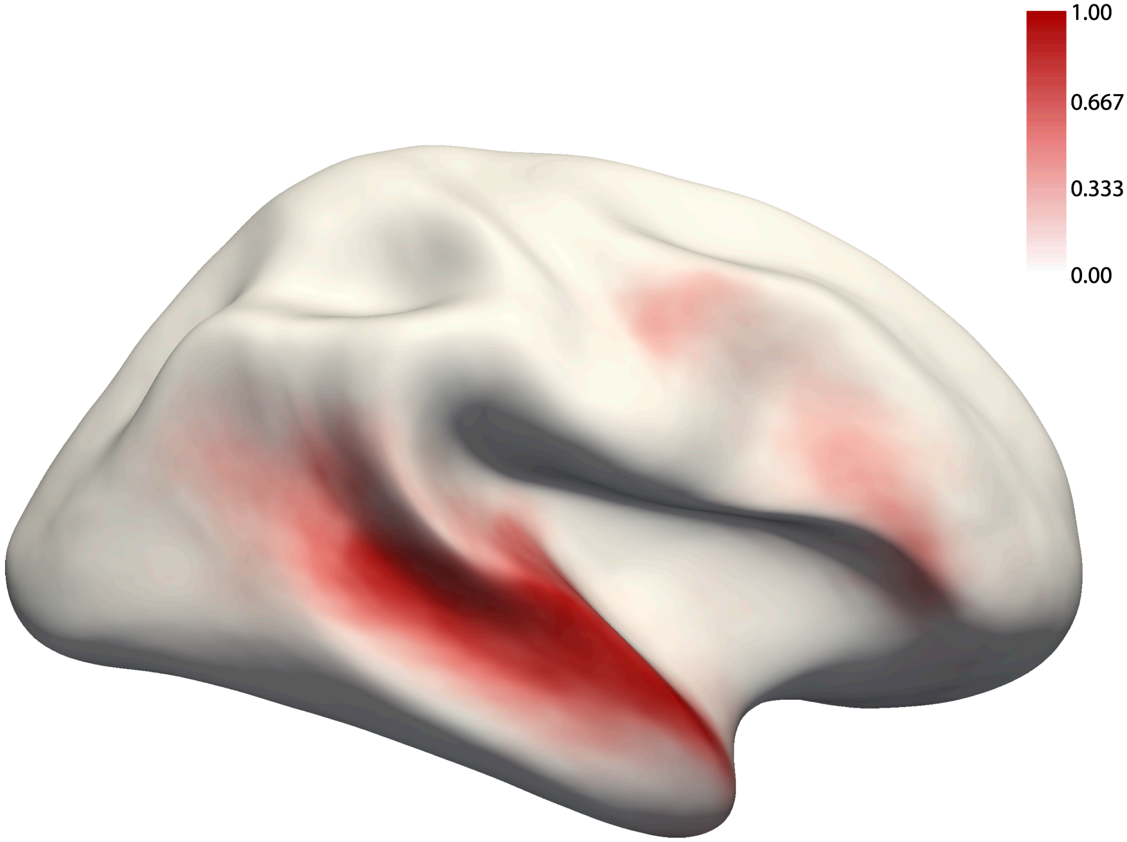
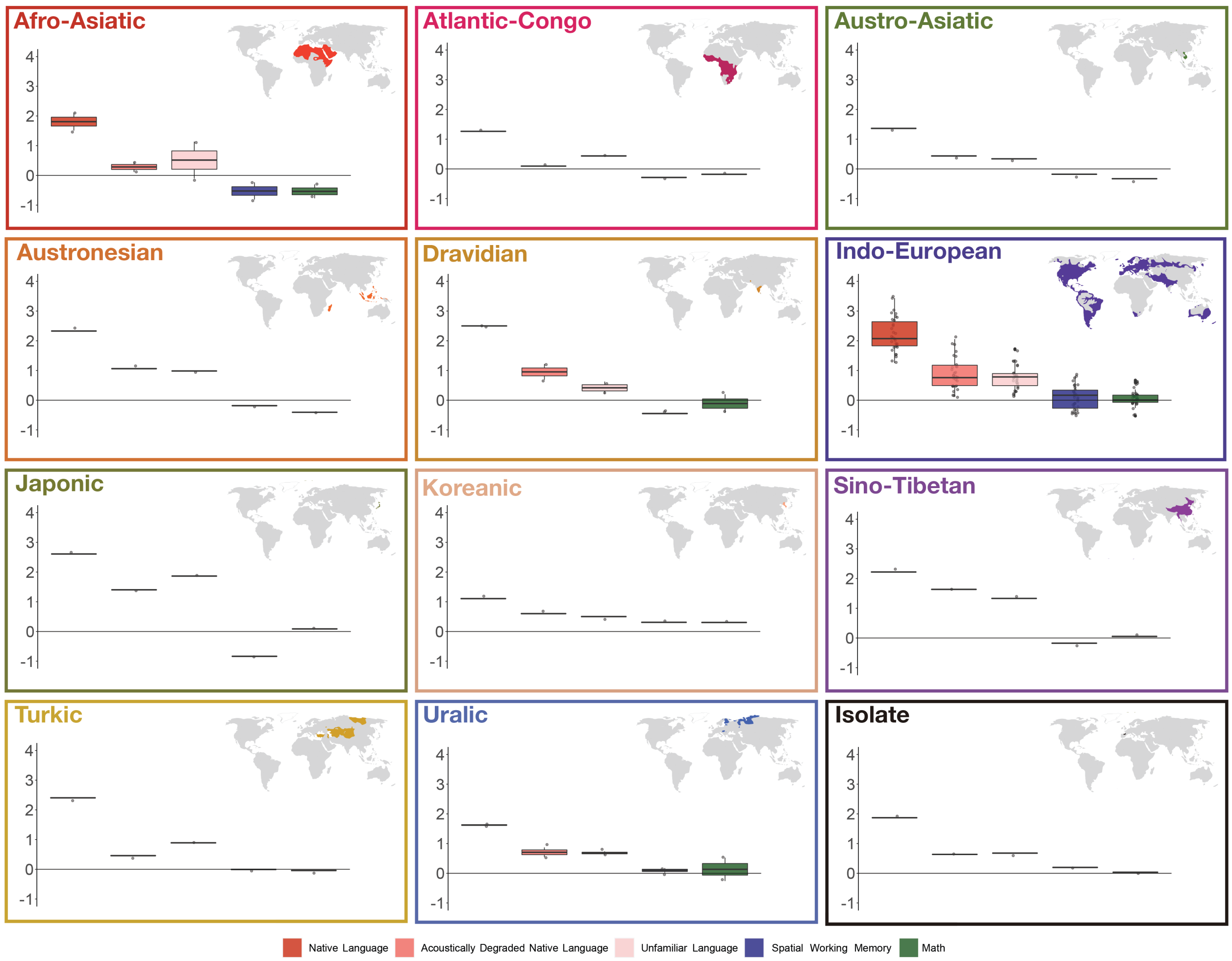
**Supplementary Information for Malik-Moraleda, Ayyash et al. (2022),** ***An investigation across 45 languages and 12 language families reveals a universal language network*****.**

