

Fractal image perception provides novel insights into hierarchical cognition



M.J. Martins^{a,b,1}, F.P. Fischmeister^{c,1}, E. Puig-Waldmüller^a, J. Oh^a, A. Geißler^c, S. Robinson^d, W.T. Fitch^{a,*}, R. Beisteiner^{c,*}

^a Department of Cognitive Biology, University of Vienna, Althanstrasse 14, A-1090 Vienna, Austria

^b Language Research Laboratory, Lisbon Faculty of Medicine, Av. Professor Egas Moniz, 1649-028 Lisboa, Portugal

^c Medical University of Vienna, Department of Neurology, High-Field MR Center of Excellence, Währinger Gürtel 18-20, A-1090 Vienna, Austria

^d Medical University of Vienna, Department of Biomedical Imaging and Image-guided Therapy, High Field MR Centre, Währinger Gürtel 18-20, A-1090 Vienna, Austria

ARTICLE INFO

Article history:

Accepted 24 March 2014

Available online 31 March 2014

Keywords:

Visual processing

fMRI

Hierarchy

Recursive embedding

Parieto-medial temporal pathway

ABSTRACT

Hierarchical structures play a central role in many aspects of human cognition, prominently including both language and music. In this study we addressed hierarchy in the visual domain, using a novel paradigm based on fractal images. Fractals are self-similar patterns generated by repeating the same simple rule at multiple hierarchical levels. Our hypothesis was that the brain uses different resources for processing hierarchies depending on whether it applies a “fractal” or a “non-fractal” cognitive strategy. We analyzed the neural circuits activated by these complex hierarchical patterns in an event-related fMRI study of 40 healthy subjects.

Brain activation was compared across three different tasks: a similarity task, and two hierarchical tasks in which subjects were asked to recognize the repetition of a rule operating transformations either within an existing hierarchical level, or generating new hierarchical levels. Similar hierarchical images were generated by both rules and target images were identical.

We found that when processing visual hierarchies, engagement in both hierarchical tasks activated the visual dorsal stream (occipito-parietal cortex, intraparietal sulcus and dorsolateral prefrontal cortex). In addition, the level-generating task specifically activated circuits related to the integration of spatial and categorical information, and with the integration of items in contexts (posterior cingulate cortex, retrosplenial cortex, and medial, ventral and anterior regions of temporal cortex). These findings provide interesting new clues about the cognitive mechanisms involved in the generation of new hierarchical levels as required for fractals.

© 2014 Elsevier Inc. All rights reserved.

Introduction

The ability to represent and generate complex hierarchical structures is one of the hallmarks of human cognition. In many domains, including language, music, problem solving, action-sequencing and spatial navigation, humans organize basic elements into higher-order groupings and structures (Badre, 2008; Chomsky, 1957; Hauser et al., 2002; Nardini et al., 2008; Unterrainer and Owen, 2006; Wohlschlagler et al., 2003). This ability to encode the relationship between items (words, people, etc.) and the broader structures in which these items are embedded (sentences, corporations, etc) affords flexibility to human behavior. For example, in action sequencing, humans are able to change, add or adapt certain basic movements to particular contexts, while keeping the overall structure (and goals) of canonical motor procedures intact (Wohlschlagler et al., 2003). Typical examples of these

actions-in-context are ‘grinding the beans’ or ‘re-filling the water container’ in the process of making coffee (Jackendoff, 2002). Individuals can evaluate the need for these actions and omit them if they are unnecessary without impairing the overall procedure of making coffee (Badre and D’Esposito, 2009). This ability is different from simple action sequencing, and seems very limited in non-human animals (Conway and Christiansen, 2001).

A promising method to represent complex hierarchical structures – realized in nature and attractive for experimental research – is the use of recursive embedding processes (Fitch, 2010; Martins, 2012). Recursive embedding refers to the incorporation of a structure inside another structure of the same sort, and it allows the generation of hierarchies with infinite depth using very simple rules. We can add several new elements to a certain hierarchical level using within-level transformation rules (Fig. 1A), but it is only possible to generate multiple hierarchical levels with a single rule if this rule involves recursive embedding (Fig. 1B). When used in association with other rule-based processes, recursive embedding allows the generation of hierarchies that are deep, structurally rich and perceived as attractive. Some examples are the fractal Mandelbrot images or fractal structures in nature such as

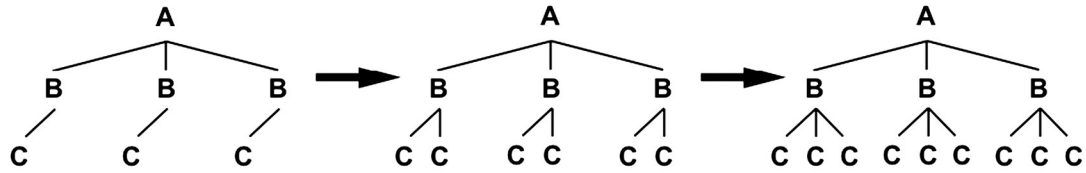
* Corresponding authors.

E-mail addresses: tecumseh.fitch@univie.ac.at (W.T. Fitch),

roland.beisteiner@meduniwien.ac.at (R. Beisteiner).

¹ These authors contributed equally to this work.

A) Within-level addition rule: Add another C to existing level under each B.



B) Cross-level recursive rule: Add three As to new level under each A.

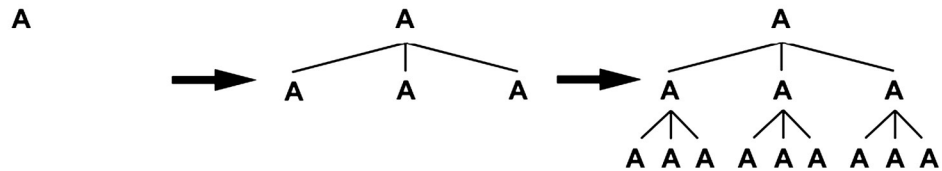


Fig. 1. Examples of processes that add elements to hierarchies. These processes can either generate new hierarchical levels (B) or simply add elements to pre-existing levels (A). Only recursive embedding (B) can add multiple hierarchical levels using a single rule.

tree branches, algae, the flower of the *Brassica oleracea*, snail shells and coastlines. These structures can be extended or sub-divided indefinitely while visual and structural similarity is retained at all scales. These kinds of structures contrast with others with simpler modes of organization such as grass or crop fields, which like bead necklaces, are formed by adding several items to a group at fixed hierarchical levels.

Here we investigate the ability to recognize well-formed visuo-spatial hierarchical structures, based on the application of rules that either operate transformations within a hierarchical level, or rules which generate new self-similar hierarchical levels (Fig. 1). For simplicity, we simply use the expression ‘recursive’ or ‘recursion’ to refer to ‘recursive embedding’.

