Cognitive Psychology 77 (2015) 20–41

Contents lists available at ScienceDirect

Cognitive Psychology

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

Representing visual recursion does not require verbal or motor resources

Maurício de Jesus Dias Martins [a,b,c,d,*], Zarja Muršič [a], Jinook Oh [a], W. Tecumseh Fitch [a]

[a] Department of Cognitive Biology, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria
[b] Berlin School of Mind and Brain, Humboldt Universität zu Berlin, Luisenstrasse 56, 10117 Berlin, Germany
[c] Max Plank Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1a, 04103 Leipzig, Germany
[d] Language Research Laboratory, Lisbon Faculty of Medicine, Av. Professor Egas Moniz, 1649-028 Lisbon, Portugal

ARTICLE INFO

Article history:
Accepted 28 January 2015

Keywords:
Hierarchy
Recursion
Self-embedding
Fractals
Language

ABSTRACT

The ability to form and use recursive representations while processing hierarchical structures has been hypothesized to rely on language abilities. If so, linguistic resources should inevitably be activated while representing recursion in non-linguistic domains. In this study we use a dual-task paradigm to assess whether verbal resources are required to perform a visual recursion task. We tested participants across 4 conditions: (1) Visual recursion only, (2) Visual recursion with motor interference (sequential finger tapping), (3) Visual recursion with verbal interference – low load, and (4) Visual recursion with verbal interference – high load. Our results show that the ability to acquire and use visual recursive representations is not affected by the presence of verbal and motor interference tasks. Our finding that visual recursion can be represented without access to verbal resources suggests that recursion is available independently of language processing abilities.

1. Introduction

Humans are exceptional creatures. Our ability to form complex social structures, and to transform our environment is unprecedented in the animal kingdom. What makes us exceptional is our cognitive

* Corresponding author at: Berlin School of Mind and Brain, Luisenstrasse 56, 10117 Berlin, Germany.
E-mail address: mauricio.martins@univie.ac.at (M. de J.D. Martins).

http://dx.doi.org/10.1016/j.cogpsych.2015.01.004
0010-0285/

power: our ability to combine actions to achieve complex goals and to represent complex structures goes well beyond what is documented in any other animal species (Badre, 2008; Badre, Hoffman, Cooney, & D'Esposito, 2009; Conway, & Christiansen, 2001; Unterrainer, & Owen, 2006; Wohlschlager, Gattis, & Bekkering, 2003). Language, for example, requires the combination of words into sentences (Chomsky, 1957). The combinatorial processes involved in language are powerful and flexible, allowing us to generate an infinite number of meaningful sentences by combining a finite set of words (Hauser, Chomsky, & Fitch, 2002; von Humboldt, 1972).

Underlying the capacity to combine individual elements to form higher order structures is the concept of hierarchy. 'Hierarchy' can be used to denote a tree-like organization in structural representations where 'higher' levels incorporate multiple 'lower' levels. Language (Chomsky, 1957; Hauser et al., 2002), complex problem solving (Unterrainer, & Owen, 2006), and complex social navigation (Nardini, Jones, Bedford, & Braddick, 2008) all require the use and production of hierarchies (Fig. 1). For example, in action sequencing (Fig. 1C), the general goal of 'making coffee' is hierarchically superior, or 'dominant' over the specific actions of 'grinding the coffee beans' and 'filling the water container' (Jackendoff, 2002). Individuals can evaluate the need for these basic actions and omit them if they are unnecessary without impairing the overall procedure of making coffee (Badre, & D'Esposito, 2009).

Hierarchies can be generated and represented using processes that establish relationships of dominance and subordination between different items (Martins, 2012). Some of these processes are depicted in Fig. 2. For instance, 'iterative rules' (Fig. 2A) can be used to represent the successive addition of items to a structure, such as the addition of beads to a string to form a necklace. 'Embedding rules' can also be used to generate hierarchies by embedding one or more items into a structure so that they depend on another item (Fig. 2B). For example, in an army hierarchy, two brigades can be incorporated into a division. Finally, we can also use 'recursive embedding rules' to generate and represent hierarchies. Recursive embedding, or simply 'recursion', is the process by which we embed one or more items as dependents of another item of the *same category* (Fig. 2C). As we can see from Fig. 2,

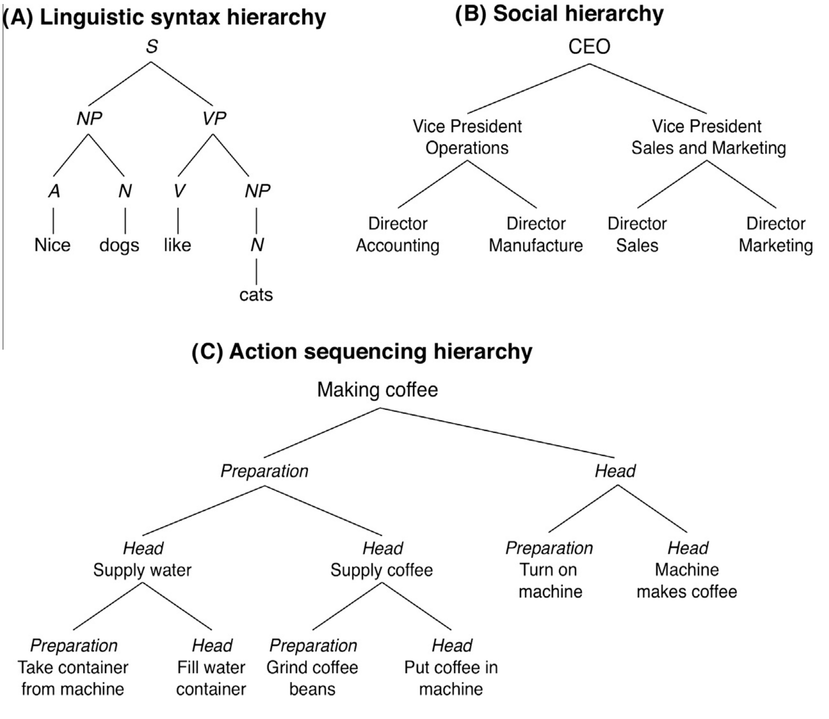

Fig. 1. Examples of linguistic (A), social (B) and action sequencing (C) hierarchies.

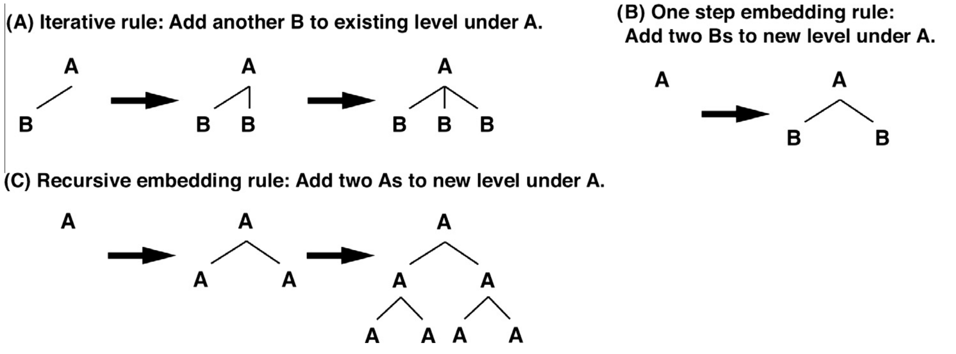

Fig. 2. Examples of processes that add elements to hierarchies. These processes can either generate new levels (b and c), or simply add elements to pre-existing levels (a). Recursion (c) can generate multiple hierarchical levels using the same single rule.

recursion is interesting and unique because it allows the generation of multiple hierarchical levels with a single rule. Hierarchies in which different levels share common properties, as in language (Chomsky, 1957, 2010; Fitch, Hauser, & Chomsky, 2005; Hauser et al., 2002), theory of mind (Miller, Kessel, & Flavell, 1970; Miller, 2009) and visuo–spatial objects (Martins, 2012), can be efficiently represented using recursion (Fig. 3). The ability to represent a recursive rule has been suggested as a necessary condition for the open-ended power of human cognition (Fitch et al., 2005; Hauser et al., 2002).

1.1. Recursion as a representational ability

In the research described here, we take an empirically oriented definition of recursion. This viewpoint is important because even though recursion has been proposed as a uniquely human capacity that gives rise to abilities such as language (Fitch et al., 2005; Hauser et al., 2002), prospective thinking (Corballis, 2011), and cooperation (Tomasello, 2008), scholars continue to disagree on its definition, and how it should be investigated.

One of the biggest sources of confusion concerning recursion derives from the fact that recursion can be defined either as a "procedure that calls itself" or as the property of "constituents that contain constituents of the same kind" (Fitch, 2010; Pinker, & Jackendoff, 2005). In general, we find an isomorphism between procedure and structure, i.e., recursive processes often generate recursive structures. However, this isomorphism does not always occur. Recursive structures can be generated by recursive processes as well as by non-recursive procedures (Fig. 4).

Given this dissociation, some authors have argued that looking at the ability to generate recursive structures is formally and empirically irrelevant (Lobina, 2011; Lobina, & Garcıa-Albea, 2009; Luuk, & Luuk, 2010; Watumull, Hauser, Roberts, & Hornstein, 2014). These authors propose that we should look at direct cognitive signatures of recursive algorithms, assuming that these are implemented in the brain in the same way that they are implemented in artificial computational systems.

