, 640 voxels). (We confirmed that parcels created based on the probabilistic overlap map for *Native-language>Degraded-language* contrast from the 86 participants in the current study are similar (Supp. Figure 7). We chose to use the 'standard' parcels for ease of comparison with past studies.) Individual fROIs were defined by selecting—within each parcel—the top 10% of most localizer-responsive voxels based on the *t*-values for the relevant contrast (*Sentences>Nonwords* for the English localizer). We then extracted the responses from these fROIs (averaging the responses across the voxels in each fROI) to each condition in the critical language localizer (native language intact, acoustically degraded native language, and unfamiliar language), and the non-linguistic tasks (averaging across the hard and easy conditions for each task). Statistical tests were then performed across languages on the percent BOLD signal change values extracted from the fROIs.

We used the English-based localizer to define the fROIs i) because we have previously observed[40] that the localizer for a language works well as long as a participant is proficient in that language (as was the case for our participants' proficiency in English (Supp. Table 3); see also Supp. Figure 8 for evidence that our participants' responses to the English localizer conditions were similar to those of native speakers), and ii) to facilitate comparisons with earlier studies[11,13]. However, in an alternative set of analyses (Extended Data Figure 4), we used the *Native-language>Degraded-language* contrast from the critical language localizer to define the fROIs. In that case, to estimate the responses to the conditions of the critical language localizer, across-runs cross-validation[17] was used to ensure independence[49]. The results were nearly identical to the ones based on the English localizer fROIs, suggesting that the two localizers pick out similar sets of voxels. Furthermore, for the two native speakers of English who participated in this study, the *Native-language>Degraded-language* contrast and the *Sentences>Nonwords* contrast are voxel-wise spatially correlated at 0.88 within the union of the language parcels (Fisher-transformed correlation[50]; Extended Data Figure 1). (Following a reviewer's suggestion, we further explored the similarity of the activation maps for the *Native-language>Degraded-language* and *Native-language>Unfamiliar-language* contrasts in the Alice localizer. These maps were similar: across the 86 participants, the average Fisher-transformed voxel-wise spatial correlation within the union of the language parcels was 0.66 (SD = 0.40; see Extended Data Figure 10 for sample individual map pairs), and the magnitudes of these effects did not differ statistically (t(44)=1.15, p=0.26). These results suggest that either contrast can be used to localize language-responsive cortex—along with the more traditional *Sentences>Nonwords* contrast—although we note that, among the two auditory contrasts, we have more and stronger evidence that the *Native-language>Degraded-language* works robustly and elicits similar responses to the *Sentences>Nonwords* contrast.)

In addition to the magnitudes of response, we estimated the degree of language lateralization in the native language localizer based on the extent of activation in the left vs. right

hemisphere. To do so, for each language tested, in each participant, we calculated the number of voxels activated for the *Native-language>Degraded-language* contrast (at the $p$<0.001 whole-brain uncorrected threshold) within the union of the six language parcels in the left hemisphere, and within the union of the homotopic parcels in the right hemisphere[13] (Figure 3b). Statistical tests were then performed across languages on the voxel count values. (We additionally performed a similar analysis considering the voxels across the brain[51].)

Finally, we calculated inter-regional functional correlations during each of the naturalistic cognition paradigms. For these analyses, in addition to the language fROIs, we examined a set of fROIs in another large-scale brain network that supports high-level cognition: the domain-general multiple demand (MD) network[15,52], which has been implicated in executive functions, like attention, working memory, and cognitive control. This was done in order to examine the degree to which the language regions are functionally dissociated from these domain-general MD regions during rich naturalistic cognition, as has been shown to be the case for native English speakers[13,53]. To define the MD fROIs, following[13,54], we used anatomical parcels[55] that correspond to brain regions linked to MD activity in prior work. These parcels included regions in the opercular IFG, MFG, including its orbital part, insular cortex, precentral gyrus, supplementary and presupplementary motor area, inferior and superior parietal cortex, and anterior cingulate cortex, for a total of 18 regions (9 per hemisphere). Individual MD fROIs were defined by selecting—within each parcel—the top 10% of most localizer-responsive voxels based on the *t*-values for the *Hard>Easy* contrast for the spatial working memory task[13] (see Extended Data Figure 9 for an analysis showing that, as expected based on prior work, this effect is highly robust in the MD fROIs, as estimated using across-runs cross-validation).

For each subject, we averaged the BOLD signal time-course across all voxels in each language and MD fROI. We then averaged the time-courses in each fROI across participants for each language where two participants were tested. For each language, we computed Pearson's moment correlation coefficient between the time-courses for each pair of fROIs. These correlations were Fisher-transformed to improve normality and decrease biases in averaging[50]. We then compared the average correlation for each language a) within the language network (the average of all 66 pairwise correlations among the 12 language fROIs), b) within the MD network (the average of all 153 pairwise correlations among the 18 MD fROIs), and c) between language and MD fROIs (the average of 240 pairwise correlations between the language fROIs and the MD fROIs). For the language network, we also computed the within-network correlations for the left and right hemisphere separately, to examine lateralization effects. (Following a reviewer's suggestion, we also explored the differences in inter-hemispheric connectivity within the language network during the two naturalistic paradigms; inter-hemispheric connectivity was higher during story listening (mean=0.33 SD=0.32) than during resting state (mean=0.20 SD=0.29; t(44)=8.11, p<0.01)) All statistical comparisons were performed across languages. The fROI-to-fROI correlations are visualized in two matrices, one for each naturalistic cognition paradigm (Figure 3c).

**Statistics and Reproducibility.**

Given the shallow sampling approach that we adopted (testing a small number of participants for a large number of languages), no statistical method was used to predetermine the sample size for each language. Because the neural markers that we examined have been previously established to be robust at the individual level[10], we expected them to hold in individual participants in the current sample (which was indeed the case). To allow for some generalizability, we recruited two participants (instead of just one) for each language (one male, one female) when possible. Five participants were excluded due to excessive in-scanner head motion or sleepiness, as is routinely done for studies in our lab and in the field in general. The order of tasks within a session was not randomized, although some variability in the orders was present (task orders for individual participants are available from the authors). Every participant performed the same set of tasks and conditions, so investigator blinding with respect to condition allocation does not apply. Data distributions were assumed to be normal (the distributions of the individual data points in the figures show that this was largely the case), but normality was not formally assessed; for the timecourse correlation analyses, a Fisher transformation[50] was applied to improve normality, as described in Methods.

## Extended Data



**a. Activations for subject 492**

Alice Localizer (*Native-Language-Degraded Language*)          English Localizer (*Sentences-Nonwords*)

**b. Activations for subject 502**

Alice Localizer (*Native-Language-Degraded Language*)          English Localizer (*Sentences-Nonwords*)
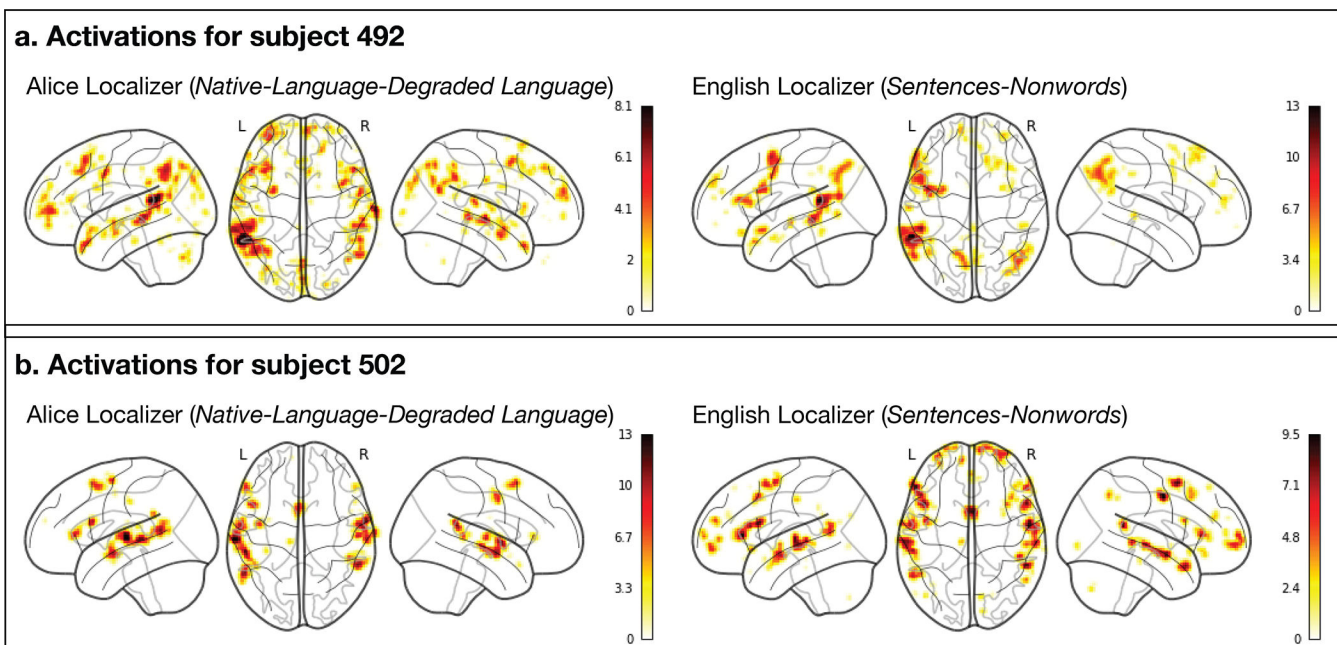
**Extended Data Fig. 1.**

