Advanced fMRI analysis - Assignments

Odysseas Papakyriakou, MSc Artificial Intelligence
o.papakyriakou@students.uu.nl (2716283)

Utrecht University

July 25, 2023

1 Assignment 1

The purpose of this assignment is to explore the connectivity between different brain areas and between task conditions using fMRI. Thus, the interest lies in exploring whether the extent of the connectivity between areas might depend on the different conditions, which is also referred to as *psychophysical interaction*. This is essentially a bilinear model of how a condition A (the psychological context) changes the influence of area B on area C, corresponding to differences in regression slopes for different conditions. Necessarily, measuring psychophysical interaction requires regression, not correlation, to establish the connectivity between the selected brain areas.

1.1 Research question

To avoid pre-processing new fMRI data, the research question for this assignment is based on the data of the course *Basic fMRI analysis*, where the goal was to explore the Penfield map for feet, tongue, and hand movement. Therefore, the research question of this assignment is whether there are differences in functional connectivity between the superior sensorimotor network, the left, and the right sensorimotor networks for the feet and tongue conditions. Particularly, the interest lies in functional connectivity between areas and not solely direct connectivity, since functional connectivity captures the collective influence of both direct and indirect connections between brain regions, thus allowing to investigate how information is integrated and transferred across different regions, even if they are not directly connected by strong anatomical pathways.

1.2 Methods

All anlyses were performed using the CONN toolbox in a matlab environment, following this pipeline: removing confounding influences, establishing individual connectivity (first-level analysis), and establishing group-wise connectivity (second-level analysis). In total, there were 18 subjects examined: 1, 3, 7, 8, 11, 14, 15, 20, 25, 26, 33, 40, 48, 50, 52, 54, 59, 60, the number of sessions (conditions) was 2, the repetition time was 2.5, the acquisition type was continuous, while the onset and duration of each condition was specified according to the assignment description.

1.2.1 Segmentation and specification of nuisance factors

Initially, the normalized anatomical images were specified for each subject with a resolution of $1 * 1 * 1$, which were the *w*.nii* files of the normalization process described in the report for the *Basic fMRI analysis* course. Then, the functional images were specified for each subject and for each session, which were the *swar*.nii* files that were produced after the last pre-processing step.

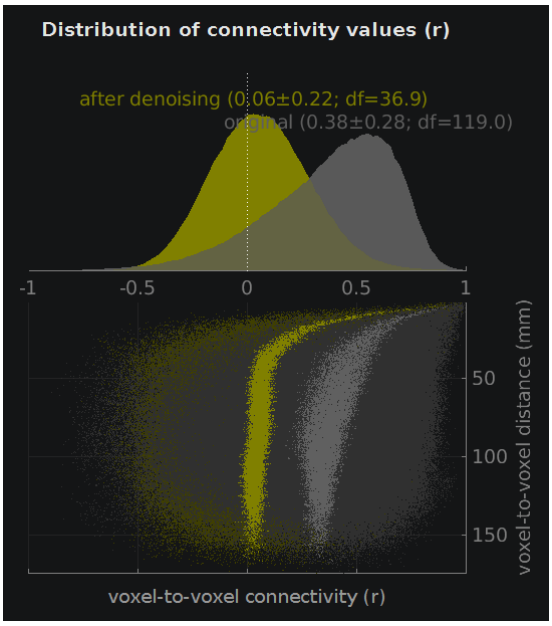
Since the carbon dioxide may affect the height of the signal and the correlation between voxels in the grey matter, grey matter can be used for global correction. This process removes the mean of the grey matter signal from all voxels by taking the T1 weighted images and segmenting them into grey matter, white matter, and cerebrospinal fluid. However, since the functional data is smoothed, the signals from the grey matter also include signals from the white matter, so the toolbox also includes the non-smoothed data, which is the *war*.nii* files. To remove the noise that may come from white matter and the cerebrospinal fluid, CONN performs principal component analysis on these covariates to establish what the most important sources of variance are and use these as a filter.

Additionally, the motion parameters in the *rp*.txt* produced by the pre-processing pipeline were included as first-level covariates for every subject and for every condition. Since there was not information about the age or the sex of the participants no other variables were included as second-level covariates.

