Published in final edited form as:

Neuroimage. 2018 December ; 183: 200–211. doi:10.1016/j.neuroimage.2018.07.063.

Local response heterogeneity indexes experience-based neural differentiation in reading

Jeremy J. Purcell^{*} and Brenda Rapp

Department of Cognitive Science, Johns Hopkins University, USA

Abstract

The ability to read requires learning letter-string representations whose neural codes would be expected to vary depending on the amount of experience that an individual has with reading them. Motivated by sparse coding theories (e.g., Rolls and Tovee, 1995; Olshausen and Field, 1996), recent work has demonstrated that better-learned relative to less well-learned neural representations are associated with more strongly differentiated, locally heterogeneous blood oxygenation level dependent (BOLD) responses (e.g., Jiang et al., 2013). Here we report a novel analysis method we call local heterogeneity regression (Local-Hreg) that quantifies the cross-voxel heterogeneity of BOLD responses, thereby providing a sensitive and methodologically flexible method for quantifying the local neural differentiation of neural representations. In a study of literate adults, we applied Local-Hreg to fMRI data obtained when participants read letter strings that varied in their frequency of occurrence in the written language. Consistent with previous research identifying the left ventral occipitotemporal cortex (vOTC) as a key site for orthographic representation in reading and spelling, we found that the cross-voxel heterogeneity of neural responses in this region varies according to the frequency with which the written letter strings have been experienced. This work provides a novel approach for examining the local differentiation of neural representations, and demonstrates that well-learned words have greater representational differentiation than less well-learned or unfamiliar words.

Keywords

VWFA; Reading; Orthographic; Hreg; Hcorr; Heterogeneity; Fusiform gyrus

1. Introduction

A key aspect of reading is the ability to access stored memories of written words. Amid the extensive literature regarding the neural basis of written language, it has been proposed that the left ventral occipito-temporal cortex (vOTC) contains orthographic neural representations of commonly occurring letter combinations such as frequent bigrams, morphemes, or short words (Dehaene et al., 2002, 2005; Dehaene and Cohen, 2011) and additionally that well-learned orthographic strings (words) have more selectively tuned

^{*}Corresponding author. Department of Cognitive Science, Johns Hopkins University, Krieger 133, 3400 N. Charles Street, Baltimore, MD, 21218, United States. purcell14@gmail.com(J.J.Purcell).

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.neuroimage.2018.07.063>.

representations than less well-learned strings (pseudowords) (Glezer et al., 2009). Thus, this area has been found to be sensitive to the experience-dependent properties of orthographic representations.

Recent work suggests that, generally, well-learned neural representations (e.g. of faces or written words) are more neurally differentiated relative to less well-learned or unfamiliar representations (e.g., Jiang et al., 2013; Glezer et al., 2015). This work has been motivated by sparse coding theories which posit that well-learned neural representations have sharply tuned, sparsely distributed neural codes, i.e. high activation in a relatively small number of neurons (e.g., Field, 1987; Rolls and Tovee, 1995; Olshausen and Field, 1996). On this basis, the expectation is that sparser representations will have more locally heterogeneous (i.e., uncorrelated) neural responses (e.g., Vinje and Gallant, 2000; Rolls and Treves, 2011). Furthermore, if one assumes orthographic information is represented in neuronal population codes that are spatially distributed across voxels (e.g., Haxby et al., 2001; Kriegeskorte et al., 2006), well-learned cross-voxel orthographic representations may be expected to have locally more differentiated and heterogeneous neural responses as compared to words that are not as well-learned.

Early attempts to quantify experience-dependent representations of letter strings examined mean signal intensity and generally reported lower activation for high frequency compared to low frequency or nonword letter strings (Kronbichler et al., 2004; Mechelli et al., 2003; but see Carreiras et al., 2006). Although mean signal measures are useful for identifying the stimulus parameters a region is sensitive to, they do not necessarily provide information regarding the “quality” or local differentiation of the representations. This is because a high mean level of signal may reflect either the responses of a few sharply tuned, high-response neuronal populations, or those of many poorly tuned, low-response populations (e.g., Jiang et al., 2006; Mahon et al., 2007; Riesenhuber, 2007). Therefore, new approaches that discount mean signal differences and instead focus on quantifying local signal heterogeneity provide a promising direction for quantifying representational differences/changes associated with learning.

Recent approaches for exploring neural representation go beyond examining mean voxel responses, and instead draw inferences from the neural response properties of adjacent voxels. The view that neural representational codes are spatially distributed across adjacent voxels is the basic premise of multi-voxel pattern analysis methods (Haxby et al., 2001; Cox and Savoy, 2003; Kriegeskorte et al., 2008; Haxby, 2012; Mahmoudi et al., 2012). Support for these approaches stems from multi-unit recording studies, which have identified correlated neural spiking due to horizontal interconnections in local areas of the cortex that can potentially span several millimeters (Ts'o et al., 1986). Building on this, if we further consider that if neural representations become sparser after learning such that neurons respond more selectively (i.e., to fewer stimuli) after training (Kobatake et al., 1998; Freedman et al., 2006) then we would expect that - the voxel level-responses of well-learned neural representations would involve a smaller set of sharply tuned neurons distributed across adjacent voxels. Furthermore, single-unit electro- physiological research finds that neurons with less similar tuning (i.e., responsive to different stimuli) have less correlated (i.e. more heterogeneous) firing patterns relative to neurons with more similar tuning (Bair et

al., 2001; Jermakowicz et al., 2009). Therefore, it is expected that as neural tuning becomes sharper (i.e. more differentiated) within a set of locally distributed neurons, their firing patterns would be correspondingly less correlated. On these bases, we might expect that, at the voxel level, well-learned representations would involve relatively low correlation of more compact cross-voxel activity while less well-learned representations would utilize a larger set of broadly tuned neurons distributed across voxels and, therefore, higher correlation of cross-voxel activity would be expected. In short, well-learned representations would be expected to have more differentiated cross-voxel neural codes compared to less well-learned representations.

Recent approaches to quantify local neural differentiation (i.e. local heterogeneity) in fMRI have examined response similarity/dissimilarity across adjacent voxels *within a brain region* much as functional connectivity approaches traditionally examined voxels *across different brain regions*. For example, Zang et al. (2004) developed a regional homogeneity (ReHo) analysis which calculates a coherence value for each voxel based on the correlations of its time series with those of its nearest neighbors. More recently, Jiang et al. (2013, 2017) developed a similar approach - heterogeneity correlation (Hcorr) - which quantifies local heterogeneity by extracting task-relevant portions of the fMRI time series from adjacent voxels, and then calculates the standard error of the mean of the cross-correlations. The assumption is that well-learned as compared to poorly-learned stimuli recruit a smaller set of sharply tuned neurons, producing relatively more heterogeneous responses across voxels.

In this report, we build upon this previous work by introducing a novel analysis termed “local heterogeneity regression” (Local-Hreg) which provides a measure of the cross-voxel heterogeneity of neural responses in fMRI. We apply Local-Hreg to fMRI data obtained from reading letter strings that vary in terms of how much they have been previously experienced. The Local-Hreg approach introduced here addresses some of the limitations of previous approaches such as ReHo or Hcorr by using a (GLM) regression approach instead of correlation. Local-Hreg is an adaptation of the general psychophysiological connectivity analysis (gPPI; McLaren et al., 2012), which was originally developed to examine task-based functional connectivity across distant brain regions, but here is adapted to examine local, cross-voxel connectivity. As such, it analyzes condition-specific variability across adjacent voxels, providing the following benefits: (1) it leaves the time series intact, identifying the task-relevant aspects of the signal through model interaction terms rather than signal segmentation. (2) By incorporating the mean task-related response into the model, it is able to set it aside, and, in this way, provide a direct method of quantifying local neural heterogeneity independent of mean neural response. This is a key innovation over previous methods (e.g. Hcorr), allowing for a direct comparison of task-specific residuals across adjacent voxels whilst discounting any task-specific mean responses they might share. (3) It accounts for shared noise (e.g., due to motion) across adjacent voxels by the incorporation of motion-specific time-series interaction regressors; these regressors quantify how much signal is shared across adjacent voxels due to motion. (4) It is a flexible method for quantifying local neural heterogeneity. For instance, as a regression approach it is easily implemented in various fMRI designs such as block or event-related. Furthermore, it is easily implemented as a spatial searchlight, allowing for analyses not limited to predefined regions.

Consistent with previous research identifying the left ventral occipitotemporal cortex (vOTC) as a key site for orthographic representation in both reading and spelling (Dehaene et al., 2002; McCandliss et al., 2003; Binder et al., 2006; Gaillard et al., 2006; e.g., Baker et al., 2007; Rapp et al., 2015; Purcell et al., 2017), in this work we report that the cross-voxel local heterogeneity of neural responses in this region varies according to the frequency with which written letter strings have been experienced (consonant strings, pseudowords, low, middle and high frequency words), with local heterogeneity increasing with increasing experience. Additionally, we compare the results of the Local-Hreg analyses to those obtained from standard GLM and Hcorr analyses. These comparisons reveal that Local-Hreg and Hcorr both measure a similar aspect of the local neural response (i.e. local heterogeneity), which is distinct from mean signal measured via the standard GLM. We also report that Local-Hreg is a more robust analysis for quantifying local heterogeneity relative to Hcorr. These findings demonstrate that Local-Hreg can be used to quantify the experience-dependent heterogeneity of local neural responses, and supports the conclusion that the degree of neural differentiation of written word representations increases with practice and learning.