**Table of contents:**

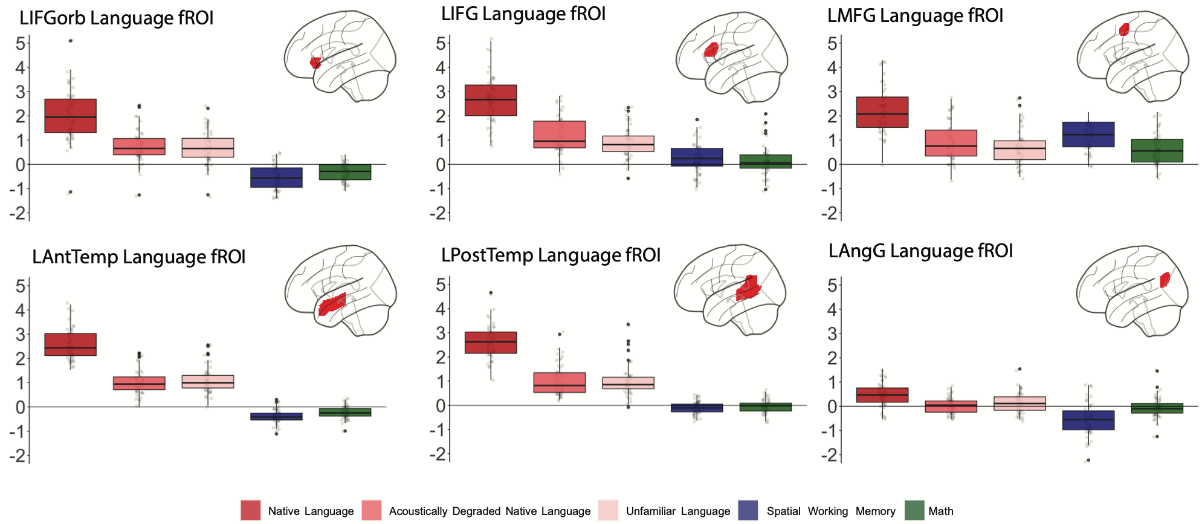
**Supplementary Figures:**



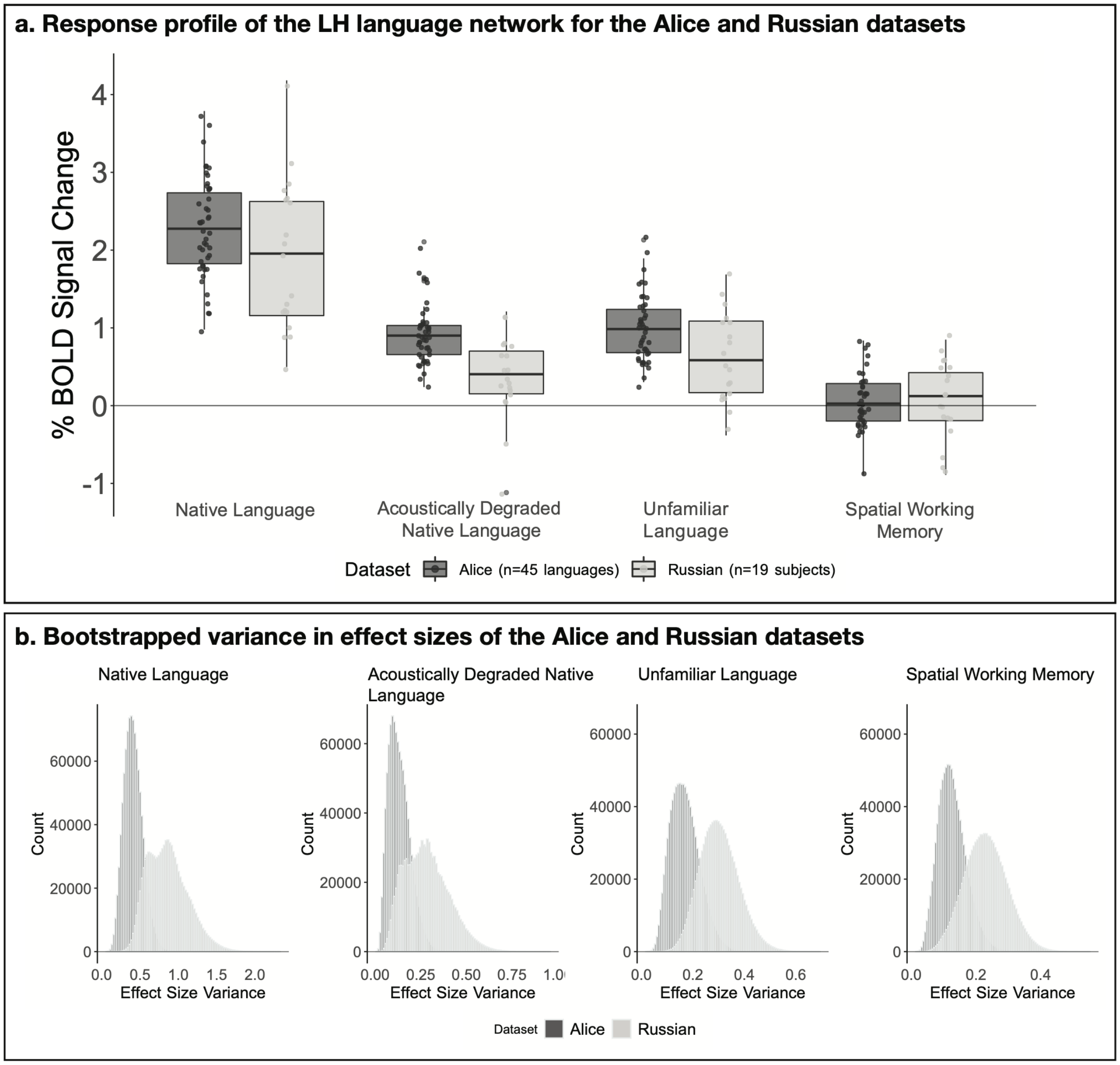

**Supplementary Figure 1:** The probabilistic overlap map for the *Native-language>Degraded-language* contrast for the right hemisphere. This map was created by binarizing and overlaying the 86 participants' individual maps (like those shown in **Extended Data Figure 2**). The value in each vertex corresponds to the proportion of participants for whom that vertex belongs to the language network (see **Extended Figure 8** for a comparison between this probabilistic atlas vs. atlases based on native speakers of the same language).



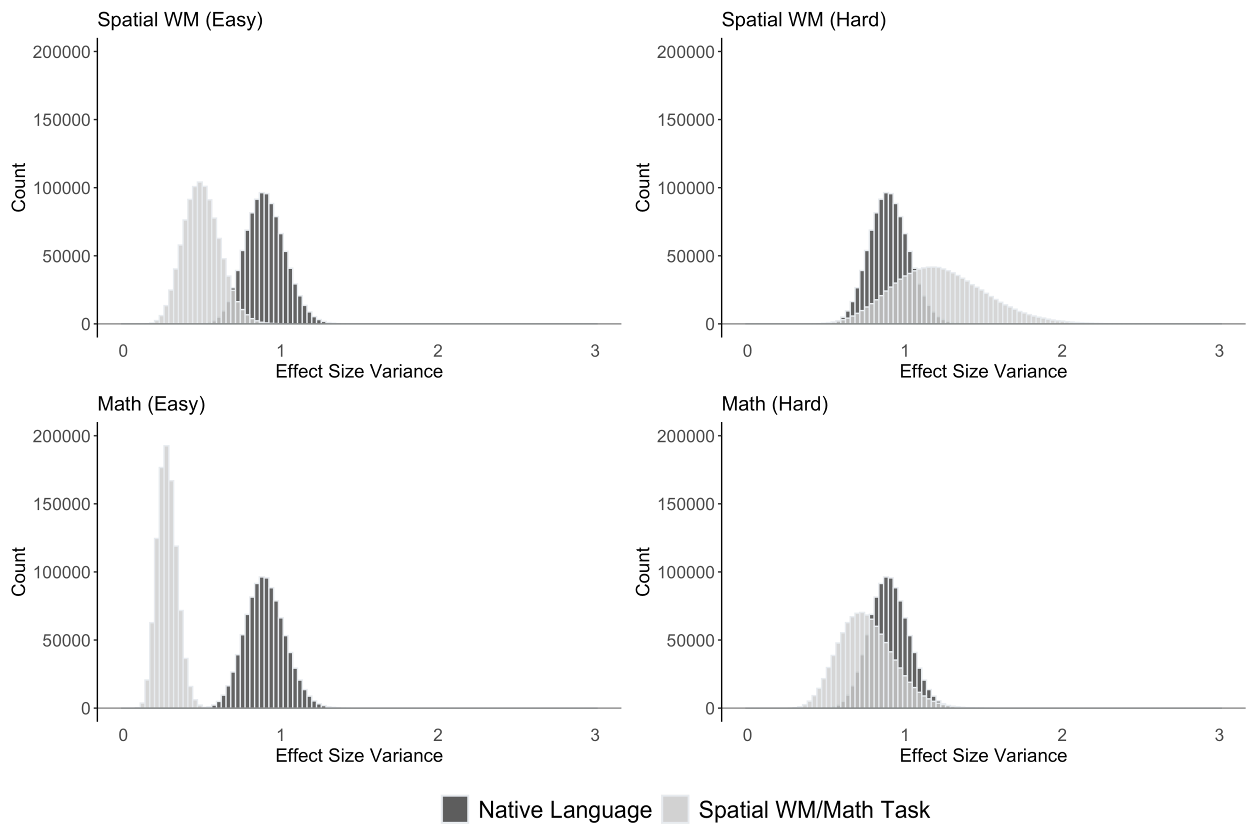

**Supplementary Figure 2:** Percent BOLD signal change across the LH language functional ROIs (defined by the *Sentences>Nonwords* contrast) for the three language conditions of the Alice localizer task (Native language, Acoustically degraded native language, and Unfamiliar language), the spatial working memory (WM) task, and the math task shown for each language family separately. 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 language families (n=12), the *Native-language* condition elicits a reliably greater response than both the *Degraded-language* condition ($t(11)=9.92$, $p<0.001$) and the *Unfamiliar-language* condition ($t(11)=9.53$, $p<0.001$). The *Native-language>Degraded-language* effect is stronger in the left hemisphere than the right hemisphere ($t(11)=3.90$, $p=0.002$), and more spatially extensive ($t(11)=4.01$, $p<0.001$). The regions of the LH language network exhibit strong correlations in their activity during story comprehension and rest, both reliably higher than zero (ts>4, ps<0.001) and phase-shuffled baselines (ts>10, ps<0.001). Further, the inter-region correlations in the LH language network are reliably stronger than those in the RH during both story comprehension ($t(11)=4.06$, $p<0.01$) and rest ($t(11)=4.78$, $p<0.001$). Responses to the *Native-language* condition are significantly higher than those to the spatial working memory task ($t(11)=10.08$, $p<0.001$) and the math task ($t(11)=11.7$, $p<0.001$). Furthermore, the language regions are dissociated in their intrinsic fluctuation patterns from the regions of the MD network: within-network correlations are reliably greater than between-network correlations both during story comprehension (ts>8, ps<0.001) and rest (ts>12, ps<0.001). All t-tests were two-tailed with no adjustment for multiple comparisons.



