# Reliable individual-level neural markers of language activity

Kyle Mahowald (MIT) & Evelina Fedorenko (MGH)

brain+cognitive sciences

## Motivation

Most language neuroimaging work aims to characterize language in an average human brain. But there has been much recent interest in the relationship among behavioral variability, genetic variability, and variability in neural activity as measured with a variety of brain imaging techniques. So far, the use of anatomical markers is more common than the use of functional markers (e.g., the relative activation of different functional regions), plausibly due to the availability of large datasets obtained by labs performing the same anatomical scans on hundreds or even thousands of participants. Although there is good reason to believe that functional markers may be more useful than anatomical ones (e.g., because function does not align well with macroanatomical landmarks, and we don't yet have the power to see microanatomy), there has been relatively little work investigating the robustness, reliability, and interactive properties of functional markers. Here, we present a large dataset of language activations of healthy adult participants (n=150) and show that these activations are robust and stable within individuals (both across runs within a session and across sessions) and critically different across individuals along several key dimensions, including effect sizes and degree of lateralization.

## Methods

**Language Localizer (Fedorenko et al., 2010)**

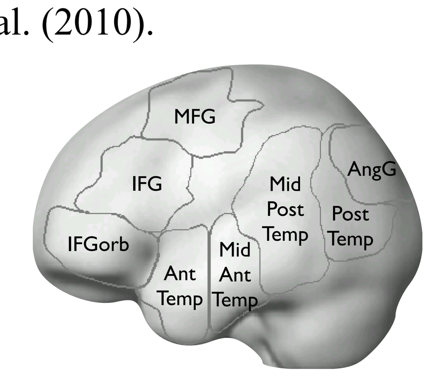
Participants read sentences (e.g., A RUSTY LOCK WAS FOUND IN THE DRAWER) and lists of unconnected pronounceable nonwords (e.g., DAP DRELLO SMOP UL PLID KAV CRE REPLODE) in a blocked design. Several slightly different versions of the localizer task were used across the 150 participants.

| Sentences | | Nonwords |
|---|---|---|
| A RUSTY LOCK WAS FOUND IN THE DRAWER | > | DAP DRELLO SMOP UL PLID KAV CRE REPLODE |

