Universidade de Lisboa
Faculdade de Medicina de Lisboa

LANGUAGE EVOLUTION AND RECURSION: AN EMPIRICAL INVESTIGATION OF HUMAN HIERARCHICAL PROCESSING

Dr. Maurício de Jesus Dias Martins

Doutoramento em Ciências Biomédicas

Especialidade de Neurociências

22 Abril 2014

Universidade de Lisboa
Faculdade de Medicina de Lisboa

LANGUAGE EVOLUTION AND RECURSION: AN EMPIRICAL INVESTIGATION OF HUMAN HIERARCHICAL PROCESSING

Dr. Maurício de Jesus Dias Martins

Tese orientada e co-orientada pela Professora Doutora Isabel Pavão Martins e pelo Professor W. Tecumseh Fitch

Doutoramento em Ciências Biomédicas

Especialidade de Neurociências

Todas as afirmações efecuadas no presente documento são da exclusiva responsabilidade do seu autor, não cabendo qualquer responsabilidade à Faculdade de Medicina de Lisboa pelos conteúdos nele apresentados.

A impressão desta dissertação foi aprovada pelo Conselho Científico da Faculdade de Medicina de Lisboa em reunião de 22 de Abril de 1014

Agradecimentos

Esta tese e os trabalhos que a compõem nunca teriam sido possíveis sem a ajuda preciosa de mentores, amigos, familiares e colaboradores. Em primeiro lugar, agradeço aos meus supervisores, Isabel Pavão Martins e Tecumseh Fitch, por terem sempre acreditado em mim, e por me terem dado a autonomia e a responsabilidade necessárias para crescer como cientista. Em segundo lugar agradeço aos meus colaboradores, colegas e amigos Piera Fillipi, Jinook Ho, Riccardo Hofer, Lisa Horn, Dan Bowling, Bruno Gingras, Andrea Ravignani, Gesche Westphal-Fitch, Estela Puig-Waldemüller, Florian Fischmeister, Roland Beisteiner, Stefan Reber, Eva Freiberger, Sabine Laaha, Soonja Choi, Andrius Pašukonis, Maxime Garcia, Marisa Hoeschele e Sarah Deventer pelas discussões e sugestões que tanto contribuíram para melhorar a qualidade do meu trabalho. Agradeço também a Petra Pesak, Nadja Navcik, Drasko Boko, Elisabete Lopes, Gabriela Leal e José Fonseca por me terem sempre ajudado a ultrapassar dificuldades institucionais e burocráticas. Agradeço ao Paolo Baldi e à Francesca Bonalumi por terem sido a minha família em Viena e por terem tentado fazer de mim uma pessoa melhor, e à Lorena Ciutacu pela capacidade de dissolver a minha ansiedade com um sorriso.

Finalmente, e acima de todos, agradeço à minha família - Susana, Maria Amélia e Maurício Martins. Para onde eu for levo comigo o seu exemplo de força, generosidade e entrega incondicional. Esta tese é sobretudo deles.

Summary

Humans generate complex hierarchical structures in a variety of domains such as in language, social organization, music, action sequencing and visual arts. One cognitive capacity associated with this extraordinary generative power is recursion. Recursion is a very efficient method to process hierarchies and it allows the generation of unbounded hierarchical depth from finite means. Recursion can be defined as the ability to represent the embedding of hierarchies within hierarchies of the same kind.

Although recursion has been hypothesized as uniquely human and primarily linguistic, the empirical investigation of these hypotheses has been hindered by the absence of methods to test for recursive capabilities outside the domain of language.

In this thesis I present a novel task that can be used to investigate the ability to represent recursion (hierarchical self-similarity) in the visuo-spatial domain. I will describe a set of experiments in which I attempt to characterize recursion as a psychological entity by describing its relationship with other cognitive abilities, as well as its developmental patterns and neural underpinnings.

The conclusions of this research program are the following: 1) humans can represent recursion in the visuo-spatial domain; 2) this ability requires the acquisition of abstract rules; 3) recursion can be efficiently used to represent information common to different levels of a hierarchy, and it enhances the ability to detect fine-grained hierarchical mistakes, 4) linguistic resources are not specifically active while processing visual recursion neither behaviorally nor at the neural level, however recursion seems to require the integration of spatial and categorical information.

The novel task and results presented here open up exciting pathways in the investigation of recursion as a cognitive ability. Because it is a visual task, not requiring verbal instructions or responses, it can also be used to test non-human primates and clinical populations with language impairment.

Keywords: Recursion, Hierarchy, Cognition, Language, Neural Correlates.

Resumo (300 palavras)

A espécie humana é capaz de produzir hierarquias complexas na linguagem, organização social, música, actividade motora e nas artes visuais. O poder generativo da cognição humana tem sido associado a um módulo computacional designado recursividade, que pode ser definido como a capacidade de representar a incorporação de hierarquias dentro de hierarquias do mesmo tipo.

A recursividade pode ser usada de modo eficiente no processamento de hierarquias, permitindo a geração de estruturas infinitamente profundas partindo de um número finito de elementos. Esta capacidade tem sido postulada como exclusivamente humana e primariamente linguística. No entanto, a investigação empírica destas hipóteses tem sido dificultada pela ausência de um método para testar capacidades recursivas fora do domínio linguístico.

Nesta tese irei apresentar um novo método para testar a capacidade de representar a recursividade no domínio visuo-espacial. Irei descrever uma série de experiências nas quais caracterizarei a recursividade como uma entidade psicológica, descrevendo de que forma se relaciona com outras capacidades cognitivas, o seu padrão de desenvolvimento e correlatos neurais.

As conclusões deste programa de investigação são as seguintes: 1) a espécie humana é capaz de representar recursividade visuo-espacial; 2) esta capacidade requer a aquisição de regras abstractas; 3) a recursividade é usada para representar informação comum a vários níveis hierárquicos e melhora a capacidade de detectar erros estruturais ao nível dos pequenos detalhes; 4) o processamento de recursividade visual não activa especificamente recursos verbais, quer ao nível do comportamento quer ao nível neural, contudo esta capacidade requer a integração de informação espacial e categorial.

A tarefa e os resultados inovadores aqui apresentados abrem novas vias de investigação relativamente à capacidade de utilizar recursividade ao nível cognitivo. Por ser uma tarefa visual não requer instruções nem respostas verbais, pelo que pode ser usada para testar primatas não humanos e populações clínicas com defeitos de linguagem.

Palavras-chave: Recursividade, Hierarquia, Cognição, Linguagem, Correlatos Neurais

Contents **Page**

1. General introduction

1. General Introduction

Humans are exceptional creatures. Our ability to form complex social structures, and to transform our environment is unprecedented in the animal kingdom. These capabilities allowed humans to spread through a wide variety of habitats, and to adopt flexible survival strategies making us the one of the most versatile species in the history of animal life.

What makes us exceptional is our cognitive power: our ability to combine actions to achieve complex goals and to represent complex structures go 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; 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 (Figure 1). For example, in action sequencing (Figure 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).

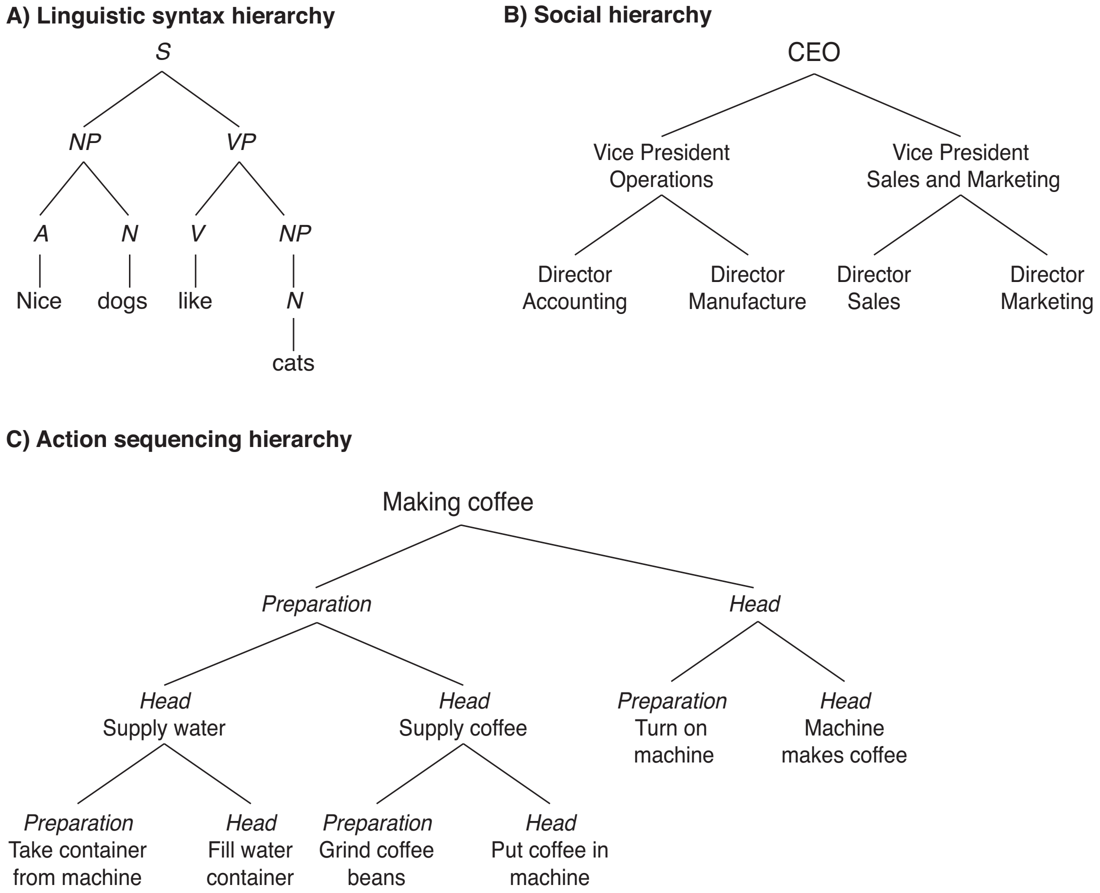
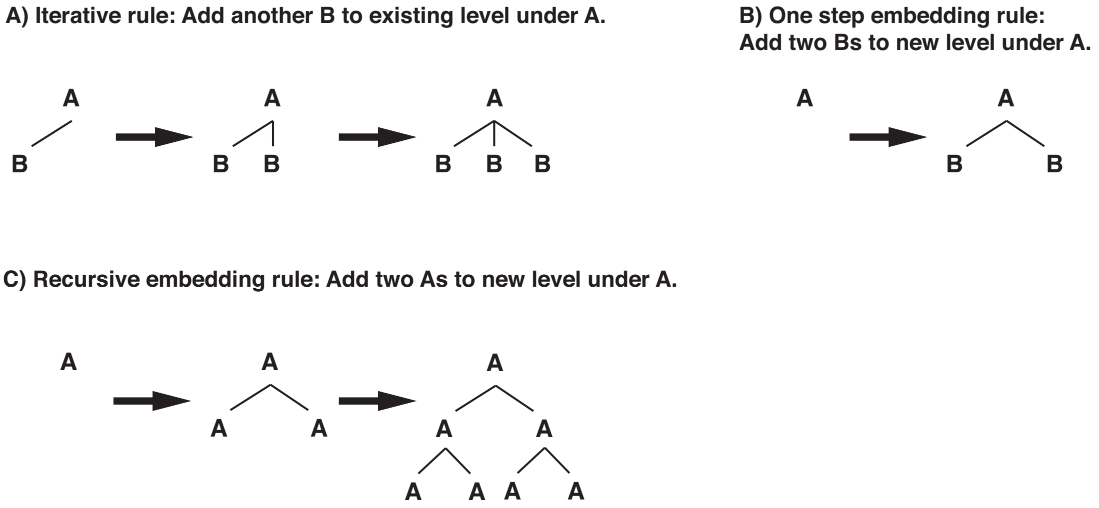


Figure 1. Examples of linguistic (A), social (B) and action sequencing (C) hierarchies.

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 Figure 2. For instance, 'iterative rules' (Figure 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 (Figure 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 (Figure 2C). As we can see from Figure 2, 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, &

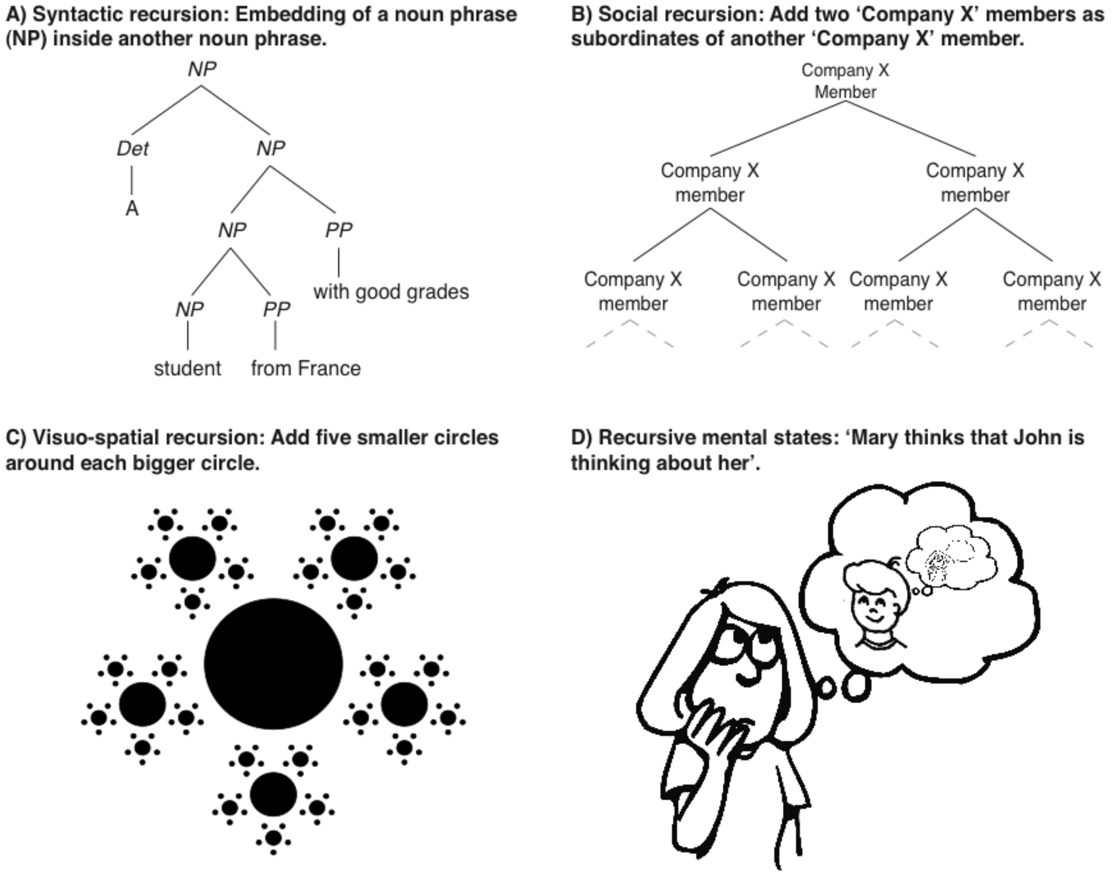
Chomsky, 2005; Hauser et al., 2002), theory of mind (P. H. Miller, Kessel, & Flavell, 1970; S. A. Miller, 2009) and visuo-spatial objects (Martins, 2012), can be efficiently represented using recursion (Figure 3). The ability to represent a recursive rule has been suggested as a necessary condition for the open-ended power of human cognition (Fitch et al., 2005; Hauser et al., 2002).


A) Iterative rule: Add another B to existing level under A.

B) One step embedding rule: Add two Bs to new level under A.

C) Recursive embedding rule: Add two As to new level under A.

Figure 2. Examples of rules used to generate hierarchical relationships.


A) Syntactic recursion: Embedding of a noun phrase (NP) inside another noun phrase.

B) Social recursion: Add two 'Company X' members as subordinates of another 'Company X' member.

C) Visuo-spatial recursion: Add five smaller circles around each bigger circle.

D) Recursive mental states: 'Mary thinks that John is thinking about her'.

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

1.1 Recursion as a representational ability

In this thesis, a series of manuscripts will be presented that attempt to characterize the cognitive phenomenon of recursion from an empirical perspective. The empirical 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. The empirical framework presented here was developed in an attempt to resolve these questions.

One of the biggest sources of confusion surrounding 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. As will be discussed in detail in the second and third chapters, recursive structures can be generated by recursive processes as well as by non-recursive procedures (Figure 4).

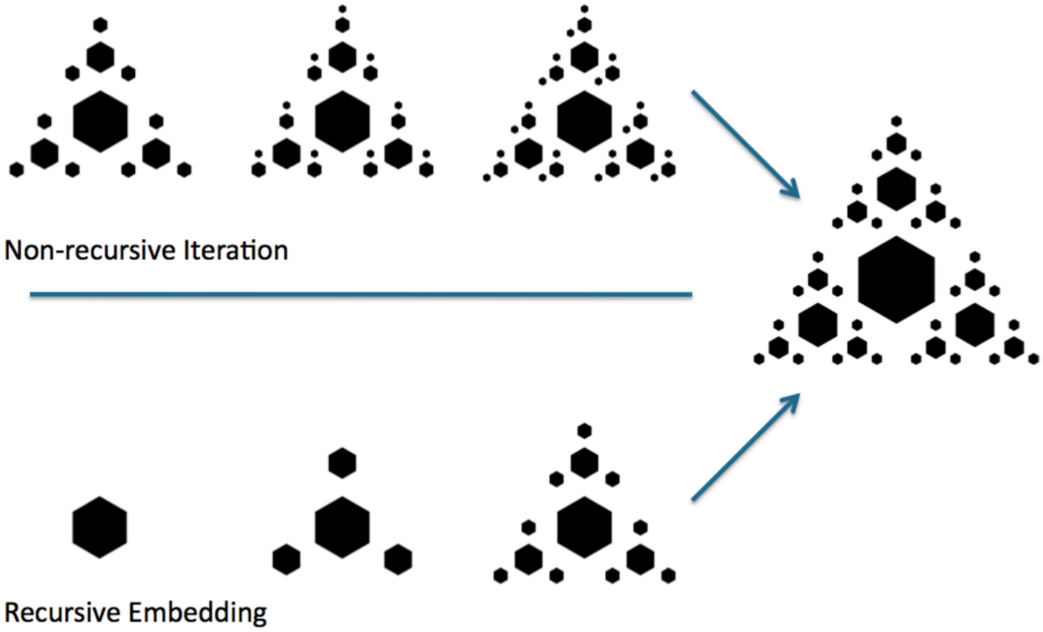


Figure 4. A visual recursive structure that could be generated by either by a recursive process or a non-recursive process.

The opposite is also true: recursive processes can generate structures that are not visibly or ostensibly recursive. For example, take a function 'round(x)' that rounds a decimal number 'x' to its closest integer unit. Here are some applications of the function: round (2.4) = 2; round (2.6) = 3, round (3.2) = 3, etc. Now, consider the following process:

$y_0 = 1$;
$y_i = \text{round}(y_{i-1} + 0.4),\ i = 1, 2, 3 \ldots n$

This process is recursive, because the function y_i calls itself. However, on each iteration i (i = 1, 2, 3 … n), the addition of 0.4 is insufficient to round the number up to the integer '2'. Therefore, the recursive nature of the generative process cannot be derived from its output sequence, which is simply 1, 1, 1, …1.

Given this double 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:

$N_0 = 1$
$N_n = N_{i-1}+1,\ i = 1, 2, 3 \ldots n$

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 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).

In the second and third chapters of this thesis I will lay out why I disagree with this approach, and argue that trying to directly measure a relatively unexplored cognitive processes 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.

This thesis will start with a defense of 'representationalism'. I will defend 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 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; 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

Within the domain of language, recursion seems to be universally used (Reboul, 2012), and although rare in common speech (Laury & Ono, 2010), most language users in the world are likely to have generated several recursive sentences in their lifetimes (for instance, compound nouns such as "[[[student] film] committee]"). Furthermore, the ability to extract the correct meaning from recursive sentences seems to be available early during ontogenetic development (Alegre & Gordon, 1996; Roeper, 2009). This interesting relationship, yet to be demonstrated in other domains, has led some authors to propose that the biological evolution of language might have been tied in with the cognitive availability of recursion. One specific hypothesis states that recursion is a domain-specific “linguistic computational system […], independent of the other systems with which it interacts and interfaces” (Fitch et al., 2005; Hauser et al., 2002). This hypothesis goes on to propose that the use of recursion in other domains might be parasitic on verbal resources. Coincidentally, the ability to perform second-order theory of mind tasks correlates with language abilities (S. A. Miller, 2009), and verbal interference tasks block the ability to use natural numbers (Gordon, 2004). These results, though interesting, are not in themselves proof that recursion is a linguistic domain-specific ability.

This hypothesis is extremely attractive, especially because it is a strongly intuitive idea that the evolution of language is at the center of the extraordinary human cognitive development. Other species lack communicative behaviors that come even close to the generative power of human language (Conway & Christiansen, 2001; ten Cate & Okanoya, 2012), and language allows the coordination of groups at large quantitative, spatial and temporal scales. However fascinating, these hypotheses have remained empirically untested for more than a decade. Numerous articles and books have been written about the role of recursion in language evolution (Chomsky, 1995; Corballis, 2011; Fitch, 2010; Fitch et al., 2005; Hauser et al., 2002; Hulst, 2010; Jackendoff & Pinker, 2005; Lobina, 2011; Lobina & Garcıa-Albea, 2009; Lowenthal & Lefebvre, 2014; Luuk & Luuk, 2010; Watumull et al., 2014), but to date the lack of a clear method to test for recursive abilities in non-linguistic domains has prevented the empirical assessment of these hypotheses.

1.3. Goals of the thesis

In this thesis I will describe a research program that addresses the issues described above. I will describe the development of a novel (non-linguistic) visual recursion task, and a series of experiments investigating how visual recursion relates with other cognitive abilities, including language. Specifically, I will aim at providing empirical answers to the following questions:

(1) Is recursion specific to the linguistic domain?

(2) Is language necessary to use recursion in non-linguistic domains?

1.4. Outline of the thesis

This thesis will be divided into 6 sections, each being an independent module of a broader research program. These sections, or manuscripts, are presented in the same format as they were submitted or published.

In Chapter 2, "Distinctive signatures of recursion", I will review the different concepts of recursion used in the literature, and argue for the usefulness of definitions focused on the ability to represent hierarchical self-similarity. Then, I will describe possible methods to test for this capacity, and speculate about the evolutionary advantages of recursion.

In Chapter 3, "Investigating recursion within a domain-general framework", I will argue that rather than assuming that recursion is language domain-specific, it is a better empirical approach to investigate its presence in a variety of domains, and then measure how these correlate or dissociate. This viewpoint frames the domain specificity of recursion as an empirical question, and not as an assumption.

In Chapter 4, "Fractal geometry and visual recursion: a novel approach to hierarchical self-similarity", I will present a novel method to test for visuo-spatial recursion, and several experiments in which the task is validated.

In Chapter 5, "Processing visual recursion does not require verbal and motor resources", I will describe an experiment using a dual task paradigm in which we investigated whether visual recursion requires verbal resources.

In Chapter 6, "How children perceive fractals: hierarchical self-similarity and cognitive development", I will describe an experiment in which we tested whether the development of grammar comprehension in children specifically correlates with the ability to represent recursion in the visuo-spatial domain. We also investigated the learning constraints of recursion in different developmental stages.

In Chapter 7, "Fractal Image Perception provides novel insights into hierarchical cognition", an fMRI experiment will be reported in which we investigated the neural correlates of recursive cognition, with a special focus on potential activations of language areas.

Finally, in the last section, I will summarize and integrate the findings from all experiments and discuss their implications for future research.

2. Distinctive signatures of recursion

Martins M.D. (2012). Distinctive signatures of recursion. *Phil. Trans. R. Soc. B 367*, 2055-2064

Distinctive signatures of recursion

Mauricio J.D. Martins [1,2,*]

[1]Department of Cognitive Biology, University of Vienna, Althanstrasse 14, 1090
Wien, Austria
[2]Language Research Laboratory, Lisbon Faculty of Medicine, Av. Professor Egas
Moniz, 1649-028 Lisboa, Portugal

Abstract

Although recursion has been hypothesized to be a necessary capacity for the evolution of language, the multiplicity of definitions being used has undermined the broader interpretation of empirical results. I propose that only a definition focused on representational abilities allows the prediction of specific behavioral traits that enable us to distinguish recursion from non-recursive iteration and from hierarchical embedding: Only subjects able to represent recursion, i.e., to represent different hierarchical dependencies (related by parenthood) with the same set of rules, are able to generalize and produce new levels of embedding beyond those specified *a priori* (in the algorithm or in the input). The ability to use such representations may be advantageous in several domains: Action sequencing, problem solving, spatial navigation, social navigation and for the emergence of conventionalized communication systems. The ability to represent contiguous hierarchical levels with the same rules may lead subjects to expect unknown levels and constituents to behave similarly and this prior knowledge may bias learning positively. Finally, a new paradigm to test for recursion is presented. Preliminary results suggest that the ability to represent recursion in the spatial domain recruits both visual and verbal resources. Implications regarding language evolution are discussed.

Keywords: Recursion; Hierarchy; Embedding; Representation; Language

1. Introduction

Recursion is one of the most controversially discussed terms in the cognitive sciences. Although it has been hypothesized as a human-unique trait and a as a necessary capacity for the evolution of language [1], the multiplicity of definitions being used [2-6] has undermined the broader interpretation of empirical results [7]. One of the major problems that stems from this multiplicity is the difficulty in drawing boundaries between recursion and similar processes such as cognitive grouping, hierarchical embedding and iteration [8].

Although there has been a proliferation of literature arguing for and against the claim of recursion as uniquely human [2, 8-14], the debate remains unresolved. Some empirical paradigms have been considered relevant to address the topic [15-18] but since they fail to capture the distinction between recursion and hierarchical embedding, over-interpreting the results may be misleading [12, 19].

Given that brain computations are opaque to observers (until behavioural correlates have been found), definitions focused on algorithmic properties (such as "a recursive function is one that calls itself") may not be entirely relevant for empirical research. On the other hand, isolated analyses of signals (such as vocalizations, social interactions etc.) may be misleading since not all structures that can modelled using recursion are produced by recursive processes, neither are these structural properties necessarily perceived by observers.

To overcome these difficulties, I propose that only a definition focused on representational abilities such as "the ability to represent self-similarity across hierarchical levels" enables the prediction of recursion-specific behavioural traits: if a subject is able to represent different hierarchical levels, i.e. different hierarchical nodes related by 'parenthood' (in the graph theory sense), with the same set of rules, then he or she may be able to generalize and generate new levels of embedding ('child' nodes) beyond those specified *a priori* (whether in the algorithm or in the input).

Defined as such, recursion may provide advantages to its users in the domains that it is available: It may provide prior knowledge regarding new or unknown hierarchical levels; and if shared by a population, it may contribute to the establishment of communicative conventions. Here it is important to make explicit that we analyse recursion as a kind of representational abstraction without considering how it could be implemented in the brain. Recursion could be a single module recruited by different modalities or it can be an umbrella term referring to a set of mechanisms that operate independently in different domains, each with its own specific constraints. The empirical research essential to support any of these hypotheses has been delayed by the shortage of tools to assess the use of recursion in non-linguistic domains.

Under this framework, a new paradigm to test for recursion in the visuo-spatial domain will be presented. Given that it can be applied independently of language and in a non-serialized modality, it has the potential to provide insights regarding the relationship between recursion and language in the evolutionary history.

2. Recursion: from operation to structure

As pointed out by Fitch [7], recursion has been many things to many people. Some definitions focus on the characterization of recursive computations; others attempt to describe which structures can be considered recursive.

In modern computer science a recursive function is one that calls itself, or one that is defined in terms of itself [7, 10]. However, in logics, 'recursive' can mean 'computable' (i.e. if membership of the function products can be determined by a Turing machine) [3, 20], or refer to the process of defining something in terms of something previously defined [2, 4-6]. As pointed out by some authors [7, 8], this latter definition is too broad since it includes computations that specify items in terms of simpler items and therefore any operation able to generate hierarchies (as occurs in cognitive grouping and different perceptual domains [11]). In the most restrictive sense of recursion ('specific recursion'), the items being combined (or embedded) should be categorized as of the same category as the ones they generate (or are embedded on) [7, 8, 10].

Although definitions focused on the process can be a good start to define which phenomena we are trying to grasp, they are not completely useful for empirical purposes. Since the implementation of a computation is opaque to the observer (at least before some behavioural correlate has been found), a better empirical approach is to search for distinctive signatures in the output that may suggest the presence of that computation. In the case of recursion those signatures are usually the presence of structural self-similarity or the embedding of constituents within constituents of the same kind [7, 10].

3. Recursion: From structure to representation

Recursive structures (in the strict sense) are ubiquitous in human activity and have been claimed in visual art [21], music [22, 23], architecture [24], humour [25], second-order theory of mind [26], problem solving [27], action sequencing [28], syntax [29-31], prosody [8, 32] and conceptual structure [33, 34]. These cultural achievements are present not only in modern societies but also in pre-industrial and ancient civilizations.

In spite of the pervasiveness of structures that can be modelled using recursive algorithms or rule sets, not all of them will be represented as such. This means that the amount of recursion in a structure will only be relevant for an observer to the extent that he can decode it meaningfully. For example, in the Kotoko architecture [24], self-similarity in different scales is built consciously, subjected to abstract representation and used to convey a meaning (e.g. social ranking). In such circumstances we can say that both producers and observers have the ability to represent the underlying recursive structure. On the other hand, although we can model the long-distance tensional structure (e.g. tonal deviation from the tonic) in music as recursive, untrained listeners may not be sensitive to such properties [22, 23]. Likewise, even if we can use recursion to model baboons' social hierarchies [35], this does not imply that baboons, in spite of their success in social navigation [36], are able to represent recursion. In the latter example, it is possible that individuals use separate rules to represent different hierarchical levels ([X is dominant over Y]; [Y is dominant over Z]) instead of using recursive rules to encode dominance ([X is dominant over Y [who is dominant over Z]]) [37].

The opacity of algorithmic processes can be further exemplified by one of the first structures described using a recursive generating rule: The Fibonacci sequence. In 1201, Leonardo de Pisa described a sequence of numbers where each member of the sequence $S_{(n)}$ could be obtained by the sum of the two previous members: $S_{(n)}=S_{(n-1)}+S_{(n-2)}$. Although the Fibonacci sequence (1 1 2 3 5 8 13 etc.) can be implemented using a recursive algorithm, with a function that calls itself:

```
def fib(n):
        if n = 1 or n = 2: return 1
        return fib(n-2) + fib(n-1)
```

it can also be implemented with a non-recursive simple iterative loop:

```
def fib_iter(n):
    if n = 1 or n = 2: return 1
    pre = 1
    prepre = 1
    for i in range(3,n):
        pre, prepre = pre+prepre, pre
        return pre+prepre
```

An isolated analysis of the output/signal is insufficient to determine the underlying computation, therefore, the fact that a certain individual can produce the Fibonacci sequence tells us little about his ability to use or represent recursion. Independently on how the sequence is produced, if a given observer is able to use recursion to represent the subset that he receives from the input, then he or she may display specific behaviours while generating further elements. In the next section we will discuss these distinctive behaviours in more detail.

In summary, not all activities that can be synthesized with recursive processes are going to be perceived as structurally meaningful by the observers. Hence, the ability to produce recursive structures and the ability to decode them do not necessarily come together [2, 38]. Given that the ability to represent self-similarity in a structure (regarding a certain feature) may result in different behaviours, questions concerning representational abilities are more tractable empirically.

To the purposes of this paper, we define representation as a relationship between 2 objects (O1 and O2), where a given set of characteristics [15] of an object (O1) can be retrieved from another (O2). A cognitive representation entails that some change at the neural level (O2) occurred due to the perception, storage or processing of certain features {C} present in the object (O1). That neural change (O2) will have a causal relationship to a secondary process (O3) that can be *measured* (for example, BOLD signal or behavior). Under this assumption one can detect whether a feature was represented while remaining agnostic regarding the nature or implementation of that representation.

Following this framework, in order to plan experiments and interpret behavioural responses we first need to make theoretical distinctions between recursion and related processes such as non-recursive iteration and non-recursive embedding.

4. Iteration, hierarchical embedding and recursion

Iterative processes involve the repetition of an operation a given number of times. These processes may or may not generate hierarchical structures and may or not create dependency relationships between different elements.

Hierarchical structures involve the embedding of constituents within other constituents. If the embedding involves constituents of the same category it is called recursive embedding (in the strict sense), otherwise it's called non-recursive. Iteration, hierarchical embedding and recursion are not mutually exclusive. Nevertheless it is possible to segregate the cognitive abilities that are necessary to represent the kind of information that they encode (Fig.1). In the next sections I will discuss how.

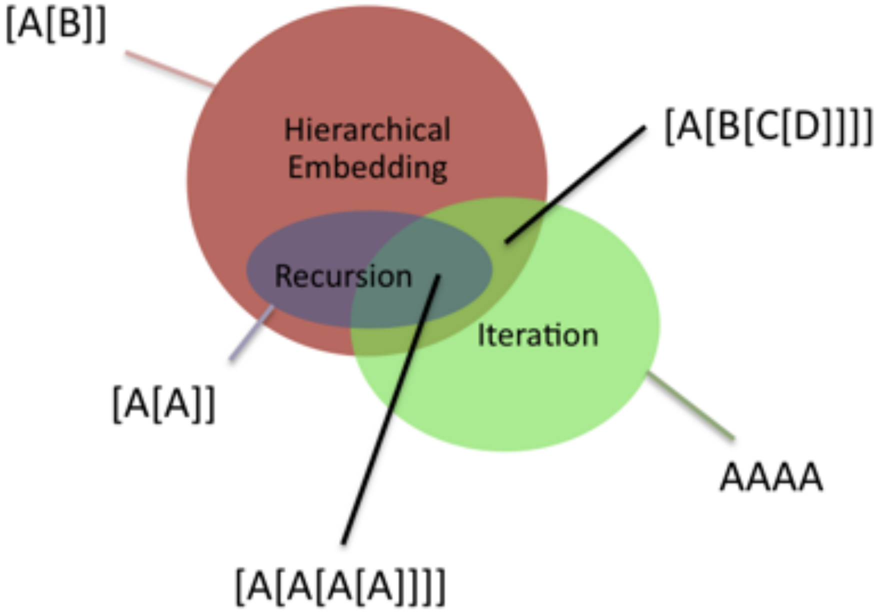


Fig 1. Examples of structures produced by iteration, hierarchical embedding and recursion and by the combination of these processes. Constituents represented with the same letter are perceived as similar regarding a certain feature of relevance to the hierarchical structure. Brackets mean embedding.

4.1 Iteration without embedding

Consider a set composed by the ordered and indexed alphabet list. Call this set ALPHA. An iterative process example could be:

(1) Add the element ALPHA(n) to the structure x(n) until n=3. Each cycle add 1 to n.

x0= A
x1= AB
x2= ABC
x3= ABCD

In this process, each iterative step is a separate act that can exist independently from the others [30, 37]. Such processes can create infinite sequences by unlimited concatenation [39] but cannot encode dependency relationships nor create new hierarchical levels [10, 35]. The encoding of dependencies requires the representation of rules that allow embedding, i.e. the representation that some constituents are dependent of other constituents, either structurally of functionally. Consider the next example:

4.2. Iteration with embedding

(2) Implement one of the following rules. Repeat the cycle 3 times:

a) Embed B on A
b) Embed C on B
c) Embed D on C

Again, well-formed structures could be:

x0= [A]
x1= [A[B]]
x2= [A[B[C]]]
x3= [A[B[C[D]]]]

For a given observer, the shape of the output in both examples, (1) and (2), could be similar: *'ABCD'*. However, if the observer is able to represent rules of embedding, he has the possibility to interpret the positional attributes of the string *'ABCD'* as containing information about dependency relationships (e.g. "right constituents are dependent over left constituents"). This perceptual decision requires access to semantic information and will be influenced by biases that can be innate, cultural or contextual (for example, prosodic cues can influence syntactic interpretations).

Considering the big picture of comparative cognition, there are three empirical questions relevant at this level: 1) Which species possess the ability to represent dependency relationships; 2) In which domains are they able to do so; and 3) Which factors influence the perception of a structure as having dependency relationships.

4.2.1 Single vs. Multi-constituent hierarchical levels

Iterative processes that allow embedding (such as (2)) can generate hierarchical structures. However, the same process that can create hierarchies with more than one constituent per level ([A[BB[C]]], [A[B[CC]]] or [A[BBB]]) can be used to create hierarchies with only one (non-empty) constituent per level (e.g. A[B[C[D]]]).

Potential differences in the processing of hierarchical nodes with one or several dependents relate to the ability to use memory to keep track of non-adjacent dependencies when the hierarchical information is presented linearly. If memory constraints are not an issue or if structures are presented non-linearly, then the same representational abilities should allow the encoding of both single and multi-constituent hierarchical information. This means that the ability to process long-distance dependencies is not a specific signature of recursion, but general to hierarchical processing when there is more than one dependent per hierarchical node.

Consider the strings used in artificial grammar learning studies with the structure A^nB^n [16]: In $A_1A_2B_2B_1$, for example, the semantic content of the inner 'A_2B_2' is not modified by the outer 'A_1__B_1'. In fact, there is no real dependency relationship between the inner and outer 'AB', since they can exist independently and without changing the properties of each. The potential 'dependency' in these structures relates to the association between 'A_n' and 'B_n' and not between [AB] and another [AB]. Thus, and contrary to what has been argued [12, 13], the ability to process A^nB^n structures (when embedded structures are not semantically related) cannot be considered as a specific trait of recursion [37]; neither are structures without long-distance dependencies necessarily deprived of recursion (such as in tail-recursion).

A more interesting issue relates to the limits of hierarchical processing with non-recursive embedding rules. If we consider the process (2) and the kind of structures it generates, we realize that each hierarchical level has to be represented individually. In these circumstances, we can embed an infinite number of constituents within the same hierarchical level [39], but we cannot create new hierarchical levels unless they are specified *a priori* (either as explicit rules in the algorithm itself, or acquired *via* the input). Recursion overcomes this limitation.

4.3. Recursive embedding

Within the same framework, a recursive generation rule would be:

(3) Embed a member of the ALPHA set in another member of the ALPHA set

Again we can obtain structures such as:

x0 = [A]
x1 = [A[B]]
x2 = [A[B[C]]]
x3 = [A[B[C[D]]]],

If a rule like (3) is used, all elements of ALPHA are represented as having the same properties (relatively to the fact that they belong to the same set, although these elements can differ in many other characteristics). Hence, the structure [A[B[C[D]]]] can be perceived as equivalent to [ALPHA[ALPHA[ALPHA[ALPHA[ALPHA]]]]]. Within this framework, new hierarchical levels can be represented without new rules being specified, as in the structure:

x4 = [A[B[C[D[E]]]]]

Moreover, with the same set of rules we can represent new design features that might be useful for dynamic hierarchies:

- Inversion of the previous order of dependency: [B[A[C]]];
- Expression of bilateral dependency relationships: [A[B[C[B]]]], etc.

Obviously, such representations rules can be useless if unconstrained since they are too general. However, the availability of these rules to represent hierarchical structures may be advantageous in terms of flexibility [7], and can be the only practical method for large and highly complex hierarchies [40].

In spite of these processing advantages of recursion, it is not clear to what extent they are relevant empirically, given that it may be difficult to distinguish between an algorithm with a large set of rules and one that uses recursion [41]. For this reason, in my opinion, the key "functionalist-cognitive accomplishment", as Harder [42] puts it, "is the ability to take one incremental step beyond the given". This means that the key empirical test for recursion is the ability to represent dependency relationships that were not previously defined, or to represent information within hierarchical levels not previously 'available'. What this ability presupposes is the knowledge (or expectation) that all nodes within a hierarchy can behave similarly and can display the same properties relatively to the way they interact with the nodes 'above' and 'below'. This allows, for example, that we embed a noun phrase inside a noun phrase already embedded in a noun phrase ($[NP_{(n-1)}[NP_{(n)}[NP_{(n+1)}]]]$); or that each individual in a social hierarchy is represented has having both dominants and dependents. In the next section I will discuss how such properties might have provided evolutionary advantages.

The main point of this section is that different behavioural signatures may enable the detection of different cognitive processes:

a) Iteration: Ability to represent repetition of constituents.
b) Hierarchical embedding: Ability to represent dependency relationships between constituents.
c) Recursive embedding: Ability to represent new hierarchical levels (or new dependency relationships) beyond the given (innately or beyond the observable).