2. Materials and methods

2.1. Experimental design and statistical analysis

FMRI single-word reading data were acquired as part of data acquisition efforts involving two different studies. We combine the data from the two studies and although the design was similar for each study, differences are described below.

2.2. Participants

Combined across the two studies, there were 30 neurotypical right-handed literate adults (15 Male/15 Female; average age = 25 years, standard deviation (sd) = 7.9). Participants did not have contraindications for functional MRI and had normal or corrected to normal vision. Research participation consent was obtained using procedures approved by the Johns Hopkins University Institutional Review Board.

2.3. Task stimuli and design

Each study involved two functional runs per participant that included 6 and 12 s fixation periods at the beginning and end of each run, respectively. Each stimulus was presented for 200 msec with a 1300 msec fixation between stimuli. Stimuli were presented in a mini-block design with blocks of 2–10 stimuli of each type. Each run contained the following conditions: High Frequency words (HF; e.g., stop), Middle frequency (MF; e.g., belt), Low frequency words (LF; e.g. grit), Pseudowords (PW; e.g., cheg), Consonant Strings (CS; e.g., hgnk), and Checkerboards. The HF, MF, and LF words were defined using the MCWord Orthographic Database (Medler and Binder, 2005) to define these word conditions: HF > 35 wpm, MF 5 < 35, and LF < 5 words per million (wpm); as reported in Table 1 the average wpm across both studies for HF words = 86, MF words = 19, and LF words = 2. The word conditions had an average of 30 trials per condition per run (a total of 60 trials across both runs), the nonword stimuli had a total of 40 trials per condition per run (at a total of 80 trials across both runs). For each word and pseudoword stimulus we computed orthographic

neighborhood frequency (ONF, i.e. cumulative written word frequency of all of the orthographic neighbors), orthographic neighborhood count (ONC, i.e. number of orthographic neighbors) (Medler and Binder, 2005), and phoneme-grapheme consistency (PGC, i.e. the minimum position specific syllabic phoneme-grapheme mapping) (Hanna et al., 1966). Given that these parameters were not strictly matched across conditions and studies, they were included as regressors of no-interest in the subsequent GLM in order to better account for neural differences driven by ONF, ONC, or PGC. These parameters are summarized in Table 1. Stimuli were black on a white background and participants were instructed to fixate on a central cross when present, and to attend carefully to all of the centrally presented stimuli. Word and pseudoword stimuli were presented only once per study, were all in lower case, and had an equal number of stimuli of 4 and 7 letter length. Checkerboard and 7 letter stimuli occupied $2 \times 5^\circ$ of visual angle and 4 letter stimuli $2 \times 3^\circ$.

2.4. Imaging parameters

Full-brain functional and structural MRI scans were obtained using a 1.5T Phillips scanner for Study 1 and a 3T Phillips scanner for Study 2. A structural MPRAGE scan was acquired for each participant; full-brain coverage was obtained with a 1 mm^2 in plane resolution and 1 mm slice thickness. Each structural scan was aligned to the AC-PC plane, segmented into gray and white matter, and then normalized into Montreal Neurological Institute (MNI) space using SPM12. For both studies, T2*-weighted functional MRI Echo Planar Imaging (EPI) scans were acquired using a repetition time (TR) of 1.5 s and 212 acquisition volumes. Differences across studies included the following: for Study 1, 29 interleaved axial slices were obtained with a 3.6 mm^2 in plane resolution and 4 mm slice thickness, a Field of View (FOV) = 230 mm^2 , a Flip Angle (FA) = 65° , and an Echo Time (TE) = 30 msec. For Study 2, 23 sequential axial slices were obtained with a 1.9 mm^2 in plane resolution and 4 mm slice thickness, a FOV = 230 mm^2 , a FA = 70° , and a TE = 40 msec.

2.5. Pre-processing

Functional and anatomical data from each study were analyzed in the same manner using an integrated processing pipeline that included BrainVoyager QX2.4, SPM12, and Matlab. Pre-processing steps included correction for motion, high-pass temporal filtering (128 Hz), and resampling to 3 mm^3 . Each participant's functional scans were co-registered to their own structural scans in native space (BrainVoyager). Each structural scan was normalized to MNI space; these parameters were used to normalize individual functional scans. All coordinates are reported in MNI space.

Means and standard deviation in parentheses for: Freq = written word frequency in words per million; orthographic neighborhood count (ONC) and orthographic neighborhood frequency (ONF) = the number of and cumulative written frequency of all one letter different orthographic neighbors for each word (Medler and Binder, 2005). Phoneme-grapheme consistency (PGC) = the minimum position-specific c phoneme-grapheme consistency mapping for each word (Hanna et al., 1966).

2.6. Analytic approaches

In this work, we use three different analysis approaches: (1) a standard general linear model (GLM), (2) Local Heterogeneity Regression (Local-Hreg), and (3) a searchlight Heterogeneity Correlation (Hcorr) analysis. In this section, we describe each approach and its general implementation in this work. Information about specific analyses and comparisons across approaches is provided in the Data Analysis section below.

2.6.1 Approach 1: standard GLM—We apply the standard GLM approach which was originally developed as a technique for quantifying the mean BOLD signal within the context of a specific task or condition (Friston et al., 1994). The pre-processed, normalized data were smoothed with a 6 mm FWHM to account for cross-subject variability. A random effects (by participant) GLM was implemented which included stimulus regressors for the conditions of interest: HF, MF, LF, PW, and CS as well as the Checkerboard condition. Fixation condition time-points were left unmodeled and thus this condition served as an implicit baseline. Each condition regressor was convolved with a canonical HRF. In addition, participant-specific nuisance regressors were derived from the motion parameters (roll, pitch yaw, x, y, z) and CompCor Principle Component Analysis (PCA) components. This PCA analysis was used to estimate physiologically plausible noise (e.g. breathing and heart rate), and was performed on voxels of highly variable signal (i.e., voxels with top 2% variance) and cerebral spinal fluid voxels. Voxels were excluded in the PCA analysis mask if they were related to any of the condition regressors as defined by a fixed effects GLM and using a liberal voxel-wise threshold of $p < 0.2$ (Soltysik et al., 2015).

2.6.2. Approach 2: local heterogeneity regression analysis (Local-Hreg)—We developed Local Heterogeneity Regression Analysis (Local-Hreg) for quantifying the relative similarity of BOLD signal across adjacent voxels within the context of a specific task or condition by using a general psychophysiological interaction approach (gPPI; McLaren et al., 2012).

Local-Hreg is a searchlight analysis that was performed on unsmoothed, pre-processed, MNI normalized data. Each searchlight contained 7 voxels; 1 central voxel and the 6 voxels which shared a face with the central voxel. Within each searchlight, a general psychophysiological interaction model (gPPI) was used to characterize the center voxel of each searchlight using the following equations (adapted from McLaren et al., 2012):

$$Y_i = \underbrace{[Y_k H(g_p) N] * \beta_G}_{\text{mean fit estimates}} + \underbrace{[(Y_k * H(g_p))(Y_k * N)] * \beta_i}_{\text{interaction fit estimates}} + \underbrace{e}_{\text{error}} \quad (1)$$

In equation (1), Y_k is the time-series obtained from the center voxel of the searchlight. Y_k is used to develop a gPPI model that will be used to predict the time series of each adjacent voxel that shares a face with the center voxel in turn; in a searchlight of seven voxels, there are six adjacent voxels: these are denoted by Y_i . Equation (1) essentially corresponds to the linear combination of the *mean fit estimates*, the *interaction fit estimates* and the *error term*. In the mean fit estimates portion of the model, $H(g_p)$ corresponds to the condition regressor

vectors g_p (i.e. column vector for the stimulus timing of each condition such as HFW, LFW, PW etc ...) convolved with the canonical hemodynamic function H . $\mathbf{N}$ corresponds to a matrix of nuisance variables (e.g., motion). β_G are the parameter estimates of $\mathbf{Y}_k$ the condition regressors and the nuisance regressors; these terms can be considered to estimate the mean response in $\mathbf{Y}_i$ for each of these regressors. Including $\mathbf{Y}_k$ allows us to account for physiological based variance in $\mathbf{Y}_i$; i.e. account for any variance that exactly matches the center voxel. Including the condition and nuisance regressors, allow us to account for the mean response to these conditions in $\mathbf{Y}_i$. Critically, this mean fit estimates portion of the model permits us to account for and ignore the physiological and psychological shared variance of the center voxel to each adjacent voxel.

The interaction fit estimate portion of equation (1), essentially takes each term used in the mean fit estimates portion and multiplies it by the time series of the center voxel $\mathbf{Y}_k$. The critical parameter estimates for Local-Hreg are these psychophysiological interaction estimates contained in the β_i matrix. These interaction estimates allow us to quantify the amount of shared variance from the center and the adjacent voxel within the context of a particular condition (e.g. reading HF words). Note, that the inclusion of interaction regressors for the nuisance variables ($\mathbf{N}$) is not a characteristic of traditional gPPI analysis (McLaren et al., 2012), but is particularly relevant to the Local-Hreg analysis as it accounts for shared variance across adjacent voxels due to sources of noise such as motion.