The processing of visuo-spatial stimuli is often described as occurring in parallel in two different systems – the ventral stream and the dorsal stream (de Haan and Cowey, 2011; Kravitz et al., 2011). The ventral stream, an occipito-temporal network, seems to process object quality or semantic information, with more abstract categories represented in more anterior portions of the temporal lobe (Kravitz et al., 2013). The dorsal stream, an occipito-parietal network, has classically been described as processing spatial information only. Recently, however, this classical view of the dorsal stream has been updated (Kravitz et al., 2011). While projections from the parietal cortex to the prefrontal cortex seem to be important for spatial working memory and visually-guided action, a third system, called the parietal-medial temporal pathway (PMT) appears to be necessary to integrate spatial and semantic information (Kravitz et al., 2011). The PMT pathway connects the dorsal stream with the medial temporal cortex (hippocampus and parahippocampus), through the posterior cingulate (PCC) and retrosplenial cortices (RSC) (Kravitz et al., 2011; Margulies et al., 2009). This pathway appears to be crucial for the retrieval of landmark information during spatial navigation and for the integration of objects in contextual frames (e.g. a mug in a date in a coffee shop) (Aguirre and D’Esposito, 1999; Buzsáki and Moser, 2013; Ino et al., 2007; Ranganath and Ritchey, 2012; Sato et al., 2010) (Fig. 2). We therefore hypothesize that the PMT may play a specific role in the representation of principles that allow the recognition and generation of well-formed hierarchical embeddings in the visuo-spatial domain.

Based on the principles depicted in Fig. 1, we developed two tasks to investigate the cognitive processes involved in the representation of visuo-spatial hierarchies: The Visual Recursion Task (VRT) and the Embedded Iteration Task (EIT). In both tasks participants are exposed to generative processes for a certain number of iterative steps and

then asked to make inferences about further iterations. This means that in both tasks participants are asked to extract simple rules from the first iterations which can then be applied to predict further transformations. In VRT, each iterative step generates a new hierarchical level according to one particular spatial rule isomorphic to the rule displayed in previous levels of the hierarchy. The brain requires only one simple rule to be able to generate large self-similar structures (fractals) with an unlimited number of levels. In EIT, new elements are embedded iteratively within a fixed hierarchical level, according to a spatial rule but without generating new levels. It is important to clarify that both tasks are iterative (i.e. a certain rule is applied a given number of times) and both may generate hierarchies of similar complexity (see Figs. 1 and 3).

Our previous research with these tasks (Martins, 2012; Martins and Fitch, 2012) suggests that, in comparison with EIT, performance in VRT is more strongly associated with abstract reasoning and less correlated with specific visuo-spatial cognitive abilities. In the current study, we investigated the neural bases involved in the representation of visuo-spatial hierarchies by comparing the brain circuits active during VRT and EIT. As a control task we introduced a ‘similarity task’ (Positional Similarity Visual Task – PSVT), in which participants were asked to match a target visuo-spatial hierarchy with a set of alternatives. The setup and images displayed were closely matched for all three tasks. As indicated above, our primary hypothesis was that the brain uses different resources for processing identical hierarchical structures depending on whether it applies a “fractal” or a “non-fractal” cognitive strategy.

Material and methods

Participants

40 healthy participants (19 males and 21 females, age range 20–32) took part in the study. All had normal or corrected-to-normal vision, no history of neurological or psychiatric disease, and no current use of psychoactive medications. All completed a short questionnaire screening for previous clinical history and a battery of cognitive tests. Participants, who were all right-handed native German speakers and mostly university students, were recruited online, and gave informed written consent prior to participation in the study, which was approved by the local ethics committee. Before the functional Magnetic Resonance Imaging (fMRI) session, each participant was explicitly debriefed about both hierarchy-generating rules and practiced one or two blocks of the

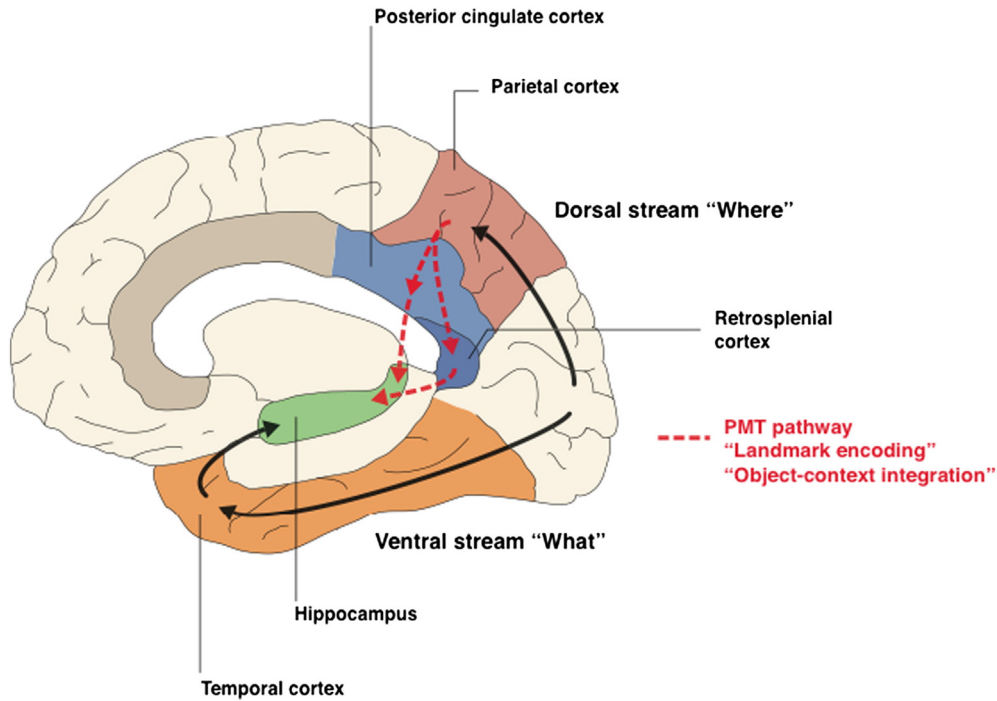


Fig. 2. Neural pathways involved in visuo-spatial processing. The dorsal stream, which includes the parietal cortex and its projections to the frontal cortex, is involved in the processing of spatial information. The ventral stream, which includes the inferior and lateral temporal cortex and their projections to the medial temporal cortex, is involved in the processing of categorical or semantic information. The parieto-medial temporal (PMT) integrates information from both pathways and is involved in the encoding of landmarks in spatial navigation and in the integration of objects into contextual frames. We hypothesize that the generation of hierarchical levels using recursive processes will recruit the PMT pathway.

experimental task (with stimuli which were different from those used in fMRI) after which they received feedback. Participants were paid 30 Euros for their participation. The overall procedure comprised 1 h of practice plus cognitive testing and approximately one and a half hours of fMRI scanning.

Task

Modified VRT and EIT tasks, described in Martins (2012) and Martins and Fitch (2012), were used. While EIT requires the representation of iterative processes occurring within a hierarchical level, VRT requires the representation of iterative processes generating new hierarchical levels (Fig. 1). For this study, we devised an additional Positional Similarity Visual Task (PSVT) to investigate the effects of observing visual fractals without rule-based reasoning. In the latter, participants attended to a set of three random images and were then asked to choose which of two new items was *identical* to one of the previous three (Fig. 3).

Participants performed 4 sessions inside a 3 Tesla MRI scanner. Each session included 14 VRT stimuli, 14 EIT stimuli and 8 PSVT stimuli (Fig. 3). We used an event-related design in which stimuli from different task categories were randomized within the same session.