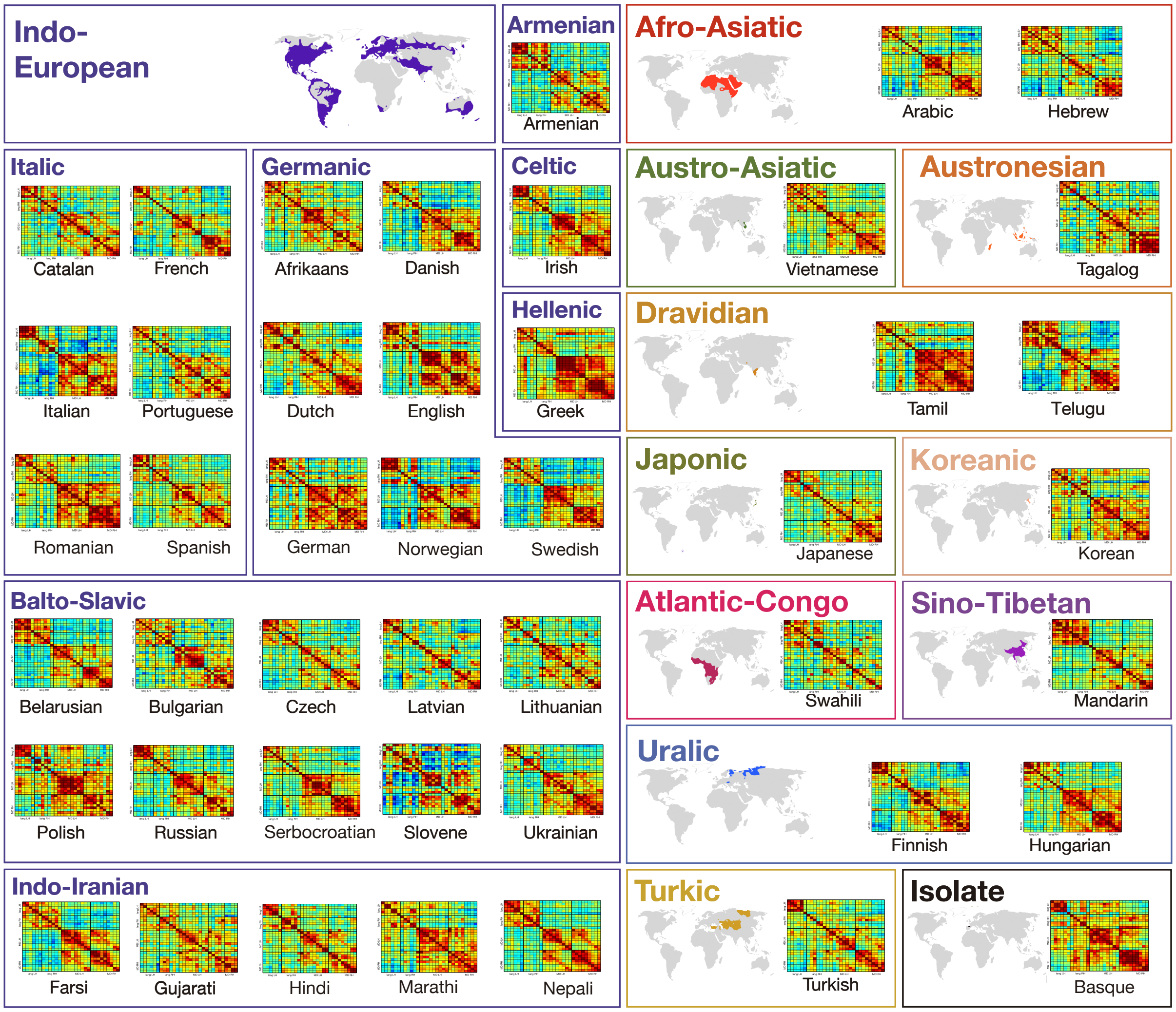

**Supplementary Figure 3:** Percent BOLD signal change for each of the six LH language functional ROIs (defined by the *Sentences>Nonwords* contrast) for the three language conditions of the Alice localizer task (Native language, Acoustically degraded native language, and Unfamiliar language), the spatial working memory task, and the math task. The dots correspond to languages (n=45). 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.



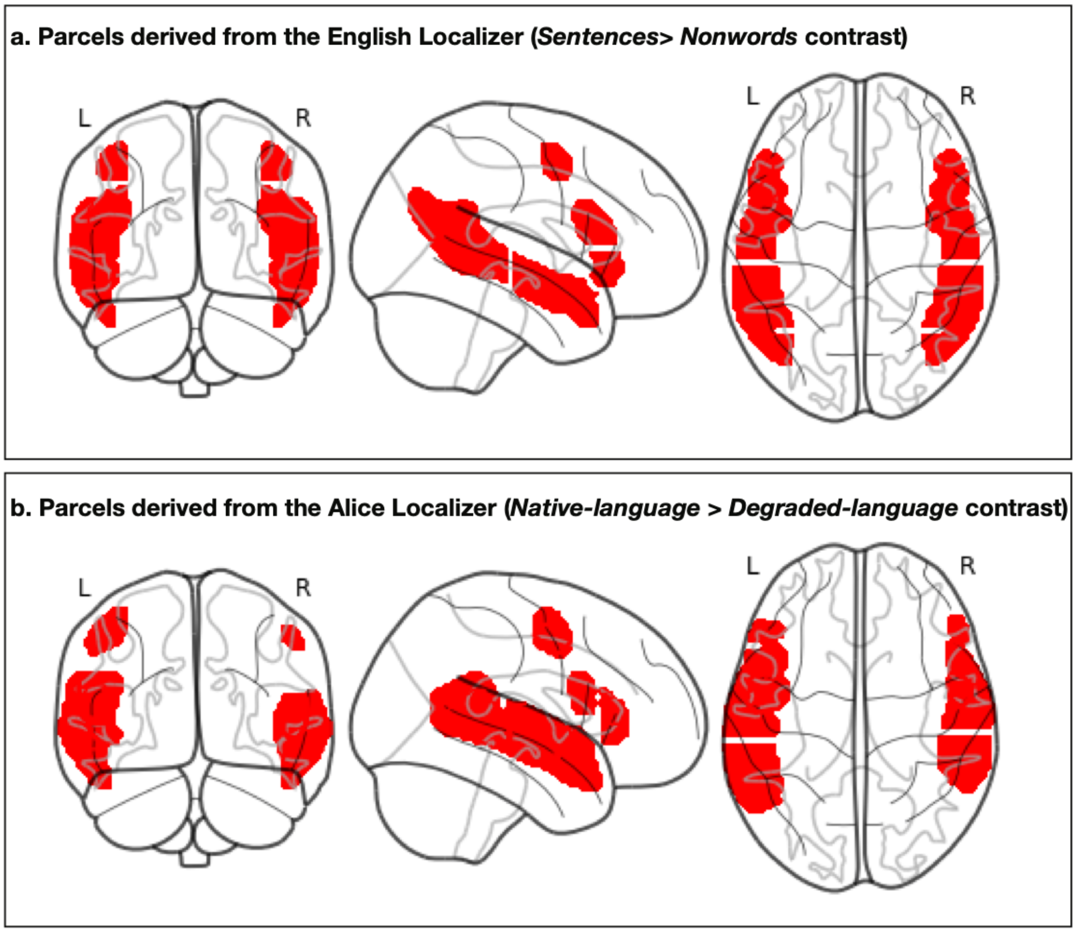

**Supplementary Figure 4**: A comparison of inter-individual variability in effect sizes for one's native language and the control conditions for speakers of diverse languages vs. for speakers of the same language (Russian). As can be seen in Figure 3a in the main text, we observed substantial variability across languages in the strength of neural response during language processing (and the control conditions). In order to compare the level of cross-linguistic variability to inter-individual variability for speakers of the same language[1,2], we leveraged an existing dataset of 19 native speakers of Russian (see also **Extended Data Figure 8**), who completed the Alice localizer (and the spatial working memory task included here for completeness; as in the main paper, we are averaging the responses across the hard and easy conditions). a) Percent BOLD signal change across the LH language functional ROIs (defined by the *Native-language>Degraded-language* contrast) for the three language conditions of the Alice localizer task (Native language, Acoustically degraded native language, and Unfamiliar language), and the spatial working memory (WM) task. Left bars (within each of the four conditions): the current dataset (n=45 languages (1-2 participants per language); dots=languages); right bars: a dataset of n=19 native Russian speakers (unfamiliar language = Tamil) (dots=individual participants). 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. Visual inspection of the distributions of the individual data points suggests that cross-linguistic and inter-individual variability are comparable. b) Bootstrapped variance in effect sizes of the Alice dataset (n=86 participants) and the Russian dataset (n=19 participants) for each of the conditions in the Alice localizer task (Native language, Acoustically degraded native language, and Unfamiliar language). To perform this analysis, we bootstrapped (n=1,000,000) the effect sizes for each of the three conditions for the 86 participants in the Alice dataset (sampling 19 participants at a time) and for the 19 participants in the Russian dataset. If cross-linguistic variability is greater than the variability that exists among individual speakers of the same language, we should see higher bootstrapped variance in the Alice dataset compared to the Russian dataset. Instead, as can be seen in panel b, the bootstrapped variance in the Alice dataset is actually lower than that in the Russian dataset for all conditions. (The reason for higher variability in the Russian dataset may have to do with a wider age range in that group.) As a result, the variability that we observe in the main Figure 3a likely reflects inter-individual rather than cross-linguistic variability. As discussed in the main text, however, future work may discover cross-linguistic differences (when a deep sampling approach is used, with large numbers of speakers tested for each language/language family)—in the measures examined here or some other ones—that would exceed inter-individual variability.