Comparison of the individual activation maps for the *Sentences>Nonwords* contrast and the *Native-language>Degraded-language* contrast in the two native-English-speaking participants. The two maps are voxel-wise (within the union of the language parcels) spatially correlated at r=0.77 and r=0.99 for participants 492 and 502, respectively (the correlations are Fisher-transformed). Across the full set of participants, the average Fisher-transformed spatial correlation between the maps for the *Sentences>Nonwords* contrast in
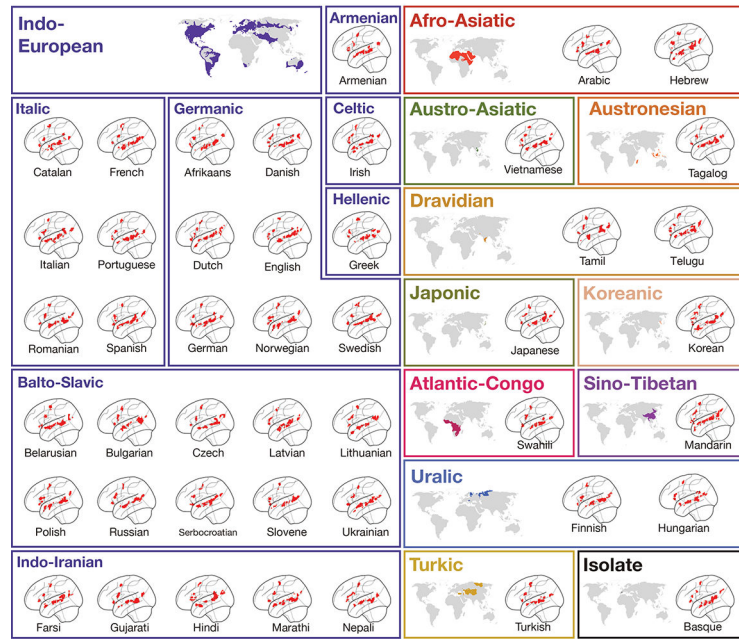
English and the *Native-language>Degraded-language* contrast in the participant's native language (again, constrained to the language parcels) is r=0.88 (SD=0.43) for the left hemisphere and 0.73 (SD=0.38) for the right hemisphere. (Note that using the union of the language parcels rather than the whole brain is conservative for computing these correlations; including all the voxels would inflate the correlations due to the large difference in activation levels between voxels that fall within the language parcels vs. outside their boundaries. Instead, we are zooming in on the activation landscape within the frontal, temporal, and parietal areas that house the language network and showing that these landscapes are spatially similar between the two contrasts in their fine-grained activation patterns.)
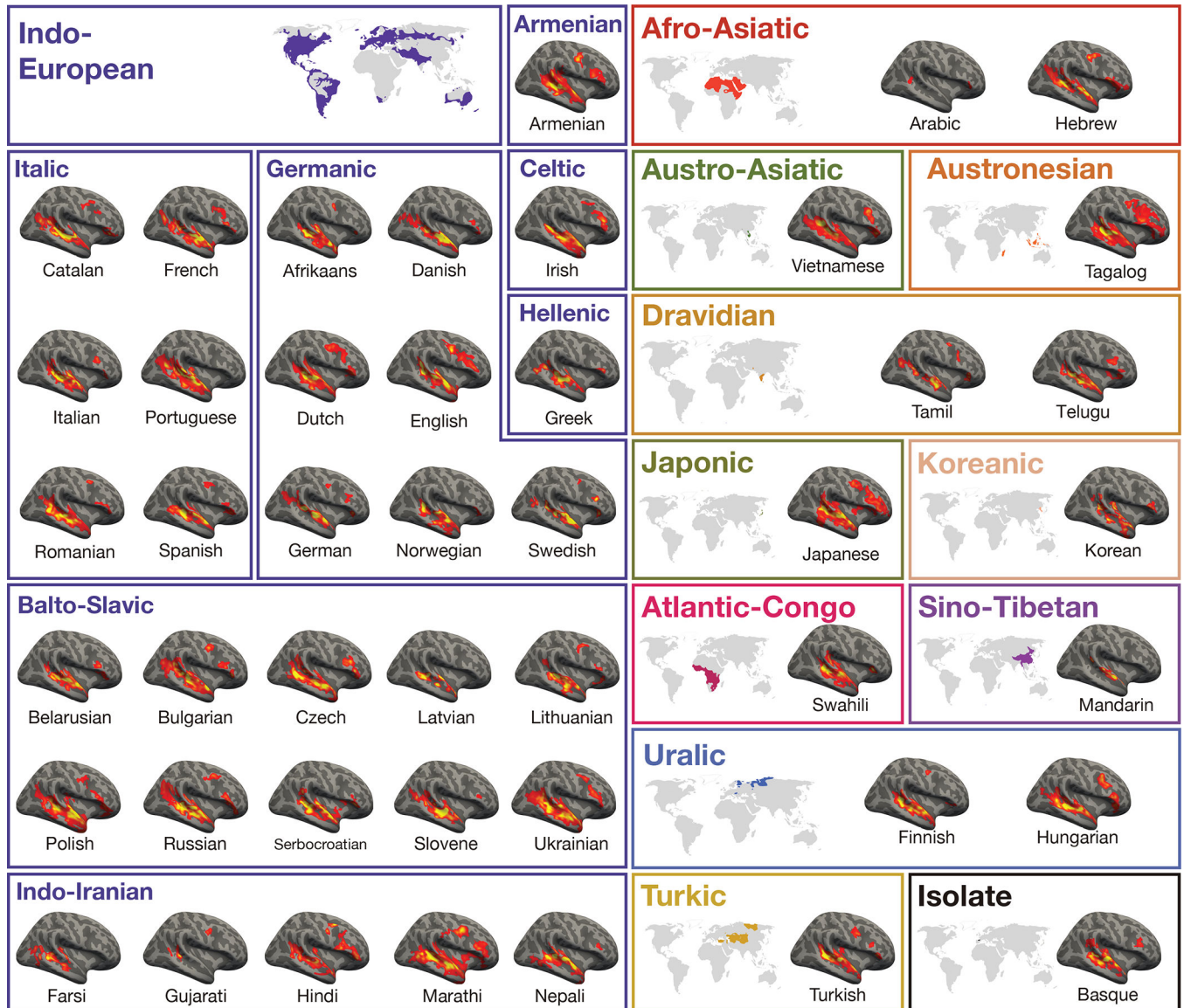


**Extended Data Fig. 2.**

Activation maps for the Alice language localizer contrast (*Native-language>Degraded-language*) in the right hemisphere of a sample participant for each language (see Figure 1 for the maps from the left hemisphere). A significance map was generated for each participant by FreeSurfer[44]; each map was smoothed using a Gaussian kernel of 4 mm full-width half-max and thresholded at the $70^{th}$ percentile of the positive contrast for each participant (this was done separately for each hemisphere; note that the same participants are used here as those used in Figure 1). The surface overlays were rendered on the 80% inflated white-gray matter boundary of the fsaverage template using FreeView/FreeSurfer. Opaque red and yellow correspond to the $80^{th}$ and $99^{th}$ percentile of positive-contrast activation for each subject, respectively. Further, here and in Figure 1, small and/or idiosyncratic bits of activation (relatively common in individual-level language maps[e.g., 9, 10]) were removed. In particular, clusters were excluded if a) their surface area was below 100 mm^2, or b) they did not overlap (by >10%) with a mask created for a large number (n=804[56]) participants by overlaying the individual maps and excluding vertices that did not show language responses in at least 5% of the cohort. (We ensured that the idiosyncrasies were individual- and not language-specific: for each cluster removed, we checked that a similar cluster was not present for the second native speaker of that language.) These maps were used solely for visualization; all the statistical analyses were performed on the data analyzed in the volume.
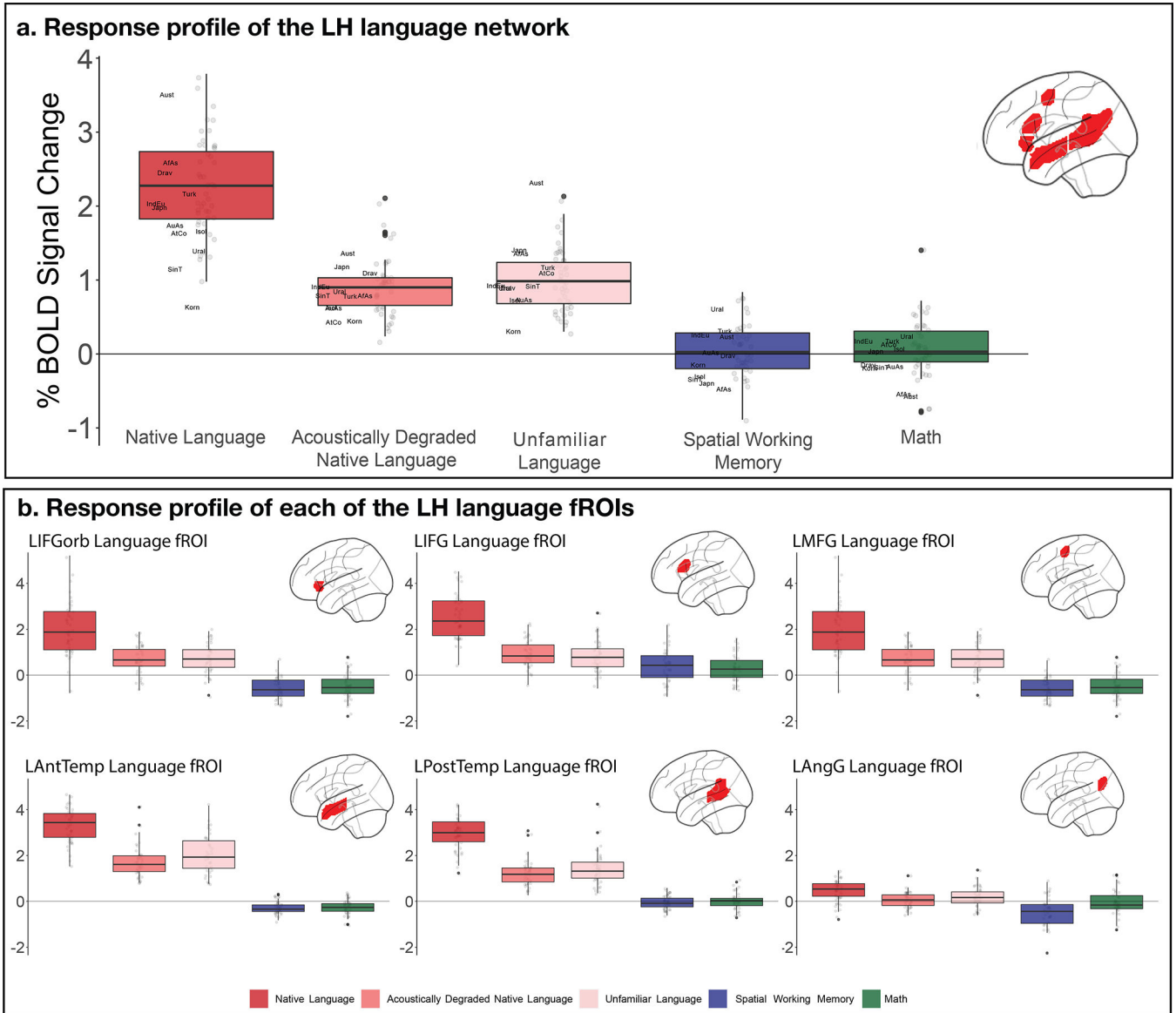
**a. Top 10% voxels within each language ROI**