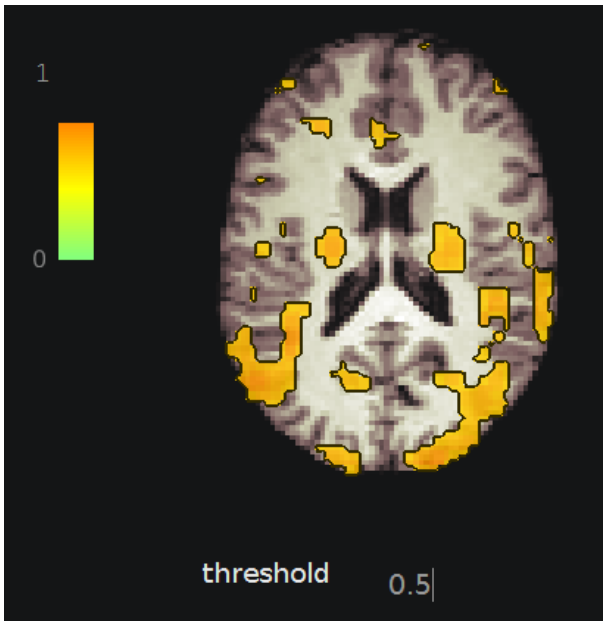
1.2.2 Denoising

Denoising allows to filter out the effect of confounding factors in connectivity that may come from head movements, white or grey matter, or the cerebrospinal fluid. Including the effect of rest as a confounder is redundant since the effect of each condition is also included. For the principal component analysis on white matter and the cerebrospinal fluid, the first five principal components were used to explain most of their variance. Moreover, the first-order derivative of the six motion parameters was also included to better distinguish task effects. So in total, there are 28 parameters that can be used as filters for noise, namely: 5 for white matter, 5 for cerebrospinal fluid, 6+6 (including the derivative) for motion, 2 for rest, 2 for the feet condition, and 2 for the tongue condition. The feet and tongue movement parameters are included as confounds because the interest lies in exploring the connectivity regardless of the task, so the variance that is explained by the effect of each condition needs to also be removed. Additionally, despiking before regression was also performed for each confounder. This allows to remove the confounding effect of spikes, which are sudden and large impulsive signal changes that are caused by sudden discharges of static electricity, thus causing electromagnetic distortions.

Figure 1 shows the result of denoising for an example subject from the tongue movement condition. Particularly, figure 1a shows the distribution of the correlations between all voxels before and after denoising, with grey and yellow color respectively, where the correlation between two voxels is used as a measure of connectivity between these voxels. The unfiltered correlations (grey) are higher, centered around 0.4, than the correlations after denoising (yellow), which are centered around 0. This is as expected, because denoising essentially removes the variance that is explained by all the specified factors. As an example, figure 1b shows the percentage of BOLD variance that is explained by white matter at a correlation threshold of 0.5. The lower the specified threshold is and the more parameters are included, the higher the percentage of explained variance should be.



(a) Connectivity distribution



(b) Variance explained by white matter

Figure 1: Denoising results example

1.2.3 First-level analysis

The analysis that was performed was seed based. This means that after specifying the region of interest (ROI), or the seed, the mean timeseries for the voxels within the ROI were calculated and regressed on the mean timeseries of all voxels. An alternative would be to perform ROI to ROI analysis which would result in higher statistical power, overall. However, the seed-voxel analysis was preferred because it provides more elaborate connectivity patterns.

Moreover, since the interest lies in exploring functional connectivity, the comparison of different areas should not include any additional parameters of other areas, so the method that was used was simple regression between the two voxels. Additionally, since the interest lies in exploring differences in connectivity patterns between conditions, and not just for a single condition, this necessarily points to the use of regression and not correlation.

Since the conditions involve motor movements, the ROIs are in the sensorimotor network, and were specified using CONN's pre-defined networks. Specifically, three ROIs were selected as seeds, namely, the superior sensorimotor network, the left lateral sensorimotor network, and the right lateral sensorimotor network.

1.2.4 Second-level analysis

While the first level analysis establishes the connectivity patterns between the selected seeds for each subject and for each condition individually, the second-level analysis establishes whether there is an average difference in connectivity between conditions when considering all subjects. More specifically, it aims to answer the research question: does the functional connectivity between the sensorimotor networks change between rest, feet, and tongue conditions?

To test for these differences, three custom contrasts for Seed-Based Connectivity (SBC) were created in the CONN environment at the 2nd-level analysis step. These were [-2, 1, 1], testing for the difference in connectivity between the rest condition on the one hand, and the feet and tongue conditions on the other; [1, -2, 1], testing for the difference in connectivity between the feet condition on the one hand, and the rest and tongue conditions on the other; and [1, 1, -2], testing for the difference in connectivity between the tongue condition on the one hand, and the rest and feet conditions on the other.

1.3 Results

This section displays the results of the first and second level connectivity analyses respectively.

1.3.1 First-level analysis

As explained under section 1.2.3, the first-level statistical analysis shows the functional connectivity patterns between the selected seed and all other voxels, for every subject individually, and for each condition. As an example, figure 2 shows the connectivity patterns between the selected seed and all other voxels for each condition, for subject 40. The more red a voxel appears, the closer the correlation coefficient is to 1, while the more blue it appears, the closer the correlation coefficient is to -1, and the closer the coefficient is to either 1 or -1, the stronger the connectivity between the seed and that voxel is.

Visually, when looking at the superior sensorimotor network as seed, the tongue condition seems to have a a wider and more intense connectivity pattern than the other conditions, while when looking at the left or the right lateral sensorimotor network as seed, the connectivity pattern in the feet condition seems quite different than that in the other conditions. Additionally, the left and right lateral sensorimotor network show a similar connectivity pattern in each condition.

