Multimodal MRI cerebral correlates of verbal fluency switching and its impairment in women with depression


Introduction

Major depressive disorder (MDD) affects 350 million people (women predominantly) worldwide and is disabling due to chronicity (James et al., 2018). In practice, it remains challenging for clinicians to predict the course of MDD given its clinical heterogeneity and the lack of biomarkers. The development of magnetic resonance neuroimaging (MRI) technologies has enabled significant advances towards the quest for such biomarkers. Specifically, anatomical MRI (e.g. structural, diffusion-weighted MRI), by allowing visualization and quantification of brain alterations in MDD have played a critical role in deciphering the pathogenesis of the disease. From structural MRI, cortical thickness (i.e. the distance between the white matter-gray matter surface) can be estimated. Cortical thickness (CT) permits a higher registration accuracy for depression cerebral bases. The search of biomarkers in the field of depression requires easy implementable tests that are biologically rooted. Qualitative analysis of verbal fluency tests (VFT) are good candidates, but its cerebral correlates are unknown.

Methods: We collected qualitative semantic and phonemic VFT scores along with grey and white matter anatomical MRI of depressed (n = 26) and healthy controls (HC, n = 25) women. Qualitative VFT variables are the “clustering score” (i.e. the ability to produce words within subcategories) and the “switching score” (i.e. the ability to switch between clusters). The clustering and switching scores were automatically calculated using a data-driven approach. Brain measures were cortical thickness (CT) and fractional anisotropy (FA). We tested for associations between CT, FA and qualitative VFT variables within each group.

Results: Patients had reduced switching VFT scores compared to HC. Thicker cortex was associated with better switching score in semantic VFT bilaterally in the frontal (superior, rostral middle and inferior gyri), parietal (inferior parietal lobule including the supramarginal gyri), temporal (transverse and fusiform gyri) and occipital (lingual gyr) lobes in the depressed group. Positive association between FA and the switching score in semantic VFT was retrieved in depressed patients within the corpus callosum, right inferior fronto-occipital fasciculus, right superior longitudinal fasciculus extending to the anterior thalamic radiation (all p < 0.05, corrected).

Conclusion: Together, these results suggest that automatic qualitative VFT scores are associated with brain anatomy and reinforce its potential use as a surrogate for depression cerebral bases.

Keywords: Cortical thickness, Verbal fluency, Switching, Automatic speech analysis, Major depressive disorder, Multimodal MRI cerebral correlates of verbal fluency switching and its impairment in women with depression.
than volume-based registration (Desai et al., 2005), reduces the risk of possible false positives (Greve and Fischl, 2018) and is considered as a sensitive (Lerch and Evans, 2005) and specific (Lerch et al., 2008) measure of brain structure. Moreover, other standard approaches such as surface and volume estimation should be normalized to the total intra cranial volume, leaving room to possible variance in the estimated effect sizes (Schwarz et al., 2016). Recent results have reported widespread cortical thinning in depressed patients in comparison to healthy controls (HC) in the orbitofrontal cortex, insula, anterior and posterior cingulate, and temporal lobes (Schmaal et al., 2017), providing candidates MRI-derived biomarkers (Jiang et al., 2020). Diffusion weighted imaging measures water molecule diffusion to estimate fractional anisotropy (FA) along the white matter bundles. Large sample analyses refer to small effect sizes but widely distributed reduced FA in MDD compared to HC within the superior and inferior longitudinal fasciculus (S/ILF), the corpus callosum (CC), the uncinate, the inferior fronto-occipital fasciculus (IFOF) (van Velzen et al., 2020). These findings have led to the suggestion that structural and diffusion-weighted MRI could be a promising tool for identifying imaging biomarkers of MDD. However, these methods have contraindications, are costly, and are currently not widely available. Therefore, there is a need for simple tools, doable in routine but which have clear associations with biological, a condition required to be considered as biomarkers (Ruggeri et al., 2014). Cognitive functions are good candidates since: i) executive dysfunction is frequent in patients with MDD (Snyder, 2013; Fossati et al., 2003), (ii) it can be performed in clinical settings. Recent development automates their assessments (https://ki-elements.de/en/start/, https://www.cambridgecognition.com/) which make it easy to quantify and standardize individual performances against population norms. Verbal fluency tasks (VFT) are among the quickest and easiest to acquire and measures both executive and language abilities (Snyder, 2013) participants are required to produce as many words belonging to a semantic or phonemic category as possible within a given time (Troyer et al., 1997). However, in a recent meta-analysis, VFT failed to predict the course of MDD (Pimontel et al., 2016). This might be related to the scoring method, as most studies use the quantitative score (i.e. simply summing correct words) which does not shine any light on the cognitive function. In contrast, the qualitative analysis of the performance provides a more fine-grained measure of cognitive function (Troyer et al., 1997). It consists in measuring the size of word subcategories or cluster (clustering score) and counting the number of transitions (switching score) between clusters (Troyer et al., 1997). Those “clustering” and “switching” scores reflects the integrity of the storage in lexico-semantic memory and executive function respectively (Troyer et al., 1997). Separating between clustering and switching should allow clinicians to distinguish the two cognitive processes occurring during VFT: accessing semantic memory to retrieve words and executive search functions active in navigating the larger semantic stock. Typically, the clusters are determined using predefined subcategories, following the approach from Troyer et al. (Troyer et al., 1997). This leaves room for subjective interpretation, and variance in cluster sizes and switch counts. Statistical model-based approaches which allow for automatic cluster identification reduce this variance and showed promising results of identifying early stages of major cognitive disorders (Linz et al., 2017; Linz et al., 2017; Troyer et al., 2017; König et al., 2015). This offers new perspectives on using qualitative VFT performance (i.e. clustering and switching scores) as reliable cognitive marker of MDD. However, this is currently unknown if qualitative VFT performance and measures of CT and FA in sample of patients with MDD and in HC. The secondary objective was to compare the groups in term of switching and clustering scores during both VFT, and in term of MRI-derived metrics (CT and FA).