**b. Thresholded statistical maps (p<0.001 uncorrected)**
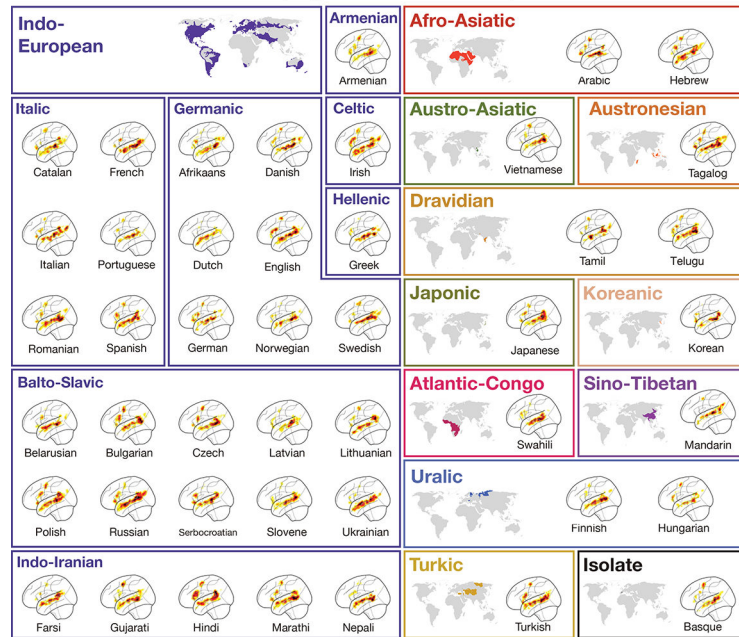


**Extended Data Fig. 3.**

Volume-based activation maps for the *Native-language>Degraded-language* contrast in the left hemisphere of a sample participant for each language (the same participants are used as those used in Figure 1 and Extended Figure 2). a) Binarized maps that were generated for each participant by selecting the top 10% most responsive (to this contrast) voxels within each language parcel. These sets of voxels correspond to the fROIs used in the analyses reported in Extended Data Figure 4 (except for the estimation of the responses to the conditions of the Alice localizer, where a subset of the runs was used to ensure

independence; the fROIs in those cases will be similar but not identical to those displayed).
b) Whole-brain maps that are thresholded at the p<0.001 uncorrected level.



**a. Response profile of the LH language network**

**b. Response profile of each of the LH language fROIs**

**Extended Data Fig. 4.**
Percent BOLD signal change across (panel a) and within each of (panel b) the LH language
functional ROIs (defined by the *Native-language>Degraded-language* contrast from the
Alice localizer, cf. the *Sentences>Nonwords* contrast from the English localizer as in the
main text and analyses; Figure 3a and Supp. Figure 3) for the three language conditions
of the Alice localizer task (Native language, Acoustically degraded native language, and
Unfamiliar language), the spatial working memory (WM) task and the math task. The
dots correspond to languages (n=45), and the labels (panel a only) mark the averages
for each language family. In all panels, box plots include the first quartile (lower hinge),
third quartile (upper hinge), and median (central line); upper and lower whiskers extend

from the hinges to the largest value no further than 1.5 times the inter-quartile range; darker-colored dots correspond to outlier data points. Across the six fROIs, the *Native-language* condition elicits a reliably greater response than both the *Degraded-language* condition (2.32 vs. 0.91 % BOLD signal change relative to the fixation baseline; t(44)=18.57, p<0.001) and the *Unfamiliar-language* condition (2.32 vs. 0.99; t(44)=18.02, p<0.001). Responses to the *Native-language* condition are also significantly higher than those to the spatial working memory task (2.32 vs. 0.06; t(44)=11.16, p<0.001) and the math task (2.32 vs. −0.02; t(40)=20.8, p<0.001). These results also hold for each fROI separately, correcting for the number of fROIs (*Native-language > Degraded-language*: ps<0.05; *Native-language > Unfamiliar-language*: ps<0.05; *Native-language > Spatial WM*: ps<0.05; and *Native-language > Math*: ps<0.05). All t-tests were two-tailed and corrected for the number of fROIs in the per-fROI analyses.
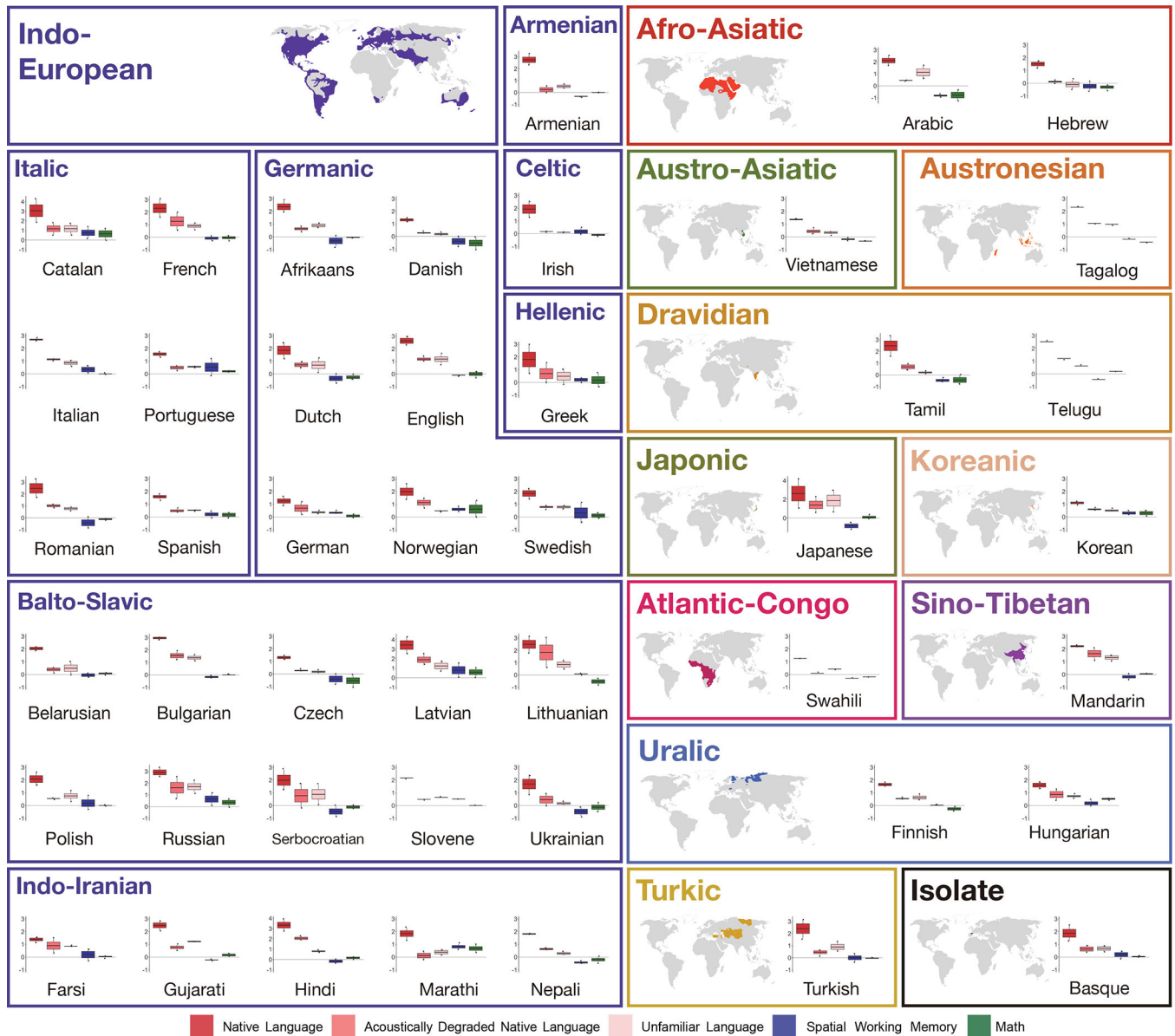
**Extended Data Fig. 5.**

Percent BOLD signal change across the LH language functional ROIs (defined by the *Sentences>Nonwords* contrast) for the three language conditions of the Alice localizer task (Native language, Acoustically degraded native language, and Unfamiliar language), the spatial working memory (WM) task, and the math task shown for each language separately. The dots correspond to participants for each language (n=2 in all languages except Slovene, Swahili, Tagalog, Telugu, where n=1). Box plots include the first quartile (lower hinge), third quartile (upper hinge), and median (central line); upper and lower whiskers extend from the hinges to the largest value no further than 1.5 times the inter-quartile range; darker-colored dots correspond to outlier data points. (Note that the scale of the y-axis differs across languages in order to allow for easier between-condition comparisons in each language.)

