

Recursive hierarchical embedding in vision is impaired by posterior middle temporal gyrus lesions

Mauricio J.D. Martins^{1,2,3}, **Carina Krause**^{2,3,4}, **David A. Neville**^{1,5}, **Daniele Pino**^{2,3},
Arno Villringer^{1,2,3} and **Hellmuth Obrig**^{1,2,3}

The generation of hierarchical structures is central to language, music and complex action. Understanding this capacity and its potential impairments requires mapping its underlying cognitive processes to the respective neuronal underpinnings. In language, left inferior frontal gyrus and left posterior temporal cortex (superior temporal sulcus/middle temporal gyrus) are considered hubs for syntactic processing. However, it is unclear whether these regions support computations specific to language or more generally support analyses of hierarchical structure. Here, we address this issue by investigating hierarchical processing in a non-linguistic task. We test the ability to represent recursive hierarchical embedding in the visual domain by contrasting a recursion task with an iteration task. The recursion task requires participants to correctly identify continuations of a hierarchy generating procedure, while the iteration task applies a serial procedure that does not generate new hierarchical levels. In a lesion-based approach, we asked 44 patients with left hemispheric chronic brain lesion to perform recursion and iteration tasks. We modelled accuracies and response times with a drift diffusion model and for each participant obtained parametric estimates for the velocity of information accumulation (drift rates) and for the amount of information accumulated before a decision (boundary separation). We then used these estimates in lesion-behaviour analyses to investigate how brain lesions affect specific aspects of recursive hierarchical embedding. We found that lesions in the posterior temporal cortex decreased drift rate in recursive hierarchical embedding, suggesting an impaired process of rule extraction from recursive structures. Moreover, lesions in inferior temporal gyrus decreased boundary separation. The latter finding does not survive conservative correction but suggests a shift in the decision criterion. As patients also participated in a grammar comprehension experiment, we performed explorative correlation-analyses and found that visual and linguistic recursive hierarchical embedding accuracies are correlated when the latter is instantiated as sentences with two nested embedding levels. While the roles of the inferior temporal gyrus and posterior temporal cortex in linguistic processes are well established, here we show that posterior temporal cortex lesions slow information accumulation (drift rate) in the visual domain. This suggests that posterior temporal cortex is essential to acquire the (knowledge) representations necessary to parse recursive hierarchical embedding in visual structures, a finding mimicking language acquisition in young children. On the contrary, inferior frontal gyrus lesions seem to affect recursive hierarchical embedding processing by interfering with more general cognitive control (boundary separation). This interesting separation of roles, rooted on a domain-general taxonomy, raises the question of whether such cognitive framing is also applicable to other domains.

- 1 Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Berlin, Germany
- 2 Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
- 3 Department of Cognitive Neurology, University Hospital Leipzig, Leipzig, Germany
- 4 Erziehungswissenschaftliche Fakultät Pädagogik im Förderschwerpunkt Sprache und Kommunikation, Leipzig University, Leipzig, Germany
- 5 Donders Institute for Brain, Cognition and Behavior, Radboud University, Nijmegen, The Netherlands

Received November 26, 2018. Revised June 11, 2019. Accepted June 16, 2019

© The Author(s) (2019). Published by Oxford University Press on behalf of the Guarantors of Brain.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Correspondence to: Mauricio Dias Martins
 Max Planck Institute for Human Cognitive and Brain Sciences
 Stephanstrasse 1a, 04103, Leipzig, Germany
 E-mail: mmartins@cbs.mpg.de

Keywords: hierarchy; visuospatial; lesion; syntax; inferior frontal gyrus

Abbreviations: DDM = drift-diffusion model; IFG = inferior frontal gyrus; ITE = Visual Iteration Task; MTG = middle temporal gyrus; pTC = posterior temporal cortex; REC = Visual Recursion Task; RHE = recursive hierarchical embedding

Introduction

Humans have the ability to process hierarchical structures (Fitch and Martins, 2014; Wilson *et al.*, 2017), and this capacity has been mostly studied in the domain of language. For instance, in the sentence ‘instinctively, birds that fly swim’, the adverb ‘instinctively’ modifies the verb ‘swim’ and not the verb ‘fly’, despite being more distant to the former in the linear structure. This is because, within the underlying syntactic structure, ‘instinctively’ is closer to ‘swim’ than to ‘fly’ regarding hierarchical depth: {instinctively, {birds {that fly} swim}} (Berwick and Chomsky, 2015). Such hierarchy in language is thought to result from an innate recursive procedure (Berwick *et al.*, 2013; Berwick and Chomsky, 2015; Everaert *et al.*, 2015), which when applied stepwise, generates multiple nested hierarchical levels. Although ‘infinite recursion’ is considered a core feature of human language, it is rare to find sentences with more than two levels of hierarchical embedding. One important limitation is that, in spoken language, working memory capacity is strongly taxed due to an increasing number of elements that have to be held active before sentence meaning integration. In other domains, these memory limitations might not be as pronounced (e.g. the visuospatial), even though the theoretical limit of recursion is never achieved.

The mechanisms supporting hierarchy in language are thought to be implemented by dedicated neural systems (Berwick *et al.*, 2013). Evidence for this view comes from neuroimaging (EEG and functional MRI) and lesion studies. The patholinguistic condition of agrammatism, for instance, interferes with the processing of complex syntax, while other linguistic abilities are largely preserved (Grodzinsky and Santi, 2008). It remains unclear, however, whether areas putatively supporting hierarchical linguistic processes could also support recursive hierarchical embedding (RHE) in other domains. Here, we aim to investigate whether lesion patterns associated with agrammatism also interfere with RHE in the visual domain, by testing patients with an acquired chronic brain lesion in the left hemisphere.

While studies on the neural correlates of RHE are scarce, there is converging evidence that syntactic processing relies on two major hubs: the inferior frontal gyrus (IFG) and the posterior temporal cortex (pTC) (Friederici, 2011; Hagoort and Indefrey, 2014; Matchin *et al.*, 2017). Within the pTC,

some studies implicate the posterior superior temporal sulcus and others implicate the posterior middle temporal gyrus (MTG) (see Hagoort and Indefrey, 2014, for a meta-analysis). The role of these areas has been discussed along two rationales: computations within IFG may be necessary to implement recursive generation of linguistic hierarchies (Friederici *et al.*, 2011; Zaccarella *et al.*, 2017), a finding that would explain central symptoms of agrammatism (Matchin and Rogalsky, 2018). Alternatively, functional MRI results are compatible with the interpretation that IFG implements domain-general ‘computations’ (e.g. relating to working memory or cognitive control), which operate on domain-specific ‘representations’ supported by pTC (Rogalsky *et al.*, 2011; Matchin, 2018; Matchin *et al.*, 2017). These representations might be both simple lexical units or complex hierarchical templates containing a set of features that dictate how basic units can be further combined (Matchin, 2018). These two models may be tested by extending research on recursion to other domains.

Here we focus on the visual domain. We study recursive abilities in individuals with a chronic circumscribed lesion in the left hemisphere. As most of them participated in another study on linguistic syntax processing, we are in the unique position to gain a first insight into potential correlations between RHE across different modalities.

RHE has been hypothesized to play a role in the processing of visuo-spatial structures (Pinker and Jackendoff, 2005). To supply experimental evidence, a visual recursion task (REC) was recently developed in which participants generate novel hierarchical levels using a recursive embedding rule (Martins *et al.*, 2016). In a control iteration task (ITE) items are added sequentially to existing hierarchies without generating new levels. As recursive competence in this task correlates with similar abilities in music and action (Martins *et al.*, 2017), we consider it ideal to tap into shared RHE resources across domains.