2. Material and methods

2.1. Participants

We included 51 women (26 patients) between July 2012 and October 2014. All patients were outpatient from an expert center in mood disorders of Rennes University Hospital (France). All participants were native French speakers. All patients were suffering from either chronic or recurrent depression. Patients met Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR) criteria for MDD and had a minimum Montgomery Åsberg Depression Rating Scale (MADRS) score of 21. Clinical characteristics of the sample is provided by the mean of the 4-factor structure of the MADRS (Quilty et al., 2013). They had no other psychiatric disorder, as measured with the Mini International Neuropsychiatric Interview (MINI) and were all native French speakers. All assessment (including MRI) were performed within the week after inclusion. Healthy volunteers did not have current psychiatric disorder (according to the MINI). Since VFT are amenable to effects of age (Tombaugh, 1999) and education (Tombaugh, 1999), the HC were matched for age and education level with the MDD group. All participants were right-handed (according to the Edinburgh Handedness Inventory) and did not exhibit any severe cognitive impairment (defined as a score < 130 on the Mattis Dementia Rating Scale) or suicide risk (assessed by the Clinical Global Impression of Severity of Suicidality Scale). Exclusion criteria were potential contraindications to MRI (pacemakers, metal implants, pregnancy, and lactation). The study was approved by the relevant institutional review board (CPP of Nancy number 2759, ID-RCB number 2019-A00111-56). All participants provided their written informed consent.

2.2. Procedure

After inclusion, each participant underwent a structural MRI scan, and psychiatric and neuropsychological assessments within the same week.

2.2.1. Neuropsychological assessment

All participants underwent the same neuropsychological battery, in the same order. They were asked to generate as many words as possible within the space of 120 s. For semantic VFT, they had to produce words according to a category criterion (“animal”). For phonemic VFT, they had to produce words beginning with the letter “P”. The words were split into 30-seconds time frames. Additionally, processing speed was assessed with the digit-symbol substitution test (DSST), cognitive switching with the TMT, cognitive inhibition with the Stroop Test, abstract reasoning and set-shifting with the modified Wisconsin Card Sorting Test (WCST) and working memory with the direct/indirect spans.

2.2.2. Fully automatic analysis of the verbal fluency tests

To compute the clusters, we used the statistical method described by Linz et al. (Linz et al., 2017). This automatic method of analysis requires words (textual information) to be converted into a numerical form (vector). This conversion method is known as word vectorization. Each
into clusters. Semantic distance is calculated as the semantic distance between all possible word pairs for each subject during the VFT, estimated based on large word corpora. Here, semantic distance was determined using the fastText word embedding pipeline (Lin et al., 2017); pretrained on the Common Crawl® and the Wikipedia® corpora (Grave and Bojanowski, 2018). It acts like a map of the semantic search performed by each subject. We then applied Troyer’s rules for scoring the cluster size and switch numbers (Troyer et al., 1997). The mean cluster size is the average number of words contained in a semantic cluster (a sequential group of words that have a semantic relation). The size of any cluster is the number of words in it minus one (as single word clusters are considered length zero). The cluster size reflects the integrity of the lexico-semantic memory, while the switching score is related to the ability to access and navigate among the semantic stock. Typically, if executive functions are impaired, patients are more likely to latch onto a semantic cue they have discovered and produce words that are closely related (i.e., fewer switches). If semantic memory is impaired, patients will move on more quickly from a given cue and therefore explore more different semantic categories (i.e., smaller clusters). Both are triggered by the time constraint and the goal of the subject to produce as many words as possible. Currently, there are no existing norms for the qualitative measures of VF in French.

2.2.3. MRI acquisition

All MRI studies were performed on a 3 T Magnetom Verio Syngo MR B17 scanner (Siemens Healthcare, Erlangen, Germany). An anatomical scan was performed using the following 3D T1-weighted, sagittal, magnetization prepared rapid gradient echo (MPRAGE) sequence: 176 sagittal slabs, voxel size 1 × 1 × 1 mm³, repetition time 1900 ms, echo time 2.26 ms, inversion time 99 ms, field of view 256 × 256 mm², number of excitations 1, GRAPPA acceleration factor 2. Diffusion weighted images were acquired with an EPI sequence (30 directions, b-value 1000 s/mm²) with repetition time 11000 ms, echo time 99 ms, field of view 256 × 256 mm², 60 slices and voxel size: 2 × 2 × 2 mm³.