**a. Comparison of individual topographies between speakers of the same language vs. speakers of different languages**



**b. Comparison of topographies between language pairs**
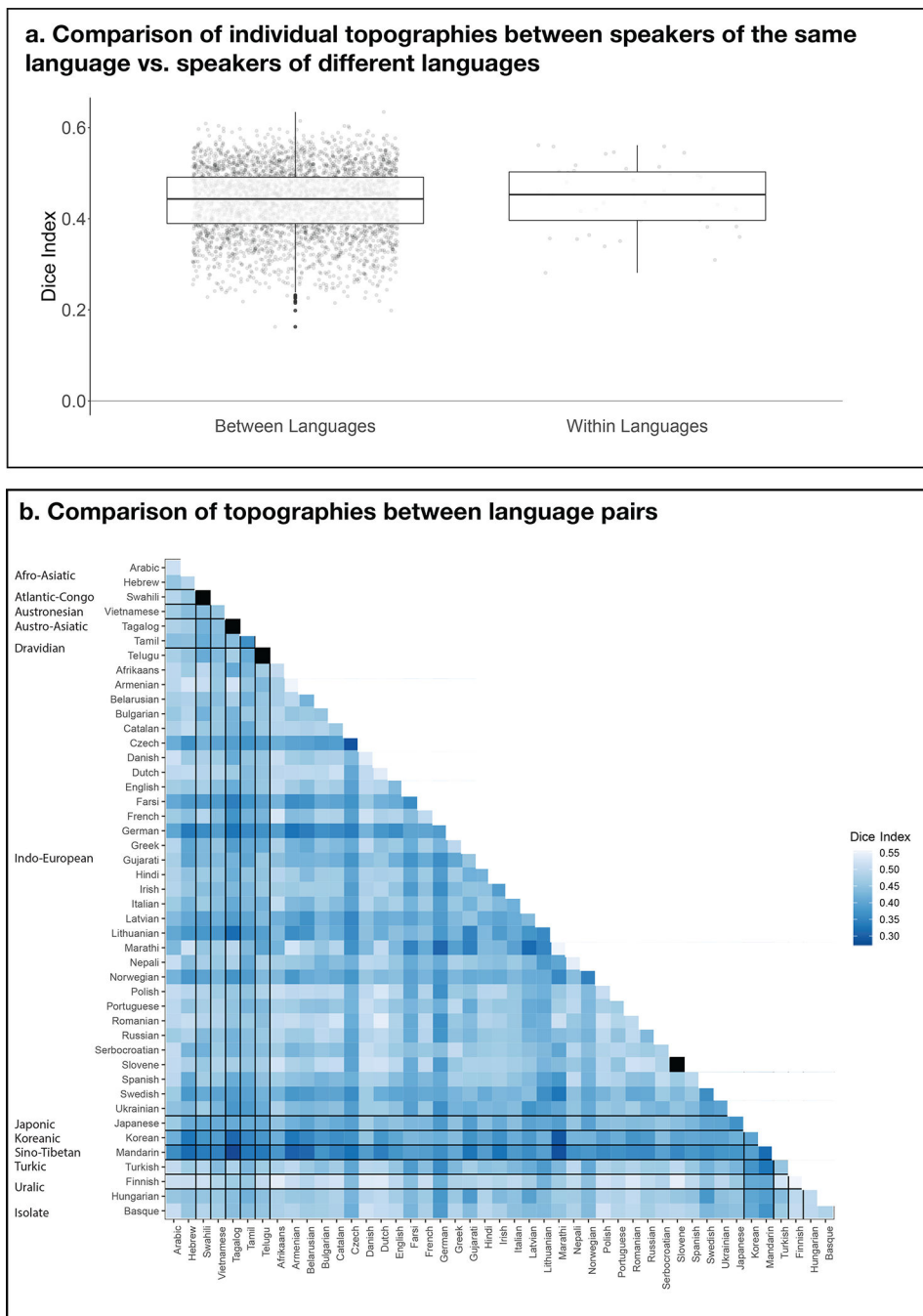


**Extended Data Fig. 6.**

A comparison of individual LH topographies between speakers of the same language vs. between speakers of different languages. The goal of this analysis was to test whether inter-language / inter-language-family similarities might be reflected in the similarity structure of the activation patterns. To perform this analysis, we computed a Dice coefficient[57] for each pair of individual activation maps for the *Intact-language>Degraded-language* contrast (a total of n=3,655 pairs across the 86 participants). To do so, we used the binarized maps like those shown in Extended Data Figure 3a, where in each LH language parcel the top 10% of

most responsive voxels were selected. Then, for each pair of images, we divided the number of overlapping voxels multiplied by 2 by the sum of the voxels across the two images (this value was always the same and equaling 1,358 given that each map had the same number of selected voxels). The resulting values can vary from 0 (no overlapping voxels) to 1 (all voxels overlap). a) A comparison of Dice coefficients for pairs of maps between languages (left, n=3,655 pairs) vs. within languages (right; this could be done for 41/45 languages for which two speakers were tested). If the activation landscapes are more similar within than between languages, then the Dice coefficients for the within-language comparisons should be higher. Instead, no reliable difference was observed by an independent-samples t-test (average within-language: 0.17 (SD=0.07), average between-language: 0.16 (SD=0.06); t(40.7)=−0.52, p= 0.61; see also Extended Data Figure 8 for evidence that the range of overlap values in probabilistic atlases created from speakers of diverse languages vs. speakers of the same language are comparable). Box plots include the first quartile (lower hinge), third quartile (upper hinge), and median (central line); upper and lower whiskers extend from the hinges to the largest value no further than 1.5 times the inter-quartile range; darker-colored dots correspond to outlier data points. b) Dice coefficient values for all pairs of within- and between-language comparisons (the squares in black on the diagonal correspond to languages with only one speaker tested). As can be seen in the figure and in line with the results in panel a, no structure is discernible that would suggest greater within-language / within-language-family topographic similarity. Similar to the results from the within- vs. between-language comparison in a, the within-language-family vs. between-language-family comparison did not reveal a difference (t(19.8)=0.71, p=0.49). In summary, in the current dataset (collected with the shallow sampling approach, i.e., a small number of speakers from a larger number of languages), no clear similarity structure is apparent that would suggest more similar topographies among speakers of the same language, or among speakers of languages that belong to the same language family.
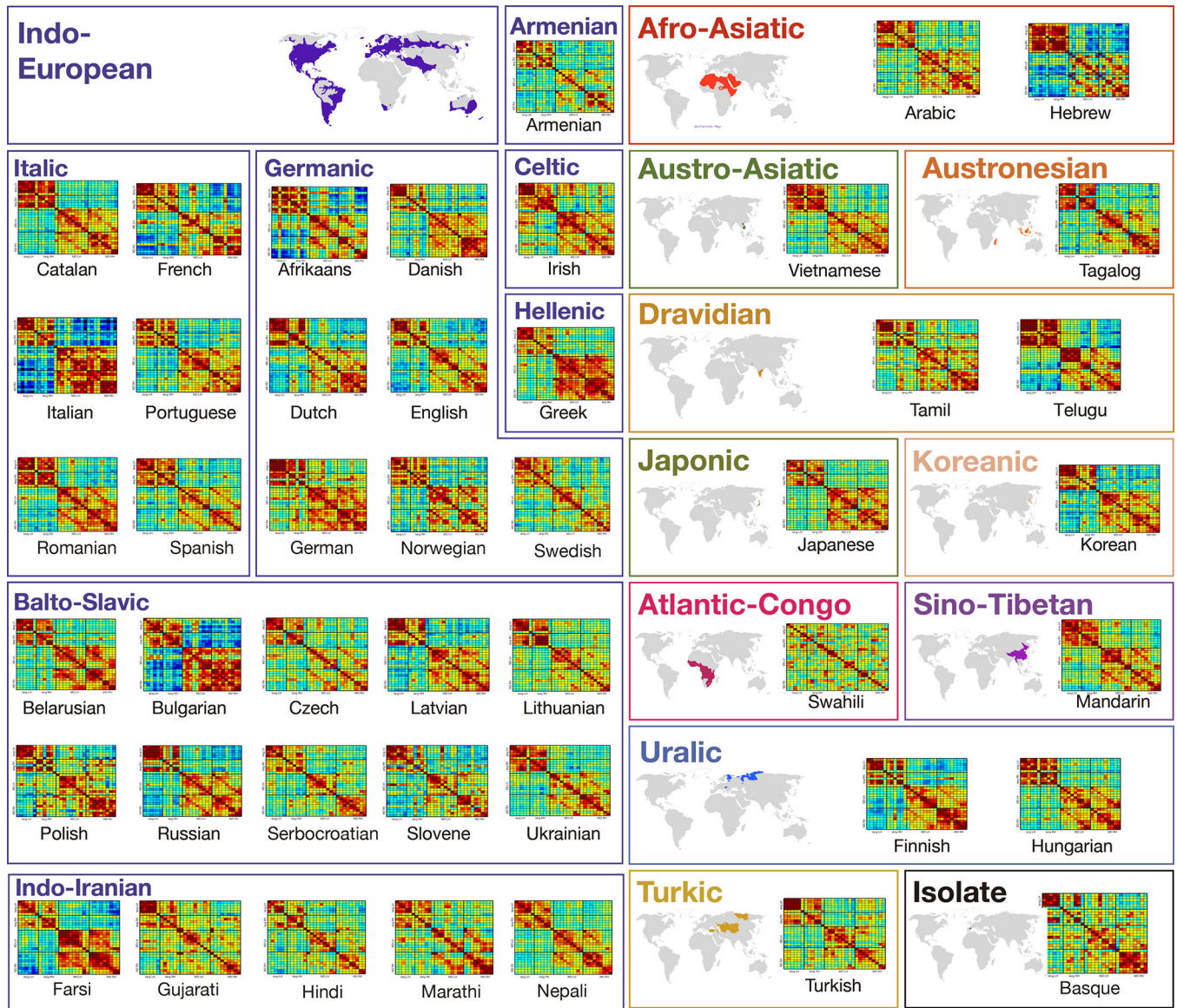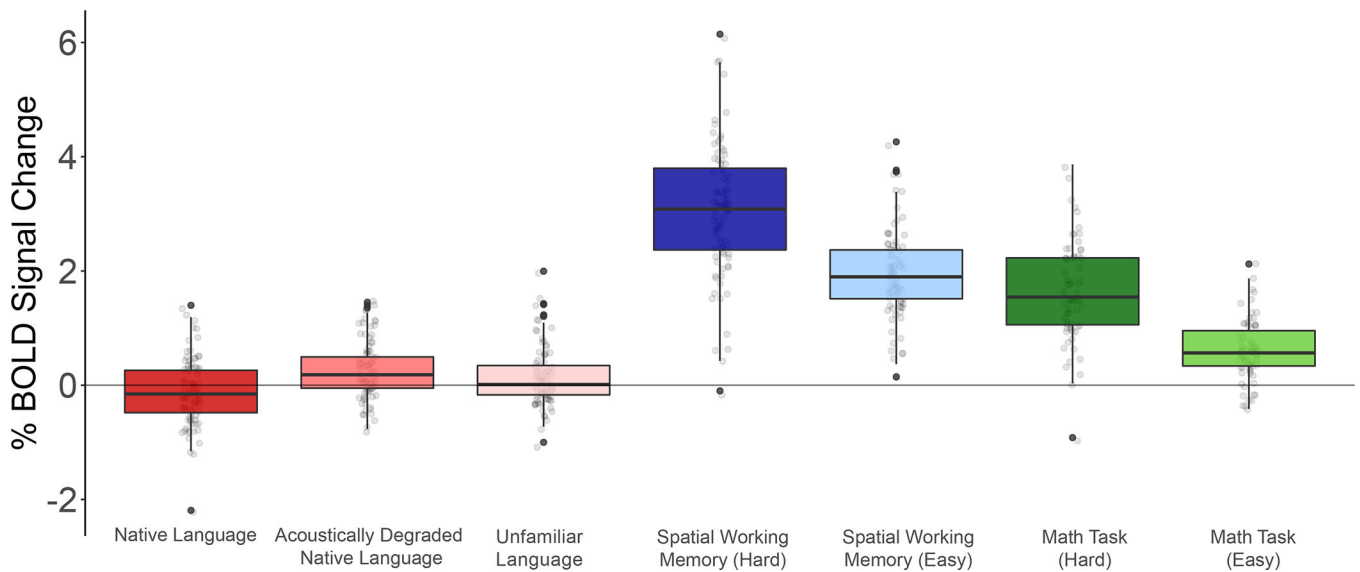
**Extended Data Fig. 7.**
Inter-region functional correlations in the language and the Multiple Demand networks during story comprehension for each of the 45 languages. Inter-region functional correlations for the LH and RH of the language and the Multiple Demand (MD) networks during a naturalistic cognition paradigm (story comprehension in the participant's native language) shown for each language separately.