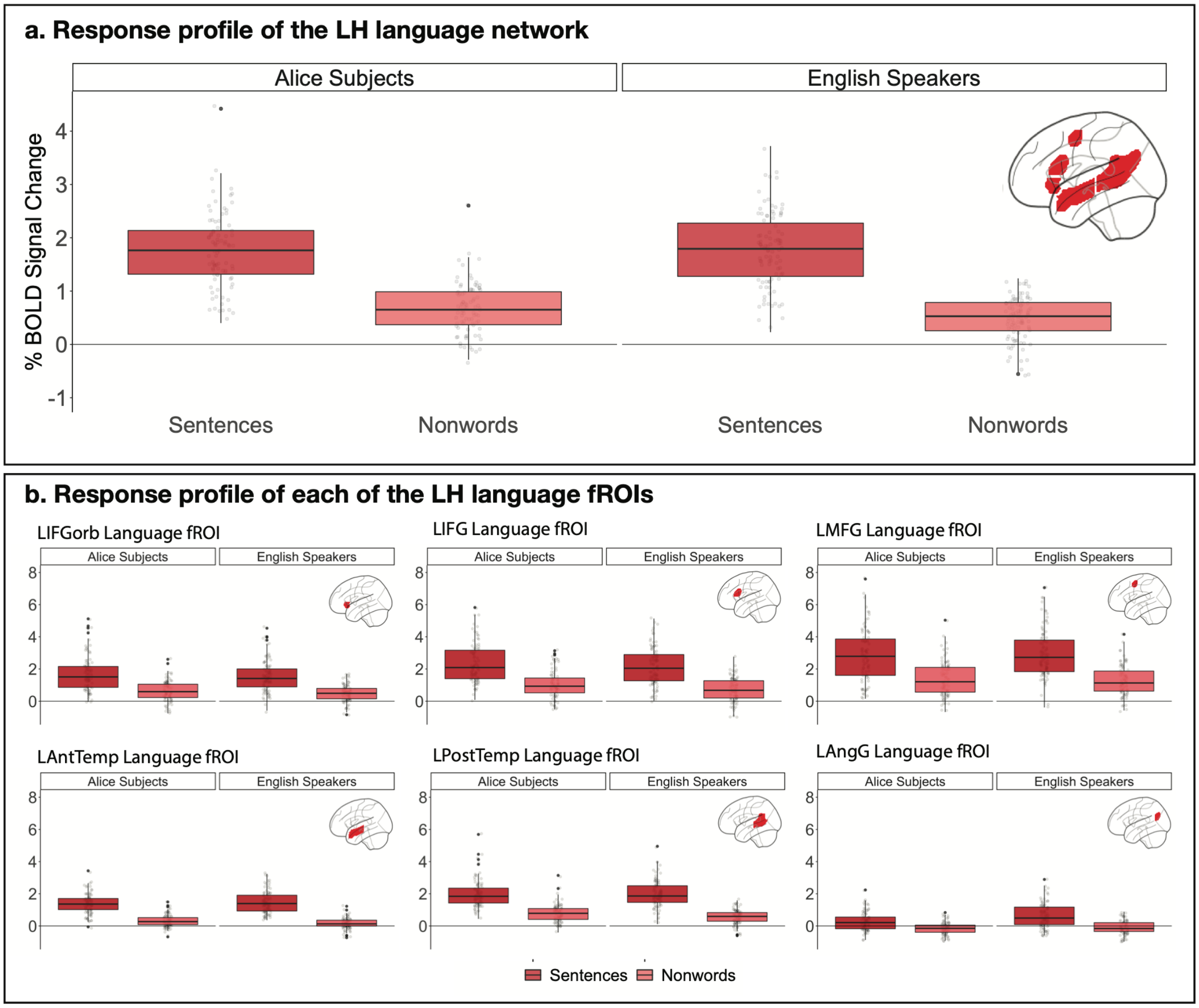

**Supplementary Figure 5**: A comparison of inter-individual variability in effect sizes during language processing and during non-linguistic cognitive tasks for the Alice dataset (n=86 participants). (Note that this between-system comparison is not straightforward because of potential between-system differences in the strength of neural responses, which, to the best of our knowledge, have not been systematically investigated before.) Bootstrapped variance in effect sizes for the Native language condition in the Alice localizer task (dark grey; same distribution across the four panels) and the non-linguistic control task (light grey; top: spatial WM task, bottom: math task; left: easy condition, right: hard condition). To perform this analysis, we bootstrapped (n=1,000,000) the effect sizes in the LH language network for the Native language condition in the Alice localizer task, and in the bilateral MD network for each of the four non-linguistic conditions (which were identical across participants, in contrast to the Alice localizer task, which differed depending on the participant's native language). If cross-linguistic variability is greater than the variability that exists in the strength of neural responses during non-linguistic tasks, we should see higher variance in response to the Native language condition compared to the responses to the different non-linguistic tasks, assuming the effect sizes are comparable (given that variance scales with effect sizes, we would generally expect to see higher variance for larger effects). As the figure shows, the bootstrapped variance for the Native language condition was comparable to the variance in the hard conditions (which elicit a strong response in the MD network), as evidenced by overlapping distributions. For the hard spatial WM condition, the bootstrapped variance is, on average, higher than that for the Native language condition, and for the hard math condition, the bootstrapped variance is, on average, a little lower than for the Native language condition; this pattern argues against uniformly higher variance in the Native language condition than in non-linguistic conditions, and—similar to what we concluded based on Supp. Figure 4—suggests that the variance in the Native language condition is likely due to inter-individual rather than cross-linguistic variability. (For the easy conditions—which elicit a relatively lower response in the MD network—the bootstrapped variance for the Native language condition was higher, as would be expected given the relatively lower response to the easy conditions in the MD network (compared to the response to the Native language condition in the language network).)




**Supplementary Figure 6:** Inter-region functional correlations for the LH and RH of the language and the Multiple Demand (MD) networks during a naturalistic cognition paradigm (resting state) shown for each language separately.




**Supplementary Figure 7:** A visual comparison of the parcels that are used in the current study (derived via a Group-constrained Subject-Specific (GSS) approach[3] from the probabilistic overlap map for the *Sentences>Nonwords* contrast in n=220 independent participants), and the parcels derived (also via GSS) from the probabilistic overlap map for the *Native-language>Degraded-language* contrast in the participants (n=86) in the current study. (Although the temporal-lobe parcels for the latter extend somewhat more superiorly, the *fROIs* selected based on contrasts between language and a perceptually-matched control condition—i.e., contrasts that target high-level language processing—are ~identical for visual and auditory contrasts[4].)




**Supplementary Figure 8:** Percent BOLD signal change across (panel a) and within each of (panel b) the LH language functional ROIs (defined by the *Sentences>Nonwords* contrast; responses were estimated using across-runs cross-validation[5], to ensure independence) for the Sentences and Nonwords conditions. The Alice subjects are the 86 participants from the current study (84 of whom are non-native but proficient speakers of English; we included the two native English speakers here for ease of comparing these results to the results in the rest of the paper where we report the results for the full set of 86 participants); the English speakers are a set of n=74 native English speakers (all learned English before the age of 5). The dots correspond to individual participants. In both 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 LH fROIs, the *Sentences* condition elicits a reliably greater response than the *Nonwords* condition in both the Alice subjects (1.23 vs. 0.49 % BOLD signal change relative to the fixation baseline; $t(85)=20.38$, $p<0.001$) and the native English speakers (1.22 vs. 0.37; $t(73)=18.8$, $p<0.001$). The magnitude of response for the sentences condition is almost identical between the two populations (1.23 vs. 1.22, $t<1$); the magnitude of response for the nonwords condition is a little higher in the Alice subjects (0.49 vs. 0.37; $t(157.36)=2.1$, $p=0.03$). (Because this difference was not predicted, we do not attempt to interpret it.) All t-tests were two-tailed with no adjustment for multiple comparisons. Critically, this supplementary analysis shows that the response during the processing of English is similar between our Alice subjects and a set of native English speakers, and the *Sentences>Nonwords* contrast is similarly robust, suggesting that the use of this contrast as a language localizer is justified (as is also clear from **Extended Data Figure 4**, which shows that similar responses obtain when the fROIs are defined by one's native language localizer).