In summary, for each searchlight we obtained a vector of six β_i estimates (from each adjacent voxel) for each of the HF, MF, LF, PW, and CS conditions of interest. For each searchlight, the Local-Hreg values correspond to the inverse of the median of these six β_i estimates, one for each of the five conditions of interest; Local-Hreg values were assigned to the center voxel of each searchlight. Medians were used (instead of means) in order to avoid assumptions of a normal distribution across the six β_i estimates. The inverse was taken for ease of communication, so that higher Local-Hreg values correspond to higher heterogeneity (less similar neural responses across adjacent voxels), and lower Local-Hreg values correspond to lower heterogeneity (more similar neural responses across adjacent voxels). This algorithm was developed and implemented in MATLAB.

Given that this searchlight approach essentially generates a smoothed result (i.e., the value at each voxel is necessarily related to its surrounding voxels), we applied additional smoothing of only 3 mm FWHM to each individual Local-Hreg condition map in order to take into account cross-subject variability. In this manner, the total smoothing of these maps roughly approximates 9–12 mm FWHM.¹

We note that the studies reported on here were part of a larger investigation and that the block design was not selected specifically with Local-Hreg analysis in mind. It is possible that there are other fMRI designs (e.g., event-related designs) that would better capture shared trial- to-trial BOLD responses across voxels. However, we expect that just as it has been necessary to optimize other designs such fMRI adaptation or multi-voxel pattern

¹ To confirm this approximation we calculated the smoothness from an unthresholded group statistical Local-Hreg map (for the HF > PW contrast discussed in analysis 3 below) using the *mapsmooth* program in the *Neuroelf* toolbox; as expected the resulting FWHM values in x, y, z direction was 11.1, 11, and 10.1 mm.

analysis (e.g., Aguirre, 2007; Abdulrahman and Henson, 2016), further work with Local-Hreg will be needed to identify optimal designs. Nonetheless, we think the mini-block design used here is appropriate given the following considerations:

First, the core algorithm of the Local-Hreg analysis is directly derived from the general PPI analysis (McLaren et al., 2012). Whereas the PPI analysis was developed to evaluate the similarity of task-dependent responses across spatially distant voxels, Local-Hreg was developed to evaluate the similarity of task-dependent responses across adjacent voxels. Given that both the PPI and Local-Hreg rely on the premise of examining psychological interaction terms in a general linear model, and that the PPI analysis traditionally employed block designs (Friston et al., 1997), the inferences drawn with respect to local heterogeneity should be just as valid as those drawn from the PPI analyses that used a block design.

Second, specifically with regard to the relationship between block and event-based (e.g., trial-based) designs for Local-Hreg, given that the Local-Hreg analysis evaluates the similarity of signal time-courses across voxels, it necessarily considers each time-point in a block of trials. The cross-voxel similarity in signal at each time point (e.g. mean difference in intensity of at each time point) is inversely proportional to the correlation of the time series such that, for two time series, if there is a mean high difference across time points, there will be a low correlation across the entire time series, and vice versa. This underscores that, in this particular context, event-related and block designs differ only in the extent to which the BOLD time-courses for a set of trials are compressed in time. In either case, the difference in signal at each point across voxels is still proportional to the correlation of the entire time series.

2.6.3. Approach 3: search light heterogeneity correlation analysis (Hcorr; adapted from Jiang et al., 2013)—We introduce an adapted version of the Heterogeneity Correlation (Hcorr) analysis (Jiang et al., 2013, 2017), which was originally developed as a technique for quantifying the relative similarity of BOLD signal across adjacent voxels within the context of a specific task or condition. Although the goal of Hcorr is the same as in the Local-Hreg analysis, the technique is algorithmically different in that it does not rely on a GLM (specifically gPPI) approach, but instead segments the BOLD time series and specifically quantifies the variability of local correlations. We include the Hcorr analysis in order to compare the results with those of Local-Hreg. To carry out this comparison, we implemented a searchlight version of heterogeneity correlation (Hcorr) analysis (Jiang et al., 2013).

The searchlight Hcorr analysis was performed on un-smoothed, pre-processed, MNI normalized data. First, the raw time series data was extracted from each voxel within a searchlight. Second, to appropriately segment the time series, the portions of the time series associated with each condition - (e.g. the trials for HF words) - were first identified. To do this while accounting for time shift of the BOLD response, we first convolved the canonical HRF with that of our design matrix (this is the same standard GLM model used in Analysis 1). We then assigned a portion of the time series to a given condition (e.g. HF) if that condition had the highest positive values at these time points in the pre-convolved HRF design matrix. These condition-specific portions of the raw time series were then extracted,

mean-normalized (within consecutive time points), and then finally concatenated for each condition separately. We avoided bias due to unequal number of time points per condition by setting the number of time points associated with each condition to the minimum across conditions (e.g., if HF words had 81 time points and the other conditions all had 76 time points, the last 5 time points of the HF time series were excluded). Third, we performed pairwise correlations on the condition-specific portions of the time series between n voxels in the searchlight (i.e. 7 voxels = 21 pairwise correlations). This is noted in equation (2) where r_{ij} is a correlation between condition specific time series across pairs of voxels (i and j) within the searchlight where i and j is an element of the set from 1 to n number of voxels (in this case 7).

$$r_{ij} = \text{corr}(Vox_i, Vox_j) \quad (2)$$

Although the Hcorr values in Jiang et al. (2013) corresponded simply to the standard error of the mean (SEM) of the condition-specific correlations, in order to determine if a given condition had *higher* variability than another did, we calculated an F statistic (F_{stat} ; equation (3)). To do so, and in order to generate maps with both the magnitude and *directionality* of variance differences, we assigned the F_{stat} value to each voxel and then assigned negative or positive polarity depending on the direction of the comparison. For example, in a comparison of HF > PW, in equation (3) *ConditionA* can be either HF or PW depending on which has the greatest variance, with the condition with the highest variance placed in the numerator. If HF has the greatest variance then it would be set as the numerator (i.e. *ConditionA* σ^2), and PW would be set as the denominator (*ConditionB* σ^2). The Hcorr value assigned to the center voxel would simply be the ratio of these two variances (i.e. the F_{stat} , which is by default a positive value). If, on the other hand, PW has the greatest variance, then the PW variance would be set as the numerator and HF in the denominator, and the F_{stat} value would be multiplied by -1 in order to define the direction in the HF > PW variance comparison (note: although not performed here, in a PW > HF comparison we would multiply the F_{stat} by -1 in searchlights for which HF had the greatest variance). Subsequently, a map is generated with the magnitude *and* direction of F_{stat} values for the HF > PW comparison. Therefore, in this work, positive Hcorr values indicate greater local heterogeneity for HF relative to say PWs; negative values depict greater local heterogeneity for PW relative to HF words. In this manner, a single Hcorr map can be generated for each contrast in each subject (e.g. a separate HF > PW map).

$$F_{stat} = \frac{\text{ConditionA } \sigma^2}{\text{ConditionB } \sigma^2}, \quad \text{where } \sigma^2 = (r_{ij} - u)^2 \quad (3)$$

As in the Local-Hreg analysis, a FWHM 3 mm smoothing kernel was applied to both individual Hcorr contrast maps (e.g. for HF > LF or HF > PW). Group-level one-tailed t-statistical tests were performed (voxel-wise $p < 0.05$).

2.7. Data analysis

We deployed the above analytic approaches to examine the data with four sets of statistical analyses: (1) Identification of an Orthographic Network; (2) Evaluation of parametric stimulus condition effects (standard GLM and Local-Hreg); (3) Local-Hreg identification of the neurotopographic distribution of stimulus condition effects; and (4) comparison of Local-Hreg, Search Light Hcorr, and the Standard GLM.

2.7.1. Analysis 1: identification of the orthographic network—An Orthographic Network was identified from an Orthographic > Fixation contrast. Specifically, this included voxels with significant mean response to Orthographic stimuli (HF, MF, LF, PW, CS) > Fixation (voxel-wise $p < 0.01$ and cluster-wise $p < 0.05$). This analysis was constrained to a gray matter mask which included all cortical regions in the Harvard-Oxford atlas (Desikan et al., 2006). The cluster-wise threshold was obtained by running a Monte Carlo simulation (Forman et al., 1995) via the AlphaSim program in the Neuroelf toolbox estimating the maximal cluster due to chance at an alpha level < 0.05 after 10,000 iterations of simulated random voxel responses that are non-isometric in the x, y, and z dimensions (Eklund et al., 2016). The final unthresholded group statistical map smoothness was estimated using the mapestsmooth program in the Neuroelf toolbox; the resulting FWHM values in x, y, z direction were 10.4, 10.6 and 9.8 mm, resulting in a cluster threshold of 107 contiguous voxels.

2.7.2. Analysis 2: evaluation of parametric stimulus condition effects (standard GLM and Local-Hreg)—We tested whether there was a linear relationship across the stimulus conditions such that the neural signal in a given voxel was linearly related to the amount of experience participants would be expected to have had with each letter string condition. To do this we used a standard GLM to generate maps for each stimulus condition: HF, MF, LF, PW, and CS. At each voxel within the Orthographic Network mask we tested whether there was a parametric linear relationship across stimulus conditions such that $HF > MF > LF > PW > CS$. We then repeated this parametric linear relationship analysis with the Local-Hreg analysis. All analyses were carried out within the spatial constraints of the Orthographic Network mask.

To correct for multiple comparisons, a stringent non-parametric permutation test was used (Nichols and Holmes, 2002) that involved label-scrambling (i.e., randomly assigning condition labels; in this case, what is the HF, MF, LF, PW or a CS condition) for each of 10,000 iterations. For each iteration, the maximum cluster size was obtained and an extent threshold was identified by identifying the cluster size that occurred in fewer than 5% of random permutations.