2.2.4. Cortical thickness

Data were preprocessed and analyzed using Statistical Parametric Mapping (SPM) Version 12 (Ashburner and Friston, 2000) (Welcome Department of Cognitive Neurology, London, UK, http://www.fil.ion.ucl.ac.uk/spm/) with MATLAB 2018b (The Mathworks Inc., Sherborn, MA, United States). First, T1-weighted images of each participant were manually reoriented to the anterior commissure. The reoriented structural images were then segmented using the Segment option of the CAT (Computational Anatomy Toolbox) Version 12.6 (Gaser and Dahmke, 2016). The resulting segmented images were then used to estimate the cortical thickness and central surface simultaneously, using the projection-based thickness. This approach does not require a reconstruction of the outer boundary or white matter surface deformations for thickness estimation within sulci where partial volume effect and blurred images

### Table 1
Clinical and Demographic Variables. Values given are means (Standard Deviation). Except for the variable “treatment” where values are proportion (%).

<table>
<thead>
<tr>
<th>Variable</th>
<th>MDD group (n = 26)</th>
<th>HC group (n = 25)</th>
<th>Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>46.69 (11)</td>
<td>49.5 (9.5)</td>
<td>W = 281, p = 0.41</td>
</tr>
<tr>
<td>Education (years)</td>
<td>12.9 (2.3)</td>
<td>13.7 (2.4)</td>
<td>W = 259, p = 0.21</td>
</tr>
<tr>
<td>Number of depressive episodes</td>
<td>4.7 (3.5)</td>
<td>11.9 (13.5)</td>
<td></td>
</tr>
<tr>
<td>Disease duration (years)</td>
<td>14.5 (6.3)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Phonic VFT

**Fig. 1.** Scatter dot plots of the switching and clustering scores stratified by groups and phonemic and semantic VFT.
2.2.5. Fractional anisotropy

et al., 2015). Vectors, ExploreDTI applies motion and eddy current-induced geometrical distortions by applying an appropriate B matrix rotation to carded. After brain extraction, we used ExploreDTI (http://www.exploredti.com), running on MATLAB 2018b, to preprocess the data. After having created a matrix which incorporates the b-values and the b-vectors, ExploreDTI applies motion and eddy current-induced geometrical distortion corrections by applying an appropriate B matrix rotation (Leemans and Jones, 2009). Finally, the tensor is estimated on a voxel-based basis using a linear least square approach embedded in the informed Robust Extraction of kurtosis INDices with linear estimation (REKINDLE) algorithm (Tax et al., 2015). Informed REKINDLE is an improved method which additionally removes physiological noise in low redundancy datasets (30 directions). Fractional anisotropy (FA) maps are then calculated accordingly (Basser and Pierpaoli, 1996).

The FA maps were processed using tract-based spatial statistics (Smith et al., 2007) part of FSL (Jenkinson et al., 2012). TBSS performs voxel-based statistics projected onto a “skeleton” of the major tracts to avoid partial volume effects and does not imply any a-priori smoothing step. Each of the 51 (MDD and HC) FA map was registered to each other and the resulting deformations fields were compared to identify the target FA map. The target FA map was non-linearly registered to the MNI152 space at 1 × 1 × 1 mm³ resolution. Finally, every FA map was non-linearly warped by combining the transformation from the subject’s FA map to the representative FA map and the transformations from the latest to the MNI152 space. The mean FA image across all participants was computed, “skeletonized” and thresholded at 0.2. TBSS then projected the maximum FA value found along the normal of the skeleton onto it.

2.3. Statistical analyses

2.3.1. Clinical and verbal fluency data analyses

We used Wilcoxon two-sample tests for between-group analyses. As qualitative VFT measures are sparsely documented, we used non-parametric Spearman correlations between qualitative FA measures and clinical characteristics, as long with other executive functions in both groups. We quantified the contribution of MDD to number switches during the semantic VFT, as post-hoc analysis, while accounting to other executive functions (abstract reasoning/set shifting (MCST-time), working memory (backward span), cognitive inhibition (Stroop interference score) and cognitive switching (TMT B-A)), using a Poisson regression.

2.3.2. Cortical thickness analyses

We tested for positive and negative associations between CT (within the 68 ROIs (Desikan et al., 2006) and clustering/switching scores after adjusting for age within each group (MDD and HC) using linear regressions with CT as the dependent variable. Between-group comparisons were performed, while accounting for age. Multiple testing correction was controlled using the False Discovery Rate (FDR) (Benjamini and Hochberg, 1995).

2.3.3. Fractional anisotropy analyses

The skeletonised FA data were analyzed using the FSL “Threshold-Free Cluster Enhancement” option in “randomise” with 5000 permutations (Nichols and Holmes, 2002; SMITH and Nichols, 2009) of the maximum statistic of the cluster as a non-parametric approach of multiple testing correction which reduce false positive for cluster inference (Eklund et al., 2016). Tracts were then identified using the John-Hopkins University (JHU) white-matter tractography atlas (Hua et al., 2008). In each group (MDD and HC), we tested for positive and negative correlations between voxel-wise FA values and the switching and clustering scores, accounting for age within MDD and HC groups separately. We also performed between-groups comparisons also accounting for age.

2.3.4. Exploratory group by brain interactions

Our primary goal was to identify associations within the MDD group only to derive VF behavioral markers of cerebral physiopathology. However, to determine that our findings are specific to MDD condition, we also conducted exploratory group by brain interactions on semantic switching scores. We performed 25 models testing for group by CT interactions and 7 models testing for group by FA interactions, all accounting for age. False Discovery rates multiple correlation were applied.

For all statistical analyses, multiple testing corrected p < 0.05 type 1 error rate was applied.

