The effect of spatial resolution on decoding accuracy in fMRI multivariate pattern analysis

NeuroImage 132 (2016) 32–42

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

The effect of spatial resolution on decoding accuracy in fMRI multivariate pattern analysis

Anna Gardumi [a,b,*], Dimo Ivanov [a,b], Lars Hausfeld [a,b], Giancarlo Valente [a,b], Elia Formisano [a,b], Kâmil Uludağ [a,b]

[a] Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands
[b] Maastricht Brain Imaging Center, Maastricht University, Maastricht, The Netherlands

ARTICLE INFO

Article history:
Received 11 September 2015
Accepted 10 February 2016
Available online 17 February 2016

Keywords:
Multivariate pattern analysis
fMRI
7 T
Spatial resolution
Spatial smoothing
Auditory cortex

ABSTRACT

Multivariate pattern analysis (MVPA) in fMRI has been used to extract information from distributed cortical activation patterns, which may go undetected in conventional univariate analysis. However, little is known about the physical and physiological underpinnings of MVPA in fMRI as well as about the effect of spatial smoothing on its performance. Several studies have addressed these issues, but their investigation was limited to the visual cortex at 3 T with conflicting results. Here, we used ultra-high field (7 T) fMRI to investigate the effect of spatial resolution and smoothing on decoding of speech content (vowels) and speaker identity from auditory cortical responses. To that end, we acquired high-resolution (1.1 mm isotropic) fMRI data and additionally reconstructed them at 2.2 and 3.3 mm in-plane spatial resolutions from the original k-space data. Furthermore, the data at each resolution were spatially smoothed with different 3D Gaussian kernel sizes (i.e. no smoothing or 1.1, 2.2, 3.3, 4.4, or 8.8 mm kernels). For all spatial resolutions and smoothing kernels, we demonstrate the feasibility of decoding speech content (vowel) and speaker identity at 7 T using support vector machine (SVM) MVPA. In addition, we found that high spatial frequencies are informative for vowel decoding and that the relative contribution of high and low spatial frequencies is different across the two decoding tasks. Moderate smoothing (up to 2.2 mm) improved the accuracies for both decoding of vowels and speakers, possibly due to reduction of noise (e.g. residual motion artifacts or instrument noise) while still preserving information at high spatial frequency. In summary, our results show that – even with the same stimuli and within the same brain areas – the optimal spatial resolution for MVPA in fMRI depends on the specific decoding task of interest.

Introduction

Functional magnetic resonance imaging (fMRI) is currently the most popular non-invasive method to investigate human brain structure and function. It indirectly measures neural activity primarily via the blood oxygenation level-dependent (BOLD) effect. Standard univariate statistical analysis (i.e. general linear model (GLM) analysis) of the task-based fMRI data has been utilized to detect voxel-wise differences of BOLD activation levels and, thus, to infer which brain areas are involved in a certain task. In recent years, multivariate pattern analysis (MVPA) has been used in fMRI to extract information from spatially distributed activation patterns, which may go undetected in conventional univariate analysis. Reliable decoding of information from fMRI data acquired at 3 T has been demonstrated from activation patterns in different brain areas (Haxby et al., 2001; Cox and Savoy, 2003; Haynes and Rees, 2005; Kamitani and Tong, 2005; Kriegeskorte and Bandettini, 2007; Formisano et al., 2008). Different biophysical hypotheses have been proposed to explain the ability of MVPA on fMRI data to detect information inaccessible with GLM. It has been suggested that MVPA is sensitive to information encoded at the sub-millimeter scale of neuronal functional columns. Such information, even if sampled at the lower resolution of standard fMRI voxel sizes (e.g. $3 \times 3 \times 3$ mm^3), may be accessible by MVPA due to local variations and irregularities in the columnar organization, resulting in weak but consistent biases in fMRI responses of the different voxels (Boynton, 2005; Kamitani and Tong, 2005; Haynes and Rees, 2006; Kamitani and Tong, 2006; Kriegeskorte and Bandettini, 2007); this mechanism is, therefore, named *hyperacuity* or *voxel biased sampling*. Alternatively, the transposition from high spatial frequency components of columns preferences to lower spatial frequency of the fMRI signal may be attributed to the cortical vasculature. This hypothesis is based on the fact that, using the standard gradient echo (GE) MRI sequences, the fMRI signal stems mostly from veins draining blood from a given tissue volume (see Uludag et al., 2009). Thus, a specific vein could be more sensitive to one neuronal population than another introducing a spatial bias. Hence, this hypothesis is known as *biased draining regions* (Kamitani and Tong, 2005; Gardner et al., 2006; Kamitani and Tong, 2006; Kriegeskorte and Bandettini, 2007; Gardner, 2010; Kamitani and Sawahata, 2010; Kriegeskorte et al., 2010; Shmuel et al., 2010).

* Corresponding author at: Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, P.O. Box 616, 6200MD Maastricht, The Netherlands.
E-mail address: anna.gardumi@maastrichtuniversity.nl (A. Gardumi).

http://dx.doi.org/10.1016/j.neuroimage.2016.02.033
1053-8119/

According to another hypothesis, MVPA may rely on large spatial scale non-columnar organization (Op de Beeck, 2010), such as radial preference maps (Freeman et al., 2011). Since MVPA represents a computational scheme to non-locally average the fMRI signal, in this framework, MVPA would be able to detect low spatial frequency information too weak to be detected with univariate analysis.

Note that these hypotheses are not mutually exclusive (see Shmuel et al., 2010; Swisher et al., 2010). Nevertheless, they do predict testable effects of spatial smoothing on decoding performance. Op de Beeck has shown that spatial smoothing does not deteriorate decoding performance of objects and orientations from activation patterns in lateral occipital cortex and V1, respectively (Op de Beeck, 2010). He interpreted these results as an argument against hyperacuity and in favor of large-scale organization. Further support for this hypothesis comes from the finding that it is possible to decode across experimental sessions performed in different days (Freeman et al., 2011). In contrast, several studies (Swisher et al., 2010; Alink et al., 2013; Misaki et al., 2013) demonstrated that spatial smoothing decreases decoding accuracies for orientation and ocular dominance from V1 data, suggesting relevant information content at the individual voxel level. The few investigations so far on the underlying mechanisms of MVPA on fMRI data and the effect of spatial smoothing have been limited to the early visual cortex. In addition, they have been restricted to a small set of stimuli and decoding tasks (e.g. decoding of orientation, ocular dominance, and direction of motion) and have yielded conflicting evidence.

The main goal of the current study is to investigate how information at different spatial resolutions contributes to MVPA decoding. We employed ultra-high field (7 T) fMRI to acquire high-resolution data (1.1 mm isotropic), which were then reconstructed at different effective spatial resolutions from original k-space data to evaluate the effects of spatial resolution on MVPA decoding performance. Based on an experimental paradigm and on stimuli that were used in a previous fMRI study at 3 T (Formisano et al., 2008), we presented speech stimuli (vowels) from different speakers and considered the single-trial decoding of vowels and speakers from auditory cortical response patterns. Compared to conventional 3 T fMRI, 7 T fMRI presents several advantages, such as higher signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR), and therefore the possibility of higher spatial resolution with lower partial volume effects and greater spatial specificity (Yacoub et al., 2005; Uludag et al., 2009; Polimeni et al., 2010). On the other hand, it presents challenges such as larger distortions, sensitivity to motion, and larger number of voxels to be handled by the decoding algorithm (Formisano and Kriegeskorte, 2012). Therefore, we also investigated the effects of temporal SNR, CNR, and head motion and of typical noise-reduction steps (spatial smoothing) on MVPA performances.

Material and methods

Subjects

Ten healthy volunteers (seven females, age range 25–32) with normal hearing took part in this experiment. Informed consent was obtained from all participants according to the approval by the Ethical Committee of the Faculty of Psychology and Neuroscience, University of Maastricht.

Stimuli and task

We used the same auditory stimuli as in the study by Formisano et al. (2008) consisting of three vowels (/a/, /i/, /u/) spoken by three different speakers (sp1: female, sp2: male, sp3: male). For each of these 9 conditions, three different tokens were included in order to introduce acoustic variability. All stimuli were equated in length to 230 ms and in sound intensity by matching their root mean square amplitude. For more details about the stimulus properties, please see (Formisano et al., 2008). Prior to the functional experiment, participants were familiarized with the stimuli and were able to recognize the corresponding vowels and speakers. During the fMRI experiment, subjects were instructed to attentively listen to the stimuli while fixating a white cross in the center of the screen. The stimuli were presented in the silent gap between two subsequent image acquisitions (see below).