5. What is recursion good for?

The ability to take steps beyond the given (regarding hierarchical embedding) and the ability to represent different hierarchical levels with the same set of rules may provide several advantages in the domains it is available:

1. Within a hierarchical system, recursion allows the same way of thinking across different levels and the generation of new levels of embedding [27]. This entails the possibility of unbounded subdivision of each constituent into further subordinate constituents (useful in problem solving [43], action sequencing [44-46] etc.); and the combination of elements creating new dominant constituents (for example, the combination of primitive concepts into new concepts [33], an important feature of human creativity [34]). In this regard it is important to refer that these properties are distinct from cognitive grouping (often taken as a synonym of recursion [11, 23, 32]): Although cognitive grouping may allow the clustering of existing constituents in supra-constituents (e.g. the organization of a visual array in clusters like 0000 → [00][00] → [[00][00]]), it does not allow the generation and recruitment of new constituents, for example, the subdivision of each "0" into "[00]", generating the structure [[[00][00]][[00][00]]].

2. Recursion is an efficient method to encode complex hierarchies whether or not they were generated recursively [27, 40, 43]: If a given observer is able to build a compressed abstract representation of an entire hierarchy, then he can focus his attention on a small subset of constituents without loosing track of the contextual ensemble. An eventual orthogonal and simultaneous representation of both the whole (abstract) and the details (perceptual) would constitute an efficient and accurate real-time strategy to parse complex hierarchical information. This kind of representation, already described for vision [47, 48], could be useful if available in other domains, such as social and spatial navigation, where unseen (or unknown) landmarks have to be implicitly represented. Although this processing seems to occur in an automatic way, it is possible that in some domains the generation of complex abstract rules implies a slower and more cognitive demanding acquisition phase. Currently, it is unknown where such a phase is required and whether its processing is domain specific.

3. By allowing an implicit abstract representation of unseen or unknown constituents, the availability of recursion may decrease the amount of uncertainty regarding the interaction with hierarchical structures [14]. If an implicit representation would be generated automatically, this could bias individuals to perceive hierarchies as self-similar, and to expect similar behaviours at different hierarchical levels (Fig.2). This predisposition could be advantageous to the extent that most hierarchies can be usefully modelled as self-similar; or to the extent that the expectation of self-similarity is better than no expectation.

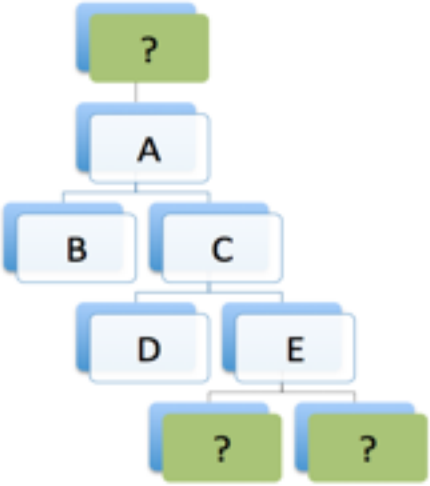


Fig.2. Expectation of self-similarity in a social hierarchy: By observing a series of interactions between the individuals (A), (B), (C) and (D), a given observer may infer the dominance relationships depicted in white boxes. If recursive rules are used to represent these relationships (e.g. 'each Y member of the hierarchy has two Y subordinates and one Y dominant'), then the observer might expect [E] to have subordinates and [A] to have dominants (green boxes).

Here it is important to refer two points: 1) If there are no priors regarding which rules are possible or likely to represent the output of a given system, it may be impossible to generalize and predict its future behaviour [28, 49]; 2) Although real objects deviate from abstractions (like apples deviate from spheres), it seems that, at least in visual processing, abstract prototypes and the amount of deviation from those prototypes can be simultaneously represented in a mutually informative way [47]: The amount of deviation allowed before a shift from the initial abstraction is dependent on the suitability of the resulting behaviours [14, 50]. Although recursive prototypes can be the initial priors, the representational schema may evolve due to functional constraints. Updated rule sets can contain non-recursive hierarchical templates that are less powerful and more restrictive. Independently of the initial state of a given representational system, the final rule set can assume a variety of forms [41].

This hypothesis raises the empirical question of whether humans (or other species) try to collapse hierarchies into recursive prototypes (until the amount of perceived deviation forces cognitive set shifting).

4. An eventual predisposition to interpret hierarchies as self-similar may be useful for the transmission of information and for the acquisition of communication systems: On the one hand, given that self-embedding processes can generate self-embedded structures (for example, in prosody and syntax), the resulting isomorphism can increase the precision of decoding [1, 38]; On the other hand, if such cognitive biases are shared by a population, this would increase the likelihood of the emergence of a conventionalized system to transmit information [51, 52]. Recent work has shown that the presence of recursive rules as Bayesian priors may enhance the acquisition of syntactic rules [53]. This seems to support the notion that the ability to represent recursion may bias learning positively.

5.1. Recursion and Communication

Much has been speculated about the relationship between recursion and language. However, in recent discussions, some convergence has started to form around the idea that recursion might have been available in other domains before the emergence of language [1, 9-11, 28, 46], and that further modifications in human cognition made it available for communication [1, 9-11].

A plausible candidate for this ‘further modification’ seems to relate to the hypothesis that the phonological output of private speech might help to serialize private cognition increasing the attention focus on one train of thought and strengthening the short-term memory capacity [7, 10]. Another plausible candidate seems to be the ability to build symbolic representations [54] (not necessarily available in other serialized domains such as prosody and motor sequencing).

Although the processing of information, for example in the visuo-spatial domain, seems to occur independently from the ability to serialize thought, it is an open question whether the generation of recursive abstract rules can occur without serialization and/or without symbolic representation. If it seems to be true that principles of perceptual abstraction may be employed in non-linguistic domains [11, 32, 47], it also seems that the usage of language and serial representations in these modalities might enhance the processing accuracy [55].

The assessment of these hypotheses must be empirical, and could be systematized with the following questions:

a) Are humans (or other species) biased to interpret hierarchies as self-similar (i.e. as structures where successive mother-child dependency relationships can be described by the same rules)?

b) Are humans (or other species) able to represent new hierarchical levels (beyond the given)?

c) Is this ability available in non-linguistic domains? If so, in which domains?

d) If available in other domains, is it dependent or enhanced by language (or by symbolic serialization)?

In the next section, I will describe a new method that could be used to address these questions and report some preliminary results.

6. Visual Recursion Task

A new task has been developed [56] to assess the ability to represent visuo-spatial hierarchies as recursive structures (fig.3); and to apply these representations in the production of new levels of embedding. This method, called Visual Recursion Task (VRT), is based on the properties of geometrical self-similar fractals, which can be generated by applying recursive embedding rules a given number of iterations.

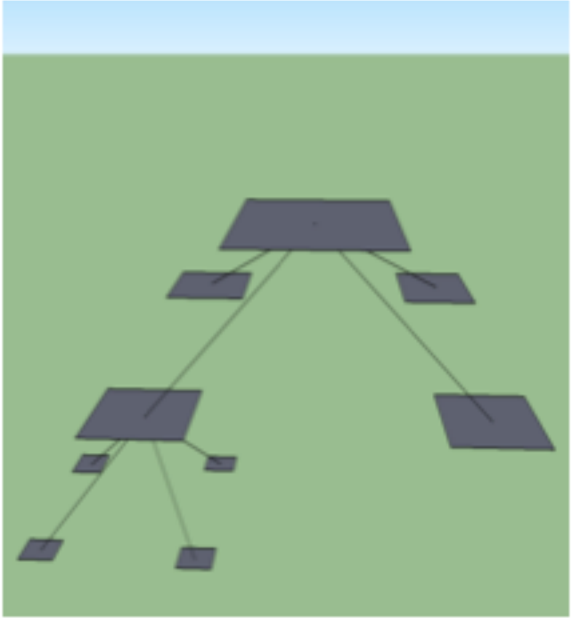


Fig.3. Recursive visuo-spatial hierarchy. A 2D visual structure can be represented as 3D hierarchy where bigger constituents are dominant over smaller constituents.

In a typical VRT stimulus, subjects are exposed to the first three iterations of a fractal structure generation. Then, they are asked to choose, from two possible alternatives, which corresponds to the ‘correct’ answer (fig.4; correct in this sense means the 4th iteration of the generating process).

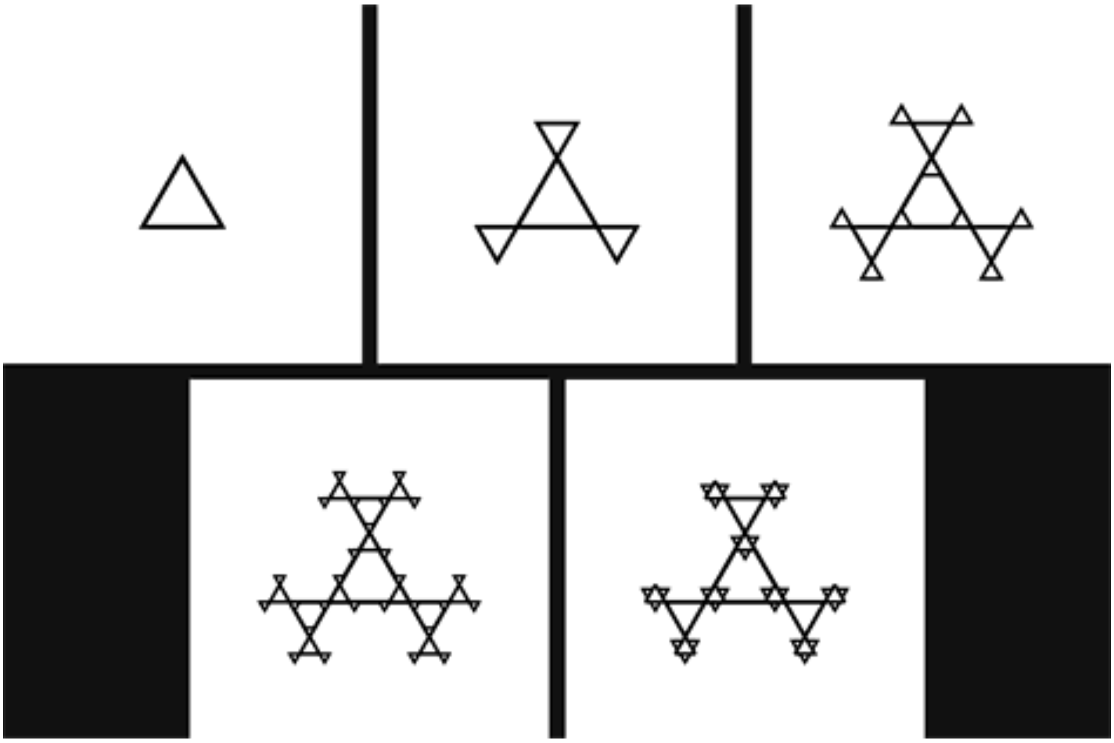


Fig.4. Example of a Visual Recursion Task stimulus. The first three iterations of a fractal generation are presented in the top row. The subject is then asked to choose, from the images in the lower row, which corresponds to the correct 4th iteration.

In theory, in order to correctly generalize a particular recursive rule to further iterations, subjects have to: 1) Acquire categorical knowledge about constituents (shape and position), 2) Recognize that constituents are structured hierarchically (with dominance and subordination relationships); 3) Recognize that constituents at different hierarchical levels display similar positional properties (e.g. 'each triangle has smaller triangles at its vertices') and 4) Apply the abstracted rule one level beyond the given.

To distinguish between recursion and embedded iteration, a non-recursive control task was also developed. In this task, iterative processes embed constituents within fixed hierarchical levels, without generating new levels (fig.5).

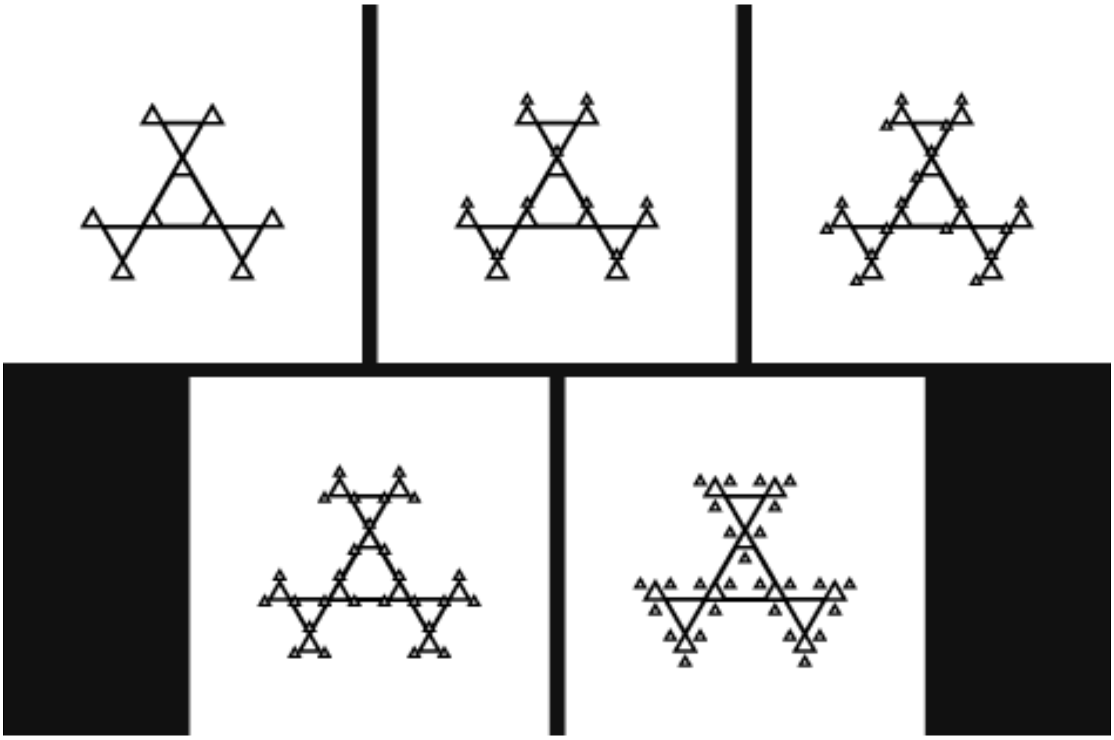


Fig.5. Example of a stimulus from the Visual Hierarchical Task. The procedure is similar to the Visual Recursion Task.

After validating the tasks, we tested different populations in both, together with some well-standardized cognitive measures of fluid intelligence and working memory (WASI matrix reasoning, digit span and corsi blocks). We found that:

a) The percentage of correct responses was lower in visual recursion than in embedding iteration (87% vs. 92%) and the response time was longer (18sec vs. 16sec).

b) Fluid intelligence was the best predictor of both Visual Recursion and Embedded Iteration (25% and 35% of the variance, respectively). However, while visual recursion accuracy was better predicted by verbal working memory than by spatial working memory; the opposite pattern was found in the embedded iteration task.

c) Taken together, embedded iteration and the processing component of verbal working memory accounted for 60.4% of visual recursion variance (*Embedded iteration*: Beta= 0.554, t=4.367, p<0.001; *Verbal working memory*: Beta= 0.404, t=3.184, p=0.004). Similarly, visual recursion and spatial working memory together predict 64.2% of embedded iteration variance (*Visual recursion*: Beta= 0.588, t=4.015, p=0.001; *spatial working memory*: Beta= 0.378, t=2,580 p=0.018).

If we take working memory as a measure of the ability to store and manipulate information, it is interesting that there is a dissociation in the modality of information processing that better predicts visual recursion and embedded iteration: 1) The ability to reverse the order of a given sequence of digits without losing track of the original sequence predicts accuracy in visual recursion even when all shared variance with embedded iteration is accounted for; 2) On the other hand, the ability to reverse a visuo-spatial sequence predicts accuracy in embedded iteration even when all shared variance with visual recursion is accounted for. This may reflect the fact that recursive hierarchies are more regular [57], hence can be better represented by compressed abstract rules. Once these rules are used to encode information across hierarchical levels, this might reduce the visual memory load necessary to represent each constituent individually [40, 47].

If these conclusions hold, the next empirical question is whether verbal processing resources are a necessary condition for recursive representations in the visual domain or whether they are recruited when available, given that they enhance reasoning in non-linguistic domains [55]. New studies are underway to assess if humans can perform above chance in VRT, under conditions of verbal and motor masking. If so, this would support the hypothesis that recursion, as an abstract representational property, can be used independently of language.

7. Recursion and syntax

The idea of universal grammar was developed to account for two facts: 1) That children are able to learn which syntactic constructions are allowed, and to apply them productively, beyond what seems possible from limited input; and 2) That there is an apparent deep syntactic similarity among different languages [58-60]. Both these facts could be explained by human cognitive biases [31, 38, 42, 45, 52, 61, 62], with origins in non-linguistic domains, provided that they would be flexible enough to be useful in a fast-changing cultural environment [51]. The ability (or predisposition) to represent hierarchical structures as containing self-similarity could be one of these biases.

A few have challenged the importance of recursion in language claiming that it is not used in all languages [63, 64] and that it is not very common in the languages where it is used [29]. These facts have lead some to the conclusion that the usage of recursion is a cultural option and not a requirement for the evolution of language. Although languages are cultural conventions (despite the fact that they might recruit innate cognitive abilities [65]), I think that the importance of recursion in the processes of acquisition and development of conventionalized communication systems should not be disregarded. There are several reasons to think so: 1) Children seem to be able to generalize recursive syntactic structures even though they are rare in the input. This seems to suggest that they are able to represent syntactic recursion *a priori* [28, 31, 53]; 2) Although recursion can be meaningfully used in prosody [32] and in discourse [30, 63], its application in syntax has the property of reducing the semantic ambiguity [31]. This expressive power might have been one of the reasons why recursion became available for syntax. However, tools that enhance expressiveness and reduce ambiguity may be less necessary in conditions where the linguistic content can be predicted by the context or by a shared cultural background.

This could explain why communication within communities [64] seems to be full of idioms and ungrammatical expressions [11].

The point is that the current rarity of recursive syntax in speech may not reflect its importance in the evolution and acquisition of language. The fact that languages can lose overt markers of clausal integration during glossogenetic history [30] seems to suggest that combinatoriality in syntax doesn't always evolve towards greater signal complexity.

Conclusion

A definition of recursion, such as "the ability to represent a succession of hierarchical dependencies (related by parenthood) with the same rules", by focusing on the representational and behavioural level, can be turned into empirical questions, which, provided with the appropriate methods, can be tested experimentally.

Under this definition, specific behavioural signatures of recursion can be outlined, namely the ability to take generative steps beyond the given (regarding hierarchical embedding). These traits not only distinguish recursion from hierarchical embedding and iteration, but also disqualify some abilities from being recursion-specific. For example, infinity can be also achieved by unlimited concatenation of finite elements using iteration [39]; Long-distance dependencies relate to hierarchical parsing and the usage of memory to process serial information [17, 18, 37, 66]; and cognitive grouping can create supra-hierarchical levels with existing constituents, but cannot recruit or create new constituents.

The availability of recursion as a cognitive ability (vs. hierarchical embedding) might have allowed the development of behavioural traits potentially advantageous in a wide range of domains: Problem solving, action sequencing, spatial navigation and social navigation [1, 9, 27, 28, 35, 45, 46]. Furthermore, since recursion can induce cognitive biases to interpret hierarchies as self-similar, if shared by a population, these biases might have increased the likelihood of the emergence of conventionalized communication systems [51, 52].

If recursion has been available in other domains before the faculty of language, it is also possible that further modifications could have occurred in the human lineage, for example, adaptations to serialized information or to symbolic representation [7, 10]. In the social and spatial domains, recursion may have been used prior to the ability to represent serial and symbolic information. This does not exclude that these processes were not further enhanced by these language-specific adaptations.

Another important issue concerns the nature of recursion. Although for operational reasons we define recursion as a monolithic construct, it can be an epiphenomenon resulting from the interaction of several cognitive abilities [30, 31, 42]. For example, if we consider the evolution from hierarchical embedding to recursion, there is one trait that seems to be crucial. That is the ability to compare (and match) relations between different hierarchical levels related by parenthood (level 1 is for level 2, as 2 is for level 3) and to further generalize the obtained rules to other ('mother' or 'child') levels or constituents [34]. Evidently, this entails that subjects are sensitive to the particular features targeted for the cross-level comparison, before similarity principles are extracted.

Differences between recursion and hierarchical processing can be addressed with the new methods presented in this manuscript – Visual Recursion Task and Embedded Iteration Task. Given that these methods are based on the representation of visuo-spatial information, they can provide insights regarding the question of whether recursion can be used independently of language. These insights will come from research with children, verbal masking, patients with aphasia, and from animal studies.

Whatever its precise role in language, recursion is an important property of human cognition and one that is used for the transmission of information. The exciting questions that it raises concerning our evolutionary history should be researched within a clear framework and one that allows the development of empirical approaches in different domains.

Acknowledgments

The author is grateful to Paolo Baldi for the insightful discussions. This work was supported by the FCT grant SFRH/BD/64206/2009 and by the ERC advanced grant SOMACCA project number 230604.

Bibliography

1 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.

2 Chomsky, N. 2010 Some simple evo devo theses: how true might they be for language? In The evolution of human language. (ed. R. K. Larson, V. M. Déprez,H. Yamakido), pp. 45-62. New York: Cambridge University Press.

3 Odifreddi, P. G. 1999 *Classical recursion theory*. Elsevier.

4 Hofstadter, D. R. 2008 *I Am a Strange Loop*. Basic Books.

5 Kilpatrick, J. 1985 Reflection and recursion. *Educational Studies in Mathematics*. **16**, 1-26.

6 Hofstadter, D. R. 1980 *Gödel, Escher, Bach: An Eternal Golden Braid*. Harmondsworth: Penguin.

7 Fitch, W. T. 2010 Three meanings of "recursion": key distinctions for biolinguistics. In The evolution of human language. (ed. R. K. Larson, V. M. Déprez,H. Yamakido), pp. 73-90. New York: Cambridge University Press.

8 Hulst, H. v. d. 2010 Re Recursion. In Recursion and Human Language. (ed. H. v. d. Hulst), pp. Berlin/New York: de Gruyter Mouton.

9 Fitch, W. T., Hauser, M.,Chomsky, N. 2005 The evolution of the language faculty: Clarifications and implications. *Cognition*. **97**, 179-210. (10.1016/j.cognition.2005.02.005)

10 Pinker, S.,Jackendoff, R. 2005 The faculty of language: what's special about it? *Cognition*. **95**, 201-236.

11 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**, 211-225. (doi: 10.1016/j.cognition.2005.04.006)

12 Gentner, T. Q., Fenn, K. M., Margoliash, D.,Nusbaum, H. C. 2006 Recursive syntactic pattern learning by songbirds. *Nature*. **440**, 1204-1207. (10.1038/nature04675)

13 Corballis, M. C. 2007 Recursion, Language, and Starlings. *Cognitive Science*. **31**, 697-704. (10.1080/15326900701399947)

14 Friston, K. 2010 The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*. **11**, 127-138.

15 Abe, K.,Watanabe, D. 2011 Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nature Neuroscience*. **14**, 1067-1074. (10.1038/nn.2869)

16 Fitch, W. T.,Hauser, M. D. 2004 Computational Constraints on Syntactic Processing in a Nonhuman Primate. *Science*. **303**, 377-280.

17 Bahlmann, J., Schubotz, R.,Friederici, A. 2008 Hierarchical artificial grammar processing engages Broca's area. *NeuroImage*. **42**, 525-534. (10.1016/j.neuroimage.2008.04.249)

18 Bahlmann, J., Schubotz, R. I., Mueller, J. L., Koester, D.,Friederici, A. D. 2009 Neural circuits of hierarchical visuo-spatial sequence processing. *Brain Research*. **1298**, 161-170. (10.1016/j.brainres.2009.08.017)

19 Bloomfield, T. C., Gentner, T. Q.,Margoliash, D. 2011 What birds have to say about language. *Nature Neuroscience*. **14**, 947-948. (10.1038/nn.2884)

20 Penrose, R. 1989 *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics.*: Oxford University Press.

21 Eglash, R. 1997 Bamana Sand Divination: Recursion in Ethnomathematics. *American anthropologist*. **99**, 112-122. (10.1525/aa.1997.99.1.112)

22 Lerdahl, F.,Jackendoff, R. 1996 *A generative theory of tonal music*. The MIT Press.

23 Jackendoff, R.,Lerdahl, F. 2006 The capacity for music: what is it, and what's special about it? *Cognition*. **100**, 33-72.

24 Eglash, R. 1998 Fractals in African settlement architecture. *Complexity*. **4**, 21-29.

25 Eisenberg, M. 2008 Recursion—or, Better Computational Thinking Through Laughter. *International Journal of Computers for Mathematical Learning*. **13**, 171-174.

26 Miller, S. A. 2009 Children's Understanding of the Second-Order Mental States. *Psychological Bulletin*. **135**, 749-773.

27 Schiemenz, B. Year Managing Complexity by Recursion. In: R. Trappl, editor. Cybernetics And Systems; 2002; Wien; 2002. p. 475–479.

28 Pulvermüller, F.,Fadiga, L. 2010 Active perception: sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience*. **11**, 351-360. (10.1038/nrn2811)

29 Karlsson, F. 2010 Syntactic recursion and iteration In Recursion and Human Language. (ed. H. v. d. Hulst), pp. 43-67. Berlin/New York: de Gruyter Mouton.

30 Mithun, M. 2010 The fluidity of recursion and its implications. In Recursion and Human Language. (ed. H. v. d. Hulst), pp. 17-41. Berlin/New York: de Gruyter Mouton.

31 Roeper, T. Year The minimalist microscope: How and where interface principles guide acquisition. Boston University Conference on Child Language; 2009; 2009.

32 Hunyady, L. 2010 Cognitive grouping and recursion in phonology. In Recursion and Human Language. (ed. H. v. d. Hulst), pp. Berlin/New York: de Gruyter Mouton.

33 Picard, O., Massé, A. B., Hanard, S., Marcotte, O., Chicoisne, G.,Gargouri, Y. Hierarchies in Dictionary Definition Space. arXiv:0911.5703v1 [cs.CL] 2010.

34 Hofstadter, D. R. 2000 Analogy as the core of cognition. In The Best American Science Writing. (ed. J. C. a. J. Gleick), pp. 116-144.

35 Flack, J., Jackendoff, R., Krakauer, D. C.,DeDeo, S. A Theory for "Grammar" for Primate Social Behaviour. 2011.

36 Cheney, D.,Seyfart, R. M. 2007 *Baboon Metaphysics*. The University of Chicago Press.

37 Kinsella, A. 2010 Was recursion the key step in the evolution of the human language faculty? In Recursion and Human Language. (ed. H. v. d. Hulst), pp. 179-191. Berlin/New York: de Gruyter Mouton.

38 Pinker, S.,Bloom, P. 1990 Natural Selection and Natural Languages. *Behavioral and Brain Sciences* **13**, 707-784.

39 Lieberman, P. 2010 The creative capacity of language, in what manner is it unique, and who had it? In The evolution of human language. (ed. R. K. Larson, V. M. Déprez,H. Yamakido), pp. 163-175. New York: Cambridge University Press.

40 Koike, H.,Yoshihara, H. Year Fractal approaches for visualization of huge hiearchies. Proceedings of the 1993 IEEE Symposium on Visual Languages; 1993: IEEE/CS; 1993. p. pp.55-60.

41 Verhagen, A. 2010 What do you think is the proper place of recursion? Conceptual and empirical issues. In Recursion and Human Language. (ed. H. v. d. Hulst), pp. 93-110. Berlin/New York: de Gruyter Mouton.

42 Harder, P. 2010 Over the top - recursion as a functional option. In Recursion and Human Language. (ed. H. v. d. Hulst), pp. 233-244. Berlin/New York: de Gruyter Mouton.

43 Maruyama, T., Takagi, M.,Hoshino, T. Hardware Implementation Techniques for Recursive Calls and Loops. Lecture Notes in Computer Science 1673—The Ninth International Workshop, FPL'99. Glasgow, UK 1999:pp. 450–455.

44 Jackendoff, R. 2009 Parallels and nonparallels between language and music. *Music Perception*. **26**, 195-204.

45 Levy, S. D. 2010 Becoming recursive: Towards a computational neuroscience account of recursion in language and thought. In Recursion and Human Language. (ed. H. v. d. Hulst), pp. 371-391. Berlin/New York: de Gruyter Mouton.

46 Gardenfors, P.,Osvath, M. 2010 Prospection as a cognitive precursor to symbolic communication. In The evolution of human language. (ed. R. K. Larson, V. M. Déprez,H. Yamakido), pp. 103-114. New York: Cambridge University Press.

47 Alvarez, G. A. 2011 Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*. **15**, 122-131.

48 Duncan, J. 2010 The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends in Cognitive Sciences*. **14**, 172-179. (10.1016/j.tics.2010.01.004)

49 Nowak, M. A., Komarova, N. L.,Niyogi, P. 2001 Evolution of universal grammar. *Science*. **291**, 114-118.

50 Shams, L. 2010 Causal inference in perception. *Trends in Cognitive Sciences*. **14**, 425-432.

51 Chater, N.,Christiansen, M. H. 2010 Language Acquisition Meets Language Evolution. *Cognitive Science*. **34**, 1131-1157.

52 Smith, K.,Kirby, S. 2008 Cultural evolution: implications for understanding the human language faculty and its evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*. **363**, 3591-3603. (10.1098/rstb.2008.0145)

53 Perfors, A., Tenenbaum, J., Gibson, E.,Regier, T. 2010 How recursive is language? A Bayesian exploration. In Recursion and Human Language. (ed. H. v. d. Hulst), pp. 159-175. Berlin/New York: de Gruyter Mouton.

54 Bickerton, D. 2003 Symbol and structure: A comprehensive framework for language evolution. In Language evolution. (ed. M. H. Christiansen,S. Kirby), pp. 77-93: Oxford University Press.

55 Carruthers, P. 2002 The cognitive functions of language. *Behavior and Brain Sciences*. **25**, 657-674.

56 Martins, M. J.,Fitch, W. T. Fractal geometry and visual recursion. Wien: University of Vienna 2011.

57 Helm, P. v. d., van Lier, R. J.,Leeuwenberg, E. J. 1992 Serial pattern complexity: irregularity and hierarchy. *Perception*. **21**, 517-544.

58 Chomsky, N. 1965 *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

59 Chomsky, N. 1980 *Rules and representations*. Oxford, England: Blackwell.

60 Pinker, S. 1994 *The language instinct: How the mind creates language*. New York: William Morrow and Company.

61 Dor, D.,Jablonska, E. 2010 Plasticity and canalization in the evolution of linguistic communication: an evolutionary developmental approach. In The evolution of human language. (ed. R. K. Larson, V. M. Déprez,H. Yamakido), pp. 135-147. New York: Cambridge University Press.

62 Endress, A. D., Nespor, M.,Mehler, J. 2009 Perceptual and memory constraints on language acquisition. *Trends in Cognitive Sciences*. **13**, 348-353.

63 Everett, D. L. 2009 Pirahã Culture and Grammar: A Response to Some Criticisms. *Language*. **85**, 405-442.

64 Sakel, J.,Stapert, E. 2010 Pirahã - In need of recursive syntax? In Recursion and Human Language. (ed. H. v. d. Hulst), pp. 3-16. Berlin/New York: de Gruyter Mouton.

65 Koster, J. 2010 Recursion and the Lexicon. In Recursion and Human Language. (ed. H. v. d. Hulst), pp. 285-297. Berlin/New York: de Gruyter Mouton.

66 Friederici, A. D. 2009 Pathways to language: fiber tracts in the human brain. *Trends in Cognitive Sciences*. **13**, 175-181.

3. Investigating recursion within a domain-general framework

Martins, M.D., Fitch, W.T. (2014). Investigating recursion within a domain-general framework. In *Language and Recursion* (Eds. F. Lowenthal & L. Lefebre), pp. 15-26. Springer, New York.

Investigating recursion within a domain-general framework

Maurício Dias Martins[1,2]
W. Tecumseh Fitch[1]

[1]Department of Cognitive Biology, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria
[2]Language Research Laboratory, Lisbon Faculty of Medicine, Av. Professor Egas Moniz, 1649-028 Lisbon, Portugal

1. Introduction

Recursion has long been an object of interest and fascination from scholars in different fields such as mathematics, computer science, linguistics and visual arts, partly because recursion allows the generation of structures that are simple and complex at the same time. Recursive structures are complex because they can contain infinite hierarchical levels, yet simple because this infinity can be achieved and represented using very simple rules.

One famous class of recursive structures are the fractals (Fig.1). Fractals are structures that display self-similarity (Mandelbrot, 1977), so that they appear geometrically similar when viewed at different scales. Fractals are produced by simple rules that, when applied iteratively to their own output, can generate complex hierarchical structures. Since the same kind of representation can be used at different levels of depth, simple rules suffice to represent the whole structure.

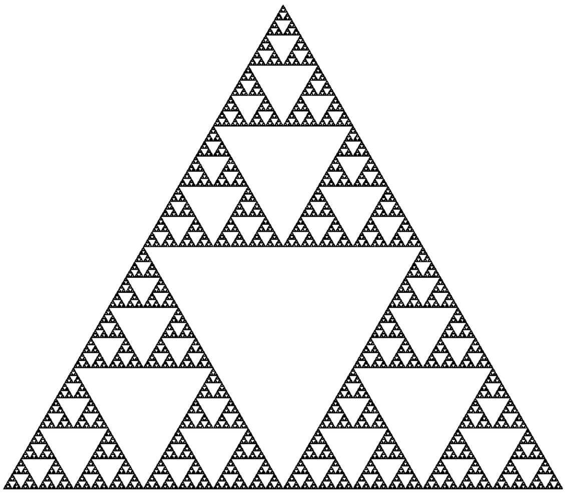


Fig.1. Example of a fractal structure – a Sierpinski gasket - exhibiting self-similarity.

Humans have long understood this generative potential at an intuitive level. In ancient Egypt we find depictions of recursive structures representing the self-generating power of the universe (Eglash, 1997) and the Sierpinki triangles shown in Fig. 1 is found in the Anagni Cathedral, Italy (Wolfram, 2002). In other contemporary, pre-industrialized cultures, we find recursive depth as a symbolic representation of status or power (Eglash, 1998).

Due to these curious properties, several contemporary thinkers have proposed that recursion, when available to the human mind, could have been associated with the emergence of our unbounded creativity (Hofstadter, 1980; Penrose, 1989). When available to different domains such as language (Hauser, Chomsky, & Fitch, 2002), problem solving (Schiemenz, 2002), spatial reasoning etc., recursion could have allowed an open-ended generative power.

Unfortunately the investigation of recursion as a cognitive ability has proven to be difficult for several reasons: First, multiple definitions of recursion exist in different fields of research. When used interchangeably in the literature, such discrepant interpretations hinder mutually-consistent interpretation of empirical findings. Second, there are multiple levels of analysis at which recursion can be measured, at least including algorithm, structure, and mental representation. Analysing one of these levels does not necessarily allow inferences about the other two. For example, both recursive and non-recursive algorithms can be used to generate recursive structures. Therefore we cannot conclude that systems able to generate recursive structures necessarily use a recursive algorithm. Furthermore, the structural properties attributed to an object depend on the representational abilities of the observer. It is certainly possible for an individual to generate structures without representing them mentally (e.g. one's own heartbeat time series). Third, if we assume that the mind is modular, we might feel tempted to discuss recursion in a restrictive and domain-specific fashion, for example, in linguistic terms (Chomsky, 2010; Hornstein & Pietroski, 2009; Roeper, 2011). But it remains an open empirical question whether recursion is domain-general or domain-specific, and attempts to address this issue must therefore start with a definition compatible with both possibilities.

Here we attempt to address these questions systematically, highlighting some crucial distinctions, and laying out a grid of empirical hypothesis. We will also provide examples of syntactic and visuo-spatial recursion, illustrating how a single framework can be applied to different domains.

2. Defining recursion

In general, the term recursion has been used to characterize to the process of embedding a constituent inside another constituent of the same kind (Fitch, 2010; Hulst, 2010; Pinker & Jackendoff, 2005). Recursive processes can generate hierarchical structures that display similar properties across different levels of embedding. This feature, called self-similarity, is a signature of recursive structures. In language, this process establishes a dependency relationship between two constituents of the same category. An example of a recursive linguistic structure is the compound noun "[[student] committee]", where we find a noun phrase embedded inside another noun phrase. In contrast, a sentence with a noun and a verb together, such as "[[trees] grow]", is hierarchical, but not recursive, because a constituent of a given type is not nested within a constituent of that same type.

Although recursion has been hypothesized as a uniquely human trait and as a necessary capacity for the evolution of language (Hauser et al., 2002), the diversity of definitions in use has prevented the consistent interpretation of empirical results (Fitch, 2010). One of the major problems is the difficulty in establishing clear distinctions between recursion and similar processes such as hierarchical embedding and iteration (Hulst, 2010).

According to the framework that we adopt (Fitch, 2010; Martins, 2012), “iteration” refers to the process of repeating an operation a given number of times. Such processes may or may not generate hierarchical structures or create dependency relationships between different elements. For example, adding one marble at a time to a bag is an iterative process, but neither hierarchical nor recursive. On the other hand, “hierarchical” structures always involve the embedding of elements within other elements. This embedding can refer to the grouping of a set of constituents within a higher order element, such as the grouping of individuals within a family; or it can refer to the establishment of dominance-subordination relationships such as in social hierarchies. If the hierarchical embedding occurs between constituents of the same category (e.g. such as a noun phrase inside a noun phrase) we classify it as recursive, otherwise as non-recursive. Iteration, hierarchical embedding and recursion are not mutually exclusive processes: recursion typically involves both hierarchy and iteration. Nevertheless, it is important to conceptually segregate the cognitive abilities necessary to represent the kind of information that each of these processes encode (Fig.2). For instance, encoding iteration requires the ability to represent the repetition of constituents/elements. Encoding hierarchical embedding requires the ability to represent dependency or grouping relationships between constituents. Finally, encoding recursive embedding requires the ability to represent successive hierarchical dependencies (hierarchical levels related by parenthood) with the same rules. A specific behaviour trait that this ability enables is the possibility to generate new hierarchical levels beyond those previously experienced or specified, maintaining consistency with existing levels at a higher level of abstraction.

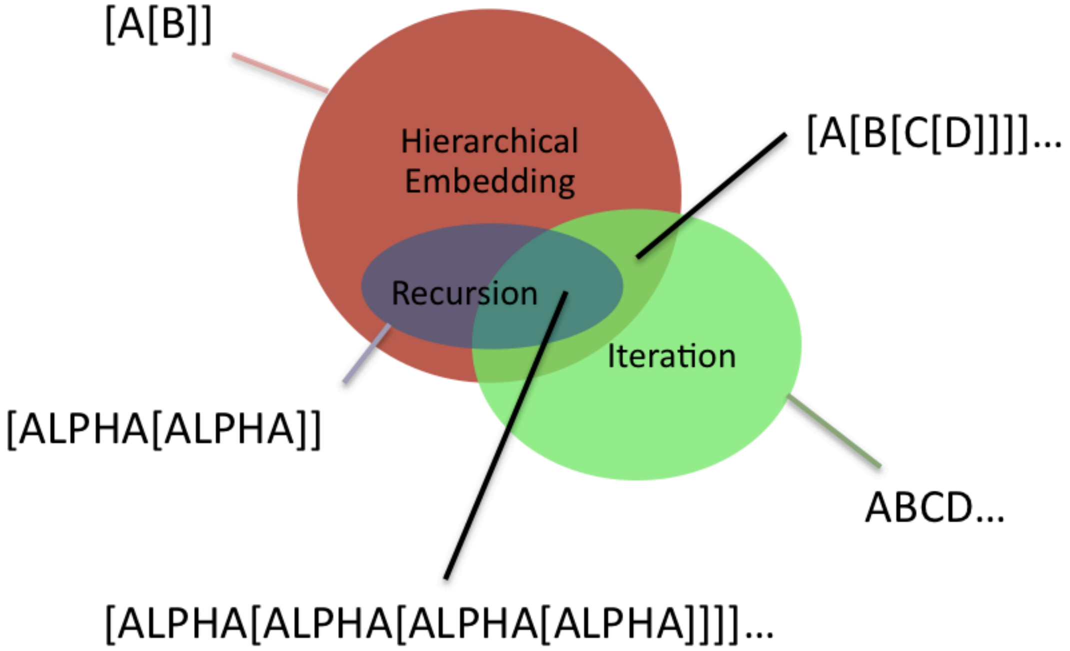


Fig.2. Examples of structures produced by iteration, hierarchical embedding and recursion and by the combination of these processes. A non-iterated hierarchical embedding corresponds to the establishment of a dependency but without repetition. The ability to represent repetition and the ability to represent dependencies may be orthogonal.

2.1. Iteration, Hierarchy and Recursion: Some Illustrative Examples

Nature provides nice illustrative examples of the distinction between iteration, hierarchy and recursion (Fig 3). Some marine algae grow in a recursive fashion, illustrating self-similarity ("self-embedding"). Multiple, hierarchical growth tips remain undifferentiated, and can in principle spawn an endless proliferation of further growth tips. Plants such as grasses can grow by propagating copies along a single extension (a *stolon* or *rhizome*) and illustrate a serial, iterative structure. Iteration can thus exist without hierarchy. Trees and shrubs provide a nice example of hierarchy, because growth occurs in parallel, at multiple growth points, but this becomes non self-embedding as soon as differentiation into leaves or flowers occurs: branches bear twigs which bear leaves, but the opposite does not occur. It is interesting to note in this case that the more primitive and ancient plants show recursion, while "advanced" plants like angiosperms do not (Niklas, 1997). Another ubiquitous example of non-recursive hierarchy is found in chemistry and atomic physics. Compounds are made of molecules, which are made of atoms made of electrons, neutrons and protons, the latter composed of quarks. As these examples show, many real-world examples of hierarchy occur without recursive self-similarity.