For example, take the following recursive process that generates the natural numbers:

$$N_0 = 1$$

$$N_n = N_{i-1} + 1, \quad i = 1, 2, 3 \ldots n$$

Using this process, the number '4' would be defined as $N_4 = (((N_0 + 1) + 1) + 1)$. In computational terms, the recursive representation of the number 4 would require a memory stack storing the three addition operations that have to be performed in sequence in order to generate the number 4. This representation requires time and memory resources. If the brain implements processes the same

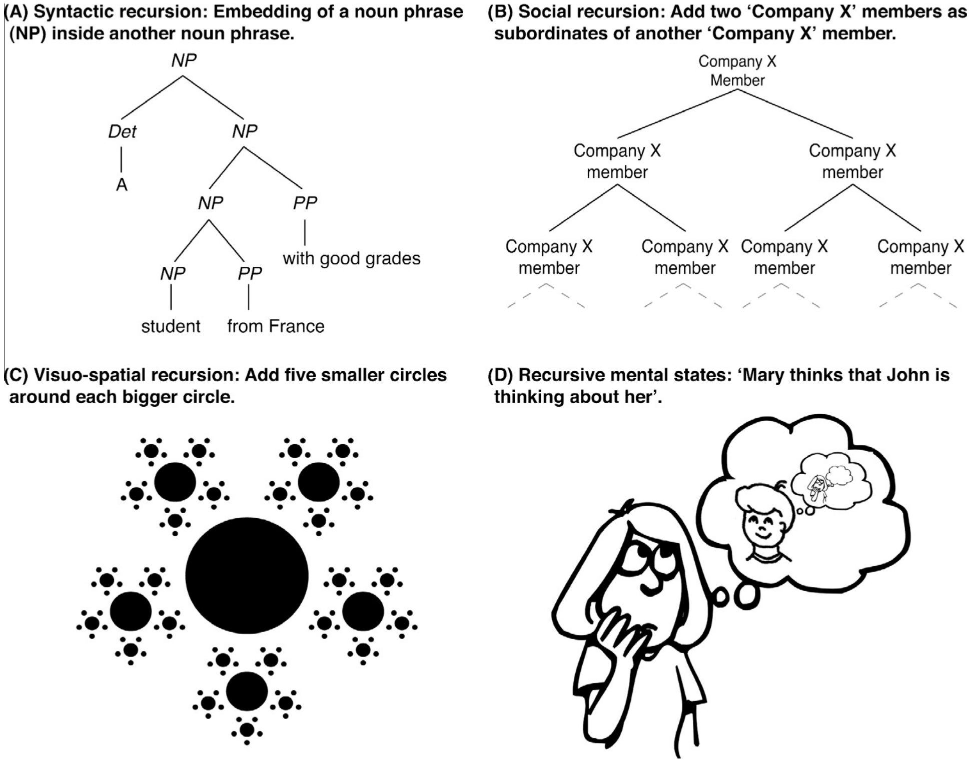

Fig. 3. Examples of structures that can be efficiently represented using recursive rules.

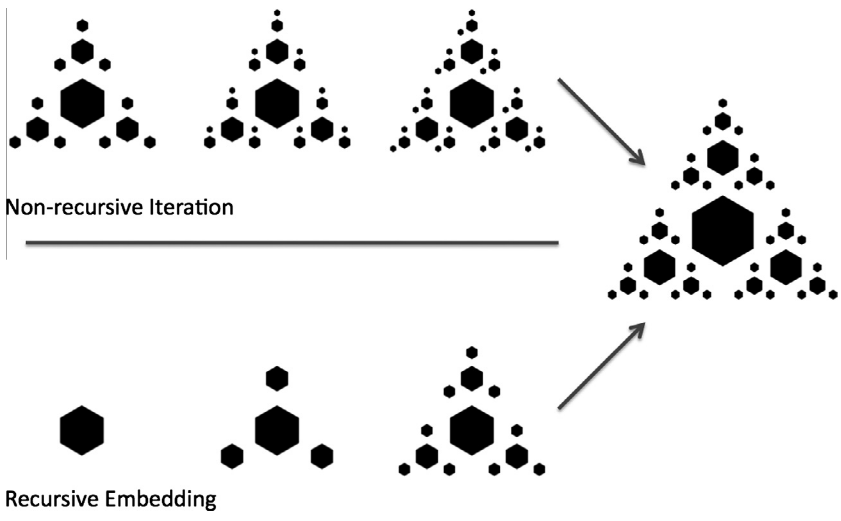

Fig. 4. Example of a recursive structure (on the right) that can be generated either by using a non-recursive procedure (top) or a recursive embedding process (bottom).

way as a computer, then we could use higher memory demands and slower reaction times as indicators of recursion (Lobina, 2011; Lobina, & Garcıa-Albea, 2009; Luuk, & Luuk, 2010; Watumull et al., 2014).

Elsewhere we have expressed our disagreement with this approach (Martins, 2012) and argued that trying to directly measure a relatively unexplored cognitive process such as recursion can be misleading. Most cognitive processes are opaque: we can have an idea of *what* kind of information is represented, but it is often the case that we cannot directly measure *how* it is represented. In fact, very few cognitive functions can be clearly assigned to specific neural and algorithmic processes.

Here, we endorse an empirical investigation of recursion based on detecting *what* kind of information individuals can represent, rather than on *how* this information is implemented. This approach is instantiated not by measuring the ability to generate recursive structures, but by detecting the ability to correctly continue unfamiliar recursive processes. In this approach, an individual able to detect hierarchical self-similarity from unfamiliar structures and to use this information to generate new hierarchical levels beyond the given is able to represent the idea of recursion and use it productively.

If individuals can represent the kind of information that allows the generation of multiple hierarchical levels using a single rule, then this would afford all the behavioral and evolutionary advantages of recursion (i.e. infinity from finite means); and this would be true even if the algorithms or physiological mechanisms used to implement this representation would not be recursive *de facto*.

1.2. Recursion and human language

Recursion has been an phenomenon of interest for scholars in many fields (Chomsky, 2010; Corballis, 2007; Eglash, 1997; Hauser et al., 2002; Hofstadter, 1980; Mandelbrot, 1977; Penrose, 1989), and has been associated with the unbounded character of human creativity and generative capacity. However, little is known about its psychological nature and biological implementation. While hierarchies in several domains, such as in music, language, motor sequencing, problem solving and architecture, can be described as being generated by recursive rules (Eglash, 1997, 1998; Eisenberg, 2008; Jackendoff, & Lerdahl, 2006; Miller, 2009; Pulvermüller, & Fadiga, 2010; Schiemenz, 2002), the human ability to represent hierarchical self-similarity in these different domains, that is, to what extent humans actually extract recursive principles while parsing self-similar structures, remains mostly untested. Furthermore, while there have been some attempts to describe recursion as a cognitive module akin to an encapsulated system in the brain, (Chomsky, 1995, 2010), there is no empirical evidence currently either supporting or challenging this view. Although the place of recursion in the broader human and animal cognitive architectures has been a topic of intense discussion, unfortunately there is little empirical data available outside of language to support any of the different current claims (Corballis, 2007; Fitch et al., 2005; Gentner, Fenn, Margoliash, & Nusbaum, 2006; Hauser et al., 2002; Jackendoff, & Pinker, 2005; Pinker, & Jackendoff, 2005).

One area in which empirical data has been collected concerns the relationship between recursion and human language. Recursion seems to be universally used in all languages (Reboul, 2012), and although rarely evidenced in common speech (Laury, & Ono, 2010), most speakers, regardless of their language, are likely to have generated several recursive structures in their lifetimes (in English, for instance, compound nouns such as "[[[student] film]] committee]"). Furthermore, children from an early age can extract the correct meaning from recursive sentences (Alegre, & Gordon, 1996; Roeper, 2009). These abilities, as yet undemonstrated in other cognitive domains, have led some authors to propose that the evolution of language may have been tightly connected with the availability of recursion. One influential hypothesis states that recursion is a domain-specific "linguistic computational system [...], independent of the other systems with which it interacts and interfaces" (Hauser et al., 2002). According to this hypothesis, although the usage of recursive rules may be available in non-linguistic domains such as visual art (Eglash, 1997), music (Jackendoff, & Lerdahl, 2006), architecture (Eglash, 1998), humor (Eisenberg, 2008), second-order theory of mind (Miller, 2009), problem solving (Schiemenz, 2002), or action sequencing (Pulvermüller, & Fadiga, 2010), these uses may rely upon a previously evolved system of abstract arbitrary symbol manipulation and may thus be dependent on the faculty of language. Alternatively, Pinker and Jackendoff have proposed that the usage of recursion in some domains, for example in visual perception, can occur independently of language (Pinker, & Jackendoff, 2005). Thus, the main hypotheses concerning the relationship between recursion and human language are the following:

Hypothesis 1. The ability to form recursive representations is specific to language and is implemented by a linguistic 'recursion module'. The representation of recursion in other domains depends on language, and therefore recruits linguistic resources.

Hypothesis 2. The ability to form recursive representations is domain-general. There is a single cognitive system implementing recursion, which can be recruited by several domains, with no primacy of language.

Hypothesis 3. The ability to build recursive representations is multiply-domain-specific, but not restricted to language. Each domain of cognition can access its own cognitive system able to implement recursive representations, independent from the other domains.

These three hypotheses are all logically possible and consistent with the scarce currently available empirical data. Although Hypothesis 2 could be criticized for being non-modular (Fodor, 1983; Hornstein, & Pietroski, 2009; Roeper, 2011), there are a number of other cognitive processes, for example those involved in central executive processing, which are implemented by specific neural systems, and yet are available for all domains of cognition (Baddeley, 1998; Fodor, 1983).

1.3. Empirical investigation of recursion

While recursion has been studied mostly within the linguistic domain, recent research has shown that humans are also able to acquire and apply recursive rules governing the generation of visuo–spatial hierarchies (Martins, 2012; Martins, Fischmeister, Puig-Waldmueller, Oh, Geissler, Robinson, & Beisteiner, 2014; Martins, & Fitch, 2012). This research suggested that the ability to acquire recursive rules in the visuo–spatial domain might crucially depend on the engagement of analytical and effortful cognitive strategies. Interestingly, compared with non-recursive iterative processes, visual recursive abilities only correlated weakly with specific visual resources (non-verbal intelligence, spatial short-term memory and spatial working memory), but correlated strongly with recursive planning tasks (Martins et al., 2014) and with the processing component of verbal working memory (Martins, & Fitch, 2012). However, the later finding does not necessarily entail that visuo–spatial recursion recruits verbal-specific resources. Instead this correlation may be driven by some third variable common to both domains, for example by cognitive resources comprising the central executive. Interestingly, a recent brain imaging study revealed that visual recursion does not specifically activate classical perisylvian language areas when compared with a simple iterative task (Martins et al., 2014). However, these data are correlational and need to be confirmed by procedures that causally manipulate the ability to use linguistic resources.