3. Results

3.1. Study population

Participants’ clinical and demographic characteristics are displayed in Table 1. Of relevance, patients had moderate to severe apathy with high levels of anxiety, without any general cognitive deficits. Groups
were comparable in terms of age, education and total intra-cranial volumes. Supplementary Table 1 displays the associations between the MDD clinical characteristics and the performances on the executive functions and with the general cognitive functions. Supplementary Table 2 displays the descriptive statistics of the cognitive performances for each group and the between group comparison. Nineteen percent of the MDD sample had working memory (backward span) considered as outside the norms, 11.5%, 27%, 15.4% had abnormal cognitive switching (TMT B-A), semantic and phonemic VF, respectively.

3.2. Between group comparison

3.2.1. Verbal fluency tests

The switching score was significantly reduced in the MDD group,
compared with HC in both VFT (Fig. 1). Quantitative (stratified by 30 s time-window) and qualitative indices for both groups and task are displayed in Table 2. Group differences on both VFT are essentially supported by number of switches without any cluster sizes difference. Raw quantitative number of words are smaller in the MMD group, essentially during the first minute in both VFT.

3.2.2. Cortical thickness

The comparison between the two groups on whole-brain CT did not reveal any significant differences (FDR corrected).

3.2.3. Fractional anisotropy

After FDR correction, we observed significantly lower FA for MDD patients in comparison with HC in widespread bilateral white matter tracts corresponding to a cluster of 55 533 voxels including: the forceps minor and major, bilateral inferior fronto-occipital fasciculus (IFOF), bilateral uncinate fasciculi (UF), inferior longitudinal fasciculi and superior longitudinal fasciculi (SLF) (Supplementary Fig. 1).

4. Clustering and switching scores correlates within MDD and HC groups

4.1. Executive functions

Tables 3A and 3B display within-groups clinical and executive correlates of VFT measures for phonemic and semantic VFT, respectively. Both semantic and phonemic number of switches are associated with processing speed (DSST and TMT A and B) and cognitive inhibition. Only number of switches during semantic VFT is associated with AES among MDD patients. Table 4 displays the standardized $\beta$, exponentiated $\beta$ and 95IC of the Poisson regression for each variable. Compared to HC, being part of the MDD group significantly have 18% the number of switches during the semantic VFT. Each increase in Stroop interference score would increase by 1.3% the number of switches during the semantic VFT.

4.2. Cortical thickness

In the MDD group, we found significant positive correlations between the switching score in the semantic VFT and CT in 25 regions including the bilateral frontal (inferior, and superior, rostral middle

<table>
<thead>
<tr>
<th>Variables</th>
<th>Beta</th>
<th>95 IC</th>
<th>Exp(Beta)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDD Group</td>
<td>-0.2</td>
<td>-0.37; -0.03</td>
<td>0.82</td>
<td>0.02</td>
</tr>
<tr>
<td>Backward span</td>
<td>-0.01</td>
<td>-0.91; 0.07</td>
<td>1</td>
<td>0.8</td>
</tr>
<tr>
<td>Stroop interference</td>
<td>0.15</td>
<td>0.004; 0.02</td>
<td>1.015</td>
<td>0.008</td>
</tr>
<tr>
<td>TM B-A</td>
<td>0</td>
<td>-0.002; 0.003</td>
<td>1</td>
<td>0.8</td>
</tr>
<tr>
<td>MCST - Time</td>
<td>0</td>
<td>-0.002; 0.001</td>
<td>1</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Fig. 2. Regions showing significant positive correlations between switching score during the semantic VFT and CT within the MDD group (age considered as covariate). FDR corrected group by CT interaction plots are also displayed for the left lingual gyrus and the right cuneus (left and right panel, respectively). Color bar depicts the p-value. L: left, R: right. Regions displayed are bilateral lingual gyrus, bilateral superior frontal gyrus, middle frontal gyrus, inferior parietal lobule, transverse and superior temporal and left posterior cingulate.
gyri), parietal (inferior parietal lobule including the supramarginal gyrri), temporal (transverse and fusiform gyrri) and occipital (pericalcarine and lingual gyrri) areas. There were no significant associations between CT and switching score during the semantic VFT within the HC group (all p > 0.05 uncorrected). Only the right cuneus (t = 3, β = 32, p = 0.047) and the left lingual gyrus (t = 3.5, β = 36 Ashburner et al., 2003, p = 0.02) showed significant group by CT interaction after FDR correction (Fig. 2 and Table 5). There was no association between the clustering score in the semantic VFT and CT neither in the MDD nor the HC groups. There was no association between either score in the phonemic VFT and mean FA neither in the MDD nor the HC groups.

5. Discussion

Here, we found that the semantic switching score correlated with CT in various regions distributed bilaterally in the frontal (superior, rostral middle and inferior gyrri), parietal (inferior parietal lobule including the supramarginal gyrri), temporal (transverse and fusiform gyrri) and occipital (lingual gyrri) lobes. After FDR correction, only the right cuneus and the left lingual gyrus showed similar pattern of group by cortical thickness interaction. This switching score also correlated with FA in the CC, right SLF extending to the ATR, and IFOF. Only the right superior longitudinal fasciculus, right arcuate, right anterior part of the CC showed group by FA interaction. Our results revealed that MDD subjects had decreased switching scores in both VFT compared to HC.