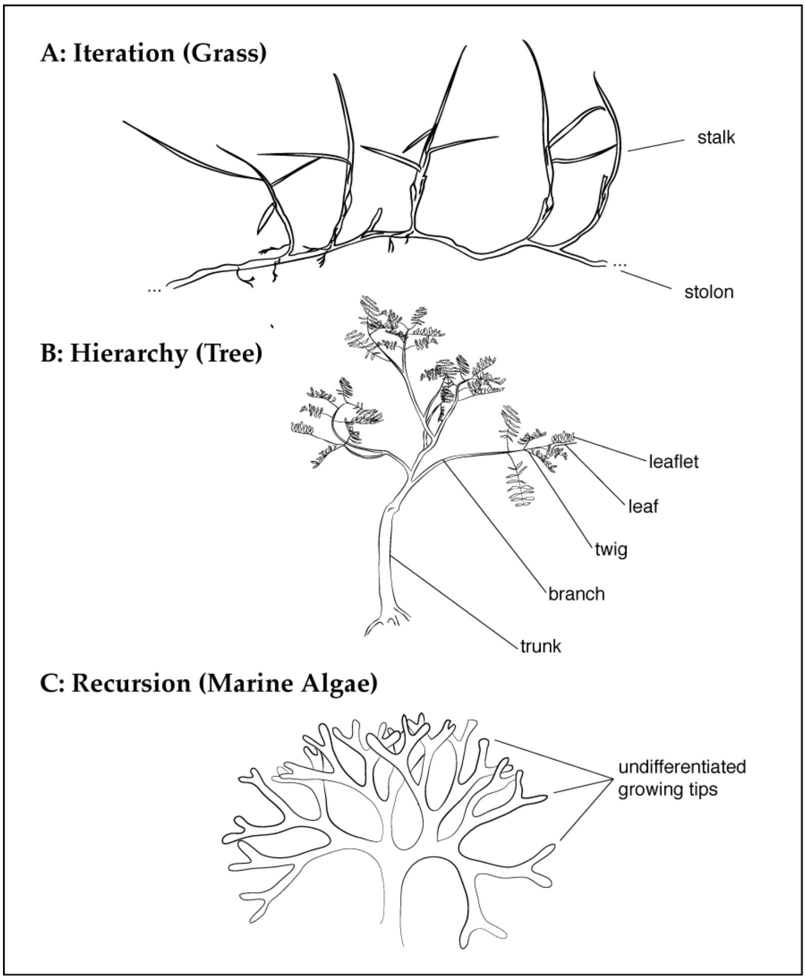


Fig.3. Botanical Examples of Iteration, Hierarchy and Recursion

Fig 3A: Grass lateral growth illustrating **Iteration**: Many grasses grow by lateral extension, via above-ground stolons or below-ground rhizomes. Growth along the stolon is iterative: each addition of a new stalk happens singly and independently, with no consequences for other stalks.

Fig 3B: Growth of a tree illustrating non-recursive **Hierarchy**: Growth occurs in parallel, at many different terminals, but is differentiated such that branches bear twigs which bear leaves (which may be compound and made up of leaflets). Leaves cannot bear twigs, and each level is of a different type: this is an example of hierarchy without recursive self-embedding.

Fig 3C: Growth of a marine algae illustrating botanical **Recursion**: As the algae grows, every growth tip can undergo unbounded further subdivision, with no necessary differentiation, into twigs or leaves, and each tip is a potential new plant.

To further exemplify the difference between hierarchical embedding and recursion consider the following algorithm, which specifies how some letters of the alphabet can be incorporated in a hierarchical structure, in dependency to other letters:

"Execute one of the following rules: {1) Incorporate one or more [B]s in dependency of [A]; 2) Incorporate [C]s in dependency of [B]; 3) Incorporate [D]s in dependency of [C]}. Repeat as desired."

With such an algorithm one could generate structures such as the one depicted in Fig.4.

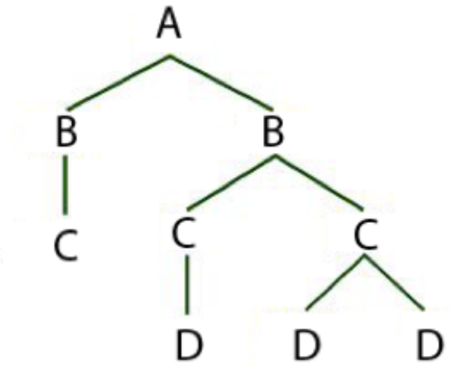


Fig.4. Example of a hierarchical structure generated by a non-recursive algorithm.

This algorithm works fine to handle a predefined number of hierarchical levels. It even allows the incorporation of infinite [C]s within the hierarchy in dependency of a certain [B]. However, if one would like to incorporate an [E] in the structure, this would not be possible without adding a new rule to the algorithm, one that specifies *a priori* how [E] can interact with the other letters. Thus, in non-recursive hierarchical embedding, for each hierarchical level that is generated, a specific rule needs to be specified. Recursion overcomes this limitation.

For instance, consider the algorithm defined as: “Embed one or more members of the ALPHABET within another member of the ALPHABET”. In this example, because all members of the alphabet ([A], [B], [C], [D], [E], etc.) are categorized as having the same properties (regarding the way they interact with the levels above and below) we could incorporate in the hierarchy elements that were not explicitly pre-specified (such as [E] or [F]). Furthermore, we can potentially generate infinite hierarchical levels with one single rule, illustrating the power of recursion to go beyond mere hierarchy.

Clearly, grammars composed only of recursive rules run the risk of being too powerful; allowing the over-generation of structures beyond what would be useful, for example, for transmitting information. The power of recursion, including in language, is only apparent when recursive rules are combined with non-recursive rules (Perfors, Tenenbaum, Gibson, & Regier, 2010). But for this combination to exist, the ability to represent recursion must be present in the first place.

In summary, we propose that the ability to generate novel hierarchical levels beyond those previously specified is a signature trait of recursion, and this trait should be an important object of empirical research aiming to tap into recursive abilities.

3. Empirical analysis

Before discussing possible empirical approaches for evaluating recursion, we alert the reader to a widespread misconception: that the formal grammar A^nB^n provides a litmus test for recursion. To our knowledge the first use of the A^nB^n grammar in empirical research was Fitch & Hauser (2004) (Fitch & Hauser, 2004), who proposed it as a test for pattern-perception abilities above the regular (= finite state grammar) level, and thus as a test for George Miller’s “supra-regular hypothesis”. Miller hypothesized that humans have a propensity to use context-free or other supra-regular grammars when perceiving patterns (G. A. Miller, 1967). This is a hypothesis about level in the formal language hierarchy (regular versus context-free), which is quite different from the issue of recursion, which can occur at any level of this hierarchy. Fitch & Hauser (2004) tested Miller’s hypothesis for humans and monkeys, and did not even mention recursion. Unfortunately, both a commentary on that paper (Premack, 2004) and several subsequent empirical papers have promulgated the misconception that success on the A^nB^n grammar reliably indicates recursive abilities (Abe & Watanabe, 2011; Gentner, Fenn, Margoliash, & Nusbaum, 2006; Marcus, 2006) Despite repeated critiques to the contrary (e.g. Corballis, 2007; Fitch, 2010; Fitch & Friederici, 2012) this seems to be an idea that will not die.

As stated in the introduction, there are 3 levels at which we can investigate the presence of "recursion": 1) The level of the algorithm, 2) the level of structure and 3) the level of representational abilities. Although all levels are interesting, we should not draw inferences about the presence of recursion in one level based on investigations of another. For instance, there are 2 different processes that can be used to generate Fig.5c. One of these processes is recursive, since it involves self-embedding, and the other is iterative and hierarchical, but not recursive.

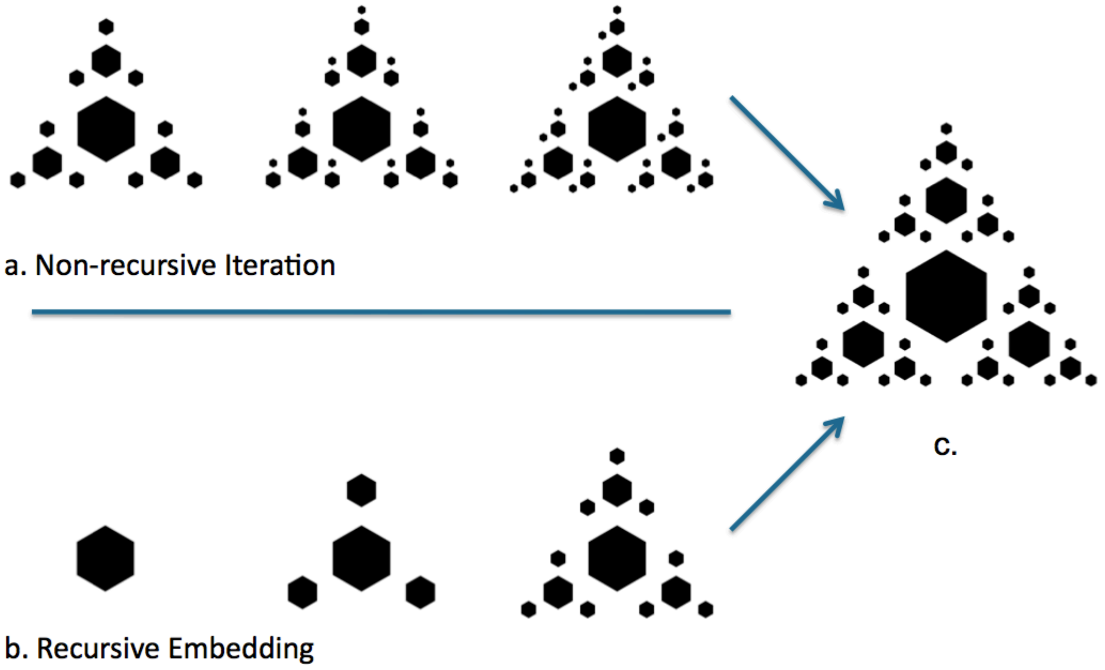


Fig5. Example of a recursive structure (c) generated by non-recursive iteration (a) and recursive embedding (b).

At the level of structure, Fig.5c can be plausibly modelled as being recursive, since it contains self-similarity. However, if a certain individual is able to generate Fig.5c, he could have used either the recursive procedure (b) or the non-recursive procedure (a). This observation implies that the ability to generate recursive structures such as (c) is neither necessarily informative regarding the ability to use recursive procedures, nor the ability to build cognitive representations of recursion. However, there is an empirical way to tackle this issue: If we expose an individual to the first 3 steps of the recursive procedure depicted in Fig.5b, and he is able to generate Fig.5c within that context, then we can assume that we was able to: 1) represent the underlying rules connecting the previous iterative steps, and 2) apply these rules productively to generate one step further. This is compatible with one clear definition of recursion: “An ability to represent multiple levels of hierarchical dependencies (hierarchical levels related by parenthood) following the same rules, entailing the possibility to generalize and produce new levels of embedding beyond those specified *a priori* (in the algorithm or in the input)”.

3. Modularity and domain-specificity issues

Recently, the development of the human ability to represent recursion has been described as an important step in the evolution of language (Fitch, Hauser, & Chomsky, 2005; Hauser et al., 2002). This ability may have allowed language to take on an unprecedented generative power. For example, recursion allows the representation and generation of sentences where a noun phrase $n+1$ is embedded inside a noun phrase n already embedded in a noun phrase $n-1$ ($[NP_{(n-1)}[NP_{(n)}[NP_{(n+1)}]]]$), as occurs in the sentence “John’s sister’s house”. Besides this gain in generative power, also the speed of acquisition of a certain language may be enhanced by the presence of prior recursive rules in the grammar system (Perfors et al., 2010).

Recursion, as it is used in language, has been hypothesized to be part of a “linguistic computational system […], independent of the other systems with which it interacts and interfaces”, potentially restricted to humans (Fitch et al., 2005; Hauser et al., 2002). According to this view, 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), humour (Eisenberg, 2008), second-order theory of mind (S. A. Miller, 2009), problem solving (Schiemenz, 2002), or action sequencing (Pulvermüller & Fadiga, 2010), these uses may all rely upon a system of abstract arbitrary symbol manipulation dependent on language (Fitch et al., 2005). Alternatively, some authors have proposed that the usage of recursion in some domains, for example in visual perception, can occur independently of language (Pinker & Jackendoff, 2005).

The idea of recursion as part of a linguistic computational system that is independent of other systems presupposes some concept of modularity. A module, according to Fodor (Fodor, 1983), implies encapsulation and domain-specificity. If recursion is a module then there are two possibilities: 1) Either recursion is a monolithic operation that doesn't recruit domain-general computations; or 2) Recursion results from the interaction of several sub-operations, all domain-specific, and none exchanges information with the external neural milieu during recursion-internal operations.

Although it can be useful to think about the mind as a composite of several encapsulated operations, we think there are some dangers that result from a strict, or "massively modular" view. If we entertain the possibility that the ability to represent recursion may result from the interaction of several abilities, some domain-specific and some domain-general, then investigations of the fine-grained structure of recursive operations require broad definitions beyond any one single domain. Currently, there are several plausible hypotheses regarding this fine grained-structure: H1) Different recursive-type operations in different domains are completely independent; H2) A single recursive module is recruited by different modalities; H3) There is some overlap between recursive operations in different domains, together with some dissociations owing to domain-specific computations and/or interface constrains.

Given that the current empirical evidence doesn't allow the elimination of any of these hypotheses, we advocate a definition of recursion that is compatible with several domains. The definition of recursion that we proposed above satisfies this desideratum. The concern about assuming domain-specificity is reinforced by recent empirical research in cognitive sciences, where domain-specific activities have been shown to depend more on domain-general operations than previously supposed. This is true, for example, in visual (Kirkham, Slemmer, & Johnson, 2002), verbal (Saffran, Aslin, & Newport, 1996) and motor pattern-extraction (Baldwin, Andersson, Saffran, & Meyer, 2008) as well as in social reasoning (McKinnon & Moscovitch, 2007), music perception (Treuhub & Hannon, 2006) and number/quantity representation (Holloway & Ansari, 2008).

It thus seems likely to us that some components of recursive mental operations may be domain-general and other components highly modularized and domain-specific. Domain-specificity may be especially true regarding operations associated with so-called ‘fast thought’ or expert behaviour.

According to (Kahneman, 2011) and many others, human cognition uses two separated systems: One fast, intuitive, and heuristic; and another slow, effortful, and abstract. It is possible that the generation of abstract recursive representation rules is domain-general. This would imply a transferability of representational knowledge across domains. However, it is also possible that the representation of recursive structures is achieved not by the generation of abstract rules but by matching the contents of perception with previously acquired (or biologically endowed) templates. Some empirical reports seem to suggest that the processing of familiar structures depends less on domain-general cognitive abilities than the processing of unfamiliar structures. This seems to be true for syntactic processing (Novik, Trueswell, & Thompson-Schill, 2005), visual perception (Sinha & Balas, 2008), social reasoning (McKinnon & Moscovitch, 2007) and may reflect a progressive automatization and “lexicalization” of structural knowledge, which complements more abstract and flexible representations (Brinton, 2008).

How exactly recursion is represented by different species, in different domains and ontogenetic stages is an open question that can be investigated empirically, for example, with dual task-paradigms.

4. Conclusion

In this chapter we outlined some difficulties concerning the assessment of recursion as a cognitive ability, and offered a framework that allows specific questions to be assessed separately:

1) Is the ability to represent recursion cognitively distinct from the ability to represent similar operations such as hierarchical embedding and iteration?
2) Are these abilities predicted by general intelligence?
3) Is the ability to represent recursion present in more than one domain?
4) If present in more than one domain, is there a single recursion module that is recruited by different modalities; or is the execution or recursion-type operations achieved by (partially or totally) different cognitive resources in different domains?

5) Are there any cognitive abilities or operations that constitute strict causal precursors of recursion, in the absence of which recursion cannot be represented (e.g. language)?

6) If recursive capacities are present in other species or in multiple domains, are they achieved via a flexible, slow, and abstract representational system; or via an automatic and rigid template-matching system?

Our recent work (Martins et al., submitted.) begins to answer some of these questions. For instance, our results demonstrate that recursion can be represented in the visuo-spatial domain, and that there is some cognitive dissociation between recursion, general intelligence, and hierarchical embedding. Interestingly, both behavioural and imaging data seem to suggest that recursion recruits visual-spatial-specific resources less than hierarchical embedding and relies more on domain-general resources. Finally, the ability to represent recursion in the visual domain seems to be associated with high reaction times initially. However, with practice, these reaction times decrease. Furthermore, performance in latter trials seem to recruit different cognitive resources in comparison with initial trials, perhaps reflecting the transition from an abstract representational strategy to a more automatic one.

However represented, and whatever its precise role in different cognitive domains, recursion is an important and powerful property of human cognition. The fundamental questions that it raises concerning human cognition and our evolutionary history require a clear framework that both allows and encourages the development of empirical approaches spanning different cognitive domains.

Acknowledgements

The authors are grateful to Nadja Kavcik, for her assistance in generating the pictures of this chapter. This work was supported by the FCT grant SFRH/BD/64206/2009 to MM and by ERC Advanced Grant SOMACCA, project number 230604 to WTF.

Bibliography

Abe, Kentaro, & Watanabe, Dai. (2011). Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nature Neuroscience, 14*(8), 1067-1074.

Baldwin, D., Andersson, A., Saffran, J., & Meyer, M. (2008). Segmenting dynamic human action via statistical structure. *Cognition, 106*(3), 1382-1407.

Brinton, L. J. (2008). Where grammar and lexis meet. In E. Seoane & M.J. López-Couso (Eds.), *Theoretical and empirical issues in grammaticalization*. Amsterdam: John Benjamins Publishing Co.

Chomsky, N. (2010). Some simple evo devo theses: How true might they be for language? In Richard K Larson, Viviane M Déprez & Hiroko Yamakido (Eds.), *The evolution of human language* (pp. 45-62). New York: Cambridge University Press.

Corballis, Michael C. (2007). Recursion, language, and starlings. *Cognitive Science, 31*(4), 697-704. doi: 10.1080/15326900701399947

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.

Fitch, W.T. (2010). Three meanings of "recursion": Key distinctions for biolinguistics. In Richard K Larson, Viviane M Déprez & Hiroko Yamakido (Eds.), *The evolution of human language* (pp. 73-90). New York: Cambridge University Press.

Fitch, W.T., & Friederici, A. (2012). Artificial grammar learning meets formal language theory: An overview. *Philosophical Transactions of the Royal Society B, 367*, 1933-1955.

Fitch, W.T., Hauser, M., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition, 97*(2), 179-210. doi: 10.1016/j.cognition.2005.02.005

Fitch, W.T., & Hauser, M.D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science, 303*, 377-280.

Fodor, J.A. (1983). *The modularity of the mind*. Cambridge, MA: Bradford.

Gentner, T.Q., Fenn, K.M., Margoliash, D., & Nusbaum, H.C. (2006). Recursive syntactic pattern learning by songbirds. *Nature, 440*(7088), 1204-1207. doi: 10.1038/nature04675

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.

Holloway, I.D., & Ansari, D. (2008). Domain-specific and domain-general changes in children's development of number comparison. *Developmental Science, 11*(5), 644-649.

Hornstein, N., & Pietroski, P. (2009). Basic operations: Minimal syntax-semantics. *Catalan Journal of Linguistics, 8*, 113-139.

Hulst, H. v.d. (2010). Re recursion. In Harry van der Hulst (Ed.), *Recursion and human language*: de Gruyter Mouton.

Jackendoff, R., & Lerdahl, F. (2006). The capacity for music: What is it, and what's special about it? *Cognition, 100*(1), 33-72.

Kahneman, D. (2011). *Thinking fast and slow*. New York: Farrar, Straus and Giroux.

Kirkham, N.Z., Slemmer, J.A., & Johnson, S.P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, B35-B42.

Mandelbrot, B. (1977). *The fractal geometry of nature*: Freeman.

Marcus, Gary. (2006). Startling starlings. *Nature, 440*, 1204-1207.

Martins, M.D. (2012). Specific signatures of recursion. *Philosophical Transactions of the Royal Society B, 367*, 2055-2064.

McKinnon, M.C., & Moscovitch, M. (2007). Domain-general constributions to social reasoning: Theory of mind and deontic reasoning re-explored. *Cognition, 102*, 179-218.

Miller, G. A. (1967). Project grammarama. In George A Miller (Ed.), *Psychology of communication*. New York: Basic Books.

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

Niklas, Karl J. (1997). *The evolutionary biology of plants*. Chicago: University of Chicaog Press.

Novik, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of broca’s area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience, 5*(3), 263-281.

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

Perfors, A., Tenenbaum, J., Gibson, E., & Regier, T. (2010). How recursive is language? A bayesian exploration. In Harry van der Hulst (Ed.), *Recursion and human language*: de Gruyter Mouton.

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

Premack, David. (2004). Is language the key to human intelligence? *Science, 303*, 318-320.

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

Roeper, T. (2011). The acquisition of recursion: How formalism

articulates the child's path. *Biolinguistics, 5*(1-2), 57-86.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926-1928.

Schiemenz, B. (2002). *Managing complexity with recursion.* Paper presented at the Cybernetics and Systems, Wien 2002.

Sinha, P., & Balas, B. J. (2008). Computational modeling of visual information processing. In R. Sun (Ed.), *The cambridge handbook of computational psychology* (pp. 612-634). Cambridge: Cambridge University Press

Treuhub, S.E., & Hannon, E.E. (2006). Infant music perception: Domain-general or domain-specific mechanisms? *Cognition, 100*, 73-99.

Wolfram, Stephen. (2002). *A new kind of science*. Champagne, IL: Wolfram Media.

4. Fractal geometry and visual recursion: a novel approach to hierarchical self-similarity

Martins M, Martins IP, Fitch WT (submitted to Psychological Methods). Fractal geometry and visual recursion: a novel approach to hierarchical self-similarity.

Fractal geometry and visual recursion: a novel approach to hierarchical self-similarity

Maurício Dias Martins[1,2]
Isabel Pavão Martins [2]
W. Tecumseh Fitch[1]

[1]Department of Cognitive Biology, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria
[2]Language Research Laboratory, Lisbon Faculty of Medicine, Av. Professor Egas Moniz, 1649-028 Lisbon, Portugal

Corresponding author:
Mauricio Martins
mauricio.martins@univie.ac.at

Abstract

We describe a new method to explore recursive cognition in the visual domain. We define recursion as the ability to represent multiple hierarchical levels using the same rule, entailing the ability to generate new levels beyond those previously encountered. With this definition recursion can be distinguished from general hierarchical embedding. To investigate this recursion/hierarchy distinction in the visual domain, we developed two novel methods: The Visual Recursion Task (VRT), in which an inferred rule is used to represent new hierarchical levels, and the Embedded Iteration Task (EIT), in which additional elements are added to an existing hierarchical level. We found that adult humans can represent recursion in the visuo-spatial domain, and that this ability is both distinct from general intelligence and from the ability to represent iterative processes embedded within hierarchical structures. Compared with embedded iteration, visual recursion correlated positively with other recursive planning tasks (Tower of Hanoi), but not with specific visuo-spatial resources (spatial short-term memory and working memory). We conclude that humans are able to use recursive representations to process complex visuo-spatial hierarchies and that our visual recursion task taps into specific cognitive resources. This method opens exciting opportunities to explore the relationship between visual recursion and language.

Keywords: recursion, cognition, vision, fractals, representation

1. Introduction

Recursion has fascinated scholars in fields as diverse as mathematics, computer science, linguistics and visual arts for many years, because it allows the generation of structures that are both simple and complex at the same time. Recursive structures are complex because they can contain infinite hierarchical levels, and yet simple because this infinity can be achieved and represented using very simple rules.

Recursion is a term that has been used to characterize the process of embedding a constituent inside another constituent of the same kind (Fitch, 2010; Hulst, 2010; Pinker & Jackendoff, 2005). Recursive processes can generate hierarchical structures that display similar properties across different levels of embedding. This feature, called self-similarity, is a signature of recursive structures. An example of a recursive linguistic structure is the compound noun "[[student] committee]", where we find a noun phrase embedded inside another noun phrase. In contrast, a sentence containing a noun and a verb, such as "[[trees] grow]", is hierarchical, but not recursive, because a constituent of one type (noun) is nested within a constituent of a *different* type (verb).

Recently, the development of the human ability to represent recursion has been considered an important step in the evolution of language (Fitch, Hauser, & Chomsky, 2005; Hauser, Chomsky, & Fitch, 2002), because recursion may have added an unprecedented generative power to language precursors. For example, recursion allows the representation and generation of sentences where a noun phrase $n+1$ is embedded inside a noun phrase n already embedded in a noun phrase $n-1$ ($[NP_{(n-1)}[NP_{(n)}[NP_{(n+1)}]]]$), as occurs in the compound noun "John's sister's house". Besides this increase in generative power, the speed of acquisition of a first language may also be enhanced by the presence of pre-existing (potentially innate) recursive rules in the grammar system (Perfors, Tenenbaum, Gibson, & Regier, 2010).

Despite considerable agreement about the importance of recursion, many different definitions of recursion are in use (Chomsky, 2010; Corballis, 2007; Gentner, Fenn, Margoliash, & Nusbaum, 2006; Hofstadter, 1980; Kilpatrick, 1985; Odifreddi, 1999; Penrose, 1989) which has hindered consistent interpretation of empirical results (Fitch, 2010). It has proven to be particularly difficult to establish clear distinctions between recursion and similar processes such as hierarchical embedding and iteration (Hulst, 2010).

Within the framework adopted here (Fitch, 2010; M. D. Martins, 2012), "iteration" refers to the process of repeating an operation a certain number of times. An iterative process may or may not generate hierarchical structures or create dependency relationships between different elements. For example, putting one marble at a time into a bag is an iterative process, but neither hierarchical nor recursive. In contrast, "hierarchical" structures always involve the embedding of elements within other elements. This embedding can refer to the grouping of constituents within a higher order set, such as the grouping of individuals within a family; or it can refer to the establishment of dominance-subordination relationships between constituents such as in social hierarchies. If the hierarchical embedding occurs between constituents of the same category (e.g. such as a noun phrase inside a noun phrase) we classify it as recursive, otherwise as non-recursive. Iteration, hierarchical embedding and recursion are not mutually exclusive processes: in fact, recursion typically involves both hierarchy and iteration. Nevertheless, it is possible to segregate the cognitive abilities necessary to represent the kind of information that each of these processes encode (Fig.1). Encoding iteration requires the ability to represent the repetition of a certain process, for instance the repeated addition of elements to a structure. Encoding hierarchical embedding requires the ability to represent dependency or grouping relationships between constituents at multiple levels. Encoding recursive embedding requires the ability to represent successive hierarchical levels generated by applying the same transformation rules. Recursion enables the generation of new hierarchical levels beyond those previously experienced, maintaining consistency with existing levels at a higher level of abstraction.

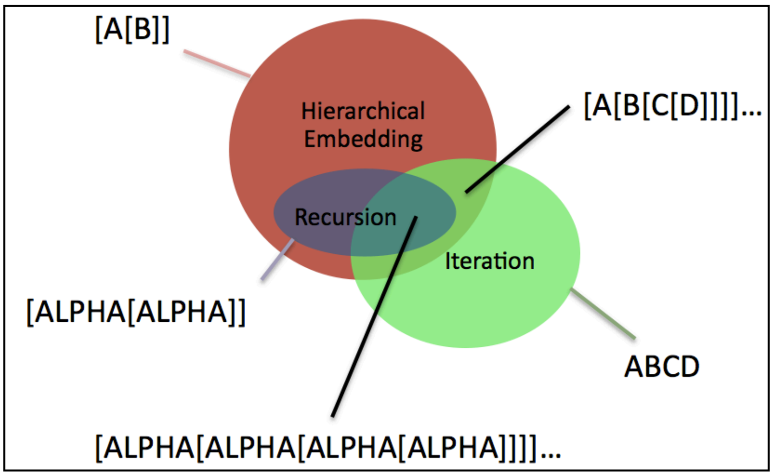


Fig.1. Examples of structures produced by iteration, hierarchical embedding and recursion and by various combinations of these processes. A non-iterated hierarchical embedding corresponds to the establishment of a dependency without repetition. The ability to represent repetition and the ability to represent dependencies may be orthogonal.

As discussed at length elsewhere (M. D. Martins, 2012), we are aware that there are other definitions of recursion, and that recursive processes do not necessarily translate into hierarchical structures. In fact, covert recursion could be involved even to generate simple unstructured sequences. However, only in the generation of multilevelled hierarchies do self-embedding processes translate into self-similar structures, making the recursive properties of both noticeable and obvious. This property makes recursive hierarchies particularly attractive for empirical research.

Recursion has been hypothesized to be part of a "linguistic computational system […], independent of the other systems with which it interacts and interfaces" (Hauser et al., 2002). According to this view, although recursive representations may be available in non-linguistic domains such as visual art (Eglash, 1997), music (Jackendoff & Lerdahl, 2006), architecture (Eglash, 1998), humour (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 linguistic system of abstract arbitrary symbol manipulation and thus be dependent on language (Fitch et al., 2005). Alternatively, some authors have proposed that the usage of recursion in some domains, for example in visual perception, can occur independently of language (Pinker & Jackendoff, 2005).

Despite considerable theoretical debate concerning recursion, very few attempts have been made to evaluate these various hypotheses empirically. A body of research using artificial grammar learning (AGL) methods to test the ability to represent hierarchical structures and long-distance dependencies (Fitch & Hauser, 2004; Gentner et al., 2006) exists, but does not directly address recursion (Fitch & Friederici, 2012). Rather, AGL methods test for memory capabilities beyond the finite state level (Corballis, 2007; Fitch, 2010; Fitch & Friederici, 2012). These experiments suggest that the ability to process hierarchies may be available in non-human species (Abe & Watanabe, 2011; Gentner et al., 2006) (but see (Beckers, Bolhuis, Okanoya, & Berwick, 2012; van Heijningen, de Visser, Zuidema, & ten Cate, 2009), which in humans occurs both in visual and auditory modalities (Bahlmann, Schubotz, Mueller, Koester, & Friederici, 2009).

Regardless of their value for AGL research, these paradigms do not assess the differences between recursive and non-recursive hierarchies. They test neither for the representation of recursive hierarchies nor for the ability to apply them productively in iterative processes. Since there are some qualitative differences between recursive and non-recursive processes, regarding computational efficiency (Ninos & Dollas, 2008; Schiemenz, 2002; Sklyarov, 2004), representational power (Mandelbrot, 1977; Pollack, 2003), speed of acquisition (Perfors et al., 2010) and parsing of complex hierarchies (Koike & Yoshihara, 1993), it is possible that some fundamental properties of human language depend precisely on recursion-specific features, beyond the more general hierarchy processing abilities that might be present in other species.

In the current study, we introduce and explore an experimental paradigm focusing specifically on recursion capabilities in the visual domain using fractal images. This paradigm allows us to distinguish between iterative, hierarchical and recursive processes empirically. Fractals are (typically visual) structures that display self-similarity (Mandelbrot, 1977), that is, they appear similar when viewed at different scales (as in the famous Mandelbrot set). Fractals can be produced by simple rules that generate complex hierarchical structures when applied iteratively to their own output. Due to these self-similar structural properties, new hierarchical levels can be predicted by generalising the production rules and projecting them to further levels. In other words, the ability to predict structural self-similarity requires the ability to represent recursive embedding.

To our knowledge, besides some work in aesthetic preferences (Taylor et al., 2005), no study has been conducted to investigate how people perceive and represent visual fractals. We developed an experimental task based on the properties of fractal geometry, allowing us to assess participants' ability to represent visuo-spatial structures via a set of recursive rules, and to use these representations to make inferences and generalizations in the visual domain.

In experiment 1 we investigated the strategies that participants applied in our visual recursion task (VRT) when feedback was provided. In experiment 2 we investigated whether VRT dissociated cognitively from embedded iteration and general intelligence. In experiment 3 we replicated experiment 2 with a different set of stimuli, with different categories of foils, and without providing response feedback. Finally, in experiment 4, we provided participants with explicit instructions concerning recursion and iteration, and tested the correlation between the application of recursive principles and other cognitive constructs (recursive planning (Tower of Hanoi), visual and spatial working memory, and general intelligence). The potential correlation between visual recursion and the Tower of Hanoi task is particularly important since the latter task has been suggested to involve recursive solutions (Goel & Grafman, 1995).

2. Experiment 1: Response paradigm and aesthetic biases

In experiment 1 we tested the ability of adult humans to make inferences about recursive embedding in the visuo-spatial domain. Since we were interested in exploring how participants would approach visual recursion, we gave minimal instructions and did not restrict response time. We assessed the strategies that participants reported after completing the task, and tested how this subjective measure correlated with objective parameters (percentage of correct responses and reaction times). We also evaluated the effects of the particular response paradigm (binary forced-choice and subjective aesthetic preferences on individuals' performance by (1) adding an additional response task (1-alternative forced-choice), and (2) testing whether an aesthetic preference for self-similar fractals could account for participants' choices, regardless of their ability to represent recursion.

2.1. Method

2.1.1. Participants

We tested 20 volunteers (undergraduates and PhD students; 14 females and 6 males) aged between 20 and 44 (*M* = 28.1, *SD* = 6) recruited at the University of Vienna. All participants were tested using the same experimental apparatus, and all reported normal or corrected-to-normal visual acuity. All participants gave their prior written consent, and were not paid for taking part. The research conformed to institutional guidelines and Austrian national legislation regarding ethics.

2.1.2. Stimuli and procedure

2.1.2.1. Stimulus generation

We based the Visual Recursion Task on the well-established properties of fractal geometry (Mandelbrot, 1977). Visual fractals can be generated from single constituents such as lines, squares or triangles (*the initiators*) by applying a simple transformation rule (*the generator*) a given number of times (iterations). The structures generated by iterating this process are hierarchical and self-similar (see Fig.2 for a schematic overview of such a process).

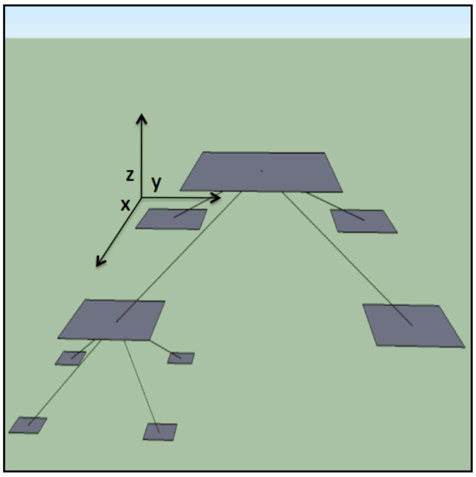


Fig.2. Visual fractals can be conceived as visuo-spatial hierarchies: Different elements (squares) are organized in a two-dimensional space, defined by the xy-axis, and different hierarchical levels are organized vertically along the z-axis. An element with a higher z value is dominant over an element with a lower z value, if the elements are connected.

We produced four successive iterations of 60 different types of fractals, generated using Python code running in Nodebox (version 1.9.5, http://nodebox.net), a visual interface. For each of these 60 fractals, we produced (1) a correct fourth continuation of the first three iterative steps and (2) an incorrect continuation as a Foil. This incorrect fourth iteration was produced by applying a different *generator* to the third stage, and had the same number and size of constituents as the correct fourth iteration.

The fractals produced for this task can be divided into 4 broad categories (see Fig.3 for examples): (1) Polygons (n = 32), (2) trees (n = 9), (3) curves (n = 11) and (4) Koch snowflakes (n = 8). Peano curves and Koch snowflakes were produced using Lindenmayer systems (Lindenmayer, 1968). In these systems, the recursive process substitutes each constituent by a set of new constituents without preserving the *initiator* across iterations. The other two categories of fractals were produced with custom Nodebox scripts.

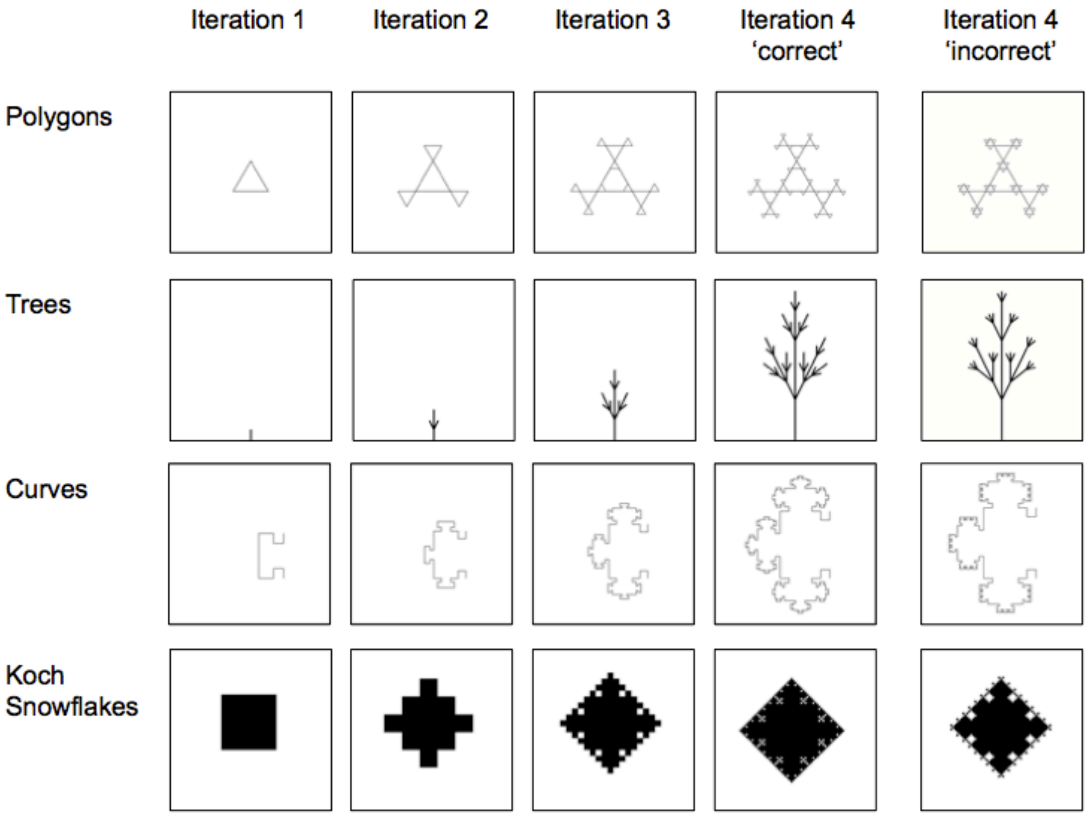


Fig.3. Fractal categories and iterations: For each fractal, we generated the first four iterations and an incorrect fourth iteration. The latter violated the embedding rule used in the previous steps. The fractals were grouped in four classes according to the generating algorithm: Polygons (n=32), trees (n=9), curves (n=11) and Koch snowflakes (n=8).

2.1.2.2. Visual Recursion Task (VRT) 2-choice

The three iterations and two test images were arranged on a panel (Fig.4). Each panel depicted five images, presented simultaneously, arranged in two rows: The first three iterations of each fractal ('sequence' images) were shown in the top row and two alternatives for the fourth iteration (“correct” vs. “incorrect” fourth iteration, henceforth 'choice' images) were shown in the bottom row. The position of the choice images (left or right) was randomized. The sequence of panels was presented on a computer screen (Elo Touchsystems) in a randomized order, which was different for each participant, using custom Python software (version 2.6, www.python.org).

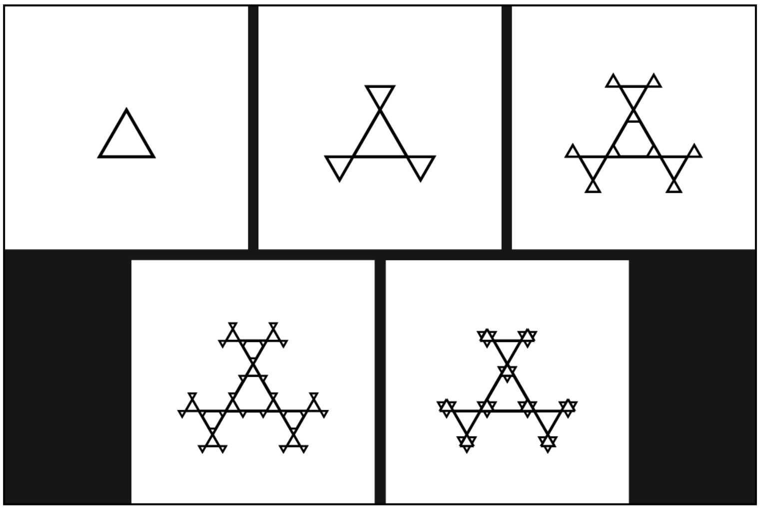


Fig.4. Example of a two-choice stimulus in the Visual Recursion Task (VRT). The first three iterations were presented in the top row. Participants had to choose which of the images in the lower row was correct. In this example, the correct image is on the left.