In order to ensure the engagement of the participants in both listening and fixating tasks, the participants performed a one back-task on the speaker dimension irrespectively of the spoken vowel: 10% of the total number of trials were catch-trials (signaled to the participants by the fixation cross turning red for 100 ms), in which the subjects were asked to report whether the speaker of the last heard sound was the same as the previous one. Subjects performed the task by pressing a button with either the index ("Yes"-answer) or the middle ("No"-answer) finger of the right hand. Catch-trials were excluded from all subsequent analyses.

The sounds were played according to a slow-event related design with a variable interstimulus interval (ISI) of 6–8 TRs (TR = 2500 ms, average ISI 17.5 s). At the beginning of the fMRI session, the volume of the stimuli was adjusted to a comfortable intensity level. The stimuli were presented in the 500 ms silent gap via MR-compatible earphones (Sensimetrics S14, Malden, MA, USA). After the experiment, all subjects reported a clear hearing of the stimuli. Every run consisted of 5 trials for each of the 9 stimulus conditions and 5 catch-trials, resulting in a total of 50 trials and a run duration of approximately 15 min. The order of stimulus presentation was randomized within and across runs. Four functional runs were acquired, leading to a total of 200 trials in the whole experiment.

Data acquisition

Functional and anatomical images were acquired with a 7 T Siemens Magnetom scanner using a 32-channel Nova Medical head coil. Four high-resolution (1.1 mm isotropic voxel size) functional runs were acquired using a gradient-echo (GE) EPI sequence (Moeller et al., 2010) with the following parameters: TR 2500 ms, TE 22 ms, Partial Fourier 5/8, GRAPPA 2, delay in TR 500 ms, multi-band acceleration factor 2 with blipped-CAIPIRINHA (1/FOV shift 4; Setsompop et al., 2012). The sequence was optimized to maximize tSNR in the auditory cortex. In two separate pilot runs of 140 volumes (~6 min, resting state), we acquired the sequence with these parameters and additionally a variant with GRAPPA 3, Partial Fourier 6/8 and TE 24.4 ms. The latter showed less distortions and signal dropout only in the anterior and posterior parts of the brain albeit with a lower tSNR in the auditory cortex (23.27 versus 34.02, respectively).

In addition to the magnitude images, phase images were collected in order to allow image reconstruction with lower voxel resolution (see below for details). 48 slices were acquired centered approximately on the superior temporal gyrus, covering the auditory cortex. One high-resolution (0.7 mm isotropic voxel size) anatomical image covering the whole brain was collected using MP2RAGE sequence (Marques et al., 2010).

Data analysis: preprocessing and univariate analysis

Functional and anatomical data were preprocessed and analyzed in BrainVoyager QX 2.8.2 (Brain Innovation). The four functional runs were 3D motion corrected and coregistered to the first volume of the first run through rigid-body transformation (3 translational and 3 rotational parameters). Neither nonlinear transformation nor distortion correction algorithm were applied to avoid interpolation confounds in our comparison across resolutions. We visually inspected every coregistered run and no large motion was observed. Linear and low-frequency non-linear drifts up to 7 cycles per time course were removed via temporal high-pass filtering. This cut-off frequency, corresponding to a cut-off period of ~128 s, was adequate to the stimulus design and analyses here employed (as estimated through spectral analysis of the class stimulus design). For each subject, the anatomical image was segmented at the gray-white matter boundary via an automatic procedure.

The obtained segmentation was visually inspected and manually corrected where necessary. Then, the anatomy was realigned in the subject-specific native space of the functional data in order to avoid any additional resampling and/or interpolation of the functional data. Finally, the segmented anatomical image was inflated to create an individual cortex mesh for each subject. The obtained inflated cortices were used for mask creation and displaying purposes.

Standard univariate statistical analysis of the fMRI data was performed for each subject using general linear modeling (GLM) of the time series. Voxels with a mean EPI value lower than 100 a.u. were excluded in order to prevent inclusion of cerebrospinal fluid (CSF) voxels in the ROI. The design matrix consisted of a predictor for each vowel and speaker combination convolved with a canonical (double gamma) hemodynamic response function (HRF) in order to account for the hemodynamic response delay. Contrast analysis was performed for each binary comparison of vowels and speakers.

Data analysis: multivariate analysis

Multivariate analysis was performed on a single-subject basis in anatomically defined masks including the superior temporal gyrus (STG), Heschl's gyrus (HG), and superior temporal sulcus (STS). The masks were delineated on the inflated cortex mesh in the native space of each subject.

A linear support vector machine (SVM) (see below) was trained to decode either the vowels (/a/ vs /i/, /a/ vs /u/, /i/ vs /u/) or the speakers (/sp1/ vs /sp2/, /sp1/ vs /sp3/, /sp2/ vs /sp3/). In case of vowel decoding, the trials were grouped according to the vowel irrespectively of the speaker dimension (e.g. the class /a/ contained all the trials corresponding to /sp1_a/, /sp2_a/, and /sp3_a/). In case of speaker decoding, trials were grouped according to the speaker irrespectively of the vowel dimension (e.g. the class /sp1/ contained all the trials corresponding to /sp1_a/, /sp1_i/, and /sp1_u/). Three classes of 60 trials were obtained in both cases. For each comparison, the classification was performed in leave-one-run-out cross-validation. At each fold of the cross-validation, 3 runs were used for feature selection, feature extraction, and training of the classifier and 1 run to test the performance. The details of each of these stages are explained in the following paragraphs.

Feature selection

For the purpose of algorithm efficiency and interpretability of the results, we further restricted the selection of voxels entering the classification by univariately choosing the n most active voxels per class. As MVPA classification accuracy may depend on number of selected voxels (Cox and Savoy, 2003), we let n vary between 50 and 10,000 voxels in discrete steps (n = [50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 3000, 4000, 5000, 6000, 8000, 10,000]) for a total of 19 levels of feature selection. The n most active voxels per class were defined as the n voxels with highest t-values resulting from the GLM of the training set. Note that the training set of a specific binary comparison and a specific cross-validation fold contained the trials corresponding to the two classes of that comparison and the 3 training runs of that fold. Considering, for example, the binary comparison "/sp1/ vs /sp2/" and the 1st cross-validation fold (in which run no. 1 is used for testing), all the trials corresponding to the stimuli /sp1_a/, /sp1_i/, and /sp1_u/ from the run nos. 2, 3, and 4 were included in the predictor of /sp1/ and all the trials corresponding to the stimuli /sp2_a/, /sp2_i/, and /sp2_u/ from the run nos. 2, 3, and 4 were included in the predictor of /sp2/. Once the GLM was performed, we selected the voxels having the n highest t-value for the predictor /sp1/ and those having the n highest t-value for the predictor /sp2/. The union of these two ensemble of voxels, which was ranging between n and 2^*n depending on the amount of overlap between the two ensembles, represented the voxels selected for that specific level of feature selection n and was passed to the *feature extraction* stage.

Feature extraction

The preprocessed time courses of the selected voxels were divided into single trials according to the time of presentation of each stimulus. Each trial was normalized to the averaged value of the signal at 2500 ms (1TR) and 0 ms before the stimulus onset to avoid biases due to global changes in the baseline signal across runs or signal drifts within and across runs. The obtained percent BOLD signal change was fitted with an optimized model of hemodynamic response function resulting in a β-value for each trial and each voxels which was then used as a feature in the classifier.

The HRF model used here was optimized to take into account potential difference in the temporal delay of the BOLD response. The optimal delay for each voxels was calculated by averaging the hemodynamic response across all trials and fitting it with a canonical HRF models whose time-to-peak parameter was varied between 4.0 s and 6.0 s in step of 0.5 s: the time-to-peak giving the best fit (as estimated by F-statistics of the regression) was chosen as optimal HRF delay of that voxel.

Classification

We used the linear SVM implemented in the Spider toolbox (http://people.kyb.tuebingen.mpg.de/spider/) as supervised classification algorithm with a fixed regularization parameter ($C = 1$; see Cortes and Vapnik, 1995 for details of the algorithm). Training and testing of the algorithm were performed in leave-one-run-out cross-validation: 3 runs (i.e. 45 trials per class) were used to train the classifier and 1 run (i.e. 15 trials per class) to test the performance. Voxel selection, feature extraction, training, and testing were repeated 4 times for all possible training and testing run combinations. The decoding performance (generalization score) was assessed by averaging the classification accuracies across the 4 folds of the cross-validation.

Finally, the entire classification procedure (including voxel selection, feature extraction, training, and testing of the classification in cross-validation) was repeated for each of the nineteen levels of feature selection.