Our results are consistent with our a-priori hypothesis of thinner cortices among prefrontal regions, supported by recent results which investigated CT basis of semantic fluency in a large sample (n = 505) of healthy older participants and found positive correlations with quantitative semantic fluency within the left superior, rostral middle and inferior frontal gyri (Vonk et al., 2019). The left superior, rostral middle and inferior frontal gyri, left supramarginal gyrri, left fusiform, left transverse temporal gyri, and left lingual gyrus were also involved, partially overlapping with our results. An increased activation within the middle and inferior frontal gyri, and bilateral parietal cortex (superior and inferior parietal lobule) was also found during both self-reported switching in comparison to constrained switching VFT among healthy individuals (Hirshorn and Thompson-Schill, 2006). Also, cognitive switching with reduced working memory component recruits both the middle frontal gyri along with bilateral inferior parietal lobule and superior/transverse temporal gyri, in line with the current results (Smith et al., 2004).

Switching consists in shifting between clusters when one cluster is exhausted (Troyer et al., 1997). It requires: (i) to identify all the category-relevant items within the semantic memory, and (ii) to retrieve and select the words accordingly to the recommendation (i.e. “animal” or “letter P”) which involves attention and working memory (Ralph et al., 2017). There is substantial evidence that those cognitive processes are supported by different and partially overlapping neural networks (Ralph et al., 2017).

(i) The organization of the semantic memory has been described by the prominent hub-and-spoke theory (Patterson et al., 2007): the knowledge of a semantic concept (e.g. cat) is stored both in a “semantic hub”, and in distributed cortical regions related to its sensorimotor attributes (e.g. shape, name, colors, motion, sound, function) called “the spokes”. The “hub” is assumed to integrate inputs from activated sensorimotor feature sets, categorize, for example, both cat and fish as “animals”. This hub is suggested to be localized in the bilateral anterior temporal lobes (ATL), including the fusiform gyrus (Pobric et al., 2010). The semantic spokes are suggested to be supported by primary and associative auditory (bilateral transverse temporal gyrri), motor (the supplementary motor area/paracentral lobule) and visual areas (lingual gyrri, middle and lateral occipital gyrri and cuneus) cortices. It has been described that the categories of concepts can be dependent on a particular modality. For example, whereas “tools” can be individuated by their associated actions (e.g. hammer/hitting), “animals” are rather individuated by virtue of their constituent sensory features (e.g. shape, colors). Therefore, naming “animals” might recruit greater responses in visual areas) compared to naming “tools” (Chouinard and Goodale, 2010).

(ii) Semantic VFT implies a controlled retrieval and selection of the semantic information, also known as “semantic control”, which has been suggested to rely on inferior frontal gyri (Ralph et al., 2017). More precisely, it seems that the most posterior part of the inferior frontal gyri

---

**Table 5**

Regions showing significant negative associations between the switching score during the semantic VFT and cortical thickness (p < 0.05, False Discovery Rate (FDR) corrected for multiple comparisons.

<table>
<thead>
<tr>
<th>Left hemisphere</th>
<th>P-value (FDR corrected)</th>
<th>T-value</th>
<th>Z-value</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.003</td>
<td>4.79948 3.95433</td>
<td></td>
<td></td>
<td>Lingual gyrus (BA 18/19)</td>
</tr>
<tr>
<td>0.03</td>
<td>3.40297 3.03058</td>
<td></td>
<td></td>
<td>Supramarginal gyrus (BA 40)</td>
</tr>
<tr>
<td>0.03</td>
<td>2.89176 2.62458</td>
<td></td>
<td></td>
<td>Insmouth of cingulate gyrus (BA 26/28)</td>
</tr>
<tr>
<td>0.03</td>
<td>2.88486 2.63715</td>
<td></td>
<td></td>
<td>Transverse Temporal gyrus (BA 41/42)</td>
</tr>
<tr>
<td>0.032</td>
<td>2.79163 2.56333</td>
<td></td>
<td></td>
<td>Posterior Cingulate (BA 24)</td>
</tr>
<tr>
<td>0.032</td>
<td>2.72878 2.51304</td>
<td></td>
<td></td>
<td>Inferior parietal lobule (BA 40/39)</td>
</tr>
<tr>
<td>0.039</td>
<td>2.56065 2.37646</td>
<td></td>
<td></td>
<td>Rostral Middle frontal gyrus (BA 10)</td>
</tr>
<tr>
<td>0.043</td>
<td>2.46341 2.29611</td>
<td></td>
<td></td>
<td>Pars Opercularis (BA 44)</td>
</tr>
<tr>
<td>0.047</td>
<td>2.32776 2.18239</td>
<td></td>
<td></td>
<td>Superior Frontal gyri (BA 6)</td>
</tr>
<tr>
<td>0.047</td>
<td>2.32177 2.17732</td>
<td></td>
<td></td>
<td>Fusiform gyrus (BA 37)</td>
</tr>
<tr>
<td>0.048</td>
<td>2.29235 2.15237</td>
<td></td>
<td></td>
<td>Pars Triangularis (BA 45)</td>
</tr>
<tr>
<td>0.048</td>
<td>2.26827 2.13191</td>
<td></td>
<td></td>
<td>Pericalcarine nuc(lus (BA 17)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Right hemisphere</th>
<th>P-value (FDR corrected)</th>
<th>T-value</th>
<th>Z-value</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.029</td>
<td>3.49092 3.09454</td>
<td></td>
<td></td>
<td>Parahippocampal gyrus (BA 28/34/35/36)</td>
</tr>
<tr>
<td>0.03</td>
<td>3.27641 2.9371</td>
<td></td>
<td></td>
<td>Supramarginal gyrus (BA 40)</td>
</tr>
<tr>
<td>0.03</td>
<td>3.15651 2.84699</td>
<td></td>
<td></td>
<td>Pariacentral lobule</td>
</tr>
<tr>
<td>0.03</td>
<td>3.01009 2.73486</td>
<td></td>
<td></td>
<td>Superior Frontal gyrus (BA 6)</td>
</tr>
<tr>
<td>0.03</td>
<td>2.91241 2.65879</td>
<td></td>
<td></td>
<td>Cuneus (BA 17)</td>
</tr>
<tr>
<td>0.03</td>
<td>2.90005 2.6491</td>
<td></td>
<td></td>
<td>Rostral Middle frontal gyrus (BA 10)</td>
</tr>
<tr>
<td>0.032</td>
<td>2.78501 2.55805</td>
<td></td>
<td></td>
<td>Lingual gyrus (BA 18/19)</td>
</tr>
<tr>
<td>0.032</td>
<td>2.71204 2.49957</td>
<td></td>
<td></td>
<td>Inferior parietal lobule (BA 40/39)</td>
</tr>
<tr>
<td>0.039</td>
<td>2.56991 2.38406</td>
<td></td>
<td></td>
<td>Fusiform gyrus (BA 37)</td>
</tr>
<tr>
<td>0.043</td>
<td>2.46611 2.29834</td>
<td></td>
<td></td>
<td>Pars Opercularis (BA 44)</td>
</tr>
<tr>
<td>0.047</td>
<td>2.37309 2.2206</td>
<td></td>
<td></td>
<td>Lateral Occipital (BA 19)</td>
</tr>
<tr>
<td>0.047</td>
<td>2.34753 2.19907</td>
<td></td>
<td></td>
<td>Transverse Temporal (BA 41)</td>
</tr>
<tr>
<td>0.047</td>
<td>2.30821 2.16583</td>
<td></td>
<td></td>
<td>Superior Temporal (BA 22)</td>
</tr>
</tbody>
</table>