**a. Alice Atlas (n=86)**

**b. English Atlas (native English speakers from the Language Atlas (LanA)) (n=629)**

**c. Russian Atlas (native Russian speakers) (n=19)**
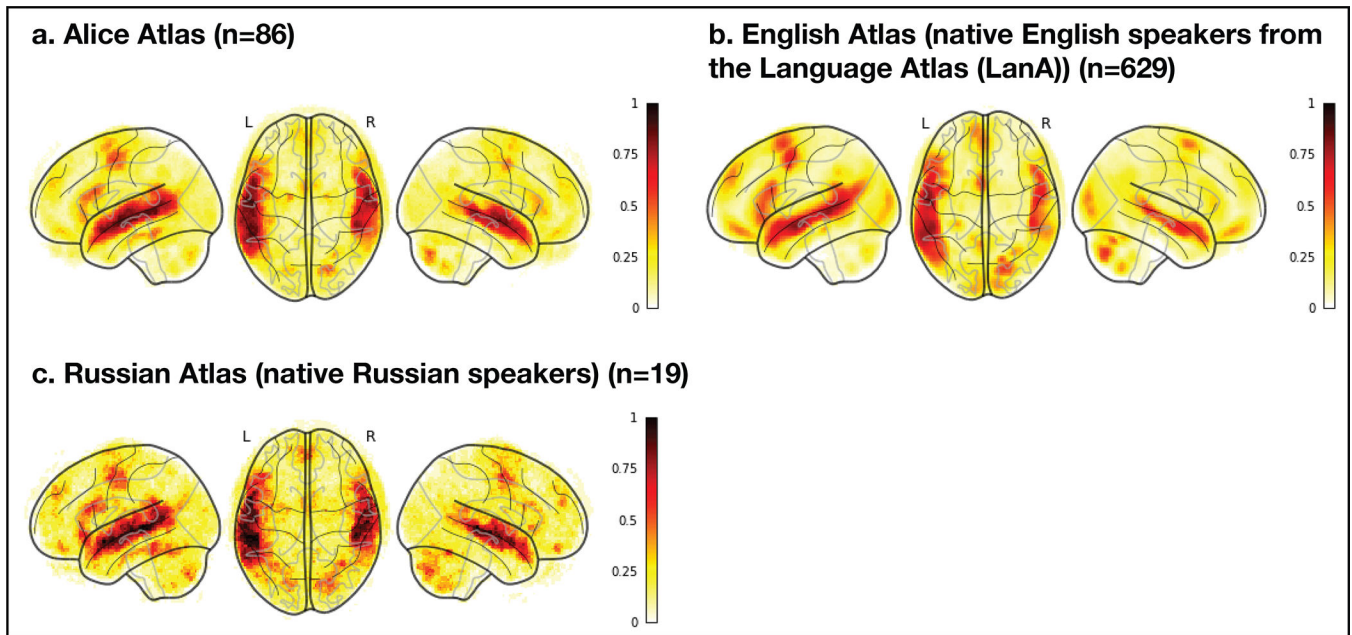
**Extended Data Fig. 8.**

Comparison of three probabilistic overlap maps (atlases). Comparison of three probabilistic overlap maps (atlases): a) the Alice atlas (n=86 native speakers of 45 languages) created from the *Native-language>Degraded-language* maps; b) the English atlas (n=629 native English speakers; this is a subset of the Fedorenko lab's Language Atlas (LanA[56]) created from the *Sentences>Nonwords* maps; and) the Russian Atlas (n=19 native Russian speakers) created from the *Native-language>Degraded-language* maps for the Russian version of the Alice localizer. All three atlases were created by selecting for each participant the top 10% of voxels (across the brain) based on the t-values for the relevant contrast in each participant, binarizing these maps, and then overlaying them in the common space. In each atlas, the value in each voxel corresponds to the proportion of participants (between 0 and 1) for whom that voxel belongs to the 10% of most language-responsive voxels. The probabilistic landscapes are similar across the atlases: within the union of the language parcels (see Extended Data Figure 1 caption for an explanation of why this approach is more conservative than performing the comparison across the brain), the Alice atlas is voxel-wise spatially correlated with both the English atlas (r=0.83) and the Russian atlas (r=0.85). Furthermore, the range of non-zero overlap values is comparable between the Alice atlas (0.1–0.87; average within the language parcels=0.08, median=0.05) and each of the other atlases (the English atlas: 0.002–0.79; average within the language parcels=0.07, median=0.03; the Russian atlas: 0.05–0.84; average within the language parcels=0.13, median=0.11). The latter result suggests that the inter-individual variability in the topographies of activation landscapes elicited in 86 participants of 45 diverse languages is comparable to the inter-individual variability observed among native speakers of the same language.