Participants were instructed in English to select the image they considered correct from the two ‘choice’ images in the bottom row and to “try to understand the right strategy and to choose correctly as often as you can”.

Participants responded by pressing one of two buttons on a button box (ioLab Systems), corresponding to the position of the correct image (left or right). Auditory and visual feedback was given for all trials. After an incorrect choice, the screen turned red for 1.5 s and a negative feedback sound (frequency = 98.0 Hz and 1.5 s duration) was played. After a correct choice, the screen turned white for 1 s and a positive feedback sound (frequency = 348.7 Hz, 1 s duration) was played. The sounds were played through Sennheiser HC 520 headphones. There was a two second inter-trial interval. There was no time limit per trial (timeout) because we did not want to constrain participants' strategies, and because we were interested in knowing how they would naturally approach the tasks when given minimal instructions.

Before the VRT began, participants were given a short training session of five trials. The training stimuli were similar to the VRT panels, except that the sequence of images was generated according to a simple non-hierarchical iterative rule (see Fig.5).

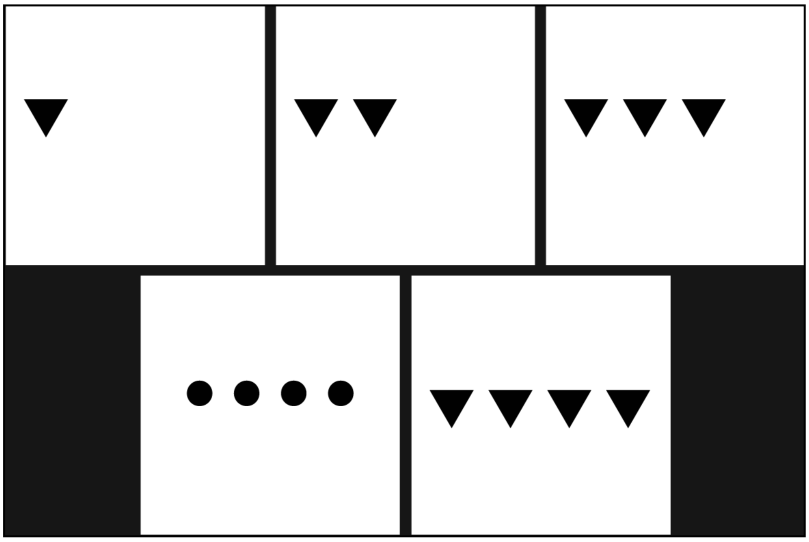


Fig.5. Example of a training stimulus: The iterations followed a number and shape rule but did not produce hierarchical structures. The correct image is on the right.

2.1.2.3. Visual Recursion Task (VRT) 1-choice

In order to evaluate possible performance effects associated with a binary forced choice paradigm, we designed a *VRT 1-choice* task. This task was identical in all aspects to the basic *VRT 2-choice*, except that only one image was presented in the center of the second row of each panel, corresponding to either the correct or incorrect fourth iteration (Fig.6). Participants were instructed to choose whether the image in the lower row was correct (right button) or incorrect (left button). The same number (n = 10) of correct and incorrect fourth iterations were presented.

Before the beginning of the task, the same five training stimuli were presented as in VRT 2-choice, but with only one ‘choice’ image. Feedback and inter-stimuli intervals were the same as in the *VRT 2-choice* task.

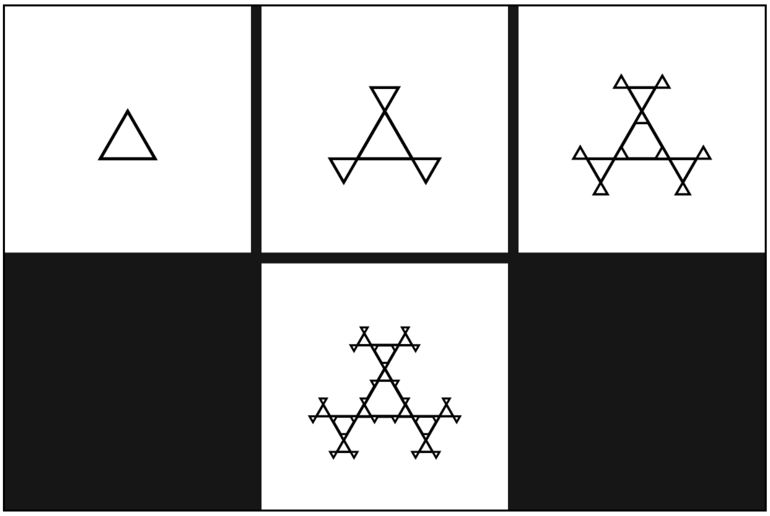


Fig.6. Example of a VRT '1-choice' stimulus. Participants had to decide whether the image presented in the bottom row was correct or incorrect. In this example, the image is correct.

2.1.2.4. Aesthetic preference task

This task was designed to assess the effects of possible preference biases in *VRT 2-choice*. Here, only the 'choice' images ("correct and "incorrect" fourth iteration) were presented on the screen (Fig.7) with no previous 'sequence' images. Participants were asked to simply select the image they preferred. No auditory or visual feedback was given.

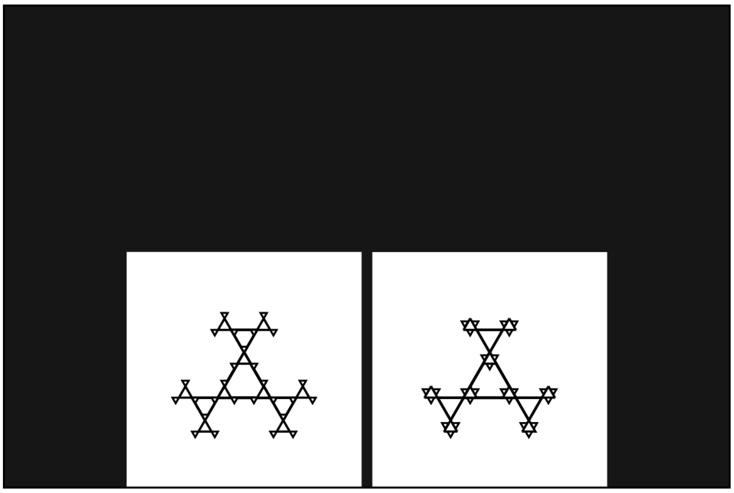


Fig.7. Example of stimuli used in the preference task. Participants were asked to choose the image they preferred.

2.1.2.5. Procedure

All participants began the experiment with the *preference task.* Participants then performed both recursion tasks in one of two possible orders: 10 participants completed *VRT 1-choice* before *VRT 2-choice* ('1-2' condition), and 10 participants performed *VRT 2-choice* before *VRT 1-choice* ('2-1'condition). Participants were randomly assigned to one of the two orders.

The same pool of sixty fractals was used in all tasks, with twenty fractals assigned randomly assigned to each of the three tasks. The distribution of fractal classes was balanced for all tasks and each fractal appeared only once in each experimental session.

Participants' choices and reaction time (RT; in milliseconds) were recorded for all stimuli and for all tasks. The performance was calculated as the percentage of correct answers. In the *preference task*, we recorded as "correct" answers the occurrences where the preferred image corresponded to the well-formed fractal, i.e to the correct fourth iteration. At the end of each task, participants were asked to assess the kind of strategy they had used on a five-point scale. The scale of possible strategies was: 1 - "mostly intuitive"; 2 – "more intuitive than analytical"; 3 – "mixed"; 4 – "more analytical than intuitive"; 5 - "mostly analytical". Intuitive answers were described to the participant as being based on a gut feeling and analytical answers being derived by looking carefully at the details and making explicit inferences.

2.1.3. Analysis

Percentages of correct responses, RT and self-reported strategies were compared between (1) *VRT 2-choice* and *VRT 1-choice* and (2) *VRT 2-choice* and *preference task*. Furthermore, we assessed performance correlations between these tasks. For percentages of correct responses and RT we tested if the data was normally distributed using the Shapiro-Wilk (S-W) test. If variables were continuous and normally distributed we used paired t-tests for comparing means and Pearson bivariate correlations, otherwise we used non-parametric tests (Mann-Whitney U and Spearman correlations).

All statistical analyses were performed using SPSS 19 (IBM).

2.3. Results

2.3.1. Performance, Reaction Time and Strategy

On average, participants scored 84% (*SD* = 12) correct in *VRT 2-choice* and 70% (*SD* = 14) correct in *VRT 1-choice* (Fig. 8). In the *preference task*, the "correct" image was preferred in 58% (*SD* = 11) of the trials. All average task scores differed significantly from chance (Binomial test, $p < .005$, for all 3 tasks). A one-way repeated measures ANOVA showed that the difference in performance between tasks was significant ($F(2, 18) = 41.7$, $p < .001$). Pairwise comparisons with Bonferroni p-value adjustment showed that performance was significantly lower in *VRT 1-choice* than in *VRT 2-choice* (*mean difference* = -15%, $p < .001$); and higher in *VRT 2-choice* than in the *preference task* (*mean difference* =27%, $p < .001$).

Analyzed by participant, the percentage of correct responses in *VRT 2-choice* was correlated with performance in *VRT 1-choice* ($r = .57$; $p = .009$), but not with the *preference task* ($r = .27$, $p = .24$). This correlation between *VRT 1-choice* and *VRT 2-choice* was significant in the group of participants that started the procedure with *VRT 2-choice* ($n = 10$; $r = .797$, $p = .006$), but not in the group that started with *VRT 1-choice* (n =10; $r = .260$, $p = .469$).

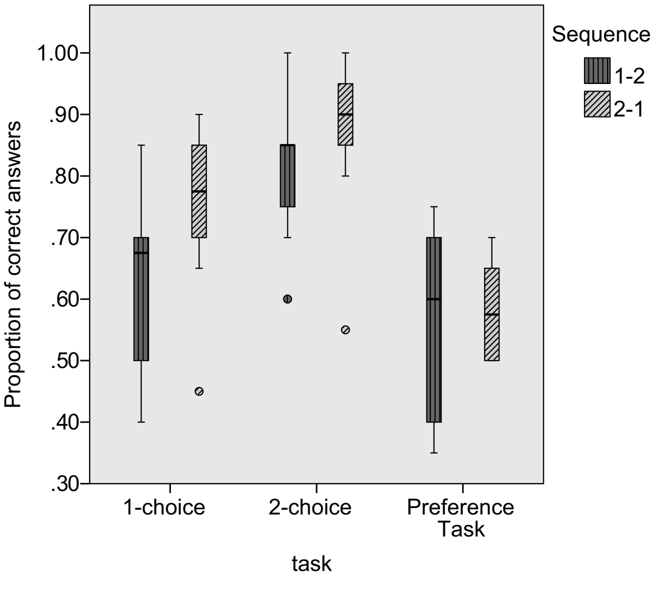


Fig.8. Percentage of correct responses in *VRT 1-choice*, *VRT 2-choice* and *preference task*, in two different task-sequence conditions: '1-2' and '2-1'. Bar charts show median (horizontal line), first quartile (lower edge) and third quartile (upper edge). Dots represent outliers.

On average, RT was 12.5s (*SD* = 1) in *VRT 1-choice*, 12.2s (*SD* = 7) in *VRT 2-choice* and 5.3s (*SD* = 3) in the *preference task* (Fig 9). A significant RT difference was found between *VRT 2-choice* and *preference task* (Wilcoxon signed ranks; $z = -3.81$; $p < .001$) but not between *VRT 2-choice* and *VRT 1-choice* ($p = .6$). We repeated the analysis excluding the extreme outlier participants (i.e. average RTs more than 2 standard deviations beyond the mean) and found the same results.

In general, participants reported a more intuitive strategy for the *preference task* ($M = 2.45$, $SD = .9$) and a more analytical strategy in both *VRT 1-choice* ($M = 4.2$, $SD = .8$) and *VRT 2-choice* ($M = 4.0$, $SD = 1.2$). Interestingly, participants who reported a more analytical strategy in *VRT 2-choice* also had longer reaction times (Spearman's $\rho = .485$, $p = .03$) and higher percentage of correct answers (Spearman's $\rho = .585$, $p = .007$) than those participants who reported intuitive strategies. This suggests that an analytical rather than an intuitive strategy was optimal for the VRT.

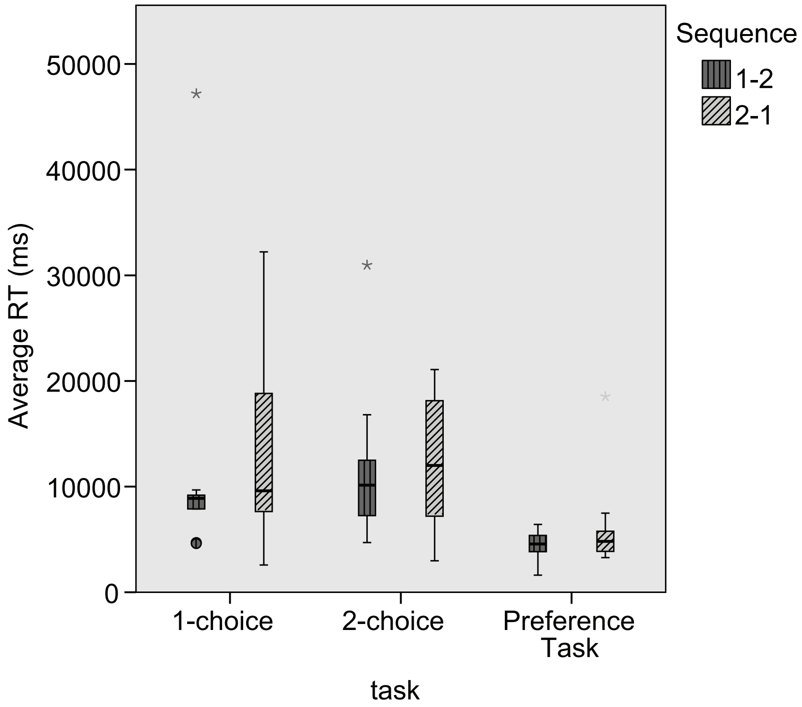


Fig.9. Average reaction time in *VRT 1-choice*, *VRT 2-choice* and *preference task*, in two different task-sequence conditions: '1-2' and '2-1'. Bar charts show median (horizontal line), first quartile (lower edge) and third quartile (upper edge). Dots are outliers.

2.3.3. Stimulus analysis

To test whether participants were using simple visual heuristic strategies to solve VRT, our experiment used four different fractal categories in the task. VRT performance across different stimulus categories is shown in Fig.10. In the task *VRT 2-choice*, the proportion of correct answers was significantly above chance for all stimulus categories (*Binomial Test*, p < .001). In *VRT* 1-*choice*, participants' performance was above chance only for the categories 'polygon' and 'Koch snowflakes' (*Binomial Test*, p < .001).

To analyse whether there were performance differences between different stimuli categories (both between and within tasks), we ran a Generalized Estimating Equations Model (GEE) with 'correctness' (correct/incorrect) as a binomial dependent variable, and 'stimulus category' and 'task' (*VRT 1-choice* vs. *VRT 2-choice*) as within-subject factors. Overall, there was a main effect of stimulus category (*Wald Chi-Square* = 18.7, p < .001). Pairwise comparisons show a significant difference between 'polygons' and 'curves' (*Mean difference* = 18%, p < .001, after Bonferroni correction), but not between the other categories.

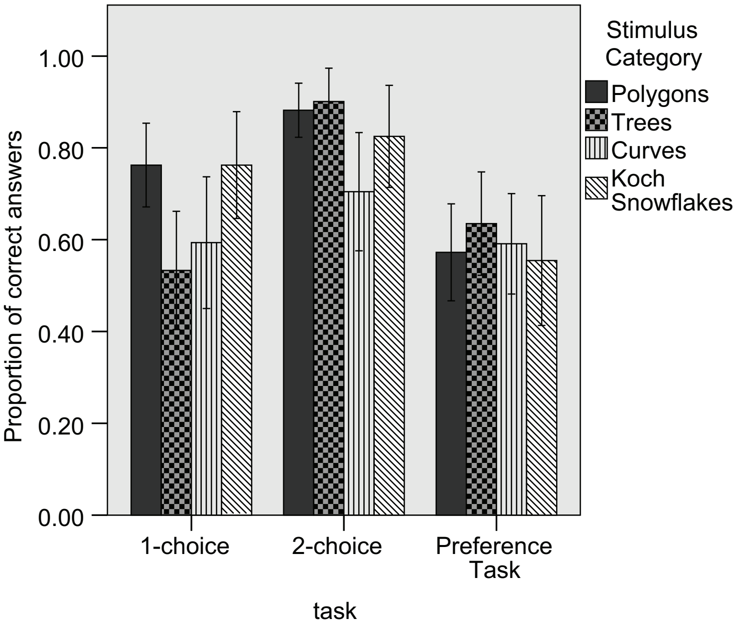


Fig.10. Percentage of correct responses in *VRT 1-choice*, *VRT 2-choice* and *preference task*, across different stimulus categories.

Another important issue was whether the decision between the choice images in the *2-choice* condition was influenced by aesthetic preferences. Given that the same 120 images were part of the pool of possible choices in *VRT 2-choice* and *preference task*, we assessed the frequency with which each image was chosen in both tasks (i.e., for each image, we counted the number of times it was chosen in *VRT 2-choice* and *preference task*). We found that these frequencies were not correlated (r = .027; p = .838), meaning that the images chosen more frequently in *VRT 2-choice* were not the images more frequently chosen in the *preference task*, suggesting that aesthetic preferences can not account for above chance performance in the recursion task.

2.4. Discussion

Our results suggest that human adults can quickly learn how to use recursive information in the visual domain without being explicitly trained or instructed about the concept of recursion. Moreover, a self-reported analytical strategy was associated with higher reaction times, and significantly correlated with better performance. Although response feedback was provided during the task, participants were required to respond to a wide variety of stimuli, with different visual and structural features. Structural recursion was the common element among these stimuli and most likely this abstract regularity was transferred across trials. We propose that the ability to represent structural self-similarity in the visual domain was a necessary condition for good performance in this experiment, regardless of how this information was represented.

Given that *VRT* performance could be influenced by the response paradigm used as well as by aesthetic biases in favour of (or against) self-similar fractals, we included three tasks: two recursive tasks (*VRT 2-choice* and *VRT 1-choice*) and a *preference task*. Our findings rule out an effect of aesthetic preferences on performance in VRT, and demonstrate that both versions of the recursive task were similar to each other: (1) Percentages of correct responses in *VRT 2-choice* and *VRT 1-choice* were correlated. (2) RT and self-reported strategy were similar in these tasks but differed significantly from the *preference task*. (3) Images preferred in the *preference task* were not the images more frequently chosen as 'correct' in the *VRT 2-choice* condition.

However, there was a significant performance difference between *VRT 1-choice* and VRT *2-choice,* depending on task order: Performance in the two tasks only correlated when *VRT 1-choice* was performed after *VRT 2-choice*. It seems that when *VRT 2-choice* was performed first in the presence of correct and incorrect information, participants learned to attend more closely to the relevant image details, thereby increasing their accuracy. This might imply that the ability to process recursion is influenced by the ability to orient attention to the relevant features of the stimuli, and that poor performance in such a task is not necessarily due to an inability to process recursion, but may arise from incorrectly focussed visual attention.

Finally, our task included stimuli from four different categories. These categories instantiated recursion using different procedures, and there was considerable visual variability within each category. As a group, participants scored above chance in all categories in the task VRT *2-choice*. This result supports the idea that participants were able to represent structural self-similarity in this task, as it was the common feature between stimuli. However, there was a significant performance difference between 'polygons' and 'curves'. We think that this difference was more likely the result of different degrees of structural complexity in the stimuli than the result of different cognitive strategies. For instance, the 'curves' category contained some Peano-like fractals, which have a characteristically high Hausdorff dimensionality (Mandelbrot, 1977). In these fractals the multiplication rate of number of segments from one iteration to the next is higher than the 'polygons', making them more complex, and potentially harder to process. Such differences in visual complexity may partially account for differences in performance.

3. Experiment 2: Recursive vs. Non-recursive iteration

Experiment 1 suggested that human adults are able to represent visual recursion successfully. However, it remains an open question whether the *VRT* measures something specific to recursion, or instead taps into a more general ability to extract visual regularities. In experiment 2, we attempted to gain more specific insight into the cognitive processes underlying *VRT*. We devised an Embedded Iteration Task *(EIT)* as a control task, which shared the 'hierarchicality' and iteration features of *VRT*, but lacked recursive embedding. We compared participants' accuracy in both *VRT* and *EIT* with a standardized measure of rule-based visual cognition (Matrix Reasoning from WASI®, see below).

To produce *EIT* images, an iterative process embedded additional elements within a pre-existing hierarchical structure, without producing new hierarchical levels

(Fig.11). To empirically validate the distinction between recursion and iteration we first assessed the behavioural response profile for both tasks. Furthermore, we tested whether different cognitive abilities (fluid intelligence and working memory) predicted accuracy in solving the two tasks. We were also particularly interested in the predictive power of verbal and spatial working memory performance on *VRT* and *EIT* performance.

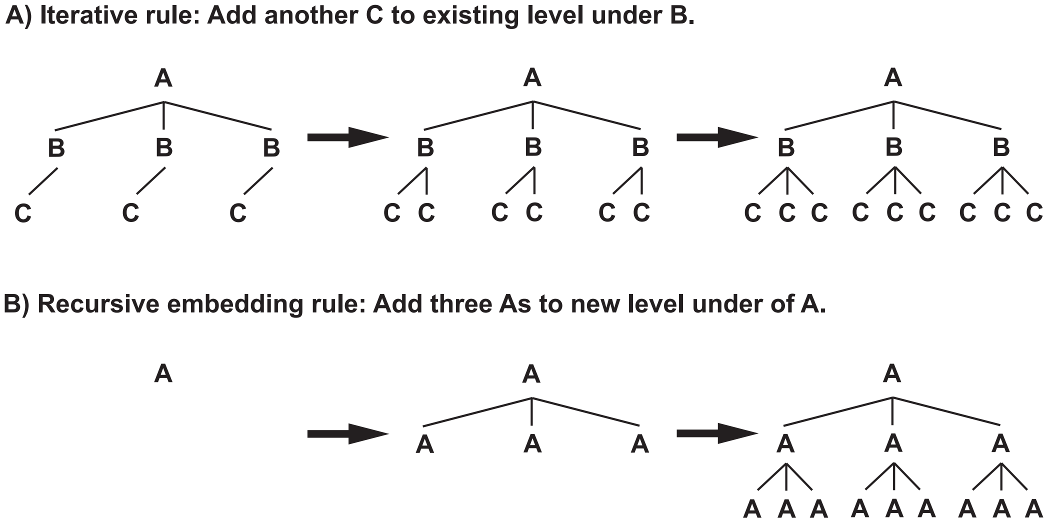


Fig.11. Principles underlying the generation of hierarchies in the Visual Recursion Task (VRT) and Embedded Iteration task (EIT). (A) In EIT we used an iterative rule that adds elements to a previous existing hierarchy, without generating new levels; (B) In VRT we used a recursive rule that adds elements to newly generated hierarchical levels.

3.1. Methods

3.1.1. Participants

We tested 30 volunteers (university undergraduates and employees; 21 females) aged between 18 and 39 (*M* = 23.6, *SD* = 5) recruited at the Lisbon Faculty of Medicine. Education ranged between 11 and 20 years of successfully completed studies (*M* = 15.6, *SD* = 2). All participants were tested in the same room, with the same experimental apparatus as experiment 1, and all reported normal or corrected-to-normal visual acuity. Participants were paid 10 Euros for participating and gave their written informed consent. The research conformed to the appropriate institutional and national legislation regarding ethics.

3.1.2. Stimuli and procedure

3.1.2.1. Visual Recursion Task (VRT)

Stimulus generation and experimental procedure were similar to *VRT 2-choice* described in Experiment 1. In this experiment only 40 test panels were presented to each participant (13 'polygons, 7 'trees', 11 'curves' and 9 'Koch snowflakes').

3.1.2.2. Embedded Iteration Task (EIT)

EIT stimuli were hierarchical structures generated by Python scripts in Nodebox and were very similar to *VRT* fractals. Each *VRT* item was modified to generate a corresponding *EIT* item, with a precise one-to-one correspondence in size, structure and element identity. In *VRT*, each iteration produced a new hierarchical level, while in *EIT* the first image was already a hierarchical structure and each iterative step merely added one additional item within a chosen hierarchical level, without generating a new level (see Fig.12). Crucially, both *VRT* and *EIT* generated hierarchies of the same number of elements and the same number of hierarchical levels.

To control for the use of a simple similarity assessment strategy in *EIT*, we included 10 stimuli ('repetition foils') requiring participants to represent the cumulative addition of constituents (Fig.12). In this subset of stimuli, one of the choice images was a simple repetition of the third iteration; there was no increase in the number of constituents from third to fourth iteration, hence this was the incorrect choice. In the remaining 30 stimuli, we used 'positional foils' in which the possible choices contained the same number of elements that differed in their overall positional scheme (Fig. 12). These 40 panels were intermixed. With these two conditions, we aimed to evaluate whether participants were able to detect both the iterative and positional properties of the hierarchical stimuli.

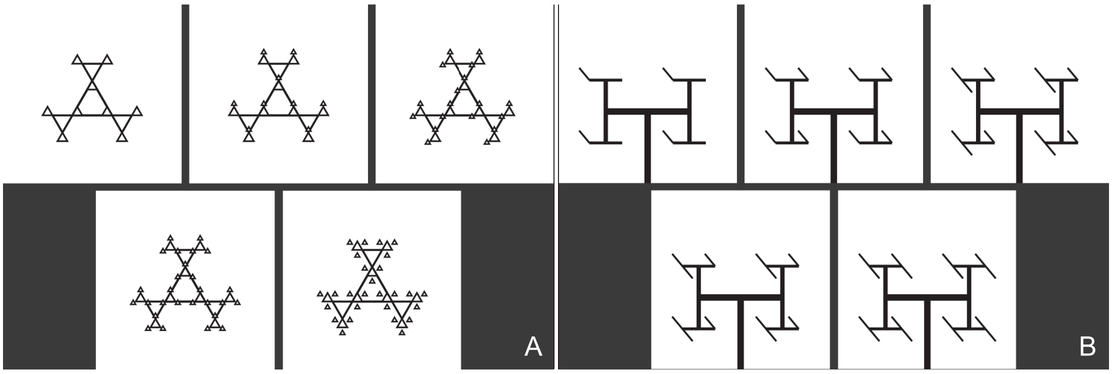


Fig.12. Examples of *EIT* stimuli. (A): "Positional" category (n=30), correct image is on the left. (B): "Number" category (n=10), correct image is on the right.

3.1.2.3. Cognitive assessment

All participants also performed a battery of standardized cognitive tasks. Verbal short-term memory (STM) and working memory (WM) were assessed using *Digit Span* (Richardson, 2007). Spatial short term memory and working memory were assessed with two sub-tests of *CANTABeclipse Spatial Span* (Owen, Morris, Sahakian, Polkey, & Robbin, 1996): (1)'forward' (the number of items successfully repeated in the same order as the example) and (2) 'backwards' (the number of items successfully repeated in the reverse order). Finally, we used un-standardized scores (number of items answered correctly) in two sub-tests of the *WASI* (Wechsler, 1999) test battery - 'vocabulary' and 'matrix reasoning' - as proxies for crystallized and fluid intelligence.

3.1.2.4. Procedure

The procedure took about 90 minutes in total. All instructions were given in Portuguese. *VRT* and *EIT* were randomly assigned either to the beginning or end of the procedure and the cognitive assessment was conducted between the two tasks. Within *VRT* and *EIT*, trial order was differently randomized for each participant. Feedback was provided as in Experiment 1, and there was no timeout limit.

3.1.3. Analysis

Accuracy and RT differences between *VRT* and *EIT* were performed using paired t-tests or non-parametric Wilcoxon tests, as appropriate. We applied the same criteria for normality and statistical decisions as in experiment 1.

We performed correlation analyses to assess whether performance in *VRT* and *EIT* provided non-redundant information relative to standardized measures of intelligence and working memory. Furthermore, to probe for cognitive differences between *VRT* and *EIT*, we performed partial correlations.

All statistical analyses were performed using SPSS 19 (IBM).

3.2. Results

3.2.1. VRT and EIT performance

With two exceptions, all participants scored above chance in both *VRT* and *EIT* (*Binomial test*, proportion = .68, p = .038; Fig.13). Overall, participants performed significantly better in *EIT* (M = 92%, SD = 8) than in *VRT* (M = 84%, SD = 7) (*Wilcoxon signed ranks*, z = -3.75, p < .001). We found no significant difference in accuracy between *EIT* trials with 'Repetition' (M = 87%, SD = 17) and 'Positional' foils (M = 93%, SD = 7) (z = -1.84, p = .066). With one exception, each participant responded adequately both to the positional and repetition stimuli (Fig.13). When we repeated the analysis excluding this outlier, the accuracy difference between 'Repetition' and 'Positional' subtasks remained non-significant (z = -1.61, p = .11).

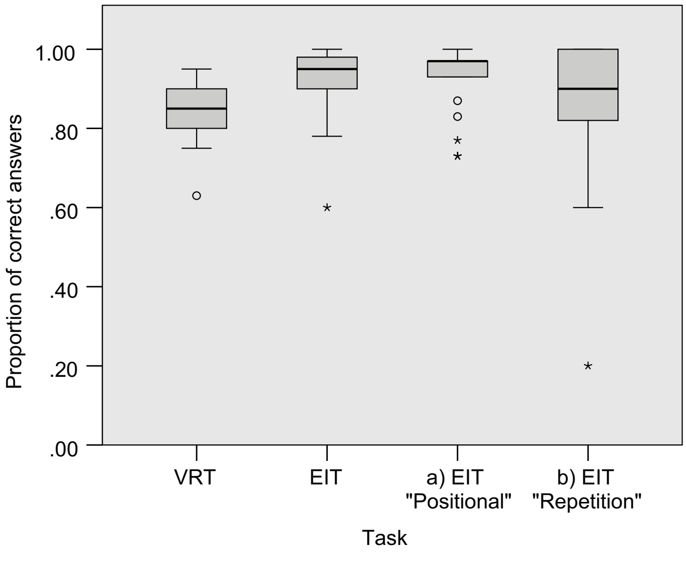


Fig.13. Accuracy across tasks: VRT (visual recursion task), EIT (embedded iteration task), and in the two sub-tasks (a, b) of EIT. The boxplot divides the scores into quartiles, the 'box' represents the distance from the 25th percentile to the 75th percentile and is called the interquartile range. Horizontal 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.

Mean response time (RT) was longer in *VRT* (*M* = 22.2s, *SD* = 12) than in *EIT* (*M* = 18.4s, *SD* = 7) (*Wilcoxon signed ranks*, $z = -2.5$, $p = .012$). Performance on both tasks correlated across participants in accuracy (*Spearman's* $\rho = .365$, $p = .048$) and reaction time (*Spearman's* $\rho = .733$, $p < .001$). Similarly to Experiment 1, participants with longer reaction times performed better in *VRT* (*Spearman's* $\rho = .451$, $p = .012$); this was not true for the EIT (*Spearman's* $\rho = .232$, $p = .217$). The order of the tasks (*VRT* or *EIT* first) did not have a significant main effect on the accuracy of *VRT* and *EIT* (repeated measures ANOVA: $F(1, 28) = .053$, $p = .819$).

3.2.2. Stimulus categories analysis

As in Experiment 1, performance across all four stimulus categories was above chance for both *VRT* and *EIT* (*Binomial test*, $p < .001$ for all categories, see Fig.14). To assess whether there were differences in performance between categories, we ran a GEE model with correctness (correct/incorrect) as the dependent variable, and task (*VRT* vs. *EIT*) and 'stimulus category' as the within-subjects factors. We found a significant interaction between 'task' and 'stimulus category' (*Wald Chi-Square* = 60.1, $p < .001$). Specifically, while there were no significant differences between 'stimulus categories' in EIT (all pairwise comparisons $p > .05$, with Bonferroni correction), in *VRT*, participants performed better with 'polygons' and 'trees' than 'curves' and 'Koch snowflakes' (Pairwise comparisons $p < .001$, with Bonferroni correction).

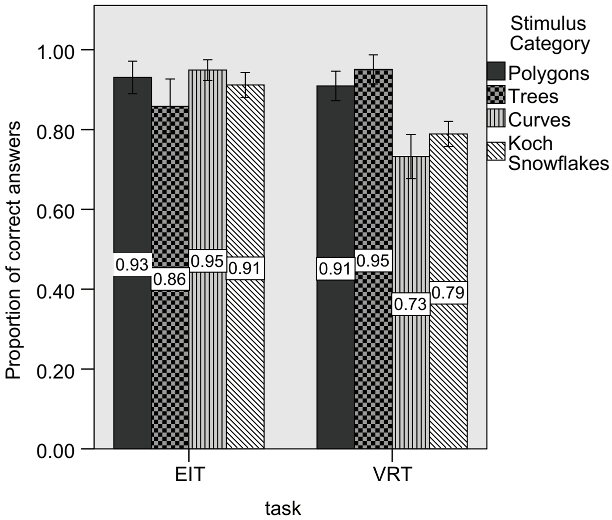


Fig.14. Percentage of correct responses in *VRT* and *EIT* across different stimulus categories.

3.2.3. Correlations with fluid intelligence and working memory

In order to assess whether *VRT* and *EIT* scores were redundant relative to measures of fluid intelligence and crystallized capacity, we compared them to participants' performance in a *Matrix Reasoning* task (*MR*) and in a *Vocabulary* task. Raw results are depicted in Table 1. All variables were normally distributed except for *EIT* scores. We applied a transformation (power of 4) to this variable to achieve normality. Overall Pearson correlations are depicted in Table 2. After p-value correction with the Bonferroni-Holm method (with family-wise error (FWE) level = .05), we found a significant correlation between *EIT* and *MR* ($r (30) = .57$, $p = .008$), and between *VRT* and *MR* ($r (30) = .52$, $p = .02$). One participant had an *MR* score that was two standard deviations below the mean. When this outlier was excluded from the analysis, the correlation between *MR* and *EIT* remained significant ($p = .008$), however, the correlation between *MR* and *VRT* did not ($p = .083$), suggesting that the latter correlation is not a stable (strong) one. The score in the 'vocabulary' task (proxy for crystallized intelligence) was not correlated with *VRT* or *EIT* ($p > .1$).

	N	Minimum	Maximum	Mean	SD
Matrix Reasoning	30	22	35	30.67	3.06
Vocabulary	30	59	79	71.43	5.44
Verbal STM	30	4	9	6.70	1.18
Verbal WM	30	3	8	5.43	1.52
Spatial STM	30	3	9	7.10	1.52
Spatial WM	22	5	9	7.09	1.34

Table 1. Summary of results in the standardized cognitive tasks. STM: Short-term memory, WM: Working Memory, SD: standard deviation.

We also wanted to assess to what extent the capacity for processing verbal and visual information influences *VRT* and *EIT* accuracy. Therefore, we assessed our participants' short-term (STM) and working memory (WM) abilities, in both the visuo-spatial and verbal domains. Due to technical problems there were 8 missing values in Spatial WM. Raw scores are depicted in Table 1 and overall correlations in Table 2.

After p-value correction with the Bonferroni-Holm method (with family-wise error (FWE) level = .05), there were significant correlations between *EIT* and *spatial STM* (*r (30)* = .49, *p* = .04), and a trend correlation with *spatial WM* (*r (22)* = .53, *p* = .055). *VRT* performance did not correlate significantly with performance in either memory task.

	1. VRT	2. EIT	3. VSTM	4. VWM	5. Vocab.	6. MR	7. SSTM
1. VRT							
2. EIT^4	0.493**						
3. Verbal STM (VSTM)	-0.219	-0.132					
4. Verbal WM (VWM)	0.376*	0.368*	0.401*				
5. Vocabulary (Vocab.)	0.107	-0.185	-0.119	0.126			
6. Matrix Reasoning (MR)	0.527**	0.573**	-0.306	0.328	0.121		
7. Spatial STM (SSTM)	0.284	0.492**	0.249*	0.518**	0.208	0.469**	
8. Spatial WM (SWM)	0.350	0.535**	-0.105	0.504*	-0.271	0.670**	0.588**

Table 2. Correlations between standardized cognitive tasks, Visual Recursion Task (VRT) and Embedded Iteration Task (EIT). STM: Short-term memory, WM: Working Memory. **p*<.05. ***p*<.01, for uncorrected p-values.

3.2.4. VRT vs. EIT: Cognitive resources

We performed partial correlation analyses to assess whether different cognitive resources predicted performance in *VRT* and *EIT*. After controlling for the overall variance explained by *VRT*, *EIT* remained significantly correlated with *spatial WM* ($r(19) = .443$, $p = .044$), *spatial STM* ($r(27) = .422$, $p = .023$) and with *Matrix Reasoning* ($r(27) = .423$, $p = .022$). This suggests that *EIT* performance may require the activation of specific visuo-spatial resources to a greater extent than *VRT*. The inverse analysis (correlations with *VRT*, controlling for *EIT*) yielded no significant correlations.

3.3. Discussion

Experiment 2 compared the processing of recursively and iteratively generated items, and sought possible correlations with other standard psychometric measures. We found that performance in *VRT* diverged from non-recursive iterative embedding and from a standardized (visual) fluid intelligence task. These results suggest that performing *VRT* activates specific cognitive resources, and that this task does not simply measure the general ability to perform rule-based visual tasks. Moreover, our results suggest that visual recursion correlates less with visual-specific resources (memory and non-verbal intelligence) than embedded iteration.

Participants' accuracy in *VRT*, but not *EIT*, was significantly correlated with longer reaction times. One explanation for this difference might be that adequate performance in *VRT* depends more on the usage of analytical strategies, perhaps associated with the generation of more abstract representation rules.

With the 'repetition foils', we were able to show that a simple visual heuristic strategy based on visual similarity is not sufficient for solving EIT. Our results demonstrate that most participants understood the iterative rules displayed in the stimuli, and thereby were able to choose the correct continuation of those rules. Crucially, they did so even when the correct continuation of the iterative process (fourth iteration) was not the response choice most similar to the third iteration.

Regarding the correlations with standardized cognitive measures, only a portion of *VRT* and *EIT* variance could be predicted by matrix reasoning and working memory performance. This suggests that our new tasks tap into distinct cognitive abilities. Matrix reasoning seemed to be a mild predictor of *VRT* (28%) and *EIT* (33%). Excluding an outlier participant eliminated the correlation between VRT and matrix reasoning, but not *EIT*.

Spatial STM and WM were better predictors of *EIT* than *VRT,* which suggests that *VRT* performance relies less in the activation of these specific visuo-spatial capacities. It has been suggested that superficial visual processing demands can be reduced by the usage of abstract representations (Alvarez, 2011). Combined with the specific correlation we found between *VRT* (but not *EIT*) performance and high reaction times, this suggests that a more analytical strategy may be used to perform *VRT*, perhaps based on the generation of abstract representations.

Finally, while participants performed above chance in all stimuli categories, there were significant differences between categories. In *VRT*, accuracy in ‘polygons’ and ‘trees’ was higher than in ‘curves’ and ‘Koch snowflakes’. One possible explanation for this difference may be due to the fact that for ‘polygons’ and ‘trees’, the visual information from a certain iteration *n* remains present in the iteration *n+1*. For example, in a ‘tree’ fractal, an iteration *n+1* contains all the branches of the previous iteration *n* plus additional new branches. In a typical ‘curve’ fractal, the whole visual contour is transformed from one iteration to the next, because every segment of the curve is transformed according to the recursive rule. Thus, while the structural ‘core’ is preserved from one iteration to the other in ‘polygons’ and ‘trees’ (analogous to ‘Mother’s bike’ → ‘John’s mother’s bike’), in ‘curves’ and ‘Koch snowflakes’ the ‘core’ constituents of a certain iteration are separated in space in the next iteration (analogous to ‘The driver drinks’ → ‘The driver that the mother loved drinks'). The fact that participants scored above chance in all stimulus categories of a task where all stimuli were intermixed and feedback was provided (facilitating learning from one trial to the next), suggests that differences in performance may be due to the differing visual processing demands of the tasks, rather than differences in the participants' understanding of structural self-similarity *per se.*

4. Experiment 3: Effects of response feedback and stimulus categories

In Experiments 1 and 2 we investigated whether human adults were able to solve a task that required them to form representations of visual recursion. We provided response feedback in both experiments. It could be argued that this training experience, giving response feedback, allowed participants to develop alternative heuristic strategies by trial-and-error, thus avoiding the need to represent hierarchical self-similarity (e.g. participants might base their choice on which image is more similar to the most recent iteration). To test for these effects we assessed performance in *VRT* and *EIT*, in a procedure without response feedback. Furthermore, here we also included repetition foils in *VRT*, in a procedure similar to *EIT* in experiment 2.