4.3. Fractional anisotropy

In the MDD group, we found positive correlation between FA values and the switching score in semantic VFT in the MDD group within 7 clusters of 5706 (right superior longitudinal fasciculus extending to the right arcuate fasciculus to the body of the CC (bilateral) and to the right inferior fronto-occipital fasciculus, 134 (the right anterior corpus callosum (CC) extending to the right anterior corona radiata), 74, 33, 11, 7 and 3 voxels including: the right anterior corpus callosum (CC) extending to the right anterior corona radiata, right IFOF, right SLF extending to the anterior thalamic radiation (ATR) (FDR corrected) (Fig. 3 and Table 6). Only voxels in the right superior longitudinal fasciculus, right arcuate, right anterior part of the CC and anterior corona radiata showed significant interaction group by FA interaction, after FDR correction of multiple testing (Table 6). After FDR correction, we did not find any significant correlation between mean FA and the clustering score in the semantic VFT neither in the MDD nor the HC groups.
The pars opercularis and pars triangularis are more critically involved in the selection process (Badre et al., 2005). While inferior frontal gyri seems to be critical, semantic control also relies on posterior middle temporal gyri, the inferior parietal lobule and the intraparietal sulcus within a “semantic control network” (Ralph et al., 2017).

Secondly, there is growing evidence that white matter tracts are involved in quantitative VFT processing. For example, poor quantitative VFT scores have been associated with greater left SLF damage in a study which used voxel-based lesion symptom mapping in patients with penetrating traumatic brain injury (Cristofori et al., 2015). Both mean FA value and lesion of the left SLF, the left IFOF, and left ATR were correlated with the score in semantic VFT in stroke patients (Li et al., 2017). Finally, associations between FA along the IFOF and semantic VFT scores were observed in patients with left diffuse low-grade glioma (Almairac et al., 2015). Although these previous findings align with our current results, they do not address their specific cognitive function in the same way as qualitative VFT scores. Controlled retrieval, selection in accordance with semantic criteria processes have not been associated with white matter MRI-derived measures. Nevertheless, the SLF, the IFOF and the ATR have been associated with other cognitive processes required during semantic switching. For example, the right SLF and IFOF have been related with semantic categorization performance (i.e. the mental operation by which the brain classifies objects and events) in a voxel-based morphometry in healthy subjects (Garcin et al., 2018). The FA of SLF and IFOF have also been found to be positively correlated with processing speed in healthy older adult (Kerchner et al., 2012). A positive association was also found between verbal working memory performance assessed by the letter-number span task and FA in bilateral SLF young healthy subjects (Peters et al., 2012).

Inspection of the interaction plots (Figs. 2 & 3) suggests that for all regions (for either the CT or the FA analyses) there might be a negative correlation between the semantic VF switching score and the MRI-derived brain features within the HC group. This is quite unexpected since the literature suggests that better cognitive functions are associated with thicker cortices and higher FA. We therefore explored if these associations were statistically significant and none of them were (all p > 0.05, without any correction for multiple testing). Adding to the fact that (i) none of the whole brain analyses in the HC group revealed significant associations and (ii) that the regions where the interaction terms were statistically significant are very similar to the ones which showed the greater effect size within the MDD group (left lingual gyrus and right SLF, CC and arcuate), it suggests that the interaction terms are related to the significant associations within the MDD group and not to unexpected associations within the HC group. This reinforces and increases our confidence in the specificity of our results within the MDD group.