2.7.3. Analysis 3: Local-Hreg evaluation of the neurotopographic distribution of stimulus condition effects—In order to explore the robustness of the novel Local-Hreg technique, we performed individual contrasts between the two pairs of conditions that were predicted to have robust differences, i.e. $HF > LF$ and $HF > PW$. The corrections for multiple comparisons were obtained in the same non-parametric permutation manner as discussed for Analysis 2.

2.7.4. Analysis 4: comparison of Local-Hreg, Search Light Hcorr, and standard GLM—Here a comparison was made between the three analytic approaches (i.e., standard GLM, Local-Hreg, and Search Light Hcorr). This was carried out by first identifying clusters with significant stimulus condition effects, and then by examining the relationship (i.e. correlation) between the values generated by the different analytic approaches within these clusters (e.g. the relationship between Standard GLM β and Local-Hreg values within the same voxels).

As reported below, Hcorr analysis within the Orthographic Network yielded no significant clusters (i.e., for HF > LF or HF > PW, even with a lenient threshold of $p < 0.3$). To be able to nonetheless compare the three approaches, the comparison was restricted to the left ventral occipital temporal cortex (vOTC) defined by the Harvard-Oxford atlas as including left inferior occipitotemporal, inferior temporal, and fusiform gyri defined within the following y and z plane bounding ranges: y between -80 and -25 and z between -5 and -37 . Corrections for multiple comparisons were obtained in the same non-parametric permutation manner as discussed for Analysis 2. In order to directly statistically evaluate the relationship between the values obtained from the three analytic approaches, we first extracted the voxel-level values from reported clusters for the standard GLM β parameter estimates, the Local- Hreg, and the Hcorr values. We then report pre-planned correlations (e.g. standard GLM vs Local Hreg) and a linear mixed-effects modeling (LMEM) in order to determine if there was a relationship between the values obtained from these approaches. For the LMEM, the “lme4” library in R[R Core Team, 2013] was used; the Satterthwait method was used to estimate the degrees of freedom (Kuznetsova et al., 2017) and calculate p-values.

3. Results

3.1. Identification of the orthographic network

The (HF, MF LF, PW, CS) > Fixation contrast identified a set of left lateralized functional clusters that have typically been associated with reading (e.g., Martin et al., 2015) and which we will refer to as the Orthographic Network. As shown in Fig. 1 and reported in Table 2, these regions include bilateral inferior frontal gyrus (IFG), bilateral ventral occipitotemporal cortex (vOTC), and left intraparietal sulcal region. This network was used as an inclusive mask for subsequent analyses.

3.2. Evaluation of parametric stimulus condition effects (standard GLM and Local-Hreg)

Within the Orthographic Network, the linear parametric analysis with the Standard GLM identified a single cluster (voxel-wise $p < 0.05$, cluster threshold $p < 0.05$) in the left vOTC (peak MNI = -47 , -66 , -14 ; Cluster size = 222; max z-value = 3.2; corrected $p = 0.04$; Fig. 2A). In order to depict this linear relationship across conditions within this cluster, mean β estimates for each condition (i.e. HF, MF, LF, PW, and CS) were extracted for each participant; the means and standard errors across participants are plotted in Fig. 2B.

This linear parametric analysis was also performed using the Local- Hreg approach. Although we did not find any clusters that survived correction for multiple comparisons, we

examined the largest cluster identified at uncorrected voxel-wise $p < 0.05$ (peak MNI = —51, —69, —19; Cluster size = 36; max z-value = 3.2; Fig. 2C). The means and standard errors of the Local-Hreg values for each condition were extracted from this cluster and plotted in Fig. 2D.

3.3. Local-Hreg evaluation of the neurotopographic distribution of stimulus condition effects

The contrast of HF with LF words within the Orthographic Network identified a single significant Local-Hreg cluster within the left vOTC (peak MNI = —53, —45, —23; Cluster size = 187; max $t(29) = 4.6$; corrected $p = 0.04$; Fig. 3A). The contrast of HF with PW using the same thresholds identified a single cluster within the left vOTC (peak MNI = —50, —44, —23; Cluster size = 98; max $t(29) = 3.4$; corrected $p = 0.09$; Fig. 3B). These two clusters intersect, sharing 66 voxels. These results reveal significant local heterogeneity differences between better-known words and less well-known words (LF) or unknown words (PW) within the same part of the left vOTC. In order to depict the relative robustness of these left vOTC clusters we also present the uncorrected contrast maps using a more lenient voxel-wise threshold ($p < 0.1$) for both the HF > LF and HF > PW contrasts within the Supplementary Materials S1.

3.4. Comparison of Local-Hreg, Search Light Hcorr, and standard GLM

First, with the standard GLM analysis within the left vOTC mask, we found significant (or near significant) clusters for both the HF > LF and HF > PW contrasts (voxel-wise $p < 0.05$; cluster threshold $p < 0.05$). HF > LF: peak MNI = —45, —45, —20; Cluster size = 68; max $t(29) = -3.26$; corrected $p = 0.06$ and HF > PW: peak MNI = —63, —46, —17; Cluster size = 182; max $t(29) = -3.9$; corrected $p = 0.02$.

Second, with the Local-Hreg analysis within the left vOTC mask, we found significant clusters for both the HF > LF and HF > PW contrasts (voxel-wise $p < 0.05$; cluster threshold $p < 0.05$). HF > LF: peak MNI = —53, —45, —15; Cluster size = 293; max $t(29) = 4.4$; corrected $p = 0.003$ and HF > PW: peak MNI = —50, —44, —23; Cluster size = 125; max $t(29) = 3.4$; corrected $p = 0.036$.

Third, for the searchlight Hcorr analysis results a more liberal threshold was required (voxel-wise $p < 0.05$; cluster threshold $p < 0.3$) in order to identify left vOTC clusters for the HF > LF and HF > PW contrasts (peak MNI = —57, —47, —23; Cluster size = 28; max $t(29) = 2.6$; corrected $p = 0.22$) and (peak MNI = —32, —60, —21; Cluster size = 44; max $t(29) = 2.98$; corrected $p = 0.13$), respectively. Each of these clusters overlaps with that of the significant Local-Hreg clusters presented in Fig. 4B, suggesting that although the response is less robust for Hcorr, both approaches identify common regions of where there is greater local heterogeneity for well-known words relative to less well-known or unknown words.

In order to directly examine the relationship between these three analysis approaches we first compared results from Local-Hreg with Standard GLM and then compared Local-Hreg with Hcorr. To start, we identified intersecting voxels from the standard GLM (Fig. 4A) and Local-Hreg (Fig. 4B), extracting and z-scoring (across participants) the β estimates from the standard GLM and the Local-Hreg values. The Pearson correlation coefficients for the

standard GLM and the Local-Hreg values across participants for both HF > LF and HF > PW contrasts were not significant (See Fig. 4D) with r -values < 0.3 and p -values > 0.1: HF > LF $r = 0.11$, $p = 0.6$ and HF > PW $r = -0.26$, $p = 0.17$. Next, we examined the relationship between standard GLM and the Local-Hreg values across the left vOTC anatomical mask including data from each voxel (1054 voxels; green outline in Fig. 4A-C) and each contrast (HF > LF and HF > PW) within a single LMEM with Local-Hreg as the dependent variable; fixed effects were the GLM β estimates and contrast type. As data from every voxel was included in the model, a fixed effect regressor that quantified the mean-neighboring value of each surrounding voxel (i.e. 6 voxels that share face) was also included to account for spatial correlation across voxels. Random effects were participant, voxel, mean-neighboring value as a random slope by voxel. The results do not indicate a significant relationship between the Local-Hreg and GLM β estimates for the HF > LF and HF > PW contrasts ($p = 0.006$, $t = 1.2$, $p = 0.24$).

To determine if the Local-Hreg and the searchlight Hcorr analysis techniques were quantifying the same aspect of the BOLD response (i.e. local heterogeneity) we extracted and z-scored (across participants) the values from overlapping voxels of the Local-Hreg (Fig. 4B) and Hcorr (Fig. 4C) clusters. Pearson correlation values (Fig. 4E) for both HF > LF and HF > PW were significant with r -values > 0.3 and p -values < 0.1. Next, we examined the relationship between Local-Hreg and Hcorr values across an anatomical left vOTC mask, including data from each voxel each contrast (HF > LF and HF > PW) in a single LMEM with Local-Hreg as the dependent variable; fixed effects were the Hcorr values, contrast type, mean-neighboring voxel value. Random effects were participant, voxel, mean-neighboring voxel value as a random slope by voxel. The results reveal a significant positive relationship between the Local-Hreg and searchlight Hcorr values for the HF > LF and HF > PW contrasts ($\beta = 0.006$, $t = 3.4$, $p = 0.0006$).

Finally, it is worth noting that although we did not observe any near significant results for the HF > MF contrast with any of the three analytic approaches, we did observe, and essentially replicated the findings reported in this section via the HF > CS contrast. These results are reported in the Supplementary Materials S2.

This analysis reveals that both the Local-Hreg and Hcorr quantify similar aspects of the local neural processing that are (1) not detected with the standard GLM analysis, and (2) presumably due to relatively high local heterogeneity/sparseness of HF word representations compared to LF words, PWs, and CSs within the left vOTC.