Verbal interference paradigms have been commonly used to block the use of linguistic resources, and to assess the causal role of language in non-linguistic cognitive abilities. These paradigms usually involve the repetition of simple syllables such as 'da' (Fatzer, & Roebers, 2012) or 'na' (Baldo, Dronkers, Wilkins, Ludy, Raskin, & Kim, 2005), the repeated production of fixed sequences such as 'a–b–c' (Emerson, & Miyake, 2003), or the repetition of arbitrary sequences of 7 or more digits (Lupyan, 2009; Winawer, Witthoft, Frank, Wu, Wade, & Boroditsky, 2007). In these studies, when the ability to use verbal resources was blocked by simple articulatory procedures, participants were impaired in tasks requiring perceptual categorization (Lupyan, 2009), exact numerosity (Frank, Fedorenko, Lai, Saxe, & Gibson, 2012), cognitive control (Emerson, & Miyake, 2003), executive functions (Fatzer, & Roebers, 2012), color discrimination (Winawer et al., 2007), and problem solving (Baldo et al., 2005). These results have been used to argue that language is either necessary or plays a significant role in these cognitive abilities. Mechanistically, verbal interference tasks might prevent lexical representations from feeding back onto lower-level representations (Lupyan, 2009). When linguistic abstractions are rendered inaccessible by verbal interference, language users fall back on non-linguistic cognitive strategies similar to those of children, animals and aphasics (Frank et al., 2012; Lupyan, 2009).

In the current experiment, our goal was to directly address the question of whether verbal resources are necessary to acquire and apply recursive rules in the visual domain. Participants had to perform a Visual Recursion Task (VRT) under conditions of verbal interference. If the ability to process recursive hierarchies in the visual domain is negatively influenced by verbal rehearsing of digits, then this would support the hypothesis that language plays a significant role in the representation of recursion in non-linguistic domains (Fig. 5, Hypothesis 1). If the ability to represent visual recursion were unaffected when linguistic resources are blocked then this would lend support to the idea that the visual domain can directly access the cognitive system of 'recursion', independently of language (Fig. 5, Hypotheses 2 and 3). To test the specificity of this potential effect, we included verbal conditions with varied memory loads, and a motor interference condition as a control, in which participants rehearsed motor sequences of finger movements.

In addition to the main experiment, we have included a secondary procedure to control for the usage of simple visual strategies based on low-frequency visual information. In this control experiment, we have used the same 48 stimuli as in the main experiment, but changed their visual features by using a spatial frequency low-pass filter. The goal was to test whether high spatial frequencies were necessary to solve the task, or whether general gestalt strategies using only "low-frequencies" sufficed. If the latter were true, that would argue for the usage of simple visual heuristics to solve VRT, and against the induction of abstract rules.

2. Methods

2.1. Participants

We tested 24 volunteers (18 females) aged between 19 and 35 ($M = 22.8$, $SD = 3.7$) who were recruited at the University of Vienna. All participants were right-handed, non-musicians, and either German native speakers ($n = 22$), or proficient in German for more than 5 years ($n = 2$). All participants were tested in the same room, with the same experimental apparatus and all reported normal or corrected-to-normal visual acuity. Participants were paid 8 Euros for participation and gave written informed consent.

2.2. Procedure

We used a dual-task paradigm in order to assess whether the recruitment of verbal resources is a necessary condition to represent recursion in the visual domain. The procedure involves a primary task (here: a visual recursion task) performed either in isolation, or simultaneously with a secondary interference task. If performance in the primary visual recursion task decreases in the presence of a

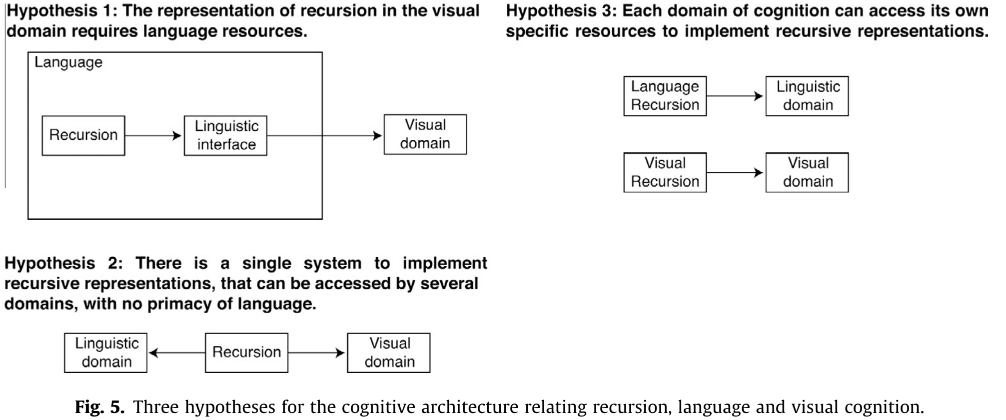

Fig. 5. Three hypotheses for the cognitive architecture relating recursion, language and visual cognition.

secondary verbal interference task, this would suggest that verbal recourses are required in order to solve visual recursion. Of course, an impaired VRT performance in the presence of verbal interference could also be due to general attention constraints. To evaluate this possibility we also included a non-verbal motor interference task (see details below).

The experiment took approximately 50 min. Initially, there was a training session, after which each participant completed four experimental sessions, each session comprising 12 trials: (1) VRT without secondary task ('none'); (2) VRT with motor interference ('motor'); (3) VRT with verbal interference – low load ('verbal low'); and (4) VRT with verbal interference – high load ('verbal high'). The order of conditions was balanced across participants. As there were 24 possible orders of conditions (see Supplemental Materials Table S1), each participant was tested using a different order. In total, each of the four conditions (none, motor, verbal low, and verbal high) appeared six times in each possible position (first, second, third, fourth).

At the beginning of each trial, participants were exposed to a specific secondary task (with different memory content): In the motor condition they were shown a series of pictures denoting finger movements, and asked to rehearse (by executing) the finger-tapping motor sequence; In the verbal conditions they were shown a sequence of digits and they were asked to continuously repeat it vocally. When ready, participants had to press a button to proceed to the VRT task. In this task, the VRT images were presented and participants had 10 s to provide an answer. After the primary response was provided, or after 10 s, a dialog box appeared, asking participants to repeat the motor or verbal sequence rehearsed throughout the trial. The responses to both primary and secondary tasks were recorded. The structure of a typical trial is depicted in Fig. 6.

The experimental apparatus is schematically depicted in Fig. 7. Participants sat in front of a computer screen on which visual stimuli were presented. With their left hand they provided VRT responses by pressing one of two buttons on a button box (ioLabs Systems®). With their right hand they gave responses to the secondary tasks: In the verbal interference condition they typed digits on a numeric keypad and in the motor interference condition they provided responses using a five-button button box (ShuttleXpress®). Each button was assigned to a specific finger.

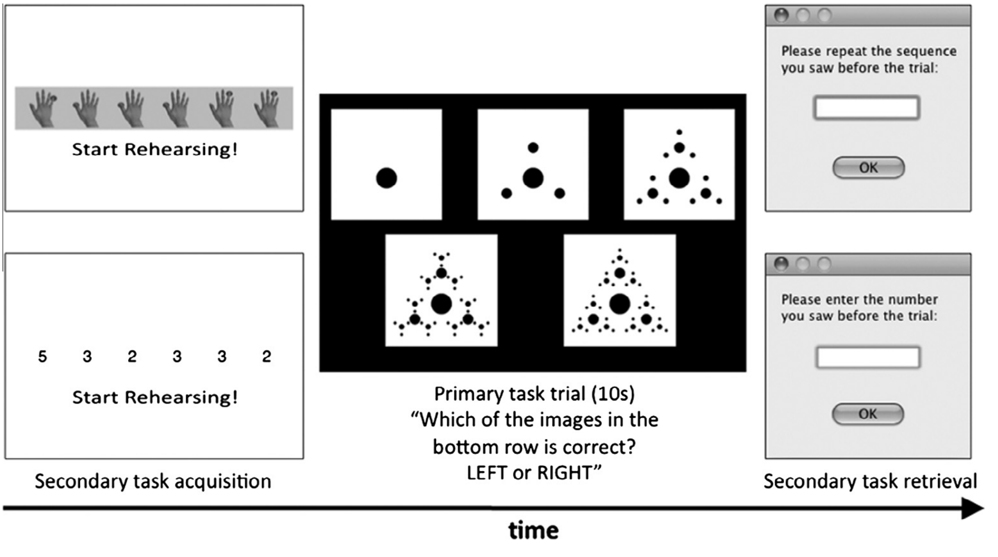

Fig. 6. Overview of the trial structure. In the beginning of the trial, participants were shown the secondary task memory content ('Secondary task acquisition'). Participants rehearsed a finger tapping motor sequence in the motor condition and a digit sequence in the verbal condition. Participants pressed a button once they were ready to advance to the primary task trial (VRT). After an answer to the recursion task was provided (or after 10 s), participants had to type the motor or verbal sequences rehearsed throughout the trial ('Secondary task retrieval').

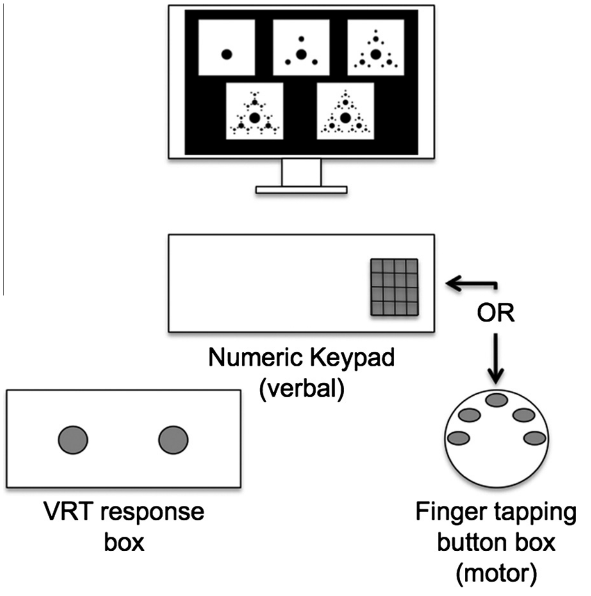

Fig. 7. Experimental setup. Participants sat in front of the screen. With their left hand they provided VRT responses by pressing one of the buttons on the VRT response box. With their right hand they gave responses to the secondary tasks. In the verbal interference condition they typed digits in a numeric keypad and in the motor interference condition they provided responses by pressing buttons on the finger-tapping button box.