In the visual domain, the contrast REC versus ITE activates anterior regions within superior temporal sulcus, along with several regions within the default mode network (DMN) and medial temporal lobe (MTL) (Martins *et al.*, 2014a; Fischmeister *et al.*, 2017). These regions fit the classical view that the visual ventral stream and the parieto-medial temporal pathway (PMT) integrate items in contextual frames (Kravitz *et al.*, 2011). Recently, this research has been extended to the domains of music and action (Martins, 2017; Martins *et al.*, 2017). While we

found no evidence for the involvement of IFG or pTC in these domains, we found areas involved in tonal sequence representation (anterior STG) in music (Martins, 2017), and areas involved in motor imagery for the motor domain (premotor cortex, basal ganglia, cerebellum; Martins *et al.*, 2019). In favour of Matchin's hypothesis (Matchin, 2018; Matchin *et al.*, 2017), this suggests that domain-specific 'representations' might be involved in the processing of RHE.

If domain-specific representations are recruited for highly trained behaviour as applies to trained musicians, this does not preclude the possibility that *de novo* analysis of recursive structure may rely on more domain-general capacities, therefore tapping into a similar neuronal network across tasks. In fact, previous behavioural research with untrained participants during the acquisition of RHE suggests communalities across visual, music and action domains (Martins *et al.*, 2017).

Importantly, the general capacity to instantiate RHE is thought to result from the interaction of a core RHE machinery with two peripheral systems: sensory-motor and conceptual-intentional (Everaert *et al.*, 2015). Crucially, the interaction of the core capacity with sensorial systems of different domains requires specialized interfaces. For instance, RHE in music and oral language hinges on auditory working memory system, while in the visual domain RHE is dependent on a visual working memory system. While these interfaces might have constraints specific to each domain, it is possible that similar neural networks are necessary to instantiate the core capacity across domains.

To elucidate the processes involved in the acquisition of RHE in the visuo-spatial domain *vis a vis* with the processes involved in syntactic processing, we tested individuals with a chronic acquired left hemisphere lesion resulting from various aetiologies. If a distinct network is required for the inference of the recursive process in a novel task, we expect that variance across participants will vary depending on lesion location. Furthermore, we can test this relationship by correlating performance in the visual recursion task with some aspects of a linguistic task involving syntactic embedding, and performed in a largely overlapping cohort of participants.

As behavioural measures of response accuracy and latency are expected to largely vary in clinical populations, we extend our behavioural analysis to a more comprehensive analytical framework. The drift-diffusion model (DDM; Smith and Ratcliff, 2004) is a sequential sampling model for the analysis of choice-reaction time data, which is the combination of reaction time and accuracy measures. The DDM model yields estimates of: (i) the velocity at which a decision is made (v' , drift); (ii) the amount of information that is required to make the decision (a' , boundary); and (iii) the amount of time required to complete non-decision processes (t' , non-decision time). Here, we used a hierarchical version of the DDM model (Wiecki *et al.*, 2013) to obtain estimates that take into account inter-subject variability.

As an example of how the DDM can explain non-linear dynamics of the decision process, consider the scenario where impairment may lead to slower responses in a specific task condition. Simple analyses of behavioural performance would not be able to compare the alternative explanations of whether the slow responses are due to more information being needed to make the decision (larger boundary separation) or due to a reduction in the rate at which information accumulates (drift rate). Changes in either of these two factors could produce responses with larger reaction times. Therefore a sequential sampling model is needed to correctly estimate the dynamics of the underlying decision process.

For the detection of brain regions involved in parsing the RHE properties of our stimuli, the v' (drift) parameter is particularly informative since it has been shown to account for information accumulation in both behaviour and neural (spike) data (Gerstein and Mandelbrot, 1964; Srinivasan and Sampath, 2013; Durstewitz *et al.*, 2016).

Here, we investigate the mechanisms supporting the acquisition of RHE in the visual domain, by assessing performance in a population with acquired focal chronic lesions of the left hemisphere. We target the effects of lesion site on interindividual variance in behavioural performance and hypothesize that lesions in brain areas central for recursive analysis in the linguistic domain (IFG and pTC) also affect RHE in vision. Regions within the right hemisphere might be relevant for rule acquisition in general and for the acquisition of linguistic competence (Dehaene-Lambertz *et al.*, 2002). However, since linguistic impairment in adults is largely caused by left hemispheric lesions, these will be the focus of our analysis. Thus, we predict that patients with lesions in these areas of the left hemisphere will be slower in the accumulation of information necessary to process RHE (modelled as drift rate). Comparing accuracy in the novel visual recursion task and accuracy in a syntax task (reported in detail elsewhere), we hypothesize that performance in REC (rather than ITE) correlates with performance in understanding multiple embedded sentences.

Materials and methods

Participants

Forty-four participants ($n = 19$ female) with an acquired chronic left hemispheric brain lesion were included. They were recruited at the Clinic for Cognitive Neurology, University Hospital Leipzig. Mean age \pm SD (range) was 50 years \pm 10.6 (24–74), and time since event 23.1 months \pm 22.3 (3–115). Participants were in the chronic stage after a vascular lesion ($n = 37$; 25 ischaemic stroke, five subarachnoid haemorrhage, seven intracerebral haemorrhage) traumatic brain injury ($n = 3$), encephalitis ($n = 1$) or suffered from brain tumours in the stage of a 'stable disease' ($n = 3$). We deliberately chose patients with lesions related to different aetiologies as this reduces the bias for specific lesion patterns

due to the nature of the disease (e.g. vascular territories in ischaemic stroke). All patients were in the chronic stage of the disease and showed no clinical or radiological sign of relevant affection of brain areas distant to the focal lesion which entered the analysis (Supplementary Table 1 for details, including some additional comments on the rationale of patient selection). Participants gave informed consent; data were collected according to the Declaration of Helsinki after approval of the local ethics committee.

Patients underwent an extensive clinical testing during their therapeutic stay at the clinic, which was used to judge eligibility for participation. Severe cognitive impairment was an exclusion criterion. A left hemispheric lesion in a structure considered to be part of the extended language network was present in all patients. However, manifest aphasia at the time of participation was not mandatory. Overall, 12 patients showed no aphasia at the time of testing, of whom seven had shown a clinically relevant aphasia in the acute stage of their disease. At the time of testing 13 participants had an aphasia that was classified according to the Aachen Aphasia Test (AAT) (Huber *et al.*, 1984), the standard diagnostic tool for aphasia in German (seven amnesic, three Broca, two Wernicke, one non-classifiable). Nineteen participants had clinically manifest language impairment termed ‘residual aphasia’ according to AAT logics. In sum, the cohort was mildly impaired, but only four patients had never shown aphasic symptoms, while 32 showed a clinically-relevant language impairment at the time of testing. For details see Supplementary Table 1.

Imaging

For all participants ($n = 44$), structural imaging was available. Thirty-nine scans were performed at in-house scanners (3T Siemens MRI system Trio[®] or Verio[®] system, Siemens Medical Systems) including 3D T₁-weighted (1 mm isovoxel), and fluid-attenuated inversion recovery (FLAIR) images. In four patients clinical MRI at a lower resolution [3–5 mm slice thickness, including FLAIR or turbo inversion recovery magnitude (TIRM) and T₁ images] was available; in one patient a cranial CT was used for lesion delineation. For the lesion-behaviour analyses, lesions were manually delineated in all three planes on each slice of the T₁ or cranial CT images using MRIcron (Rorden and Brett, 2000), for MRI FLAIR/TIRM-images served as a reference. Images were then transformed into standard stereotactic space (MNI) using SPM8 (www.fil.ion.ucl.ac.uk/spm) and the ‘clinical toolbox’ (nitrc.org/projects/clinicaltbx/), which allows for normalizing images from different modalities into the same space. The unified segmentation approach was applied (Ashburner and Friston, 2005) and estimation of normalization parameters was restricted to healthy tissue using predefined lesion masks (Brett *et al.*, 2001). The resulting normalized binary lesion maps were next analysed in NiiStat (<https://github.com/neuro-labusc/NiiStat>), including a ‘traditional’ voxel-wise analysis but also providing options for a region-based analysis. While the former can be considered more sensitive to smaller lesion foci correlating with behaviour, the latter is less susceptible to false negatives, since the issue of multiple comparison correction is greatly reduced.