4. Discussion

Our primary goal was to determine if there is a relationship between the familiarity of letter strings and the degree of local heterogeneity of task-related BOLD responses across adjacent voxels that are involved in their representation. To achieve this, we developed a novel method for examining local neural heterogeneity: Local-Heterogeneity Regression (Local-Hreg). We report 4 key findings within the left vOTC: (1) In confirmation of previous findings (e.g., Kronbichler et al., 2004) that the greater the orthographic frequency of presented alphabetic stimuli (high frequency words, middle frequency, low frequency words,

pseudowords and consonant strings) the *lower* the relative mean BOLD response within the left vOTC. (2) Conversely, via the novel Local-Hreg technique we determined that the greater the orthographic frequency of presented alphabetic stimuli the *greater* the relative local heterogeneity of local BOLD responses. This provides support for the hypothesis that, with learning, letter strings develop better-differentiated and more compact representations. (3) Although there was spatial overlap in the left vOTC between the results of mean-signal standard GLM and Local-Hreg analyses, the two measures were not correlated. This indicates that Local-Hreg quantifies an aspect of the local BOLD response - local heterogeneity - that is not easily detected via standard univariate approaches. (4) There was a clear relationship between Local-Hreg and Hcorr analysis results. Specifically, there was spatial overlap of the results (although only if a lenient corrected threshold was applied for the Hcorr analysis) and, furthermore, individual participant Local-Hreg and Hcorr values were correlated. These findings support the conclusion that both approaches constitute measures of the local heterogeneity, although Local-Hreg provides a more robust approach for doing so. Combined, these findings demonstrate that there are word frequency-based differences in BOLD heterogeneity within the left vOTC such that of well-learned letter strings exhibit greater local heterogeneity than those of less well-learned or unfamiliar letter strings.

4.1. Implications for understanding the neural bases of orthographic representation and processing

A great many neuroimaging findings indicate that the left vOTC encodes orthographic neural representations, and that the neural responses within this region are sensitive to the extent of subjects' experiences with letter strings. For instance, mean BOLD response within this region has been found to be linearly dependent on the written frequency of stimuli (Fiez et al., 1999; Kronbichler et al., 2004), with well-learned alphabetic stimuli being associated with lower mean BOLD responses within the left vOTC relative to less well-learned stimuli. Furthermore, it has been reported there is a response gradient of increasing string size and familiarity along the posterior to anterior axis of the left vOTC (Dehaene et al., 2005; Vinckier et al., 2007) such that more posterior left vOTC is responsive to less familiar, smaller size alphabetic strings (e.g. infrequent letters or bigrams) and more anterior vOTC is more responsive to more familiar, larger unit sizes (e.g., frequent trigrams and whole words). In addition, fMRI adaptation studies indicate that the more anterior vOTC region includes narrowly tuned representations for words that contrast with more distributed representations for pseudowords (Glezer et al., 2009, 2015). Also consistent with the orthographic experience-dependent sensitivity of left vOTC, research has found that the responsivity of this region changes with literacy acquisition (Dehaene et al., 2010) and, more specifically, that the precision of the tuning changes during the course of word learning (Glezer et al., 2015).

The Local-Hreg (and Hcorr) findings reported here support and build on this previous work. This work supports the proposal that there are highly tuned orthographic representations within the left vOTC, because across the orthographic network, it was specifically within the left vOTC that the Local-Hreg analysis identified an area with relatively high local neural heterogeneity for familiar relative to unfamiliar letter strings. The findings significantly

extend previous results by furthering our understanding of the nature of the experience-dependent neural code. By quantifying the similarity of neural responses across adjacent voxels, Local-Hreg provides a more direct evaluation of the prediction of the sparse coding hypothesis; i.e., highly-learned representations should yield greater heterogeneity of the local neural response.

A number of previous studies have reported that the left dorsal IFG is relevant to orthographic processing for both reading and spelling (Rapp and Lipka, 2011; Purcell et al., 2017). However, we did not find experience-dependent sensitivity in this area, as measured by Local-Hreg. One interpretation of the role of the left IFG is that it contains task dependent representations which coordinate orthographic and phonological representations found in more posterior brain areas (e.g. left vOTC) (Mesulam, 1998; Booth et al., 2002; Bitan et al., 2005). Specifically, this region could be involved in selecting among competing orthographic representations (e.g., Rapp and Lipka, 2011; Purcell et al., 2017; for similar claims in speech production see: Thompson-Schill et al., 1998; Schnur et al., 2009). Given these interpretations, we might not expect to be able to quantify the relative differentiation of representations within the IFG using the same Local-Hreg approach used here (e.g. other approaches could use different search light sizes/shapes, disjoint voxels, or stimulus conditions).

Furthermore, it is worth considering the potential relevance of Local-Hreg in examining the neural bases of competing hypotheses regarding orthographic representations that are predicted by current models of reading. For instance, models such as the Dual Route Cascaded model (Coltheart et al., 2001) or the Triangle model (Plaut et al., 1996; Harm and Seidenberg, 2004) make different claims regarding the degree to which orthographic representations are local or distributed. Local-Hreg may provide a tool for examining predictions (if any) regarding the nature of the underlying neural representations that might be generated by these models. For example, the degree to which local heterogeneity can be identified in different sizes of searchlight (or regions of interest) could be relevant. The work reported here was limited to the smallest searchlight and explored the clustering of representational heterogeneity across spans of three adjacent voxels; searchlights that cover a larger volume may reveal differences in local heterogeneity for orthographic representations that are relatively more distributed (i.e. that span a larger number of voxels).

Local-heterogeneity may provide a new way to test theories and examine predictions of computational models of reading/spelling as they relate to the brain and, more generally, a novel approach for using functional neuroimaging to inform cognitive models (Coltheart, 2013).

4.2. Local-Hreg and mean BOLD response

Given that there were mean BOLD response differences between the HF, LF, PW and CS in voxels that also exhibited Local-Hreg differences between the same conditions, the question arises regarding whether standard within-voxel, mean signal-based GLM approaches might also be sensitive to the relative heterogeneity of local representations for written words. To address this we explicitly examined the relationship between mean BOLD response differences and Local-Hreg differences in multiple ways. Due to the fact that Local-Hreg

analysis incorporates (and removes) the mean BOLD response, as expected we did not observe a significant relationship between mean BOLD response difference values and the Local-Hreg difference values, even when incorporating data from every voxel in a region of interest (left vOTC) into a LMEM that used subject and voxels as random effects. This supports the conclusion that Local-Hreg measures an aspect of local neural representation that is not readily detected via traditional within-voxel, mean signal-based approaches. Of course, it is possible that there is still a relationship that was not detected in this work (i.e. a null result does not prove there is no relationship).

4.3. Local-Hreg: an adaptation of gPPI

As a regression based approach, gPPI analysis allows for the “selection” of variance associated specific portions of a time series (McLaren et al., 2012). Critically, the interaction terms quantify the condition-specific activation relationships between a seed region and another target region in a way that is independent of general task-related signal fluctuations and presumed physiologically-based components. Local-Hreg, by implementing gPPI to examine local connectivity, has access to the interaction terms for different stimulus conditions, allowing - in this context - for a comparison of the degree of local, cross-voxel similarity (heterogeneity) associated with letter strings of different frequency. Furthermore, the implementation of Local-Hreg as a searchlight analysis allows for this without the need to identify specific source and target voxels.

Local-Hreg has advantages over previous heterogeneity analysis approaches such as the Hcorr. Whereas Hcorr segments out condition-specific portions of the time series and then quantifies the variability of their cross-voxel correlations (Jiang et al., 2013), Local-Hreg uses a regression-based approach to analytically select the condition-specific variance of interest and explicitly model the amount of shared condition-specific variance across voxels. Although both approaches quantify the relative amount of local condition-specific heterogeneity in algorithmically very different ways, high Local-Hreg and Hcorr values were observed in the same left vOTC region and were correlated. This confirms that they both quantify the same aspect of the local BOLD signal (i.e., local heterogeneity), although Local-Hreg is a more robust method for doing so.

4.4. Comparison of Local-Hreg to other approaches to investigating neural representation: MVPA and fMRI-A

Multi-voxel pattern analysis (MVPA) assumes that neural representations are spatially distributed across multiple voxels and that the content of these representations can be inferred when cross-voxel patterns distinguish specific representational features/dimensions (Haxby et al., 2001; Cox and Savoy, 2003; Kriegeskorte et al., 2008; Haxby, 2012; Mahmoudi et al., 2012). Although algorithmically very different from Local-Hreg, one key similarity between the approaches is that both are insensitive to the mean level of activity in a brain region (Davis et al., 2014) since both examine whether neural patterns differentiate two conditions independently of the average neural response. The fundamental difference between the methods is that whereas MVPA first computes the mean BOLD response for each voxel and then statistically evaluates the similarity of the cross-voxel patterns, Local-Hreg first computes cross-voxel similarity (each voxel with its neighbors) and then

statistically evaluates these values for each voxel. By first calculating the mean neural response in each voxel via a standard univariate GLM, the MVPA approach reduces the dimensionality of the data from high dimensional voxel time-series to low-dimensional voxel activation values (i.e., GLM beta values). Local-Hreg, instead, quantifies the shared condition-specific similarity of voxel time series prior to reducing the dimensionality of the data. The Local-Hreg approach considers subtle similarities/differences of the neural responses within given conditions over the time-series, although it does not evaluate representational content that may be distributed across voxels.