**Supplementary Tables:**

| **A characterization of the languages included in the current study—as well studied, somewhat studied, or understudied/not studied—with respect to past fMRI work.** | |
|---|---|
| **i. Well Studied Languages (>100 papers per language)** | |
| **Language** | **Sample Citation** |
| Dutch | Snijders, T. M., Vosse, T., Kempen, G., Van Berkum, J. J., Petersson, K. M., & Hagoort, P. (2009). Retrieval and unification of syntactic structure in sentence comprehension: an fMRI study using word-category ambiguity. *Cerebral Cortex*, *19*(7), 1493-1503. |
| English | Fedorenko, E., Behr, M. K., & Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences*, *108*(39), 16428-16433. |
| French | Pallier, C., Devauchelle, A. D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences, 108*(6), 2522-2527. |
| German | Friederici, A. D., Meyer, M., & Von Cramon, D. Y. (2000). Auditory language comprehension: an event-related fMRI study on the processing of syntactic and lexical information. *Brain and Language*, *74*(2), 289-300. |
| Japanese | Kim, J., Koizumi, M., Ikuta, N., Fukumitsu, Y., Kimura, N., Iwata, K., Watanabe, J., Yokoyama, S., Sato, S., Horie, K., & Kawashima, R. (2009). Scrambling effects on the processing of Japanese sentences: An fMRI study. *Journal of Neurolinguistics*, *22*(2), 151-166. |
| Korean | Pallier, C., Dehaene, S., Poline, J. B., LeBihan, D., Argenti, A. M., Dupoux, E., & Mehler, J. (2003). Brain imaging of language plasticity in adopted adults: Can a second language replace the first? *Cerebral Cortex*, *13*(2), 155-161. |
| Mandarin | Chee, M. W., Caplan, D., Soon, C. S., Sriram, N., Tan, E. W., Thiel, T., & Weekes, B. (1999). Processing of visually presented sentences in Mandarin and English studied with fMRI. *Neuron*, *23*(1), 127-137. |
| Spanish | Brignoni-Perez, E., Jamal, N. I., & Eden, G. F. (2020). An fMRI study of English and Spanish word reading in bilingual adults. *Brain and Language*, *202*, 104725. |
| **ii. Somewhat Studied Languages (>10 but <100 papers)** | |
| Arabic | Mohtasib, R. S., Alghamdi, J. S., Baz, S. M., Aljoudi, H. F., Masawi, A. M., & Jobeir, A. A. (2021). Developing fMRI protocol for clinical use Comparison of 6 Arabic paradigms for brain language mapping in native Arabic speakers. *Neurosciences Journal*, *26*(1), 45-55. |
| Basque | Quiñones, I., Amoruso, L., Pomposo Gastelu, I. C., Gil-Robles, S., & Carreiras, M. (2021). What can glioma patients teach us about language (re)organization in the bilingual brain: Evidence from fMRI and MEG. *Cancers*, *13*(11), 2593. |
| Catalan | Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S. F., & Fazio, F. (2003). The role of age of acquisition and language usage in early, high-proficient bilinguals: An fMRI study during verbal fluency. *Human Brain Mapping*, *19*(3), 170-182. |

| | |
|---|---|
| Danish | Buchweitz, A., Mason, R. A., Tomitch, L., & Just, M. A. (2009). Brain activation for reading and listening comprehension: An fMRI study of modality effects and individual differences in language comprehension. *Psychology and Neuroscience, 2*(2), 111-123. |
| Finnish | Hugdahl, K., Thomsen, T., Ersland, L., Rimol, L. M., & Niemi, J. (2003). The effects of attention on speech perception: an fMRI study. *Brain and Language*, *85*(1), 37-48. |
| Greek | Kokkinos, V., Selviaridis, P., & Seimenis, I. (2021). Feasibility, contrast sensitivity and network specificity of language fMRI in presurgical evaluation for epilepsy and brain tumor surgery. *Brain Topography*, *34*(4), 511-524. |
| Hebrew | Bitan, T., Kaftory, A., Meiri-Leib, A., Eviatar, Z., & Peleg, O. (2017). Phonological ambiguity modulates resolution of semantic ambiguity during reading: An fMRI study of Hebrew. *Neuropsychology*, *31*(7), 759. |
| Hindi | Kumar, U., Padakannaya, P., Mishra, R. K., & Khetrapal, C. L. (2013). Distinctive neural signatures for negative sentences in Hindi: an fMRI study. *Brain Imaging and Behavior*, *7*(2), 91-101. |
| Italian | Carota, F., Bozic, M., & Marslen-Wilson, W. (2016). Decompositional representation of morphological complexity: Multivariate fMRI evidence from Italian. *Journal of Cognitive Neuroscience*, *28*(12), 1878-1896. |
| Norwegian | Lehtonen, M. H., Laine, M., Niemi, J., Thomsen, T., Vorobyev, V. A., & Hugdahl, K. (2005). Brain correlates of sentence translation in Finnish–Norwegian bilinguals. *NeuroReport*, *16*(6), 607-610. |
| Polish | Bozic, M., Szlachta, Z., & Marslen-Wilson, W. D. (2013). Cross-linguistic parallels in processing derivational morphology: Evidence from Polish. *Brain and Language*, *127*(3), 533-538. |
| Portuguese | Buchweitz, A., Mason, R. A., Tomitch, L., & Just, M. A. (2009). Brain activation for reading and listening comprehension: An fMRI study of modality effects and individual differences in language comprehension. *Psychology and Neuroscience, 2*(2), 111-123. |
| Russian | Axelrod, V., Bar, M., Rees, G., & Yovel, G. (2015). Neural correlates of subliminal language processing. *Cerebral Cortex*, *25*(8), 2160-2169. |
| Swedish | Ettinger-Veenstra, V., McAllister, A., Lundberg, P., Karlsson, T., & Engström, M. (2016). Higher language ability is related to angular gyrus activation increase during semantic processing, independent of sentence incongruency. *Frontiers in Human Neuroscience*, *10*, 110. |
| **iii. Understudied/Not Studied Languages (<10 papers)** | |
| Afrikaans | Benjamin, C.F., Dhingra, I., Li, A.X., Blumenfeld, H., Alkawadri, R., Bickel, S., Helmstaedter, C., Meletti, S., Bronen, R.A., Warfield, S.K., & Spencer, D. D. (2018). Presurgical language fMRI: Current technical practices in epilepsy surgical planning. *bioRxiv*, 279117. |
| Armenian | No studies found. |
| Belarussian | No studies found. |
| Bulgarian | Kaiser, A., Kuenzli, E., Zappatore, D., & Nitsch, C. (2007). On females' lateral and males' bilateral activation during language production: a fMRI study. *International Journal of Psychophysiology*, *63*(2), 192-198. |

| | |
|---|---|
| Czech | Brázdil, M., Chlebus, P., Mikl, M., Pažourková, M., Krupa, P., & Rektor, I. (2005). Reorganization of language-related neuronal networks in patients with left temporal lobe epilepsy–an fMRI study. *European Journal of Neurology*, *12*(4), 268-275. |
| Farsi | Dehghani, M., Boghrati, R., Man, K., Hoover, J., Gimbel, S.I., Vaswani, A., Zevin, J.D., Immordino-Yang, M.H., Gordon, A.S., Damasio, A., & Kaplan, J. T. (2017). Decoding the neural representation of story meanings across languages. *Human Brain Mapping*, *38*(12), 6096-6106. |
| Gujarati | Gupta, S. S. (2014). fMRI for mapping language networks in neurosurgical cases. *The Indian Journal of Radiology & Imaging*, *24*(1), 37. |
| Hungarian | Kiss, M., Rudas, G., & Kozak, L. R. (2016, March). The outcome of fMRI language mapping is affected by patient fatigue. *European Congress of Radiology*-ECR 2016. |
| Irish | No studies found. |
| Latvian | No studies found. |
| Lithuanian | No studies found. |
| Marathi | No studies found. |
| Nepali | Mu, J., Xie, P., Yang, Z. S., Lu, F. J., Li, Y., & Luo, T. Y. (2006). Functional magnetic resonance image study on the brain areas involved in reading Chinese, English, and Nepali in Nepalese. *Zhong nan da xue xue bao. Yi xue ban.* (*Journal of Central South University. Medical Sciences.)*, *31*(5), 759-762. |
| Romanian | No studies found. |
| Serbocroatian | Progovac, L., Rakhlin, N., Angell, W., Liddane, R., Tang, L., & Ofen, N. (2018). Diversity of grammars and their diverging evolutionary and processing paths: evidence from functional MRI study of Serbian. *Frontiers in Psychology*, *9*, 278. |
| Slovene | Benjamin, C.F., Dhingra, I., Li, A.X., Blumenfeld, H., Alkawadri, R., Bickel, S., Helmstaedter, C., Meletti, S., Bronen, R.A., Warfield, S.K., & Spencer, D. D. (2018). Presurgical language fMRI: Current technical practices in epilepsy surgical planning. *bioRxiv*, 279117. |
| Swahili | No studies found. |
| Tagalog | No studies found. |
| Telugu | Agrawal, A., Hari, K. V. S., & Arun, S. P. (2018). How does reading expertise influence letter representations in the brain? An fMRI study. *Journal of Vision*, *18*(10), 1161-1161. |
| Tamil | No studies found. |
| Turkish | No studies found. |
| Ukranian | No studies found. |
| Vietnamese | No studies found. |