Experimental tasks

Visual Recursion Task

For the Visual Recursion Task (REC), participants were shown a sequence of three images (steps 1–3), which depicted a process generating a visual fractal. After the first three images, participants were asked to discriminate, from two choices, the image corresponding to the correct continuation of the previous sequence of three (i.e. the fourth step). One of the choices was the correct image, and the other was a foil. The task is an adaptation of the one used and described in detail elsewhere (Martins, 2012; Martins *et al.*, 2014a, b). REC was composed of 27 trials, nine of each foil category. Variability was achieved by varying the number of constituents composing the visual fractal, as described in (Martins *et al.*, 2014b).

Each trial began with the presentation of three images of a fractal generation in the top half of the screen, sequentially from left to right (Fig. 1A) at a rate of 2 s between image onsets. After the presentation of the first three steps, two new images were presented simultaneously in the bottom half of the screen. One image corresponded to the correct continuation of the recursive process that generated the first three fractals, and the other corresponded to a foil (or ‘incorrect’ continuation). Participants were asked to select the image that continued the recursive process. The position of the ‘correct’ image (left or right) was randomized. After the initial instructions, each trial had a maximum duration of 30 s before a timeout. No feedback was given regarding the correctness of choice. Total duration of the task (27 trials) was ~12 min.

To control for effects of information processing demands, we included stimuli with different degrees of visual complexity (complexity ‘3’, ‘4’, and ‘5’). Furthermore, to control for the usage of simple visual heuristic strategies in REC performance, we included several categories of foils (‘Odd’, ‘Position’ and ‘Repetition’; Fig 1B). Complexity and foil-type allow for nine types of stimuli (three complexity levels × three foil types). Three examples of each type of stimuli were generated using the programming language Python, resulting in the total set of 27 stimuli.

Visual Iteration Task

The second task was iterative but non-recursive (Martins *et al.*, 2016). The principle underlying ITE is similar to REC in that it involves a stepwise procedure applied to hierarchical structures. However, ITE lacks recursive embedding. Instead, in ITE, additional elements are added to one pre-existing hierarchical structure, without producing new hierarchical levels (Fig. 1A, bottom right). As for REC, an understanding of this stepwise procedure is necessary to correctly predict the next step. Number of trials, visual complexity and foil categories and distributions were equivalent to REC.

Procedure

The visual recursion/iteration task and a task assessing comprehension of multiple degrees of sentential embedding (Fig. 4A) were part of a larger assessment battery. The linguistic task was performed one day prior to the visual tasks. For the visual task half of the participants started the procedure with ITE [order: Iteration (I)→Recursion (R)] and half started the procedure with REC (order: R→I).

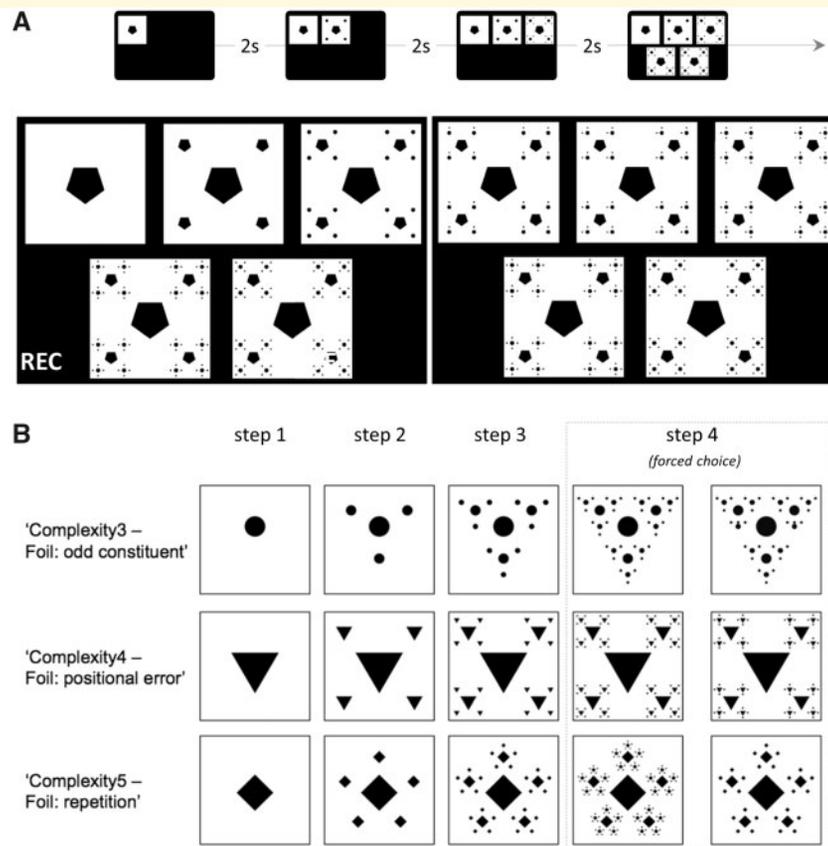


Figure 1 Experimental paradigm. (A) The presentation of the visual recursion/iteration task (REC, ITE) comprised four steps including a successive presentation of the steps 1–3 at the top of the screen to then present the two options for a forced choice at the bottom. Examples for REC and ITE screen shots are provided for step 4; note that the final choice images are identical for both tasks. Location of correct image was randomized (e.g. left in the ITE and right in the REC example provided). (B) Examples of fractals used in REC. There were different categories of 'visual complexity'—3, 4 and 5—and different categories of foils. In 'odd constituent' foils, two elements within the whole hierarchy were misplaced; in 'positional error' foils, all elements within new hierarchical levels were internally consistent, but inconsistent with the previous iterations; in 'repetition' foils, no additional iterative step was performed after the third iteration.

Statistical analyses

Visual task data allows the analysis of two factors: Task (REC/ITE) and Order (I→R/R→I) and their interaction. We analysed the base parameters of the behavioural experiment (reaction time and accuracy) and then fed these into a drift-diffusion-model yielding a measure for the velocity of information accumulation (drift, v') and the amount of information needed to make a decision (boundary, a'). The third parameter of the latter analysis, non-decision time (t'), is not analysed further here. Posterior group-level distributions for all of the parameters can be inspected in [Supplementary Fig. 8](#).

Reaction time and accuracy

Statistical analyses were performed in R studio (1.1.453). We ran linear mixed models with the function `lmer()` with package `lme4` (Bates *et al.*, 2014), with participants as random factor. Best lambda transformation was found using `boxcox()` with package `MASS` (Venables and Ripley, 2002). All residual distributions reported in this manuscript were normal, calculated using Shapiro-Wilk test (all P 's > 0.2). Models are reported using ANOVA (type = II) and the R package `Anova()` for P -

values. When main effects were found, we tested for pairwise differences with `emmeans()` (Russell, 2018), using Kenward-Roger methods to calculate the degrees of freedom, and Tukey P -value adjustment when comparing three parameters. Finally, we ran spearman correlations since variables were not normally distributed, and one-tailed tests since we expected grammar comprehension to positively correlate with visual recursion accuracy and negatively with response time.

Drift diffusion analysis

Choice reaction time data were fitted to a hierarchical version of drift diffusion model using customized scripts implemented in the HDDM toolbox (Wiecki *et al.*, 2013) for Python. Since the dataset presented very long responses (> 10 s) we scaled all of the reaction times by a constant factor (10) so that all of the reaction times would be in the range of 0–10 s. The analyses proceeded as follows. First, we fitted to the data models with different combinations of free parameters over the two factors of experiment (Order and Task). Each of the models was fitted to the data using standard MCMC minimization routines with 50 000 iterations, a burn-in period of 5000 and thinning of 1. For all of the chains, the results converged to stable estimates

as indicated by the relative diagnostic plots and by the Rhat statistics (<1) which is a measure of chain convergence. For each of the models we then computed the relative deviance information criterion (DIC) which is a measure of the goodness of fit of the model to the data penalized by the complexity of the model (i.e. functional form of the free parameters). A model with a lower DIC score is to be preferred as the most parsimonious account of the data. Comparisons of the DIC scores for all models indicated a model with drift rate (ν') and non-decision time (t') free to vary over both experimental conditions and boundary separation (a') fixed over task-order conditions as the most parsimonious account of the data. For brevity, results of the model comparison are reported in the supplementary information and only results from the best-fitting model are reported in the main text. Estimated parameters for the best fitting model were finally tested for significant differences using linear regression analyses with permutation based-calculation of significance levels implemented in R[®] with the package `lmPerm` (Wheeler and Torchiano, 2016).