Statistical testing

The statistical significance of the decoding accuracies was assessed at group level using non-parametric statistics. For each subject, we estimated the mean accuracy under the null hypothesis by performing a permutation test with 200 permutations. Note that this limited number of permutations does not allow reliable testing for significance at the single subject level; however, it provides a stable estimate of the mean of the null distribution, which is then used in the second-level analysis. For each permutation, the labels of the training dataset were scrambled using the *randperm* function of MATLAB (MathWorks, Natick, MA, USA) and the classification procedure repeated at the different feature selection levels. Using the same labels at different feature selection levels allows correctly taking into account the correlations of decoding performance among different levels of feature selection.

In the second-level analyses, the accuracy curve (i.e. the set of accuracy values at all nineteen levels of feature selection) was considered in a permutation test. In each permutation, we randomly swapped the accuracy curve of one (or more) subject(s) between the tested model and its null distribution mean and we counted the number of occurrences where the distance between the newly obtained group average curves was larger than the original one. The test was performed with an exact number of permutations $nr\_perm = 2^N = 1024$ (with N standing for number of subjects), which allowed to assess a minimum p-value of 2^{-10}. For details of the permutation test at group level, please, see the supplementary material.

Effect of spatial smoothing

In order to assess the effect of spatial smoothing of fMRI data on the decoding performance, we spatially smoothed the data with a three-

dimensional Gaussian kernel of full-width-half-maximum (FWHM) of 1.1, 2.2, 3.3, 4.4, and 8.8 mm. We performed the multivariate analysis on the unsmoothed data and after smoothing with each kernel size.

Reconstruction of data at lower effective spatial resolution

The raw data of the EPI images are acquired in the k-space where each k-value represents a sinusoidal spatial frequency. The inverse discrete Fourier transform (DFT) of the k-space data yields a complex image with both real and imagery components. These complex components are generally converted into a magnitude and phase image. Typical fMRI studies analyze only the magnitude image, while the phase image is usually discarded (but see Balla et al., 2014 and references therein). The original k-space data can be recovered by applying DFT on the combined information of the magnitude and phase data. The spatial resolution in the image space corresponds to the field-of-view (FOV) in the k-space and vice versa.

In order to determine the influence of spatial resolution on the decoding performance, we reconstructed the data at different lower effective resolutions, as follows (see Fig. 1). First, for each slice, we used both the magnitude and phase images collected at the original high resolution (1.1 mm isotropic) to reconstruct the complex k-space via 2-dimensional fast Fourier transform (FFT). Second, we retained a central submatrix of the k-space and we substituted the rest of the k-values with zero. In this way, we effectively removed the high spatial frequency information while leaving the low spatial frequency information unaffected. Finally, we applied the inverse 2D FFT to obtain an image with the same nominal resolution albeit with high spatial frequency information removed, i.e. lower effective resolution.

The original matrix size was 182×182 and the reconstructed ones had non-zero k-space information in the central 91×91 and 61×61 submatrices leading to an effective in-plane resolution of 2.2×2.2 mm^2 and 3.3×3.3 mm^2, respectively. This approach mimics the acquisition of fMRI data at different in-plane resolutions since the downsampling of the data is performed in the complex k-space. However, zeroing the external part of the k-space matrix, instead of cropping it, allowed us to preserve the same matrix size and, therefore, the same number of voxels in all analyses.

Note that the downsampling approach here presented is different from smoothing techniques in the image space (such as Gaussian smoothing), although both are filtering techniques. Gaussian smoothing is an invertible operation, which attenuates high spatial frequency components, but it does not eliminate them entirely (Kamitani and Sawahata, 2010; Swisher et al., 2010). Our downsampling methodology is a non-invertible operation, which directly manipulates the information in the spatial frequency space by complete cut-off of the high-frequency components. As a control analysis, the effect of resolution was tested also using downsampling of the complex data through cropping (see supplementary material for details and results). Importantly, not only the smoothing kernel is different for both approaches but also the data on which these operations are applied (magnitude vs. magnitude and phase data). Fourier transforming the magnitude data would not yield the complex k-space sampled during fMRI acquisition and would, therefore, not allow mimicking acquisition at lower (effective) spatial resolutions via zeroing/cropping its external part.

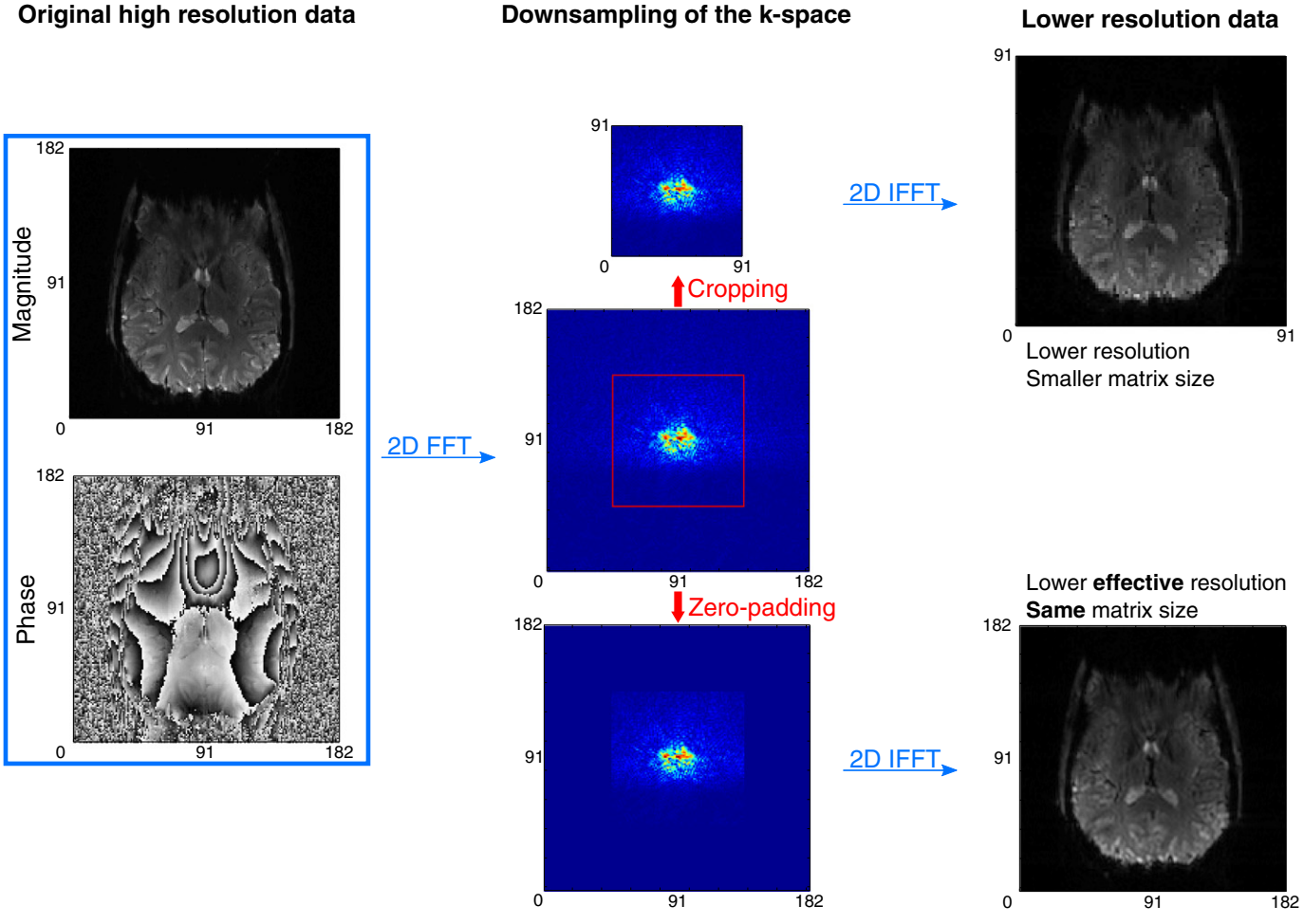

Fig. 1. Pipeline of the reconstruction of the fMRI data at different spatial resolutions (see text for details).

We preprocessed (motion correction, high-pass filtering, and spatial smoothing) and analyzed the datasets at lower effective resolutions following the same procedures as described above for the original data. The effects of resolution and smoothing on decoding accuracies were tested via a two-way repeated measures analysis of variance using R (package ez, http://CRAN.R-project.org/package=ez).