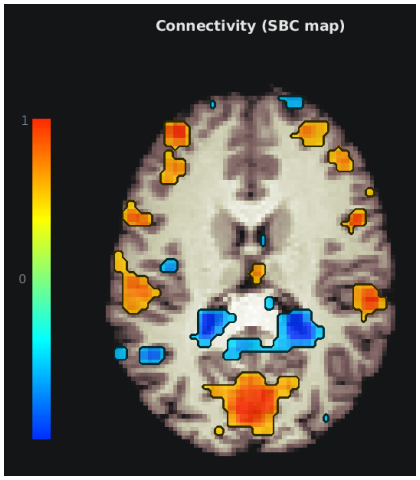
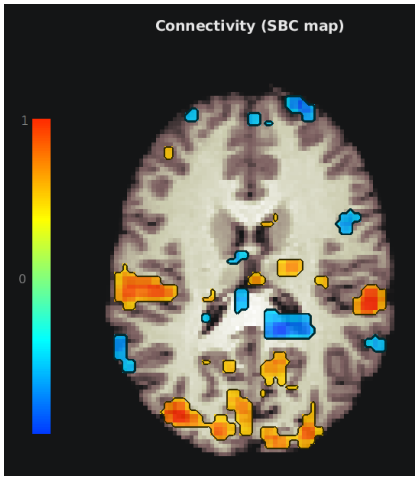
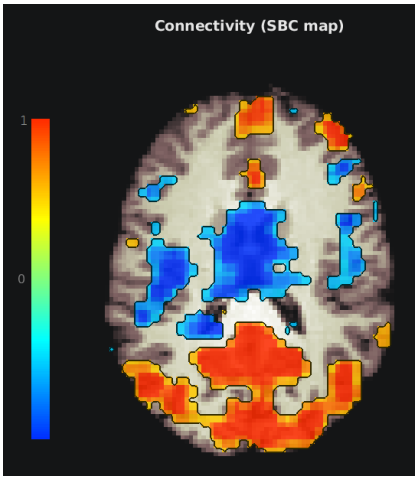
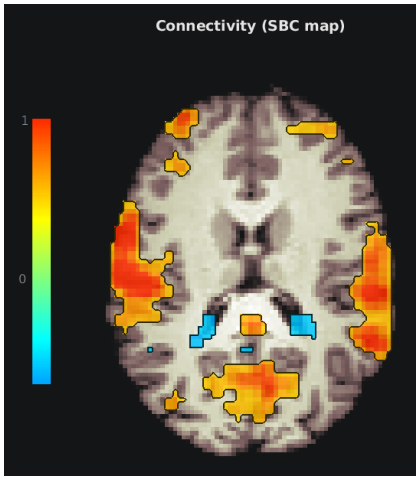
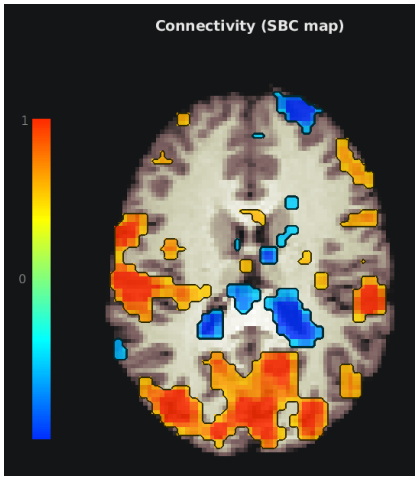
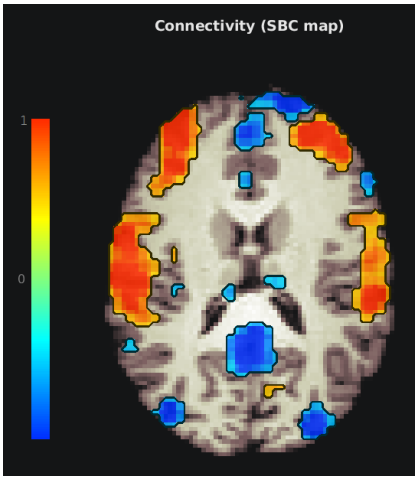
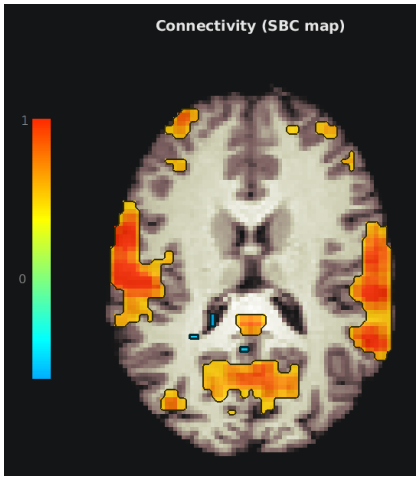
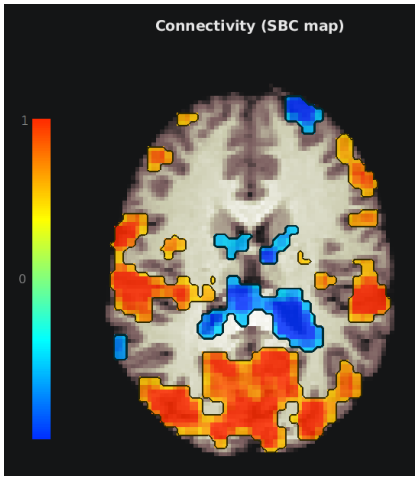
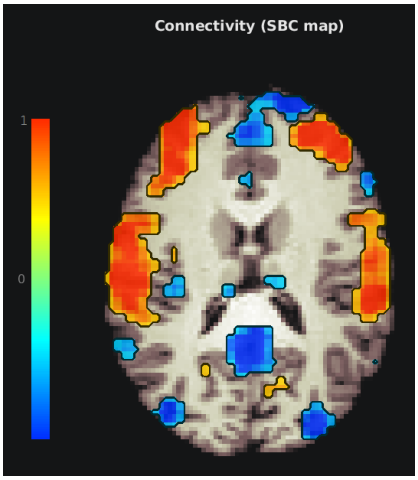
Seed	Rest condition	Feet condition	Tongue condition
Superior sensorimotor network			
Left lateral sensorimotor network			
Right lateral sensorimotor network			