**Extended Data Fig. 9.**

Responses in the domain-general Multiple Demand network to the conditions of the Alice localizer task, the spatial working memory task, and the math task. Percent BOLD signal change across the domain-general Multiple Demand (MD) network[15,52] functional ROIs for the three language conditions of the Alice localizer task (Native language, Acoustically degraded native language, and Unfamiliar language), the hard and easy conditions of the spatial working memory (WM) task, and the hard and easy conditions of the math task. The dots correspond to languages (n=45 except for the Math Task, where n=41). Box plots include the first quartile (lower hinge), third quartile (upper hinge), and median (central line); upper and lower whiskers extend from the hinges to the largest value no further than 1.5 times the inter-quartile range; darker-colored dots correspond to outlier data points. As in the main analyses (Figure 3c), the individual MD fROIs were defined by the *Hard>Easy* contrast in the spatial WM task (see [54] for evidence that other *Hard>Easy* contrasts activate similar areas). As expected given past work[e.g., 54], the MD fROIs show strong responses to both the spatial WM task and the math task, with stronger responses to the harder condition in each (3.05 vs. 1.93 for the spatial WM task, t(44)=23.1, p<0.001; and 1.68 vs. 0.62 for the math task, t(40)=8.87, p<0.001). These robust responses in the MD network suggest that the lack of responses to the spatial WM and math tasks in the language areas can be meaningfully interpreted. Furthermore, in line with past work[e.g., 58–60], MD fROIs show a stronger response to the acoustically degraded condition than the native language condition (0.26 vs. −0.10, t(44)=4.92, p<0.01), and to the unfamiliar language condition than the native language condition (0.15 vs. −0.10, t(44)=4.96, p<0.01). All t-tests were two-tailed with no adjustment for multiple comparisons.
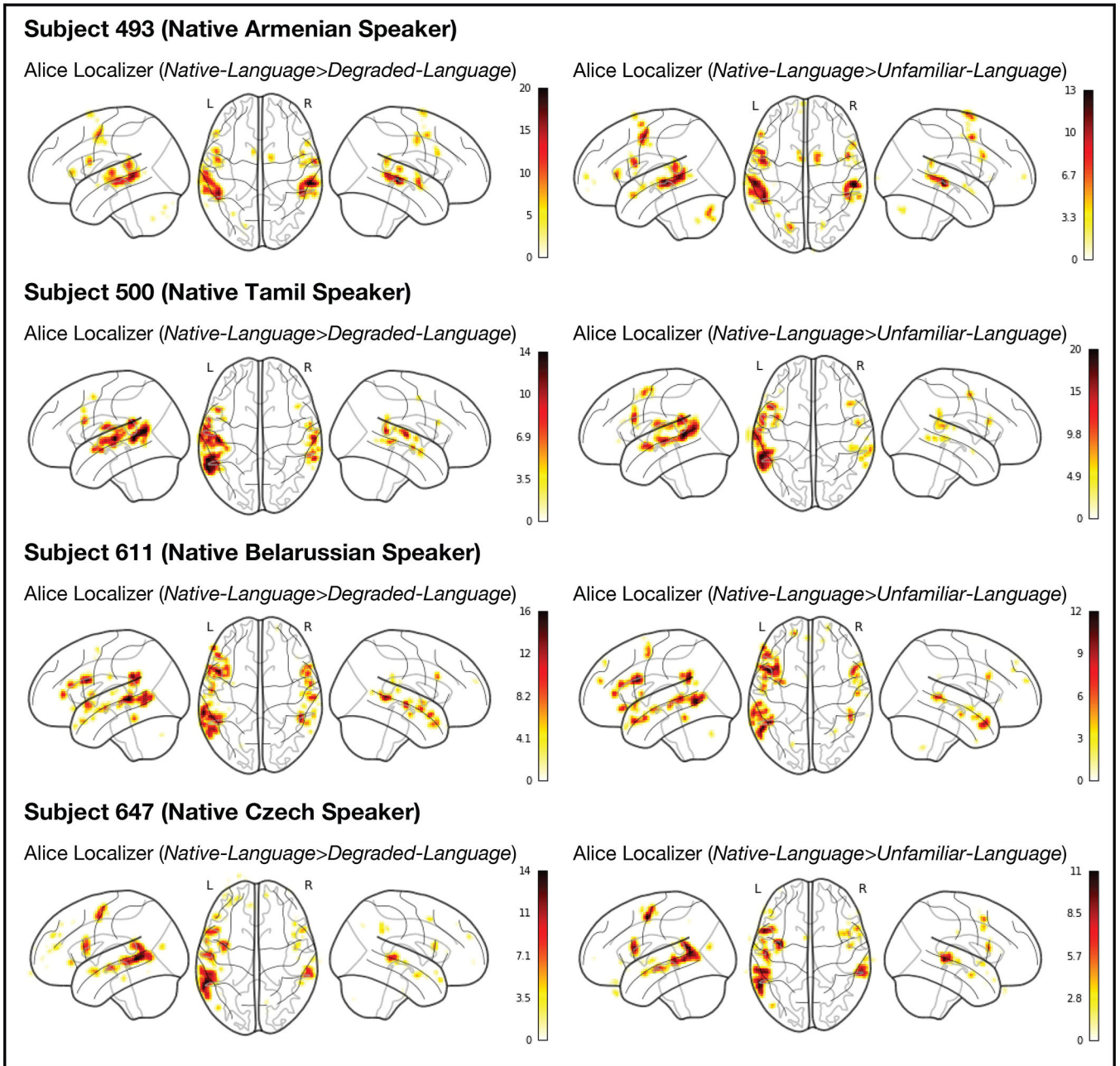
**Extended Data Fig. 10.**

Comparison of the individual activation maps for the *Native-language>Degraded-language* contrast and the *Native-language>Unfamiliar-language* contrast in four sample participants. The activation landscapes are broadly similar: across the full set of 86 participants, the average Fisher-transformed voxel-wise spatial correlation within the union of the language parcels between the maps for the two contrasts is r=0.66 (SD=0.40). (Note that this correlation is lower than the correlation between the *Native-language>Degraded-language* contrast and the *Sentences>Nonwords* contrast in English (see Extended Data Figure 1). This difference may be due to the greater variability in the participants' responses to an unfamiliar language.) Furthermore, across the language fROIs, the magnitudes of the

*Native-language>Degraded-language* and the *Native-language>Unfamiliar-language* effects are similar (mean = 1.02, SD(across languages)=0.41 vs. mean=1.07, SD=0.37, respectively; t(44)=1.15, p=0.26).

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

## Data availability:

The data that support the findings of this study are available at: https://osf.io/cw89s.

## References

1. Lewis MP Ethnologue: Languages of the world. (International, SIL, 2009).

2. Chomsky N Knowledge of language: Its nature, origin, and use. (Greenwood Publishing Group, 1986).

3. Gibson E et al. How efficiency shapes human language. Trends Cogn. Sci. (2019).

4. Evans N & Levinson SC The myth of language universals: Language diversity and its importance for cognitive science. Behav. Brain Sci. 32, 429–448 (2009). [PubMed: 19857320]

5. Bates E, McNew S, MacWhinney B, Devescovi A & Smith S Functional constraints on sentence processing: A cross-linguistic study. Cognition 11, 245–299 (1982). [PubMed: 7199413]

6. Bornkessel-Schlesewsky I & Schlesewsky M The importance of linguistic typology for the neurobiology of language. Linguist. Typology 3, (2016).

7. Hudley AHC, Mallinson C & Bucholtz M Toward Racial Justice in Linguistics: Interdisciplinary Insights into Theorizing Race in the Discipline and Diversifying the Profession. (2020).

8. Rueckl JG et al. Universal brain signature of proficient reading: Evidence from four contrasting languages. Proc. Natl. Acad. Sci. U. S. A. 112, 15510–15515 (2015). [PubMed: 26621710]

9. Fedorenko E, Hsieh P-J, Nieto-Castañón A, Whitfield-Gabrieli S & Kanwisher N New Method for fMRI Investigations of Language: Defining ROIs Functionally in Individual Subjects. J. Neurophysiol. 104, 1177–1194 (2010). [PubMed: 20410363]

10. Mahowald K & Fedorenko E 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 (2016). [PubMed: 27261158]

11. Fedorenko E, Behr MK & Kanwisher N Functional specificity for high-level linguistic processing in the human brain. Proc. Natl. Acad. Sci. 108, 16428–16433 (2011). [PubMed: 21885736]

12. Scott TL, Gallée J & Fedorenko E A new fun and robust version of an fMRI localizer for the frontotemporal language system. Cogn. Neurosci. 8, 167–176 (2017). [PubMed: 27386919]

13. Blank IA, Kanwisher N & Fedorenko E A functional dissociation between language and multiple-demand systems revealed in patterns of BOLD signal fluctuations. J. Neurophysiol. 112, 1105–1118 (2014). [PubMed: 24872535]

14. Fedorenko E & Blank IA Broca's Area Is Not a Natural Kind. Trends in Cognitive Sciences vol. 24 270–284 (2020). [PubMed: 32160565]

15. Duncan J The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. Trends in Cognitive Sciences vol. 14 172–179 (2010). [PubMed: 20171926]

16. Gurunandan K, Arnaez-Telleria J, Carreiras M & Paz-Alonso PM Converging Evidence for Differential Specialization and Plasticity of Language Systems. J. Neurosci. 40, 9715–9724 (2020). [PubMed: 33168623]

17. Nieto-Castañón A & Fedorenko E Subject-specific functional localizers increase sensitivity and functional resolution of multi-subject analyses. Neuroimage 63, 1646–1669 (2012). [PubMed: 22784644]

18. Bornkessel-Schlesewsky I et al. Think globally: Cross-linguistic variation in electrophysiological activity during sentence comprehension. Brain Lang. 117, 133–152 (2011). [PubMed: 20970843]

19. Bickel B, Witzlack-Makarevich A, Choudhary KK, Schlesewsky M & Bornkessel-Schlesewsky I The Neurophysiology of Language Processing Shapes the Evolution of Grammar: Evidence from Case Marking. PLoS One 10, e0132819 (2015). [PubMed: 26267884]

20. Kemmerer D Do language-specific word meanings shape sensory and motor brain systems? the relevance of semantic typology to cognitive neuroscience. Linguist. Typology 20, 623–634 (2016).

21. Albert ML Auditory sequencing and left cerebral dominance for language. Neuropsychologia 10, 245–248 (1972). [PubMed: 5055231]

22. Grafton ST, Hazeltine E & Ivry RB Motor sequence learning with the nondominant left hand. A PET functional imaging study. Exp. brain Res. 146, 369–378 (2002). [PubMed: 12232693]

23. Bornkessel-Schlesewsky I, Schlesewsky M, Small SL & Rauschecker JP Neurobiological roots of language in primate audition: common computational properties. Trends Cogn. Sci. 19, 142–150 (2015). [PubMed: 25600585]

24. Norman-Haignere S, Kanwisher N & McDermott JH Cortical Pitch Regions in Humans Respond Primarily to Resolved Harmonics and Are Located in Specific Tonotopic Regions of Anterior Auditory Cortex. J. Neurosci. 33, 19451–19469 (2013). [PubMed: 24336712]

25. Li Y, Tang C, Lu J, Wu J & Chang EF Human cortical encoding of pitch in tonal and non-tonal languages. Nat. Commun. 2021 121 12, 1–12 (2021).

26. Gil D Riau Indonesian: a language without nouns and verbs. Flex. Word Classes 89–130 (2013) doi:10.1093/ACPROF:OSO/9780199668441.003.0004.

27. Beeman M Semantic processing in the right hemisphere may contribute to drawing inferences from discourse. Brain Lang. 44, 80–120 (1993). [PubMed: 8467379]

28. Saxe R & Kanwisher N People thinking about thinking people: The role of the temporo-parietal junction in 'theory of mind'. Neuroimage 19, 1835–1842 (2003). [PubMed: 12948738]

29. Radford A et al. Language Models are Unsupervised Multitask Learners.

30. Schrimpf M et al. The neural architecture of language: Integrative modeling converges on predictive processing. Proc. Natl. Acad. Sci. 118, e2105646118 (2021). [PubMed: 34737231]

31. Bender EM Linguistically NäiveNäive != Language Independent: Why NLP Needs Linguistic Typology. in Proceedings of the EACL 2009 Workshop on the Interaction between Linguistics and Computational Linguistics: Virtuous, Vicious or Vacuous? 26–32 (2009).

32. Chi EA, Hewitt J & Manning CD Finding Universal Grammatical Relations in Multilingual BERT. 5564–5577 (2020) doi:10.18653/v1/2020.acl-main.493.