**Supplementary Table 1:** A partial selective review of past fMRI studies on the languages included in the current investigation. For each language, SMM performed searches (on Google, GoogleScholar, PubMed, etc.) for “fMRI [language]” (e.g., fMRI Ukranian) and extracted the relevant citations where available. All papers dealing with speech (perception and articulation), reading, and language (comprehension and production) were considered (i.e., we did not restrict our search to only papers that focus on high-level linguistic processing). Further, we included papers from the clinical literature (that simply used the language in question to facilitate pre-surgical planning rather than asking scientific questions about the particular language or language processing mechanisms in general) and papers where the language in question was used as a control condition. We classified languages into three groups: well-studied languages (with more than100 papers per language), somewhat studied languages (with more than 10 but fewer than 100 papers), and understudied / not studied languages (with fewer than 10 papers, several with not a single paper that we could find; note that for some of these, there exist EEG/MEG studies). This table is not meant to serve as a comprehensive literature review, but to highlight the fact that for many, especially non-‘dominant’, languages, no fMRI investigations have been conducted, and if they have been, they tend to be clinical in nature (e.g., developing tools for pre-surgical mapping), to use the language as a control condition, and/or to be published in lower-impact journals.

| | Effect Significance | Participant Variance | Language Variance | Lang Family Variance | ROI Variance |
|---|---|---|---|---|---|
| ***Response Strength Measures:*** *EffectSize ~ Condition + (1 \| Participant) + (1 \| Language) + (1 \| Lang. Family) + (1 \| fROI)* | | | | | |
| *Native-language > Degraded-language* | p<0.001 | 0.42 | 0.23 | 0.22 | 0.40 |
| *Native-language > Unfamiliar-language* | p<0.001 | 0.37 | 0.17 | 0.22 | 0.37 |
| *Native-language > Spatial Working Memory (Hard)* | p<0.001 | 0.16 | 0.00063 | 0.019 | 0.48 |
| *Native-language > Math (Hard)* | p<0.001 | 0.14 | 0.0017 | 0.017 | 0.25 |
| ***Lateralization measures (response strength and activation extent):*** *EffectSize ~ Hemisphere + (1 \| Participant) (1 \| Language) + (1 \| Lang. Family) + (1 \| fROI)* | | | | | |
| *Left Hemisphere > Right Hemisphere Response Strength* | p<0.01 | 0.33 | 0.043 | 0.010 | 0.77 |
| *Left Hemisphere > Right Hemisphere Activation Extent* | p<0.01 | 11,701 | 2,383 | 8,486 | 91,582 |
| ***Lateralization measures (functional correlations):*** *EffectSize ~ Hemisphere + (1 \| Participant) + (1 \| Language) + (1 \| Lang. Family)* | | | | | |
| *Story compreh.: Left Hemisphere > Right Hemisphere Functional Correlations* | p<0.01 | 0.026 | 0.0066 | 0.011 | NA |
| *Resting state: Left Hemisphere > Right Hemisphere Functional Correlations* | p<0.01 | 0.0044 | 0.0050 | 0.0053 | NA |
| ***Within- and between-network correlation measures:*** *EffectSize ~ Networks + (1 \| Participant) + (1 \| Language) + (1 \| Lang. Family)* Here, networks (either language-language (pairs of fROIs within the language network) or language-MD (pairs of fROIs straddling network boundaries)) were modeled as fixed effects: | | | | | |
| *Story compreh.: Within language network > lang-MD correlations* | p<0.01 | 0.0041 | 0.00074 | 0.00057 | NA |
| *Resting state: Within language network > lang-MD correlations* | p<0.01 | 0.0035 | 0.00063 | 0.00054 | NA |

**Supplementary Table 2:** Results of linear mixed effects models. The analyses reported in the main text were supplemented with linear mixed effects models to ensure the robustness of the results to the analytic procedure. These models included condition (as specified in column 1 for each measure) as a fixed effect and random intercepts for participant (n=86), language (n=45), language family (n=12), and ROI (n=6). The significance of the critical effects is shown in column 2, the 'Effect Significance' column (shaded cells). (In several cases, the model with the full random effects structure resulted in a singular fit. In such cases, we simplified the random effects structure by removing the language family effect and, if needed, the ROI and language random effects, until convergence was achieved. The details of all models are included on OSF: https://osf.io/5bzmc/.) In columns 3-6, we provide information on the variance associated with each random effect. R packages lme4[6] and lmerTest[7] were used.

| Participant Number | Native language(s) | Language(s) spoken fluently | Language(s) with some familiarity |
|---|---|---|---|
| 1 (544) | Arabic (0; 5; 5; home/class) | English (4; 5; 5; class) | French (15; 3; 3.5; class)<br>German (25; 2; 3; class) |
| 2 (561) | Arabic (0; 5; 5; home/class) | English (5; 5; 5; class) | French (6; 2; 2; class)<br>German (19; 3.5; 3.5; class)<br>Spanish (19; 3.5; 3.5; class)<br>Italian (19; 2; 2; class) |
| 3 (182) | Hebrew (0; 5; 5; home/class) | English (8; 5; 5; class) | French (25; 2; 2; class)<br>Arabic (14; 1; 1.5; class)<br>American Sign Language (30; 2; 1; class) |
| 4 (506) | Hebrew (0; 5; 5; home/class)<br>English (0; 5; 5; home/class) | Spanish (7; 4; 4; class) | German (21; 2; 2; class)<br>Swedish (21; 2; 2; class) |
| 5 (458) | Vietnamese (0;5; 5; home/class) | English (10; 3.5; 4; class) | |
| 6 (570) | Vietnamese (0;5;4.5; home/class) | English (7; 5; 5; home/class) | French (13; 2.5; 2.5; home/class)<br>Spanish (13; 1.5; 2; home/class) |
| 7 (580) | Tagalog (0; 5; 5; home)<br>English (0; 5; 5; home/class) | | |
| 8 (467) | Tamil (1; 5; 5; home/class)<br>English (3; 5; 5; home/class) | | Japanese (21; 3.5; 4; class)<br>German (18; 2; 2; class) |
| 9 (500) | Tamil (1; 5; 5; home)<br>English (3; 5; 5; home/class) | | |
| 10 (800) | Telugu (0; 5; 1; home)<br>English (1; 5; 5; home/class) | | Hindi (1; 5; 4; home/class)<br>French (14; 2; 2; class)<br>Gujarati (4; 3; 1; home/class) |
| 11 (451) | Afrikaans (0; 5; 5; home/class)<br>English (2; 5; 5; home/class) | | Dutch (9; 2; 2.5; class)<br>German (16; 2; 2; class) |
| 12 (455) | Afrikaans (0; 5; 5; home/class) | English (5; 4.5; 5; class) | Greek (20; 1.5; 1.5; class)<br>Hebrew (20; 1.5; 1.5; class) |
| 13 (454) | Armenian (0; 5; 5; home/class) | Russian (6; 5; 4.5; class)<br>English (9; 5; 5; home/ class) | French (21; 2; 2; class) |
| 14 (493) | Armenian (0; 5; 4; home/class) | English (10; 5; 5; home/class)<br>Russian (0; 4; 4; class) | French (15; 2; 1.5; class) |
| 15 (543) | Belarusian (5; 4; 4; home/class)<br>Russian (0; 5; 5; home/class) | English (10; 4; 4.5; home/class) | German (19; 2; 2; class) |
| 16 (611) | Belarusian (1; 5; 5; home/class)<br>Russian (0; 5; 5; home/class)<br>English (3; 4.5; 4.5; home/ class) | | Lithuanian (11; 3.5; 3.5; home/ class)<br>French (18; 3.5; 3; home/ class)<br>Polish (10; 3; 3; home/ class)<br>German (12; 2.5; 2.5; class)<br>Latvian (2; 2.5; 2.5; class)<br>Georgian (24; 2; 1.5; class)<br>Old Church Slavic (21; 1; 2.5; class)<br>Latin (18; 1; 2; class)<br>Sanskrit (21; 1; 2; class) |
| 17 (513) | Bulgarian (0; 5; 5; home/class)<br>English (3; 5; 5; home/ class) | Spanish (21; 4; 4; home/ class) | Russian (6; 3.5; 3.5; class)<br>French (14; 3; 3; class)<br>German (19; 2; 2; class) |
| 18 (517) | Bulgarian (0; 5; 5; home/class)<br>Russian (0; 5; 5; home/class)<br>Spanish (0; 5; 5; home/class)<br>German (0; 5; 5; home/class) | English (10; 4.5; 5; class) | |