4.1. Method

4.1.1. Participants

We recruited 24 volunteers (university undergraduates and employees, 12 females) aged between 19 and 47 (*M* = 26.6, *SD* = 5.5) at the University of Vienna. All participants were tested in the same room, with the same experimental apparatus as Experiment 2, and all reported normal or corrected-to-normal visual acuity. Participants were paid 7 Euros for participating and gave their written informed consent. The research conformed to the appropriate institutional and national legislation regarding ethics.

4.1.2. Stimuli and procedure

4.1.2.1. Visual Recursion Task (VRT)

Stimulus generation and experimental procedure were similar to experiment 2. In this experiment, 40 test panels were presented to each participant. We divided the stimuli into two complexity categories: 'core preservation' stimuli ('polygons' and 'trees', n=20) and 'core transformation' stimuli ('curves' and 'Koch snowflakes', n=20). To test for the use of similarity-based heuristic strategies we included ten *VRT* stimuli with 'repetition' foils (5 core preserving and 5 core transforming), and thirty stimuli with 'positional' foils (15 core preserving and 15 core transforming), see Fig.15 for examples.

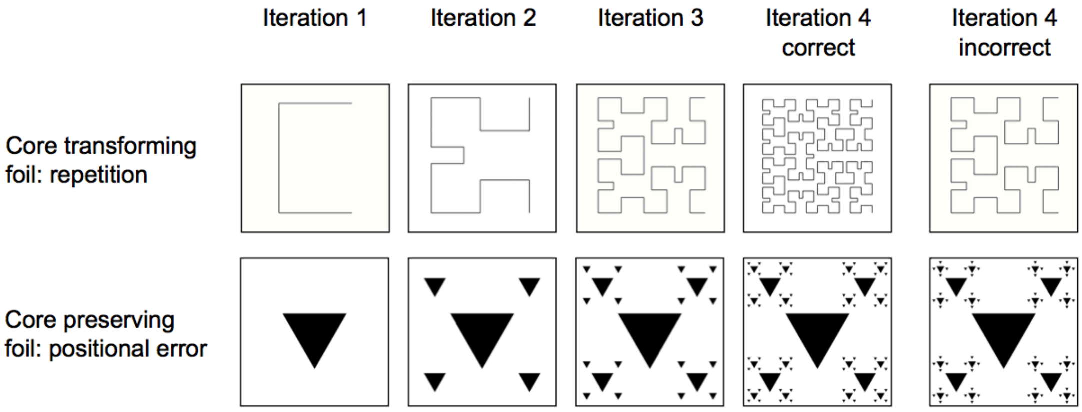


Fig. 15. Examples of fractals used in the Visual Recursion Task: The first four iterations of a fractal generation, as well as one foil ('incorrect' fourth iteration), were produced. There were two categories of rule complexity: core preserving and core transforming; and two categories of foils: 'positional' and 'repetition' (see text for details).

4.1.2.2. Embedded Iteration Task (EIT)

As in *VRT*, there were forty *EIT* stimuli, ten with 'repetition' foils (5 core preserving and 5 core transforming), and thirty stimuli with 'positional' foils (15 core preserving and 15 core transforming), see Fig.16 for examples.

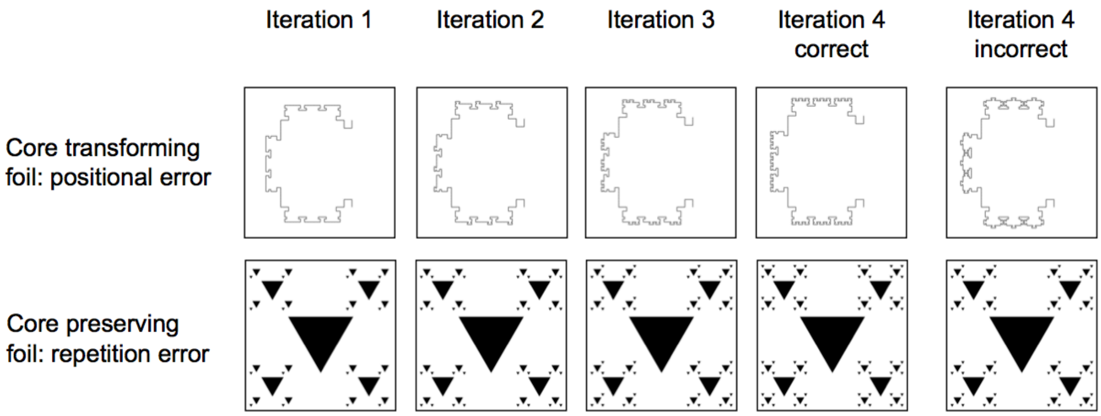


Fig. 16. Examples of fractals used in the Embedded Iteration Task: The first four iterations of a fractal generation using a non-recursive process, as well as one foil (incorrect fourth iteration), were produced. There were two categories of 'visual complexity', matching *VRT* core preserving and core transforming. There were also two categories of foils: 'positional' and 'repetition' (see text for details).

4.1.2.3. Procedure and analysis

Participants were tested in a procedure that took about 60 minutes. The order of *VRT* and *EIT* was balanced across participants. Responses and reaction times were recorded.

General accuracy scores were computed for *VRT* and *EIT*, and specific accuracy scores were computed for the categories 'core preservation', 'core transformation', 'repetition' foils, and 'positional' foils. Principles for statistical analysis were the same as previous experiments.

4.2. Results

With one exception, all participants performed above chance in *VRT* and *EIT* (*Binomial test, proportion* = .68, p = .038), even without feedback. On average, the percentage of correct answers was 86% (*SD* = 1) in *VRT* and 89% in *EIT* (*SD* = 1). This difference was not significant (*paired t-test:* t = -1.4, p = .2). In order to test for our tasks' internal consistency, we performed internal reliability analyses (Cronbach, 1951). Both tasks presented acceptable levels of reliability (*Cronbach's alpha* = .71 for *VRT* and *Cronbach's alpha* = .88 for *EIT*), suggesting they were measuring internally consistent constructs.

Performance for different stimuli categories is depicted in Fig.17. At the group level, performance was above chance for all foil and stimuli categories (*Binomial test*:

$p < .05$). Interestingly, although overall performance in *VRT* was similar to *EIT*, there were differences in the patterns of response. In *EIT*, participants scored significantly better in trials with positional foils (*M* = 92%, *SD* = 1) than in trials with repetition foils (*M* = 78%, *SD* = 2) (*paired t-test:* $t = -3.9$, $p = .001$). The opposite pattern was found for *VRT*, in which participants scored significantly better in trials with repetition foils (*M* = 91%, *SD* = 1) than in trials with positional foils (*M* = 84%, *SD* = 1, *paired t-test:* $t = 3.2$, $p = .003$).

Regarding visual stimulus complexity (core preservation vs. core transformation), we found an effect specific to *VRT*: participants scored lower in core transformation trials (*M* = 82%, *SD* = 1) than in core-preservation trials (*M* = 90%, *SD* = 1, *paired t-test*: $t = 3.7$, $p = .001$). In *EIT*, this difference was not significant (90% vs. 87%, *paired t-test:* $t = 1.0$, $p = .3$).

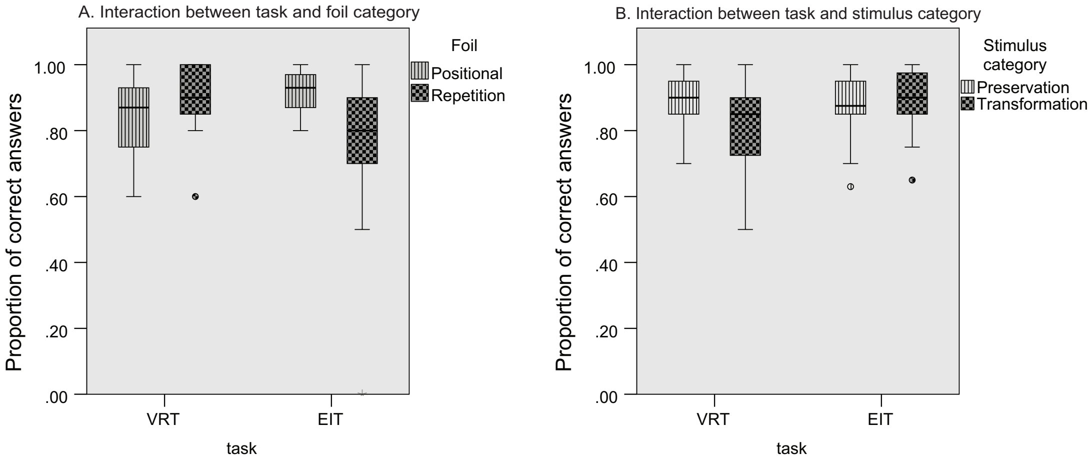


Fig.17. Percentage of correct responses for different stimulus categories. A: performance for repetition and positional foils. Performing adequately in repetition foils means correctly rejecting the image more similar to the third iteration, when this image is not the correct continuation of the iterative process. B: performance for core transforming and core preserving stimuli. In VRT, “transformation” stimuli are generated by a more complex rule than “preservation” stimuli.

4.3. Discussion

In Experiment 3, we replicated the results of Experiment 2 without providing response feedback to the participants. We also investigated whether participants were using simple visual heuristic strategies to solve *VRT*. We found that participants could solve recursion (and iteration) without response feedback, suggesting that they were able to induce abstract principles common across trials (Dewar & Xu, 2010). Participants rejected the repetition foils consistently, that is, the choice image identical to the third iteration, which suggests that they were not simply performing an assessment of similarity between choice images and the first three iterations. In *EIT*, even though performance was above chance in both repetition and positional foils, participants had some difficulty in rejecting the repetition foils. In line with the results of experiment 2, participants seemed to use different cognitive resources to solve *VRT* and *EIT*, even though overall performance was balanced across tasks.

Further evidence for specifically hierarchical processing in *VRT* is suggested by the performance difference found between core preservation and core transformation stimuli. Participants scored lower in stimuli where hierarchical transformations from one iteration (n) to the next ($n+1$) were more complex (See Fig.14, and discussion of Experiment 2). Even though the choice images used in *VRT* and *EIT* were of the same degree of visual complexity, we found no performance differences between core-preservation and core-transformation stimuli in *EIT*. This suggests that these results were not due to the intrinsic complexity of the images, but rather to the complexity of the processes used to generate them (only in *VRT* was there a new hierarchical level, rendering evident the difference in the complexity of processes generating core-preservation and core-transformation hierarchies).

Taken together these results suggest that our participants were sensitive to the processes generating new hierarchical levels, that they were able to learn abstract principles and generalize this information across trials without any feedback or training, and that they were not using simple visual heuristic strategies. These findings provide further support to the hypothesis that human adults can represent recursive principles underlying self-similar hierarchies in the visual domain.

5. Experiment 4: Cognitive correlates of recursive and iterative rules with explicit training

In this experiment we again assessed how performance on recursive and iterative tasks correlated with different cognitive variables, but unlike in Experiment 2, we explicitly instructed our participants about the concepts of recursion and iteration. Participants thus did not have infer these concepts themselves while performing the tasks. This study provided the behavioural data component of an fMRI study, published elsewhere (M. J. Martins et al., 2014). In this experiment we also added the Tower of Hanoi (ToH) task to our test battery. ToH involves hierarchical processing of a sequence of movements and is best solved using a recursive strategy (Goel & Grafman, 1995). A correlation between *VRT* and ToH performance would thus lend support to the hypothesis that the *VRT* taps into cognitive resources associated with recursive processing.

5.1. Method

5.1.1. Participants

We tested 40 volunteers (university undergraduates and employees, 21 females) aged between 20 and 32, who were recruited at the University of Vienna. All participants were tested in the same room with the same experimental apparatus as in experiment 2, and all reported normal or corrected-to-normal visual acuity. Participants were paid 30 Euros for participating[1] and all gave their written informed consent. The research conformed to the appropriate institutional and national legislation regarding ethics.

5.1.2. VRT and EIT

We used shortened versions of the tasks already described in experiment 3. Both *VRT* and *EIT* were composed of 14 items each (7 items each of the two foil categories). We reduced the number of items because participants were explicitly instructed regarding the recursive and iterative rules and thus were expected to need fewer trials to perform adequately. In the instruction phase, participants were shown examples of sequences of images depicting the generation of hierarchies using recursive or iterative processes. In the recursive condition they were told that at each step new elements were added to new hierarchical levels according to a spatial rule that was constant across levels; in the iterative condition they were told that new elements were added within an existing hierarchical level, according to a predictable spatial rule. All items were of the simpler 'core preserving' category. We restricted our test items to this category because in the previous experiments performance was more consistent and rule application was clearer than for items of the ‘core transforming’ category.

---

[1] This experiment was the first part of a longer experimental procedure, including an FMRI session.

5.1.3. Cognitive assessment

We applied a neuropsychological test battery which was composed of computerized versions of digit span backwards (DSPAN, a task of verbal working memory), Corsi block tapping backwards (CORSI, a task of spatial working memory), Tower of Hanoi (a task of recursive planning in action sequencing, computer software retrieved from http://pebl.sf.net/battery.html) (Mueller, 2011)), and a paper-and-pencil version of Raven's progressive matrices (RAVEN, a test of non-verbal intelligence)(Raven, Raven, & Court, 2004). We recorded and analyzed the maximum number of elements correctly reproduced in DSPAN and CORSI, the maximum length (in number of steps) of ToH problems that participants were able to complete without errors, and the number of correct answers in RAVEN.

4.1.4. Procedure and Analysis

Participants were tested in a procedure that took 60 minutes. The order of *VRT* and *EIT* was balanced across participants. The neuro-cognitive test battery was performed after *VRT* and *EIT*. Principles of statistical analysis were the same as for the previous experiments.

5.2. Results

The average percentage of correct answers was 83% in *VRT* (*SD* = 2), and 81% in *EIT* (*SD* = 2). Results in the neuropsychological tests are depicted in table 3 and correlation results are depicted in table 4. The percentage of correct answers in *VRT* was significantly correlated with the maximum length of Tower of Hanoi problems that participants were able to complete without errors (r (36) = .42, p = .011), while the percentage of correct answers in *EIT* was significantly correlated with spatial working memory (r (37) = .43, p = .009). These correlations remained significant after p-value correction with the Bonferroni-Holm method (with family-wise error (FWE) level = .05). There was also a correlation that approached significance between *VRT* and verbal working memory (r (35) = .31, p = .07).

	M	*SD*	*Minimum*	*Maximum*
ToH	5.3	1.6	0	7
Verbal WM	6.9	1.2	4	9
Spatial WM	6.5	1.3	3	9
RAVEN	29.7	2.7	17	32

Table 3. Summary of neuro-psychological pre-testing results. M: Mean, SD: Standard deviation, WM: Working memory, ToH: Tower of Hanoi. RAVEN: Raven's progressive matrices.

	1. VRT	2. EIT	3. ToH	4. VWM	5. SWM
1. VRT					
2. EIT	0.44**				
3. ToH	0.42*	0.20			
4. Verbal WM (VWM)	0.31	0.28	0.303		
5. Spatial WM (SWM)	0.21	0.43**	0.09	0.58**	
6. RAVEN	0.26	0.09	0.13	0.33*	0.24

Table 4. Correlations between VRT (Visual Recursion Task), EIT (Embedded Iteration Task) and other neuro-psychological tasks: WM: Working memory, ToH: Tower of Hanoi, RAVEN, Raven's progressive matrices. **p*<.05. ***p*<.01, for uncorrected p-values.

5.3. Discussion

This experiment replicated the findings of experiment 2 concerning the correlations between multiple cognitive tasks and the application of recursive and iterative rules. Explicit instructions were found to have little effect, either negative or positive. We confirmed that *EIT* is more correlated with specific spatial resources than *VRT*. Furthermore, we showed that *VRT,* but not *EIT,* correlates with Tower of Hanoi (ToH), a hierarchical planning task inviting a recursive solution (Goel & Grafman, 1995). Crucially, we used a measure of ToH that forced participants to plan the complete solution of each problem before starting a trial. This required the representation of a chain of sub-goals embedded within other goals (Anderson & Douglass, 2001), which some have argued to be recursive (Pulvermüller & Fadiga, 2010). Taken together, these results strongly again that our novel visual recursion task taps into cognitive resources associated with recursive representations.

6. General Discussion

Recursion has been hypothesized to be an important ability for the evolution of human language (Fitch et al., 2005; Hauser et al., 2002). However, despite considerable debate surrounding this hypothesis (Corballis, 2007; Fitch, 2010; Fitch et al., 2005; Gentner et al., 2006; Hulst, 2010; Jackendoff & Pinker, 2005; Pinker & Jackendoff, 2005), three crucial questions that have not yet been answered empirically, will ultimately determine our understanding of the role of recursion in the emergence of language: (1) Is recursion a necessary cognitive ability for language? (2) Is recursion only available in the linguistic domain? (3) Is recursion only available to humans? To date, the lack of an empirical method to test for recursion outside the language domain has hampered our ability to address these questions separately. Without a non-linguistic method to test for recursion, investigating other species' recursive abilities is impossible. Likewise, it would not be possible to investigate whether humans use recursion in non-linguistic domains, independently of linguistic recursion.

To begin to resolve these issues, we have developed a new method – the Visual Recursion Task (*VRT*) - testing whether individuals can learn and apply recursive rules in the visual domain. Because our task does not necessarily require linguistic instructions or responses, it is well suited for non-linguistic populations (e.g. young children, aphasia patients, and non-human animals), and for experimental designs in which linguistic resources cannot be used or are specifically blocked (e.g. verbal interference paradigms).

We conducted four experiments to characterize the cognitive resources associated with our visual recursion task. In general, participants had little problem understanding or executing the task.

In the first experiment we showed that human adults can represent and use recursion in the visuo-spatial domain. The results support the hypothesis initially put forth (without empirical evidence) that recursion is not restricted to the linguistic domain (Pinker & Jackendoff, 2005). In our Experiment 1, the ability to represent visual recursion seemed to require analytical strategies, and was not influenced by aesthetic biases towards well-formed fractals. Crucially, our participants were able to perform adequately in different trials with very distinct visual patterns, both in a two-alternative forced-choice task and in a single-choice paradigm.

In the second experiment, we tested whether the cognitive resources used in visual recursion were somehow distinct from visual iteration and general intelligence. The results suggested that performance in *VRT* was more strongly associated with the use of analytical strategies and correlated less with visuo-spatial memory and general intelligence than *EIT*.

In experiment 3, we replicated the findings of the first two experiments without providing response feedback or training. We also used a more homogeneous set of stimuli to achieve good internal reliability. We found that participants did not use simple similarity assessment strategies to solve *VRT* or *EIT*, and that they were able to generalize information across trials, without response feedback, suggesting that they were inducing and applying abstract rules (Dewar & Xu, 2010). Furthermore, our results confirm that even when accuracy was similar, *VRT* and *EIT* showed very different response profiles: 1) Iterative and recursive representations were associated with better performance in different kinds of foil categories; and 2) Participants seemed to be sensitive to the complexity of the processes used to generate new hierarchical levels in *VRT*, which confirms the assumption that in this task participants were able to encode cross-level hierarchical information.

Finally, in experiment 4 we explicitly instructed our participants on the concept of recursion and iteration prior to the procedure, and assessed the cognitive correlates of the application of recursive and iterative rules. We found that the application of recursive rules in the visual domain correlated with performance in another potentially recursive task (Tower of Hanoi), and confirmed that *EIT* correlates more strongly with visuo-spatial memory resources than *VRT*.

All of these results are clearly consistent with the suggestion that our novel visual task measures a cognitive construct associated with recursive cognition, and show that human adults are easliy able to encode information regarding hierarchical self-similarity.

These results also provide some insights into the cognitive nature of recursive visual representations. In comparison to *EIT*, performance in *VRT* seems to be better predicted by tasks requiring prospective thinking (e.g. Tower of Hanoi), and less associated with specific spatial working memory tasks. The nature of this dissociation is consistent with the proposal that recursive representations involve more abstract and parsimonious rules than non-recursive representations (Helm, van Lier, & Leeuwenberg, 1992). The ability to generate compressed and more abstract representations of hierarchical structures may thus decrease the processing demands of visuo-spatial resources (Alvarez, 2011; Brady & Alvarez, 2011). The greater the regularity of a visual structure, the better people are in building abstract representations of it (Brady & Alvarez, 2011). This process of abstraction could then decrease the need to store item-based representations, reducing the storage and processing load upon visual working memory.

The four studies presented here clearly show that a cognitive capacity for recursion is not limited to language, but is also available in the visual domain. Our new task opens new methodological and conceptual paths to empirical investigations into the nature of recursive representations. We predict that extending this research to include language-impaired populations, verbal-interference paradigms, participants at different developmental stages or cultures, and to non-human animals will provide rich and varied experimental evidence that can help to resolve ongoing debates concerning the role of recursion in the evolution of human language.

Acknowledgments

We are grateful to Paolo Baldi for the insightful discussions, and to Gesche Westphal-Fitch, Drasko Boko, Bruno Gingras, Andrea Ravignani, Nina Stobbe, and Piera Filippi for useful comments. This work was supported by the FCT grant SFRH/BD/64206/2009 to MM and by ERC Advanced Grant SOMACCA, project number 230604 to WTF.

References

Abe, K., & Watanabe, D. (2011). Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nature Neuroscience, 14*(8), 1067-1074.

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences, 15*(3), 122-131. doi: 10.1016/j.tics.2011.01.003

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. doi: 10.1037/0278-7393.27.6.1331

Bahlmann, J., Schubotz, R. I., Mueller, J. L., Koester, D., & Friederici, A. D. (2009). Neural circuits of hierarchical visuo-spatial sequence processing. *Brain Research, 1298*, 161-170. doi: 10.1016/j.brainres.2009.08.017

Beckers, G. l. J. L., Bolhuis, J. J., Okanoya, K., & Berwick, R. C. (2012). Birdsong neurolinguistics: Songbird context-free grammar claim is premature. *NeuroReport, 23*(3), 139-145 110.1097/WNR.1090b1013e32834f31765.

Brady, T. F., & Alvarez, G. A. (2011). Hierarchical encoding in visual working memory: Ensemble statistics bias memory for individual items. *Psychological Science, 22*(3), 384-392. doi: 10.1177/0956797610397956

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.

Corballis, M. C. (2007). Recursion, language, and starlings. *Cognitive Science, 31*(4), 697-704. doi: 10.1080/15326900701399947

Cronbach, L. J. (1951). Coeficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-333.

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.

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.

Fitch, W. T., & Friederici, A. D. (2012). Artificial grammar learning meets formal language theory: An overview. *Philosophical Transactions of the Royal Society B, 367*, 1933-1955.

Fitch, W. T., Hauser, M., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition, 97*(2), 179-210. doi: 10.1016/j.cognition.2005.02.005

Fitch, W. T., & Hauser, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science, 303*, 377-280.

Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature, 440*(7088), 1204-1207. doi: 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.

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.

Helm, P. v. d., van Lier, R. J., & Leeuwenberg, E. J. (1992). Serial pattern complexity: Irregularity and hierarchy. *Perception, 21*, 517-544.

Hofstadter, D. R. (1980). *Gödel, escher, bach: An eternal golden braid*: Harmondsworth: Penguin.

Hulst, H. v. d. (2010). Re recursion. In H. v. d. Hulst (Ed.), *Recursion and human language*: de Gruyter Mouton.

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. doi: doi: 10.1016/j.cognition.2005.04.006

Kilpatrick, J. (1985). Reflection and recursion. *Educational Studies in Mathematics, 16*(1), 1-26.

Koike, H., & Yoshihara, H. (1993). *Fractal approaches for visualization of huge hiearchies.* Paper presented at the Proceedings of the 1993 IEEE Symposium on Visual Languages.

Lindenmayer, A. (1968). Mathematical models for cellular interaction in development i. Filaments with one-sided inputs. *Journal of Theoretical Biology, 18*, 280-289.

Mandelbrot, B. (1977). *The fractal geometry of nature*: Freeman.

Martins, M. D. (2012). Specific signatures of recursion. *Philosophical Transactions of the Royal Society B, 367*, 2055-2064.

Martins, M. J., Fischmeister, F. P., Puig-Waldmueller, E., Oh, J., Geißler, A., Robinson, S., . . . Beisteiner, R. (2014). Fractal image perception provides novel insights into hierarchical cognition. *NeuroImage*. doi: 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.

Mueller, S. T. (2011). *The psychology experiment building language*.

Ninos, S., & Dollas, A. (2008, 2008). *Modeling recursion data structures for fpga-based implementation.* Paper presented at the Field Programmable Logic and Applications, 2008. FPL 2008. International Conference on Heidelberg.

Odifreddi, P. G. (1999). *Classical recursion theory*: Elsevier.

Owen, A. M., Morris, R. G., Sahakian, B. J., Polkey, C. E., & Robbin, T. W. (1996). Double dissociations of memory and executive functions in working memory tasks following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampectomy in man. *Brain 119*, 1597-1615.

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

Perfors, A., Tenenbaum, J., Gibson, E., & Regier, T. (2010). How recursive is language? A bayesian exploration. In H. v. d. Hulst (Ed.), *Recursion and human language*: de Gruyter Mouton.

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

Pollack, J. B. (2003). Recursive distributed representations. *Artificial Intelligence, 46*(1-2), 77-105.

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

Raven, J., Raven, J. C., & Court, J. H. (2004). *Manual for raven's progressive matrices and vocabulary scales.* San Antonio: Harcourt Assessment.

Richardson, J. T. (2007). Measures of short-term memory: A historical review. *Cortex, 43*(5), 635-665.

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

Sklyarov, V. (2004). Fpga-based implementation of recursive algorithms. *Microprocessors and Microsystems, 28*(5-6), 197-211. doi: 10.1016/j.micpro.2004.03.008

Taylor, R. P., Spehar, B., Wise, J. A., Clifford, C. W., Newell, B. R., Hagerhall, C. M., . . . Martin, T. P. (2005). Perceptual and physiological responses to the visual complexity of fractal patterns. *Nonlinear Dynamics, Psychology and Life Scences, 9*(1), 89-114.

van Heijningen, C. A. A., de Visser, J., Zuidema, W., & ten Cate, C. (2009). Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. *Proceedings of the National Academy of Sciences, 106*(48), 20538-20543. doi: 10.1073/pnas.0908113106

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence*. New York.

5. Processing visual recursion does not require verbal and motor resources

Martins M, Mursic Z, Fitch WT (under review in Cognitive Psychology). Visual recursion is independent from verbal and motor domains of cognition.

Processing visual recursion does not require verbal or motor resources

Maurício de Jesus Dias Martins[1,2]
Zarja Muršič[1]
Jinook Oh[1]
W. Tecumseh Fitch[1]

[1]Department of Cognitive Biology, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria

[2]Language Research Laboratory, Lisbon Faculty of Medicine, Av. Professor Egas Moniz, 1649-028 Lisbon, Portugal

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 processing 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 processed without access to verbal resources suggests that recursion is available independently of language processing abilities.

Keywords: Hierarchy, Recursion, Self-embedding, Fractals, Language

1. Introduction

The ability to represent and generate complex hierarchical structures is one of the hallmarks of human cognition. In many domains, including language, music, problem-solving, action-sequencing, and spatial navigation, humans organize basic elements into higher-order groupings and structures (Badre, Hoffman, Cooney, &

D'Esposito, 2009; Chomsky, 1957; Hauser, Chomsky, & Fitch, 2002; Nardini, Jones, Bedford, & Braddick, 2008; Unterrainer & Owen, 2006; Wohlschlager, Gattis, & Bekkering, 2003). The human ability to simultaneously represent basic elements and higher-order structures results in a high degree of behavioral flexibility. For example, in action sequencing, humans are able to change, add, or adapt certain basic movements to particular contexts, while keeping the overall structure (and goals) of canonical motor procedures intact. For instance, while preparing coffee with a coffee machine we can interrupt the normal sequence of events when the water container is empty, and refill it, without losing track of the main goal and procedure. Non-human animals are much less flexible in performing sequences of actions (Conway & Christiansen, 2001)

In the research described here, we understand recursion as one particular procedure for generating hierarchies. According to our definition, this procedure requires the ability to represent hierarchical self-similarity, and to use this knowledge to make inferences about new or previously absent hierarchical levels (Fitch, 2010; Martins, 2012)[2]. Defined as such, recursion allows the generation of potentially infinite hierarchical levels using a finite set of rules. The unbounded generative power of recursion is distinct from other hierarchy generating processes. For instance, we can add several new elements to an existing hierarchical level using iteration (Fig. 1A), and we can generate a single hierarchical level using a non-recursive embedding rule (Fig. 1B); but it is only possible to generate multiple hierarchical levels with a single rule, if this rule is recursive (Fig.1C), i.e., if elements of same category occur in both sides of a transformation rule.

---

[2] Many definitions of recursion have been offered (Arsenijević & Wolfram, 2010; Chomsky, 2010; Fitch, 2010; Friederici, Bahlmann, Friedrich, & Makuuchi, 2011; Hulst, 2010; Karlsson, 2010; Levinson, 2013; Lobina, 2011; Luuk & Luuk, 2010; Martins, 2012; Odifreddi, 1999). For theoretical and practical reasons that we reviewed elsewhere (Martins, 2012), we focus here on the ability to represent hierarchical self-similarity. Although hierarchical output is not a strict requirement of recursive processes (e.g. natural numbers are not hierarchical), self-embedded hierarchies are useful because they explicitly convey structural recursion. This explicitness facilitates the detection and use of recursive principles, independently of how they are implemented algorithmically.

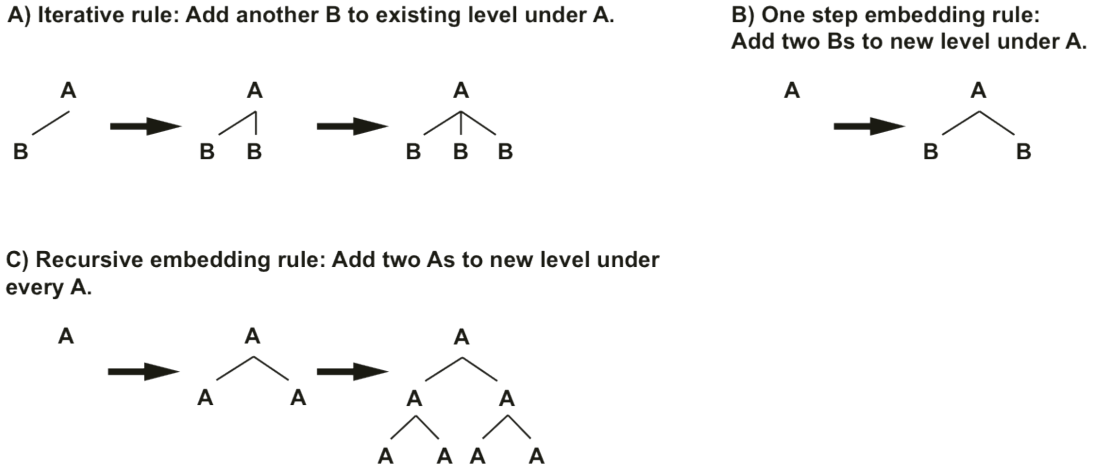


Fig.1. Examples of processes that add elements to hierarchies. These processes can either generate new levels (b and c), or simply add elements to pre-existing levels (a). Recursion (c) can generate multiple hierarchical levels using the same single rule.

An example of a recursive visuo-spatial rule is shown in Fig.2: At each step, the spatial location of groups of 3 added small hexagons depends on the location of an arbitrary number of pre-existing larger hexagons. If individuals are able to perceive that this rule stays constant across multiple hierarchical levels, then they should be able to extrapolate to further levels following the same principle (Martins, 2012). The self-similar properties of the recursive process means that the analysis of information present in existing stimuli is sufficient to extract rules governing the generation of hierarchies, and subsequently allows the application of these rules in the generation of new hierarchical levels. Self-similarity and open-endedness make recursive rules useful for building compressed and memory-efficient representations of complex hierarchies (Koike & Yoshihara, 1993), and allow humans to parse information deeply nested within hierarchies (Martins, unpublished data).

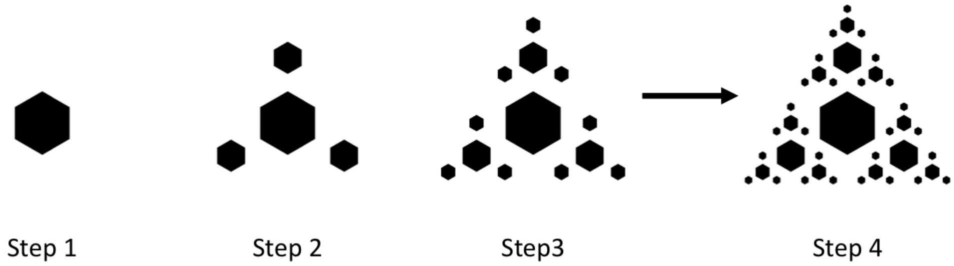


Fig.2. Example of a recursive rule governing the generation of a visuo-spatial hierarchy. The exposure to the first three steps allows the extraction of the recursive rule, which can then be applied to generate the fourth step.

1.1. Recursion and 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, architecture, etc, 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 recursive 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, Hauser, & Chomsky, 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' (or by a recursive operation). The representation of recursion in other domains depends on language, and therefore recruits language-specific resources.

Hypothesis 2: The ability to form recursive representations is domain-general. There is a single 'recursion module' (or a single recursive operation) 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 domain-specific recursion operation, independent from the other domains. This implies that there may be multiple independent processes that can be used to implement recursion.

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).

1.2. 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 & Fitch, 2012). This research suggested that the ability to acquire recursive rules in the visuo-spatial domain may 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), and more strongly with the processing component of verbal working memory (Martins & Fitch, 2012). However, this 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.

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 verbal (i.e., language-specific) resources are necessary for the use of recursion in other non-linguistic domains. 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.

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.

2.1.1. Ethics Statement. The experiment reported in this article was conducted in accordance with Austrian law and the policies of the University of Vienna. According to the Austrian Universities Act 2002, the appointment of ethics committees is required only for medical universities engaged in clinical tests, the application of new medical methods, and/or applied medical research on human subjects. Accordingly, ethical approval was not required for the present study. Nevertheless, all participants gave written informed consent and were aware that they could withdraw from the experiment at any time without further consequences. All data was stored anonymously.

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 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 fifty minutes. 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 seconds to provide an answer. After the primary response was provided, or after 10 seconds, 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.3.

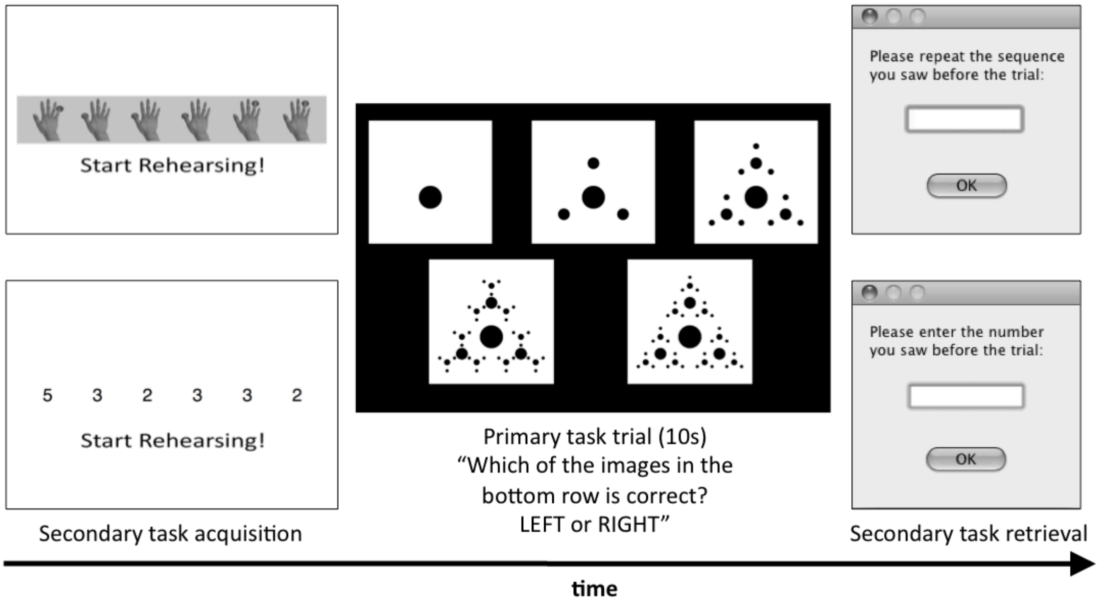


Fig.3. Overview of the trial structure. In the beginning of the trial, participants were shown the secondary task memory content ('Secondary task acquisition'). Participants rehearsed a finger tapping motor sequence in the motor condition and a digit sequence in the verbal condition. Participants pressed 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 seconds), participants had to type the motor or verbal sequences rehearsed throughout the trial ('Secondary task retrieval').

The experimental apparatus is schematically depicted in Fig.4. 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.

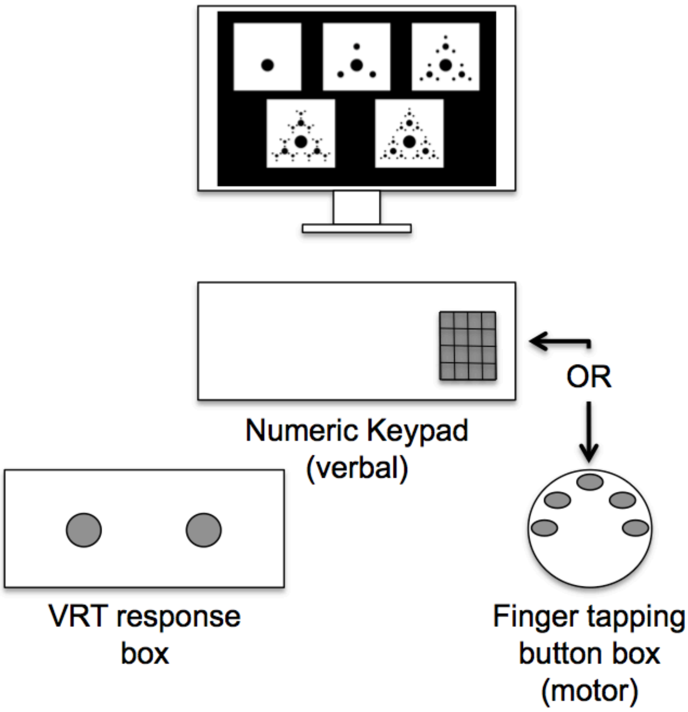


Fig.4. Experimental setup. Participants sat in front of the screen. With their left hand they provided VRT responses by pressing one of the buttons on the VRT response box. With their right hand they gave responses to the secondary tasks. In the verbal interference condition they typed digits in a numeric 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.5). 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.5). 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.5): 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 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.

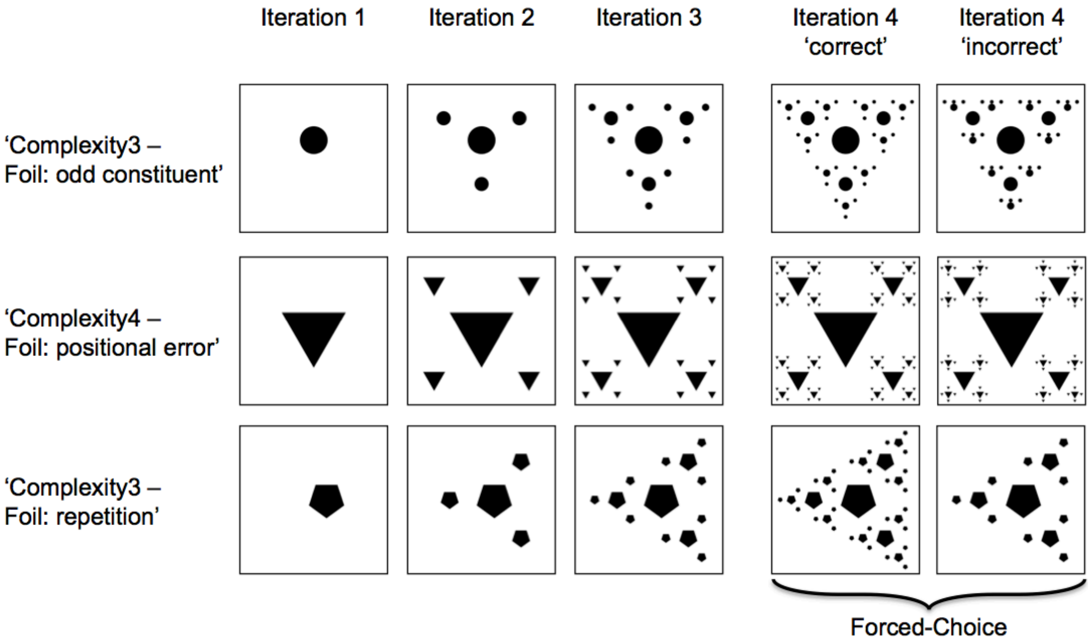


Fig.5. Examples of fractals used in the Visual Recursion Task: The first four iterations of a fractal generation, as well as one foil ('incorrect' fourth iteration), were produced. There were two categories of 'visual complexity' (using either 3 or 4 elements in a set) - and different categories of foils: 'Odd constituent', 'Positional error' and 'Repetition'.