Figure 2: Example results from subject 40, from a slice towards the middle of the axial plane, with a correlation threshold of 0.5

1.3.2 Second-level analysis

The second-level analysis allows to answer whether the overall functional connectivity patterns between the seed areas are different between the rest, feet, and tongue conditions. When interpreting the results, the default settings for cluster-based inferences were kept, namely, the random field theory for parametric statistics. Specifically, the uncorrected voxel threshold was set at $p < 0.001$, while the cluster threshold was set at $p < 0.05$ for the p-FDR correcter cluster size. Statistically significant and positively related connectivity patterns are shown with red, while statistically significant and negatively related connectivity patterns are shown with blue.

Figure 3 shows the differences in functional connectivity for each condition compared to the other two, for every network as a seed. When looking at the results, it becomes apparent that tongue connectivity patterns are not significantly different than these of the other two conditions for any of the three seeds. Moreover, none of the conditions seems to have a significantly different connectivity pattern than the other two when looking at the right lateral sensorimotor network as a seed. Additionally, the differences in connectivity patterns in the feet condition compared to the other two seem to be more pronounced between the left lateral sensorimotor network than between the superior sensorimotor network, while the connectivity during rest seems to be different than that during the other two conditions between both the superior and the left lateral sensorimotor networks.

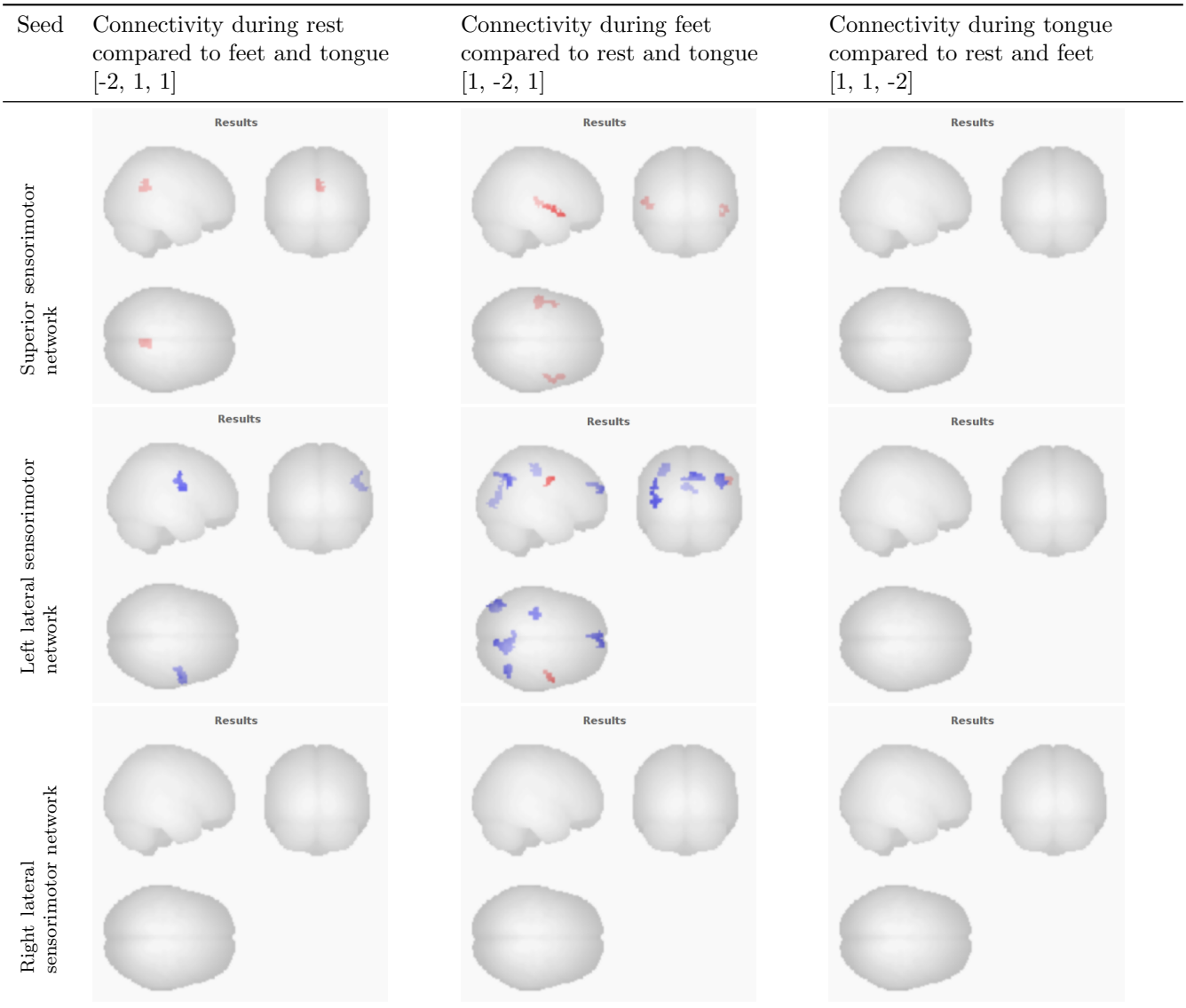


Figure 3: Significant voxel-based connectivity differences across conditions for different seeds

1.4 Conclusion

Conclusively, there is some difference in functional connectivity patterns for the different conditions, but these depend on the network that is taken as seed and on the condition. Notably, there is no significant difference in functional connectivity during the tongue task compared to the feet and the rest tasks, while the rest and feet tasks show differences in functional connectivity only when taking the superior or the left lateral sensorimotor network as a seed, and not the right lateral sensorimotor network.

2 Assignment 2

The purpose of this assignment is to perform Independent Component Analysis (ICA) with 20 output components on a single subject performing one task, and to reason about whether each component represents BOLD signal or an artefact. Specifically, the task that was selected was the feet task, and ICA was performed on the smoothed normalised data, with the *swar* prefix, as these were pre-processed for the course *Basic fMRI analysis*.

2.1 General information about Independent Component Analysis