2.3. Visual recursion task (VRT)

2.3.1. Stimulus generation

The stimuli used here were generated by the method described in (Martins, & Fitch, 2012) and were based on the properties of fractals. A pool of simple geometrical shapes served as *initiators*. Then, different kinds of recursive embedding rules (*generators*) were applied over these shapes in order to generate fractal structures. In our task, four iterative steps were generated for each fractal (Fig. 8). Different *generators* were used, which determined (a) the symmetry (bilaterally symmetrical vs. asymmetrical) and (b) the complexity of the resulting structures. Here 'visual complexity' refers to the number of elements added to the visual fractal (sets of three or four elements). The spatial coordinates of each set of elements were calculated, based on the coordinates of a previously existing "higher-order" element (Fig. 8). Symmetrical and asymmetrical stimuli were included to increase the visual variability and prevent a strategy based only on symmetry.

In addition to the first four iterations, a foil structure was generated for each fractal. This foil structure corresponded to an "incorrect" fourth iteration, generated by applying a rule for the fourth iteration that differed from the one used to generate the first three iterations. There were 3 types of foils, depending on the process used in their generation (Fig. 8): (i) 'Odd constituent foil': one element within each set of 3 or 4 elements within the lower visual scale was misplaced; (ii) 'Positional error foil': a novel positional scheme for all new added elements of the fourth iteration was employed; (iii) 'Repetition foil': The third image was simply repeated.

We used different foils in order to discourage participants from applying simple heuristic strategies based on the comparison between the 'correct' and 'incorrect' fourth iterations, strategies which could be unrelated to the recursive rule itself. For example, a simple-minded similarity-based comparison strategy would not allow participants to correctly solve the task in the 'repetition foil' condition, as the incorrect image was identical to the third iteration.

The combination of symmetry (symmetrical and asymmetrical), visual complexity (3 and 4) and foil categories (positional, odd, repetition) resulted in 12 types of stimuli. Exactly 4 examples of each type of stimuli were generated using Nobebox (http://nodebox.net/), an open source application using

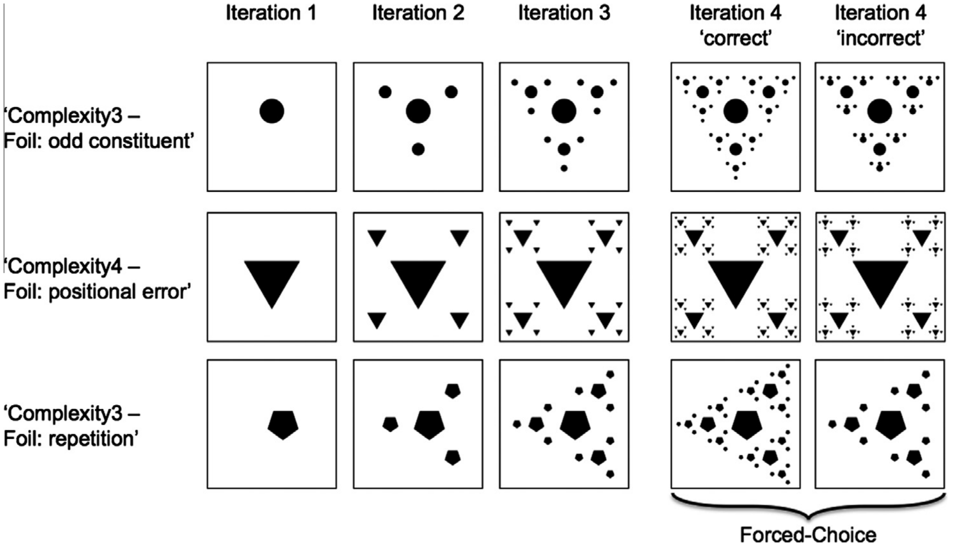

Fig. 8. Examples of fractals used in the visual recursion task. The first four iterations of a fractal generation, as well as one foil ('incorrect' fourth iteration), were produced. There were two categories of 'visual complexity' (using either 3 or 4 elements in a set) – and different categories of foils: 'Odd constituent', 'Positional error' and 'Repetition'.

Python programming code (www.python.org), resulting in a total of 48 stimuli. Stimulus categories were balanced across testing conditions (none, motor, verbal low, and verbal high). Each testing condition contained exactly one example of each stimulus type.

2.3.2. The visual recursion task

Each trial began with the sequential presentation of three images corresponding to the first three iterations (steps) of a fractal generation on the top half of the screen (Fig. 9, top), appearing with an interval of 500 ms between images. After this sequence, two images were presented simultaneously on the bottom half of the screen (Fig. 9, bottom) for forced choice, and the previous three remained visible. One choice image always corresponded to the correct continuation of the recursive process that generated the first three fractals, and the other corresponded to a foil. Participants were asked to press one of two buttons in a button box (ioLabs Systems®), corresponding to the position of the image they considered to be the correct continuation of the recursive process. The image positions on the screen (LEFT or RIGHT) were randomized. No response feedback was given during testing. In order to control for the use of global visual strategies, based solely on low-frequency spatial information, we decomposed choice images using Python Fast Fourier Transform functions (Bradski, 2000), and compared them regarding their average power density at "low frequencies" below 6 cycles/image (standard cutoff, e.g. (Mahon, Kumar, & Almeida, 2013; Vuilleumier, Armony, Driver, & Dolan, 2003). We found that correct and incorrect images did not significantly differ in their average power at low frequencies ($p > 0.1$, Supplementary Materials Fig. S2.1), rendering the use of gestalt strategies implausible.

2.4. Secondary verbal and motor tasks

Participants performed VRT either alone or with one of three interference tasks: motor interference, verbal interference – low load, and verbal interference – high load. In the 'sequential motor tapping task' subjects were shown a sequence of 6 pictures denoting 6 finger-tapping movements (Fig. 6, top left), which included all five fingers of the right hand. These images were simultaneously presented on the screen. Tapping sequences were randomly generated for each trial. Participants were instructed to repeatedly execute the sequence and to press a button when ready to proceed to VRT.

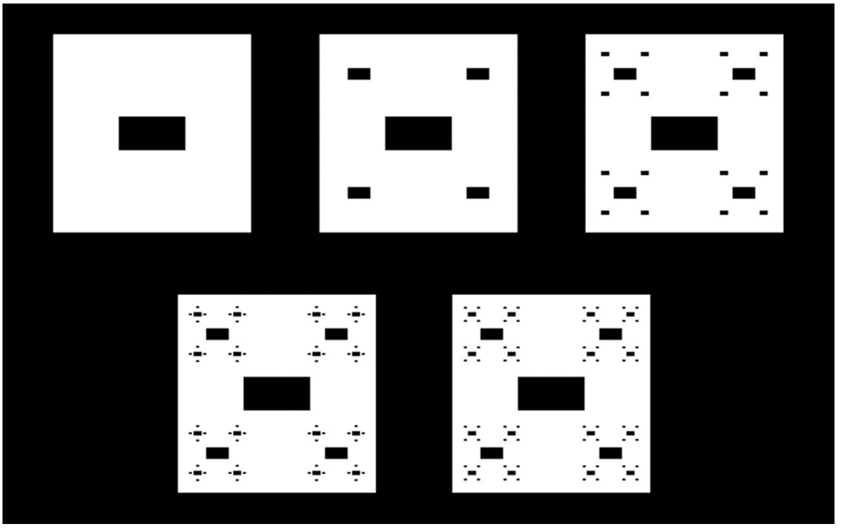

Fig. 9. A visual recursion task trial. Initially, the first three iterations of a fractal generation were depicted sequentially from left to right (top). Then, two images were presented simultaneously on the bottom half of the screen, corresponding to the 'correct' fourth iteration (bottom right) and a foil (bottom left). From these images, participants had to choose which corresponded to the correct fourth iteration (LEFT or RIGHT).

Here we did not restrict the rehearsal time, since was a great variability in the speed of learning motor sequences across participants. On average, participants rehearsed 13 s, 24 s and 38 s in verbal low, verbal high and motor conditions, respectively. A similar procedure was used by (Lupyan, 2009), with positive interference results, which suggests that providing unlimited rehearsal time does not prevent the interference of the secondary task on the primary task. Our participants were instructed to repeat the sequence with their right hand during the total duration of a VRT trial, using no other cognitive (e.g. verbal) or physical resources but their fingers. After an answer was provided to the VRT, or after the 10 s timeout, participants were instructed to type the motor sequence they had been rehearsing.

The verbal interference task was based on digit span, a verbal working memory task. In this task, participants were visually presented a sequence of digits, and asked to vocally repeat the sequence, while performing a VRT trial (Fig. 6 bottom). After each trial they were asked to type the sequence in the keyboard. A new random digit sequence was generated for each VRT trial.

In the 'verbal low' condition, participants had to memorize a randomly-generated sequence of 6 digits, ranging from '1' to '5', which matched the information load presented in the 'sequential motor tapping task'. In the 'verbal high' condition, participants had to memorize a sequence of 7 digits, ranging from '1' to '9', increasing memory load.

2.5. Training

Participants underwent two short training sessions prior to the experimental procedure to become familiar with the experimental apparatus, and the task requirements.

The training session for VRT consisted of four trials with a series of images, similar to VRT. However, the first three items followed simple non-recursive iterative rules of incremental complexity, and the last item followed a recursive rule (see Fig. 10 for an example). During training, no visual or auditory feedback was provided, however.

Training for both digit span and sequential motor tapping consisted of ten trials of each condition in which participants performed the procedure described in Fig. 3, but without the primary task (VRT). After the sequence of digits (or finger movements) was presented on the screen, participants rehearsed (repeated) the sequence while attending to a blank screen for 10 s. Then, they were asked

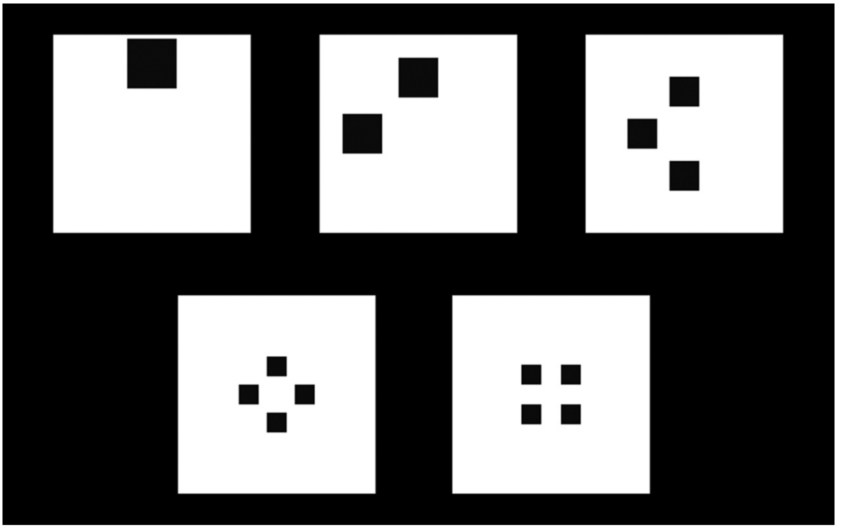

Fig. 10. Example of an item included in the training task for VRT. Images follow an iterative, not recursive, rule. The correct choice is on the bottom left.

to type the sequence. The order of the training was the same for all participants (VRT first, digits span second, and sequential motor tapping third).