Lesion-behaviour analyses

Analyses were conducted with NiiStat (<https://github.com/neuroabusc/NiiStat>). Anatomical correlates of REC and ITE accuracy (with Task-Order as nuisance variable) were assessed using both a region- and a voxel-based approach (Bates *et al.*, 2003; Rorden *et al.*, 2007).

For the statistical region of interest approach two atlases were used: the Atlas of Intrinsic Connectivity of Homotopic Areas (AICHA), which contains 384 grey matter regions of interest (Joliot *et al.*, 2015), and the Brodmann atlas, containing 82 grey matter areas. To find out about the relevance of a specific region of interest for task performance the proportion of damage to a given region was computed in each participant and entered into a general linear model (GLM). This statistically assesses a correlation between the amount of damage in a given region and the behaviour in question. The result was converted to a z-score for each region. To control for family-wise error, the data were permuted 5000 times to establish a significance threshold. Only those regions with z-scores above the permutation threshold $P < 0.05$ are reported (Rorden *et al.*, 2007). Relative to voxelwise approaches (see below), this method increases statistical power by both averaging data and limiting the number of statistical comparisons. An initial statistical region of interest analysis was conducted to examine each parameter individually. A second analysis used each parameter as a nuisance factor for the other variables using the Freedman-Lane method (Freedman and Lane, 1983), this is a natural extension of GLM that allows us to compute permutation (Winkler *et al.*, 2014). We were especially interested in the correlates of REC when taking ITE and order as nuisance regressors.

Additionally, we performed the 'traditional' voxel-wise approach, conducting independent statistical tests for each voxel that is covered by a lesion overlap of at least four participants. To control for multiple comparisons only voxels surviving 5000 permutation correction are reported in the statistical maps. Using NiiStat, *t*-tests were computed for every voxel to see if those individuals with a lesion at that location exhibited reliably different behavioural performance (using our continuous indices) than those without a lesion. The *t*-tests were confined to voxels that sustained damage in at least

10% ($n = 4$ subjects) of the sample, defining the areas for which the analysis can provide statistical inference.

Grammar task

All participants of the visual recursion/iteration-task also performed another experiment targeting the comprehension of complex syntax (Krause *et al.*, submitted for publication). In that experiment participants listened to sentences containing three propositions regarding two animals interacting with each other. Propositions were: (i) the action (e.g. *X washes Y*); (ii) the colour of one animal (e.g. *X is brown*); and (iii) the mood of one animal (e.g. *X laughs*). After the auditory presentation of the sentence (overall set: 132 sentences), participants had to choose the correct image from a set of four images (one correct, three incorrect containing a distractor for each proposition). Syntactic complexity of the sentences was manipulated by the embedding depth in that the three propositions were serially linked by a conjunction or embedded using embedded relative clauses. An additional manipulation was introduced by varying argument order (i.e. subject first or object first relative clause). Here, we only use the differences in embedding depth, for which we supply an example in Fig. 4A. Note that only the difference between EMB1 and EMB2 enters the analysis. EMB0 sentences contain crossed-dependencies, meaning that the personal pronoun 'er' [he] in the last proposition can relate to the animal mentioned first and the one mentioned second because they share grammatical gender. The difference in argument order is illustrated in an example in the Supplementary material including an example for the four-image choice.

Data availability

Anonymized data are available on request.

Results

Reaction time and accuracy

Accuracy showed a main effect of Order and an interaction between Task \times Order. The main effect for Task did not reach statistical significance. For reaction time neither main nor interaction effects were significant.

Results for REC and ITE are presented in Fig. 2. To test for differences between ITE and REC, we performed two independent analyses, one with accuracy and another with response time as dependent variables. As predictors, we included Task (REC/ITE), Order (R \rightarrow I versus I \rightarrow R, balanced across participants), and the interaction Task \times Order.

For accuracy, we found a significant effect of Order [$F(1,85) = 6.4$, $P = 0.01$] and Task \times Order [$F(1,85) = 7.4$, $P = 0.007$]. The main effect of Task was not significant ($P = 0.7$). The proportion of correct responses in REC was lower in the Order when this task was performed first (R \rightarrow I versus I \rightarrow R) [$t(61.7) = -3.5$; $P = 0.0009$]. Performance in ITE did not differ between the two orders [$t(61.7) = 1.0$; $P = 0.3$]. We repeated the procedure for

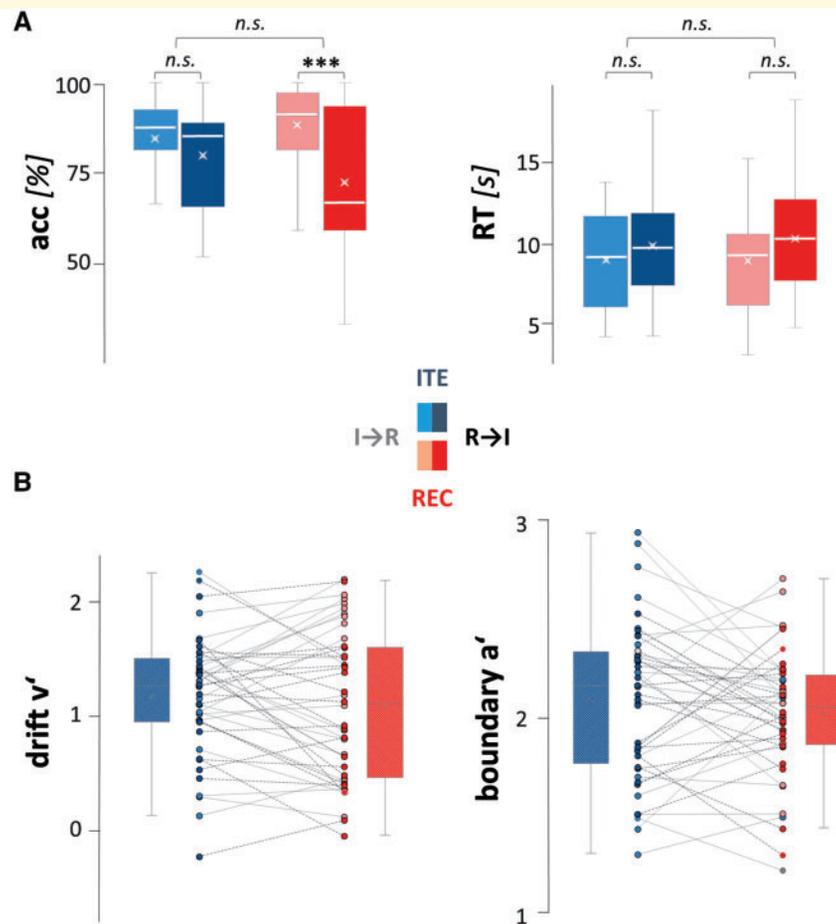


Figure 2 Behavioural data. (A) Percentage of correct answers (acc [%], left) and response time (RT [s], right) in the ITE (blue), and REC (red). The order of visual tasks was either 'I→R' or 'R→I' as indicated by the light or darker shading. (B) We combined these data into a hierarchical DDM (text for details) and obtained posterior estimates for drift rate (drift v') and boundary separation (boundary a'). Note that order no longer influences the performance since roughly equal numbers of patients showed values REC > ITE and ITE > REC for these measures (colour coding of individual data points as in A). The variability can be used across participants in the lesion-behaviour analysis (main text and Fig. 3). For detailed descriptive statistics, see Supplementary Table 2.

reaction time and the best fit was the intercept-only model [with restricted maximum likelihood (REML) = 180], meaning that Task, Order, and the interaction were not significant (all P 's > 0.2).