When applied to fMRI data, ICA, along with Principal Component Analysis (PCA), can be used to draw a more comprehensive picture of the brain's connectivity structure, since they are not hypothesis-driven connectivity measures, like the seed to voxel analysis that was performed in section 1. Instead, they are data-driven methods aiming to uncover statistically independent components from the observed fMRI data. With ICA in particular, these components are also non-Gaussian and represent distinct sources of neural activity.

The output of ICA consists of the independent component timescourses, as well as the independent component maps. The timecourses represent the activity fluctuations in each component over time, while the spatial maps represent to what extent each component is represented in every voxel, resulting in a visualization of the brain region of each component. Additionally, the component timecourse can be transformed to create a frequency power spectrum, which is a representation of the power, or amplitude, of neural activity at different frequencies within the brain. Specifically, it is obtained with Fourier transform by decomposing the time-domain fMRI signal into its constituent frequency components, thus showing the amplitude of the fMRI signal over frequency.

2.2 Methods

ICA was performed using the GIFT toolbox in a matlab environment. Specifically, when setting up the ICA analysis, the smoothed normalised fMRI data that were selected were from subject 7 during the feet task, and the number of components was set to 20. All other options were left at default. When running the analysis all results were selected, and a results summary on a PDF file was created without regressors for temporal sort. The graphs and the information in this results summary are used to evaluate the 20 components and can be found under this link.

Since ICA is a statistically-driven method, the components can represent anything that produces enough change in the signal, from task-related activity, to respiration and heartbeat artefacts. When identifying each component as a BOLD signal or an artefact, the following guidelines were taken into consideration:

- BOLD signal is only found in grey matter
- BOLD signal usually has high power between 0.01 and 0.1 Hz in the frequency power spectrum
- activity at the edges of the brain may indicate motion artefacts and usually contain abrupt jumps in signal
- activity near large vessels with a high fixed frequency may indicate heartbeat artefacts
- activity throughout the brain with rhythmic and high frequency may indicate scanner artefacts

2.3 Results

This section attempts to identify whether each of the 20 components is a BOLD signal or an artefact, based on the output of the generated results summary.

Component 1 could be a BOLD signal because there is a peak power at lower frequencies, but the component also loads on the edges of the brain and on non grey matter areas.

Component 2 is an artefact and probably cerebrospinal fluid (CSF) since the component is found mostly in the ventricles.

Component 3 is an artefact and probably a cardiac cycle because it looks like it loads on different blood vessels.

Component 4 is probably an artefact because there is no clear peak in low frequencies and the component loads a lot on non grey matter areas.

Component 5 looks like an artefact, mostly because there is no clear peak in lower frequencies.