Each trial comprised two main phases (Fig. 4) – the rule acquisition phase, and the rule application phase. Before the rule acquisition phase, at the beginning of each trial, a white letter was presented on a black background in the center of the screen for a duration which ranged between 1000 and 1750 ms. This letter indicated the task of the trial: “R” for VRT, “I” for EIT, and “S” for PSVT. Then, in the rule acquisition phase, three images, corresponding to the first three iterations of either a within-level or recursive process were presented simultaneously in the top half of the screen. In the case of the similarity task these were three different images selected quasi-randomly from the large pool of fractal images. This phase had a fixed duration of 3 s. Between the rule acquisition and rule application phases, a white crosshair was presented in the center of the screen for a duration which ranged between 1000

and 3000 ms. Finally, in the rule application phase, two additional images were presented in the bottom half of the screen, simultaneously and side-by-side. One of these corresponded to the correct fourth iteration of the previous iterative process and the other was a foil. In the case of the similarity task, the correct image was identical to one of the previously presented three images (Fig. 3). In this rule application phase participants were asked to choose the image they considered correct by pressing either the left or right button with the thumb of the left or right hand. No visual or auditory feedback was provided. The maximum duration of this phase was 6 s. The inter-trial interval (ITI) ranged from 500 ms to 14,000 ms and during this period participants were exposed to a black screen. The position of the correct and foil images (LEFT or RIGHT) was random and counterbalanced. To control for luminance effects, all stimuli had the same number of black and white pixels, both globally and for each trial phase. For more details on the generation of the stimuli, see the Supplementary methods.

One week before the fMRI session, participants had a first experimental session where they were instructed about the hierarchical rules involved in VRT and EIT. They were shown examples of sequences of images depicting the generation of hierarchies. In VRT they were told that at each step new elements were added to new hierarchical levels according to a spatial rule that was constant across levels; in EIT they were told that new elements were added to an existing hierarchical level according to a predictable spatial rule. Then they performed a training session using a sequence which was identical in the representation of item types to that later applied inside the scanner (14 VRT items, 14 EIT items and 8 PSVT items) but which used different stimuli.

Data acquisition

Data acquisition was performed with a 3 Tesla TIM Trio system (Siemens, Erlangen, Germany) using a 32-channel Siemens head coil. Functional magnetic resonance images (fMRI) were acquired using an optimized 2D single-shot echo planar imaging (EPI) sequence which

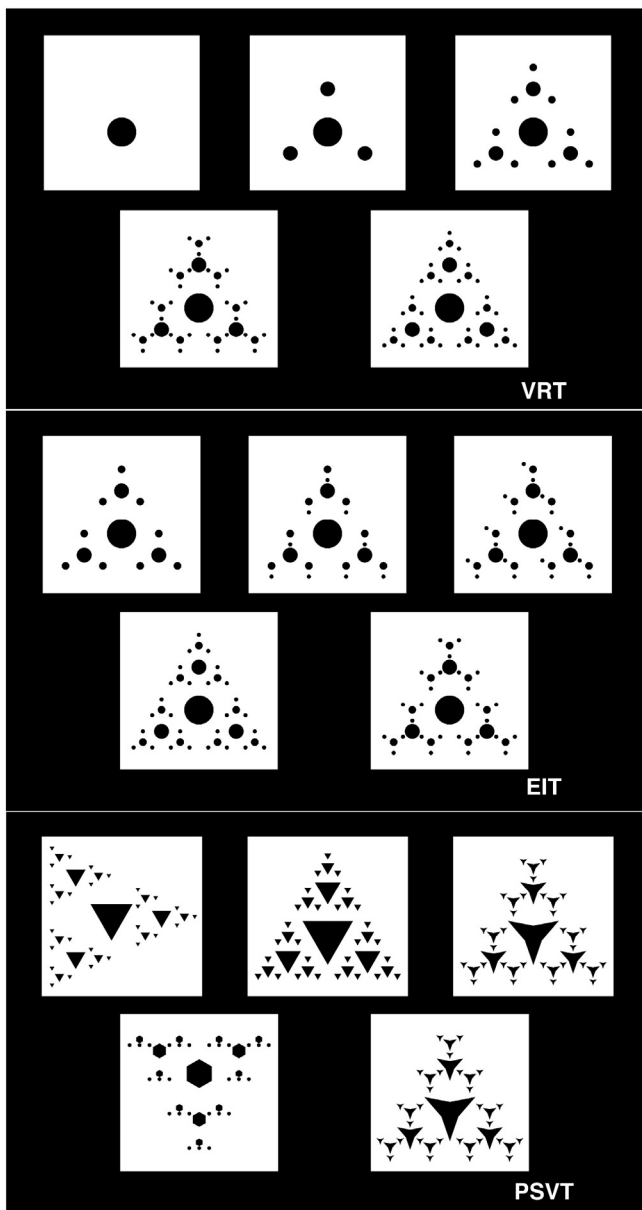


Fig. 3. Examples of VRT (Visual Recursion Task) stimuli (top), EIT (Embedded Iteration Task) stimuli (middle) and PSVT (Positional Similarity Visual Task) stimuli (bottom). In the “rule acquisition” phase (see [Material and methods](#) for details), the first three iterations of a process were presented in the top half of the screen. Afterwards, in the “rule application” phase, two images were presented in the bottom half, from which participants were asked to choose the one corresponding to the fourth iteration of the same process. In the case of the similarity task, the images in the top half were randomly chosen from a pool of fractals and participants were asked to choose which of the lower images was *identical* to one of the images in the top row. The right bottom image is *CORRECT* and the left image is *INCORRECT* in the examples in this figure. Note that our fMRI data were recorded during the processing of identical target stimuli (bottom half of VRT, EIT stimuli). Crucially, the same image can be correct or incorrect depending on the rule used to generate the fractal.

included online EPI distortion correction with PSF mapping (Zaitsev et al., 2004). 350 EPI volumes per session were acquired with a square FOV of 220 mm, an in-plane matrix size of 128×128 , with 36 slices of 2.7 mm thickness and 20% gap (i.e. $2.3 \text{ mm} \times 2.3 \text{ mm} \times 2.7 \text{ mm}$ voxel size) aligned parallel to the AC-PC plane, a repetition time (TR) of 2000 ms, echo time (TE) 32 ms, and a flip angle of 73° . For anatomical registration, high-resolution T1-weighted MR images were acquired using a 3D MPRAGE sequence (TE = 3.02 ms, TR = 2190 ms, inversion time [TI] = 1300 ms) with a matrix size of $250 \times 250 \times 256$, with

isometric voxels with a nominal side length of 0.9 mm, flip angle of 9° and GRAPPA acceleration factor 2.

Data preprocessing

Image preprocessing and statistical analysis at individual and group levels were performed using SPM 8 (<http://www.fil.ion.ucl.ac.uk/spm/>). Data were first slice-time and then motion corrected. These corrected data were then spatially normalized using New Segment (SPM manual, FIL Group) and finally smoothed using a 5 mm full-width-at-half-maximum Gaussian filter. For single-subject analyses, evoked hemodynamic responses for the different event types were modeled within a general linear model using delta functions corresponding to the stimulus presentation length convolved with a canonical hemodynamic response function. This way, the model captured differences in reaction time. To this design matrix we added 24 nuisance regressors of no interest, corresponding to the motion realignment parameters and their Volterra expansion (Lund et al., 2005), to regress out residual motion artifacts. In addition, a 356 s cutoff high-pass filter was applied to account for low-frequency drifts and signal fluctuations. Block regressors were used to correct for session-related mean and scaling effects (added as confounds). Responses corresponding to the rule application phase of the three stimuli types were then summarized across the four sessions and entered into a second-level GLM.