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.6, 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.6, 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.

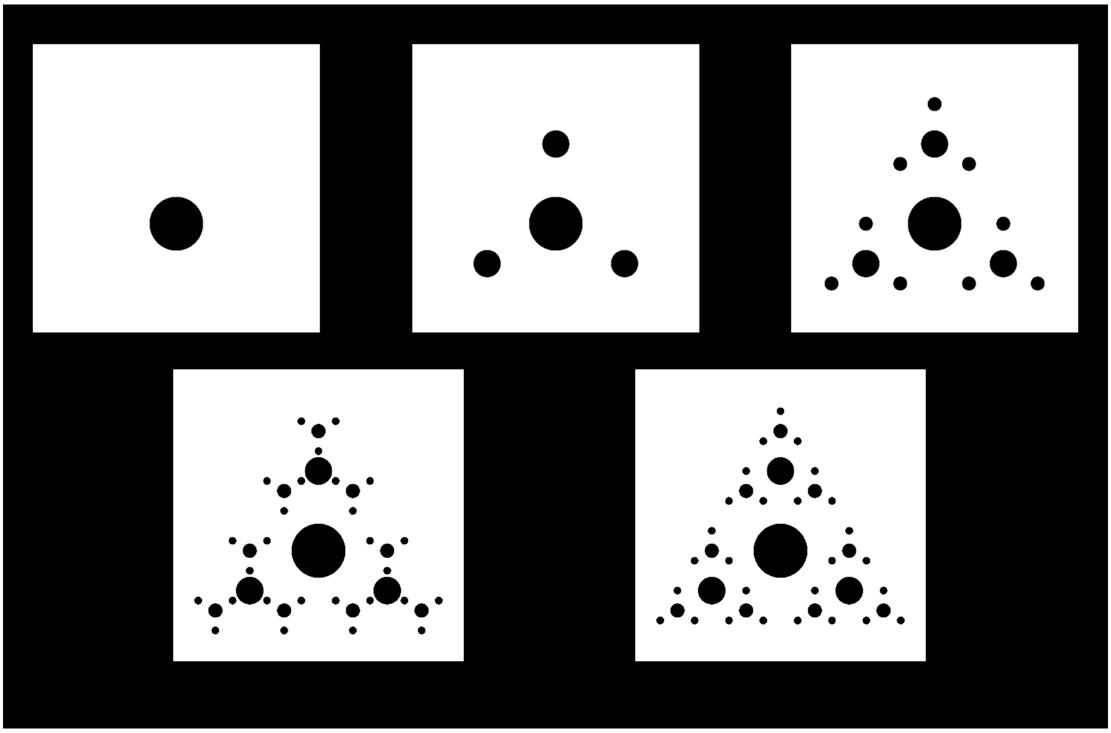


Fig.6. A Visual Recursion Task trial: Initially, the first three iterations of a fractal generation were depicted sequentially from left to right (top). Then, two images were presented simultaneously on the bottom half of the screen, corresponding to the ‘correct’ fourth iteration (bottom right) and a foil (bottom left). From these images, participants had to choose which corresponded to the correct fourth iteration (LEFT or RIGHT).

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.3, 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. 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 13s, 24s and 38s 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 seconds 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.3 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 beginning the experimental procedure to familiarize them 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.7 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 seconds. Then, they were asked 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).

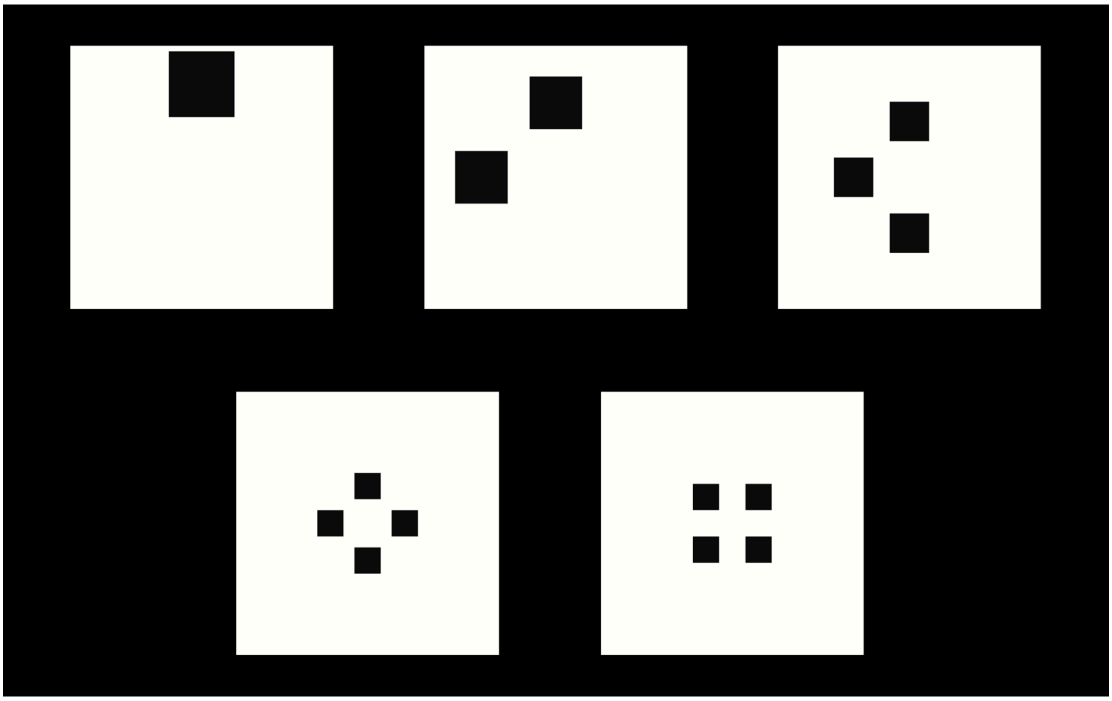


Fig.7. Example of an item included in the training task for VRT. Images follow an iterative, not recursive, rule. The correct choice is on the bottom left.

2.5. 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 x 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 x 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.

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. 8. At the group level, participants scored above chance in all sessions (Binomial test: all *p-values* < 0.001), see Fig. 8. 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).

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.8), 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).

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* $\chi 2 = 3.6$, $p = 0.3$).

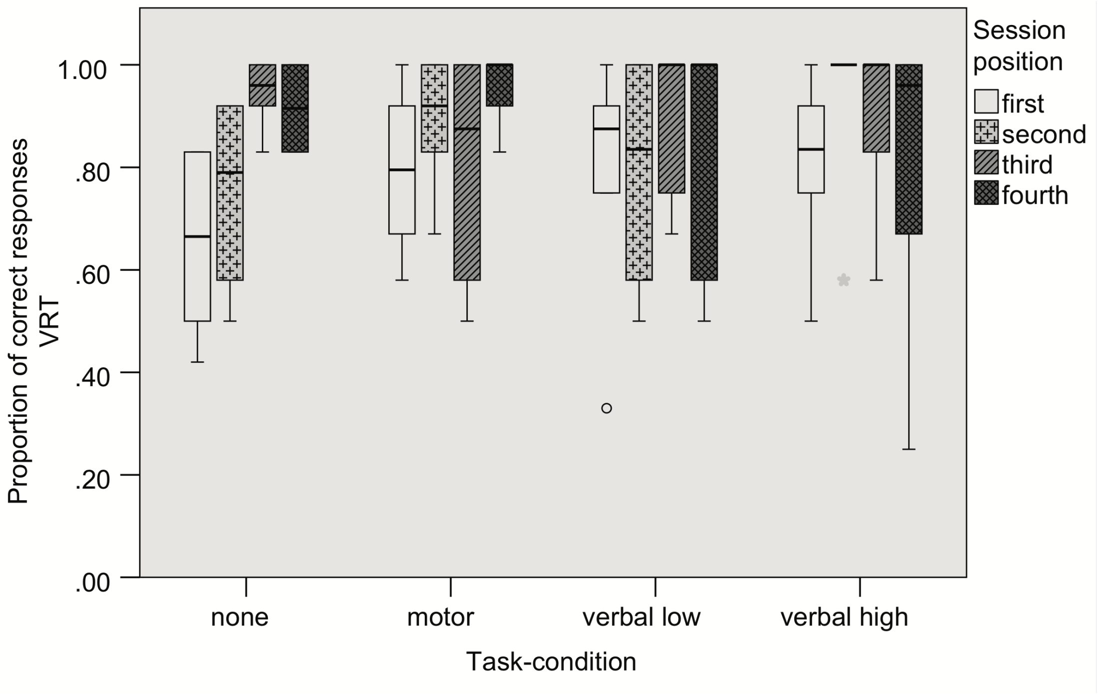


Fig.8. Visual recursion task (VRT) performance across task-conditions and sessions. The presence of motor and verbal interference tasks did not impair VRT performance. However, participants scored lower in the first session than in the last three sessions. The boxplot divides the scores into quartiles, 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.9. 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.9.

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 $p > 0.05$).

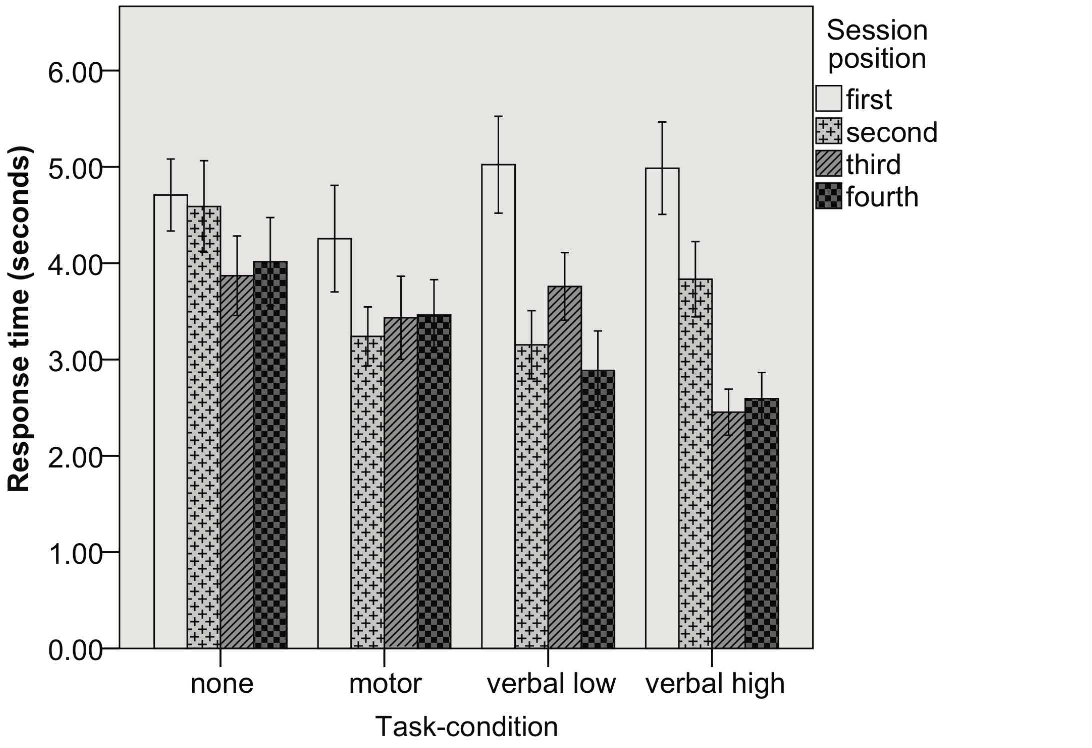


Fig.9. Response times across task-conditions and across sessions. Overall, motor and verbal interference tasks decreased VRT response times significantly and participants had longer response times in the first of the four sessions.

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.10), 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.

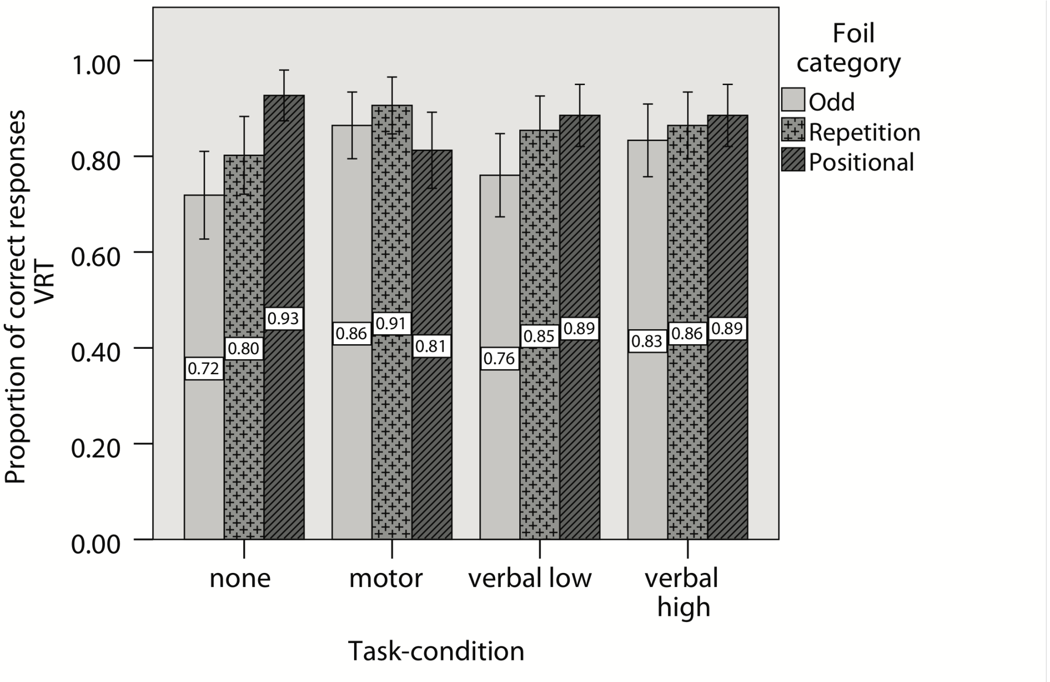


Fig.10. Average VRT performance across task-conditions for all foil categories. Participants scored above chance in all foil categories across all conditions, suggesting that, on a group level, no single heuristic was used to solve VRT.

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. incorrect secondary task responses. Results are depicted in Fig.11. Surprisingly, the proportion of correct trials in VRT was *higher* when the secondary task response was also correct ($\chi2 = 19.9, p < 0.001$). 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* $\chi2 = 6.1$, $p = 0.1$), while there was a similar significant main effect of session position (*Wald* $\chi2 = 25.5, p < 0.001$).

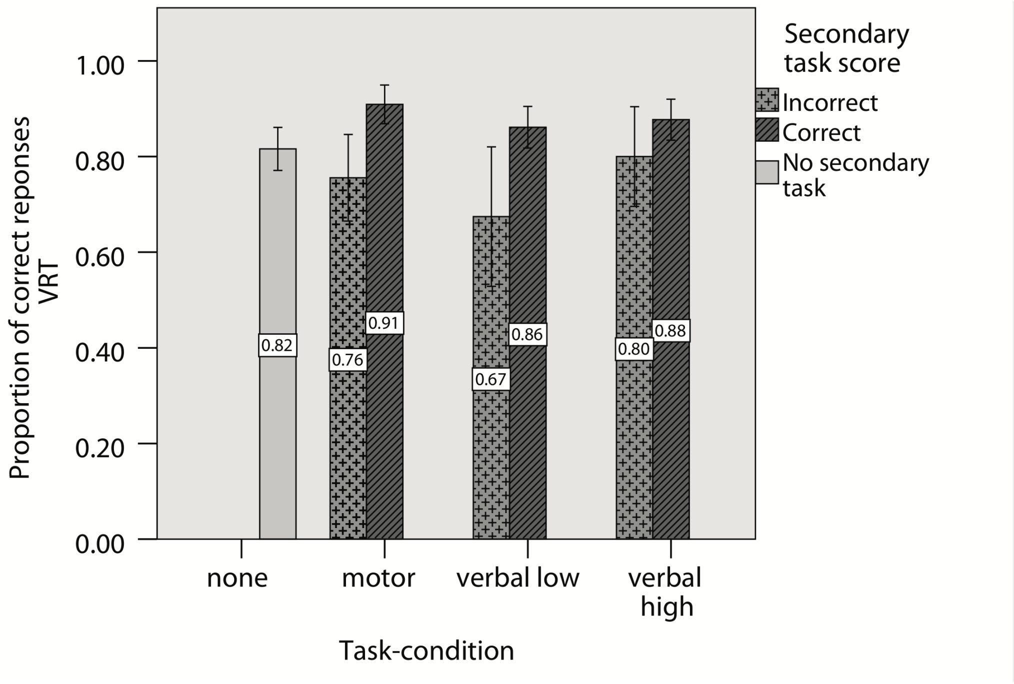


Fig.11. VRT performance across task-conditions, showing trials with correct and incorrect secondary task scores separately. Trials with correct secondary tasks responses yielded better performance in VRT.

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 rules 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 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 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 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 language evolution

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, Fedorenko, Lai, Saxe, & Gibson, 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). These effects have been shown for representations of exact numerosity (Frank et al., 2012) and for taxonomic categorization along perceptual dimensions (Druks, 2000). 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 session we did not find any differences between conditions with and without interference.

Taken together, these previous findings of verbal interference 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 a recursive interference task, 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.

Acknowledgements

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.

References

Alegre , M. A., & Gordon, P. (1996). Red rats eater exposes recursion in children's word formation. *Cognition, 60*, 65-82.

Arsenijević, B., & Wolfram, H. (2010). Recursion as a human universal and as a primitive. *Biolinguistics, 4*(2-3), 165-173.

Baddeley, A. (1998). The central executive: A concept and some misconceptions. *Journal of the International Neuropsychological Society, 4*, 523–526.

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. doi: 10.1038/nn.2277

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 Richard K Larson, Viviane M Déprez & Hiroko 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. doi: 10.1080/15326900701399947

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.

Druks, J. (2000). Selective preservation of naming from description and the "restricted preverbal message". *Brain and Language, 72*(2), 100-128. doi: 10.1006/brln.1999.2165

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.

Fitch, W.T. (2010). Three meanings of "recursion": Key distinctions for biolinguistics. In Richard K Larson, Viviane M Déprez & Hiroko Yamakido (Eds.), *The evolution of human language* (pp. 73-90). New York: Cambridge University Press.

Fitch, W.T., Hauser, M., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition, 97*(2), 179-210. doi: 10.1016/j.cognition.2005.02.005

Fodor, J. A. (1983). *The modularity of the 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. doi: 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. doi: 10.1016/j.cogpsych.2011.10.004

Friederici, A.D. , Bahlmann, J., Friedrich, R., & Makuuchi, M. (2011). The neural basis of recursion and complex syntactic hierarchy. *Biolinguistics, 5*(1-2), 087–104.

Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature, 440*(7088), 1204-1207. doi: 10.1038/nature04675

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. doi: 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.

Hanley, J. A. (2003). Statistical analysis of correlated data using generalized estimating equations: An orientation. *American Journal of Epidemiology, 157*(4), 364-375. doi: 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.

Hulst, H. v.d. (2010). Re recursion. In Harry van der Hulst (Ed.), *Recursion and human language*: de Gruyter Mouton.

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. doi: doi: 10.1016/j.cognition.2005.04.006

Kahneman, D. (2011). *Thinking fast and slow*. New York: Farrar, Straus and Giroux.

Karlsson, F. (2010). Syntactic recursion and iteration In Harry van der Hulst (Ed.), *Recursion and human language* (pp. 43-67). Berlin/New York: de Gruyter Mouton.

Koike, H., & Yoshihara, H. (1993). *Fractal approaches for visualization of huge hiearchies.* Paper presented at the Proceedings of the 1993 IEEE Symposium on Visual Languages.

Laury, R., & Ono, T. (2010). Recursion in conversation: What speakers of finnish and japanaese know how to do. In Harry van der Hulst (Ed.), *Recursion and human language*: de Gruyter Mouton.

Levinson, S. (2013). Recursion in pragmatics. *Language, 89*(1), 149-162. doi: 10.1353/lan.2013.0005

Lobina, D.J. (2011). Recursion and the competence/performance distinction in agl tasks. *Language and Cognitive Processes, 26*(10), 1563-1586. doi: 10.1080/01690965.2011.560006

Lupyan, G. (2009). Extracommunicative functions of language: Verbal interference causes selective categorization impairments. *Psychonomic Bulletin & Review, 16*(4), 711-718. doi: 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. doi: 10.1007/s10339-010-0368-6

Mandelbrot, B. (1977). *The fractal geometry of nature*: Freeman.

Martins, M.D. (2012). Distinctive signatures of recursion. *Philosophical Transactions of the Royal Society B, 367*, 2055-2064.

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.

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

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. doi: 10.1016/j.cub.2008.04.021

Odifreddi, P.G. (1999). *Classical recursion theory*: Elsevier.

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. doi: 10.1126/science.1102085

Pinker, S., & Jackendoff, R. (2005). The faculty of language: What's special about it? *Cognition, 95*(2), 201-236. doi: 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. doi: 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. doi: 10.1111/j.1467-9280.2009.02377.x

Reboul, A. (2012). Language: Between cognition, communication and culture. *Pragmatics & Cognition, 20*(2), 295-316. doi: 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, Wien 2002.

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

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. doi: 10.1016/j.jphysparis.2006.03.014

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. doi: 10.1098/rstb.2002.1257

6. How children perceive fractals: hierarchical self-similarity and cognitive development

Martins M, Laaha S, Freiberger E, Choi S, Fitch WT (in press). How children perceive fractals: hierarchical self-similarity and cognitive development. *Cognition*

How children perceive fractals: hierarchical self-similarity and cognitive development.

Maurício Dias Martins[a,b]
Sabine Laaha[c]
Eva Maria Freiberger[c]
Soonja Choi[c,d]
W. Tecumseh Fitch[a]

[a]Department of Cognitive Biology, University of Vienna, Althanstrasse 14, A-1090 Vienna, Austria,
[b]Language Research Laboratory, Lisbon Faculty of Medicine, Av. Professor Egas Moniz, 1649-028 Lisboa, Portugal
[c]Department of Linguistics, University of Vienna, Porzellangasse 4, A-1090 Vienna, Austria
[d]Department of Linguistics & Asian/Middle Eastern Languages, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182-7727, USA

Abstract

The ability to understand and generate hierarchical structures is a crucial component of human cognition, available in language, music, mathematics and problem solving. Recursion is a particularly useful mechanism for generating complex hierarchies by means of self-embedding rules. In the visual domain, fractals are recursive structures in which simple transformation rules generate hierarchies of infinite depth. Research on how children acquire these rules can provide valuable insight into the cognitive requirements and learning constraints of recursion.

Here, we used fractals to investigate the acquisition of recursion in the visual domain, and probed for correlations with grammar comprehension and general intelligence. We compared second (n=26) and fourth graders (n=26) in their ability to represent two types of rules for generating hierarchical structures: Recursive rules, on the one hand, which generate new hierarchical levels; and iterative rules, on the other hand, which merely insert items within hierarchies without generating new levels. We found that the majority of fourth graders, but not second graders, were able to represent both recursive and iterative rules. This difference was partially accounted by second graders' impairment in detecting hierarchical mistakes, and correlated with between-grade differences in grammar comprehension tasks. Empirically, recursion and iteration also differed in at least one crucial aspect: While the ability to learn recursive rules seemed to depend on the previous acquisition of simple iterative representations, the opposite was not true, i.e., children were able to acquire iterative rules before they acquired recursive representations. These results suggest that the acquisition of recursion in vision follows learning constraints similar to the acquisition of recursion in language, and that both domains share cognitive resources involved in hierarchical processing.

Keywords: Recursion; Iteration; Hierarchy; Language Evolution; Development; Visuo-spatial

Introduction

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

The ability to process hierarchical structures develops in an interesting way. Young children seem to have a strong bias to focus on the local information contained within hierarchies. For instance, in the visual-spatial domain, while attending to a big square composed of small circles, children have a tendency to identify the small circles faster and easier than they can identify the big square (Harrison & Stiles, 2009; Poirel, Mellet, Houdé, & Pineau, 2008). This local-oriented strategy to process hierarchical stimuli is similar to non-human primates (Fagot & Tomonaga, 1999; Spinozzi, De Lillo, & Truppa, 2003), and it usually precludes adequate hierarchical processing. Conversely, in human adults a global bias develops, in which global aspects of hierarchical structures are processed first, and where the contents of global information interfere with the processing of local information (Bouvet, Rousset, Valdois, & Donnadieu, 2011; Hopkins & Washburn, 2002). This ability to represent items-in-context is one of the pre-requisites of hierarchical processing. In other domains such as in language, children display equivalent impairments: they seem to grasp the meaning of individual words, and of simple adjacent relationships between them, but display difficulties in extracting the correct meaning of sentences containing more complex constructions (Dąbrowska, Rowland, & Theakston, 2009; Friederici, 2009; Roeper, 2011). This progressive development in the ability to integrate local and global information within hierarchies seems to be associated with brain maturational factors (Friederici, 2009; Moses et al., 2002), but also with the amount of exposure to the particular kinds of structures that children are asked to process (Roeper, 2011).

In this study, we are interested in investigating a particular aspect of hierarchical processing, which is the ability to encode hierarchical self-similarity. Hierarchies contain different levels that relate to each other in dominance-subordination relationships. For instance, in a 'Company X' with three hierarchical levels (C.E.O→Manager→Employee), there are two hierarchical relationships (C.E.O→Manager; and Manager→Employee). However, the same hierarchy could be characterized using more general concepts that highlight the similarities between different hierarchical relationships. For instance, the general characterization 'Member of Company X→ Member of Company X → Member of Company X' would require a single rule to encode the whole hierarchy (Member of Company X→Member of Company X). The ability to perceive similarities across hierarchical levels (i.e. hierarchical self-similarity) can be advantageous in parsing complex structures (Koike & Yoshihara, 1993). On the one hand, representing several levels with a single rule obviously reduces memory demands. On the other hand, this property allows the generation of new (previously absent) hierarchical levels without the need to learn or develop new rules or representations. This ability to represent hierarchical self-similarity, and to use this information to make inferences is closely associated with the concept of 'recursion'(Fitch, 2010; Hofstadter, 1980; Martins, 2012; Penrose, 1989).

One famous class of recursive structures is the fractals. Fractals are structures that display self-similarity (Mandelbrot, 1977), so that they appear geometrically similar when viewed at different scales. Fractals are produced by simple rules that, when applied iteratively to their own output, can generate complex hierarchical structures. Since the same kind of representation can be used at different levels of depth, simple rules suffice to represent the entirety of the structure. An example of a process generating a visuo-spatial fractal is depicted in Fig.1. Here, a simple recursive rule adds a triad of smaller hexagons around each bigger hexagon. Since the relations between successive hierarchical levels are kept constant, individuals representing recursion can make inferences about new (previously absent) hierarchical levels (Martins, 2012). This is the principle that we use in our investigation (For a more details, see Appendix A). Our goal was to investigate how the ability to represent hierarchical self-similarity develops in the visual domain, and how this ability can be predicted by individual differences in intelligence, grammar comprehension and general visual processing.

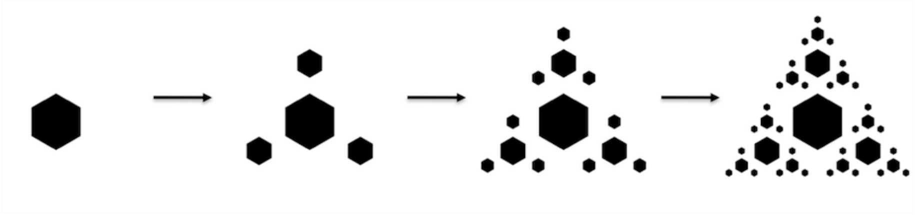


Figure 1. Recursive process generating a visual fractal.

The ability to represent hierarchical self-similarity has been empirically tested in the syntactic domain and in the visual domain (Martins & Fitch, 2012; Roeper, 2007). However, the developmental aspects of this ability have only been investigated in language (Roeper, 2011). In the next sections we briefly review what is currently known, and why it is important to extend this analysis to the visual-spatial domain.

Hierarchical self-similarity and language

Within the domain of language, recursion seems to be universally used (Reboul, 2012), and although rare in common speech (Laury & Ono, 2010), most language users are likely to have generated recursive sentences (for instance, compound nouns such as "[[[student] film]] committee]"). The widespread use of recursion in syntax has lead to the influential hypothesis that recursion would be part of a computational module specific to language (Hauser et al., 2002). In its strongest version, the thesis 'minimalist program' postulates recursion as the central operation of most syntactic processes (Chomsky, 2010). Within this theory, the usage of recursion in other domains would be parasitic on language. It is thus essential to empirically investigate the ability to acquire recursion in non-linguistic domains.

The development of recursion remains controversial. In English, children as young as 7-years-old are able to generate novel recursive structures, despite being exposed to a very limited recursive input (Berwick, Pietroski, Yankama, & Chomsky, 2011; Roeper, 2009). They can also discriminate well-formed recursive constructions at the age of 3 (Alegre & Gordon, 1996). This has been taken as evidence that children are able to represent recursion *a priori.* Studies concerning the processing of child directed speech suggest that the presence of recursive rules as Bayesian priors better explain the acquisition of language than priors without recursion (Perfors, Tenenbaum, Gibson, & Regier, 2010). In this Bayesian framework, although the ability to represent recursion is assumed to be present in the cognitive repertoire of young children, its explicit use in particular kinds of constructions may require experience with enough examples from those specific kinds. This experience may rapidly lead to the development of abstract representations, if a process of overgeneralization occurs (Perfors, Tenenbaum, Griffiths, & Xu, 2011; Perfors, Tenenbaum, & Regier, 2011). Consistent with this framework, the ability to represent recursion becomes available at different ontogenetic stages for different syntactic categories (Alegre & Gordon, 1996; Roeper, 2007, 2011). Initially, children tend to interpret linguistic hierarchies as non-recursive (Roeper, 2011), before they substitute these representations with more abstract (recursive) ones (Dickinson, 1987). This substitution process occurs if non-recursive representations become insufficient.

In sum, there are two main factors which can influence the ontogenetic development of the ability to represent hierarchical self-similarity. The first factor is a general process of brain maturation, which could impose hard limits on the kinds of information children are able to encode. Adult-like brain connectivity does not occur until the age of 8-9 (Friederici, 2009; Power, Fair, Schlaggar, & Petersen, 2010), and this brain connectivity pattern seems to enhance the ability to understand hierarchical structures (both recursive and non recursive). The second factor concerns experience, and the cumulative acquisition of constructions of increased abstraction (from non-recursive to recursive). In the current study we were interested in investigating the contribution of these factors in the acquisition of recursion in a non-linguistic domain. We developed a visuo-spatial paradigm using fractal stimuli to which children are not normally exposed. Thus, we could assess the ability to acquire novel recursive representations independently of the effects of previous exposure.

Current study

Here, we investigated whether the ability to represent structural self-similarity in visual hierarchies (fractals) followed a developmental time course similar to recursion in language, and occurred under similar learning constraints. We decided to compare two groups of children - second graders (7- to 8-year-olds) and fourth graders (9- to 10-year-olds) – which seem to differ in their ability to understand hierarchical and recursive structures in the linguistic domain (Friederici, 2009; Miller, Kessel, & Flavell, 1970). Differences between these groups have also been reported within the visual domain: children below the age of 9 seem to have a strong bias to focus on local visual information (Harrison & Stiles, 2009; Poirel et al., 2008), which as we have discussed, can affect normal hierarchical processing. Interestingly, also adults seem to display a strong local bias when exposed to novel and complex structures (Harrison & Stiles, 2009). This suggests that both maturational and experience factors play a role in determining visual processing strategies.

The paradigm that we used in this experiment was based on (Martins & Fitch, 2012): we present a series of images that build up a particular type of structure, incrementally, and the participants are asked to choose between two possible "completion" images that continue the pattern. In all cases, one of these two images is the "correct" continuation of the pattern in the first three images, and the other is a foil, quite similar but differing in some crucial respect. In the current experiment we did not provide response feedback, hence we could assess the natural cognitive abilities of the children, whether they were able to generalize the structural features of recursive stimuli. In this version of the task we also included stimuli with different levels of visual complexity, to evaluate the role of this factor, which is orthogonal to recursion itself, in the ability to extract hierarchical self-similarity principles in the visual domain. We included several categories of foils in order to prevent the use of simple heuristic strategies, and we added a second, non-recursive iterative task, with the same apparatus and experimental conditions as the ones described for the recursive task. Finally, we included a grammar comprehension and a non-verbal intelligence task in the test battery. With this setup we could investigate not only whether there are age differences in the ability to represent visual recursion and non-recursive iteration, but also the influence of several factors potentially related with these differences, namely: grammar comprehension, general intelligence and sensitivity to visual complexity. The inclusion of a grammar comprehension task in the procedure is also interesting to investigate whether there are domain-general factors involved in the processing of hierarchical structures. If recursion is the core computational operation of syntactic operations (Chomsky, 2010), and if open-ended representations of self-similar hierarchies are parasitic on the linguistic domain (Fitch, Hauser, & Chomsky, 2005; Hauser et al., 2002), we would expect to find a strong and specific correlation between grammar comprehension and visual recursion.

Methods

Participants

A total sample of 52 children took part in the study. They were all monolingual native speakers of German and were recruited from an elementary school in a middle-to-high socioeconomic neighborhood in Vienna (Austria). They were divided into two grade groups: 26 children (14 males) attended the second grade and aged 7 to 8 years (*M* = 8;2, range = 7;7–8;8); and 26 children (15 males) attended the fourth grade and were 9 and 10 years (*M* = 10;2, range = 9;8–10;4). Exclusion criteria included bilingualism, known neurological and psychiatric medical history, developmental learning disorders, and visual or auditory impairment. Children's participation was conditional upon approval by their head teachers and teachers, and their own willingness to take part in the experiment. They were aware that they could withdraw from the experiment at any time without further consequences. Moreover, all parents provided written informed consent for their children's participation in the study, and all data was stored anonymously.

Procedure

Children were tested individually in a quiet room at their school, in a single session of approximately 45 min. During this session, participants performed 4 tasks: 1) The Visual Recursion Task (VRT), designed to assess the ability to represent recursive iterative processes in the visuo-spatial domain (Martins & Fitch, 2012); 2) The Embedded Iteration Task (EIT), designed to test the ability to represent non-recursive iterative processes in the visuo-spatial domain (Martins & Fitch, 2012); 3) The Test for Reception of Grammar (TROG-D), a grammatical comprehension task (Bishop, 2003; Fox, 2007).; and 4) The Raven's Coloured Progressive Matrices (CPM), a non-verbal intelligence task (Raven, Raven, & Court, 2010)

The whole testing procedure was divided into two parts, with a break of 5 minutes in between. The first part included VRT and EIT, as well as a specific training for these tasks, and the second part included TROG-D and CPM. The order of tasks in the first part was randomized and equally distributed: Within each grade group 13 children started the procedure with VRT and 13 children started the procedure with EIT. The order of tasks in the second part was fixed (TROG-D first and then CPM).

Both VRT and EIT were binary forced-choice paradigms, where children were asked to choose between two images. After the completion of the first two tasks, we asked 42 out of 52 children the following question: "How frequently were the two images in the bottom different? a) Almost never, b) Sometimes, or c) Almost always?" We initiated this systematic questioning after the experiment had begun, due to the feedback that we got from some children, reporting perceiving no differences between the choice images. Unfortunately, it was not possible to retrieve the answer from the first 10 children.

Visual Recursion Task (VRT)

Test procedure. This task was adapted from (Martins & Fitch, 2012). In VRT, each trial began with the presentation of three images corresponding to the first three iterations (steps) of a fractal generation. These images were presented non-simultaneously in the top half of the screen, sequentially from left to right ('Sequence images'; Fig.2), with an interval of 2 seconds between the presentation of one image and the next. After the presentation of the first three iterations, two additional images were presented simultaneously in the bottom half of the screen ('Choice images'; Fig.2). 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 touch the image they considered as the correct continuation of the recursive process, and their response was captured using a touch-screen (Elo Touchsystems). The position of the 'correct' image (LEFT or RIGHT) was randomized. The same instructions were given (in German, and during training only) to all participants:

Instructions in German: "*Schau mal, das Bilderrätsel da geht so: Da oben sind drei Bilder. Und da unten sind zwei Bilder. Und du sollst dann unten auf das richtige Bild drücken. Das ist das erste Bild, das ist das zweite Bild und das ist das dritte Bild.*

Welches ist das richtige ***<u>nächste</u>*** *Bild: das oder das? [*Feedback: *Super, das hast du richtig gemacht.* *<u>(or)</u>* *Nein, das war jetzt nicht richtig. Schau mal, das ist das richtige Bild."*

Instructions (English translation): "*Look, this picture puzzle works like this: Up at the top there are three pictures. And down below there are two pictures. You have to press on the correct picture down below. This is the first picture, this is the second picture, and this is the third picture. What is the correct* ***<u>next</u>*** *picture: this or that? [Feedback: Great, you got it right. (or) No, that was not correct. Look, this is the correct picture.]"*

After the initial instructions, each trial had a maximum duration of 30 seconds before a timeout. No visual or auditory feedback was given regarding whether the answer was correct or incorrect. The task comprised 27 trials, and had a total duration of about 12 minutes.

To test for effects of information processing constraints, we included stimuli with different degrees of visual complexity (complexity '3','4', and '5'). Furthermore, in order to control for the usage of simple visual heuristic strategies in VRT performance, we included several categories of foils ('Odd', 'Position' and 'Repetition'). For details on stimuli generation and stimuli categories, see appendix A and Fig.3. Overall, the combination of both 'visual complexity' and 'foils' categories resulted in 9 types of stimuli: Complexity 3, 4 and 5 with odd constituent foils; Complexity 3, 4 and 5 with positional error foils and Complexity 3, 4 and 5 with repetition foils. Exactly three examples of each type of stimuli were generated using the programming language Python, resulting in a total of 27 stimuli.

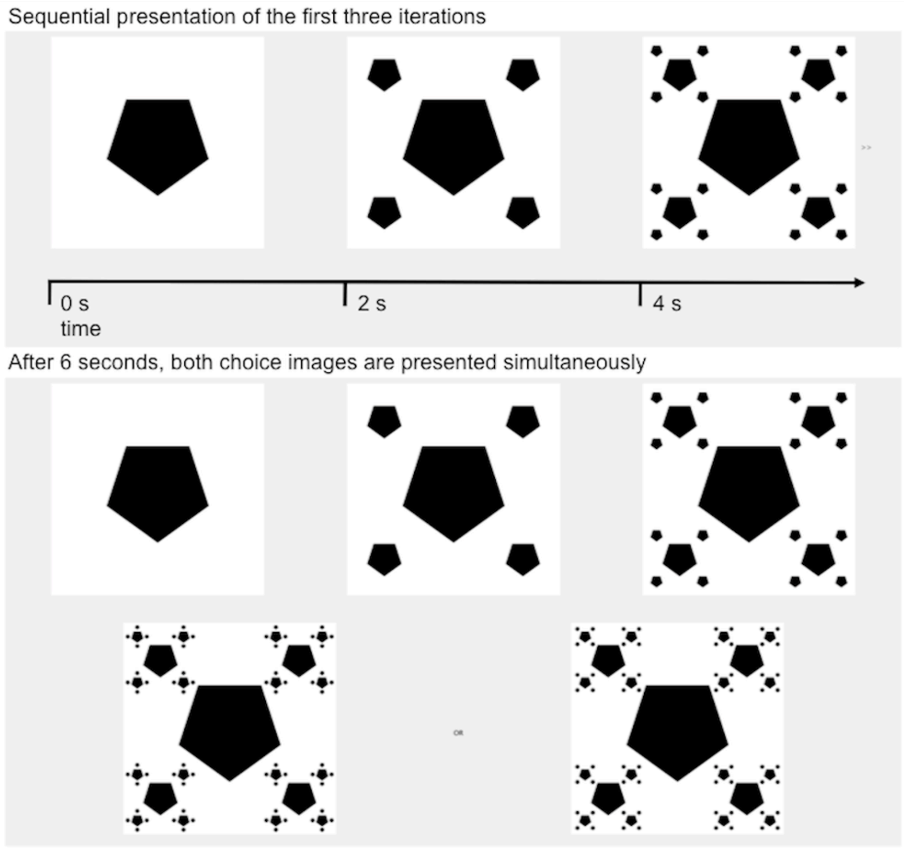


Figure 2. Example of a typical Visual Recursion Task trial. Initially, the first three iterations of a fractal generation are depicted, sequentially, from left to right (top). Then, two images are presented simultaneously in the bottom part of the screen, corresponding to the 'correct' fourth iteration (bottom right) and a foil (bottom left), and the participant chooses between them. In this example, the foil is a 'positional foil' (see Fig.3 for details on foils).