In the third place, we investigated both CT and FA correlates of switching and clustering scores in phonemic VFT. We did not find any significant association between those scores and the CT, neither the FA. We postulate here that this could be due to a weaker association between the CT and the phonemic performances than with the semantic ones, as previously found in old-age population (Vonk et al., 2019). In addition, the implication of the pars triangularis in both semantic and phonemic performances has been described (Costafreda et al., 2006; Wagner et al., 2014). Focusing on this cortical area, we ran additional post-hoc analysis to explore its association with the semantic and phonemic quantitative scores in the HC (Supplementary Table 3). We retrieved that the CT of the left pars triangularis is associated with semantic and phonemic fluency quantitative scores (effect size: −0.39 and −0.06 respectively), suggesting that we might suffer from low power to reveal significant associations between the brain features and phonemic qualitative measures.

Additionally, to explore the specificity of the association between the qualitative VFT score and the measures of the FA and the CT, we ran additional post-hoc analysis in each group. We tested for associations...
significant results. biomarkers for MDD, because easily doable, interpretable and, if repli
test whether automated qualitative indices of VFT could be possible
of latent variables difficult to interpret. Our purpose here was rather to
markers. However, data reduction approaches render the interpretation
of MRI-derived brain features measures are greater for the automated
ted whether the TMT B-A score, which reflect cognitive switching, is
-between the standard VFT score (i.e. the quantitative score) and the CT,
switches than HC with similar clustering sizes, consistent with executive
mental Fasciculus, SLF: Superior Longitudinal Fasciculus; 95CI: 95% confidence
interval; FDR: False Discovery Rate.

<table>
<thead>
<tr>
<th>Cluster size (voxels)</th>
<th>Tracts</th>
<th>Group by FA interaction on semantic switching score (t, β [95CI])</th>
<th>FDR corrected p-values</th>
<th>Healthy control group correlation with semantic switching score (FDR corrected p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5706</td>
<td>Right SLF</td>
<td>T = 4.18, β = 210 [109–311]</td>
<td>0.0008190</td>
<td>T = -2.4, β = -134 [-252–16], p = 0.17</td>
</tr>
<tr>
<td>134</td>
<td>Right anterior body CC</td>
<td>T = 3.4, β = 123.4 [51–195]</td>
<td>0.0040305</td>
<td>T = -2.1, β = -63 [-125–0.5], p = 0.17</td>
</tr>
<tr>
<td>74</td>
<td>Right SLF</td>
<td>T = 3, β = 99 [32–165]</td>
<td>0.0095900</td>
<td>T = -1.6, β = -47 [-106–121], p = 0.2</td>
</tr>
<tr>
<td>33</td>
<td>Right corticospinal tract</td>
<td>T = 0.9, β = 31 [41–104]</td>
<td>0.4515000</td>
<td>T = 0.7, β = 24 [-45–92], p = 0.5</td>
</tr>
<tr>
<td>11</td>
<td>Right anterior corona radiata</td>
<td>T = 2.6, β = 80 [18–142]</td>
<td>0.0216125</td>
<td>T = -1.6, β = -37 [-88–14], p = 0.2</td>
</tr>
<tr>
<td>7</td>
<td>Right superior thalamic radiation</td>
<td>T = 2.3, β = 39 [5–73]</td>
<td>0.0333200</td>
<td>T = -0.5, β = -7 [-35–20], p = 0.6</td>
</tr>
<tr>
<td>3</td>
<td>Body CC</td>
<td>T = 0.2, β = 5 [56–67]</td>
<td>0.8621000</td>
<td>T = 1.7, β = 43 [-9–94], p = 0.2</td>
</tr>
</tbody>
</table>

between the standard VFT score (i.e. the quantitative score) and the CT, and the FA. In both
groups, the quantitative score was not significantly associated with the CT or the FA (p < 0.05 corrected). Moreover, we also
tested whether the TMT B-A score, which reflect cognitive switching, is
associated with CT and FA among the MDD and we could not find any significant
results.

These post-hoc exploratory analyses suggest that i) the associations of
MRI-derived brain features measures are greater for the automated
qualitative indices of semantic VFT than raw quantitative word count. ii) It also suggests that semantic VFT switching is more specific to robust
anatomic MRI measures than other measures of cognitive switching, in
MDD.

As expected, we found that MDD participants performed fewer
switches than HC with similar clustering sizes, consistent with executive
dysfunctions in MDD without impaired lexico-semantic stocks (Fossati
et al., 2003). In this sample, we found moderate to strong correlations
between number of switching, processing speed (DSST), attention (TMT-
A) and cognitive inhibition and apathy (dor semantic VFT) suggesting
shared variance, reinforcing our view that switching during VFT relies
on executive functions. Future studies with larger samples might apply
data reduction methods to capture the shared variance of cognitive,
motivational and possibly other features set and test for cerebral bio-
markers. However, data reduction approaches render the interpretation
of latent variables difficult to interpret. Our purpose here was rather to

-analyses, we also found that while accounting for abstract reasoning/
set shifting, working memory, cognitive inhibition and switching, MDD
decreases estimation of number of switches during semantic VFT by
18%, suggesting that MDD specifically affects VFT switching, above and
beyond other executive functions (Tariot, 1986).