Hierarchical drift diffusion model results

As reaction time and accuracy interact in a complex manner even in binary decision tasks, we performed an additional analysis based on the values that were derived from a DDM.

We ran a model selection procedure (Supplementary material) and found that a model with drift and non-decision time free to vary over Order and Task factors and boundary distance over Task provided the most parsimonious account of the data (best model). We then used the posterior estimates for each Task (REC and ITE) while controlling for order effects. We thus obtained measures of

performance for REC and ITE independent of the (arbitrarily assigned) order and the overall performance, the latter depending on a large number of individual differences between patients which are of no specific interest here. As illustrated in Fig. 2B the analysis showed a large variance across participants and no significant differences between Tasks. The fact that nearly equal numbers of participants showed REC > ITE and vice versa for both orders supports the effective cancellation of the order effect for this analysis, and thereby allows for the lesion-behaviour analysis across all participants.

Lesion-behaviour correlations: statistical region of interest and voxelwise analyses

Lesion-behaviour analyses performed on the 'base-parameters' reaction time and accuracy yielded no statistically robust results.

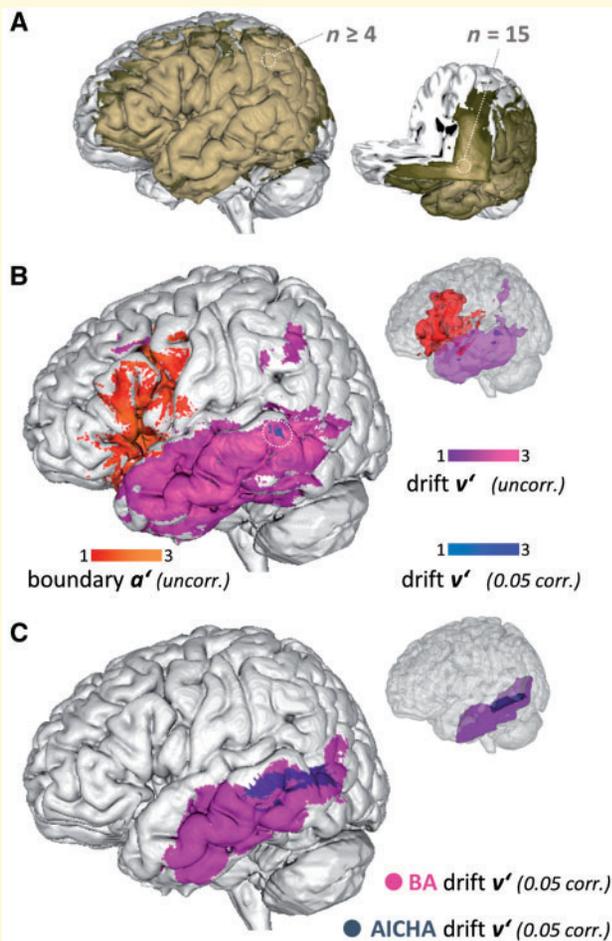


Figure 3 Lesion-behaviour analyses. (A) The area covered. *Left:* Coloured areas show a lesion in at least one patient, in the lighter area at least four lesions overlap representing the area in which the analysis was performed; *right:* area of maximal overlap ($n = 15$) projecting to the insular cortex as is typically seen in stroke dominated lesion studies. (B) Voxel-wise approach: Uncorrected (uncorr.) maps are shown for boundary separation (a' , red) and drift rate (v' , purple) for the REC, with ITE as nuisance variable. IFG lesions were associated with lower a' , meaning that participants collected less information before reaching a decision. On the other hand, lesions in the MTG and STG were associated with lower drift rate, meaning that these patients collected information slower. Only 39 voxels in the MTG (blue area circled for illustration purpose) survived correction for REC v' . (C) Statistical region of interest-symptom mapping, shows significant correlations between REC v' and MTG for the AICHA (purple) and the Brodmann atlas (BA21, blue).

On the contrary, for the drift rate (v') and boundary separation (a') both statistical region of interest-based and voxelwise analyses yielded lesion patterns which correlated significantly with the variability of the parameters across participants.

For the statistical region of interest analysis, temporal cortex areas correlated with the variation of v' when factoring out ITE as nuisance factor. Drift rate v' (speed of information integration during REC decision-making

processes) was lower when the participant's lesion included: the posterior MTG ($z = 3.0$), according to the AICHA atlas, and BA 21 according to the Brodmann atlas ($z = 3.3$). The results are depicted in Fig. 3C.

Using the voxelwise approach (Fig. 3B), lesions in the left MTG correlated with decreased drift rate v' . Interestingly, lesions in parts of the IFG were associated with a decrease in the boundary separation a' , indicating that participants with lesions in these areas acquire less information before they make a decision. Only 39 voxels in the posterior MTG survived thresholding in the drift results ($z > 3.7$, 5000 permutations, $P = 0.05$), converging with the statistical region of interest approach. No voxels survived thresholding for the boundary separation.

Additional analysis regarding the correlation with a Grammar task

As most of the participants ($n = 41$) also performed a task on complex grammar comprehension, which is reported in detail elsewhere (Krause *et al.*, submitted for publication), we performed a correlation analysis between the performance in the visual tasks (REC and ITE) and an aspect of the grammar experiment which can be considered a linguistic counterpart of the REC/ITE-learning task reported here. As detailed in the 'Materials and methods' section, the linguistic task requires that participants judge the meaning of sentences with increasing levels of embedding. Here, we use the performance for single- and double-embedded sentences (EMB1 and EMB2, for an example see Fig. 4A). The percentage of correct responses in the Grammar task are provided in Supplementary Table 3.

To test for communalities between our visual task and the performance in the different embedding levels of the grammar tasks, we ran Spearman correlations between accuracy in REC and ITE, and accuracy in EMB1 and EMB2. For this behavioural analysis we chose accuracy and not the DDM parameters v' and a' in order to compare similar constructs. The DDM is a model for 2-forced choice tasks and therefore cannot be directly applied to the 4-choice grammar tasks. Therefore, since we cannot obtain a 'grammar drift' to compare with the 'visual drift', we decided to compare grammar accuracy with visual accuracy.

Scatterplots are depicted in Fig. 4B. P -values are given for one-tailed tests, as we expected a positive correlation between the visual and grammar tasks. We found that, for accuracy, the correlation with EMB1 was marginal for ITE ($r_s = 0.24$, $P = 0.066$) and significant for REC ($r_s = 0.38$, $P = 0.008$); EMB2 correlated only with REC ($r_s = 0.27$, $P = 0.042$) but not with ITE ($r_s = 0.01$, $P = 0.5$). The full correlation matrix (including EMB0) is depicted in Supplementary Fig. 3.

To test whether these differences between REC and ITE were consistent, i.e. to test if EMB2 was more correlated

Table 1 LMM Dependent variable: ITE and REC accuracy (%)

	Df	SS	F	P
Model 1: EMB1				
Task	1	0.02	0.6	0.45
EMB1 accuracy (%)	1	0.17	5.4	0.02*
EMB1* Task	1	0.06	2.0	0.15
Model 2: EMB2				
Task	1	0.02	0.6	0.45
EMB2 accuracy (%)	1	0.06	1.9	0.17
EMB2* Task	1	0.12	3.9	0.05*

We ran two Linear Mixed Models (LMM), one for EMB1 and another for EMB2, to test whether the visual tasks differed in how much they are predicted by EMB1 and EMB2. We found that EMB1 predicted both REC and ITE, with no significant difference between tasks (top). Conversely, EMB2 predicted better REC than ITE (bottom, see main text for details).

SS = sum of squares.

*P < 0.05.

with REC than with ITE, we ran linear mixed models with accuracy (%) in the visual task as the dependent variable and predicted by Task (ITE versus REC), and by the covariate 'accuracy' in grammar comprehension EMB(x) (for $x = 1$ and 2), and the two-way interactions EMB(x) \times Task. We ran two similar models, one for each grammar embedding depth (EMB1 and EMB2) (Table 1).