**Extracting functional measures**

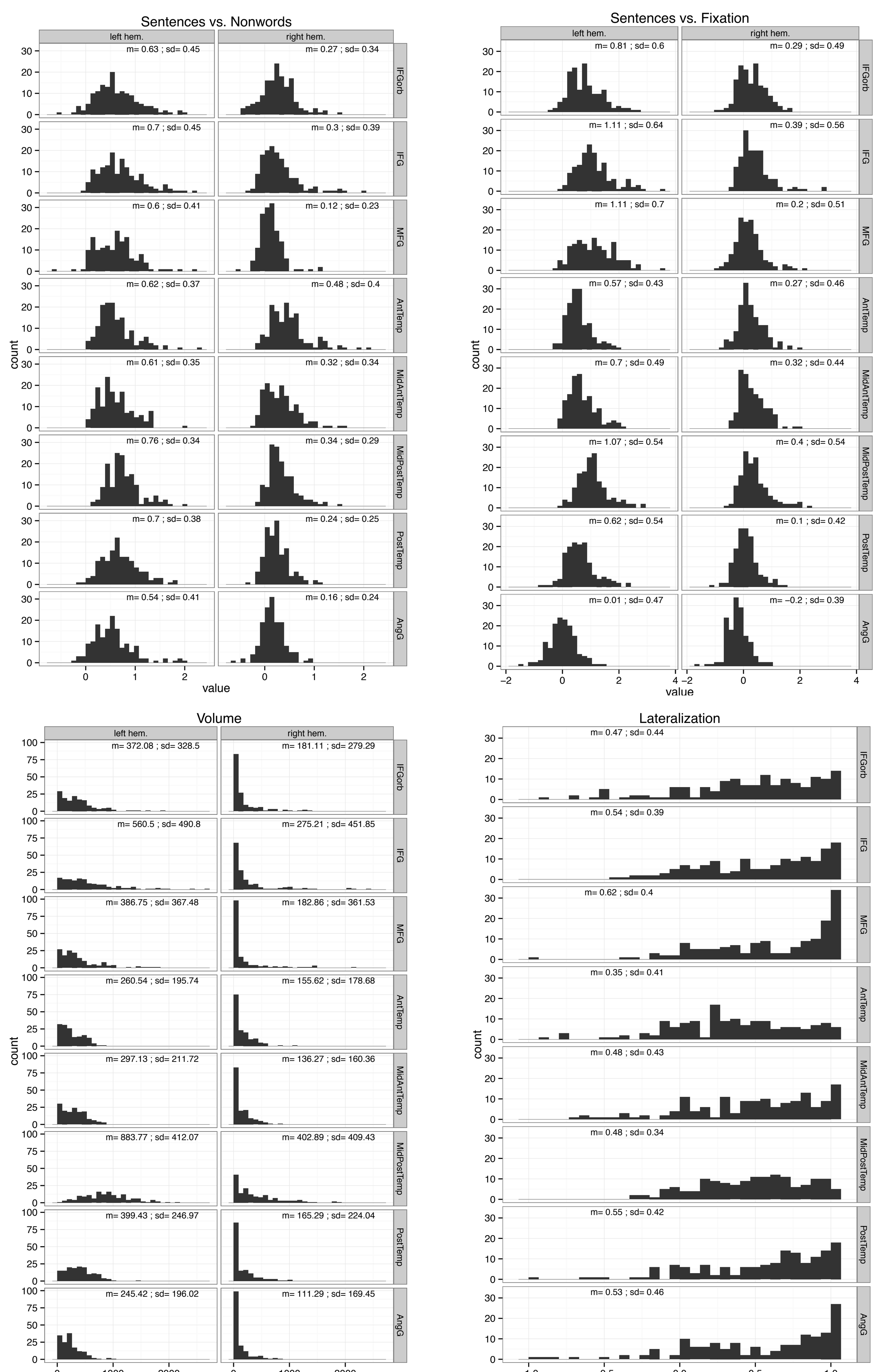
Using the individual activation maps for the Sentences > Nonwords contrast, we extracted 4 measures from each of the 16 regions of interest (ROIs), 8 in each hemisphere: a measure of activation volume (number of voxels for the Sentences > Nonwords contrast at the threshold of $p<0.001$, uncorrected, within each parcel), two measures of effect size (for the Sentences>Nonwords and Sentences>Fixation contrasts, extracted from the top 10% of the voxels within each parcel, using across-runs cross-validation), and a measure of lateralization (based on the volume measures, computed as (LH voxels - RH voxels)/(LH voxels + RH voxels)).

ROIs were defined functionally in each individual participant using the *Sentences > Nonwords* contrast, as developed in Fedorenko, et al. (2010).