Component 6 looks like BOLD signal since it is mostly on grey matter and is quite symmetrical. Moreover, there is high power in the lower frequencies between 0.03 and 0.1.

Component 7 is an artefact and probably CSF since the component loads mostly in the area of the ventricles.

Component 8 is an artefact because there is no peak in power in lower frequencies and the component loads a lot on the edges of the brain and on non grey matter areas.

Component 9 is an artefact and probably shows a cardiac cycle because it looks like it loads a lot on blood vessels.

Component 10 is a BOLD signal since it loads a lot on grey matter areas and the power shows a peak in lower frequencies. Moreover, the component loads a lot on brain areas that have been shown to be related the feet task in the assignment of the course *Basic fMRI analysis*.

Component 11 is an artefact indicating eye movement, since the component loads on the location of the eyes, outside the brain.

Component 12 is an artefact, probably showing CSF since it loads mostly on the ventricles.

Component 13 looks like a BOLD signal because there is a high peak of power in lower frequencies, and because the component loads symmetrically on grey matter areas related to motor activity. However, the abrupt jump in the timecourse graph may also indicate a motion artefact.

Component 14 is a BOLD signal and it is probably related to the task since the component loads on task-relevant brain areas.

Component 15 is also a BOLD signal since it is on grey matter and pretty symmetrical, while also showing high power in low frequencies.

Component 16 is an artefact and it probably is a cardiac cycle or CSF, because it loads a lot on the ventricles and on blood vessels.

Component 17 is an artefact, probably indicating a cardiac cycle.

Component 18 show a lot of symmetry in grey matter areas, but also in blood vessels. Since there is also a small peak in lower frequencies, it might be a combination of BOLD signal and cardiac cycle.

Component 19 is probably an artifact and it looks like a cardiac cycle since it loads a lot on blood vessels.

Component 20 is probably a BOLD signal because of the peak in power in low frequencies. It also looks like the visual network.

3 Assignment 3

The purpose of this assignment is to map the statistical fMRI results of a single subject to the surface representation of the brain. Specifically, the fMRI data that were selected were from the tongue task, since this showed a clear bilateral activation in the course *Basic fMRI analysis*, that can potentially be evident in both hemispheres of the surface representation. To perform this mapping of the functional data to the surface, a specific pre-processing pipeline was performed on the raw fMRI data that is described in the methodology section, 3.2.1, along with the remaining methodological details.

3.1 General information about surface-based analysis

Surface-based analysis entails the change of the representation of the brain from a volume to a surface. Specifically, a surface, or also called a polygon mesh, is a collection of vertices, edges, and faces, or also called triangles, that collectively define the shape of a polyhedral object. Vertices represent the coordinates in three-dimensional space, edges represent the connections between the vertices, and triangles describe the interconnectedness of the vertices.

There are several different brain surfaces that can be computed from a given brain volume, provided there is a complete grey and white matter segmentation. The pial surface represents the outer grey matter surface, while the white matter surface represents the border between grey and white matter surface, or also called the inner grey matter surface. The generation of the pial and the white matter surfaces is done by algorithms that are designed to describe the edges of the grey and white matter segmentations in vertices and triangles, resulting in an equal

number of vertices and the same polygons for the pial and the white matter surfaces. The only difference between the pial and the white matter surfaces is the coordinates of the vertices, but every vertex in the pial surface has a corresponding vertex in the white matter surface. Additionally, there can be another surface, called the midgrey surface, which is computed by averaging the coordinates of the pial and the white matter surfaces. It thus has the same polygons and runs through the middle grey matter layer.

Transforming the brain volume into a surface allows to assign features to the vertices, thus gaining additional information. For example, curvature describes how the cortex is folded locally into sulci and guri, thickness describes the local thickness of the grey matter, and sulcal depth describes how far a vertex is away from the surface of the cortex, with higher values indicating higher depth on the sulcus and lower values indicating lower depth on the gyrus.

Additionally, groups of vertices can represent specific regions of interests, or functional data. To accurately assign functional data to the surface, the surface must be aligned with the volumetric data, which is handled by specific algorithms. For this assignment, an algorithm that generates the surface from the middle grey matter layer has already been applied, so what remains to be done is the mapping from the functional data to the surface.

Even though surface-based analysis does not include subcortical structures and it is difficult to apply in subjects with neurological abnormalities, it has several advantages. Firstly it improves intersubject alignment. Due to the folded nature of the cortex, when doing volumetric-based normalization, a small error in the three-dimensional space may translate into a very large error across the cortical surface. However, with surface-based normalization, the brain of all individuals is inflated and transformed into a spherical representation, thus allowing to align the sulci and the gyri based on sulcal depth information. Any errors that may still occur with this process are found across the cortical surface, and not in the three-dimensional space. Secondly, surface-based analysis improves smoothing, as volumetric-based smoothing smears the data across the sulci resulting in smoothing way further than the surface of the cortex. In contrast, smoothing across the surface is more effective as it only includes neighbouring grey matter. Thirdly, surface-based analysis allows for more informative visualizations, while it is also better for interpreting topographic activity.