Statistical analysis

On the group level, a repeated-measures GLM with partitioned error variances (rm-GLMflex: between- and within-subject error terms are modeled separately) was used to model activity during the application phase. This one-way rm-GLMflex (with each task’s application phase being one level) allowed us to identify hemodynamic responses solely related to the tasks of interest by constructing planned contrasts to answer the different research questions within one model. Rule-based related activation was obtained by contrasting VRT and EIT with the control condition task (PSVT). The differences between recursion-related processes and those resulting from embedded iteration were assessed by directly comparing VRT and EIT (implicitly a comparison of VRT-PSVT versus EIT-PSVT) within the rm-GLMflex model. Additionally, to test for regions commonly activated during the application phase in the VRT and the EIT, a conjunction analysis across the contrasts VRT-PSVT and EIT-PSVT using the ‘conjunction null’ hypothesis was performed (Nichols et al., 2005). All comparisons were masked with the main effect of the one-way rm-GLMflex and subsequently thresholded at a voxel-wise FDR-adjusted $p < 0.05$ with a 10-voxel extent threshold.

In order to test for possible sequence effects showing learning or carry-over effects from one session to the next, a 3×4 rm-GLMflex model similar to the first one with the factor task (VRT, EIT and PSVT) and the four sessions was estimated. A comparable approach was made to test for possible gender effects by introducing a between-subjects factor ‘gender’. No significant main effects nor interaction effects were found for ‘sequence’ and ‘gender’, even adopting a more lenient uncorrected threshold of $p < 0.01$.

Neuro-cognitive battery

We applied a brief neuropsychological battery to screen our participants for possible cognitive impairments. This battery included computerized versions of Digit Span backwards (DSPAN, a verbal working memory task), Corsi block tapping backwards (CORSI, a spatial working memory task), Tower of Hanoi (ToH, a recursive planning in action sequencing task) (Mueller, 2011) (retrieved from <http://pebl.sf.net/battery.html>) and a paper-and-pencil version of the progressive matrices of RAVEN (a test of non-verbal intelligence). We recorded the maximum number of elements correctly reproduced in DPSAN and CORSI,

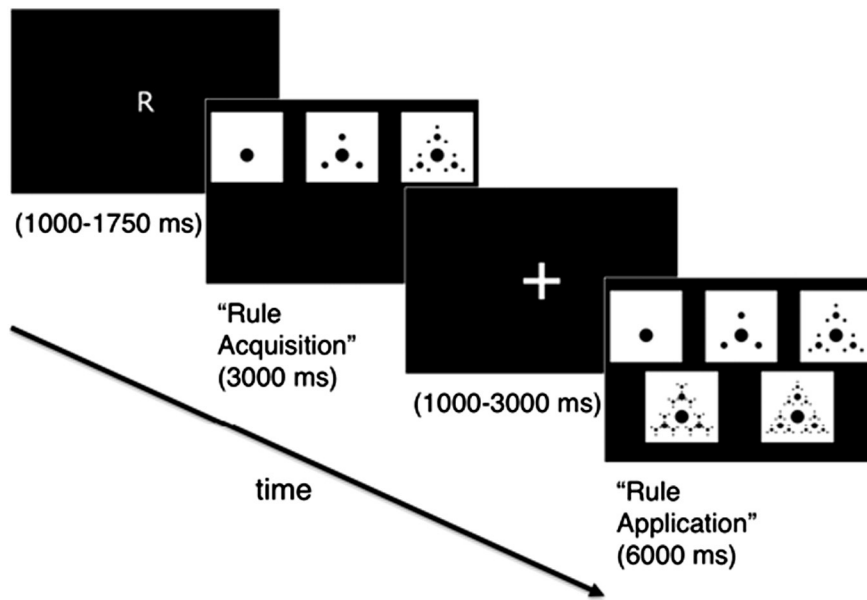


Fig. 4. Trial structure: At the beginning of each trial a letter was displayed indicating the stimulus category to be presented ('R' for VRT, 'T' for EIT and 'S' for PSVT). fMRI data were acquired in the 'Rule Application' phase.

the maximum length (viz. number of steps) of ToH problems that participants were able to complete without errors, and the number of correct answers in RAVEN.

Results

Behavioral results

All 40 participants performed well within the scanner, and reported no problems in solving the tasks. Behavioral data collected during the fMRI runs showed a high rate of correct responses in VRT ($M = 96\%$, $SD = 8\%$), EIT ($M = 91\%$, $SD = 5\%$) and PSVT ($M = 95\%$, $SD = 8\%$). The percentage of correct answers differed between tasks (repeated-measures ANOVA: $F_{1,39} = 7.1$, $p = 0.011$): participants scored lower in EIT than in VRT ($p < 0.01$) and PSVT ($p = 0.03$). Mean response time was 2.34 s in VRT, 2.56 s in EIT, and 2.59 s in PSVT. There was a significant main effect of task in response time (repeated-measures ANOVA: $F_{1,39} = 27.4$, $p < 0.001$): participants responded faster in VRT than in EIT ($p < 0.001$) and PSVT ($p = 0.012$).

In order to prevent participants from using simple heuristic strategies we included different foil categories ("ODD foil" and "POSITIONAL foil") in both VRT and EIT (see Supplementary methods). Participants

performed adequately ($>90\%$) in all foil categories (see Supplementary Table S1).

During a pre-testing session, participants were screened with a neuro-cognitive battery. All participants performed adequately in at least three out of four of these tests (Supplementary Table S2). In the pre-testing training session, participants performed the EIT, VRT and PSVT. Mean scores in the training session were as follows: VRT ($M = 83\%$, $SD = 2\%$), EIT ($M = 81\%$, $SD = 2\%$) and PSVT ($M = 80\%$, $SD = 28\%$). No significant difference was found between tasks during training (repeated-measures ANOVA: $F_{1,35} = 0.2$, $p = 0.6$). Previous research suggests that once learnt, "fractal" rules lead to more accurate judgments about hierarchies than "non-fractal" rules (Martins et al., in review). In the data presented here, a power curve fits VRT data better ($R^2 = 0.33$) than EIT data ($R^2 = 0.15$), suggesting that the learning effect is stronger in VRT. This explains why behavioral VRT–EIT differences were absent in pre-testing.

Rule-based iterative processes (within and across hierarchical levels) versus similarity assessment

While VRT and EIT both involve rule-based iterative processes, PSVT involves a simple similarity assessment between images.

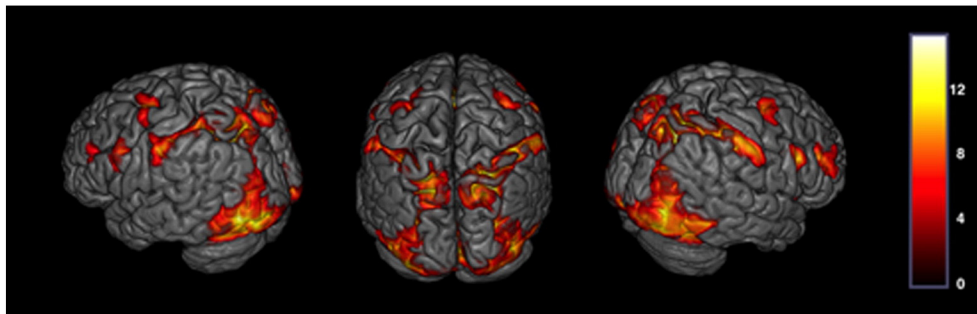


Fig. 5. Brain activations specific to both Visual Recursion and Embedded Iteration Tasks (VRT and EIT), in comparison with a simple similarity task. Both recursive and within-level iterative tasks showed activations within the visual 'dorsal stream', a bilateral network including regions from the occipital cortex along the intraparietal sulcus (also including superior and inferior parietal cortex regions), and extending to areas within the pre-motor and prefrontal cortex (Supplementary Table S3). 'Ventral' activations were also found within the inferior occipital cortex and right hippocampus, and 'midline' activations along the anterior and medial cingulate cortices. Activations within Broca's area (peak $(x, y, z) = -50, 5, 28$, $t = 7.36$) and its right hemispheric homologue were found for both VRT and EIT. Results are presented at $p < 0.05$ with FDR correction.