MVPA classification approaches have also been used to index the integrity or strength of neural representations by considering the classifiability of distributed activation patterns. For example, this has been applied to examine the integrity of visual object representations in schizophrenia (Yoon et al., 2008), face/object classification in congenital prosopagnosia (Rivolta et al., 2014) and face processing in autism spectrum disorder (Coutanche et al., 2011). Although a useful tool, it does not provide a clear picture of the basis for the classification performance. For example, in Rivolta et al. (2014) poor classification of face/object cross-voxel neural patterns from congenital prosopagnosics could have occurred because either faces, objects, or both were “poorly” represented; there is no information about the relative differentiability of the categories themselves, only that there is poor differentiation between them. On the other hand, Local-Hreg provides a more direct measure of the representational integrity of individual categories, assuming that higher integrity is associated with greater differentiability/heterogeneity of the local neural representations. Importantly, the two approaches complement each other, with MVPA identifying representational content and Local-Hreg indexing neural differentiability.

Finally, it is worth pointing out that although Local-Hreg was developed to quantify the local differentiation of BOLD responses for different stimulus categories, presumably it could be used to evaluate differentiation across learning/exposure time-points for individual stimuli. For instance, Xue et al. (2010) examined patterns of neural responses across individual items as they were learned, observing that local patterns for better learned items had more similar (more consistent) patterns of BOLD responses across the learning trajectory compared to patterns for poorly learned items. For instance the pattern of BOLD responses associated with learning a specific face stimulus was more similar upon repeated presentations of that specific subsequently well-learned face (Xue et al., 2010). We would predict that with learning, representations would not only be more consistent across trials, but would also have correspondingly higher local differentiation. It would be important in future work to determine if there is a relationship between the consistency of neural response patterns across time and their degree of differentiation.

fMRI adaptation (fMRI-A) examines neural tuning properties based on the well-established finding in single-unit recording that repeated firing leads to decreased firing rates (Desimone, 1996). In fMRI-A, the similarity of stimulus features is manipulated and then the degree of adaptation of the BOLD response in a brain area is used to infer its representational content (Grill-Spector and Malach, 2001; Henson and Rugg, 2003; Grill-Spector, 2006; Grill-Spector et al., 2006; Malach, 2012; Barron et al., 2016; Larsson et al., 2016). For example, based on fMRI-A results, Glezer et al. (2009, 2015) argued that words

have a higher degree of tuning/selectivity than pseudowords. Specifically, they found that neural adaptation for words occurs only for repeated presentations of the same word strings, whereas pseudoword adaptation was observed for word strings that differed by one letter. On this basis, they argued that there are local, highly tuned orthographic word representations in left vOTC. This claim was apparently at odds with the Baeck et al. (2015) MVPA-based finding (2015; see also Stevens et al., 2017) that although the left vOTC can distinguish the categories of words from pseudowords, it cannot make within-category distinctions between individual items. However, we suggest that MPVA may not be well-suited to address this type of question or for quantifying the degree of neural tuning of specific representations because it relies on reducing the BOLD responses to a single beta-value in each voxel, making it less sensitive to more fine-grained within or across-voxel differences, something that both Local-Hreg and neural adaptation techniques are better suited for.

Finally, although successfully used to examine neuronal tuning, the neural basis of the adaptation is not entirely clear; different proposals include: 1) the neuronal fatigue model: the originally responding neurons respond less, 2) the facilitation model: the original responding neurons respond more quickly, or 3) the sharpening model: fewer neurons respond as the neural response becomes more sharply tuned (Grill-Spector et al., 2006). We propose that the Local-Hreg method could be used to test the sharpening model hypothesis assuming that a key prediction is that heterogeneity should increase with increasing adaptation effects. This is an example of a way in which Local-Hreg could be used to examine and complement fMRI adaptation results.

4.5. Limitations and future studies

We have shown that Local-Hreg (similar to other multi-voxel approaches) quantifies a unique aspect of BOLD activity relative to the mean response. Furthermore, the pattern of Local-Hreg results is consistent with the interpretation that representational sparseness increases with learning. However, other accounts of these observations may be possible and further work is required to establish a better understanding of the underlying neural processes that give rise to the observed association of local, cross-voxel heterogeneity with stimulus familiarity. In this regard, further empirical work involving a range of learning situations (e.g., adults and children and relearning after brain injury; see Purcell et al., In Preparation) will be necessary for testing predictions established in this work. Furthermore, computational simulation would be especially important for assessing the viability of a learning-based account that assumes that exposure-based neural changes lead to local heterogeneity at the voxel level. Additionally, simulations incorporating estimates of both mean BOLD and Local-Hreg could help clarify under what, if any, circumstances there is a relationship between these measures. Furthermore, optimal designs for quantifying condition-specific local heterogeneity could be derived from simulations. That is, different designs (e.g. jittered timing, counterbalance schemes) may be optimally suited to quantify one measure or the other (or both simultaneously). Despite these clear limitations in our understanding of the underlying mechanisms of learning-related local heterogeneity, this study provides a strong empirical basis for this future work.

5. Summary

Local-Hreg constitutes a novel approach for quantifying the similarity of task-based, condition-specific BOLD responses across voxels within a local region. The work reported here specifically supports the claim that the experience-dependent orthographic representations observed in the left vOTC vary in terms of the local heterogeneity of the BOLD response across voxels, which could, in turn, reflect the relative sparseness of their representations. That is, word strings with which we have more experience may yield sparser representations and more heterogeneous response patterns across adjacent voxels. Local-Hreg is an approach that can be used to explore experience-dependent learning across a wide range of cognitive domains, and adds to our analytic toolbox by providing a general method for quantifying local differentiation of neural representations across adjacent voxels.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We acknowledge NIH support (DC012283) support for the multi-site project examining the neurobiology of language recovery in aphasia that this work forms a part of. The authors declare no competing financial interests.

References

References

- Abdulrahman H, Henson RN, 2016 Effect of trial-to-trial variability on optimal event-related fMRI design: implications for Beta-series correlation and multi-voxel pattern analysis. *Neuroimage* 125, 756–766. [PubMed: 26549299]
- Aguirre GK, 2007 Continuous carry-over designs for fMRI. *Neuroimage* 35, 1480–1494. [PubMed: 17376705]
- Baech A, Kravitz D, Baker C, Op de Beeck HP, 2015 Influence of lexical status and orthographic similarity on the multi-voxel response of the Visual Word Form Area. *Neuroimage* 111, 321–328. [PubMed: 25665965]
- Bair W, Zohary E, Newsome WT, 2001 Correlated firing in macaque visual area MT: time scales and relationship to behavior. *J Neurosci Off J Soc Neurosci* 21 1676–1697.
- Baker CI, Liu J, Wald LL, Kwong KK, Benner T, Kanwisher N, 2007 Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proc. Natl. Acad. Sci. U. S. A.* 104, 9087–9092. [PubMed: 17502592]
- Barron HC, Garvert MM, Behrens TEJ, 2016 Repetition suppression: a means to index neural representations using BOLD? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 371.
- Binder JR, Medler DA, Westbury CF, Liebenthal E, Buchanan L, 2006 Tuning of the human left fusiform gyrus to sublexical orthographic structure. *Neuroimage* 33, 739–748. [PubMed: 16956773]
- Bitan T, Booth JR, Choy J, Burman DD, Gitelman DR, Mesulam MM, 2005 Shifts of effective connectivity within a language network during rhyming and spelling. *J. Neurosci.* 25, 5397–5403. [PubMed: 15930389]
- Booth JR, Burman DD, Meyer JR, Gitelman DR, Parrish TB, Mesulam MM, 2002 Functional anatomy of intra- and cross-modal lexical tasks. *Neuroimage* 16, 7–22. [PubMed: 11969313]
- Carreiras M, Mechelli A, Price CJ, 2006 Effect of word and syllable frequency on activation during lexical decision and reading aloud. *Hum. Brain Mapp.* 27, 963–972. [PubMed: 16628608]
- Coltheart M, 2013 How can functional neuroimaging inform cognitive theories? *Perspect. Psychol. Sci.* 8 (1), 98–103. 10.1177/1745691612469208. [PubMed: 26172256]