Influence of tSNR, CNR, and head motion on MVPA

Standard machine learning techniques rely on estimating features from a training set in order to predict labels in an unseen testing set. In fMRI MVPA, this translates in the assumption that the experimental stimuli drive the same pattern of activation through the entire duration of the experiment from which training and testing dataset are extracted. However, stimulus-related signal is modulated by instrumental noise and subject-related noise, such as physiological fluctuations and motion artifacts. In order to assess the influence of these factors on MVPA performances, we investigated the relationship between decoding accuracies and a) temporal signal-to-noise ratio (tSNR), a measure of both instrumental and physiological noise; b) functional contrast-to-noise ratio (CNR), a measure of the strength of the activation signal with respect to noise, and c) head motion.

tSNR of the fMRI signal in the auditory cortex was calculated as the ratio between the mean fMRI signal and its temporal standard deviation and averaged across all voxels in the mask. CNR was calculated as the ratio between the standard deviation of the fMRI signal response and the standard deviation of the baseline (Welvaert and Rosseel, 2013). CNR was calculated only on the active voxels as defined in the feature selection step of the decoding procedure. Head motion was quantified on the basis of the rotation and translation parameters obtained from the motion correction algorithm: each motion parameter was linearly detrended; then, its standard deviation was calculated; finally, the average of the standard deviations of all parameters was taken as motion index.

Voxel-wise tissue type ranking

To characterize the physiological origin of multivariate information in fMRI, voxels were ranked according to their likelihood of containing gray matter (i.e. micro-vasculature) or large blood vessels (i.e. macro-vasculature). Such likelihood was calculated following the method introduced by Shmuel et al. (2010). First, voxels were ranked separately according to two different quantities: their EPI value and their BOLD percent signal change. Both lower EPI values (with no apparent field distortions and susceptibility artifacts in the ROI) and higher BOLD percent changes indicate large veins (see Uludag et al., 2009). Then, a global rank was built by averaging the "BOLD percent change rank" and the inverse of the "EPI value rank." Voxels were, thus, sorted according to their most likely tissue type: from gray matter (low global rank) to large vessels (high global rank) regions. For the rest of the paper, we will call this global rank "blood vessels (BV) likelihood."

Results

Univariate analysis

Univariate analysis showed significant BOLD signal activation (Q[FDR] < 0.05 and cluster size threshold of 4 voxels; FDR stands for false discovery rate) in the auditory cortex in response to the stimulus sounds (Fig. 2), similar to Formisano et al. (2008). Outside the temporal lobe, significant activation was observed in the frontal (Broca's area) and parietal lobe. Pair-wise contrast analysis between speakers and/or vowels did not show significant effects for any comparison (data not shown).

Locally averaged BOLD responses to the sound stimuli exhibit large variability in the shape and latency of the hemodynamic response in different areas of the auditory cortex. In general, the fMRI responses from the Heschl's gyrus and proximities have narrower response width than the responses from the auditory belt areas (see Fig. 3).

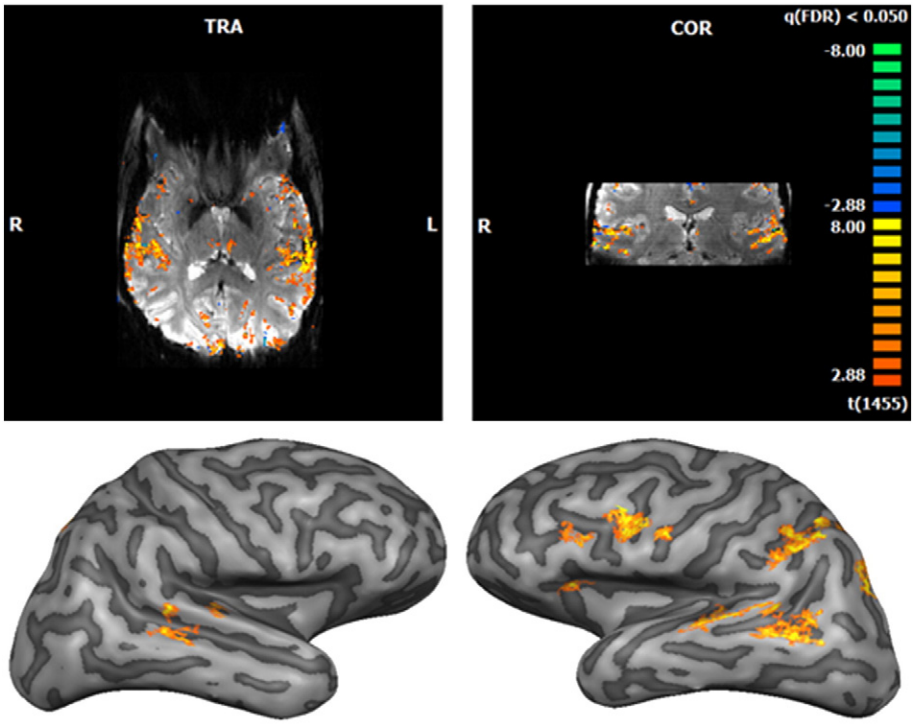

Fig. 2. BOLD activation in response to speech sounds as resulting from GLM analysis of unsmoothed data (single subject, statistical parametric F-map, Q[FDR] < 0.05, and cluster size threshold of 4 voxels). Top row, transversal (TRA), and coronal (COR) view; bottom row, projection of the F-map on the inflated cortex (light and dark gray represents gyri and sulci, respectively).

Vowel and speaker decoding from the original data

Multivariate analysis was performed on the original fMRI data (1.1 mm isotropic spatial resolution) after preprocessing with 2.2 mm Gaussian smoothing. In the section "Effects of spatial resolution and smoothing on decoding," we will motivate this choice demonstrating that smoothing using a kernel size of 2.2 mm yields optimal preprocessing of the data.

Fig. 4 shows the group level results of vowel and speaker decoding (left and right, respectively) performed on the data with resolution of 1.1 mm isotropic. For each binary comparison, we plotted the average accuracy as a function of the number of voxels selected. The accuracy curves of all models show a similar dependency on the number of voxels selected: accuracy first increases with the number of voxels and then reaches a plateau. In some cases, accuracy decreases for extremely large number of voxels.

Every model was decoded above chance level as verified with permutation tests at group level (chance level is represented in the plots with thinner lines). Vowels could be robustly decoded reaching an average accuracy of 64%, 62%, and 63% for /a/ vs /i/, /a/ vs /u/, and /i/ vs /u/, respectively ($p < 0.001$ for all comparisons, permutation test). Speakers were decoded with a lower but significant average accuracy of 57% for Sp1 vs Sp2 ($p = 0.0029$), 56% for Sp1 vs Sp3 ($p < 0.001$), and 53% for Sp2 vs Sp3 ($p = 0.0049$).

All vowels were decoded with comparable performances, while decoding between the female speaker (Sp1) versus one male speaker (Sp2 and Sp3) resulted in significantly higher accuracy than decoding between the two male speakers: 'Sp1 vs Sp2' > 'Sp2 vs Sp3' with p-value = 0.0137, 'Sp1 vs Sp3' > 'Sp2 vs Sp3' with p-value = 0.0059 (as assessed via permutation test).

Effects of spatial resolution and smoothing on decoding

For each combination of spatial resolution and Gaussian smoothing preprocessing, we repeated the entire multivariate procedure including all levels of feature selection. For display purposes, we averaged the group accuracies across the levels of feature selection and present them in Fig. 5 by a surface as a function of spatial resolution and smoothing kernel size.

The effects of resolution and smoothing on decoding accuracies were tested via a two-way repeated measures analysis of variance. In the vowel decoding case, both main effects and their interaction were significant (main effect of resolution: $F(2,18) = 10.14$, $p = 0.001$; main effect of smoothing: $F(5,45) = 13.65$, $p < 0.001$; interaction effect between resolution and smoothing: $F(10,90) = 2.67$, $p = 0.007$). In the speaker decoding case, only the main effect of smoothing was significant (main effect of resolution: $F(2,18) = 1.58$, $p = 0.233$; main effect of smoothing: $F(5,45) = 4.45$, $p = 0.002$; interaction effect between resolution and smoothing: $F(10,90) = 1.02$, $p = 0.433$). Control analyses (see supplementary material) were performed to assure that these results were not biased by considering the average accuracy across the levels of features selection, as they do not depend on the specific level of feature selection considered.