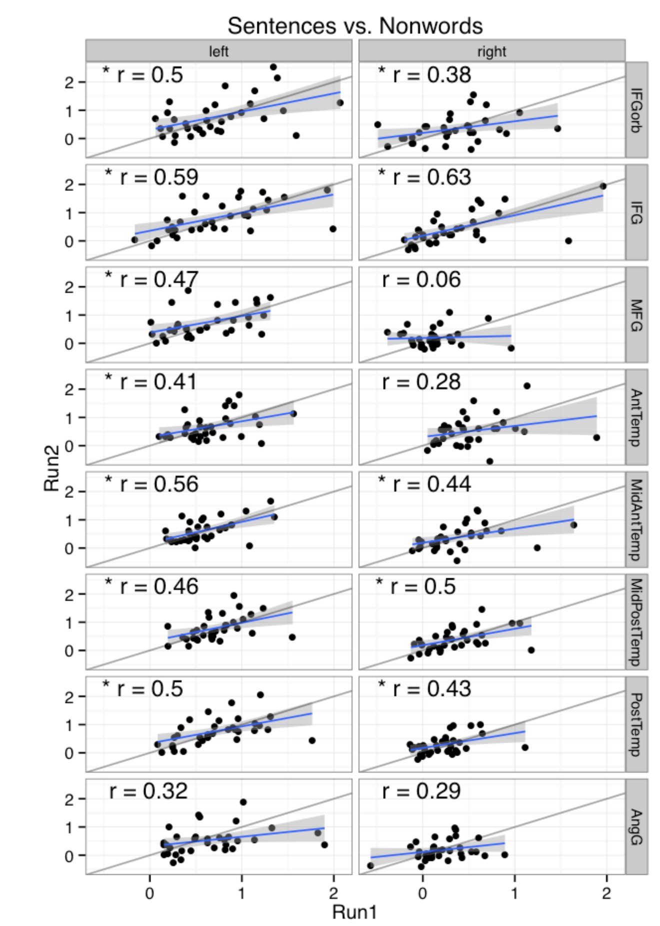

## Individual variation

These plots show the variability in each of the 4 functional measures of interest. S>N and S>F effect sizes are roughly normally distributed, while the others have longer tails.