To investigate whether there were brain activations specific for rule-based iterative tasks, we performed a conjunction analysis relative to PSVT. We found significant activations ($p < 0.05$ with FDR correction) in a network of areas including the visual ‘dorsal stream’, prefrontal and pre-motor cortices, and ‘midline structures’ (Fig. 5, Supplementary Table S3). This network comprised: 1) A large cluster extending from left inferior and right middle occipital gyri to the intraparietal sulcus (hIP1/hIP3) and superior parietal cortex (BA 7A). This large cluster also included portions of the right inferior parietal cortex, cerebellum, thalamus, and the right hippocampus; 2) Regions within the inferior and middle frontal gyrus, bilaterally, including portions of BA 6, BA 44 and BA 45. Notably activations within Broca’s area (peak $(x, y, z) = (-50, 5, 28)$, $t = 7.36$) and its right hemispheric homologue were found for both VRT and EIT; 3) A number of pre-motor areas along the pre-central gyrus (BA 6 and 44), supplementary motor area (SMA) bilaterally, and right superior frontal gyrus (BA 6); 4) Finally, we found activations within the anterior and middle cingulate cortex, and bilateral insula.

Visuo-spatial hierarchy differences: within-level transformations versus recursion

To assess whether the processing required for VRT and EIT dissociated at the neural level, we performed contrasts between these two tasks. Compared with the application of within-level rules in EIT, the application of cross-level rules in VRT yielded larger hemodynamic responses in an extensive bilateral network of brain areas associated with the visual ‘ventral stream’, the parietal–medial temporal pathway (PMT), the medial temporal lobe and the rostro-medial prefrontal cortex (Fig. 6, Supplementary Table S4). This network included 1) lateral and ventral occipito-temporal regions (from middle and superior occipital gyrus to lingual and fusiform gyri); 2) medial temporal lobe (including hippocampus and parahippocampus); 3) middle and superior temporal gyri; 4) left superior frontal gyrus (BA 9); 5) peri-rolandic areas (post-central gyrus bilaterally and right rolandic operculum); and 6) a number of midline structures including the calcarine sulcus, cuneus, precuneus, anterior, middle and posterior cingulate cortices, retrosplenial cortex (BA 29), left superior medial frontal cortex (BA 10) and left middle

orbital gyrus (BA 10). Some portions of thalamus and cerebellum were active, bilaterally.

Conversely, compared with VRT, EIT yielded greater hemodynamic responses in a bilateral network comprising fronto-parietal regions (the ‘dorsal stream’ and inferior frontal gyrus) and basal ganglia (Fig. 6, Supplementary Table S5). This network included: 1) bilateral inferior parietal cortex (including PF and PG areas); 2) right superior parietal cortex (BA7), with bilateral extensions to the dorsal portions of precuneus; 3) right superior frontal gyrus (BA6) with bilateral extensions to the medial portion of BA6 (including left SMA); 4) middle frontal gyrus (including portions of right BA44/45 and left BA6); 5) bilateral inferior frontal gyrus (BA44/45); and 6) right insula. Furthermore, there were significant activations in the basal ganglia, including bilateral caudate and left pallidum, and small foci of activations in the right middle temporal gyrus, right middle occipital gyrus, and cerebellum.

Discussion

In this study we contrasted the brain networks active during the representation of processes allowing the generation of new hierarchical levels (as required for generating fractals) with the representation of processes that may generate structures of equal complexity but do this without creating new levels. The rationale was that many attractive structures in nature are fractals and, based on our previous research, we hypothesized that these are processed in a specific and very efficient way with a ‘fractal’ cognitive strategy. Both our tasks (VRT and EIT) are innovative in that they assess the ability to form representations, using previously existing hierarchical information, which allow the discrimination of new predictable hierarchical transformations.

Our main findings were the following: 1) Both rule-based processes (within and between levels) activated a bilateral network (the dorsal stream) which includes visual association areas and fronto-parietal circuits associated with spatial reasoning (Kravitz et al., 2011). Additionally, both rule-based tasks activate the inferior frontal gyrus (IFG, including parts of Broca’s area), insula, cingulate cortex and right hippocampus; 2) Compared to within-level transformations, the representation of recursive processes generating new hierarchical levels (i.e. fractals) recruited regions within the parieto-medial temporal pathway

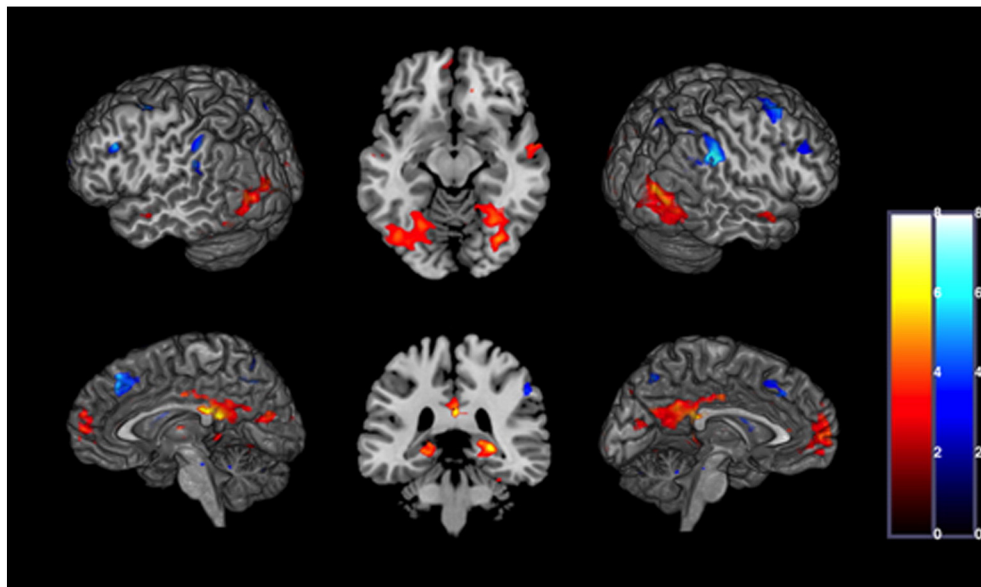


Fig. 6. Brain activation contrast between Visual Recursion Task (VRT) and Embedded Iteration Task (EIT). For VRT (shown in red and summarized in Supplementary Table S4) larger responses were found in regions related with (i) the visual ‘ventral stream’ (including fusiform, lingual, and middle temporal gyri, bilaterally); (ii) the parieto-medial temporal (PMT) pathway (including posterior cingulate and retrosplenial cortices); (iii) PMT projections to the medial temporal lobe (hippocampus and parahippocampus); and (iv) anterior portions of prefrontal cortex (BA10). For EIT, (shown in blue, Supplementary Table S5), larger responses were found in regions comprising the visual ‘dorsal stream’ (including superior and inferior parietal cortex), and these areas’ projections to the pre-frontal cortex (including areas BA 44 and 45) and pre-motor cortex (including BA 6 and supplementary motor area), bilaterally. Results are presented at $p < 0.05$ with FDR correction.