- Coltheart M, Rastle K, Perry C, Langdon R, Ziegler J, 2001 DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychol. Rev.* 108 (1), 204–256. [PubMed: 11212628]
- Coutanche MN, Thompson-Schill SL, Schultz RT, 2011 Multi-voxel pattern analysis of fMRI data predicts clinical symptom severity. *Neuroimage* 57, 113–123. [PubMed: 21513803]
- Cox DD, Savoy RL, 2003 Functional magnetic resonance imaging (fMRI) “brain reading”: detecting and classifying distributed patterns of fMRI activity in human visual cortex. *Neuroimage* 19, 261–270. [PubMed: 12814577]
- Davis T, LaRocque KF, Mumford JA, Norman KA, Wagner AD, Poldrack RA, 2014 What do differences between multi-voxel and univariate analysis mean? How subject-, voxel-, and trial-level variance impact fMRI analysis. *Neuroimage* 97, 271–283. [PubMed: 24768930]
- Dehaene S, Cohen L, 2011 The unique role of the visual word form area in reading. *Trends Cognit. Sci.* 15, 254–262. [PubMed: 21592844]
- Dehaene S, Cohen L, Sigman M, Vinckier F, 2005 The neural code for written words: a proposal. *Trends Cognit. Sci.* 9, 335–341. [PubMed: 15951224]
- Dehaene S, Le Clec HG, Poline JB, Le Bihan D, Cohen L, 2002 The visual word form area: a prelexical representation of visual words in the fusiform gyrus. *Neuroreport* 13, 321–325. [PubMed: 11930131]
- Dehaene S, Pegado F, Braga LW, Ventura P, Nunes Filho G, Jobert A, Dehaene- Lambertz G, Kolinsky R, Morais J, Cohen L, 2010 How learning to read changes the cortical networks for vision and language. *Science* 330, 1359–1364. [PubMed: 21071632]
- Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS, Killiany RJ, 2006 An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31, 968–980. [PubMed: 16530430]
- Desimone R, 1996 Neural mechanisms for visual memory and their role in attention. *Proc Natl Acad Sci U A* 93, 13494–13499.
- Eklund A, Nichols TE, Knutsson H, 2016 Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Proc. Natl. Acad. Sci. Unit. States Am.* 113, 7900–7905.
- Field DJ, 1987 Relations between the statistics of natural images and the response properties of cortical cells. *JOSA A* 4, 2379–2394.
- Fiez JA, Balota DA, Raichle ME, Petersen SE, 1999 Effects of lexicality, frequency, and spelling-to-sound consistency on the functional anatomy of reading. *Neuron* 24, 205–218. [PubMed: 10677038]
- Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC, 1995 Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn Reson Med Off J Soc Magn Reson Med Soc Magn Reson Med* 33, 636–647.
- Freedman DJ, Riesenhuber M, Poggio T, Miller EK, 2006 Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex. *Cerebr. Cortex* 16, 1631–1644.
- Friston KJ, Buechel C, Fink GR, Morris J, Rolls E, Dolan RJ, 1997 Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage* 6, 218–229. [PubMed: 9344826]
- Friston KJ, Holmes AP, Worsley KJ, Poline J-P, Frith CD, Frackowiak RSJ, 1994 Statistical parametric maps in functional imaging: a general linear approach. *Hum. Brain Mapp.* 2, 189–210.
- Gaillard R, Naccache L, Pinel P, Clemenceau S, Volle E, Hasboun D, Dupont S, Baulac M, Dehaene S, Adam C, Cohen L, 2006 Direct intracranial, FMRI, and lesion evidence for the causal role of left inferotemporal cortex in reading. *Neuron* 50, 191–204. [PubMed: 16630832]
- Glezer LS, Jiang X, Riesenhuber M, 2009 Evidence for highly selective neuronal tuning to whole words in the “visual word form area”. *Neuron* 62, 199–204. [PubMed: 19409265]
- Glezer LS, Kim J, Rule J, Jiang X, Riesenhuber M, 2015 Adding words to the Brain’s visual dictionary: novel word learning selectively sharpens orthographic representations in the VWFA. *J. Neurosci.* 35, 4965–4972. [PubMed: 25810526]
- Grill-Spector K, 2006 Selectivity of adaptation in single units: implications for FMRI experiments. *Neuron* 49, 170–171. [PubMed: 16423690]
- Grill-Spector K, Henson RN, Martin A, 2006 Repetition and the brain: neural models of stimulus-specific effects. *Trends Cognit. Sci.* 10, 14–23. [PubMed: 16321563]

- Grill-Spector K, Malach R, 2001 fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychol Amst* 107, 293–321. [PubMed: 11388140]
- Hanna PR, Hanna JS, Hodges RE, Rudorf EH, 1966 Phoneme-grapheme Correspondences as Cues to Spelling Improvement.
- Harm Michael W., Seidenberg Mark S., 2004 Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychol. Rev.* 111 (3), 662–720. 10.1037/0033-295X.111.3.662. [PubMed: 15250780]
- Haxby JV, 2012 Multivariate pattern analysis of fMRI: the early beginnings. *Neuroimage* 62, 852–855. [PubMed: 22425670]
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P, 2001 Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293, 2425–2430. [PubMed: 11577229]
- Henson RN, Rugg MD, 2003 Neural response suppression, haemodynamic repetition effects, and behavioural priming. *Neuropsychologia* 41, 263–270. [PubMed: 12457752]
- Jermakowicz WJ, Chen X, Khaytin I, Bonds AB, Casagrande VA, 2009 Relationship between spontaneous and evoked spike-time correlations in primate visual cortex. *J. Neurophysiol.* 101, 2279–2289. [PubMed: 19211656]
- Jiang X, Bollich A, Cox P, Hyder E, James J, Gowani SA, Hadjikhani N, Blanz V, Manoach DS, Barton JJS, Gaillard WD, Riesenhuber M, 2013 A quantitative link between face discrimination deficits and neuronal selectivity for faces in autism. *NeuroImage Clin* 2, 320–331. [PubMed: 24179786]
- Jiang X, Petok JR, Howard DV, Howard JH, 2017 Individual differences in cognitive function in older adults predicted by neuronal selectivity at corresponding brain regions. *Front. Aging Neurosci.* 9 Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5394166/>. (Accessed 25 May 2018).
- Jiang X, Rosen E, Zeffiro TA, VanMeter J, Blanz V, Riesenhuber M, 2006 Evaluation of a shape-based model of human face discrimination using FMRI and behavioral techniques. *Neuron* 50, 159–172. [PubMed: 16600863]
- Kobatake E, Wang G, Tanaka K, 1998 Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. *J. Neurophysiol.* 80, 324–330. [PubMed: 9658053]
- Kriegeskorte N, Goebel R, Bandettini P, 2006 Information-based functional brain mapping. *Proc. Natl. Acad. Sci. U. S. A.* 103, 3863–3868. [PubMed: 16537458]
- Kriegeskorte N, Mur M, Bandettini P, 2008 Representational similarity analysis - connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* 2 Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2605405/>. (Accessed 15 December 2016).
- Kronbichler M, Hutzler F, Wimmer H, Mair A, Staffen W, Ladurner G, 2004 The visual word form area and the frequency with which words are encountered: evidence from a parametric fMRI study. *Neuroimage* 21, 946–953. [PubMed: 15006661]
- Kuznetsova A, Brockhoff PB, Christensen RHB, 2017 lmerTest package: tests in linear mixed effects models. *J. Stat. Software* 82 Available at: <https://www.jstatsoft.org/article/view/v082i13>. (Accessed 23 July 2018).
- Larsson J, Solomon SG, Kohn A, 2016 fMRI adaptation revisited. *Cortex J Devoted Study Nerv Syst Behav* 80, 154–160.
- Mahmoudi A, Takerkart S, Regragui F, Boussaoud D, Brovelli A, 2012 Multivoxel pattern analysis for FMRI data: a review. *Comput Math Methods Med* 2012, 961257. [PubMed: 23401720]
- Mahon BZ, Milleville SC, Negri GAL, Rumiati RI, Caramazza A, Martin A, 2007 Action-related properties shape object representations in the ventral stream. *Neuron* 55, 507–520. [PubMed: 17678861]
- Malach R, 2012 Targeting the functional properties of cortical neurons using fMR- adaptation. *Neuroimage* 62, 1163–1169. [PubMed: 22245340]
- Martin A, Schurz M, Kronbichler M, Richlan F, 2015 Reading in the brain of children and adults: a meta-analysis of 40 functional magnetic resonance imaging studies. *Hum. Brain Mapp.* 36, 1963–1981. [PubMed: 25628041]
- McCandliss BD, Cohen L, Dehaene S, 2003 The visual word form area: expertise for reading in the fusiform gyrus. *Trends Cognit. Sci.* 7, 293–299. [PubMed: 12860187]

- McLaren DG, Ries ML, Xu G, Johnson SC, 2012 A generalized form of context- dependent psychophysiological interactions (gPPI): a comparison to standard approaches. *Neuroimage* 61, 1277–1286. [PubMed: 22484411]
- Mechelli A, Gorno-Tempini ML, Price CJ, 2003 Neuroimaging studies of word and pseudoword reading: consistencies, inconsistencies, and limitations. *J. Cognit. Neurosci.* 15, 260–271. [PubMed: 12676063]
- Medler DA, Binder JR, 2005 MCWord: an on-line orthographic database of the English language. <http://www.neuro.mcw.edu/mcword/>.
- Mesulam MM, 1998 From sensation to cognition. *Brain* 121 (Pt 6), 1013–1052. [PubMed: 9648540]
- Nichols TE, Holmes AP, 2002 Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum. Brain Mapp.* 15, 1–25. [PubMed: 11747097]
- Olshausen BA, Field DJ, 1996 Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* 381, 607–609. [PubMed: 8637596]
- Plaut DC, McClelland JL, Seidenberg MS, Patterson K, 1996 Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychol. Rev.* 103 (1), 56–115. [PubMed: 8650300]
- Purcell JJ, Jiang X, Eden GF, 2017 Shared orthographic neuronal representations for spelling and reading. *Neuroimage* 147, 554–567. [PubMed: 28011250]
- Purcell J.J., Wiley R.W., Rapp B., (In Preparation) Increased Differentiation of Neural Representations Supports Post-stroke Language.
- Rapp B, Lipka K, 2011 The literate brain: the relationship between spelling and reading. *J. Cognit. Neurosci.* 23, 1180–1197. [PubMed: 20433242]
- Rapp B, Purcell JJ, Hillis AE, Capasso R, Miceli M, 2015 Neural bases of orthographic long-term memory and working memory in dysgraphia. *Brain* awv 348.
- Riesenhuber M, 2007 Appearance Isn't everything: news on object representation in cortex. *Neuron* 55, 341–344. [PubMed: 17678848]
- Rivolta D, Woolgar A, Palermo R, Butko M, Schmalzl L, Williams MA, 2014 Multi-voxel pattern analysis (MVPA) reveals abnormal fMRI activity in both the “core” and “extended” face network in congenital prosopagnosia. *Front. Hum. Neurosci.* 8 Available at. <http://journal.frontiersin.org/article/10.3389/fnhum.2014.00925/full>. (Accessed 15 December 2016).
- Rolls ET, Tovee MJ, 1995 Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. *J. Neurophysiol.* 73, 713–726. [PubMed: 7760130]
- Rolls ET, Treves A, 2011 The neuronal encoding of information in the brain. *Prog. Neurobiol.* 95, 448–490. [PubMed: 21907758]
- Schnur TT, Schwartz MF, Kimberg DY, Hirshorn E, Coslett HB, Thompson- Schill SL, 2009 Localizing interference during naming: convergent neuroimaging and neuropsychological evidence for the function of Broca's area. *Proc. Natl. Acad. Sci. U. S. A.* 106, 322–327. [PubMed: 19118194]
- Soltysik David A., Thomasson David, Rajan Sunder, Biassou Nadia, 2015 Improving the use of principal component analysis to reduce physiological noise and motion artifacts to increase the sensitivity of task-based fMRI. *J. Neurosci. Meth.* 241, 18–29. 10.1016/j.jneumeth.2014.11.015.
- Stevens WD, Kravitz DJ, Peng CS, Tessler MH, Martin A, 2017 Privileged functional connectivity between the visual word form area and the language system. *J Neurosci Off J Soc Neurosci* 37, 5288–5297.
- Thompson-Schill SL, Swick D, Farah MJ, D'Esposito M, Kan IP, Knight RT, 1998 Verb generation in patients with focal frontal lesions: a neuropsychological test of neuroimaging findings. *Proc. Natl. Acad. Sci. U. S. A.* 95, 15855–15860. [PubMed: 9861060]
- Ts'o DY, Gilbert CD, Wiesel TN, 1986 Relationships between horizontal interactions and functional architecture in cat striate cortex as revealed by cross-correlation analysis. *J Neurosci Off J Soc Neurosci* 6, 1160–1170.
- Vinckier F, Dehaene S, Jobert A, Dubus JP, Sigman M, Cohen L, 2007 Hierarchical coding of letter strings in the ventral stream: dissecting the inner organization of the visual word-form system. *Neuron* 55, 143–156. [PubMed: 17610823]