| | | | |
|---|---|---|---|
| 19 (450) | Catalan (0; 5; 5; home/class)<br>Spanish (0; 5; 5; home/class) | English (4; 5; 5; class) | Serbocroatian (23; 1; 1; class) |
| 20 (464) | Catalan (0; 5; 5; home/class)<br>Spanish (0; 5; 5; home/class) | English (0; 5; 5; class) | German (13; 2; 2; class) |
| 21 (638) | Czech (details missing) | English (15; 4.5; 4.5; class) | German (8; 2; 3; class)<br>Russian (17; 2; 1.5; class)<br>Lithuanian (23; 1.5; 1.5; class) |
| 22 (647) | Czech (0; 5; 5; home/class) | English (14; 5; 5; class) | Russian (6; 3.5; 3.5; class)<br>Spanish (42; 3; 3; class)<br>German (0; 3; 3; class)<br>French (42; 2; 2; class) |
| 23 (507) | Danish (0; 5; 5; home/class)<br>English (0; 5; 5; home/class) | French (6; 5; 35; home/class)<br>Spanish (15;4; 4; home/class) | Norwegian (10; 2.5; 2.5; home/class)<br>Swedish (19; 2.5; 2.5; home)<br>German (12; 2; 1.5; home/class) |
| 24 (508) | Danish (0; 5; 5; home/class) | English (7; 5; 5; home/class) | Norwegian (11; 2; 3; class)<br>Swedish (11; 2.5; 1.5; class)<br>German (7; 2; 1.5; class)<br>French (15; 1.5; 1; class)<br>Italian (20; 1; 1; class) |
| 25 (463) | Dutch (0; 5; 5; home/class) | English (10;4.5; 4.5; class)<br>Portuguese (15;4.5;4.5; class) | French (0; 5; 5; class)<br>German (0; 5; 5; class) |
| 26 (481) | Dutch (0; 5; 5; home/class)<br>English (0; 5; 5; home/class) | | German (0; 2.5; 3; home/class)<br>French (12; 1.5; 1.5; class)<br>Spanish (18; 1.5; 1.5; class) |
| 27 (492) | English (0; 5; 5; home/class) | Spanish (13; 2; 2; class) | |
| 28 (502) | English (0; 5; 5; home/class) | | German (11; 2.5; 2; class)<br>French (10; 1; 1; class)<br>Latin (4; 1; 1; class) |
| 29 (443) | Farsi (0; 5; 5; home/class) | English (10; 5; 5; home) | German (4; 3.5; 3.5; home/class)<br>Spanish (18; 2; 2; class)<br>Turkish (2; 2; 1; class)<br>Greek (27; 1.5; 2; class)<br>Arabic (12; 1.5; 2; class)<br>Hungarian (2; 2; 1; class) |
| 30 (617) | Farsi (0; 5; 5; home/class) | English (details missing) | |
| 31 (462) | French (0; 5; 5; home/class)<br>English (0; 5; 5; home/class) | | German (0; 5; 5; home/class) |
| 32 (480) | French (0; 5; 5; home/class) | English (7; 5; 5; class) | Spanish (12; 1.5; 3; class)<br>German (23; 1.5; 2; class) |
| 33 (457) | German (0; 5; 5; home/class) | English (7; 5; 5; home/class) | Mandarin (0; 3; 2; home/class)<br>Latin (10; 1; 1; class) |
| 34 (482) | German (0; 5; 5; home/class)<br>Romanian (3; 5; 4.5; home) | English (11; 4; 4; home/class) | Hungarian (3; 2; 2; home) |
| 35 (496) | Greek (0; 5; 5; home/class) | English (8; 5; 4.5; home/class) | German (6; 3; 3.5; class) |
| 36 (548) | Greek (0; 5; 5; home/class) | English (7; 4; 4; home/class) | French (11; 2.5; 3; class)<br>German (14; 2; 2; class)<br>Portuguese (17; 2; 2; class)<br>Italian (24; 1; 2; class) |
| 37 (799) | Gujarati (0; 4.5; 2; home/class)<br>English (0; 5; 5; home/class) | | Hindi (5; 4.5; 3; home/class)<br>Arabic (19; 2; 2; class)<br>Bengali (20; 2; 2; class)<br>Latin (11; 2; 2; class) |
| 38 (808) | Gujarati (0; 4; 3; home)<br>Italian (0; 5; 5; home) | English (11; 5; 5; class) | French (3; 5; 5; home/class)<br>Spanish (14; 4; 4; class) |
| 39 (470) | Hindi (0; 5; 5; home/class) | English (4; 5; 5; home/class) | |

| | | | |
|---|---|---|---|
| 40 (504) | Hindi (2; 5; 5; home/class) | English (5; 5; 5; home/class) | Marathi (5; 3.5; 3.5; home/class)<br>Marwadi (NA; 3; 3; class) |
| 41 (618) | Irish (1; 5; 5; home/class)<br>English (1; 4; 4; home/class) | | |
| 42 (620) | Irish (0; 5; 4; home/class)<br>English (0; 5; 5; home/class) | | French (12; 3; 3; home/class)<br>German (10; 2; 2; class) |
| 43 (437) | Italian (0; 5; 5; home) | English (0; 5; 5; class) | Chinese (23; 1; 1; class) |
| 44 (444) | Italian (0; 5; 5; home/class) | English (11; 4; 4; class) | French (11; 2; 3; class)<br>Spanish (29; 2; 2.5; class)<br>German (27; 1; 1.5; class) |
| 45 (634) | Latvian (0; 5; 5; home/class)<br>English (0; 5; 5; home/class) | | French (4; 3; 4; class)<br>Italian (0; 1; 1; class) |
| 46 (635) | Latvian (0; 5; 5; home/class) | English (17; 5; 4; class) | Russian (7; 3.5; 2; class) |
| 47 (565) | Lithuanian (0; 5; 5; home/class) | English (4; 5; 5; class)<br>French (4; 5; 4.5; class) | German (11; 4; 3.5; class)<br>Spanish (13; 3.5; 3.5; class) |
| 48 (579) | Lithuanian (0; 5; 5; home/class) | English (6; 5; 5; home/class) | Spanish (19; 3; 2.5; class)<br>Russian (10; 2; 1; class) |
| 49 (810) | Marathi (0; 5; 5; home/class)<br>English (2; 5; 5; home/class)<br>Hindi (2; 5; 5; home/class) | | French (13; 1.5; 1; class) |
| 50 (813) | Marathi (0; 5; 5; home/class) | English (3; 5; 5; class) | Urdu (1; 3; 1; home/class)<br>Punjabi (10; 2.5; 1; class)<br>French (25; 1; 1; class) |
| 51 (515) | Nepali (0; 5; 3; home/class)<br>English (2; 5; 5; home/class) | Hindi (0; 5; 3; class) | |
| 52 (581) | Nepali (0; 5; 5; home/class)<br>English (2; 5; 5; home/class) | | Hindi (8; 4.5; 2.5; other) |
| 53 (460) | Norwegian (0; 5; 5; home/class) | English (7; 4.5; 4.5; class) | Swedish (5; 3.5; 4; class)<br>Danish (10; 2.5; 3.5; class)<br>Spanish (13; 2; 2.5; class) |
| 54 (469) | Norwegian (0; 5; 5; home)<br>Swedish (0; 4.5; 4; home) | English (10; 5; 5; class) | Spanish (13; 2; 2; class) |
| 55 (446) | Polish (0; 5; 5; home/class) | English (5; 5; 5; home/class)<br>German (15; 4; 4; class)<br>Slovene (19; 4; 4; class)<br>Serbocroatian (21; 4; 4; class) | Lithuanian (23; 3; 4; class)<br>Russian (21; 3; 4; class)<br>Bulgarian (26; 3; 3; class)<br>Albanian (28; 3; 3; class)<br>Latvian (25; 2; 3; class)<br>Czech (27; 3; 3; class)<br>Ukrainian (25; 2; 2.5; class)<br>Upper Sorbian (26; 2; 2.5; class)<br>French (24; 2; 2.5; class)<br>Norwegian (23; 1.5; 1.5; class)<br>Macedonian (29; 2; 2.5; class) |
| 56 (445) | Polish (0; 5; 5; home/class)<br>English (0; 4; 4; home/class) | Serbocroatian (18; 4; 4; class) | Slovene (20; 3.5; 3.5; class)<br>German (20; 3.5; 3.5; home/class)<br>French (25; 1.5; 2.5; class) |
| 57 (459) | Portuguese (0; 5; 5; home/class) | English (0; 4.5; 5; class)<br>Spanish (0; 4.5; 4.5; home/class) | French (10; 2; 2; class)<br>Mandarin (16; 3; 3; class) |
| 58 (484) | Portuguese (0; 5; 5; home/class) | English (10; 4; 4; class) | Spanish (13; 3; 2; class)<br>French (25; 2; 1; class)<br>Russian (28; 1; 1; class) |
| 59 (501) | Romanian (0; 5; 5; home) | English (10; 4.5; 4; class) | French (12; 1.5; 1.5; class) |
| 60 (509) | Romanian (0; 5; 5; home/class) | English (6; 5; 5; class) | French (11; 2; 2.5; class) |
| 61 (440) | Russian (0; 5; 5; home/class) | English (10; 3; 3; class) | |