2.6. Analysis

In order to assess the effects of task-condition (none, motor, verbal low, and verbal high), while controlling for the effects of the session-position in the procedure (first, second, third, fourth), we used Generalized Estimating Equations (GEE), a semiparametric regression technique. This technique is useful when analyzing binomial data with within-subjects effects (Hanley, 2003).

Our goal was to assess whether human adults could represent visual recursion under conditions of verbal and motor interference. To do so, we compared the number of correct responses and reaction times as dependent variables, across conditions. As responses were binary, we pooled subjects and used a Binomial test to assess whether the overall performance in each condition significantly differed from chance. With 288 trials per condition (12 trials $\times$ 24 participants), a number of correct trials of at least 162 (i.e. a proportion of 0.56) was required for performance to be significantly above chance (Binomial test, $p = 0.04$). We performed the same analysis for foil categories within each condition (96 trials: 4 trials of each foil per condition $\times$ 24 participants). In this case, 59 or more correct trials (i.e. a proportion of 0.62) had to be attained for performance to deviate significantly from chance (Binomial test, $p = 0.03$). We also assessed performance, via recall accuracy, on the secondary task.

In order to test whether VRT trials tapped into a unified construct, we performed an internal reliability analysis using Cronbach's alpha (Cronbach, 1951). We also tested whether there was an effect of learning across trials by fitting accuracy and reaction time data to power curves (Anderson, 1982).

All analyses were performed with SPSS 19 (IBM).

3. Results

3.1. Visual recursion task analysis

Across all sessions only 42 out of 1152 trials (4%) timed out, and were classified as 'incorrect' in the analysis. At the group level, scores for all conditions were significantly above chance (Binomial test: all *p-values* < 0.001). The mean percentage of correct responses in VRT was 82% in the 'none' condition ($SD = 18$); 86% in the 'motor' condition ($SD = 16$); 83% in the 'verbal low' condition ($SD = 21$); and

86% in the ‘verbal high’ condition (*SD* = 21). Task-condition results are depicted in Fig. 11. At the group level, participants scored above chance in all sessions (Binomial test: all *p-values* < 0.001), see Fig. 11. The average percentage of correct answers in VRT was 76% in the first session (*SD* = 19); 84% in the second session (*SD* = 18); 89% in the third session (*SD* = 16); and 88% in the fourth session (*SD* = 20).

Pooling together all trials across all conditions, we found a very high level of internal consistency (*Cronbach’s alpha* = 0.91), suggesting that different trials were highly correlated with each other, across all conditions.

To test whether the presence of a secondary task had a significant effect on VRT performance (while controlling for the effects of session-position), we ran a GEE model with the binomial variable ‘trial correctness’ (correct/incorrect) as the dependent variable, and task-condition (none, motor, verbal low and verbal high) and session position (first, second, third, fourth) as within-subjects factors. Crucially, we found no effect of task-condition (*Wald* $\chi^2 = 4.9$, $p = 0.18$), indicating that the motor and verbal interference tasks did not significantly reduce VRT performance. However, we found a main effect of session position (*Wald* $\chi^2 = 13.8$, $p = 0.003$). Specifically, performance in the first session of the procedure was lower than in the other three sessions (all pair-wise comparisons $p < 0.015$, after sequential Bonferroni correction), indicating improved performance as the experiment went on. We also found a significant interaction between task condition and session position (*Wald* $\chi^2 = 33.2$, $p < 0.001$). Pairwise comparisons showed that scores in the first session of the procedure were low in the condition without interference (Fig. 11), both in comparison with the same condition in other positions (none-first vs. none-third: *mean difference* = 0.14, $p = 0.011$, after sequential Bonferroni correction) and in comparison with several other task conditions: motor-fourth, verbal low-third, verbal low-fourth, verbal high-third, verbal high-fourth (all $p < 0.05$, after sequential Bonferroni correction).

To further test for a learning effect, we assessed whether performance improved across all 48 trials (12 trials $*$ 4 conditions). We found that accuracy increased following a power curve ($F_{1,46} = 48.0$, $p < .001$; Supplementary Materials II, Fig. S3.1), suggesting that information was being consistently transferred from one trial to the next, and from one session to the other.

The effects of verbal and motor interference in visual recursion might have been masked due to a ceiling effect, since subjects scored very high in our task. Since response accuracy was on average lower in the first session, we compared performance between interference conditions including in the analysis only trials performed within the first session. We still found no significant differences between conditions (*Wald* $\chi^2 = 3.6$, $p = 0.3$).

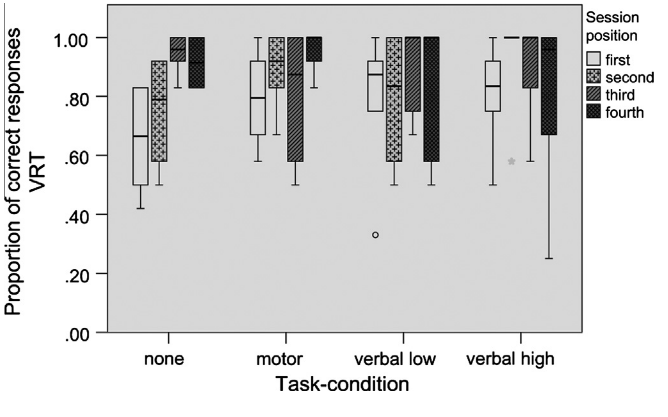

Fig. 11. Visual recursion task (VRT) performance across task-conditions and sessions. The presence of motor and verbal interference tasks did not impair VRT performance. However, participants scored lower in the first session than in the last three sessions. The boxplot divides the scores into quartiles, the ‘box’ represents the distance from the 25th percentile to the 75th percentile (interquartile range). Horizontal dark line is the median. ○ are outliers deviating from the box between 1.5 and 3 times the interquartile range; $*$ are outliers deviating from the box more than 3 times the interquartile range.

3.2. Visual recursion response time (RT) analysis

Timeouts were excluded from the analysis. On average in VRT trials, participants took 4.3 s [median = 3.9 s] to respond in the 'none' condition (*SD* = 0.6 s); 3.6 s [median = 3.0 s] in the 'motor' condition (*SD* = 1.8 s); 3.7 s [median = 3.2 s] in the 'verbal low' condition (*SD* = 1.8 s); and 3.4 s [median = 3.0 s] in the 'verbal high' condition (*SD* = 1.8 s), see Fig. 12. In relation to session position, the average RT was 4.7 s [median = 4.5 s] in the 'first' session (*SD* = 1.9 s); 3.7 s [median = 3.2 s] in the 'second' session (*SD* = 1.7 s); 3.4 s [median = 2.9 s] in the 'third' session (*SD* = 1.6 s); and 3.2 s [median = 2.8 s] in the 'fourth' session (*SD* = 1.7 s), see Fig. 12.

RT data was right skewed (skewness = 1.0) and not normally distributed (*Kolmogorov–Smirnov test* = 0.105, $p < 0.01$). We computed a new RT variable (logRT) by applying a log transformation, and achieved normality (*Kolmogorov–Smirnov test* = 0.025, $p = 0.11$). To analyze whether interference tasks influenced response time, while controlling for session position, we performed a GEE analysis with logRT as the dependent variable. We found an effect of task condition (*Wald* $\chi^2 = 21.9$, $p < 0.01$): specifically, RT was *longer* in the condition without interference than in the conditions with verbal and motor interference (all *p*-values < 0.01, after sequential Bonferroni correction). We also found an effect of session position (*Wald* $\chi^2 = 68.1$, $p < 0.001$). Specifically, RT in the first session was longer than in the other three sessions, and RT in the second session was longer than in the fourth session (all pair-wise comparisons $p < 0.05$, after sequential Bonferroni correction). There was an interaction between task condition and session position (*Wald* $\chi^2 = 18.8$, $p = 0.03$). In the 'verbal low' condition, RT in the first session was significantly higher than in the other three sessions (all *p*-values < 0.01, after sequential Bonferroni correction); and in the 'verbal high' condition, RT in the first session was significantly higher than in the third and fourth sessions (all *p*-values < 0.01, after sequential Bonferroni correction). We did not find differences between session positions within 'none' and 'motor' conditions (all $p > 0.05$). Finally, in the third session, RT was lower in 'verbal high' than in the condition without interference ($p = 0.044$). Crucially, interference conditions did not affect RT within first, second or fourth sessions (all $p > 0.05$).

To further test for a learning effect, we assessed whether RT decreased across all 48 trials (12 trials $*$ 4 conditions). We found that RT followed a power curve ($F_{1,46} = 230.8$, $p < .001$; Supplementary Materials II, Fig. S3.2), again suggesting that information was being consistently transferred from one trial to the next, and from one session to the other.

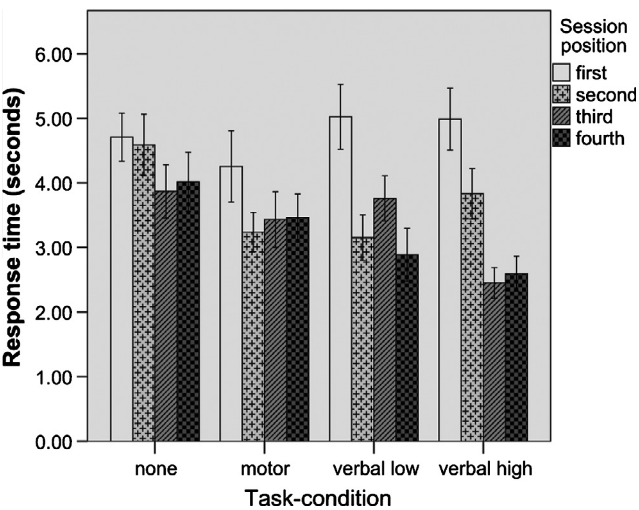

Fig. 12. Response times across task-conditions and across sessions. Overall, motor and verbal interference tasks decreased VRT response times significantly and participants had longer response times in the first of the four sessions. Bars reflect standard error.