For EMB1, the best model (REML = 27.2) included grammar comprehension only, but no effect of Task, and no interaction. This means that comprehension of sentences with one level of embedding predicted equally well both REC and ITE accuracy, as suggested by the correlation analyses (a similar result was found for EMB0, as reported in Supplementary Table 4). On the other hand, the best model for EMB2 was the full model, including the interaction EMB2: Task. This means that the correlation between EMB2 and visual task accuracy differed between REC and ITE, with this relationship being stronger for REC (REC:EMB2, $B = 0.004$, $SD = 0.002$, $t = 1.9$). To investigate whether this effect could be caused by outliers, we calculated Cook's distances (Cook and Weisberg, 1982) and found all were lower than 1.06. We repeated the analysis removing the data points with highest Cook's distances (threshold of four times the mean) and obtained the same results (Supplementary Fig. 4 and Supplementary Table 5).

Finally, considering that general cognitive abilities, such as attention, could potentially account for these differences between REC and ITE, we inspected the correlation matrix between our visual tasks and a number of standard cognitive measures, including verbal and spatial working memory and measure of alertness as a basic function of attention (see Supplementary Fig. 5 for details). Including these variables in our model had no influence on the specific relationship between REC and EMB2 (Supplementary Table 6).

Together, these results suggest that while grammar comprehension correlates with both REC and ITE, as previously shown in (Martins *et al*, 2014b), for higher-levels of

sentence centre-embedding this correlation becomes specific for REC and not for ITE.

Discussion

The ability to represent hierarchies with multiple levels of embedding is an essential component of human cognition. In language, this capacity has been extensively investigated with both functional MRI and lesion studies highlighting the importance of a network comprising an anterior and a posterior 'hub' (in IFG and posterior temporal lobe, respectively) (Friederici, 2009, 2011; Hagoort and Indefrey, 2014; Matchin *et al.*, 2017). The exact function of these hubs remains ambiguous, and it is unclear whether they support mechanisms specific to language or more generally the processing of hierarchical structures (Matchin *et al.*, 2017; Matchin and Rogalsky, 2018).

Here, we report the first study investigating the acquisition of RHE in the visual domain in patients with an acquired chronic brain lesion in the language network. We find that, despite the brain lesion, participants were able to acquire a recursive regularity in a sequence of four steps. The generation of new hierarchical levels in visuo-spatial structures was compared to the ability to acquire an iterative rule that sequentially added visual elements within a fixed hierarchy, without generation of new levels. The presence of a circumscribed chronic left hemispheric brain lesion in all participants enabled us to perform lesion-behaviour analyses probing into whether left hemispheric neuronal structures support the ability to infer recursive visual processes, as they do for language.

Participants performed the task rather well with substantial inter-individual variance. With regard to the neuronal underpinnings, lesions in the left (posterior) middle temporal lobe correlated with lower performance in the detection of the recursive process. As this applies to the drift parameter of our analysis, the critical deficits affected information accumulation and integration during decision-making. Less robustly, lesions in IFG decreased boundary separation, in other words, patients with a lesion in this area tended to acquire less information before deciding how the hierarchy generating rule continued. The latter can be conceived as a lower threshold at which participants are confident to respond correctly. Regarding the issue of supramodality of RHE, we compared the visual task to a linguistic task performed by most participants. We found that the ability to process single embedded sentences correlated equally well with REC and ITE performance. Conversely, patients performing worse on the comprehension of double embedded sentences performed worse specifically in REC.

These data, together with previous literature, suggest that pTC is important for the formation of RHE representations across different domains. In the next sections, we will discuss these findings in the broader context of hierarchical cognition.

The contrast REC–ITE isolates recursive hierarchical embedding

The essential difference between REC and ITE is that only for REC a self-similarity between the global structure and subordinate elements evolves along increasingly complex images. On the contrary, for ITE, the addition of more elements follows a simple sequential rule. This parallels differences in language when multiple embedded propositions are compared to a serial presentation of the same propositions. [As an example: ‘The mouse the cat the dog bit chased escaped’ versus ‘the dog bit the cat, the cat chased the mouse, the mouse escaped’.]

Previous research has demonstrated that performance for REC is correlated with the ability to represent recursive embedding in tonal hierarchies and in action planning, when factoring out overall performance in both tasks (Martins *et al.*, 2017). This suggests that the cognitive resources used in the acquisition of RHE representations are shared across domains. This hypothesis would be consistent with our behavioural results showing that accuracy in REC correlates (more strongly than ITE) with the ability to adequately parse sentences with two centre-embedded clauses. Moreover, the areas which correlate with the derived parameters of drift and boundary (Fig. 4 and Table 1)—IFG

and posterior temporal—are considered hubs for the processing of complex grammar (Friederici, 2011; Hagoort and Indefrey, 2014; Zaccarella *et al.*, 2017).

Additional support for the fundamental difference between ITE and REC comes from the analysis of order effects in the current experiment. The alternating order of the two tasks across patients was introduced to counterbalance mere learning effects; however, we find a strong task-order effect for REC only. Performance in REC, when performed prior to ITE, was significantly worse than for the inverse task order. On the contrary, ITE performance did not change depending on task-order. These findings replicate previous results with children (Martins *et al.*, 2014b). The fact that task-order effects are not present for ITE suggests that there is a ‘natural sequence’ in the acquisition of RHE representations in that acquisition of recursion requires previous acquisition of more simple iterative representations. This has also been shown for the domain of language in which children need to acquire conjunctive representations before they are able to acquire the construction of subordinate clauses (Roeper, 2011). This effect is influenced by exposure and inductive processes and not only by natural ontogenetic development (Dewar and Xu, 2010; Perfors *et al.*, 2011a, b). Our data showing an asymmetric effect of Order for REC and ITE is additional evidence that REC