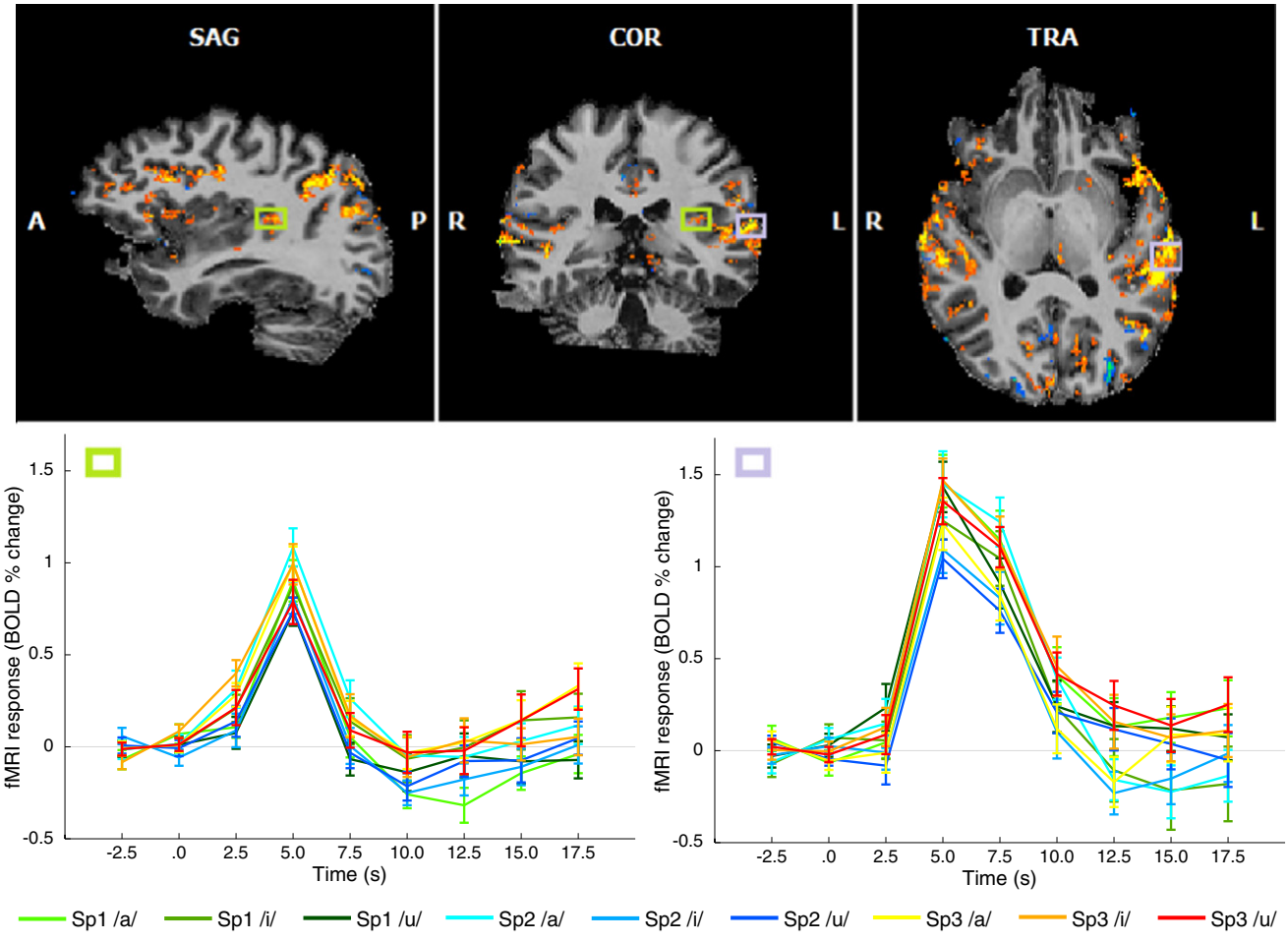

Fig. 3. Examples of mean BOLD response for different areas of the auditory cortex: right Heschl's gyrus (green ROI) and right superior temporal sulcus (violet ROI). Curves of different colors represent the mean BOLD response activated by the 9 different stimuli.

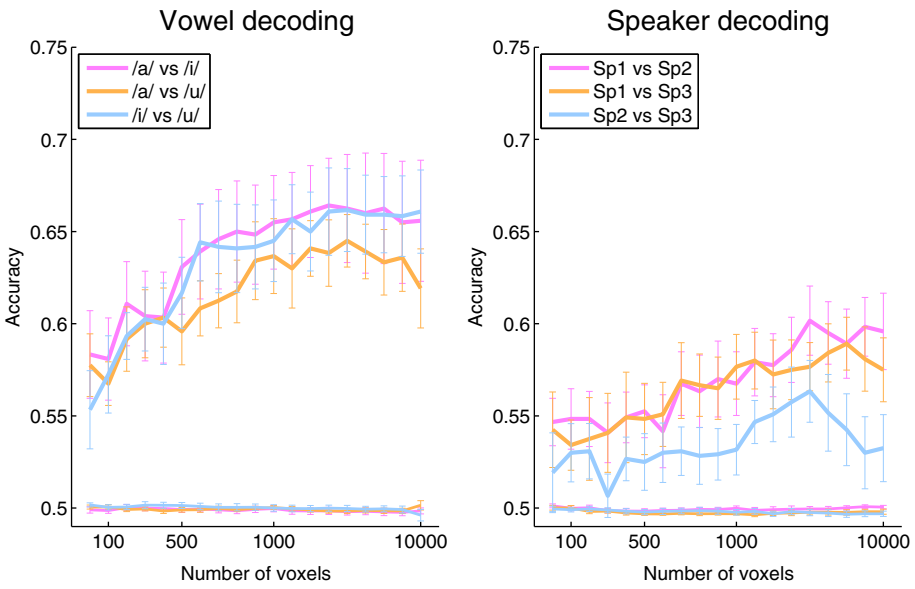

Fig. 4. Group level decoding accuracy for the data at the original resolution of 1.1 mm isotropic after smoothing with 2.2 mm Gaussian kernel. Accuracies are plotted as a function of the number of voxels selected. Each line represents the mean accuracy averaged across all the subjects and error bars represent the standard error of the mean. Thinner lines around 0.5 represent the corresponding results for the decoding with scrambled labels.

For both vowel and speaker decoding, the spatial smoothing improved the decoding accuracies in case of moderate smoothing (up to 2.2 mm), while more smoothing decreased them. This suggests that the optimal preprocessing of the fMRI data for decoding analysis, at least for the present data set, is obtained with a spatial smoothing of 2.2 mm FWHM 3D Gaussian kernel. For this reason, analyses in the rest of the paper were performed on data preprocessed with 2.2 mm FWHM Gaussian smoothing, unless otherwise specified.

Effect of spatial resolution on decoding

Fig. 6 illustrates in more detail the impact of acquiring fMRI data at different spatial resolutions on decoding performances. Each curve of the plots represents the group level decoding accuracy for each spatial resolution, calculated as the average across the three models of vowel or speaker decoding (left and right panel, respectively) and subjects. The stimuli were decoded significantly above chance level ($p < 0.01$, permutation test) at every resolution for both vowels and speakers. However, the decoding performances for vowels and speakers showed a different dependency on the effective spatial resolution of the analyzed data. Vowel decoding accuracies significantly increased with the increase in effective spatial resolution (OR1 > ER2 $p = 0.0078$, OR1 > ER3 $p = 0.0020$, ER2 > ER3 $p = 0.0498$; permutation test; OR1 stands for original resolution 1.1 mm, ER2 for effective resolution 2.2 mm, and ER3 for effective resolution 3.3 mm). In contrast, speaker decoding showed no significant accuracy differences between the tested resolutions, although a tendency of higher accuracies was observed for the effective resolution of 2.2 mm with respect of the 3.3 mm dataset (ER2 > ER3 $p = 0.0508$; permutation test).

The different resolution dependency for vowel and speaker decoding was confirmed by a permutation test ($nr\_perm = 2^{10}$) on the differences between accuracy curves at the investigated resolutions. Note that considering the difference between accuracy curves allows to account for a potential confound due to global performance differences in vowel and speaker decoding. The resolution effect on vowel decoding was significantly larger than that on speaker decoding for the difference between original and 2.2 mm in-plane resolution ($p = 0.0088$), close to significance for the difference between 1.1 and 3.3 mm in-plane resolution ($p = 0.0762$) and not significant for the difference between 2.2 and 3.3 mm in-plane resolution ($p > 0.1$).

Accuracy curves for the two lower effective resolutions showed a dependency on the number of voxels selected similar to the one of the original resolution (described in the section "*Vowel and speaker decoding from the original data*" and depicted in Fig. 6 by the blue line).

Influence of tSNR, CNR, and head motion on MVPA

We studied the relationship between averaged accuracies and temporal signal-to-noise ratio (*t*SNR), functional contrast-to-noise ratio (CNR) and head motion. Although no correlation was significant, we found that in the case of vowel decoding (and similarly for the speaker decoding case), the averaged accuracy tends to be positively correlated with tSNR ($r = 0.55$, $p = 0.09$) and CNR ($r = 0.37$, $p > 0.1$), while negatively correlated with head movement ($r = -0.34$, $p > 0.1$). Note that both increasing spatial smoothing and lowering spatial resolution of the fMRI data resulted in a monotonous increase of tSNR and CNR (data not shown).