3.2 Methods

The mapping of the functional data to the surface was done using the *Connectome Workbench* software from the Human Connectome Project. Specifically, the fMRI data from the tongue task of subject 1 were selected, after first running a pre-processing pipeline. The reconstruction of the surface in native space had already been performed with the *FreeSurfer* software and had been converted to Gifti (*.gii*) format, so that it is compatible with the *Connectome Workbench*. For each hemisphere, it included pial, white matter, and midgrey surfaces, Aparc and Aparc.2009a labels, sulcal depth, cortical thickness, and curvature, as well as the T1-weighted image aligned with the surface reconstruction.

3.2.1 Pre-processing

Pre-processing of the functional data was performed with the SPM toolbox in a matlab environment. Firstly, the data were realigned and resliced using the $Realign : Estimate \ \& \ Reslice$ function, to remove noisy movement that is not of interest to the task at hand. All parameters were left at default. Secondly, coregistration was performed to align and register the functional brain images acquired during the fMRI tasks with the T1-weighted image of the surface. This ensures that the functional data that are created are aligned with the surface. Using SPM's $Coregister : Estimate$ function, the reference image was defined as the T1-weighted image of the surface, the source image was the mean image that was generated when applying realignment, and the other images were the functional data after applying realignment and reslicing.

Moreover, the first-level statistical model was specified to estimate the task-related brain activity per voxel for this subject. This was done using SPM's $fMRI\ model\ specification$ function, where the interscan interval was set to 2.5 seconds, the microtime resolutions was set to 40, and the microtime onset was set to 21. The cue onset and durations were specified as given in the assignment details of the course *Basic fMRI analhysis*, while no extra regressors were included, and the masking threshold was set to 0.5. Next, the parameters of the model were estimated using SPM's $Model\ estimation$ function, where the input model was specified by a dependency on the model specification *SPM.mat* file that was generated from the previous step. The method used for parameter estimation was the restricted maximum likelihood method, also referred to as "classical" in SPM. Additionally, SPM's $Contrast\ Manager$ function was used to specify the t-contrast using the *SPM.mat* file with the estimated parameters of the previous step by a dependency. Since the interest lies merely in estimating the statistically significant voxels and not in doing any comparison, the weight vector was just set to 1.

Lastly, the functional results were resilced to the T1-weighted image of the surface using SPM's *Coregister* : *Reslice* function. This is a necessary step because the mapping procedure assumes that the functional images that are mapped to the surface have the same dimensions, sizes, and orientation as the T1-weighted image in the surface reconstruction.

The above procedure produces the functional output that is required to map the functional data to the surface. This is the *rspmT_0001.nii* file that is given to the command described in the following section, 3.2.2.

3.2.2 Mapping the functional data to the surface

After having pre-processed the raw fMRI data, and downloaded the surface reconstruction data in Gifti files, mapping the functional data to the surface was done using the following commands in a Linux system with the *Connectome Workbench* software installed.

For the left hemisphere:

```
wb_command -volume-to-surface-mapping Desktop/UU/advanced_fMRI/assignment3/
    sub01/tongue_output/rspmT_0001.nii Desktop/UU/advanced_fMRI/assignment3/
    sub01_surfdat/sub01_t1/lh/sub01_t1_lh.graymid.surf.gii Desktop/UU/
    advanced_fMRI/assignment3/sub01_surfdat/sub01_t1/lh/sub01_tongue_tmap_lh.
    func.gii -enclosing
```

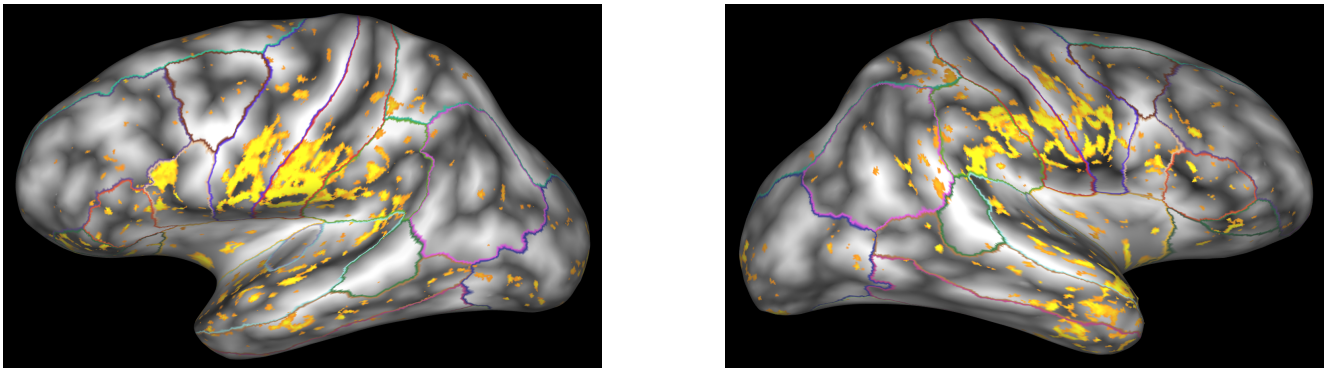
And for the right hemisphere:

```
wb_command -volume-to-surface-mapping Desktop/UU/advanced_fMRI/assignment3/
    sub01/tongue_output/rspmT_0001.nii Desktop/UU/advanced_fMRI/assignment3/
    sub01_surfdat/sub01_t1/rh/sub01_t1_rh.graymid.surf.gii Desktop/UU/
    advanced_fMRI/assignment3/sub01_surfdat/sub01_t1/rh/sub01_tongue_tmap_rh.
    func.gii -enclosing
```