| | | | |
|---|---|---|---|
| 62 (538) | Russian (0; 5; 5; home/class) | English (7; 5; 4.5; home/class) | German (18; 3; 2.5; home/class) |
| 63 (468) | Serbocroatian (0; 5; 5; home/class) | English (5; 5; 5; class) | Slovene (5; 3.5; 3.5; class)<br>Italian (8; 3; 2.5; class)<br>Spanish (8; 3; 2; class)<br>German (10; 2; 2; class) |
| 64 (546) | Serbocroatian (0; 5; 5; home/class) | English (4; 5; 5; class) | Italian (15; 3; 3; class)<br>German (15; 2; 2; class) |
| 65 (503) | Slovene (0; 5; 5; home/class) | English (14; 3; 3.5; class) | Serbocroatian (1; 5; 5; home/class)<br>German (8; 2; 1; class)<br>Italian (2; 2; 1; home/class)<br>Dutch (24; 1.5; 1; class) |
| 66 (497) | Spanish (0; 5; 5; home/class)<br>English (3; 4; 4; home/class) | | French (19; 3; 3; class)<br>German (4; 3; 3; class) |
| 67 (547) | Spanish (0; 5; 5; home/class)<br>English (5; 3.5; 3.5; home/class) | | |
| 68 (514) | Swedish (0; 5; 5; home/class) | English (9; 5; 5; home/class) | Spanish (0; 4; 4; class)<br>Mandarin (31; 1; 1; class) |
| 69 (542) | Swedish (0; 5; 5; home/class) | English (1; 5; 5; home/class) | French (11; 3; 2.5; class)<br>Mandarin (19; 2; 2; class)<br>Spanish (24; 1.5; 1; class) |
| 70 (490) | Ukrainian (0; 4.5; 4.5; home)<br>Russian (0; 5; 5; home) | English (6; 5; 5; class) | French (9; 2; 2; class) |
| 71 (495) | Ukrainian (0; 5; 5; home/class)<br>Russian (0; 5; 5; home/class)<br>English (3; 4; 4; home/class) | | German (11; 1.5; 1.5; class) |
| 72 (628) | Basque (0; 5; 5; home/class)<br>Spanish (0; 5; 5; home/class) | English (11; 4; 4; home/class) | French (13; 3; 2.5; class) |
| 73 (809) | Basque (2; 5; 5; home/class)<br>Spanish (0; 5; 5; home/class) | English (5; 5; 5; class) | Italian (24; 4; 4; class)<br>French (12; 2; 2; class)<br>Portuguese (24; 3.5; 3.5; class) |
| 74 (465) | Japanese (0; 5; 5; home/class) | English (12; 5; 5; class) | German (18; 2; 2; class) |
| 75 (512) | Japanese (1; 5; 4; home/class) | English (10; 5; 5; home/class) | Mandarin (1; 3.5; 2.5; class) |
| 76 (466) | Korean (1; 5; 5; home/class) | English (10; 3.5; 4; class) | |
| 77 (488) | Korean (0; 4.5; 3.5; home/class) | English (13; 4; 4; class) | |
| 78 (804) | Swahili (0; 5; 5; home/class)<br>English (0; 5; 5; home/class)<br>Kimeru (0; 5; 1.5; home) | | |
| 79 (478) | Mandarin (0; 4; 3; home/class)<br>English (3; 5; 5; home/class) | | Japanese (15; 2; 2; class) |
| 80 (491) | Mandarin (0; 5; 5; home/class) | English (7; 4.5; 4.5; home/class) | Japanese (21; 2; 2; class) |
| 81 (510) | Turkish (0; 5; 5; home/class) | English (10; 4.5; 5; class) | German (11; 4; 4; class)<br>Spanish (19; 3; 3; class) |
| 82 (533) | Turkish (0; 5; 5; home/class) | Arabic (15; 5; 5; class)<br>English (15; 5; 5; class) | French (28; 2; 3; class)<br>Farsi (32; 2; 2; class) |
| 83 (619) | Finnish (0; 5; 5; home/class)<br>Swedish (0; 5; 5; home/class) | English (20; 4; 4; class) | |
| 84 (648) | Finnish (0; 5; 5; home/class) | English (9; 4; 5; class) | Swedish (4; 3; 3.5; class)<br>German (14; 3; 3; class) |
| 85 (471) | Hungarian (0; 5; 5; home)<br>English (3; 5; 5; home) | German (7; 5; 5; home) | French (12; 3; 3; class)<br>Spanish (21; 2; 2.5; class) |

| | | | |
|---|---|---|---|
| 86 (518) | Hungarian (0; 5; 5; home/class) | English (8; 4; 4.5; class) | German (22; 2.5; 2; class)<br>Spanish (24; 2; 2.5; class)<br>Latin (14; 1; 2; class) |

**Supplementary Table 3**: Information on the language background of all participants. Participants are numbered 1-86 in column 1 (the number in parentheses is the UID (unique ID)—the internal lab identifier that is used in all the data tables and files on OSF: https://osf.io/cw89s/.). For each language listed in columns 2-4, we report in parentheses i) age of acquisition, ii) self-reported spoken proficiency (the average of self-reported spoken comprehension proficiency and speaking proficiency) on a scale from 1 (very basic proficiency) to 5 (native-like proficiency), iii) self-reported written proficiency (the average of self-reported written comprehension proficiency and writing proficiency) on the same 1-5 scale, and iv) environment in which the language was acquired ('home' indicates that one or both parents speak the language, 'class' indicates a formal language class either in high school or university). Listed under 'Native language(s)' is/are the language(s) that the participant listed as having learnt before the age of 6, with one or both parents speaking the language. Listed under 'Language(s) spoken fluently' is/are the language(s) with a self-reported spoken proficiency of 3 and above. Listed under 'Language(s) with some familiarity' is/are the rest of the languages reported by the participant.

| Native Language | Language Family | Number of Participants | Participant Sex and Age |
|---|---|---|---|
| Arabic | Afro-Asiatic | 2 | Male (28), Female (27) |
| Hebrew | Afro-Asiatic | 2 | Male (31), Female (26) |
| Swahili | Atlantic-Congo | 1 | Female (19) |
| Tagalog | Austronesian | 1 | Male (22) |
| Vietnamese | Austroasiatic | 2 | Male (21), Female (20) |
| Tamil | Dravidian | 2 | Male (25), Female (22) |
| Telugu | Dravidian | 1 | Male (28) |
| Afrikaans | Indo-European | 2 | Male (37), Female (25) |
| Armenian | Indo-European | 2 | Male (23), Female (30) |
| Belarusian | Indo-European | 2 | Male (23), Female (27) |
| Bulgarian | Indo-European | 2 | Male (37), Female (36) |
| Catalan | Indo-European | 2 | Male (25), Female (27) |
| Czech | Indo-European | 2 | Male (44), Female (27) |
| Danish | Indo-European | 2 | Male (32), Female (26) |
| Dutch | Indo-European | 2 | Male (32), Female (25) |
| English | Indo-European | 2 | Male (23), Female (25) |
| Farsi | Indo-European | 2 | Male (30), Female (32) |
| French | Indo-European | 2 | Male (29), Female (25) |
| German | Indo-European | 2 | Male (23), Female (30) |
| Greek | Indo-European | 2 | Male (26), Female (25) |
| Gujarati | Indo-European | 2 | Male (27), Female (27) |
| Hindi | Indo-European | 2 | Male (27), Female (22) |
| Irish | Indo-European | 2 | Male (26), Female (30) |
| Italian | Indo-European | 2 | Male (29), Female (29) |
| Latvian | Indo-European | 2 | Male (45), Female (25) |
| Lithuanian | Indo-European | 2 | Male (22), Female (19) |
| Marathi | Indo-European | 2 | Male (31), Female (28) |
| Nepali | Indo-European | 2 | Male (21), Female (24) |
| Norwegian | Indo-European | 2 | Male (22), Female (25) |
| Polish | Indo-European | 2 | Male (31), Female (31) |
| Portuguese | Indo-European | 2 | Male (19), Female (34) |
| Romanian | Indo-European | 2 | Male (19), Female (20) |
| Russian | Indo-European | 2 | Male (32), Female (23) |
| Serbocroatian | Indo-European | 2 | Male (28), Female (33) |
| Slovene | Indo-European | 1 | Female (24) |
| Spanish | Indo-European | 2 | Male (31), Female (41) |
| Swedish | Indo-European | 2 | Male (32), Female (31) |
| Ukrainian | Indo-European | 2 | Male (20), Female (19) |
| Basque | Isolate | 2 | Male (28), Female (25) |
| Japanese | Japonic | 2 | Male (29), Female (19) |
| Korean | Koreanic | 2 | Male (31), Female (29) |
| Mandarin | Sino-Tibetan | 2 | Male (25), Female (20) |
| Turkish | Turkic | 2 | Male (33), Female (30) |
| Finnish | Uralic | 2 | Male (37), Female (34) |
| Hungarian | Uralic | 2 | Male (25), Female (30) |

**Supplementary Table 4:** Information on the gender and age of the participants (at testing), as well as the number of participants tested per language. The table is sorted alphabetically by language family, and then by language.

**Supplementary Information-only References:**


1. 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).
2. Mineroff, Z., Blank, I. A., Mahowald, K. & Fedorenko, E. A robust dissociation among the language, multiple demand, and default mode networks: Evidence from inter-region correlations in effect size. *Neuropsychologia* **119**, 501–511 (2018).
3. 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).
4. Scott, T. L., 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).
5. Nieto-Castañón, A. & Fedorenko, E. Subject-specific functional localizers increase sensitivity and functional resolution of multi-subject analyses. *Neuroimage* **63**, 1646–1669 (2012).
6. Bates, D., Mächler, M., Bolker, B. M. & Walker, S. C. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, (2015).
7. Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* **82**, 1–26 (2017).