Influence of tissue type on decoding

Fig. 7 shows the weights given by the classifier to each voxel as a function of BV likelihood (from 0% indicating most likely gray matter voxels to 100% indicating most likely large blood vessel voxels) for the case of vowel decoding. Weights are on average slightly, but significantly, higher for the 10% "more GM" voxels than for the 10% "more large BV" voxels (0.040 ± 0.006 versus 0.034 ± 0.007, respectively; *t*-test, *p*-value $= 0.0044$). Similar results are obtained in case of speaker decoding (0.036 ± 0.005 versus 0.031 ± 0.007, respectively; t-test, *p*-value $= 0.0035$).

Discussion

In this study, we used high-resolution 7 T fMRI data and their reconstruction at different effective spatial resolutions to investigate the relevant spatial scale of information content of the fMRI signal in the auditory cortex for vowel and speaker identity. To the best of our knowledge, this work is the first investigating the relation between spatial resolution and MVPA with fMRI at ultra-high magnetic field strength. Instead of acquiring additional data at lower spatial

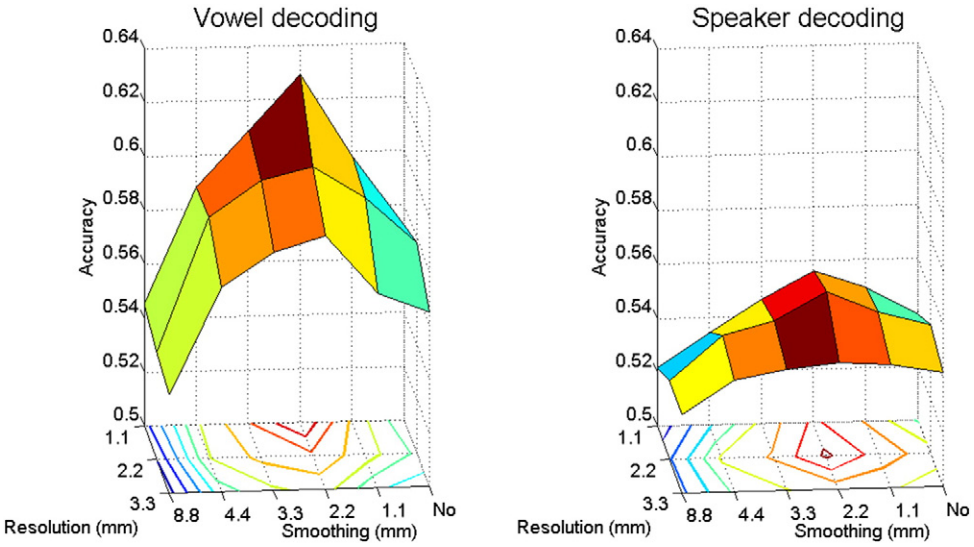

Fig. 5. Resolution and smoothing effect on decoding performances. Each point of the surface reports the group level accuracy of vowel (left plot) and speaker (right plot) decoding as a function of the spatial resolution of the data (from back to front: 1.1 × 1.1, 2.2 × 2.2, and 3.3 × 3.3 mm^2 in-plane) and the kernel size of spatial smoothing applied them (from right to left: unsmoothed and smoothed with Gaussian kernel size 1.1, 2.2, 3.3, 4.4, 8.8 mm). For display purposes, the accuracy is shown as the result of averaging across all levels of feature selection.

resolutions, we simulated their acquisition starting from the high-resolution data set. This allowed us to compare decoding performance for the very same dataset as a function of effective spatial resolution. In this way, we could exclude variability across runs and sessions related to subject performance (e.g. attention to the stimuli, movement) and acquisition differences (e.g. coil sensitivity, scanner drift).

We performed multivariate analysis of the original high-resolution data (1.1 mm isotropic) and demonstrated the feasibility of decoding speech content and speaker identity from 7 T fMRI data in the auditory cortex. These results are in agreement with the findings by Formisano et al. (2008), which demonstrated at 3 T the feasibility to decode vowels and speakers from auditory cortical response patterns. Note that a direct comparison of accuracies as obtained in this study with those in Formisano et al. (2008) is difficult due to the different methods of cross-validation and feature selection employed. Formisano and colleagues used a leave-k-stimulus-pair-out cross-validation (see Formisano et al., 2008 for details), while we employed a leave-one-run-out cross-validation. This latter approach is more conservative because it controls for false-positive rate at the nominal level; thus, it may result in lower accuracy values (Misaki et al., 2010). Moreover, Formisano and colleagues used a combination of univariate and multivariate feature selection strategies to improve sensitivity (see De Martino et al., 2008; Formisano et al., 2008 for details). Here, we chose a univariate-only feature selection method because of computational

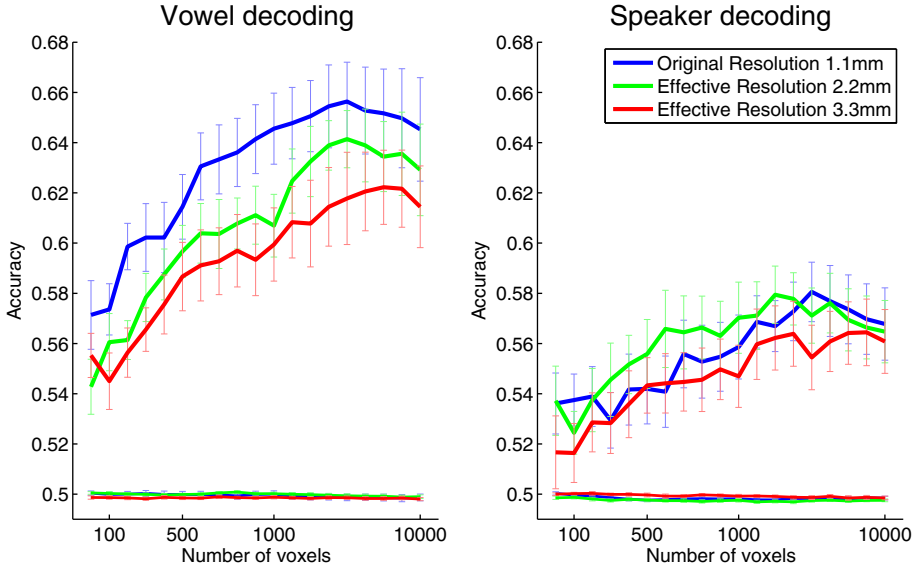

Fig. 6. Resolution effect on decoding accuracy. Each line represents the decoding accuracy averaged across all models of vowel or speaker decoding and all subjects for a certain resolution. Thinner lines around 0.5 represent the corresponding results for the decoding with scrambled labels. Error bars represent the standard error of the mean across the subjects. Both vowels and speakers are decoded above chance for all resolutions. In the left panel, vowel decoding performance significantly increases with increasing resolution: OR1 > ER2 $p = 0.0078$, OR1 > ER3 $p = 0.0020$, ER2 > ER3 $p = 0.0498$ (OR1 stands for original resolution 1.1 mm, ER2 for effective resolution 2.2 mm, and ER3 for effective resolution 3.3 mm). In the right panel, speaker decoding performance was almost significantly better for the effective resolution of 2.2 mm with respect to the effective resolution of 3.3 mm: OR1 > ER2 $p = 0.6777$, OR1 > ER3 $p = 0.1748$, ER2 > ER3 $p = 0.0508$. *p*-values were calculated via permutation test as explained in the section "Statistical testing" with the only difference that in this case we randomly swapped accuracy curves between the tested resolutions (instead of between "real" accuracy curve and "chance" accuracy curve).