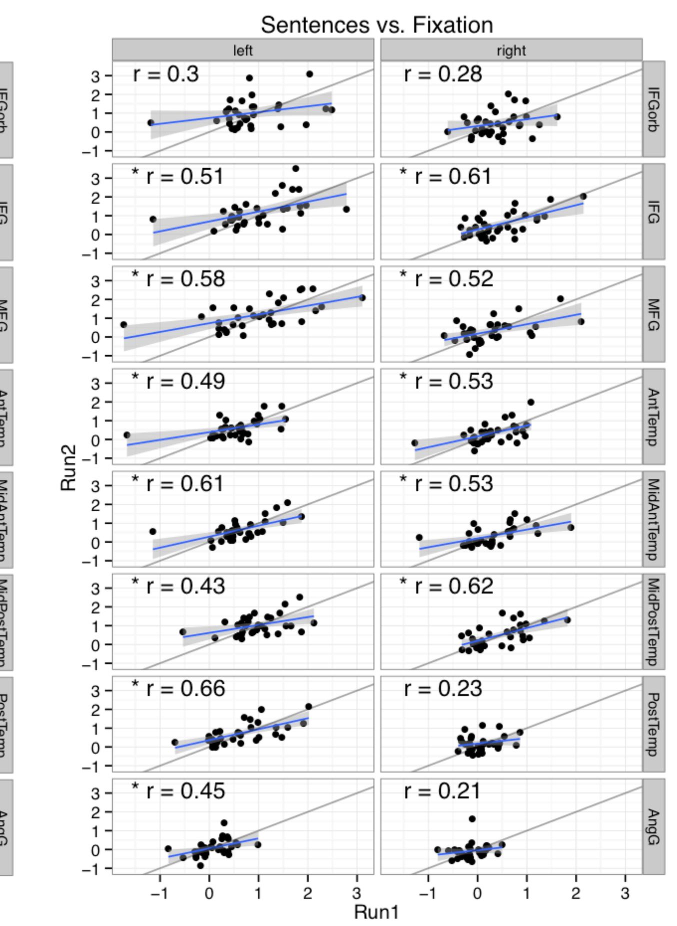



## Reliability of different neural measures

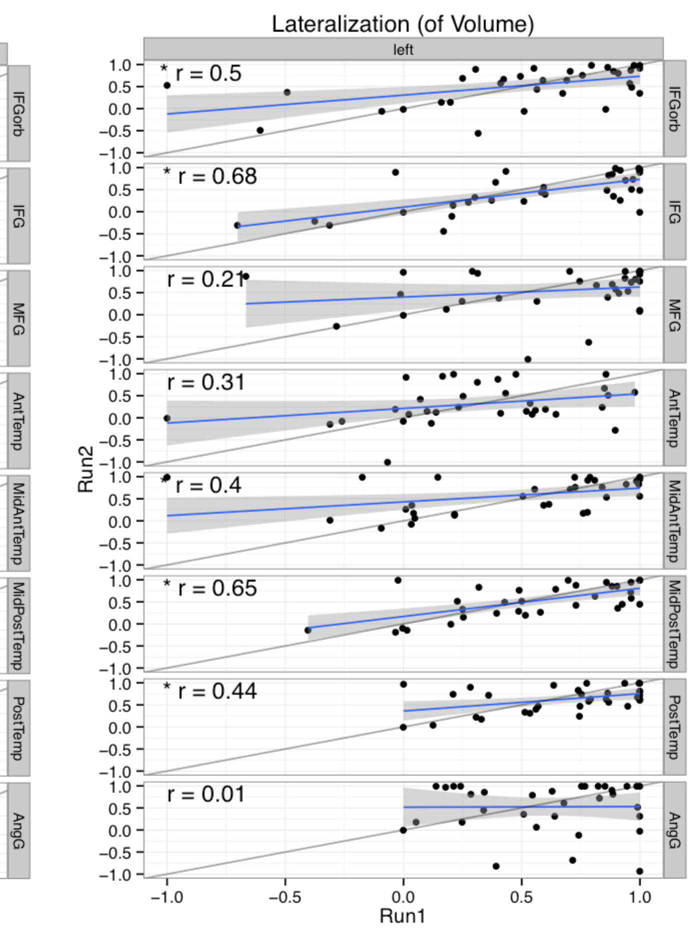
For a subset of 33 participants who were tested across 2 separate scanning sessions, we correlated first-session values and second-session values for each region separately, to test how stable effect size, volume and lateralization measures are. It appears that effect sizes are most reliable within subjects, and volume least reliable. Lateralization is reliable in a subset of the language network.



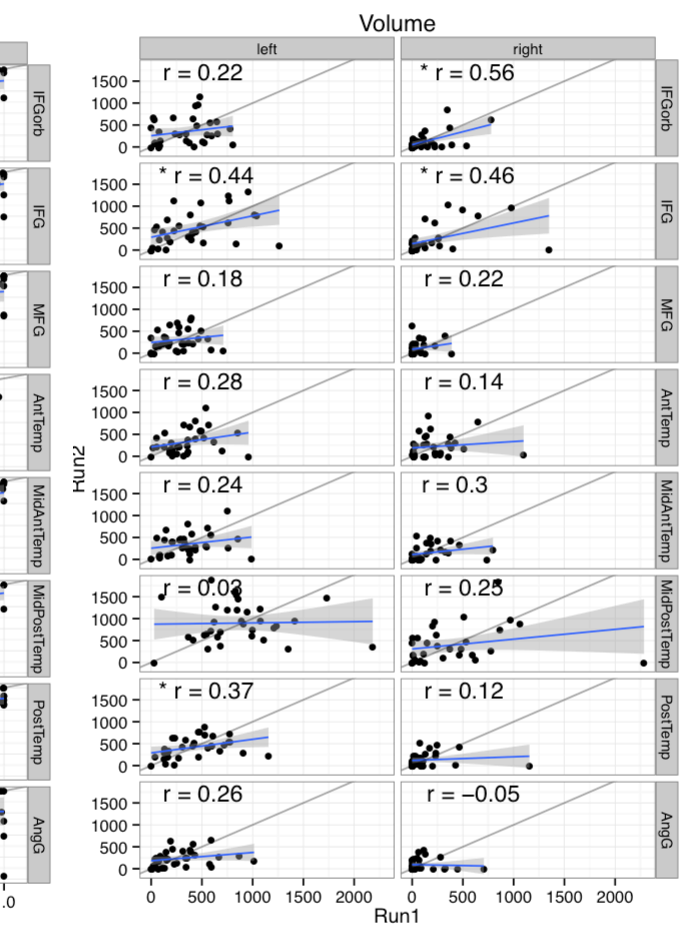

For the *Sentences > Nonwords* effect size, all 8 LH regions showed a correlation greater than .3, and these correlations were significantly different from 0 (*ps* < .05) in all regions except LAngG.




For the *Sentences > Fixation* effect size, all 8 LH regions showed a correlation greater than .29, and all were significantly greater than 0 (*ps* < . 01).