(PMT; Fig. 2) – including the posterior cingulate cortex (PCC) and retrosplenial cortex (RSC) – and their projections to the medial temporal cortex (MTL), which have been associated with the integration of spatial and semantic information (Kravitz et al., 2011). We also found activations in the anterior portions of superior and middle temporal gyri (STG and MTG, respectively); 3) In contrast, within-level iterative rules activated the following regions more strongly; the dorsal stream, the dorsal fronto-parietal network (FPN), IFG, and basal ganglia. We now elaborate on these three basic findings.

Iterative processes generating hierarchies activate the dorsal stream and IFG

Compared to simple assessment of visual similarity, the cognitive processes involved in the representation of iterative rules correlate with greater activation of visual association areas, including bilateral activations in the intra-parietal sulcus (extending to portions of superior and inferior parietal cortex). These areas comprise the so-called ‘dorsal stream’, and are involved in the processing of information relating to the location of objects in visual-spatial structures (‘where’ information) (Kravitz et al., 2011). Furthermore, we found activations in the supplementary motor area (SMA), pre-motor cortex (PMC), and prefrontal cortex (PFC). These areas have been described as projections of the dorsal stream and have been implicated in the control of eye movements, spatial working memory and executive control of visual-spatial processing (Kravitz et al., 2011). We also found activations in the insula and anterior/middle cingulate gyrus, often described as part of a ‘salience network’ (Sridharan et al., 2008). This network allows switching between external and internal modes of representation (correlated, respectively, with the activation of central executive and default-mode networks) and plays a crucial role in maintenance and update of predictions and expectations (Sridharan et al., 2008).

Finally, conjunction analysis revealed activations within Broca’s area (and its right hemisphere homologue) for both within-level and cross-level transformations. Broca’s area has been shown to be active in the processing of sequential hierarchies in natural language (Friederici et al., 2006; Makuuchi et al., 2009), artificial grammars with sound sequences (Bahlmann et al., 2008; Petersson et al., 2012), artificial grammars with image sequences (Bahlmann et al., 2009), music patterns (Maess et al., 2001; Patel, 2003) and during the processing of action sequences (Fazio et al., 2009). However the precise role of this structure is still uncertain, hypotheses ranging from it supporting sequential working memory, to participating in hierarchical or structure unification (Baddeley, 2003; Berwick et al., 2013; Braver et al., 1997; de Vries et al., 2011; Fadiga et al., 2009; Forkstam et al., 2006; Hagoort, 2005; Koechlin and Jubault, 2006; Petersson, 2004; Petersson and Hagoort, 2012; Uddén et al., 2008). Our results support the hypothesis that Broca’s area may be generally involved in maintaining online information or rules supporting iterative/sequential processes (de Vries et al., 2011; Fazio et al., 2009; Rajah et al., 2008), rather than in the integration of multiple hierarchical-levels *per se* (Berwick et al., 2013).

Representation of self-similar hierarchies (fractals) requires integration of spatial and categorical information

In addition to requiring the participation of the ‘dorsal stream’, the representation of processes generating new hierarchical levels recruited a bilateral network involving the ventral occipito-temporal cortex, including fusiform and lingual gyri – all parts of the visual ‘ventral stream’ (Kravitz et al., 2013). This network has been associated with the representation of categorical or semantic information. Furthermore, these rules recruited the anterior regions of STG and MTG, which appear to correlate with the retrieval of abstract categories (Lehmann et al., 2010; Wang et al., 2010). Interestingly, VRT also specifically activated areas within the PMT pathway (RSC and PCC), which have been described as intermediary projections of the dorsal stream to the MTL.

These areas are involved in the integration of objects in contextual frames (Kravitz et al., 2011) and in the integration of spatial and categorical/semantic information. Lesions in these areas are associated with spatial navigation deficits, in particular with an inability to use spatial landmarks (despite an intact ability to retrieve landmark location) (Aguirre and D’Esposito, 1999; Ino et al., 2007).

The generation of novel self-similar hierarchical levels also bilaterally activates the PMT projections into the MTL (hippocampus and parahippocampal cortex). These areas have been associated with episodic memory and with the formation of unified representations of items and contexts (Eichenbaum et al., 2012; Ranganath, 2010). The recruitment of the MTL has previously been reported as being crucial in the processing of spatial and social hierarchies (Aminoff et al., 2006; Kravitz et al., 2011; Kumaran et al., 2012; Zaretskaya et al., 2013), and in studies investigating the processing of novel (vs. well-trained) linguistic hierarchies (Opitz and Friederici, 2007).

Taken together, these results suggest that episodic memory and the integration of items in contexts are crucial mechanisms in the processing of rule-based generation of novel hierarchical levels using recursive principles. Furthermore, this process requires the integration of spatial and categorical information. This finding is particularly intriguing since the visuo-spatial hierarchies employed in this study do not convey “semantic” information *per se*. We hypothesize that the representation of hierarchical dependencies may require the retrieval of “semantic” information of a rather abstract sort. In order to utilize a spatial landmark one needs both to know its location (where), and to know what it is a landmark of (what) – a type of referential relationship. Processing this abstract relationship between reference and referent may require the activation of traditional ‘semantic’ networks which would therefore be necessary for the integration of multiple hierarchical levels. Consistent with this supposition, in other domains, such as language, the processing of hierarchies is also associated with the activation of areas related with semantic retrieval (e.g. STG) (Friederici et al., 2006).

Finally, our behavioral results suggest a specific correlation between VRT (but not EIT) and Tower of Hanoi, which requires hierarchical planning of actions and invites a recursive solution (Goel and Grafman, 1995) (Supplementary results). Crucially, we used a score of Tower of Hanoi (longest sequence performed without mistakes) that cannot easily be explained by simple iterative mechanisms. This behavioral correlation lends support to the hypothesis that our Visual Recursion Task may tap into cognitive resources associated with the processing of recursion.

Within-level transformations are more specifically spatial

Compared with VRT, the representation of iterative processes transforming hierarchies within a fixed level correlated more strongly with the activation of areas in the visual dorsal stream (Kravitz et al., 2011). This suggests, in agreement with previous research (Martins, 2012; Martins and Fitch, 2012), that these within-level transformations may rely on specific spatial resources (‘where’ information), to a greater extent than recursive transformations. Our behavioral results (see Supplementary results) confirm that both the acquisition and application of within level rules correlate more strongly with working memory abilities than do rules generating novel hierarchical levels. Interestingly, small foci within Broca’s area and its right homologue seemed to be more active in within-level transformations than in recursive transformations. These findings suggest that Broca’s area is not specifically active for the processing of cross-level hierarchical integration (Bahlmann et al., 2008, 2009; Berwick et al., 2013; Fadiga et al., 2009; Friederici et al., 2011), but may be more generally involved in the storage and maintenance of rule-based iterative information, or in working memory processes (Baddeley, 2003; Bengtsson et al., 2008; de Vries et al., 2011; Fazio et al., 2009; Rajah et al., 2008). These findings also suggest that recursive embedding is a more memory-efficient method to generate complex hierarchies.

Limitations of the current study

It could be argued that participants may have used simple heuristic strategies, comparing items according to their similarity, to solve our tasks. We tried to minimize this problem in three ways: 1) All neuroimaging analyses and comparisons were implicitly performed against a 'similarity task'; 2) We included different 'foil item categories' to block any specific heuristic strategies; and 3) We explicitly instructed and trained participants in the usage of within-level and recursive rules while solving EIT and VRT. Furthermore, VRT performance both inside and outside of the scanner was specifically correlated with Tower of Hanoi, which is considered a recursive planning task (Goel and Grafman, 1995) and processing of VRT stimuli was more efficient than processing of EIT stimuli, despite both tasks using identical targets. Taken together, this suggests that our experiment design and analysis tapped into the representation of recursive principles rather than the application of simple heuristic strategies.