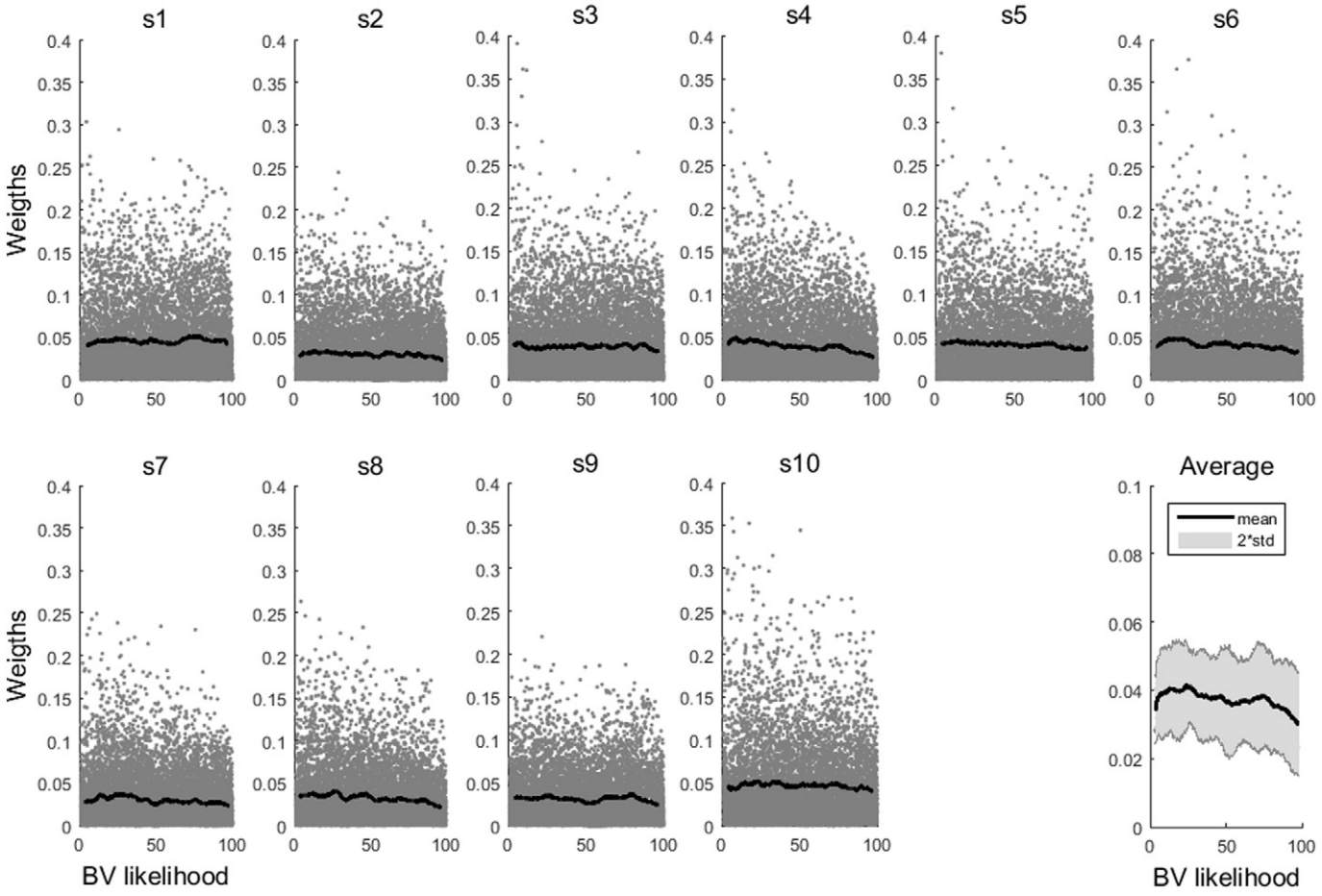

Fig. 7. Distribution of classification weights versus BV likelihood for each subject in the case of vowel decoding. Each gray dot represents one voxel and the black line the moving average of the weights calculated excluding the 5% more extreme values in order to avoid boundary effects. The bottom right plot shows the group moving average (black line) and its standard deviation (light gray shadow).

complexity, which is of particular concern for high-resolution data and large number of features (voxels).

To overcome an arbitrary choice of number of features selected, the multivariate analysis was repeated for several feature selection levels (from 50 to 10,000 voxels) and decoding accuracies of different models were compared by considering the resulting decoding accuracy curves. The accuracy curves consistently showed a dependency on the number of voxels selected: accuracy increased with the number of voxels, reached a plateau, and generally decreased for extremely high number of voxels. This observation is in agreement with previous studies (Cox and Savoy, 2003; Fan et al., 2007; De Martino et al., 2008) and highlights that our results do not critically depend on a specific feature selection level. In general, the optimal number of voxels to be included in the MVPA analysis may depend on the spatial resolution, the decoding task, and image characteristics. Additionally, we observed that speaker decoding was performed with higher accuracy for comparison between the female and a male speaker than between the two male speakers. These results reflect the smaller difference in fundamental frequency among the two male speakers (compared to the female speaker, see also Formisano et al., 2008, cfr. Fig. 2b).

The comparison of MVPA performances at different spatial resolutions showed that decoding accuracies increased with increasing resolution in case of vowel decoding. In case of speaker decoding, however, such behavior was not observed. Instead, there was a tendency to observe the highest accuracy for the intermediate spatial resolution. These results show that high spatial frequencies are informative for vowel decoding and that the relative contribution of high and low spatial frequencies is different across the two decoding tasks. In turn, this suggests a distinct spatial arrangement of the distributed neural/neurovascular sources underlying vowel and speaker representation and processing in auditory cortex.

These outcomes are consistent with previous observations that informative voxels in the auditory cortex are widely distributed for vowel decoding, while more clustered for speaker decoding (Formisano et al., 2008; Bonte et al., 2014). In addition, they support the hypothesis that the brain response patterns enabling the decoding of vowels and speakers reflect the representation and processing of distinct type of acoustic information (e.g. the representation of "formants combination" for vowel decoding and of "fundamental frequency" for speaker decoding, see Formisano et al., 2008). Furthermore, the finding that the maximum information (MVPA accuracy) is decoded optimally at different resolutions for vowels than for speakers provides a clear example of the complex dependency of MVPA results on the underlying physiological basis and thus of the need to address the issue of optimal acquisition strategy case-by-case through empirical studies (Formisano and Kriegeskorte, 2012).

In contrast to the resolution effect, we noted that the effect of smoothing is similar for vowel and speaker decoding and for different resolutions. This result suggests that the optimal smoothing kernel depends mainly on intrinsic noise properties of the data. This is in agreement with previous findings by Kamitani and Sawahata (2010). They showed that smoothing (e.g. with a Gaussian kernel) and more generally spatial convolution does not remove information content as it is an invertible transformation. However, such equivalence between original and smoothed data in terms of information content, and therefore classification results, holds only in ideal cases where the signal from

neighboring voxels in entirely independent, no selection of voxels is made and no motion (or other sources of) artifacts are present. Thus, in experimental fMRI data, the influence of smoothing on decoding might be attributed to the attenuation of (thermal) noise due to the signal correlation among neighboring voxels. Moreover, we observed that tSNR and CNR monotonously increased with larger smoothing kernels, while the decoding accuracies peaked at 2.2 mm 3D Gaussian kernel size. Therefore, the smoothing effect cannot be explained only on the basis of these two factors. Interestingly, such optimal spatial smoothing was reached for a kernel size of 2.2 mm which is similar to the experimentally determined point spread function at 7 T (Shmuel et al., 2007). Another reason why moderate smoothing is beneficial for MVPA could be the residual head motion present in the data even after motion correction. This hypothesis is consistent with previous studies showing that head motion reduces MVPA performances (Kriegeskorte et al., 2010; Swisher et al., 2010; Alink et al., 2013; Misaki et al., 2013). On the other hand, smoothing with large kernel sizes results in poorer MVPA performances due to partial volume effects with white matter and CSF or gray matter on an opposing sulcal bank (Andrade et al., 2001). For this reason, smoothing along the 2D cortical sheet may be more beneficial than performing spatial smoothing in the volume. Our results showing that large smoothing kernels are detrimental for MVPA decoding are in line with previous studies (Kriegeskorte et al., 2010; Swisher et al., 2010; Alink et al., 2013; Misaki et al., 2013). However, Op de Beeck (2010) reported no influence of spatial smoothing on decoding performances. A possible explanation of these different outcomes may be the different preprocessing of the functional data. In Op de Beeck (2010), data were coregistered in the anatomical space of the subject and then normalized to MNI template (Montreal Neurological Institute). These two transformations are achieved via re-sampling and interpolation of the original functional data. Therefore, the smoothness introduced in the transformed data could make the manipulation through different spatial smoothing kernels less sensitive, especially in experiments where decoding accuracies are at ceiling level (100% decoding accuracy for the object decoding experiment).

One of the advantages of using a linear SVM algorithm is the possibility to interpret the weights attributed to each voxel in terms of discriminability power. In order to shed further light on the physiological origin of the multivariate information in fMRI, we ranked the voxels according to their tissue properties. We showed that discriminability power is distributed over the continuum of tissue type ranging from gray matter to large blood vessels regions, with a slightly higher contribution from gray matter than from blood vessel voxels. This result suggests that decodable information resides both in the tissue and large veins. Future studies, possibly including venograms and/or using different type of functional contrast such as spin echo, arterial spin labeling, or vascular space occupancy, are needed to understand the role of tissue types in fMRI MVPA.