- Vinje WE, Gallant JL, 2000 Sparse coding and decorrelation in primary visual cortex during natural vision. *Science* 287, 1273–1276. [PubMed: 10678835]
- Xia M, Wang J, He Y, 2013 BrainNet viewer: a network visualization tool for human brain connectomics. *PLoS One* 8, e68910. [PubMed: 23861951]
- Xue G, Dong Q, Chen C, Lu Z, Mumford JA, Poldrack RA, 2010 Greater neural pattern similarity across repetitions is associated with better memory. *Science* 330, 97–101. [PubMed: 20829453]
- Yoon JH, Tamir D, Minzenberg MJ, Ragland JD, Ursu S, Carter CS, 2008 Multivariate pattern analysis of fMRI data reveals deficits in distributed representations in schizophrenia. *Biol. Psychiatr.* 64, 1035–1041.
- Zang Y, Jiang T, Lu Y, He Y, Tian L, 2004 Regional homogeneity approach to fMRI data analysis. *Neuroimage* 22, 394–400. [PubMed: 15110032]

Fig. 1. Functionally defined Orthographic Network based on the Standard GLM contrast (HF, MF, LF, PW, CS) > Fixation projected onto a smoothed inflated surface rendering using BrainNet viewer (Xia et al., 2013); voxel-wise corrected $p < 0.01$, cluster size corrected $p < 0.05$. This functional map was used as an inclusive functional mask in subsequent analyses

Fig. 2.

Results obtained from the Parametric Linear Analyses within the Orthographic Network (Fig. 1) projected onto left hemisphere ventral surfaces using BrainNet viewer (Xia et al., 2013). (A) Results of standard GLM HF > MF > LF > PW > CS linear relationship (voxel-wise $p < 0.05$; cluster threshold $p < 0.05$) left vOTC cluster (size = 222 voxels; MNI peak = 47, -66, -14). (B) Bars reports the β estimates for each participant and condition extracted from the cluster reported in A. Error bars are the standard error of the mean (SEM), and the dotted line depict the positive linear relationship across the conditions reported at the bottom of bar plot. (C) Results of Local-Hreg HF > MF > LF > PW > CS linear relationship (voxel-wise $p < 0.05$) left vOTC cluster (size = 36 voxels; MNI peak = -51, -69, -19). (D) Bars reports the Local-Hreg values for each participant and condition extracted from the cluster reported in C. Error bars are the SEM, and the dotted line depict the negative linear relationship across the conditions reported at the bottom of bar plot.

Fig. 3.

Results from the Local-Hreg analysis of HF vs. LF and HF vs. PW conditions within the Orthographic Network, projected onto left hemisphere ventral surfaces using BrainNet viewer (Xia et al., 2013). (A) HF > LF words (MNI peak = $-53, -45, -23$; cluster size = 187 voxels; corrected $p = 0.04$). (B) HF words > PWs (MNI peak = $-50, -44, -23$; cluster size = 98 voxels; voxels; corrected $p = 0.09$).

Fig. 4.

Direct comparison of Local-Hreg with Hcorr and with standard GLM analysis. A-C results constrained inclusively by a left vOTC anatomical mask (green outline). Statistical maps are presented on left hemisphere only axial slices (z-plane listed to the left); voxel wise $p < 0.05$ for all maps and cluster corrected p values are reported in the upper right of each slice. (A) Standard GLM (Analysis 1) for the contrasts: HF > LF (top) and HF > PW (bottom). (B) Local-Hreg clusters (Analysis 2) clusters for the following contrasts: HF > LF (top) and HF > PW (bottom). (C) Hcorr clusters (Analysis 3) for the contrasts: HF > LF (top) and HF > PW (bottom). (D) Scatter plot with trend-line of Local-Hreg and GLM β estimate extracted for each participant from the intersecting clusters in (A) and (B); each dot is a participant. Correlation coefficients and p-values are reported in the upper right. (E) Scatter plots with trend-lines of average Local-Hreg and Hcorr values extracted for each participant from the intersecting clusters in (B) and (C); each dot is a participant. Correlation coefficients and p-values are reported in the upper right. Overall, these data indicate that Local-Hreg is a robust measure of local heterogeneity, and Local-Hreg measures are related to Hcorr values, but not to the standard GLM β estimate values.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1:

Summary of stimuli for each study.

	High Frequency (HF) Words				Mid Frequency (MF) Words				Low Frequency (LF) Words				Pseudowords (PW)		Consonant Strings (CS)	
	Freq	ONC	ONF	PGC	PGC	Freq	ONC	ONF	PGC	PGC	ONC	ONF	ONC	ONF	ONC	ONF
Combined	86	2	18	62	62	19	5	25	63	63	2	4	22	50	2.5	15.5
Study 1	93 (64)	2 (4)	21 (68)	59 (19)	59 (19)	20 (10)	5(6)	33 (76)	67 (20)	62 (25)	1 (1)	3 (4)	33 (76)	62 (25)	2 (3)	13 (24)
Study 2	78 (41)	2 (4)	15 (46)	65 (19)	65 (19)	18 (10)	4(5)	17 (30)	58 (24)	58 (29)	2 (1)	5 (5)	31 (78)	58 (29)	3(4)	18 (62)

Table 2:

Peak coordinates for the Orthographic Network defined via the standard RFX- GLM (HF, MF, LF, PW, CS) > Fixation contrast. N = 30 (Voxel-wise $p < 0.01$, Cluster level corrected $p < 0.05$).

Region (Estimated from Harvard/Oxford Atlas)		voxels	Tmax	MNI Coordinates		
				x	y	z
Occipitotemporal	L mid vOTC	1153	9.41	-45	-63	-15
	L inf LOC		8.22	-39	-87	-9
	L mid vOTC		7.08	-48	-51	-21
	L post vOTC		6.48	-21	-90	-9
	L inf LOC		4.97	-42	-69	3
	R mid vOTC	841	7.18	42	-66	-18
	R inf LOC		7.17	30	-90	-3
	R post vOTC		5.46	33	-78	-18
	R mid vOTC		4.89	39	-54	-24
	R mid vOTC		4.85	39	-42	-21
Frontal	SMA	534	6.88	0	3	66
	Precentral Gyrus		5.17	0	-30	63
	L Precentral Gyrus		4.25	-3	-18	75
	R Precentral Gyrus		4.16	9	-15	75
	L Postcentral Gyrus		4.11	-9	-36	72
	L Precentral Gyrus	1405	6.47	-48	6	30
	L Precentral Gyrus		6.38	-60	6	21
	L Precentral Gyrus		5.94	-48	-3	57
	L IFG pars triangularis		5.80	-51	27	0
	L Precentral Gyrus		5.80	-51	-6	42
	R Precentral Gyrus	222	5.02	60	0	21
	R Precentral Gyrus		5.00	54	-3	33
	R Precentral Gyrus		4.86	51	0	48
Parietal	L SPL	174	5.56	-30	-57	54
	L superior LOC		3.62	-27	-66	42
	L superior LOC		3.56	-24	-69	27

MNI-Montreal Neurological Institute; L - Left; R - Right; IFG-Inferior Frontal Gyrus; MTG-Middle Temporal Gyrus; vOTC - ventral occipitotemporal cortex; LOC - Lateral Occipital Cortex; SPL-Superior Parietal Lobe; IFG-Inferior Frontal Gyrus.