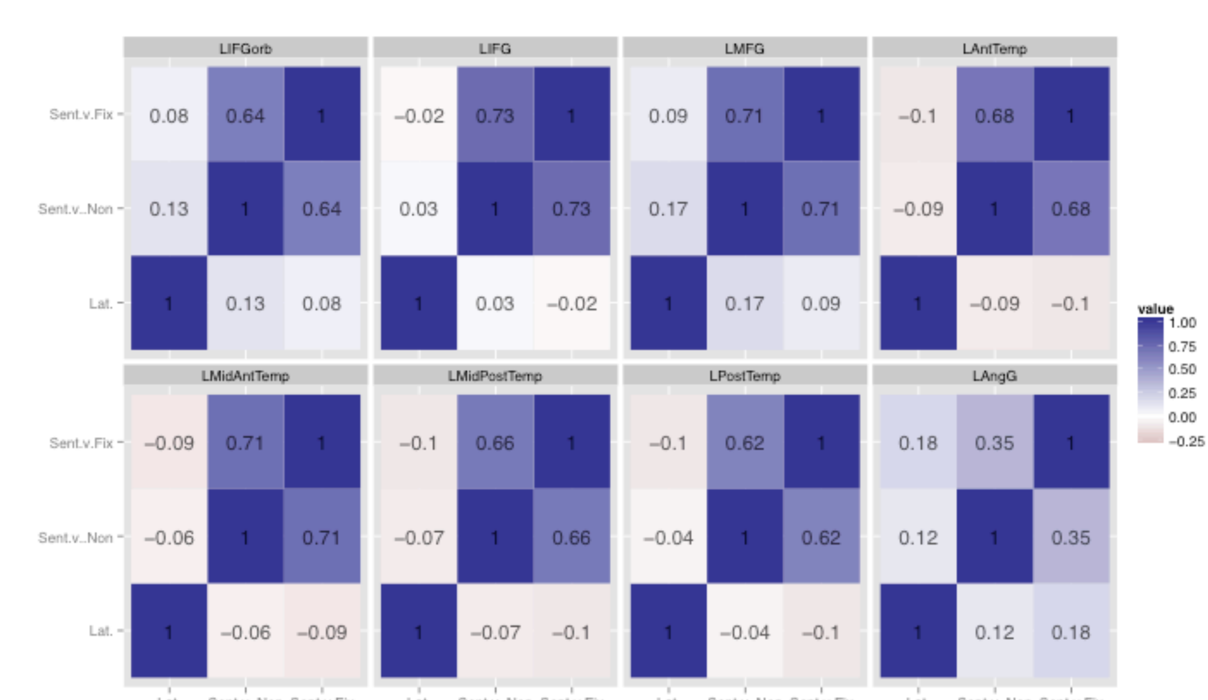

Lateralization showed a significant positive correlation in 5 of the 8 regions (*ps* < .05).