Another factor that could influence MVPA performance in fMRI is HRF variability. In this study, we observed important changes in the HRF shape in different regions. However, a systematic investigation of this phenomenon and its influence on MVPA is beyond the scope of this paper and needs to be addressed in future developments.

In summary, in this study, we investigated MVPA in fMRI at 7 T. Our results suggest that different features of complex (auditory) stimuli are spatially distributed in the brain and may be represented at different spatial scales. As a general consequence, there is no optimal spatial resolution for all decoding tasks but the acquisition needs to be tailored to the stimuli and brain area investigated. However, our findings for vowel decoding are compatible with the hypothesis that MVPA information partly originates from spatial arrangements of neuronal populations at a spatial scale below fMRI resolution (<1 mm). This spatial information may become available at resolutions accessible for fMRI due to local irregularities and/or transposition by hemodynamic filtering (Chaimow et al., 2011). Under this hypothesis and in the absence of precise knowledge of the spatial scale of interest, we recommend acquisition of data at high spatial resolution (1 mm or below). This enables the detection of stimulus-related information at all measurable spatial frequencies, which is required for inferring the spatial organization of the underlying neuronal architecture.

Acknowledgments

This work was financially supported by the Marie Curie Initial Training Network grant (PITN-GA-2009-238593) of EU and VIDI grant (KU, 452-11-002) of the Netherlands Organization for Scientific Research (NWO). EF was supported by an NWO VICI grant (453-12-002).

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2016.02.033.

References

Alink, A., Krugliak, A., Walther, A., Kriegeskorte, N., 2013. fMRI orientation decoding in V1 does not require global maps or globally coherent orientation stimuli. Front. Psychol. 4, 493.

Andrade, A., Kherif, F., Mangin, J.F., Worsley, K.J., Paradis, A.L., Simon, O., Dehaene, S., Le Bihan, D., Poline, J.B., 2001. Detection of fMRI activation using cortical surface mapping. Hum. Brain Mapp. 12, 79–93.

Balla, D.Z., Sanchez-Panchuelo, R.M., Wharton, S.J., Hagberg, G.E., Scheffler, K., Francis, S.T., Bowtell, R., 2014. Functional quantitative susceptibility mapping (fQSM). NeuroImage 100, 112–124.

Bonte, M., Hausfeld, L., Scharke, W., Valente, G., Formisano, E., 2014. Task-dependent decoding of speaker and vowel identity from auditory cortical response patterns. J. Neurosci. Off. J. Soc. Neurosci. 34, 4548–4557.

Boynton, G.M., 2005. Imaging orientation selectivity: decoding conscious perception in V1. Nat. Neurosci. 8, 541–542.

Chaimow, D., Yacoub, E., Ugurbil, K., Shmuel, A., 2011. Modeling and analysis of mechanisms underlying fMRI-based decoding of information conveyed in cortical columns. NeuroImage 56, 627–642.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20, 273–297.

Cox, D.D., Savoy, R.L., 2003. Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. NeuroImage 19, 261–270.

De Martino, F., Valente, G., Staeren, N., Ashburner, J., Goebel, R., Formisano, E., 2008. Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns. NeuroImage 43, 44–58.

Fan, Y., Rao, H., Hurt, H., Giannetta, J., Korczykowski, M., Shera, D., Avants, B.B., Gee, J.C., Wang, J., Shen, D., 2007. Multivariate examination of brain abnormality using both structural and functional MRI. NeuroImage 36, 1189–1199.

Formisano, E., Kriegeskorte, N., 2012. Seeing patterns through the hemodynamic veil—the future of pattern-information fMRI. NeuroImage 62, 1249–1256.

Formisano, E., De Martino, F., Bonte, M., Goebel, R., 2008. "Who" is saying "what"? Brain-based decoding of human voice and speech. Science 322, 970–973.

Freeman, J., Brouwer, G.J., Heeger, D.J., Merriam, E.P., 2011. Orientation decoding depends on maps, not columns. J. Neurosci. Off. J. Soc. Neurosci. 31, 4792–4804.

Gardner, J.L., 2010. Is cortical vasculature functionally organized? NeuroImage 49, 1953–1956.

Gardner, J.L., Sun, P., Tanaka, D., Heeger, D.J., Cheng, K., 2006. Classification Analysis with High Spatial Resolution fMRI Reveals Large Draining Veins with Orientation Specific Responses. Society for Neuroscience Meeting, Atlanta, GA, USA.

Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., Pietrini, P., 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 293, 2425–2430.

Haynes, J.D., Rees, G., 2005. Predicting the orientation of invisible stimuli from activity in human primary visual cortex. Nat. Neurosci. 8, 686–691.

Haynes, J.D., Rees, G., 2006. Decoding mental states from brain activity in humans. Nat. Rev. Neurosci. 7, 523–534.

Kamitani, Y., Sawahata, Y., 2010. Spatial smoothing hurts localization but not information: pitfalls for brain mappers. NeuroImage 49, 1949–1952.

Kamitani, Y., Tong, F., 2005. Decoding the visual and subjective contents of the human brain. Nat. Neurosci. 8, 679–685.

Kamitani, Y., Tong, F., 2006. Decoding seen and attended motion directions from activity in the human visual cortex. Curr. Biol. 16, 1096–1102.

Kriegeskorte, N., Bandettini, P., 2007. Analyzing for information, not activation, to exploit high-resolution fMRI. NeuroImage 38, 649–662.

Kriegeskorte, N., Cusack, R., Bandettini, P., 2010. How does an fMRI voxel sample the neuronal activity pattern: compact-kernel or complex spatiotemporal filter? NeuroImage 49, 1965–1976.

Marques, J.P., Kober, T., Krueger, G., van der Zwaag, W., Van de Moortele, P.F., Gruetter, R., 2010. MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. NeuroImage 49, 1271–1281.

Misaki, M., Kim, Y., Bandettini, P.A., Kriegeskorte, N., 2010. Comparison of multivariate classifiers and response normalizations for pattern-information fMRI. NeuroImage 53, 103–118.

Misaki, M., Luh, W.M., Bandettini, P.A., 2013. The effect of spatial smoothing on fMRI decoding of columnar-level organization with linear support vector machine. J. Neurosci. Methods 212, 355–361.

Moeller, S., Yacoub, E., Olman, C.A., Auerbach, E., Strupp, J., Harel, N., Ugurbil, K., 2010. Multiband multislice GE-EPI at 7 Tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine/Society of Magnetic Resonance in Medicine 63, pp. 1144–1153.

Op de Beeck, H.P., 2010. Against hyperacuity in brain reading: spatial smoothing does not hurt multivariate fMRI analyses? NeuroImage 49, 1943–1948.

Polimeni, J.R., Fischl, B., Greve, D.N., Wald, L.L., 2010. Laminar analysis of 7 T BOLD using an imposed spatial activation pattern in human V1. NeuroImage 52, 1334–1346.

Setsompop, K., Gagoski, B.A., Polimeni, J.R., Witzel, T., Wedeen, V.J., Wald, L.L., 2012. Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty. Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine/Society of Magnetic Resonance in Medicine 67, pp. 1210–1224.

Shmuel, A., Yacoub, E., Chaimow, D., Logothetis, N.K., Ugurbil, K., 2007. Spatio-temporal point-spread function of fMRI signal in human gray matter at 7 tesla. NeuroImage 35, 539–552.

Shmuel, A., Chaimow, D., Raddatz, G., Ugurbil, K., Yacoub, E., 2010. Mechanisms underlying decoding at 7 T: ocular dominance columns, broad structures, and macroscopic blood vessels in V1 convey information on the stimulated eye. NeuroImage 49, 1957–1964.

Swisher, J.D., Gatenby, J.C., Gore, J.C., Wolfe, B.A., Moon, C.H., Kim, S.G., Tong, F., 2010. Multiscale pattern analysis of orientation-selective activity in the primary visual cortex. J. Neurosci. Off. J. Soc. Neurosci. 30, 325–330.

Uludag, K., Muller-Bierl, B., Ugurbil, K., 2009. An integrative model for neuronal activity-induced signal changes for gradient and spin echo functional imaging. NeuroImage 48, 150–165.

Welvaert, M., Rosseel, Y., 2013. On the definition of signal-to-noise ratio and contrast-to-noise ratio for FMRI data. PLoS One 8, e77089.

Yacoub, E., Van De Moortele, P.F., Shmuel, A., Ugurbil, K., 2005. Signal and noise characteristics of Hahn SE and GE BOLD fMRI at 7 T in humans. NeuroImage 24, 738–750.