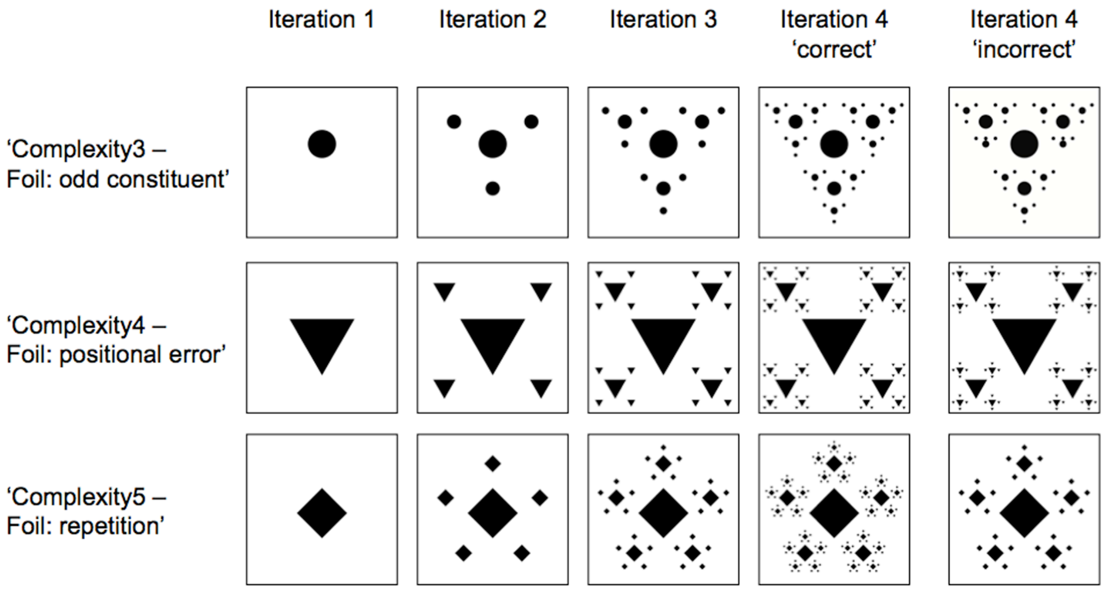


Figure 3. Examples of fractals used in the Visual Recursion Task. The first four iterations of a fractal generation, as well as one foil ('incorrect' fourth iteration), were produced. There were different categories of 'visual complexity' – 3, 4 and 5 - and different categories of foils. In 'Odd constituent' 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.

Embedded Iteration Task (EIT)

The second task was hierarchical but non-recursive, and was adapted from (Martins & Fitch, 2012). The principle underlying EIT is similar to VRT in the sense that it involves an iterative procedure applied to hierarchical structures. However, EIT lacks recursive embedding. Instead, in EIT, additional elements are added to one pre-existing hierarchical structure, without producing new hierarchical levels (Fig.4). As for VRT, an understanding of this iterative procedure is necessary to correctly predict the next iteration.

All the apparatus and experimental conditions for EIT were identical to the ones described for VRT, including number of trials, duration of each trial, 'visual complexity' categories, foil categories, and feedback conditions (see Appendix A).

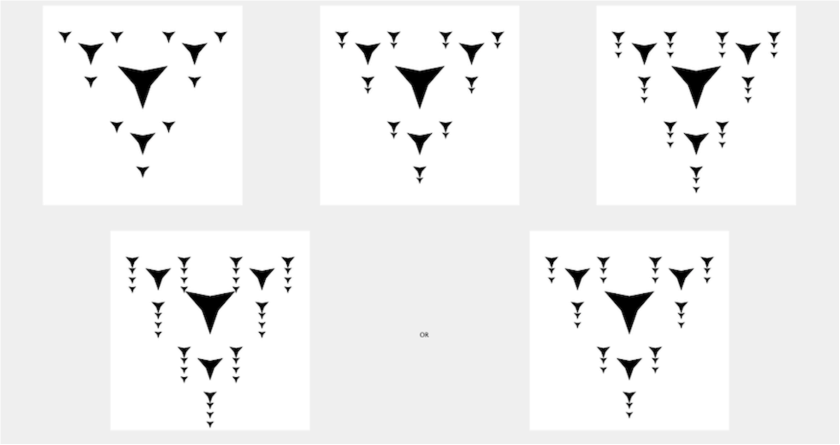


Figure 4. Example of a typical Embedded Iteration Task trial. In this example, the correct answer is on the bottom left. The foil on the bottom right is a "Repetition foil" (see Fig.3 for details on foils).

Training

Prior to the beginning of the first part of the procedure (composed by VRT and EIT), a short training session was given. The goal of this training was to give children the opportunity to manipulate the touch-screen, and to introduce them to the specific environment of VRT and EIT trials before testing. Four training items were given: Two items followed an iterative rule, which was not hierarchical (see Appendix B for an example); one item was iterative and hierarchical, but not recursive (similar to the items of EIT); and the last item was iterative, hierarchical and recursive (similar to VRT). If participants provided an incorrect response, the same item was presented again until a correct response was provided. In case of repeated failure, the experimenter tried to motivate the child (during training only) by drawing his/her attention to the structure of the trial, and repeating the instructions if necessary.

TROG-D

TROG-D is a grammatical comprehension task designed for children aged 3 to 11 years. It is the German adaptation of the English *Test for Reception of Grammar* - TROG (Bishop, 2003) and was standardized using the data from 870 monolingual German-speaking children (Fox, 2007). The test consists of 84 test items grouped into 21 test blocks, with increasing difficulty: nouns, verbs, adjectives, 2-element sentences (SV), 3-element sentences (SVO), negation, prepositions ('in/on'), perfect tense, plural, prepositions ('above/below'), passive, personal pronouns (nominative), relative clauses (nominative), personal pronouns (accusative/dative), double object constructions, subordination ('while/after'), topicalization, disjunctive conjunctions ('neither-nor'), relative clauses (accusative/dative), coordination ('and'), subordination ('that'). Test items are presented in a four picture multiple-choice format with lexical and grammatical foils. The test procedure is as follows: The investigator reads aloud the test item to the child (e.g. relative clause (nominative): *Der Junge, **der** das Pferd jagt, ist dick* 'The boy, who is chasing the horse, is chubby'), and the task of the child is to point at the appropriate picture in the test booklet. Participants' responses are analyzed by test block (N=21); in order for a test block to be classified as correct, all responses within the test block have to be correct.

CPM

Raven's Coloured Progressive Matrices (CPM) is a non-verbal intelligence task (with a focus on logical reasoning) designed for children aged 5 to 11 years (Raven et al., 2010). The test consists of 36 test items grouped into 3 test sets (A, Ab, B), with 12 test items each. Test sets are arranged in a way so as to allow development of a consistent method of thinking; set A: completion of a single, continuous pattern, sets Ab and B: completion of discrete patterns. Test items are presented in a six-picture multiple-choice format. In each test item, the task of the child is to identify the missing element that completes a pattern and to point at it in the test booklet. Participants' responses are analyzed by test item (N=36).

Predictions

Based on the previous discussion, our working hypothesis was that the ability to represent recursion becomes available at later ontogenetic stages than the ability to represent iteration, and that this difference is partially explained by biological development factors. Consequentially, our predictions were the following: 1) Fourth graders were expected to perform adequately in both recursive and iterative tasks, while second graders might be expected to do so in the non-recursive iterative task only; 2) Visual complexity was expected to play a role in performance, especially among the second graders; 3) The ability to perform adequately in the visual recursion task was expected to correlate in general with grammar comprehension abilities, and specifically with the comprehension of sentences with embedded clauses.

Alternatively, if no differences in the ability to represent recursion and iteration were found within grades, this would lend support to the hypothesis that cumulative exposure, rather than biological development, is the main limiting factor on the ability of children to represent recursion, once iteration is already available.

Analyses

Our overall goal was to assess children's ability to represent recursion and embedded iteration in the visual domain and to compare performance between second and fourth grade. Furthermore we investigated the effects of visual complexity, visual strategies (foil categories), task-order, grammar abilities and non-verbal intelligence.

In our data, we used the binomial variable VRT and EIT 'trial correctness' (correct/incorrect) as the dependent variable for regression models. When overall response data were not normally distributed (assessed using a Shapiro-Wilk test), we used non-parametric statistics. Simple response accuracy comparison between grades was performed with an unpaired Mann-Whitney *U* test. To assess whether each participant had VRT and EIT scores above chance, we first calculated the proportion of correct (and incorrect) answers that deviated significantly from chance using a Binomial test. Since we used a binary forced-choice task, the probability to score correctly due to chance was 50%. In a total of 27 test items, a number of correct answers equal or superior to 20 (i.e. a proportion of 0.74), or equal or inferior to 7 (i.e. a proportion of 0.26), is the number which differs significantly from chance (Binomial test, $p = 0.019$). The comparison between second and fourth grades, regarding children that scored above chance, was performed using a Chi-square test.

Finally, to assess the effects of visual strategies (foil categories), visual complexity, task-order, grammar abilities and non-verbal intelligence, we used a semiparametric regression technique called Generalized Estimating Equations (GEE), a technique useful when analyzing binomial data with within-subjects effects (Hanley, 2003). We created several models containing different variables: 'grade' and 'task-order' as between-subjects variables; 'task', 'foil category' and 'visual complexity' as within-subjects variables; and 'grammar' and 'intelligence' raw scores as covariates.

All analyses were performed with SPSS® 19.

Results

VRT

General overview: correct responses by grade. On average, the 26 children attending the fourth grade ($M = 0.80$, $SD = 0.21$) had a significantly higher proportion of correct responses in VRT than children attending the second grade ($M = 0.59$, $SD = 0.17$) (*Mann-Whitney U:* $z = -3.70$, $p < 0.001$; Fig.5). Moreover, while 69.2% of fourth graders had a proportion of correct answers above chance, only 26.9% of the second graders had so. This difference was also significant ($\chi 2 = 9.43$, $p = 0.002$). One child in the fourth grade and one in the second grade had performance scores lower than predicted by chance (i.e. equal or lower than 0.26). Since this highly non-random pattern cannot be explained by a failure in discriminating the recursive rule, these two participants were excluded from further regression and correlation analyses involving VRT.

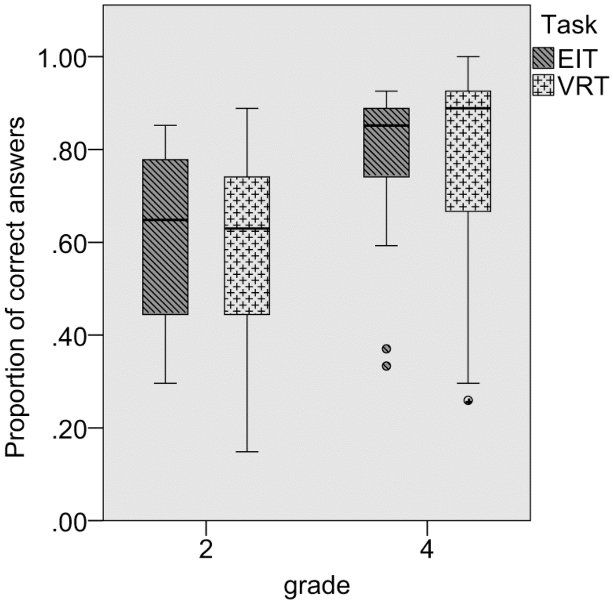


Figure 5. Performance in Visual Recursion Task (VRT) and Embedded Iteration Task (EIT), across grades. Fourth graders had higher scores than second graders, in both VRT and EIT. Within each grade, the difference between tasks was not significant.

Visual strategies. A central issue concerning our method is the question of whether participants were able to represent the structural self-similarity present in the recursive images; and to apply this knowledge throughout different VRT trials. One possible alternative to the representation of self-similarity would be the usage of heuristic strategies, based on the detection of simple salient features within the foils, which would allow their exclusion without an understanding of the underlying structure. In order to prevent the emergence of a systematic 'choice-by-exclusion' strategy, we used different categories of foils. Our assumption was that, if individuals were able to represent self-similarity, they would perform adequately in all different foil categories.

At the group level, the number of correct choices was significantly above chance for all foil categories and for both grade groups (Binomial test, $p < 0.005$). For detailed analyses comparing performance across categories see Appendix C.

Visual complexity. Another important issue concerns the role of visual complexity. It is possible that the ability to perform adequately in VRT is limited by the capacity to cope with the amount of visual information. In our experiment, fractals of 'complexity 5' contained a higher number of elements (for instance, squares) than stimuli of 'complexity 3' (Fig.3), and greater amount of visual information may be harder to process. To analyse this effect we compared the performance between trials displaying different amounts of visual complexity using a GEE with 'grade' as a between-subjects factor, and 'visual complexity' as a within-subjects factor. We found that visual complexity had a significant main effect on VRT performance (*Wald* $\chi2 = 6.5$, $p = 0.039$). Specifically, the proportion of correct answers in the category 'complexity4' was higher than in the category 'complexity5' (*estimated marginal mean (EMM) difference* = 0.06, $p = 0.026$). All p-values were corrected using sequential Bonferroni correction. Detailed grade*visual complexity interaction analyses and figures are presented in Appendix D. Overall, higher levels of visual complexity yielded worse results, especially within second graders.

EIT

General overview: correct responses by grade. On average, children attending the fourth grade ($M = 0.78$, $SD = 0.18$) had a higher proportion of correct responses in EIT than children attending the second grade ($M = 0.62$, $SD = 0.17$). This was a significant difference (*Mann-Whitney U:* $z = -3.70$, $p < 0.001$; Fig.5). While 77% of fourth graders had a proportion of correct answers above chance, only 35% of the second graders had so. This difference was also significant ($\chi2 = 5.2$, $p = 0.023$).

Visual strategies. We repeated the analysis described for VRT, now with the proportion of correct answers in EIT as the dependent variable. Our results suggest that, at the group level, second graders performed randomly in the foil category 'odd constituent' (Proportion = 0.52, Binomial test, p = 0.556). For all other foil categories and for both grade groups, performance was significantly above chance (Binomial test, p < 0.005). Detailed comparisons across categories are presented in appendix C.

Visual complexity. We repeated the complexity analysis described for VRT, with the proportion of correct answers in EIT as the dependent variable. We again found that visual complexity had a significant main effect on performance (*Wald* $\chi2 = 12.6$, $p = 0.002$): The proportion of correct answers in the category 'complexity3' was higher than in the categories 'complexity4' (*EMM difference* = 0.06, $p = 0.012$) and 'complexity5' (*EMM difference* = 0.07, $p = 0.06$). All p-values were corrected using sequential Bonferroni correction. Detailed figures, interaction analyses, and subsequent pair-wise comparisons are presented in Appendix D. Overall, results suggest that visual complexity also plays a role in the ability to perform adequately in EIT, with fewer constituents easier to process

VRT vs. EIT and effects of task order

In order to compare children’s performance in VRT and EIT, we ran a GEE model with ‘grade’ as a between-subjects factor, and ‘task’ as a within-subjects factor. There was a significant main effect of grade (*Wald* $\chi2 = 12.9$, $p < 0.001$), but no difference between tasks ($p = 0.9$) and no interaction between grade and task (*Wald* $\chi2 = 1.4$, $p = 0.24$), suggesting the grade effects were not specific to recursion (Fig.5).

To assure the validity of comparisons between VRT and EIT, we balanced the order of the tasks in the procedure. However, we noticed that one of the ‘task-order’ conditions yielded lower performance than the other. Specifically, participants starting the procedure with VRT had a significantly lower response-accuracy (on both tasks VRT and EIT combined; $M = 0.63$, $SD = 0.21$) than participants that started with EIT ($M = 0.72$, $SD = 0.17$; *Mann Whitney* $U = 851$, $z = -3.2$, $p = 0.001$). To further explore this, we first investigated whether performance was differently affected in different tasks and in different grades (Fig.6).

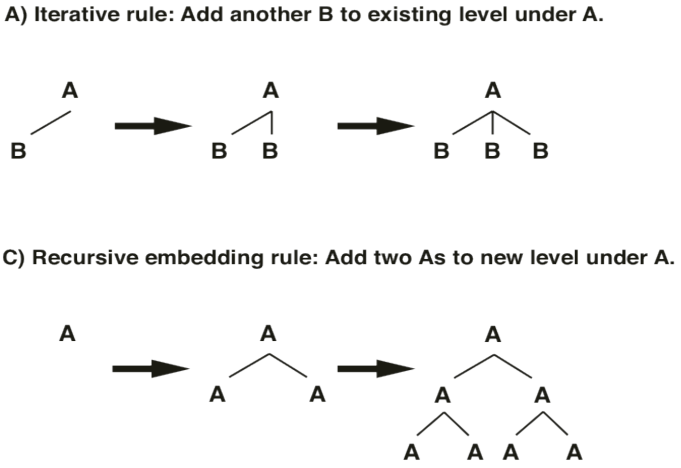


Figure 6. Performance across different task-sequence conditions. In the sequence condition ‘VRT-EIT’ (left) participants performed the Visual Recursion Task (VRT) first; in the condition ‘EIT-VRT’ (right) participants performed the Embedded Iteration Task (EIT) first. Children who performed the iterative task first scored globally better than those who started with recursion. Crucially, starting the procedure with VRT decreased EIT accuracy. This suggests that children transferred knowledge from simple iteration to recursion, but not the other way around.

Before testing the effect of task-order, and to better interpret potential interactions between ‘task-order’ (‘VRT-EIT’ vs. ‘EIT-VRT’) and ‘task’ (VRT vs.

EIT), we recoded the former variable on a trial-by-trial basis. The new variable, called ‘position’, can be understood as the position of the task in the procedure. For instance, in trials where the task is ‘VRT’ and the order of tasks is ‘VRT-EIT’, the ‘position’ variable is coded as ‘FIRST’. Likewise, in trials where the task is ‘EIT’ and the order of tasks is ‘EIT-VRT’, the ‘position’ variable is coded as ‘FIRST’, etc.

We ran a GEE model with ‘task’ (VRT vs. EIT) and position (FIRST vs. SECOND) as within-subjects effects, and ‘grade’ (second vs. fourth) as a between-subjects variable. We analyzed ‘task’, ‘grade’ and ‘position’ main effects, and all possible interactions. The summary of the model is depicted in Table1.

Model (*QICC* = 3252)	Type III	
	Wald $\chi 2$	p
Intercept	87.8	0.000
Task (VRT vs. EIT)	0.5	0.464
Position (First vs. Second)	7.7	0.005
Grade (Second vs. Fourth)	25.9	0.000
Task * Position	16.2	0.000
Task * Grade	3.6	0.057
Position * Grade	2.2	0.138
Task * Position * Grade	5.2	0.022

Table 1. Effects of task, position, grade and all interactions in the processing of visual hierarchies. Here we present the results of a General Estimating Equations Model with the ‘correctness’ (correct/incorrect) of each trial from Visual Recursion Task (VRT) and Embedded Iteration Task (EIT) as the dependent variable. Overall, fourth graders scored better in both tasks, and accuracy was better in the second task of the procedure than in the first task. However, there was a strong interaction between task and position: Performance in VRT was higher when this task was performed after EIT (i.e. in the second position of the procedure), but not vice-versa (see Fig. 6). This suggests that children transfer knowledge from simple iteration to recursion, but not the other way around. We conclude that the overall similarity in accuracy between EIT and VRT masks interesting differences in learning constraints. QICC (Corrected Quasi Likelihood under Independence Model Criterion). The asterisk (*) denotes interaction.

We found significant main effects of ‘position’ and ‘grade’ on performance ($p < 0.001$), in agreement with the previous analyses. Furthermore, we found a significant interaction between ‘task’ and ‘position’. Performance in EIT-FIRST position was better than performance in VRT-FIRST position (*EMM difference* = 0.15, $p = 0.004$). Conversely, VRT-SECOND position yielded better performance than EIT-SECOND position (*EMM difference* = 0.17, $p = 0.001$).

Within VRT, the proportion of correct answers was higher when this task was performed in the <u>SECOND</u> position of the procedure than when the same task was performed in the first position (*EMM difference* = 0.21, $p < 0.001$). Within EIT, there was also a trend towards higher accuracy when this task was performed in the <u>FIRST</u> position than when it was performed in the second position (*EMM difference* = 0.11, $p = 0.052$). All p-values were corrected with sequential Bonferroni.

Additional interaction analyses are presented in Appendix E.

Overall, results suggest that the order of the task in the procedure had a strong influence on task performance. Specifically, VRT accuracy is increased by previous experience with EIT. However, this effect of task-order was not due to a practice effect during the experiment, since EIT performance decreased when this task was performed in the second position of the procedure.

Role of grammar comprehension ability and non-verbal intelligence

To assess whether the ability to represent visual recursion was predicted by language abilities, we tested all participants in the TROG-D, a test of grammar comprehension. Furthermore, to assess whether the potential effect of grammar comprehension was independent of general capacity factors, we tested the same participants in a non-verbal intelligence task – The Raven’s coloured progressive matrices (CPM). Participants’ raw score in TROG-D was $M = 16.9$, $SD = 2.0$ (*minimum:* 13, *maximum:* 20), while CPM raw score was $M = 29.2$, $SD = 3.6$ (*minimum:* 21, *maximum:* 34). Segregated by grade group, results were the following: Second graders’ score in TROG-D was $M = 15.9$, $SD = 2.0$ (*minimum:* 13, *maximum:* 20), while CPM raw score was $M = 27.9$, $SD = 3.6$ (*minimum:* 21, *maximum:* 34); Fourth graders’ score in TROG-D was $M = 18.0$, $SD = 1.4$ (*minimum:* 16, *maximum:* 20), while CPM raw score was $M = 30.5$, $SD = 3.0$ (*minimum:* 23, *maximum:* 34).

Overall, fourth graders scored significantly higher than second graders in both TROG-D ($t(50) = -4.5$, $p < 0.001$) and CPM ($t(50) = -2.9$, $p = 0.006$).

The overall proportion of correct answers in VRT was positively correlated with both CPM ($\rho(50) = 0.52$, $p < 0.001$) and TROG-D ($\rho(50) = 0.43$, $p = 0.002$) scores. Likewise, the proportion of correct answers in EIT was positively correlated with both CPM ($\rho(50) = 0.58$, $p < 0.001$) and TROG-D ($\rho(50) = 0.41$, $p = 0.003$) scores. To test whether grammar comprehension effects were specific to VRT and independent of general intelligence, we ran a GEE model with 'task' (VRT vs. EIT) as the within-subjects factors, and TROG-D and CPM scores as covariates. The summary of the model is depicted in Table 2. Our results suggest that grammar comprehension predicts performance of both VRT and EIT (main effect of TROG-D: $Wald\ \chi2 = 6.7$, $p = 0.01$), and that this effect is partially independent from non-verbal intelligence since both main effects are significant. However these effects were neither specific for VRT nor for EIT (no interaction between task and TROG-D: $p = 0.54$. We repeated this analysis using the more specific variable 'embedded clauses' (number of TROG-D blocks containing embedded clauses which were answered correctly; maximum score = 5). The results were similar: There was a main effect of 'embedded clauses' ($Wald\ \chi2 = 5.4$, $p = 0.02$), independent of intelligence, but not specific to VRT (interaction task*embedded clauses: $p = 0.9$).

Finally, we analysed the effects of grammar and intelligence within each grade group. We ran two GEE models, one for each grade (second and fourth). We found that CPM score was a predictor of both VRT and EIT within the second grade ($Wald\ \chi2 = 10.1$, $p = 0.001$), and fourth grade ($Wald\ \chi2 = 4.9$, $p = 0.03$); and that TROG-D score was not an independent predictor of VRT and EIT performance within each grade group ($p > 0.1$), i.e. only CPM predicted performance within each grade group.

Model (*QICC* = 3093)	Type III	
	Wald χ2	*p*
Intercept	25.9	0.000
Task (VRT vs. EIT)	0.3	0.600
TROG-D (grammar)	6.7	0.010
CPM (intelligence)	22.3	0.000
Task * TROG-D	0.4	0.542
Task * CPM	0.0	0.971

Table 2. Grammar comprehension is an independent predictor of visual hierarchical processing, but not specific of recursion. Here we present the results of a General Estimating Equations Model with the 'correctness' (correct/incorrect) of each trial from Visual Recursion Task (VRT) and Embedded Iteration Task (EIT) as the dependent variable. Grammar comprehension predicts performance in VRT, even after accounting for the variability explained by general intelligence. However, this effect is not specific to VRT but general to both VRT and EIT. QICC (Corrected Quasi Likelihood under Independence Model Criterion). The asterisk (*) denotes interaction.

Discussion

In this study, we investigated for the first time the ability of children to represent structural self-similarity in visuo-spatial hierarchies. In this experiment we used visual fractals, which children are very rarely exposed to. Hence, we could investigate the ability to acquire new recursive representations without the potential confound of previous exposure. Here, we aimed at investigating not only whether the ability to acquire recursive rules in vision followed a development course somehow similar to language, but also wanted to assess whether this acquisition process was subject to similar learning constraints. In order to explain potential inter-grade differences we explored the individual variation in visual processing efficiency, grammar comprehension and general intelligence.

We found that: A) the majority of fourth graders performed adequately in both recursive and iterative tasks, while many second graders failed in both; B) higher degrees of visual complexity reduced the ability to instantiate either recursive and non-recursive rules, but specially among the second graders; C) recursive representations of hierarchical structures yielded better results than non-recursive representations in the detection of errors nested within lower visual scales; D) we found an unexpected task-order effect: performance in visual recursion improved with previous experience with non-recursive iteration, but not vice-versa; E) both general grammatical abilities and first-order clause embedding were independent predictors of accuracy in the visual tasks. However, this effect was general to hierarchical iteration, and not specific to recursion.

Taken together, these results suggest that the ability to represent recursion and iteration may become available at similar stages during the ontogenetic development (around 9 years-old). However, once this potential is present, other factors related with cumulative exposure to hierarchical structures may play a role in the representation of hierarchical self-similarity. For instance, in our study, the prior acquisition of iterative rules was fundamental to the later acquisition of recursion (but not vice-versa). These results mimic the findings of language research (Roeper, 2011). Our results also suggest that age differences can be partially explained by differences in visual processing efficiency, since the effects of visual complexity are more pronounced in second graders, and this group is especially impaired in the detection of 'odd' foils. Finally, also grammar comprehension abilities partially account for these grade differences, independently of general intelligence. This suggests that the ability to process hierarchical structures in the linguistic and visual domains partially recruit similar cognitive resources, although these resources are not specific to recursion. If recursion were central to all syntactic processes in language, we would expect to find a specific correlation between visual and linguistic recursion, instead of a general correlation with hierarchical processing. Thus, our results seem to challenge Chomsky thesis (Chomsky, 2010).

Performance across grade

Our first important result was a demonstration that 9- to 10-year-old children are well able to represent recursion in the visual domain. The fact that they are able to do so without instructions and with no response feedback, suggests that they are spontaneously able to generalize the knowledge of structural self-similarity across test items. Furthermore, we used different categories of foils, and found no performance differences between them. This suggests that children who passed VRT did not rely on simple heuristic strategies, and were probably able to perceive all features necessary to represent hierarchical self-similarity. The fourth graders were also able to correctly continue non-recursive iteration and there were no significant differences between recursive and non-recursive tasks, although more fourth graders tended to perform above chance in EIT than in VRT (77% vs. 69%).

Perhaps more surprising was the finding that many second graders performed poorly in both recursive and non-recursive tasks. Since second graders are able to handle conjunctions (e.g. "John, Bill, Fred, and Susan arrived.") and to some extent syntactic structures like "What is the color of Bill's dog's balloon?" (Roeper, 2007, 2011), we might expect them to perform adequately in a visual task that requires the representation of iterative processes embedded within hierarchical structures. However, only 35% of second graders scored above chance in EIT (and only 27% performed adequately in VRT). There are several possible interpretations for these results: On the one hand, it is possible that the ability to represent iterative processes and hierarchical structures in the visual domain is not within the cognitive repertoire of second graders. On the other hand, it is possible that even though the potential to represent these structures is available, other factors related to our particular instantiations of iteration (or recursion) impaired their ability to make explicit judgements. One such factor might be the amount of visual complexity. Another factor may be that these children likely had little or no previous experience with visuo-spatial fractals before performing our experiment.

Effects of visual complexity

Overall, we found that higher levels of visual complexity reduced participants' ability to extract recursive and iterative principles. This effect seems to be more pronounced in the second grade group. Incidentally, we asked the majority of children (18 second graders and 24 fourth graders) how frequently they had detected differences between the choice images during the realization of our tasks (i.e. between foil and correct fourth iteration). While 17.6% of the questioned second graders reported perceiving no differences between 'correct' fourth iteration and foil most of the time, only 4.5% of the fourth graders did so. This provides additional evidence that younger children may have had difficulties detecting (or retrieving) information relevant to process the test stimuli. Previous research on the development of hierarchical processing suggests that before the age of 9 children seem to have a strong bias to focus on local visual information (Harrison & Stiles, 2009; Poirel et al., 2008), which as we have discussed, can affect normal hierarchical processing. Thus, further research will be necessary to determine whether the potential to represent recursion in vision is not part of the cognitive repertoire of many younger children; or whether inadequate performance was caused by inefficient visual processing mechanisms.

Dissociations between VRT and EIT: ‘Odd foil’ detection and task-order

Although we found no significant performance differences between VRT and EIT in overall, a closer analysis revealed two interesting dissociations:

First, unlike in VRT, children seemed to have difficulty in rejecting the ‘Odd constituent’ foils in EIT, though performance was adequate in trials containing other foils categories (‘Positional error’ and ‘Repetition’). Since they were able to respond adequately to this foil category while executing VRT, it seems unlikely that this result was caused by a general inability to perceive ’odd constituent’ mistakes. Instead, we suspect that there may be differences in the way recursive and non-recursive representations are cognitively implemented. These differences might have led subjects to detect errors of the ‘odd constituent’ type more efficiently in VRT. Previous studies (Martins & Fitch, 2012) suggest that EIT may be more demanding of visual processing resources than VRT. Moreover, we found here that the effects of visual complexity in EIT were broader than in VRT, extending not only to the second grade, but also to the fourth grade (see Appendix D). If performance in EIT is more dependent on bottom-up perceptual resources, and more sensitive to variations in low-level visual information, then it is plausible that subtle errors are harder to detect in this task than in VRT. In the ‘Odd constituent’ foils, these errors occur deeply nested within the hierarchical structure (i.e. at the smallest size scale), and only in a subset of hierarchical nodes. Elsewhere, it has been argued that recursive representations may be more efficient than non-recursive representations at encoding of hierarchical structures (Koike & Yoshihara, 1993; Martins, 2012). This greater efficiency might derive from the fact that the same “rules” can be used to represent different hierarchical levels, hence allowing a simultaneous encoding of the whole and of the details. Particularly in the visual domain, there is evidence that compressed representations lead to a better perception of fine-grained details (Alvarez, 2011).

A second difference found between VRT and EIT was the effect of task-order. Previous experience with EIT seemed to help children to perform adequately in VRT.

However, the inverse effect was not found, i.e. previous exposure to VRT did not enhance EIT accuracy. This asymmetry suggests that VRT performance enhancement after EIT was not due to a general learning effect. Instead, we think that this finding reflects different characteristics of recursive and non-recursive representations.

Iteration refers to the process of repeating an operation. If this operation involves the embedding of a constituent within constituents (e.g. putting a noun and a verb together to form the sentence "[[trees] grow]"), then it is termed hierarchical embedding (Martins, 2012), Recursion is a particular subset of hierarchical embedding, where both elements of a transformation rule are perceived as belonging to the same category (e.g. the embedding of a noun phrase within another noun phrase, as in "[[student] committee]"). It seems possible that children may require exposure to simpler iterative processes before they are able to identify hierarchical self-similarity. The reason why recursion may be harder to acquire could be related to the fact that constituents within recursive representations are at a higher level of abstraction. For instance, in our EIT stimuli, it suffices to build a representation of the initial structure [A], and of the constituents [B] being added into that structure: [A] →[A[B]] → [A[BB]] → [A[BBB]]. In recursion, in order to predict the next iteration, participants are required to encode successive hierarchical levels with the same rules. A typical representation of a recursive hierarchy would be [ALPHA[ALPHA[ALPHA]]]. This requires the formation of an abstract category [ALPHA], which incorporates the features of both [A] and [B]. In order to generate a representation of [ALPHA], previous experience with [A] and [B] may be required. This explanation is consistent with the previous findings on language recursion (Roeper, 2011), and lends further support to the alternative hypothesis that biological maturational factors are not the main factor limiting the ability to represent recursion, once the ability to represent iteration is available.

Visual recursion and grammar

A final hypothesis tested in our study was that grammar comprehension and visual recursion would be correlated. We found that the ability to represent recursion in the visual domain was correlated with grammar comprehension, and that this correlation was partially independent from general intelligence. However this effect was not specific to recursion, since grammar comprehension also correlated with embedded iteration. This suggests that grammar comprehension abilities were correlated with a more general ability to represent and process hierarchical structures generated iteratively, independently of whether these were recursive or not. This result is not completely surprising given that not all syntactic structures in TROG-D are recursive, although all are hierarchical.

We also assessed whether there was a more specific correlation between visual recursion and embedded clauses, but found again only a general association with both EIT and VRT. However, it is important to note that TROG-D only includes sentences with one level of embedding, e.g. relative clause (nominative): *Der Junge,* ***der*** *das Pferd jagt, ist dick* 'The boy, who is chasing the horse, is chubby'. Children may potentially use non-recursive representations for these kind of sentences (Roeper, 2011). Only a task focussed on sentences with several levels of recursive embedding would allow a direct comparison between visual recursion and syntactic recursion. Despite this limitation, it is interesting that performance on our novel visual tasks was correlated with grammar abilities, even when the effects of non-verbal intelligence were taken into account. These correlations could be explained by the existence of shared cognitive resources, independent from non-verbal intelligence, used for the processing of hierarchical structures in both language and visuo-spatial reasoning, or even by the effects of literacy (which are partially independent of intelligence) in the processing of hierarchical structures. Interestingly, while individual differences in intelligence predicted VRT and EIT scores both between and within grades, grammatical comprehension abilities accounted only for differences between grades. Again, this argues in favour of a general age-related maturational influencing the processing of hierarchical structures, occurring between second and fourth grade, which is partially independent from non-verbal intelligence. Furthermore, in our sample, grammar comprehension and non-verbal intelligence were only weakly correlated ($r = 0.25$, $p = 0.09$). Hence, this general maturation process in hierarchical processing cannot be explained solely by the increase of intelligence with age.

Future studies with a more comprehensive assessment of grammar (that includes recursion at several levels), and the inclusion of more cognitive tests (assessing cognitive control, attention, etc.) in the experimental procedure could potentially shed more light on a possible relationship between grammar and processing of complex visual structures.

Conclusion

In this study we assessed for the first time the ability of children to represent hierarchical self-similarity in an unambiguously non-linguistic domain. Consistently with previous findings on language (Miller et al., 1970) and visual-spatial research (Harrison & Stiles, 2009; Poirel et al., 2008), we found that the majority of fourth graders, but not second graders, were able to adequately process visual fractals generated using both recursive and iterative rules. This difference is partially accounted by distinct visual processing efficiency levels, but it is also predicted by grammar comprehension. Two crucial differences seem to emerge between the representation of recursive and iterative processes: 1) While the ability to acquire recursion seems to require previous learning of non-recursive representations, the opposite is not true; 2) Though recursive representations are harder to learn, once acquired, they seem to enhance the processing of hierarchical details.

In sum, we have found an interesting developmental path in the ability to represent hierarchy and recursion in the visuo-spatial domain. This path might be influenced by biological (maturational) factors, and by the exposure to particular kinds of stimuli. On the one hand, the re-organization of brain networks (Power et al., 2010), for instance, the myelination of the superior longitudinal fasciculus (occurring around the ages 7-8), seems to increase the efficiency of hierarchical processing (Friederici, 2009); on the other hand, the acquisition of certain hierarchical categories might depend on a gradual exposure, from concrete to abstract, where knowledge builds up incrementally (Dickinson, 1987; Roeper, 2011; Tomasello, 2003). Children may be born with a latent innate ability to detect and represent hierarchical structures (Berwick et al., 2011), but the development and precise tuning of this ability may require experience with enough examples to allow inductive generalizations (Dewar & Xu, 2010) and to allow acquisition of domain-specific constraints (Perfors, Tenenbaum, Griffiths, et al., 2011; Perfors, Tenenbaum, & Regier, 2011). Although the developmental time course of recursion in language and vision seem to obey similar constraints, this study does not provide direct evidence that the same cognitive machinery is used in both domains. However, it does provide a crucial method and important results, which offer a clear path for further investigation on the interface between language and visual aspects of cognition.

References

Alegre, M.A., & Gordon, P. (1996). Red rats eater exposes recursion in children's word formation. *Cognition, 60*, 65-82.

Alvarez, G.A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences, 15*(3), 122-131.

Badre, D. (2008). Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes. *Trends in Cognitive Sciences, 12*(5), 193-200. doi: 10.1016/j.tics.2008.02.004

Berwick, R.C., Pietroski, P., Yankama, B., & Chomsky, N. (2011). Poverty of the stimulus revisited. *Cognitive Science, 35*(7), 1207-1242.

Bishop, D.V.M. (2003). *The test for reception of grammar, version 2 (trog-2)*. London: Pearson Assessment.

Bouvet, L., Rousset, S., Valdois, S., & Donnadieu, S. (2011). Global precedence effect in audition and vision: Evidence for similar cognitive styles across modalities. *Acta Psychologica, 138*(2), 329-335. doi: 10.1016/j.actpsy.2011.08.004

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.

Chomsky, N. (2010). Some simple evo devo theses: How true might they be for language? In Richard K Larson, Viviane M Déprez & Hiroko Yamakido (Eds.), *The evolution of human language* (pp. 45-62). New York: Cambridge University Press.

Dąbrowska, E., Rowland, C., & Theakston, A. (2009). The acquisition of questions with long-distance dependencies. *Cognitive Linguistics, 20*(3). doi: 10.1515/cogl.2009.025

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.

Dickinson, S. (1987). Recursion in development: Support for a biological model of language. *Language and Speech, 30*(3), 239-249.

Fagot, J., & Tomonaga, M. (1999). Global and local processing in humans (homo sapiens) and chimpanzees (pan troglodytes): Use of a visual search task with compound stimuli. *Journal of Comparative Psychology, 113*(1), 3-12. doi: 10.1037/0735-7036.113.1.3

Fitch, W.T. (2010). Three meanings of "recursion": Key distinctions for biolinguistics. In Richard K Larson, Viviane M Déprez & Hiroko Yamakido (Eds.), *The evolution of human language* (pp. 73-90). New York: Cambridge University Press.

Fitch, W.T., Hauser, M., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition, 97*(2), 179-210. doi: 10.1016/j.cognition.2005.02.005

Fox, A.V. (2007). *Trog-d: Test zur überprüfung des grammatikverständnisses* (2 ed.). Idstein: Schultz-Kirchner.

Friederici, A.D. (2009). Pathways to language: Fiber tracts in the human brain. *Trends in Cognitive Sciences, 13*(4), 175-181.

Hanley, J.A. (2003). Statistical analysis of correlated data using generalized estimating equations: An orientation. *American Journal of Epidemiology, 157*(4), 364-375.

Harrison, T.B., & Stiles, J. (2009). Hierarchical forms processing in adults and children. *Journal of Experimental Child Psychology, 103*(2), 222-240.

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.

Hopkins, W., & Washburn, D. (2002). Matching visual stimuli on the basis of global and local features by chimpanzees (pan troglodytes) and rhesus monkeys (macaca mulatta). *Animal Cognition, 5*(1), 27-31. doi: 10.1007/s10071-001-0121-8

Koike, H. , & Yoshihara, H. (1993). *Fractal approaches for visualization of huge hiearchies.* Paper presented at the Proceedings of the 1993 IEEE Symposium on Visual Languages.

Laury, R., & Ono, T. (2010). Recursion in conversation: What speakers of finnish and japanaese know how to do. In Harry van der Hulst (Ed.), *Recursion and human language*: de Gruyter Mouton.

Mandelbrot, B. (1977). *The fractal geometry of nature*: Freeman.

Martins, M.D. (2012). Distinctive signatures of recursion. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1598), 2055-2064.

Martins, M.D., & Fitch, W.T. (2012). *Empirical approaches to recursion.* Paper presented at the The Evolution of Language - Proceedings of the 9th International conference, Kyoto.

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.

Moses, P., Roe, K., Buxton, R.B., Wong, E.C., Frank, L.R., & Stiles, J. (2002). Functional mri of global and local processing in children. *NeuroImage, 16*(2), 415-424. doi: http://dx.doi.org/10.1006/nimg.2002.1064

Nardini, M., Jones, P., Bedford, R., & Braddick, O. (2008). Development of cue integration in human navigation. *Current Biology, 18*(9), 689-693. doi: 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.

Perfors, A., Tenenbaum, J., Gibson, E., & Regier, T. (2010). How recursive is language? A bayesian exploration. In Harry van der Hulst (Ed.), *Recursion and human language* (pp. 159-175). Berlin/New York: de Gruyter Mouton.

Perfors, A., Tenenbaum, J.B., Griffiths, T.L., & Xu, F. (2011). A tutorial introduction to bayesian models of cognitive development. *Cognition, 120*(3), 302-321.

Perfors, A., Tenenbaum, J.B., & Regier, T. (2011). The learnability of abstract syntactic principles. *Cognition, 118*(3), 306-338.