33. Kovelman I, Baker SA & Petitto LA Bilingual and monolingual brains compared: A functional magnetic resonance imaging investigation of syntactic processing and a possible 'neural signature' of bilingualism. J. Cogn. Neurosci. 20, 153–169 (2008). [PubMed: 17919083]

34. Costa A & Sebastián-Gallés N How does the bilingual experience sculpt the brain? Nature Reviews Neuroscience vol. 15 336–345 (2014). [PubMed: 24739788]

35. Oldfield RC The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9, 97–113 (1971). [PubMed: 5146491]

36. Braga RM, DiNicola LM, Becker HC & Buckner RL Situating the left-lateralized language network in the broader organization of multiple specialized large-scale distributed networks. J. Neurophysiol. 124, 1415–1448 (2020). [PubMed: 32965153]

37. Fedorenko E, Nieto-Castanon A & Kanwisher N Lexical and syntactic representations in the brain: an fMRI investigation with multi-voxel pattern analyses. Neuropsychologia 50, 499–513 (2012). [PubMed: 21945850]

38. Blank I, Balewski Z, Mahowald K & Fedorenko. E Syntactic processing is distributed across the language system. Neuroimage 127, 307–323 (2016). [PubMed: 26666896]

39. Fedorenko E, Blank IA, Siegelman M & Mineroff Z Lack of selectivity for syntax relative to word meanings throughout the language network. Cognition 203, (2020).

40. Chen X et al. The human language system does not support music processing. bioRxiv.

41. Carroll L Alice's adventures in wonderland. (Broadview Press, 2011).

42. Lindseth J & Tannenbaumxf A Alice in a World of Wonderlands: The Translations of Lewis Carroll's Masterpiece. (Oak Knoll Press, 2015).

43. Wolff P Observations on the early development of smiling. Determ. infant Behav. 2, 113–138 (1963).

44. Thesen S, Heid O, Mueller E & Schad LR Prospective acquisition correction for head motion with image-based tracking for real-time fMRI. Magn. Reson. Med. 44, 457–465 (2000). [PubMed: 10975899]

45. Whitfield-Gabrieli S & Nieto-Castanon A Conn : A Functional Connectivity Toolbox for Correlated and Anticorrelated Brain Networks. Brain Connect. 2, 125–141 (2012). [PubMed: 22642651]

46. Behzadi Y, Restom K, Liau J & Liu TT A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. Neuroimage 37, 90–101 (2007). [PubMed: 17560126]

47. Cordes D et al. Frequencies contributing to functional connectivity in the cerebral cortex in &quot;resting-state&quot; data. AJNR. Am. J. Neuroradiol. 22, 1326–33 (2001). [PubMed: 11498421]

48. Dale A, Fischl B & Sereno M Cortical surface-based analysis: I. Segmentation and surface reconstruction. Neuroimage 179–194 (1999). [PubMed: 9931268]

49. Kriegeskorte N, Simmons WK, Bellgowan PS & Baker CI Circular inference in neuroscience: The dangers of double dipping. J. Vis. 8, 88–88 (2010).

50. Silver NC & Dunlap WP Averaging correlation coefficients: Should Fisher's z transformation be used? J. Appl. Psychol 72, 146–148 (1987).

51. Seghier ML Laterality index in functional MRI: methodological issues. Magn. Reson. Imaging 26, 594–601 (2008). [PubMed: 18158224]

52. Duncan J The Structure of Cognition: Attentional Episodes in Mind and Brain. Neuron 80, 35–50 (2013). [PubMed: 24094101]

53. Paunov AM, Blank IA & Fedorenko E Functionally distinct language and Theory of Mind networks are synchronized at rest and during language comprehension. J. Neurophysiol. 121, 1244–1265 (2019). [PubMed: 30601693]

54. Fedorenko E, Duncan J & Kanwisher N Broad domain generality in focal regions of frontal and parietal cortex. Proc. Natl. Acad. Sci. 110, 16616–16621 (2013). [PubMed: 24062451]

55. Tzourio-Mazoyer N et al. Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. Neuroimage 15, 273–289 (2002). [PubMed: 11771995]

56. Lipkin B et al. LanA (Language Atlas): A probabilistic atlas for the language network based on fMRI data from &gt;800 individuals. bioRxiv 2022.03.06.483177 (2022) doi:10.1101/2022.03.06.483177.

57. Rombouts SARB et al. Test-retest analysis with functional MR of the activated area in the human visual cortex. Am Soc Neuroradiol. 18, 195–6108 (1997).

58. Davis M & Johnsrude I Hierarchical processing in spoken language comprehension. J. Neurosci. 23, 3423–3431 (2003). [PubMed: 12716950]

59. Hervais-Adelman AG, Carlyon RP, Johnsrude IS & Davis MH Brain regions recruited for the effortful comprehension of noise-vocoded words. Taylor Fr. 27, 1145–1166 (2012).

60. Erb J, Henry M, Eisner F, Neuroscience, J. O.-J. of & 2013, undefined. The brain dynamics of rapid perceptual adaptation to adverse listening conditions. Soc Neurosci. (2013) doi:10.1523/JNEUROSCI.4596-12.2013.
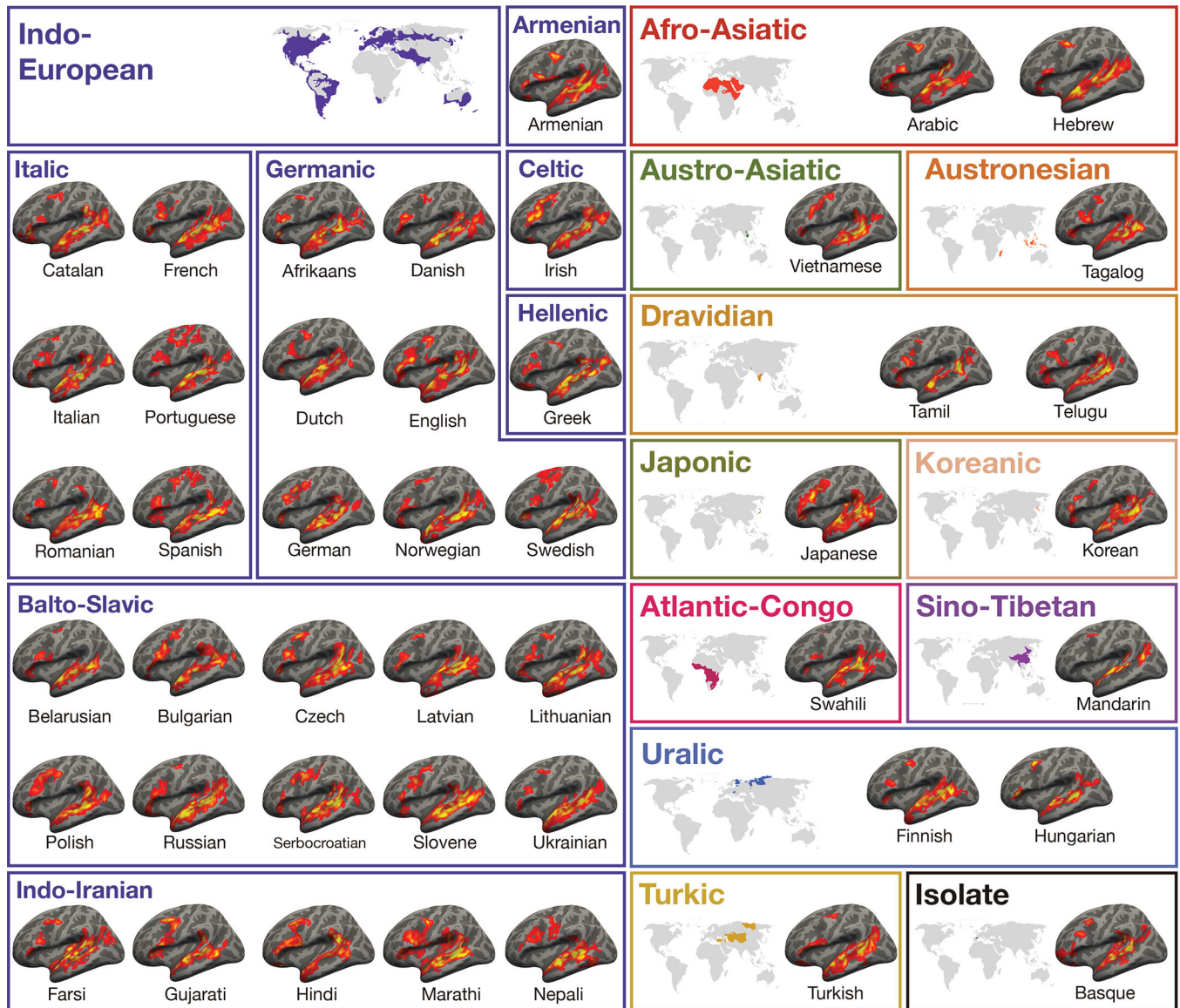
Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Figure 1.**