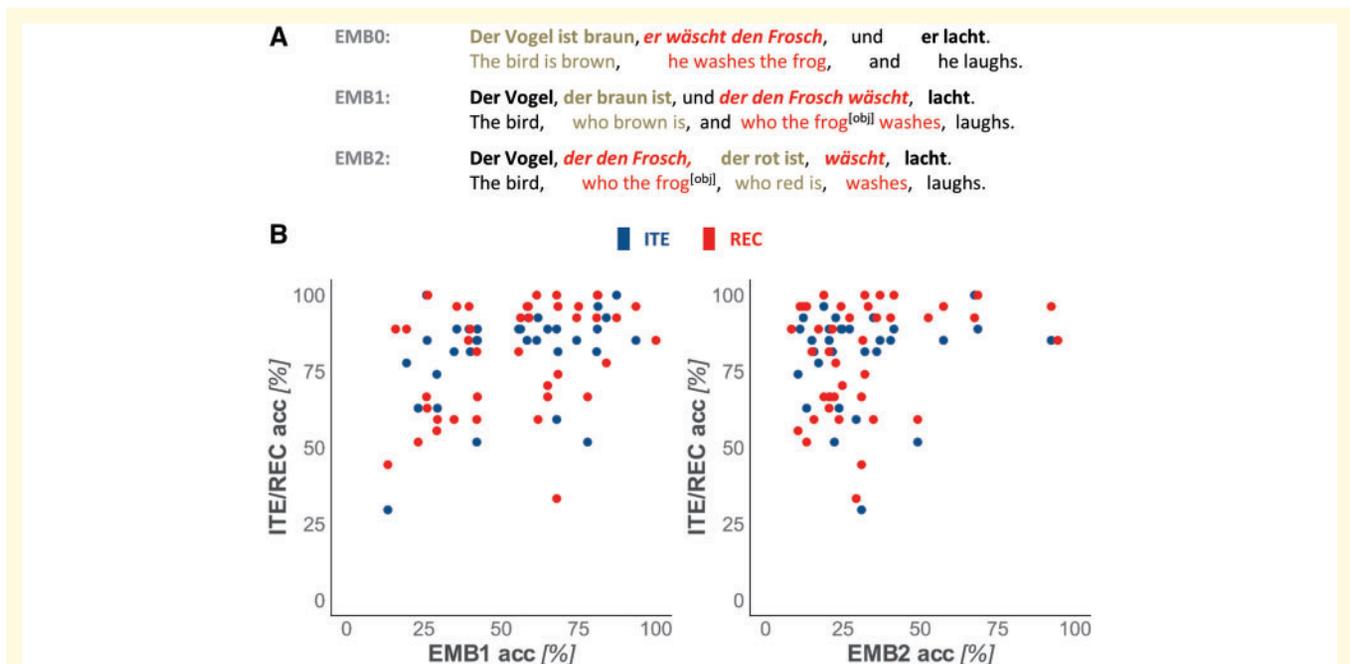


Figure 4 Relationship between REC, ITE and grammar task. (A) Example of the linguistic task of a different study in which the majority of the participants took part. Regarding the tasks reported here (ITE/REC) performance for one aspect of syntactic complexity, namely embedding, was correlated with results in the visual task. The three syntactic propositions were presented either sequentially (E0, i.e. no embedding) or with an embedded relative clause (EMB1, one embedding containing two propositions) or with a 2-fold centre-embedded structure (EMB2). Note that for the three example sentences (out of $n = 132$) the same image would be correct in the successive picture selection task (see screenshot in Supplementary Fig. 2). The task required selection of the correct picture from a set of four (one correct and three distractors for each proposition). (B) Scatterplots depicting the relationship between accuracy in the visual tasks (REC and ITE) and in the grammar tasks EMB1 (left) and EMB2 (right). Correlation coefficients are presented in the text. See also Table 1.

builds on shared resources with ITE while also adding a specific representation layer. The current results and previous research suggest that this layer pertains to the simultaneous representation of multiple hierarchical levels.

The temporal cortex supports recursive hierarchical embedding representations

Lesions in posterior temporal cortex were associated with slower integration of information during processing of RHE. The velocity of information integration during decision making (drift rate) has been proposed to reflect the availability of memory representations that guide the accumulation of perceptual information in a top-down manner, as shown for instance in lexical priming (Voss *et al.*, 2013; Mulder *et al.*, 2014). As our participants were naïve to the visual stimuli and rules, this suggests that the formation of RHE representations is crucial to process multiple hierarchical levels simultaneously. A central neuronal hub for this capacity seems to reside in posterior temporal cortex.

Our findings are partially consistent with language research showing an involvement of the posterior temporal cortex in top-down syntactic prediction and lexical access (Matchin *et al.*, 2017). A recent study also demonstrated that the posterior MTG is a common area of activity during syntactic processing in language and music (Musso *et al.*, 2015). Overall, different parts of the temporal cortex have been shown to support processing of RHE in well-trained participants in different domains: the representation of tonal hierarchies mostly activates the anterior STG (Martins, 2017), language the left posterior STG (Friederici, 2011; Hagoort and Indefrey, 2014; Matchin *et al.*, 2017), and the visual domain more anterior portions of the MTG (Martins *et al.*, 2014a).

As the MTG is associated with semantic memory, these findings invite the speculation that the bottleneck for the capacity to acquire RHE is less dependent on general multi-demand fronto-parietal systems (as also shown in Duncan, 2010; Fedorenko *et al.*, 2012, 2013), or on specialized areas instantiating RHE ‘computations’ such as those proposed for language (e.g. BA44; Friederici *et al.*, 2011; Zaccarella *et al.*, 2017), but rather on the formation of RHE ‘representations’. In support to this interpretation, posterior temporal cortex is found to be active in both semantics and syntax comprehension in children younger than 7 years of age, while IFG becomes active only at the age of 10 (Skeide *et al.*, 2014). Thus, while posterior temporal gyrus plays a significant role in the acquisition of RHE, IFG is active during automatic and expert processing (Jeon and Friederici, 2013).

Finally, we found IFG to be associated with boundary separation, which could be more related with controlled retrieval of existing representations. Interestingly, our results suggest a division of labour in the processing of RHE between the anterior and the posterior hubs, which

is somewhat consistent with the language model in which domain-general cognitive control systems operate with domain-specific representations (Matchin *et al.*, 2017).

Limitations and perspectives

First, in this study we have included only patients with lesions in cortical areas in the left hemisphere. Therefore, we are not able to determine whether the right MTG is equally important in the acquisition of RHE representations. In our previous functional MRI results with similar visual recursive tasks we found bilateral brain activity, but no hemisphere-specific regions. While we cannot make claims about the uniqueness of left hemisphere in the processing of RHE, we can conclude that the left pTC is crucial to instantiate RHE representations in both vision and language. Future studies should address potential differences between hemispheres by our results with performance in a comparable sample of participants with an acquired right hemispheric lesion.

Second, while the MTG is important in the acquisition of RHE representations, we cannot determine whether learning mechanisms supported by subcortical structures are involved in this process. It is possible that the episodic memory (supported by hippocampus) or procedural systems (supported by basal ganglia) are also fundamental to build RHE representations. Future research including patients with lesions or functional impairment in these systems will be crucial to evaluate these hypotheses.

Finally, the current theory-driven investigation in mildly affected patients does not serve an apparent clinical goal. Nonetheless, the demonstration of supramodality of a cognitive process such as RHE supports integrative interdisciplinary cognitive rehabilitation to be a promising and exciting avenue in research (Cahana-amitay and Albert, 2015).

Conclusion

In this study, we hypothesized that the acquisition of RHE representations in the visual domain were supported by neural areas known to be involved in the processing of hierarchical structures in language (IFG and pTC). We tested a group of patients with chronic acquired left hemisphere lesions with a set of tasks designed to isolate the ability to acquire RHE in vision. Crucially, these patients had not been exposed to these tasks prior to this study. We found that lesions in posterior MTG specifically impaired the ability to adequately integrate information about RHE during task decision making. This area might be fundamental for the acquisition of RHE representations in vision and across domains.

Acknowledgement

We thank Paolo Baldi for his valuable support and suggestions.

Funding

This study was supported by the Max Planck Society.

Competing interests

The authors report no competing interests.

Supplementary material

Supplementary material is available at *Brain* online.