3.3. Visual strategies

One question in this experiment was whether participants would be able to represent the structural self-similarity of the recursive images and apply this knowledge throughout VRT trials, or whether they would resort to alternative strategies that did not involve recursion. A possible alternative strategy to representing self-similarity would be to use visual heuristic strategies based on the detection of simple salient features within the foils, rather than recursive features within the fractals. In order to prevent the emergence of any systematic 'choice-by-exclusion' strategy, we used different categories of foils. If participants were able to represent self-similarity, they should perform adequately across all different foil categories.

At the group level, the number of correct choices was significantly above chance for all foil categories and for all task-conditions (Binomial test: all p-values < 0.01, see Fig. 13), which suggests that no single heuristic was used to solve VRT in any of the task-conditions. Crucially, our participants were clearly not applying a simple similarity analysis between choice-images and previous iterations, since they correctly rejected the repetition foils, which are identical to the third and final exemplar, but nonetheless incorrect.

3.4. Secondary task analyses

On average, participants' rate of correctly recalled sequences was 69% in the 'motor' condition ($SD = 46$), 85% in the 'verbal low' condition ($SD = 36$), and 79% in the 'verbal high' condition ($SD = 41$). The likelihood of correct recall differed significantly between task-conditions (*Wald* $\chi^2 = 17.2$, $p < 0.001$). All pairwise comparisons were statistically significant ($p < 0.01$, after Bonferroni correction). Overall, participants found it harder to recall finger-tapping sequences than verbal sequences, and sequences with higher verbal load were harder to recall than sequences with low verbal load.

It is possible that some participants either ignored the instructions to rehearse motor or verbal sequences during the secondary tasks or were unable to correctly recall the sequences. To control for potential effects of secondary task performance, we analyzed whether trials with correct secondary responses (in which we could be certain that verbal and motor rehearsal occurred) were associated with lower VRT performance. We calculated separate mean scores for VRT trials with correct vs.

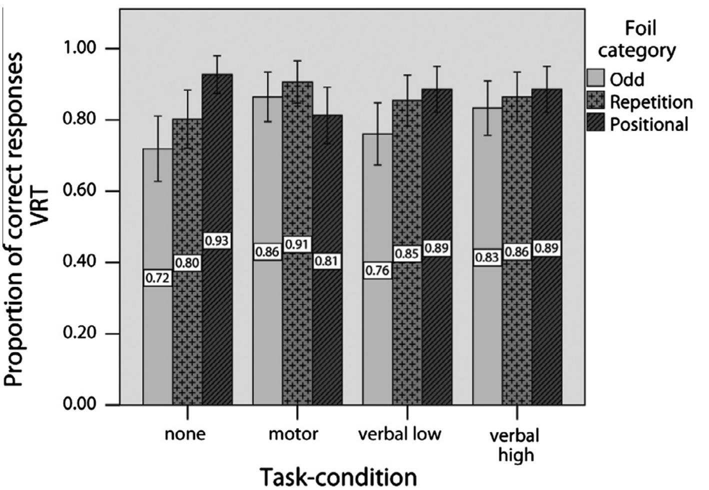

Fig. 13. Average VRT performance across task-conditions for all foil categories. Participants scored above chance in all foil categories across all conditions, suggesting that, on a group level, no single heuristic was used to solve VRT. Bars reflect standard error.

incorrect secondary task responses. Results are depicted in Fig. 14. Surprisingly, the proportion of correct trials in VRT was *higher* when the secondary task response was also correct ($\chi^2 = 19.9$, $p < 0.001$). Thus, participants continued to perform well in the VRT while successfully executing the secondary and motor tasks, and their ability to correctly repeat digit and finger tapping sequences correlated with *increased* performance in visual recursion.

We also repeated the analysis of VRT performance across conditions, on the subset of trials with correct secondary task responses. Results were similar to the previous model: The main effect of task-condition was still not significant (*Wald* $\chi^2 = 6.1$, $p = 0.1$), while there was a similar significant main effect of session position (*Wald* $\chi^2 = 25.5$, $p < 0.001$).

3.5. Control procedure for low-spatial frequencies

We performed a control procedure to assess whether it was possible to solve VRT without abstract rules, using a strategy based on the assessment of gestalt information, or overall spatial configuration. For this purpose we took all 48 stimuli of our pool and filtered out their high-frequency spatial information. Using standard Python scripts for image processing (Bradski, 2000), we transformed our stimuli using a low-pass filter (keeping only spatial $\leqslant 6$ cycles/image, similar to the procedure used in Mahon et al., 2013; Vuilleumier et al., 2003). Then, we tested 6 participants across 4 sessions of 12 stimuli each, using these transformed stimuli (e.g. Fig. 15). All sessions were similar to the condition "no interference" above.

On average, the percentage of correct answers was 46.2% ($SD = 12.5$). Performance was at chance both at the group level and at the individual level (all *Binomial tests*, $p > .2$). This was true for all experimental sessions (all *Binomial tests* $p > .1$; Fig. 16), suggesting that no learning occurred.

These results clearly demonstrate that the use of low-frequency spatial information does not account for VRT performance in the main experiment.

4. Discussion

We found no evidence of interference from either verbal or motor secondary tasks on our primary visual recursion task. This clearly suggests that the ability to acquire and use principles governing the generation of recursive (self-similar) visuo–spatial hierarchies is uninfluenced by secondary tasks in the verbal and motor domains of cognition. We controlled for potential confounds such as cross-trial

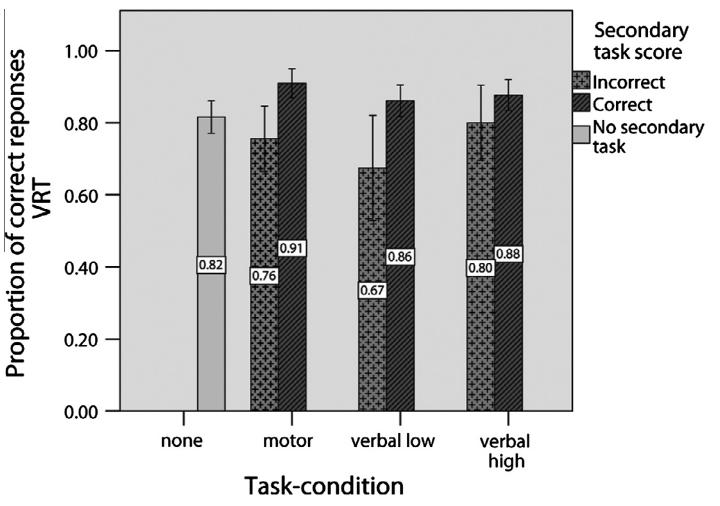

Fig. 14. VRT performance across task-conditions, showing trials with correct and incorrect secondary task scores separately. Trials with correct secondary tasks responses yielded better performance in VRT. Bars reflect standard error.

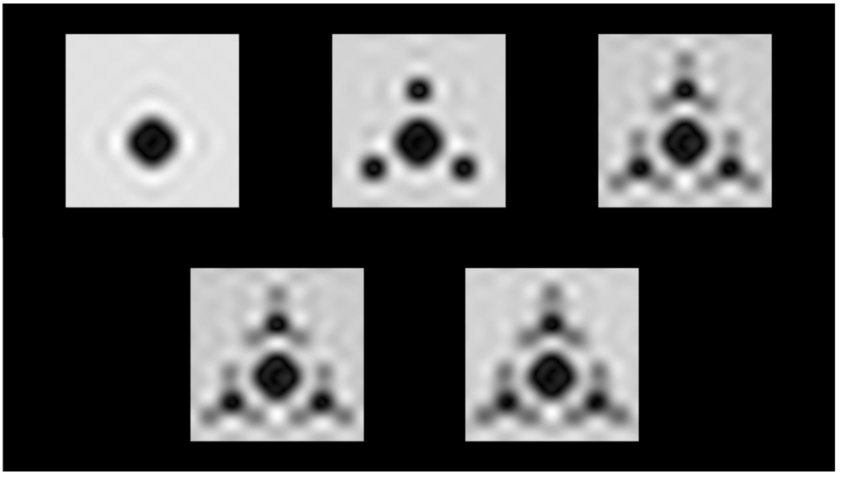

Fig. 15. Example of VRT stimulus after removing high-frequency spatial information.

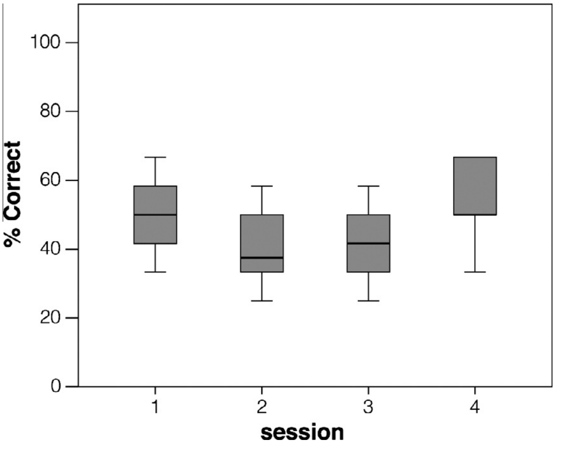

Fig. 16. Average percentage of correct answers in VRT across four sessions, using stimuli without high-frequency spatial information.

learning effects, usage of simple visual heuristic strategies, or potential impairments in the ability to rehearse verbal and motor sequences.

These results have several implications for our understanding of recursion.

First, human adults were able to extract principles governing the recursive generation of visuo–spatial hierarchies, and generalize this structural information to other recursive examples. The fact that performance increased with practice and with no response feedback, argues strongly in favor of some rule induction or a generalization process (Gordon, & Holyoak, 1983). Furthermore, each participant was exposed to 48 different stimuli, organized across 12 categories, differing in symmetry, visual complexity and foil categories. The fact that performance was consistent across different foil categories, and that these different items were well correlated, suggests that no single simple-minded visual heuristic strategy was used.