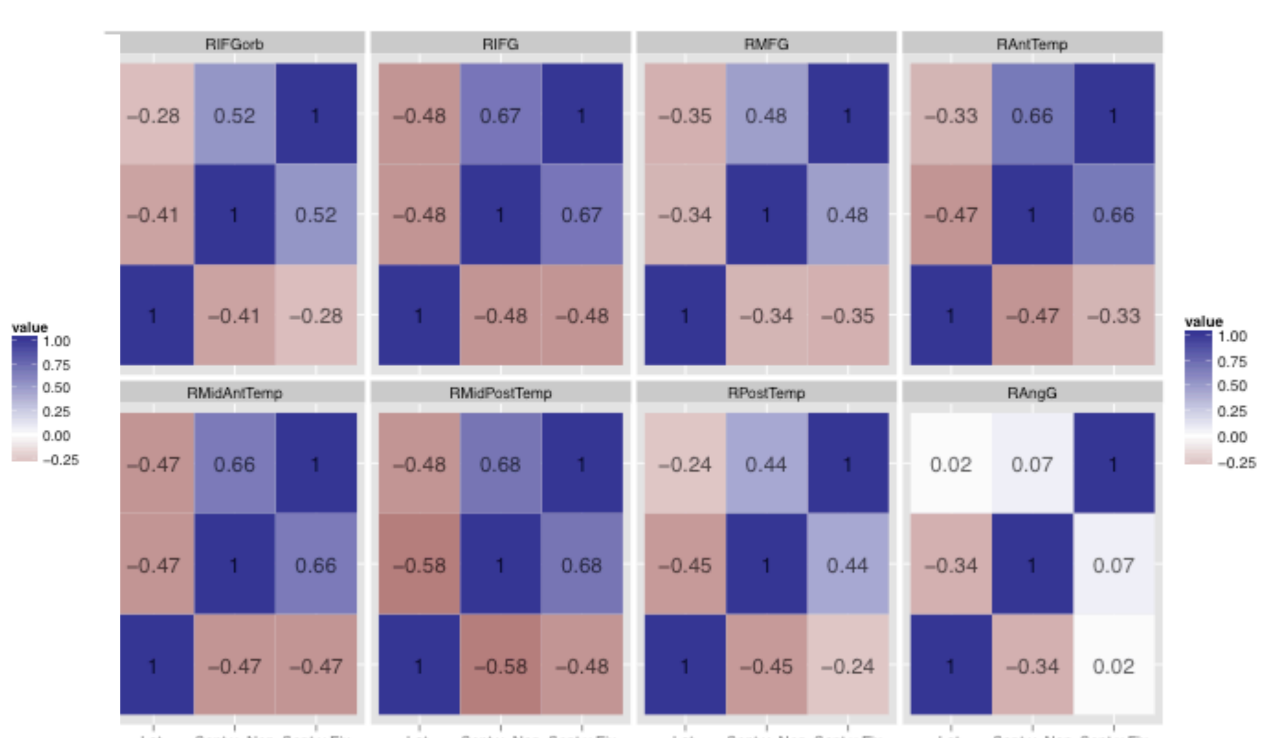



Volume showed a significant correlation in only 2 of the 8 regions (LIFG and LPostTemp, *ps* < . 05).

## Correlations among measures

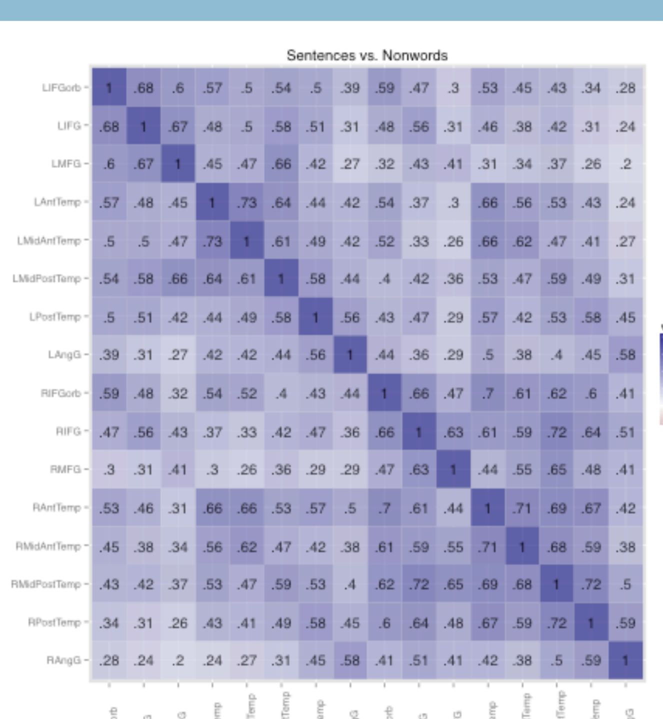
The effect size measures (*S > N* and *S > F*) are highly correlated in most regions. With the exceptions of LAngG and RAngG, for instance, the correlation between these measures is always .44 or above. The lateralization measure is almost entirely uncorrelated with the two types of effect size in the left hemisphere. We see robust inverse correlations between lateralization and effect sizes in the right hemisphere.