Activation maps for the Alice language localizer contrast (*Native-language>Degraded-language*) in the left hemisphere (LH) of a sample participant for each language (see Extended Data Figure 2 for RH maps and details of the image generation procedure). The general topography of the language network in speakers of 45 languages is similar, and the variability observed is comparable to the variability that has been reported for the speakers of the same language[10] (Extended Data Figure 8). A significance map was generated for each participant by FreeSurfer; each map was smoothed using a Gaussian kernel of 4 mm full-width half-max and thresholded at the 70th percentile of the positive contrast for each participant. The surface overlays were rendered on the 80% inflated white-gray matter boundary of the fsaverage template using FreeView/FreeSurfer. Opaque red and yellow correspond to the 80th and 99th percentile of positive-contrast activation for each

subject, respectively. (These maps were used solely for visualization; all the analyses were performed on the data analyzed in the volume (see Extended Data Figure 3).)
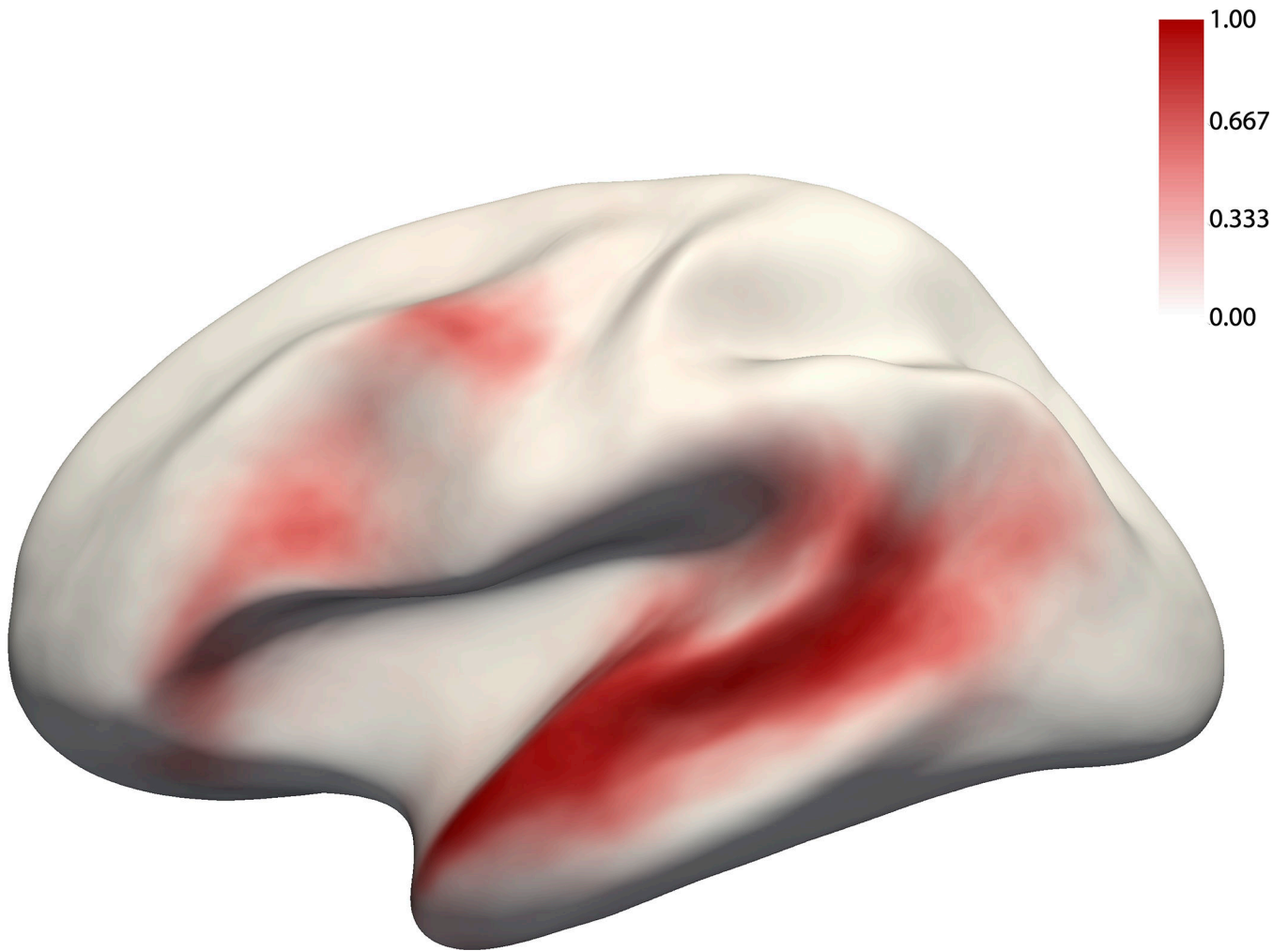
**Figure 2.**
The probabilistic overlap map for the *Native-language>Degraded-language* contrast. This map was created by binarizing and overlaying the 86 participants' individual maps (like those shown in Figure 1). The value in each vertex corresponds to the proportion of participants for whom that vertex belongs to the language network (see Extended Data Figure 8 for a comparison between this probabilistic atlas vs. atlases based on native speakers of the same language).
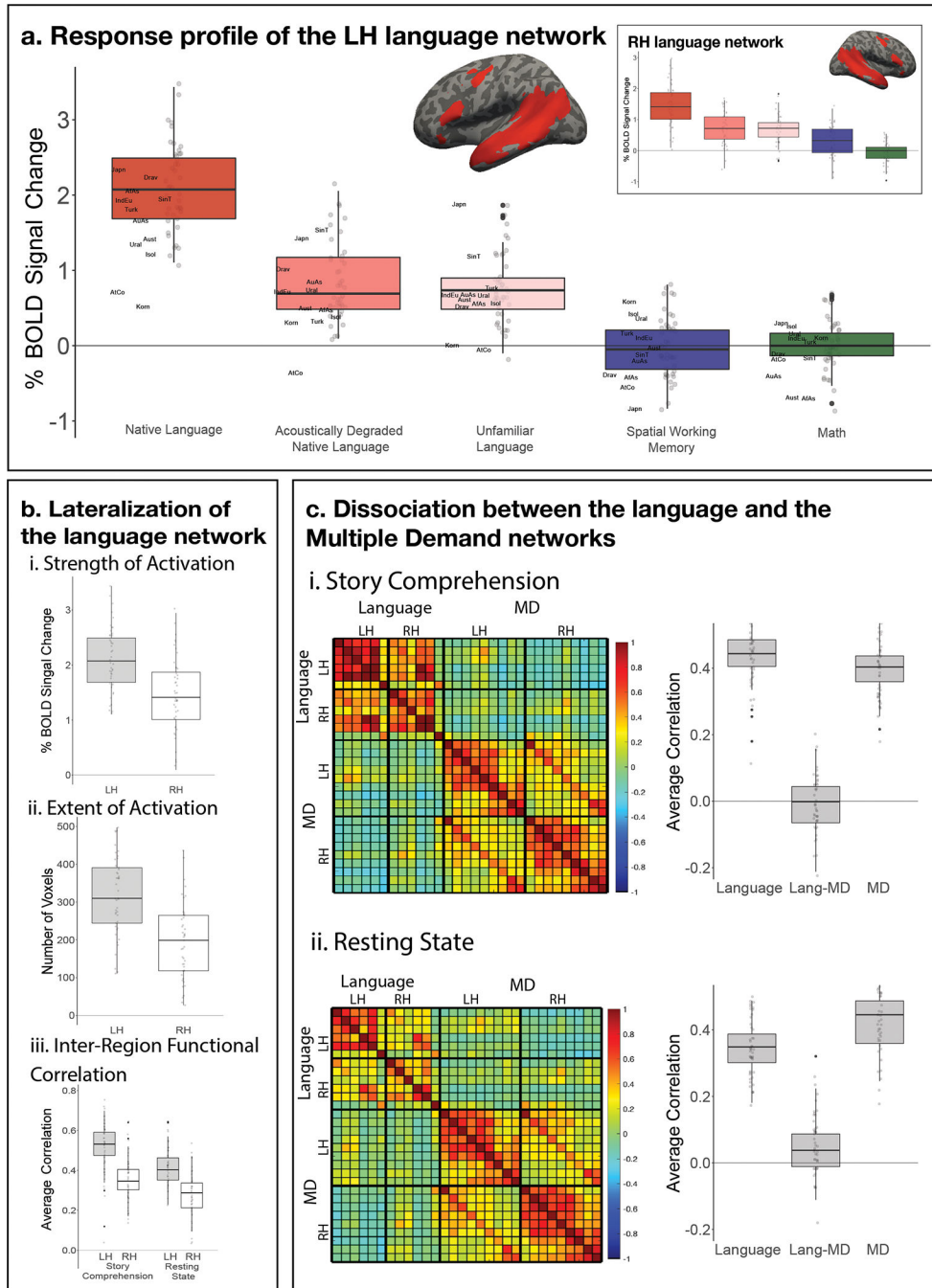
**Figure 3.**

a) Percent BOLD signal change across the LH language functional ROIs (see inset for the RH language fROIs) for the three language conditions of the Alice localizer task (Native language, Acoustically degraded native language, and Unfamiliar language), the spatial working memory (WM) task, and the math task. The language fROIs show robust functional selectivity for language processing. Here and in the other panels, the dots correspond to languages (n=45 in all panels), and the labels mark the averages for each language family (n=12; AfAs=Afro-Asiatic, AuAs=Austro-Asiatic, Aust=Austronesian, Drav=Dravidian,

IndEu=Indo-European, Japn=Japonic, Korn=Koreanic, AtCo=Atlantic-Congo, SinT=Sino-Tibetan, Turk=Turkic, Ural=Uralic, Isol=Isolate). Here and in other panels, box plots include the first quartile (lower hinge), third quartile (upper hinge), and median (central line); upper and lower whiskers extend from the hinges to the largest value no further than 1.5 times the inter-quartile range; darker-colored dots correspond to outlier data points. For each statistical comparison reported in the text, a two-tailed t-test was used (see Suppl. Table 2 for results of linear mixed effects models); no correction for the number of comparisons was applied (because each test addressed a distinct question). b) Three measures that reflect LH lateralization of the language network: i-strength of activation (effect sizes for the *Native-language>Degraded-language* contrast); ii-extent of activation (number of voxels within the union of the language parcels at a fixed threshold for the *Native-language>Degraded-language* contrast; a whole-brain version of this analysis yielded a similar result: $t(44)=5.79$, $p<0.001$); and iii-inter-region functional correlations during two naturalistic cognition paradigms (i-story comprehension in the participant's native language; ii-resting state). The LH language network shows greater selectivity for language processing relative to a control condition, is more spatially extensive, and is more strongly functionally integrated than the RH language network. c) Inter-region functional correlations for the LH and RH language network and the Multiple Demand (MD) network during two naturalistic cognition paradigms (i-story comprehension in the participant's native language; ii-resting state). The language and the MD networks are each strongly functionally integrated but are robustly dissociated from each other (pairs of fROIs straddling network boundaries show little/no correlated activity).