References

- Ashburner J, Friston KJ. Unified segmentation. *Neuroimage* 2005; 26: 839–51.
- Bates D, Martin M, Bolker B, Walker S. lme4: Linear Mixed-Effects Models Using Eigen and S4. R Package Version 1.0–6 [Internet]. 2014. <http://cran.r-project.org/package=lme4>.
- Bates E, Wilson SM, Saygin AP, Dick F, Sereno MI, Knight RT, et al. Voxel-based lesion-symptom mapping. *Nat Neurosci* 2003; 6: 448.
- Berwick RC, Chomsky N. *Why only us: language and evolution*. Cambridge: The MIT Press; 2015.
- Berwick RC, Friederici AD, Chomsky N, Bolhuis JJ. Evolution, brain, and the nature of language. *Trends Cogn Sci* 2013; 17: 98.
- Brett M, Leff AP, Rorden C, Ashburner J. Spatial normalization of brain images with focal lesions using cost function masking. *Neuroimage* 2001; 14: 486–500.
- Cahana-Amitay D, Albert ML. Neuroscience of aphasia recovery: the concept of neural multifunctionality. *Curr Neurol Neurosci Rep* 2015; 15: 41.
- Cook RD, Weisberg S. *Residuals and influence in regression*. New York: Chapman and Hall; 1982.
- Dehaene-Lambertz G, Dehaene S, Hertz-Pannier L. Functional neuroimaging of speech perception in infants. *Science* 2002; 298: 1013–15.
- Dewar KM, Xu F. Induction, overhypothesis, and the origin of abstract knowledge: evidence from 9-month-old infants. *Psychol Sci* 2010; 21: 1871–7.
- Duncan J. The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends Cogn Sci* 2010; 14: 172–9.
- Durstewitz D, Koppe G, Toutounji H. Computational models as statistical tools. *Curr Opin Behav Sci* 2016; 11: 93–9.
- Everaert MBH, Huybregts MAC, Chomsky N, Berwick RC, Bolhuis JJ. Structures, not strings: linguistics as part of the cognitive sciences. *Trends Cogn Sci* 2015; 19: 729–43.
- Fedorenko E, Duncan J, Kanwisher N. Report language-selective and domain-general regions lie side by side within Broca's area. *Curr Biol* 2012; 22: 2059–62.
- Fedorenko E, Duncan J, Kanwisher N. Broad domain generality in focal regions of frontal and parietal cortex. *Proc Natl Acad Sci* 2013; 110: 16616–21.
- Fischmeister FP, Martins MD, Beisteiner R, Fitch WT. Self-similarity and recursion as default modes in human cognition. *Cortex* 2017; 97: 183–201.
- Fitch WT, Martins MD. Hierarchical processing in music, language, and action: Lashley revisited. *Ann N Y Acad Sci* 2014; 1316: 87–104.
- Freedman D, Lane D. A nonstochastic interpretation of reported significance levels. *J Bus Econ Stat* 1983; 1: 292–8.
- Friederici AD. Pathways to language: fiber tracts in the human brain. *Trends Cogn Sci* 2009; 13: 175–81.
- Friederici AD. The brain basis of language processing: from structure to function. *Physiol Rev* 2011; 91: 1357–92.
- Friederici AD, Bahlmann J, Friedrich R, Makuuchi M. The neural basis of recursion and complex syntactic hierarchy. *Biolinguistics* 2011; 5: 87–104.
- Gerstein GL, Mandelbrot B. Random walk models for the spike activity of a single neuron. *Biophys J* 1964; 4: 41–68.
- Grodzinsky Y, Santi A. The battle for Broca's region. *Trends Cogn Sci* 2008; 12: 474–80.
- Hagoort P, Indefrey P. The neurobiology of language beyond single words. *Annu Rev Neurosci* 2014; 37: 347–62.
- Huber W, Poeck K, Willmes K. The Aachen aphasia test. *Adv Neurol* 1984; 42: 291–303.
- Jeon H-A, Friederici AD. Two principles of organization in the prefrontal cortex are cognitive hierarchy and degree of automaticity. *Nat Commun* 2013; 4: 2041.
- Joliot M, Jobard G, Naveau M, Delcroix N, Petit L, Zago L, et al. AICHA: an atlas of intrinsic connectivity of homotopic areas. *J Neurosci Methods* 2015; 254: 46–59.
- Kravitz DJ, Saleem KS, Baker CI, Mishkin M. A new neural framework for visuospatial processing. *Nat Rev Neurosci* 2011; 12: 217–30.
- Martins MD. Distinctive signatures of recursion. *Philos Trans R Soc Lond B Biol Sci* 2012; 367: 2055–64.
- Martins MD. The cognitive architecture of recursion: behavioral and fMRI evidence from the visual, musical and motor domains. In: *Proceedings of the 39th Annual Conference of the Cognitive Science Society*. London, 2017.
- Martins MJ, Bianco R, Sammler D, Villringer A. Recursion in action: An fMRI study on the generation of new hierarchical levels in motor sequences. *Hum Brain Mapp* 2019; 40: 2623–38.
- Martins MD, Fischmeister FP, Puig-Waldmüller E, Oh J, Geißler A, Robinson S, et al. Fractal image perception provides novel insights into hierarchical cognition. *Neuroimage* 2014a; 96: 300–8.
- Martins MD, Gingras B, Puig-Waldmüller E, Fitch WT. Cognitive representation of “musical fractals”: Processing hierarchy and recursion in the auditory domain. *Cognition* 2017; 161: 31–45.
- Martins MD, Laaha S, Freiberger EMEM, Choi S, Fitch WT. How children perceive fractals: hierarchical self-similarity and cognitive development. *Cognition* 2014b; 133: 10–24.
- Martins MD, Martins IP, Fitch WT. A novel approach to investigate recursion and iteration in visual hierarchical processing. *Behav Res Methods* 2016; 48: 1421–42.
- Matchin WG. A neuronal retuning hypothesis of sentence-specificity in Broca's area. *Psychon Bull Rev* 2018; 25: 1682–94.
- Matchin WG, Hammerly C, Lau E. The role of the IFG and pSTS in syntactic prediction: evidence from a parametric study of hierarchical structure in fMRI. *Cortex* 2017; 88: 106–123.
- Matchin WG, Rogalsky C. *Aphasia and Syntax*. In: *The Handbook of Experimental Syntax*; 2018. p. 1–54.
- Mulder MJ, van Maanen L, Forstmann BU. Perceptual decision neurosciences—a model-based review. *Neuroscience* 2014; 277: 872–84.
- Musso M, Weiller C, Horn A, Glauche V, Umarova R, Hennig J, et al. A single dual-stream framework for syntactic computations in music and language. *Neuroimage* 2015; 117: 267–83.
- Perfors A, Tenenbaum JB, Griffiths TL. A tutorial introduction to Bayesian models of cognitive development. *Cognition* 2011a; 120: 1–61.
- Perfors A, Tenenbaum JB, Regier T. The learnability of abstract syntactic principles. *Cognition* 2011b; 118: 306–38.
- Pinker S, Jackendoff R. The faculty of language: what's special about it? *Cognition* 2005; 95: 201–36.
- Ratcliff R, McKoon G. The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Comput* 2008; 20: 873–922.
- Roeper T. The acquisition of recursion: how formalism articulates the child's path. *Biolinguistics* 2011; 5: 57–86.
- Rorden C, Brett M. Stereotaxic display of brain lesions. *Behav Neurol* 2000; 12: 191–200.

- Rogalsky C, Rong F, Saberi K, Hickok G. Functional anatomy of language and music perception: temporal and structural factors investigated using functional magnetic resonance imaging. *J Neurosci* 2011; 31: 3843–52.
- Rorden C, Karnath HO, Bonilha L. Improving lesion-symptom mapping. *J Cogn Neurosci* 2007; 19: 1081–8.
- Russell L. emmeans: Estimated Marginal Means, Aka Least-Squares Means. R Package Version 1.2.1 [Internet]. 2018. <https://cran.r-project.org/package=emmeans>.
- Skeide MA, Brauer J, Friederici AD. Syntax gradually segregates from semantics in the developing brain. *Neuroimage* 2014; 100: 106–111.
- Smith PL, Ratcliff R. Psychology and neurobiology of simple decisions. *Trends Neurosci* 2004; 27: 161–8.
- Srinivasan SK, Sampath G. Stochastic models for spike trains of single neurons (Vol. 16). Berlin: Springer Science & Business Media; 2013.
- Venables WN, Ripley BD. Modern Applied Statistics with S [Internet]. 2002. <http://www.stats.ox.ac.uk/pub/MASS4>.
- Voss A, Rothermund K, Jena F, Gast A, Wentura D, Saarlands U. Cognitive processes in associative and categorical priming: a diffusion model analysis. 2013; 142: 536–59.
- Wheeler B, Torchiano M. lmPerm: Permutation Tests For Linear Models [Internet]. 2016. <https://github.com/mtorchiano/lmPerm>.
- Wiecki T V., Sofer I, Frank MJ. HDDM: hierarchical Bayesian estimation of the drift-diffusion model in python. *Front Neuroinform* 2013; 7: 1–10.
- Wilson B, Marslen-Wilson WD, Petkov CI. Conserved sequence processing in primate frontal cortex. *Trends Neurosci* 2017; 40: 72–82.
- Winkler AM, Ridgway GR, Webster MA, Smith SM, Nichols TE. Permutation inference for the general linear model. *Neuroimage* 2014; 92: 381–97.
- Zaccarella E, Meyer L, Makuuchi M, Friederici AD. Building by syntax: the neural basis of minimal linguistic structures. *Cereb Cortex* 2017; 27: 411–21.