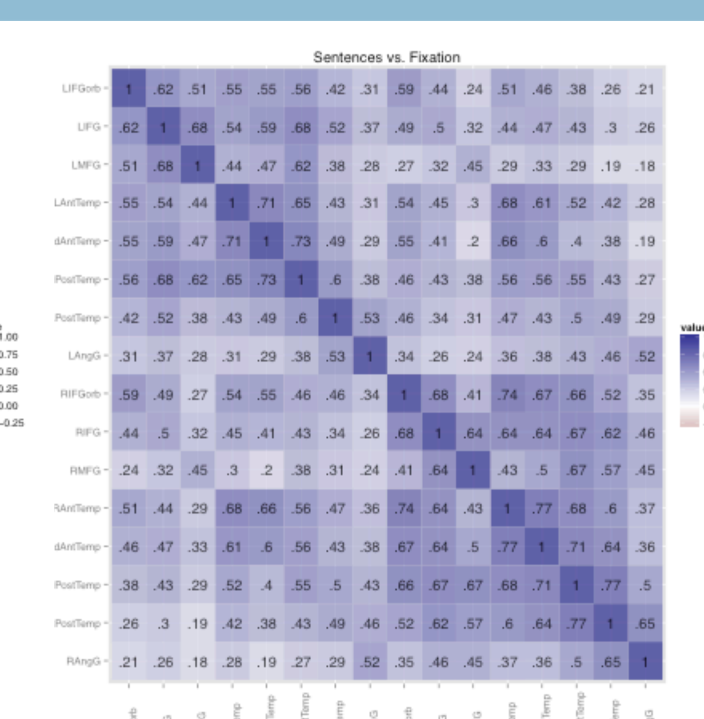


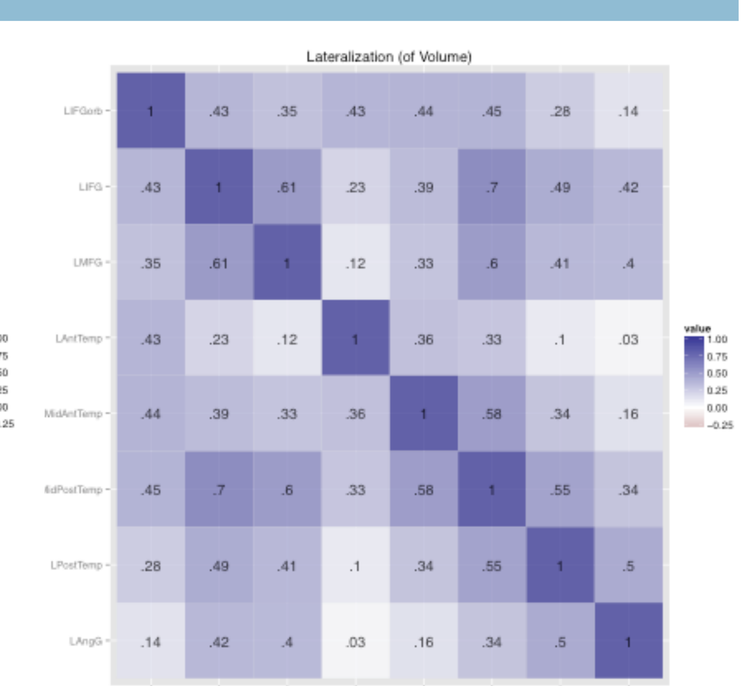


## Correlations among ROIs

There are robust correlations among the ROIs with respect to our measures of language activity, and some ROIs are more strongly correlated with each other than with other ROIs. We found the correlation across subjects between each fROI and every other fROI. All four measures show similar mean levels of correlation across all regions. The mean correlation among regions for *Sentences > Nonwords* is .53; .52 for *Sentences > Fixation*; .49 for *Volume*; and .44 for *Lateralization*.




## Conclusions

-Measures of neural activity can serve as reliable markers of individual differences that can be linked to behavioral and genetic variability in the population.
-Effect size measures (like Sentences > Nonwords) are more stable measures than volume measures.
-Effect size measures in the LH are not correlated with the lateralization measures, suggesting that the two may reflect different aspects of individual variability.
-There are strong correlations between neural activity across different ROIs, consistent with the idea that these brain regions form a functionally integrated system (e.g., Blank et al., 2013), as well as among different functional measures. These correlations should be taken into account when testing hypotheses about brain-behavior and brain-genes relationships.al measures.

**Kyle's contact info:** kylemaho@mit.edu
**Ev's contact info:** evelina9@mit.edu