Afterwards, the anatomical surface image, along with the surface data, and the generated functional data, specified as **.func.gii* files, were loaded into the *Connectome Workbench* user interface for each hemisphere. The results were visualized on an inflated brain surface, showing the aparc anatomical labels with a drawing type of *Outline Label Color*, as well as the sulcal depth on a grey scale. The functional data were visualized on a red-yellow scale, thresholded at a minimum t-value of 5. The high threshold was set to the maximum t-value for each hemisphere, while selecting to only show the data inside the thresholds.

3.3 Results

The following figures show the functional data of the tongue task for subject 1 mapped to the inflated brain surface. The *aparc* anatomical labels are also superimposed, while sulcal depth information is also included.



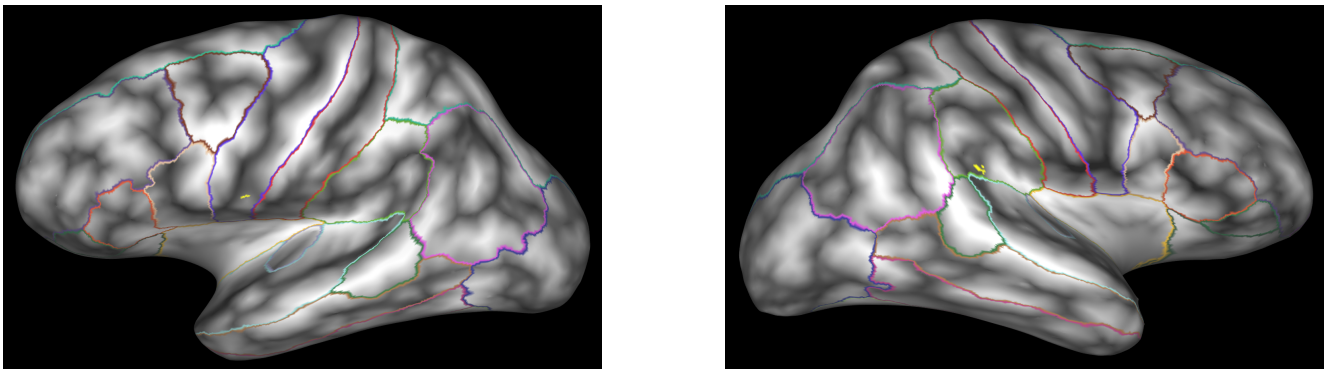
(a) Left hemisphere thresholded at a minimum t-value of 5

(b) Right hemisphere thresholded at a minimum t-value of 5

Figure 4: Results thresholded at a minimum t-value of 5

Figure 4 shows the results when thresholded at a minimum t-value value of 5. The more yellow the color, the higher the t-value, and the more red/orange, the lower the t-value. It becomes apparent that there is quite

widespread activation for the tongue task, but most of the activity seems to be located at the supramarginal, precentral, and postcentral areas.



(a) Left hemisphere showing the highest t-values

(b) Right hemisphere showing the highest t-values

Figure 5: Results showing the highest t-values

Figure 5 shows the highest t-values for the functional data, superimposed on the inflated brain surface. For the left hemisphere, this is on the precentral area, while for the right hemisphere this is on the supramarginal area. These results are not surprising, as these areas seem to correspond to the areas that were found to be more strongly activated during the tongue task in the course *Basic fMRI analysis*.

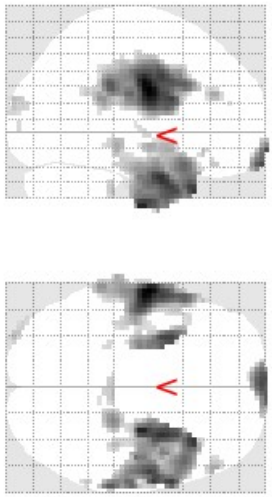


Figure 6: Volumetric activation related to the tongue task from the Basic fMRI analysis course

Figure 6 shows the volumetric areas that are more strongly activated for a single subject during the tongue task. Even though this is not the same subject, it becomes clear that the overall pattern looks similar, while most intense activity seems to be in the precentral, postcentral, and supramarginal areas.

Interestingly, the area with the highest t-value is not the same in the two hemispheres for this subject. However, both these areas, the precentral and the supramarginal, contain voxels with very high t-values, so the results do not seem to indicate an abnormality, but rather a widespread activation during the tongue task.