Second, we used an interference-task procedure to assess whether the recruitment of verbal and motor resources was necessary to acquire and use recursive hierarchical information in the visuo–spatial domain. Each participant performed a task of visual recursion under four conditions, including verbal and motor interference (none, motor, verbal – low load, verbal – high load). We found equally high performance in the VRT with no interference as with secondary verbal and motor tasks, suggesting that success in our visual task does not require the usage of online verbal or motor cognitive resources. We obtained the same results when only trials with correct motor and verbal responses to the secondary task were analyzed. Interestingly, the correct rehearsal of verbal and motor interference content seemed to promote rather than diminish performance in visual recursion. We think that the presence of a secondary task may force participants to more consciously direct their focus on the primary task, perhaps priming them to engage in a more effortful and analytical mode of cognition (Kahneman, 2011). Previous research with VRT (Martins, & Fitch, 2012) suggests that slow conscious engagement enhances the acquisition of recursive rules and the subsequent performance on our VRT. Alternatively, simultaneous failure of primary and secondary tasks may result from a general cognitive control failure, i.e., if participants try and fail to rehearse the primary task content, this may result in a general disruption of attention resources. (Morey, & Cowan, 2004) report similar effects of secondary tasks in other visuo–spatial experiments.

4.1. Differences between conditions with and without interference

Response times in the first session were similar across conditions. However, these decreased markedly with practice in verbal conditions, but not in the condition without interference. As VRT became easier, participants might have shortened their response times in the conditions with interference in order to reduce the rehearsing effort. Importantly, response time reduction was not associated with a decrease in VRT performance, which suggests that verbal content does not interfere with the acquisition and application of recursive information. Crucially, in the session in which participants were initially naïve to recursive information (first session), there were no differences between interference conditions in either response time or accuracy levels: the lack of interference was not due to a ceiling effect.

One interpretation of these findings is that participants might have used different cognitive strategies to perform VRT with and without interference. Indeed performance seemed to be more homogeneous across foil categories with than without interference. Increased cognitive load has been shown to affect the strategies used in the acquisition of schemas (Sweller, 1988), which might have relevance for our task. Our experiment was not designed to address this question, but this is an interesting topic for future research.

4.2. Theoretical implications for models of recursive cognition

Our goal was to address the question whether verbal resources are required for the acquisition and application of recursive rules in the visual domain. Underlying this question was Hypothesis 1: the ability to build recursive representations is language-domain-specific, the representation of recursion in other domains being parasitic on language (Fitch et al., 2005; Hauser et al., 2002). In evolutionary terms, this view would entail that the emergence of a language domain-specific module of recursion would be tightly related to the emergence of the faculty of language.

Recursion is not the only cognitive operation potentially dependent on linguistic resources. It has been proposed that having words for particular concepts helps to explain cognitive differences between different human populations with different languages, in a variety of domains. These include color, number, navigation, theory of mind, and object individuation (Frank, Everett, Fedorenko, & Gibson, 2008; Frank et al., 2012; Gordon, 2004; Pica, 2004; Pyers, & Senghas, 2009). Languages may help their speakers to create abstractions for the efficient processing and storage of information (Frank et al., 2012). In addition to these long-lasting transformations that language acquisition may induce, it appears that linguistic resources continue to be directly accessed in non-linguistic tasks through adulthood (Lupyan, 2009). These effects could be mediated by enhancement of thought fixation (Goldstein, 1948), modulation of conceptual representations (Deak, 2003), etc.

Previous research indicates that verbal interference tasks can prevent lexical representations from feeding back onto lower-level representations (Lupyan, 2009). When linguistic abstractions are rendered inaccessible by verbal interference, language users fall back on non-linguistic cognitive strategies similar to those of children, animals and aphasics (Frank et al., 2012; Lupyan, 2009). As reviewed in the introduction, these effects have been shown for a variety of domains (Baldo et al., 2005; Emerson, & Miyake, 2003; Fatzer, & Roebers, 2012; Frank et al., 2012; Lupyan, 2009; Winawer et al., 2007). However, our results here show that this is not the case for visuo–spatial recursion.

One potential criticism of this study could be that the tasks used in this experiment (both primary and secondary) were not challenging enough to generate visible effects. However, repeating the letters ‘a–b–c’ have been shown to interfere with cognitive control and task-switching (Emerson, & Miyake, 2003). Since we used a standard 7-digits memory load, our results are not likely to be explained by insufficient secondary memory load. Second, we found low accuracy scores and high response times in the first session of the procedure. However, even in this challenging session we did not find any differences between conditions with and without interference.

Taken together, previous positive findings of verbal interference in several domains and our negative results for visual recursion make Hypothesis 1 seem unlikely. Instead, our results support Hypotheses 2 or 3, which both state that recursion can be represented independently of language. The question of whether recursion is a single domain-general cognitive system (Hypothesis 2), or simply an umbrella term for several domain-specific independent modules (Hypothesis 3), is an interesting topic for future research. Using methods similar to those presented here, but using recursive interference tasks (instead of simple articulatory suppression), it should be possible to investigate whether the ability to represent recursion is similar in different domains (linguistic, visuo–spatial, etc.) or whether different domains compete for access to the same cognitive resources.

Recursion is an exciting concept to study that can benefit greatly from continued systematic investigation of its psychological bases in the future. Whether recursion turns out to be multiply-domain-specific or domain-general, broader insights into the nature of recursion will contribute to an improved understanding of the cognitive and biological origins of human generative capacity.

4.3. Limitations

In the development of a novel task purporting to tap into a specific cognitive construct for which there is no gold standard (here visual recursion), it is always hard to assure adequate internal and external validity. These methodological issues and our approaches to address them have been reported elsewhere (Martins, 2014). However, in the current study, we have used similar control techniques.

For instance, it could argued that solving VRT might not require the representation of recursion (Lobina, 2014; Watumull et al., 2014), or that recursion as measured here is of a quite different nature from language recursion (therefore not licensing conclusions about the interaction between the two domains).

Regarding the first issue, participants might have solved our task either by using simple visual heuristic strategies associated with the processing of gestalt or low-frequency spatial information, or by using idiosyncratic strategies for each trial. Both these results would argue against our conclusions.

Here we have tried to control for these factors in a number of ways. First, we have included different foil categories to prevent simple minded-strategies, and shown that participants scored adequately in all categories. Second, we have shown that VRT trials, including all foil and stimulus categories, were highly correlated with each other (Cronbach’s alpha = 0.91), suggesting that the task as a whole taps into a single cognitive construct. This argues against the use of idiosyncratic strategies for each trial. Third, we have shown a strong learning effect across sessions and across trials, both for accuracy and reaction time. Again this strongly indicates that participants induced a common rule across VRT and utilized this rule to solve new trials (Dewar, & Xu, 2010). Fourth, we performed spatial frequency analyses in our choice stimuli, and found that correct and incorrect stimuli had the same power for low spatial frequencies. This precludes the use of low-frequency information to solve the task, and excludes the hypothesis that simple gestalt strategies could suffice to perform adequately in VRT. In addition, we performed a control study using stimuli without high-frequency spatial information and shown that participants scored at chance, again demonstrating that fined-grained visual information was necessary. Finally, we know from previous research (Martins, & Fitch, 2012; Martins et al., 2014) that the ability to perform adequately in VRT is neither explained by general intelligence, nor by general visual resources such as spatial working memory and non-recursive iterative cognition (even though the latter task uses the exact same fractal patterns). However, VRT is specifically correlated with Tower of Hanoi (Martins et al., 2014), which has been argued to invite a recursive solution (Anderson, & Douglass, 2001; Goel, & Grafman, 1995; Halford, Wilson, & Phillips, 1998, 2010).

Even though we cannot be certain of the precise strategy our participants used to solve VRT, taken together, our results strongly argue in favor of a common rule being induced to solve different trials. In addition, previous research suggests this rule might be specifically related with the representation of recursion. This tell us that even if recursion as measured here is of a different nature than the mathematical notion put forward by (Hauser et al., 2002), and re-iterated by (Lobina, 2014; Watumull et al., 2014), our results can still be problematic to the hypothesis that the human open-ended generative capacity is afforded by recursion and requires language resources. If we accept our result that humans can extend visual hierarchies through several (potentially infinite) levels using a single rule, then two possible interpretations follow. Either we take that recursion can exist independently in language and vision, or that "infinity from finite means" (von Humboldt, 1972) can be achieved in vision by processes other than recursion (*sensu* Hauser et al., 2002; Lobina, 2014; Watumull et al., 2014). Either interpretation would falsify Hypothesis 1 or severely reduce its importance in the understanding of *hominin* cognitive evolution.

Acknowledgments

This research was supported by the FCT Grant SFRH/BD/64206/2009 to MM and by ERC Advanced Grant SOMACCA, project number 230604, and Grant "Research Cluster: Shared Neural Resources for Music and Language" to WTF.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cogpsych.2015.01.004.

References

Alegre, M. A., & Gordon, P. (1996). Red rats eater exposes recursion in children's word formation. *Cognition, 60*, 65–82.
Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*, 369–406.
Anderson, J. R., & Douglass, S. (2001). Tower of Hanoi: Evidence for the cost of goal retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(6), 1331–1346. http://dx.doi.org/10.1037/0278-7393.27.6.1331.
Baddeley, A. (1998). The central executive: A concept and some misconceptions. *Journal of the International Neuropsychological Society, 4*, 523–526.
Badre, D. (2008). Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes. *Trends in Cognitive Sciences, 12*(5), 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? *Nature Reviews Neuroscience, 10*(9), 659–669. http://dx.doi.org/10.1038/nrn2667.
Badre, D., Hoffman, J., Cooney, J. W., & D'Esposito, M. (2009). Hierarchical cognitive control deficits following damage to the human frontal lobe. *Nature Neuroscience, 12*(4), 515–522. http://dx.doi.org/10.1038/nn.2277.
Baldo, J., Dronkers, N., Wilkins, D., Ludy, C., Raskin, P., & Kim, J. (2005). Is problem solving dependent on language? *Brain and Language, 92*(3), 240–250. http://dx.doi.org/10.1016/j.bandl.2004.06.103.
Bradski, G. (2000). The opencv library. *Dr. Dobb's Journal of Software Tools*. <http://docs.opencv.org/trunk/doc/py_tutorials/py_imgproc/py_transforms/py_fourier_transform/py_fourier_transform.html>.
Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
Chomsky, N. (1995). *The minimalist program*. MIT Press.
Chomsky, N. (2010). Some simple evo devo theses: How true might they be for language? In R. K. Larson, V. M. Déprez, & H. Yamakido (Eds.), *The evolution of human language* (pp. 45–62). New York: Cambridge University Press.
Conway, C. M., & Christiansen, M. H. (2001). Sequential learning in non-human primates. *Trends in Cognitive Sciences, 5*(12), 539–546.
Corballis, M. C. (2007). Recursion, language, and starlings. *Cognitive Science, 31*(4), 697–704. http://dx.doi.org/10.1080/15326900701399947.
Corballis, M. C. (2011). *The recursive mind*. Princeton: Princeton University Press.
Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–333.