Conclusion

In the visual-spatial domain, the brain uses different resources when processing identical images with a "fractal" or a "non-fractal" cognitive strategy. The representation of recursive principles allowing the generation of new hierarchical levels appears to recruit resources associated with the integration of spatial and abstract semantic information, and with the integration of items in contexts. Rather than being tightly localized, this mechanism is implemented in a widely distributed brain network, including regions associated with specific visual-spatial processes and also regions subserving domain-general functions. Although Broca's area might be important for the processing of iterative and hierarchical information, it did not play a specific role in the representation of recursive embedding principles. Future research contrasting different domains (music, language etc.) will be required to determine whether localized, domain-specific computational processes are required for the generation of hierarchies. The methods presented here, based on the properties of fractal geometry, provide novel tools to investigate the ability to represent hierarchies of unbounded depth.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.neuroimage.2014.03.064>.

Acknowledgments

This work was supported by a research cluster grant "Shared Neural Resources for Music and Language" to WTF and RB (University of Vienna - internal project number FG761002 - and Medical University of Vienna - internal project number FA103FC003), an FCT grant SFRH/BD/64206/2009 to MM, and by ERC Advanced Grant SOMACCA, project number 230604 to WTF. The authors declare no competing financial interests.

References

- Aguirre, G., D'Esposito, M., 1999. Topographical disorientation: a synthesis and taxonomy. *Brain* 122, 1613–1628.
- Aminoff, E., Gronau, N., Bar, M., 2006. The parahippocampal cortex mediates spatial and nonspatial associations. *Cereb. Cortex* 17, 1493–1503. <http://dx.doi.org/10.1093/cercor/bhl078>.
- Baddeley, A., 2003. Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4, 829–839. <http://dx.doi.org/10.1038/nrn1201>.
- Badre, D., 2008. Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends Cogn. Sci.* 12, 193–200. <http://dx.doi.org/10.1016/j.tics.2008.02.004>.
- Badre, D., D'Esposito, M., 2009. Is the rostro-caudal axis of the frontal lobe hierarchical? *Nat. Rev. Neurosci.* 10, 659–669. <http://dx.doi.org/10.1038/nrn2667>.
- Bahlmann, J., Schubotz, R., Friederici, A., 2008. Hierarchical artificial grammar processing engages Broca's area. *NeuroImage* 42, 525–534. <http://dx.doi.org/10.1016/j.neuroimage.2008.04.249>.
- Bahlmann, J., Schubotz, R.I., Mueller, J.L., Koester, D., Friederici, A.D., 2009. Neural circuits of hierarchical visuo-spatial sequence processing. *Brain Res.* 1298, 161–170. <http://dx.doi.org/10.1016/j.brainres.2009.08.017>.
- Bengtsson, S.L., Haynes, J.D., Sakai, K., Buckley, M.J., Passingham, R.E., 2008. The representation of abstract task rules in the human prefrontal cortex. *Cereb. Cortex* 19, 1929–1936. <http://dx.doi.org/10.1093/cercor/bhn222>.
- Berwick, R.C., Friederici, A.D., Chomsky, N., Bolhuis, J.J., 2013. Evolution, brain, and the nature of language. *Trends Cogn. Sci.* <http://dx.doi.org/10.1016/j.tics.2012.12.002>.
- Braver, T.S., Cohen, J.D., Nystrom, L.E., Jonides, J., Smith, E.E., Noll, D.C., 1997. A parametric study of prefrontal cortex involvement in human working memory. *NeuroImage* 5, 49–52.
- Buzsáki, G., Moser, E.I., 2013. Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nat. Neurosci.* 16, 130–138. <http://dx.doi.org/10.1038/nn.3304>.
- Chomsky, N., 1957. *Syntactic Structures*. Mouton, The Hague.
- Conway, C.M., Christiansen, M.H., 2001. Sequential learning in non-human primates. *Trends Cogn. Sci.* 5, 539–546.
- de Haan, E.H.F., Cowey, A., 2011. On the usefulness of 'what' and 'where' pathways in vision. *Trends Cogn. Sci.* 15, 460–466. <http://dx.doi.org/10.1016/j.tics.2011.08.005>.
- de Vries, M.H., Christiansen, M.H., Petersson, K.M., 2011. Learning recursion: multiple nested and crossed dependencies. *Biolinguistics* 5, 10–35.
- Eichenbaum, H., Sauvage, M., Fortin, N., Komorowski, R., Lipton, P., 2012. Towards a functional organization of episodic memory in the medial temporal lobe. *Neurosci. Biobehav. Rev.* 36, 1597–1608. <http://dx.doi.org/10.1016/j.neubiorev.2011.07.006>.
- Fadiga, L., Craighero, L., D'Ausilio, A., 2009. Broca's area in language, action, and music. *Ann. N. Y. Acad. Sci.* 1169, 448–458. <http://dx.doi.org/10.1111/j.1749-6632.2009.04582.x>.
- Fazio, P., Cantagallo, A., Craighero, L., D'Ausilio, A., Roy, A.C., Pozzo, T., Calzolari, F., Granieri, E., Fadiga, L., 2009. Encoding of human action in Broca's area. *Brain* 132, 1980–1988. <http://dx.doi.org/10.1093/brain/awp118>.
- Fitch, W.T., 2010. Three meanings of "recursion": key distinctions for biolinguistics. In: Larson, R.K., Déprez, V.M., Yamakido, H. (Eds.), *The Evolution of Human Language*. Cambridge University Press, New York, pp. 73–90.
- Forkstam, C., Hagoort, P., Fernandez, G., Ingvar, M., Petersson, K.M., 2006. Neural correlates of artificial syntactic structure classification. *NeuroImage* 32, 956–967. <http://dx.doi.org/10.1016/j.neuroimage.2006.03.057>.
- Friederici, A.D., Bahlmann, J., Heim, S., Schubotz, S.I., Anwander, A., 2006. The brain differentiates human and non-human grammars: Functional localization and structural connectivity. *Proc. Natl. Acad. Sci.* 103, 2458–2463. <http://dx.doi.org/10.1073/pnas.0509389103>.
- Friederici, A.D., Bahlmann, J., Friedrich, R., Makuuchi, M., 2011. The Neural Basis of Recursion and Complex Syntactic Hierarchy. *Biolinguistics* 5, 087–104.
- Goel, V., Grafman, J., 1995. Are the frontal lobes implicated in "planning" functions? Interpreting data from the Tower of Hanoi. *Neuropsychologia* 33, 623–642.
- Hagoort, P., 2005. On Broca, brain, and binding: a new framework. *Trends Cogn. Sci.* 9, 416–423. <http://dx.doi.org/10.1016/j.tics.2005.07.004>.
- Hauser, M.D., Chomsky, N., Fitch, W.T., 2002. The faculty of language: what is it, who has it, and how did it evolve? *Science* 298, 1569–1579.
- Ino, T., Doi, T., Hirose, S., Kimura, T., Ito, J., Fukuyama, H., 2007. Directional disorientation following left retrosplenial hemorrhage: a case report with fMRI studies. *Cortex* 43, 248–254.
- Jackendoff, R., 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford University Press.
- Koechlin, E., Jubault, T., 2006. Broca's area and the hierarchical organization of human behavior. *Neuron* 50, 963–974. <http://dx.doi.org/10.1016/j.neuron.2006.05.017>.
- Kravitz, D.J., Saleem, K.S., Baker, C.I., Mishkin, M., 2011. A new neural framework for visuospatial processing. *Nat. Rev. Neurosci.* 12, 217–230. <http://dx.doi.org/10.1038/nrn3008>.
- Kravitz, D.J., Saleem, K.S., Baker, C.I., Ungerleider, L.G., Mishkin, M., 2013. The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends Cogn. Sci.* 17, 26–49. <http://dx.doi.org/10.1016/j.tics.2012.10.011>.
- Kumaran, D., Melo, H.L., Duzel, E., 2012. The emergence and representation of knowledge about social and nonsocial hierarchies. *Neuron* 76, 653–666. <http://dx.doi.org/10.1016/j.neuron.2012.09.035>.
- Lehmann, D., Pascual-Marqui, R.D., Strik, W.K., Koenig, T., 2010. Core networks for visual-concrete and abstract thought content: a brain electric microstate analysis. *NeuroImage* 49, 1073–1079. <http://dx.doi.org/10.1016/j.neuroimage.2009.07.054>.
- Lund, T.E., Nørgaard, M.D., Rostrop, E., Rowe, J.B., Paulson, O.B., 2005. Motion or activity: their role in intra- and inter-subject variation in fMRI. *NeuroImage* 26, 960–964. <http://dx.doi.org/10.1016/j.neuroimage.2005.02.021>.
- Maess, B., Koelsch, S., Gunter, T.C., Friederici, A.D., 2001. Musical syntax is processed in Broca's area: an MEG study. *Nat. Neurosci.* 4, 540–545 (http://www.nature.com/neuro/journal/v4/n5/supplinfo/nn0501_540_S1.html).
- Makuuchi, M., Bahlmann, J., Anwander, A., Friederici, A.D., 2009. Segregating the core computational faculty of human language from working memory. *Proc. Natl. Acad. Sci.* 106, 8362–8367. <http://dx.doi.org/10.1073/pnas.0810928106>.
- Margulies, D.S., Vincent, J.L., Kelly, C., Lohmann, G., Uddin, L.Q., Biswal, B.B., Villringer, A., Castellanos, F.X., Milham, M.P., Petrides, M., 2009. Precuneus shares intrinsic functional architecture in humans and monkeys. *Proc. Natl. Acad. Sci.* 106, 20069–20074.
- Martins, M.D., 2012. Distinctive signatures of recursion. *Philos. Trans. R. Soc. B* 367, 2055–2064. <http://dx.doi.org/10.1098/rstb.2012.0097>.
- Martins, M.D., Fitch, W.T., 2012. Empirical approaches to recursion. In: Scott-Phillips, T.C., Tamariz, M., Cartmill, E.A., Hurford, J.R. (Eds.), *The Evolution of Language – Proceedings of the 9th International conference*. World Scientific, Kyoto, Singapore, pp. 219–225.
- Mueller, S.T., 2011. *The Psychology Experiment Building Language*.
- Nardini, M., Jones, P., Bedford, R., Braddick, O., 2008. Development of cue integration in human navigation. *Curr. Biol.* 18, 689–693. <http://dx.doi.org/10.1016/j.cub.2008.04.021>.
- Nichols, T., Brett, M., Andersson, J., Wager, T., Poline, J.-B., 2005. Valid conjunction inference with the minimum statistic. *NeuroImage* 25, 653–660. <http://dx.doi.org/10.1016/j.neuroimage.2004.12.005>.