We found reduced FA in CC, bilateral SLF, IFOF, uncinate and the
ATR, which concords with recent results in large sample of MDD, where
greater effect sizes were observed among recurrent depression, as in the
current sample (van Velzen et al., 2020). In contrast, there were no
significant differences in CT between the MDD and the HC groups. This
might be due to low power given the small sample size and small effect
sizes recently shown when comparing CT between MDD and HC (Schmaal
et al., 2017).

6. Limitations

The present study had several potential limitations. We only
included women in our protocol. This choice was guided by the sug-
gested existence of sex differences in the subtypes and course of MDD
(Kuehner, 2017). Sex-related CT (Sowell et al., 2007) and FA (van
Hemmen et al., 2016) differences have also been demonstrated, as well
as an interaction between sex and performance level in VFT (Scheu-
inger et al., 2017). Nevertheless, unique gender as well as small sample
sizes might reduce the generalizability of the findings. We did not use a
standardized task to measure language ability. However, it is unlikely
that our findings are related to broad language ability for several rea-
sons. All participants were French native speakers, muddled aged and
free of cognitive disorders. Moreover, both groups were matched on
education level (a strong predictor of language ability). Also, we used
the “animal” category of the semantic task, which reduce the impact of
education on the performance (Ardila et al. 2016). Finally, the cluster
score (which reflects the storage of the lexico-semantic memory) was
similar between the groups and not associated with any cerebral mea-
sures. Additionally, we did not control for type of treatment in our MDD
group. Although previous studies did not find any influence of antide-
pressants on executive function (Killian et al., 1984), there is some ev-
idence that drugs with an anticholinergic effect may have a cognitive
impact (Orzechowska et al., 2015). Antidepressant may also have influ-
cenced the CT measurements (Schmaal et al., 2017) and the white
matter microstructure as well (Chouiter et al., 2016). We used a SPM
technique whereas it has been suggested that left basal ganglia sustain
initiation abilities in VFT tasks (van Velzen et al., 2020). While tensor-
derived measures of white matter integrity might lack specificity, it is
considered as a reasonable measure of axonal density when tracts pre-
sumably contain single fiber population, such as long antero-posterior
tracts (SLF and the IFOF) (De Santis et al., 2014).

7. Conclusion

We elucidated specific cortical and FA signature of switching per-
formance in semantic VFT in a sample of depressed women. Those
results provide additional insight about the cognitive impairment which
underly VFT deficit in MDD. Moreover, automatic qualitative VFT scores
are associated with brain anatomy and reinforces its potential use as a
surrogate for depression cerebral bases. Given the brevity and cost
profile of the semantic VFT, its use in day-to-day clinical decision
making could be recommended.

Funding

This work was supported by the Fondation de l’Avenir (ET1-628).

CRediT authorship contribution statement

L. Domain: Data curation, Formal analysis, Writing – original draft.
M. Guillery: Methodology, Validation, Writing – review & editing. N.
L. Domain et al.

NeuroImage: Clinical 33 (2022) 102910

10

Linz: Conceptualization, Data curation, Formal analysis, Methodology, Writing – review & editing. A. König: Data curation, Formal analysis, Methodology, Writing – review & editing. J.M. Batail: Data curation, Writing – review & editing. R. David: Methodology, Writing – review & editing. I. Corouge: Data curation, Writing – review & editing. E. Bannier: Data curation, Writing – review & editing. J.C. Ferré: Data curation, Writing – review & editing. T. Dondaine: Data curation, Writing – review & editing. D. Draper: Conceptualization, Project administration, Supervision, Funding acquisition, Writing – review & editing. G.H. Robert: Conceptualization, Funding acquisition, Formal analysis, Project administration, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.nicl.2021.102910.

References


Cristofori, I., Zhong, W., Chau, A., Solomon, J., Krueger, F., Grafman, J., 2015. White and
Jenkinson, M., Beckmann, C.F., Smith, S.M. 2012. FSL.
Patterson, K., Nestor, P.J., Rogers, T.T., 2007. Where do you know what you know? The
Hua, K., Zhang, J., Wakana, S., Jiang, H., Li, X., Reich, D.S., Calabresi, P.A., Pekar, J.J.,
Smith, S., Nichols, T., 2009. Threshold-free cluster enhancement: addressing problems of
Hirshorn, E.A., Thompson-Schill, S.L., 2006. Role of the left inferior frontal gyrus in
De Santis, S., Drakesmith, M., Bells, S., Assaf, Y., Jones, D.K., 2014. Why diffusion tensor
Chouiter, L., Holmberg, J., Manuel, A.L., Colombo, F., Clarke, S., Annoni, J.-M.,
Giorgio, A., Qiu, D., Tapert, S.F., Brauer, J., Anato, M.R., Khong, P.L., James, A.C.,
Scheuringer, A., Wittig, R., Pletzer, B., 2017. Sex differences in verbal fluency: the role of
Thompson-Peatfield, S., Tariot, P.N., 1986. A psychobiologic analysis of cognitive failures: structure and
Peters, B.D., Szeszko, P.R., Radua, J., Buta, T., Gruner, P., DeRosse, P., Zhang, J.-P.,
Giorgio, A., Qiu, D., Tapert, S.F., Brauer, J., Anato, M.R., Khong, P.L., James, A.C.,
van Hemmen, J., Saris, I.M.J., Cohen-Kettenis, P.T., Veltman, D.J., Pouwels, P.J.W.,
Choutier, L., Holmberg, J., Manuel, A.L., Colombo, F., Clarke, S., Annoni, J.-M.,