Deak, G. (2003). The development of cognitive flexibility and language abilities. In R. Kail (Ed.), *Advances in child development and behavior* (pp. 271–327). San Diego: Academic Press.
Dewar, K. M., & Xu, F. (2010). Induction, overhypothesis, and the origin of abstract knowledge: Evidence from 9-month-old infants. *Psychological Science, 21*(12), 1871–1877.
Eglash, R. (1997). Bamana sand divination: Recursion in ethnomathematics. *American Anthropologist, 99*(1), 112–122.
Eglash, R. (1998). Fractals in African settlement architecture. *Complexity, 4*(2), 21–29.
Eisenberg, M. (2008). Recursion—or, better computational thinking through laughter. *International Journal of Computers for Mathematical Learning, 13*, 171–174.
Emerson, M., & Miyake, A. (2003). The role of inner speech in task switching: A dual-task investigation. *Journal of Memory and Language, 48*, 148–168.
Fatzer, S. T., & Roebers, C. M. (2012). Language and executive functions: The effect of articulatory suppression on executive functioning in children. *Journal of Cognition and Development, 13*(4), 454–472. http://dx.doi.org/10.1080/15248372.2011.608322.
Fitch, W. T., Hauser, M., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition, 97*(2), 179–210. http://dx.doi.org/10.1016/j.cognition.2005.02.005.
Fitch, W. T. (2010). Three meanings of "recursion": Key distinctions for biolinguistics. In R. K. Larson, V. M. Déprez, & H. Yamakido (Eds.), *The evolution of human language* (pp. 73–90). New York: Cambridge University Press.
Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: Bradford.
Frank, M. C., Everett, D. L., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: Evidence from pirahã language and cognition. *Cognition, 108*(3), 819–824. http://dx.doi.org/10.1016/j.cognition.2008.04.007.
Frank, M. C., Fedorenko, E., Lai, P., Saxe, R., & Gibson, E. (2012). Verbal interference suppresses exact numerical representation. *Cognitive Psychology, 64*(1–2), 74–92. http://dx.doi.org/10.1016/j.cogpsych.2011.10.004.
Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature, 440*(7088), 1204–1207. http://dx.doi.org/10.1038/nature04675.
Goel, V., & Grafman, J. (1995). Are the frontal lobes implicated in "planning" functions? Interpreting data from the tower of Hanoi. *Neuropsychologia, 33*(5), 623–642.
Goldstein, K. (1948). *Language and language disturbances: Aphasic symptom complexes and their significance for medicine and theory of language*. New York: Grune & Stratton.
Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science, 306*(5695), 496–499. http://dx.doi.org/10.1126/science.1094492.
Gordon, P., & Holyoak, K. J. (1983). Implicit learning and generalization of the "mere exposure" effect. *Journal of Personality and Social Psychology, 45*(3), 492–500.
Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences, 21*, 803–865.
Halford, G. S., Wilson, W. H., & Phillips, S. (2010). Relational knowledge: The foundation of higher cognition. *Trends in Cognitive Sciences, 14*(11), 497–505. http://dx.doi.org/10.1016/j.tics.2010.08.005.
Hanley, J. A. (2003). Statistical analysis of correlated data using generalized estimating equations: An orientation. *American Journal of Epidemiology, 157*(4), 364–375. http://dx.doi.org/10.1093/aje/kwf215.
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.
Hofstadter, D. R. (1980). *Gödel, escher, bach: An eternal golden braid*. Harmondsworth: Penguin.
Hornstein, N., & Pietroski, P. (2009). Basic operations: Minimal syntax-semantics. *Catalan Journal of Linguistics, 8*, 113–139.
Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford University Press.
Jackendoff, R., & Lerdahl, F. (2006). The capacity for music: What is it, and what's special about it? *Cognition, 100*(1), 33–72.
Jackendoff, R., & Pinker, S. (2005). The nature of the language faculty and its implications for evolution of language (reply to fitch, hauser, and chomsky). *Cognition, 97*(2), 211–225. http://dx.doi.org/10.1016/j.cognition.2005.04.006.
Kahneman, D. (2011). *Thinking fast and slow*. New York: Farrar, Straus and Giroux.
Laury, R., & Ono, T. (2010). Recursion in conversation: What speakers of finnish and Japanaese know how to do. In H. van der Hulst (Ed.), *Recursion and human language*. de Gruyter Mouton.
Lobina, D. J. (2011). Recursion and the competence/performance distinction in AGL tasks. *Language and Cognitive Processes, 26*(10), 1563–1586. http://dx.doi.org/10.1080/01690965.2011.560006.
Lobina, D. J. (2014). Probing recursion. *Cognitive Processing*. http://dx.doi.org/10.1007/s10339-014-0619-z.
Lobina, D. J., & Garcıa-Albea, J. (2009). *Recursion and cognitive science. Data structures and mechanisms*. Paper presented at the proceedings of the 31st annual conference of the cognitive science society, Austin, TX.
Lupyan, G. (2009). Extracommunicative functions of language: Verbal interference causes selective categorization impairments. *Psychonomic Bulletin & Review, 16*(4), 711–718. http://dx.doi.org/10.3758/pbr.16.4.711.
Luuk, E., & Luuk, H. (2010). The redundancy of recursion and infinity for natural language. *Cognitive Processing, 12*(1), 1–11. http://dx.doi.org/10.1007/s10339-010-0368-6.
Mahon, B. Z., Kumar, N., & Almeida, J. (2013). Spatial frequency tuning reveals interactions between the dorsal and ventral visual systems. *Journal of Cognitive Neuroscience, 25*(6), 862–871. http://dx.doi.org/10.1162/jocn_a_00370.
Mandelbrot, B. (1977). *The fractal geometry of nature*. New York: Freeman.
Martins, M. D. (2012). Distinctive signatures of recursion. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1598), 2055–2064. http://dx.doi.org/10.1098/rstb.2012.0097.
Martins, M. D. (2014). *Language evolution and recursion: An empirical investigation of human hierarchical processing*. PhD, University of Lisbon, Lisbon.
Martins, M. D., & Fitch, W. T. (2012). Empirical approaches to recursion. In T. C. Scott-Phillips, M. Tamariz, E. A. Cartmill, & J. R. Hurford (Eds.), *The evolution of language – Proceedings of the 9th international conference* (pp. 219–225). Kyoto: Singapore: World Scientific.

Martins, M. D., Fischmeister, F. P., Puig-Waldmueller, E., Oh, J., Geissler, A., Robinson, S., et al (2014). Fractal image perception provides novel insights into hierarchical cognition. *NeuroImage, 96*, 300–308. http://dx.doi.org/10.1016/j.neuroimage.2014.03.064.

Miller, S. A. (2009). Children's understanding of the second-order mental states. *Psychological Bulletin, 135*(5), 749–773.

Miller, P. H., Kessel, F. S., & Flavell, J. H. (1970). Thinking about people thinking about people thinking about... A study of social cognitive development. *Child Development, 41*(3), 613–623.

Morey, C., & Cowan, N. (2004). When visual and verbal memories compete: Evidence of cross-domain limits in working memory. *Psychonomic Bulletin & Review, 11*(2). 296–230.

Nardini, M., Jones, P., Bedford, R., & Braddick, O. (2008). Development of cue integration in human navigation. *Current Biology, 18*(9), 689–693. http://dx.doi.org/10.1016/j.cub.2008.04.021.

Penrose, R. (1989). *The emperor's new mind: Concerning computers, minds, and the laws of physics*. Oxford University Press.

Pica, P. (2004). Exact and approximate arithmetic in an amazonian indigene group. *Science, 306*(5695), 499–503. http://dx.doi.org/10.1126/science.1102085.

Pinker, S., & Jackendoff, R. (2005). The faculty of language: What's special about it? *Cognition, 95*(2), 201–236. http://dx.doi.org/10.1016/j.cognition.2004.08.004.

Pulvermüller, F., & Fadiga, L. (2010). Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience, 11*(5), 351–360. http://dx.doi.org/10.1038/nrn2811.

Pyers, J. E., & Senghas, A. (2009). Language promotes false-belief understanding: Evidence from learners of a new sign language. *Psychological Science, 20*(7), 805–812. http://dx.doi.org/10.1111/j.1467-9280.2009.02377.x.

Reboul, A. (2012). Language: Between cognition, communication and culture. *Pragmatics & Cognition, 20*(2), 295–316. http://dx.doi.org/10.1075/pc.20.2.06reb.

Roeper, T. (2009). *The minimalist microscope: How and where interface principles guide acquisition*. Paper presented at the Boston university conference on child language.

Roeper, T. (2011). The acquisition of recursion: How formalism articulates the child's path. *Biolinguistics, 5*(1–2), 57–86.

Schiemenz, B. (2002). *Managing complexity with recursion*. Paper presented at the cybernetics and systems, Vienna 2002.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285. http://dx.doi.org/10.1016/0364-0213(88)90023-7.

Tomasello, M. (2008). *Origins of human communication*. MIT press Cambridge.

Unterrainer, J. M., & Owen, A. M. (2006). Planning and problem solving: From neuropsychology to functional neuroimaging. *Journal of Physiology-Paris, 99*(4–6), 308–317. http://dx.doi.org/10.1016/j.jphysparis.2006.03.014.

von Humboldt, W. (1972). *Linguistic variability and intellectual development*. Philadephia: University of Pennsylvania Press.

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience, 6*(6), 624–631.

Watumull, J., Hauser, M. D., Roberts, I. G, & Hornstein, N. (2014). On recursion. *Frontiers in Psychology* (Vol. 4). http://dx.doi.org/10.3389/fpsyg.2013.01017 (In preparation).

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences, 104*(19), 7780–7785. http://dx.doi.org/10.1073/pnas.0701644104.

Wohlschlager, A., Gattis, M., & Bekkering, H. (2003). Action generation and action perception in imitation: An instance of the ideomotor principle. *Philosophical Transactions of the Royal Society B: Biological Sciences, 358*(1431), 501–515. http://dx.doi.org/10.1098/rstb.2002.1257.