- Opitz, B., Friederici, A.D., 2007. Neural basis of processing sequential and hierarchical syntactic structures. *Hum. Brain Mapp.* 28, 585–592. <http://dx.doi.org/10.1002/hbm.20287>.
- Patel, A.D., 2003. Language, music, syntax and the brain. *Nat. Neurosci.* 6, 674–681.
- Petersson, K.M., 2004. Artificial syntactic violations activate Broca's region. *Cogn. Sci.* 28, 383–407. <http://dx.doi.org/10.1016/j.cogsci.2003.12.003>.
- Petersson, K.M., Hagoort, P., 2012. The neurobiology of syntax: beyond string sets. *Phil. Trans. R. Soc. B* 367, 1971–1983.
- Petersson, K.M., Folia, V., Hagoort, P., 2012. What artificial grammar learning reveals about the neurobiology of syntax. *Brain Lang.* 120, 83–95. <http://dx.doi.org/10.1016/j.bandl.2010.08.003>.
- Rajah, M.N., Ames, B., D'Esposito, M., 2008. Prefrontal contributions to domain-general executive control processes during temporal context retrieval. *Neuropsychologia* 46, 1088–1103. <http://dx.doi.org/10.1016/j.neuropsychologia.2007.10.023>.
- Ranganath, C., 2010. A unified framework for the functional organization of the medial temporal lobes and the phenomenology of episodic memory. *Hippocampus* 20, 1263–1290. <http://dx.doi.org/10.1002/hipo.20852>.
- Ranganath, C., Ritchey, M., 2012. Two cortical systems for memory-guided behaviour. *Nat. Rev. Neurosci.* 13, 713–726. <http://dx.doi.org/10.1038/nrn3338>.
- Sato, N., Sakata, H., Tanaka, Y.L., Taira, M., 2010. Context-dependent place-selective responses of the neurons in the medial parietal region of macaque monkeys. *Cereb. Cortex* 20, 846–858.
- Sridharan, D., Levitin, D.J., Menon, V., 2008. A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proc. Natl. Acad. Sci.* 105, 12569–12574. <http://dx.doi.org/10.1073/pnas.0800005105>.
- Uddén, J., Folia, V., Forkstam, C., Ingvar, M., Fernandez, G., Overeem, S., van Elswijk, G., Hagoort, P., Petersson, K.M., 2008. The inferior frontal cortex in artificial syntax processing: an rTMS study. *Brain Res.* 1224, 69–78. <http://dx.doi.org/10.1016/j.brainres.2008.05.070>.
- Unterrainer, J.M., Owen, A.M., 2006. Planning and problem solving: from neuropsychology to functional neuroimaging. *J. Physiol. Paris* 99, 308–317. <http://dx.doi.org/10.1016/j.jphysparis.2006.03.014>.
- Wang, J., Conder, J.A., Blitzer, D.N., Shinkareva, S.V., 2010. Neural representation of abstract and concrete concepts: a meta-analysis of neuroimaging studies. *Hum. Brain Mapp.* 31, 1459–1468. <http://dx.doi.org/10.1002/hbm.20950>.
- Wohlschlagel, A., Gattis, M., Bekkering, H., 2003. Action generation and action perception in imitation: an instance of the ideomotor principle. *Philos. Trans. R. Soc. B* 358, 501–515. <http://dx.doi.org/10.1098/rstb.2002.1257>.
- Zaitsev, M., Hennig, J., Speck, O., 2004. Point spread function mapping with parallel imaging techniques and high acceleration factors: fast, robust, and flexible method for echo-planar imaging distortion correction. *Soc. Magn. Reson. Med.* 52, 1156–1166. <http://dx.doi.org/10.1002/mrm.20261>.
- Zaretskaya, N., Anstis, S., Bartels, A., 2013. Parietal cortex mediates conscious perception of illusory gestalt. *J. Neurosci.* 33, 523–531. <http://dx.doi.org/10.1523/jneurosci.2905-12.2013>.