Poirel, N. , Mellet, E. , Houdé, O., & Pineau, A. (2008). First came the trees, then the forest: Developmental changes during childhood in the processing of visual local-global patterns according to the meaningfulness of the stimuli. *Developmental Psychology, 44*(1), 245-253.

Power, J.D., Fair, D.A., Schlaggar, B.L., & Petersen, S.E. (2010). The development of human functional brain networks. *Neuron, 67*(5), 735-748.

Raven, J.C. , Raven, J., & Court, J.H. (2010). *Coloured progressive matrices (cpm)* (German Edition ed.). Frankfurt: Pearson.

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

Roeper, T. (2007). *The prism of grammar: How child language illuminates humanism*. Cambridge, MA.: MIT Press.

Roeper, T. (2009). *The minimalist microscope: How and where interface principles guide acquisition.* Paper presented at the XX Boston University Conference on Language Development.

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

Spinozzi, G., De Lillo, C., & Truppa, V. (2003). Global and local processing of hierarchical visual stimuli in tufted capuchin monkeys (cebus apella). *Journal of Comparative Psychology, 117*(1), 15-23. doi: 10.1037/0735-7036.117.1.15

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard: Harvard University Press.

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. doi: 10.1016/j.jphysparis.2006.03.014

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. doi: 10.1098/rstb.2002.1257

Appendix A

VRT Stimulus generation

Stimuli were based on the properties of fractal geometry. Visual fractals are structures that exhibit self-similarity (Fig. A1), meaning that the organization of elements within smaller scales is similar to the organization of elements within larger scales. Visual fractals are hierarchical in the sense that each element within a higher level (larger scale) will determine the position of many elements within a lower level (smaller scale). In the example of Fig. A2, the topmost (and larger) square (called *the initiator*) determines the position of four smaller squares whose locations follow a certain rule of transformation (called *the generator*). This rule of transformation, which could be described as "add a smaller square below the vertices of each square", can also be applied to the newly embedded squares, each determining the position of four additional smaller squares. This process is theoretically infinite and allows the production of hierarchies with unbounded depth.

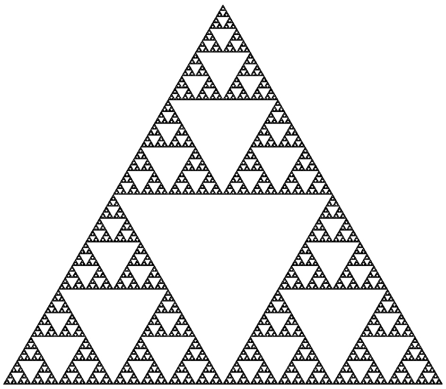


Figure A1. *Example of a Fractal Structure – a Sierpinski Gasket - Exhibiting Self-similarity*

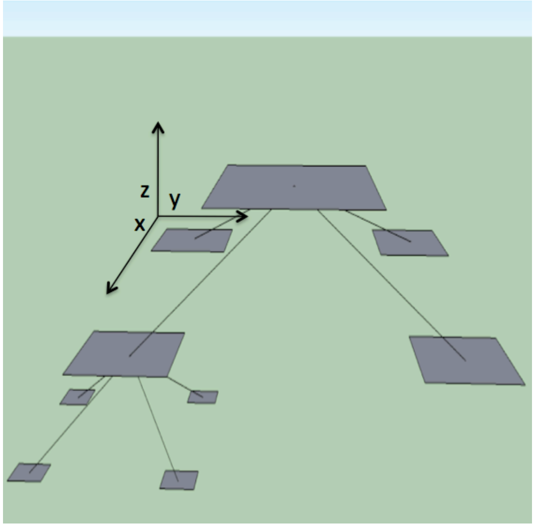


Figure A2. *Visual Fractals can be perceived as Visuo-spatial Hierarchies:* The larger squares (upper in the z-axis) are dominant over the smaller squares, in the sense that they determine their spatial coordinates in the xy-axis according to a fixed rule, similar for all hierarchical levels.

The principle underlying the generation of the stimuli was the following: There was a pool of 7 geometrical shapes (rectangle, square, circle, star, triangle, pentagon and hexagon), which were used as *initiators*. Then, different kinds of recursive embedding rules (*generators*) were applied over these shapes in order to generate hierarchical structures. Four iterative steps were generated for each fractal (Fig. 2, main text). There were 3 main categories of *generators* which determined the visual complexity of the structures that resulted from their application. Here 'visual complexity' refers to the number of elements being added to the structure, dependent on the same previously existing element, "dependent" in the sense that the spatial coordinates of the dominant element determine the spatial coordinates of the subordinate elements. For instance, in Fig. A1, the 'visual complexity' is 4, since 4 squares are added dependent on each initiator square. In this study, the 'visual complexity' utilized was 3, 4 or 5 (Fig.2, main text).

Foils. Together with the first four iterations of a fractal, foil structures were generated. This foil structure corresponded to an "incorrect" fourth iteration, and was generated by applying a different *generator* for the fourth iteration, from the one used to generate the previous three iterations. There were 3 categories of foils, relative to the kind of process used in their generation (Fig.2, main text). These processes were the following: i) Misplacement of a few elements within the whole structure ('Odd constituent foil'); ii) Utilization of a novel positional scheme for all new added elements, different from the scheme used for the previous iterations ('Positional error foil'); iii) Repetition of the image depicted as the 3rd iteration ('Repetition foil'). The rationale of using different kind of foils was to prevent the development of simple heuristic strategies based on the comparison of the 'correct' and 'incorrect' fourth iterations, strategies which could be unrelated to the recursive rule itself. For instance, if participants correctly reject the 'repetition foil', then a simple strategy to determine the fourth iteration based on a pure similarity assessment is unlikely, since the 'repetition foil' is identical to the third iteration. Likewise, the consistent rejection of the 'odd constituent foils' rules out the possibility that participants attend only to small portions of the 'choice' images, since most regions within these foils are identical to the correct fourth iteration.

EIT

EIT stimulus categories were similar to VRT, and are depicted in figure A3.

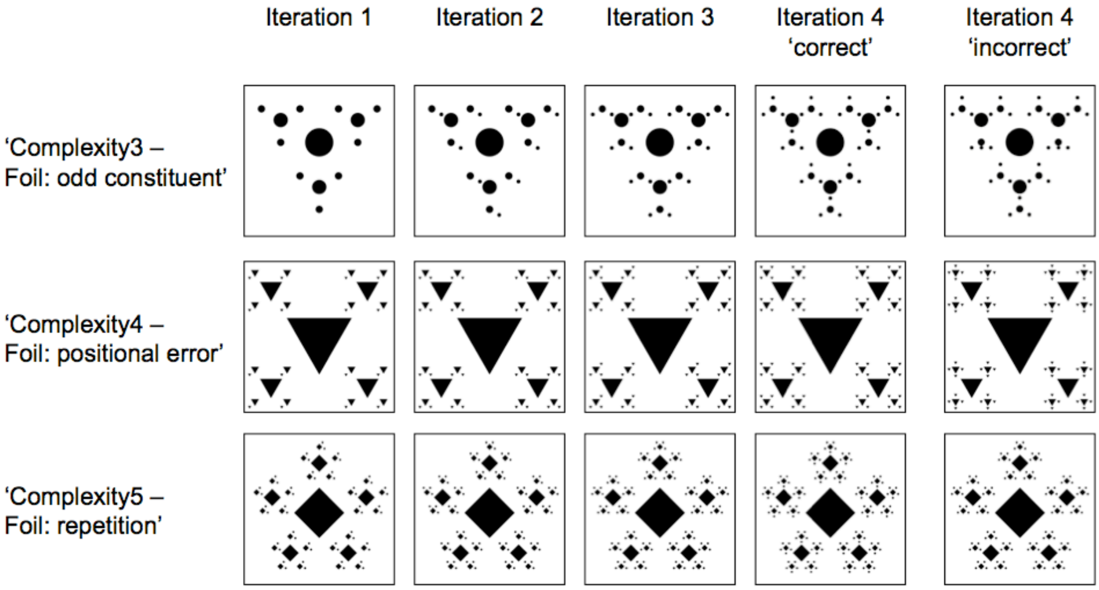


Figure A3. *Examples of Visual Structures Used in the Embedded Iteration Task:* Similarly to VRT, there were different categories of foils (see text for details) – 'Odd constituent', 'Positional error' and 'Repetition' – and different categories of 'visual complexity' – 3, 4 and 5. Some of the combinations of these categories are depicted.

Appendix B

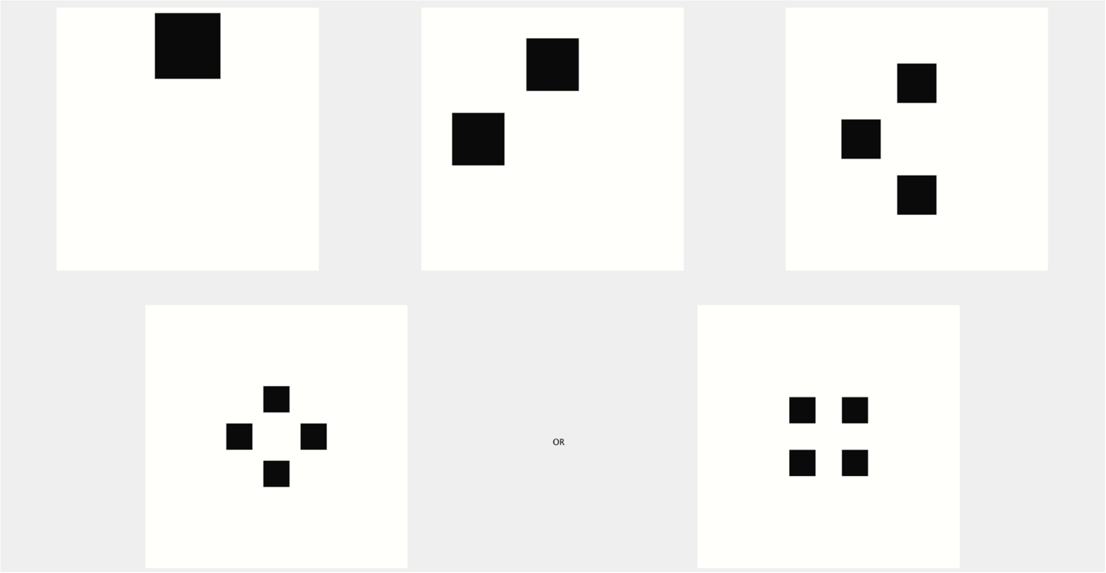


Figure Appendix B1. *Example of a Training Trial which Depicted a Non-hierarchical Iterative Process*

Appendix C

VRT visual strategies – Regression analysis

In order to compare performance across different foil categories, we ran a GEE model with ‘grade’ as a between-subjects factor, and ‘foil category’ as a within-subjects factor. We found no performance differences between different ‘foil categories’ (*Wald* $\chi 2 = 0.52$, $p = 0.77$) and no interaction between ‘grade’ and ‘foil category’ (*Wald* $\chi 2 = 0.16$, $p = 0.92$; Fig. C1). These results suggest that no simple heuristic strategy was utilized to solve the task.

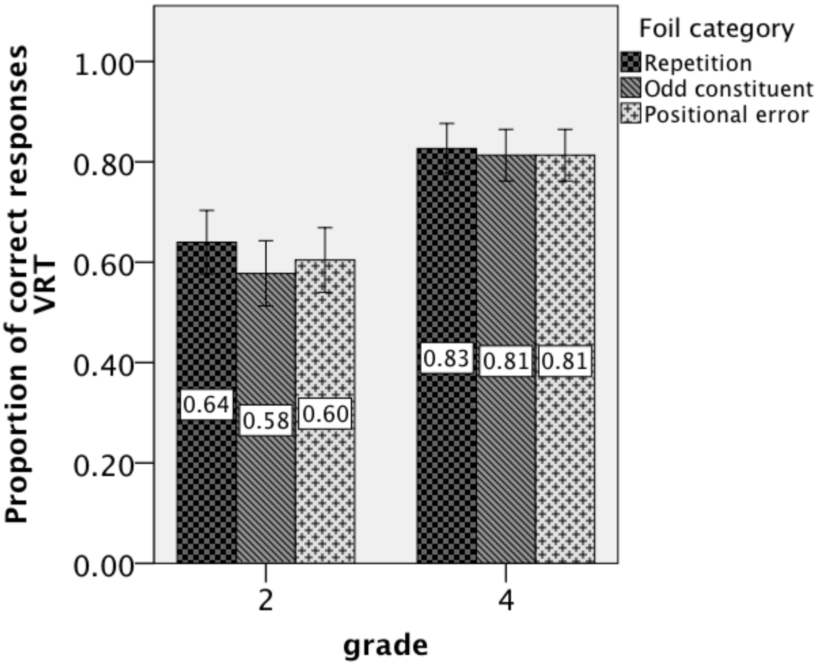


Figure C1. *VRT Performance in Trials Using Different Foil Categories (see text for details), both in Second and Fourth Grades*

EIT visual strategies – Regression analysis

In order to compare performance across different foil categories, we ran a GEE model with ‘grade’ as a between-subjects factor, and ‘foil category’ as a within-subjects factor. As to performance across different foil categories, we found a main effect of ‘foil categories’ (*Wald* $\chi 2 = 45.0$, $p < 0.001$; Fig. C2). Specifically, performance in the foil category ‘Odd constituent’ was significantly lower than in the categories ‘Repetition’ (*EMM difference* = 0.17, $p < 0.001$) and ‘Positional error’ (*EMM difference* = 0.23, $p < 0.001$). Furthermore there was an interaction between ‘grade’ and ‘foil category’ (*Wald* $\chi 2$ = 5.68, *p* = 0.06). Pairwise comparisons reveal that: 1) In the second grade, the proportion of correct answers in the ‘Odd constituent’ category was significantly lower than in the category ‘Positional error’ (*EMM difference* = 0.17, *p* = 0.007); and 2) Fourth graders were significantly worse in the foil category ‘Odd constituent’ than in ‘Repetition’ (*EMM difference* = 0.21, *p* = 0.002), and ‘Postional error’ (*EMM difference* = 0.26, *p* < 0.001). All p-values were corrected with sequential Bonferroni. No other significant differences were found (*p* > 0.1). These results suggest that children of both grades found the detection of ‘Odd constituent’ errors difficult, although they were more easily able to perceive both the ‘Positional’ and iterative (‘Number’) properties of the EIT stimuli.

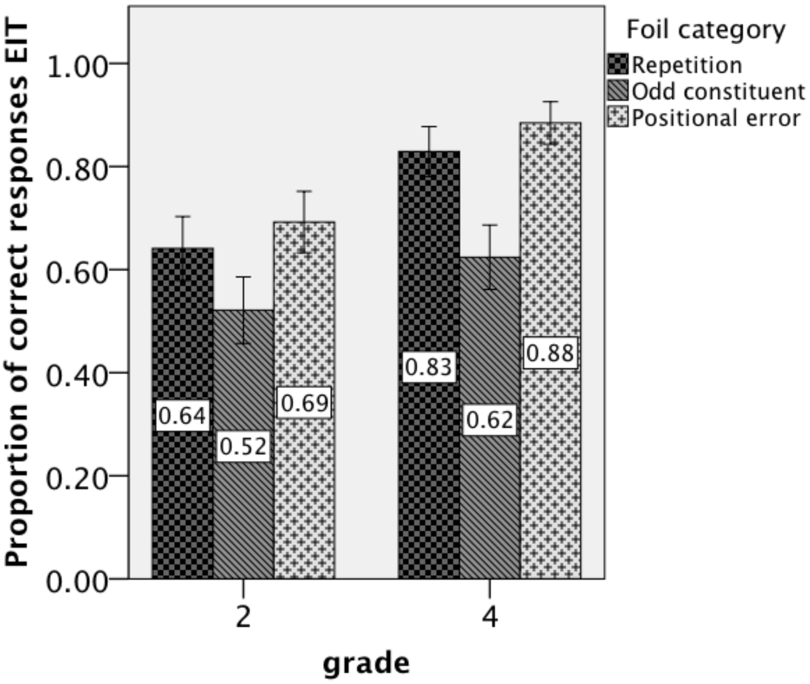


Figure C2. *EIT Performance in Trials Using Different Foil Categories (see text for details), both in Second and Fourth Grades*

Appendix D

VRT visual complexity - interaction analyses

Overall results are depicted in Fig. D1. We found that visual complexity had a significant main effect on VRT performance (*Wald χ2* = 6.5, *p* = 0.039). Specifically, the proportion of correct answers in the category 'complexity4' was higher than in the category 'complexity5' (*estimated marginal mean (EMM) difference* = 0.06, *p* = 0.026). Furthermore, there was a trend towards an interaction between 'grade' and 'visual complexity' (*Wald χ2* = 5.38, *p* = 0.068). Pairwise comparisons revealed that second graders were significantly better in the simpler 'complexity 3' trials than in 'complexity 5' trials (*EMM difference* = 0.10, *p* = 0.037). All p-values were corrected using sequential Bonferroni correction. No other significant differences were found (*p* > 0.1). These results suggest that higher visual complexity may hinder the ability of children to perform adequately in VRT, and this effect may be especially pronounced in the second graders.

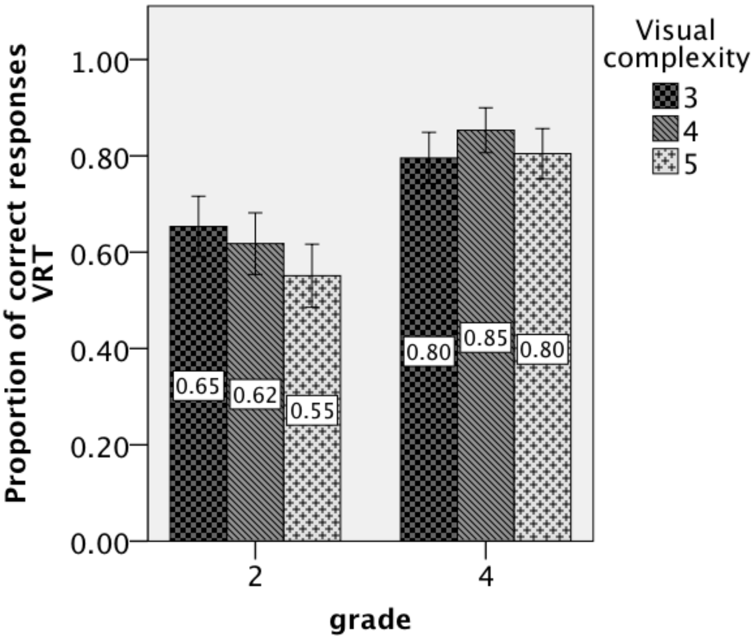


Figure D1. *VRT Performance in Trials Displaying Different Levels of Visual Complexity (3 refers to lower complexity and 5 to higher complexity, see text for details), both in Second and Fourth Grades*

EIT visual complexity - interaction analyses

Overall results are depicted in Fig. D2. We found that visual complexity had a significant main effect on performance (*Wald* $\chi 2$ = 12.6, *p* = 0.002). Pairwise comparisons revealed that: 1) The proportion of correct answers in the category ‘complexity3’ was higher than in the categories ‘complexity4’ (*EMM difference* = 0.06, *p* = 0.012) and ‘complexity5’ (*EMM difference* = 0.07, *p* = 0.06). Furthermore, there was a significant interaction between ‘grade’ and ‘visual complexity’ (*Wald* $\chi 2$ = 6.31, *p* = 0.039). Pairwise comparisons revealed that: 1) Second graders were significantly better in the simpler ‘complexity 3’ trials than in ‘complexity 5’ trials (*EMM difference* = 0.09, *p* = 0.028); and 2) Fourth graders were significantly better in the simpler ‘complexity 3’ trials than in ‘complexity 4’ trials (*EMM difference* = 0.09, *p* = 0.006). All p-values were corrected using sequential Bonferroni correction. All p-values were corrected using sequential Bonferroni correction. No other significant differences were found (*p* > 0.1). These results suggest that visual complexity also plays a role in the ability to perform adequately in EIT, with fewer constituents easier to process.

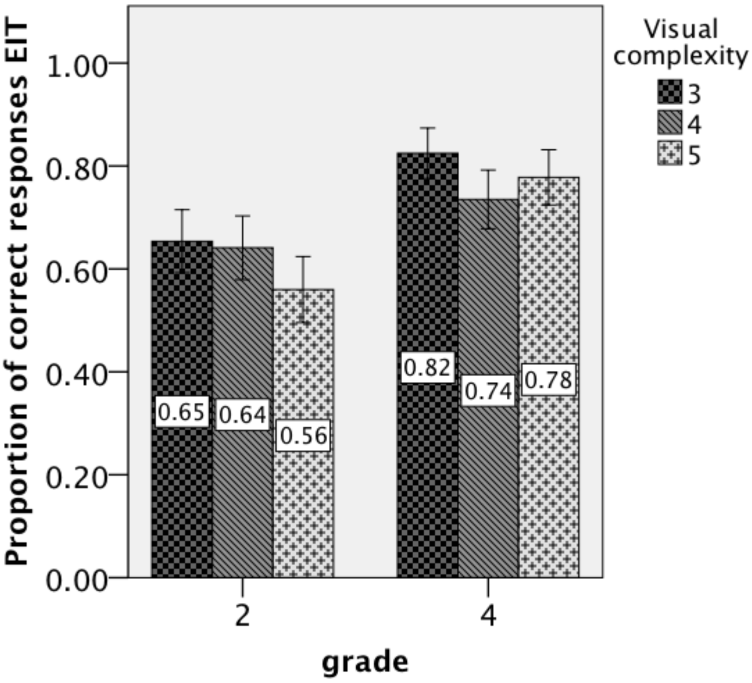


Figure D2. *EIT performance in Trials Displaying Different Levels of Visual Complexity (3 refers to lower complexity and 5 to higher complexity, see text for details), both in Second and Fourth Grades.*

Appendix E

VRT vs. EIT and effects of task order – interaction analyses

We found that when the effect of task-order was included in the model, the interaction between ‘task’ and ‘grade’ neared significance (*Wald* $\chi 2 = 3.6$, $p = 0.057$). Since there were no differences between VRT and EIT within each grade, this interaction may reflect a higher average difference between second and fourth grade for VRT (*EMM difference* = 0.24) than for EIT (*EMM difference* = 0.17).

However, more importantly, there was a significant 3-way interaction between ‘task’, ‘position’ and ‘grade’. The analysis of pair-wise comparisons reveal a finer structure in the data: 1) Fourth graders had a higher proportion of correct answers than second graders in VRT, but ONLY when VRT was performed in the second position in the procedure (*EMM difference* = 0.27, $p < 0.001$); 2) Fourth graders had a higher proportion of correct answers than second graders in EIT, but ONLY when EIT was performed in the first position in the procedure (*EMM difference* = 0.21, $p < 0.001$); 3) Fourth graders had a higher proportion of correct answers in VRT than in EIT when both tasks were performed in the second position in the procedure (*EMM difference* = 0.21, p = 0.019); 4) VRT accuracy was higher when performed in the second position than when performed in the first position, but only within the fourth grade group (*EMM difference* = 0.24, $p = 0.005$). All p-values were corrected with sequential Bonferroni.

Taken together, these results suggest that the order of the task in the procedure had a strong influence on task performance, especially for VRT and for the fourth grade group. However, this effect of task-order was not due to a practice effect during the experiment, since EIT performance decreased when this task was performed in the second position of the procedure.

7. Fractal Image Perception provides novel insights into hierarchical cognition

Martins M, Fishmeister F, Puig Waldmüller E, Oh J, Geissler A, Fitch WT, Beisteiner R (2014). Fractal Image Perception provides Novel Insights into Hierarchical Cognition. *NeuroImage.* doi: 10.1016/j.neuroimage.2014.03.064.

Fractal Image Perception provides Novel Insights into Hierarchical Cognition

Martins M.J. (MD, MSc)[1,2,†], Fischmeister F.P. (PhD)[3,†], Puig-Waldmueller E. (PhD)[1], Oh J. (MSc)[1], Geißler A. (PhD)[3], Robinson S. (PhD)[4], Fitch W.T. (PhD)[1], Beisteiner R. (MD, PhD)[3]

[1]Department of Cognitive Biology, University of Vienna, Althanstrasse 14, A-1090 Vienna, Austria.

[2]Language Research Laboratory, Lisbon Faculty of Medicine, Av. Professor Egas Moniz, 1649-028 Lisboa, Portugal.

[3]Medical University of Vienna, Department of Neurology, High-Field MR Center of Excellence, Währinger Gürtel 18-20, A-1090 Vienna, Austria.

[4]Medical University of Vienna, Department of Radiology, High-Field MR Center of Excellence, Währinger Gürtel 18-20, A-1090 Vienna, Austria.

[†]These authors contributed equally to this work

ABSTRACT

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

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

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

Keywords: Visual processing, fMRI, Hierarchy, Recursive embedding, Parieto-medial temporal pathway

Introduction

The ability to represent and generate complex hierarchical structures is one of the hallmarks of human cognition. In many domains, including language, music, problem solving, action-sequencing and spatial navigation, humans organize basic elements into higher-order groupings and structures [1-6]. This ability to encode the relationship between items (words, people, etc.) and the broader structures in which these items are embedded (sentences, corporations, etc) affords flexibility to human behavior. For example, in action sequencing, humans are able to change, add or adapt certain basic movements to particular contexts, while keeping the overall structure (and goals) of canonical motor procedures intact [5]. Typical examples of these actions-in-context are 'grinding the beans' or 're-filling the water container' in the process of making coffee [7]. Individuals can evaluate the need for these actions and omit them if they are unnecessary without impairing the overall procedure of making coffee [8]. This ability is different from simple action sequencing, and seems very limited in non-human animals [9].

A promising method to represent complex hierarchical structures - realized in nature and attractive for experimental research – is the use of recursive embedding processes [10, 11]. Recursive embedding refers to the incorporation of a structure inside another structure of the same sort, and it allows the generation of hierarchies with infinite depth using very simple rules. We can add several new elements to a certain hierarchical level using within-level transformation rules (Fig. 1A), but it is only possible to generate multiple hierarchical levels with a single rule if this rule involves recursive embedding (Fig.1B). When used in association with other rule-based processes, recursive embedding allows the generation of hierarchies that are deep, structurally rich and perceived as attractive. Some examples are the fractal Mandelbrot images or fractal structures in nature such as tree branches, algae, the flower of the *Brassica oleracea*, snail shells and coastlines. These structures can be extended or sub-divided indefinitely whilst visual and structural similarity is retained at all scales. These kinds of structures contrast with others with simpler modes of organization such as grass or crop fields, which like bead necklaces, are formed by adding several items to a group at fixed hierarchical levels.

Here we investigate the ability to recognize well-formed visuo-spatial hierarchical structures, based on the application of rules that either operate transformations within a hierarchical level, or rules which generate new self-similar hierarchical levels (Fig.

1). For simplicity, we simply use the expression ‘recursive’ or ’recursion’ to refer to ‘recursive embedding’.

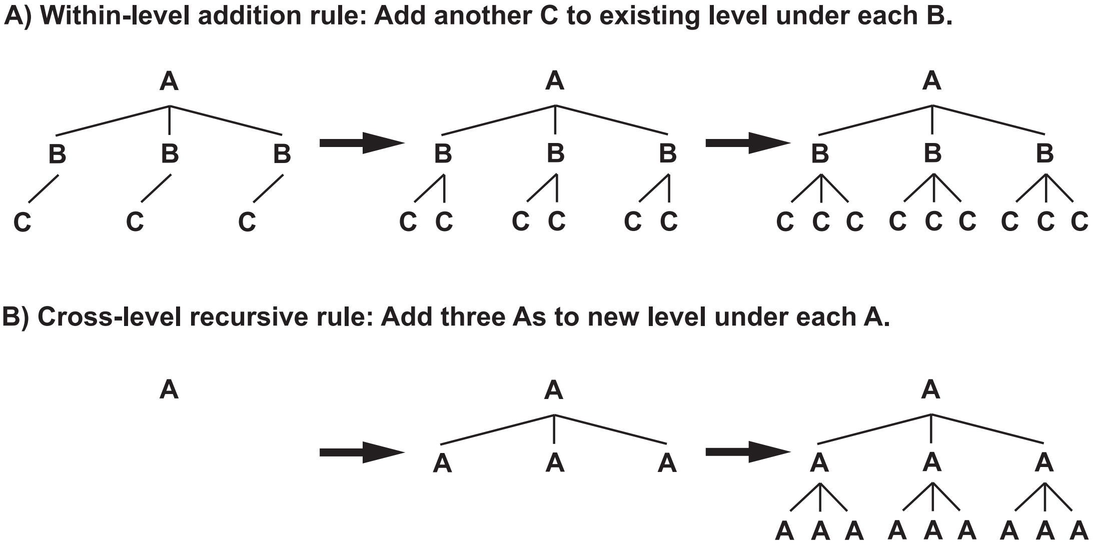


Fig.1. Examples of processes that add elements to hierarchies. These processes can either generate new hierarchical levels (B) or simply add elements to pre-existing levels (A). Only recursive embedding (B) can add multiple hierarchical levels using a single rule.

The processing of visuo-spatial stimuli is often described as occurring in parallel in two different systems - the ventral stream and the dorsal stream [12, 13]. The ventral stream, an occipito-temporal network, seems to process object quality or semantic information, with more abstract categories represented in more anterior portions of the temporal lobe [14]. The dorsal stream, an occipito-parietal network, has classically been described as processing spatial information only. Recently, however, this classical view of the dorsal stream has been updated [13]. While projections from the parietal cortex to the prefrontal cortex seem to be important for spatial working memory and visually-guided action, a third system, called the parietal-medial temporal pathway (PMT) appears to be necessary to integrate spatial and semantic information [13]. The PMT pathway connects the dorsal stream with the medial temporal cortex (hippocampus and parahippocampus), through the posterior cingulate (PCC) and retrosplenial cortices (RSC) [13, 15]. This pathway appears to be crucial for the retrieval of landmark information during spatial navigation and for the integration of objects in contextual frames (e.g. a mug in a a date in a coffee shop)

[16-20] (Fig.2). We therefore hypothesize that the PMT may play a specific role in the representation of principles that allow the recognition and generation of well-formed hierarchical embeddings in the visuo-spatial domain.

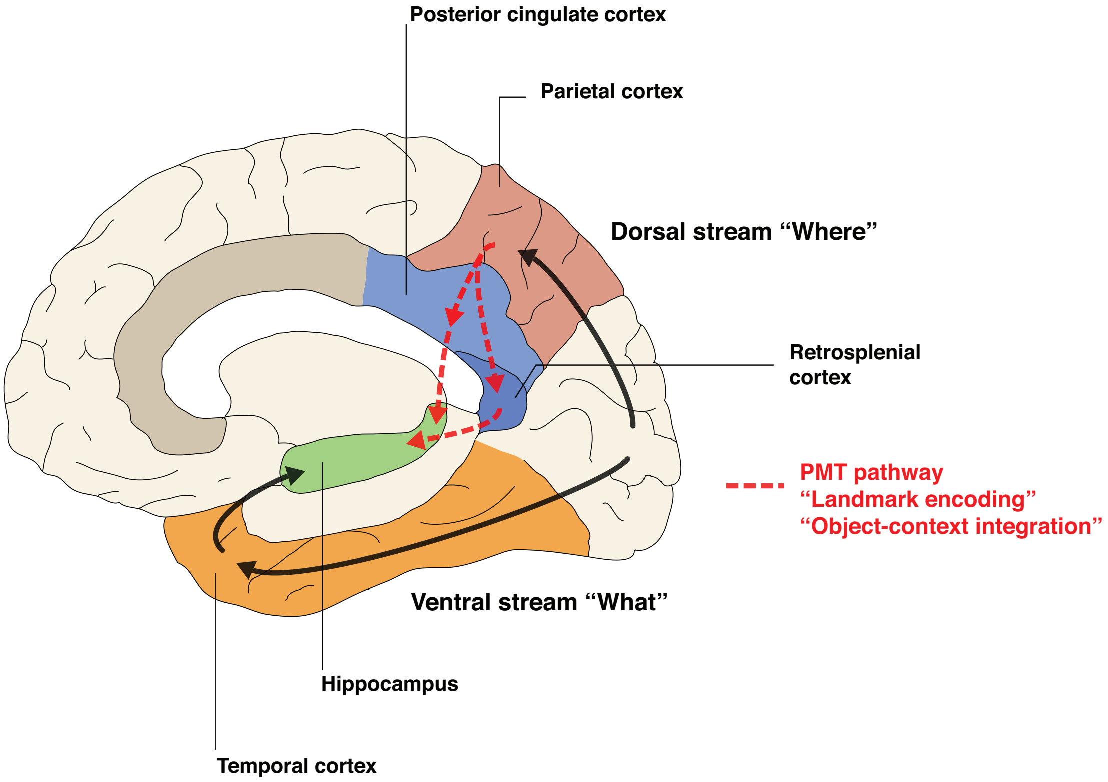


Fig.2. Neural pathways involved in visuo-spatial processing. The dorsal stream, which includes the parietal cortex and its projections to the frontal cortex, is involved in the processing of spatial information. The ventral stream, which includes the inferior and lateral temporal cortex and their projections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

Based on the principles depicted in Fig.1, we developed two tasks to investigate the cognitive processes involved in the representation of visuo-spatial hierarchies: The Visual Recursion Task (VRT) and the Embedded Iteration Task (EIT). In both tasks participants are exposed to generative processes for a certain number of iterative steps and then asked to make inferences about further iterations. This means that in both tasks participants are asked to extract simple rules from the first iterations which can then be applied to predict further transformations. In VRT, each iterative step generates a new hierarchical level according to one particular spatial rule isomorphic to the rule displayed in previous levels of the hierarchy. The brain requires only one simple rule to be able to generate large self-similar structures (fractals) with an unlimited number of levels. In EIT, new elements are embedded iteratively within a fixed hierarchical level, according to a spatial rule but without generating new levels. It is important to clarify that both tasks are iterative (i.e. a certain rule is applied a given number of times) and both may generate hierarchies of similar complexity (see Fig.1 and Fig. 3).

Our previous research with these tasks [10, 21] suggests that, in comparison with EIT, performance in VRT is more strongly associated with abstract reasoning and less correlated with specific visuo-spatial cognitive abilities. In the current study, we investigated the neural bases involved in the representation of visuo-spatial hierarchies by comparing the brain circuits active during VRT and EIT. As a control task we introduced a ‘similarity task’ (positional similarity visual task – PSVT), in which participants were asked to match a target visuo-spatial hierarchy with a set of alternatives. The setup and images displayed were closely matched for all three tasks. As indicated above, our primary hypothesis was that the brain uses different resources for processing identical hierarchical structures depending on whether it applies a “fractal” or a “non-fractal” cognitive strategy.

Material and Methods

Participants

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

fMRI) after which they received feedback. Participants were paid 30 Euros for their participation. The overall procedure comprised one hour of practice plus cognitive testing and approximately one and a half hours of fMRI scanning.

Task

Modified VRT and EIT tasks, described in [10, 21], were used. While EIT requires the representation of iterative processes occurring within a hierarchical level, VRT requires the representation of iterative processes generating new hierarchical levels (Fig.1). For this study, we devised an additional Positional Similarity Visual Task (PSVT) to investigate the effects of observing visual fractals without rule-based reasoning. In the latter, participants attended to a set of three random images and were then asked to chose which of two new items was *identical* to one of the previous three (Fig.3).

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

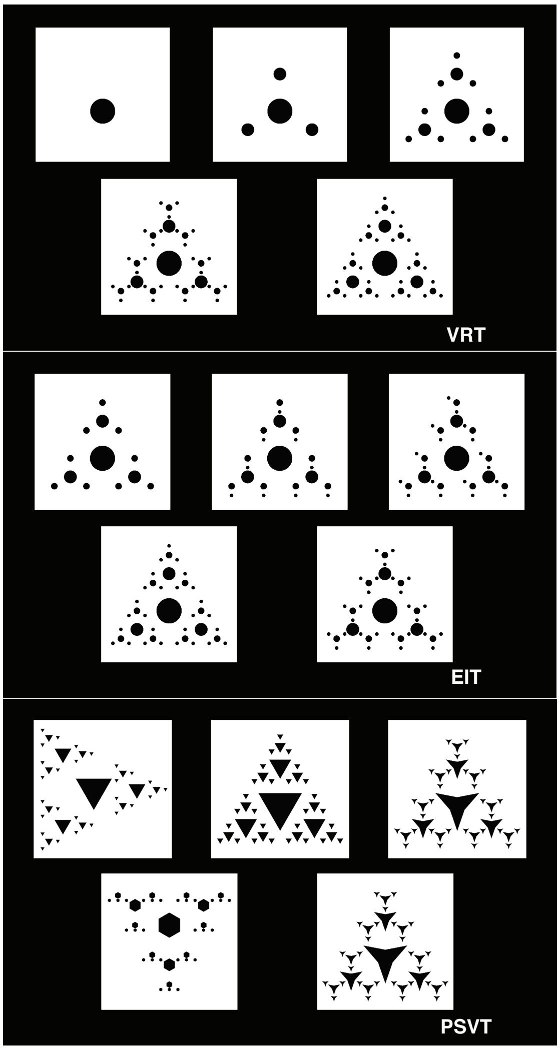


Fig.3. Examples of VRT (Visual Recursion Task) stimuli (top), EIT (Embedded Iteration Task) stimuli (middle) and PSVT (Positional Similarity Visual Task) stimuli (bottom). In the "rule acquisition" phase (see methods for details), the first three iterations of a process were presented in the top half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

Each trial comprised two main phases (Fig.4) - the rule acquisition phase, and the rule application phase. Before the rule acquisition phase, at the beginning of each trial, a white letter was presented on a black background in the center of the screen for a duration which ranged between 1000 and 1750 ms. This letter indicated the task of the trial: “R” for VRT, “I” for EIT, and “S” for PSVT. Then, in the rule acquisition phase, three images, corresponding to the first three iterations of either a within-level or recursive process were presented simultaneously in the top half of the screen. In the case of the similarity task these were three different images selected quasi-randomly from the large pool of fractal images. This phase had a fixed duration of 3 seconds. Between the rule acquisition and rule application phases, a white crosshair was presented in the center of the screen for a duration which ranged between 1000 and 3000 ms. Finally, in the rule application phase, two additional images were presented in the bottom half of the screen, simultaneously and side-by-side. One of these corresponded to the correct fourth iteration of the previous iterative process and the other was a foil. In the case of the similarity task, the correct image was identical to one of the previously presented three images (Fig.3). In this rule application phase participants were asked to choose the image they considered correct by pressing either the left or right button with the thumb of the left or right hand. No visual or auditory feedback was provided. The maximum duration of this phase was 6 seconds. The inter-trial interval (ITI) ranged from 500 ms to 14000 ms and during this period participants were exposed to a black screen. The position of the correct and foil images (LEFT or RIGHT) was random and counterbalanced. To control for luminance effects, all stimuli had the same number of black and white pixels, both globally and for each trial phase. For more details on the generation of the stimuli, see the Supplementary Methods.

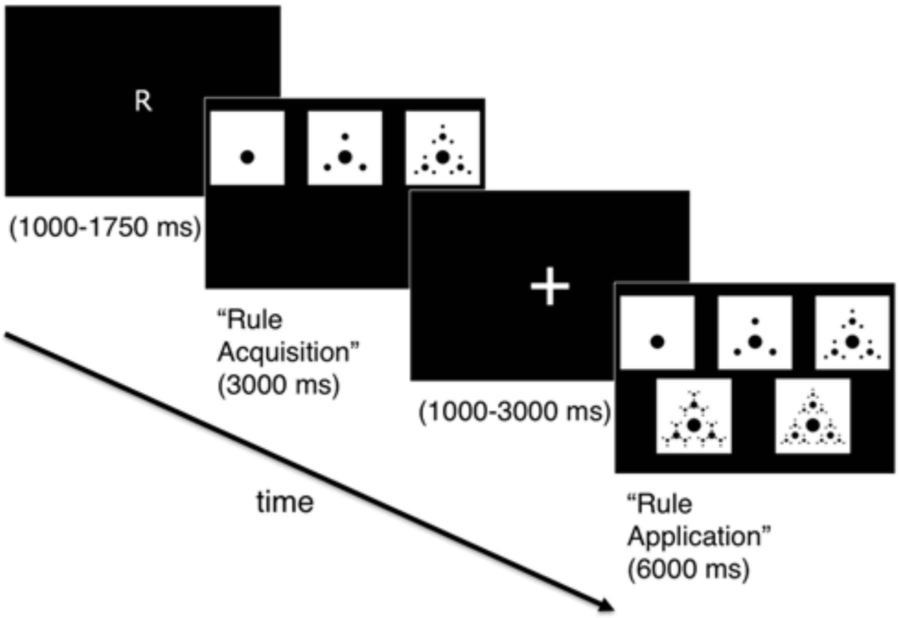


Fig.4. Trial structure: At the beginning of each trial a letter was displayed indicating the stimulus category to be presented (‘R’ for VRT, ‘I’ for EIT and ‘S’ for PSVT). fMRI data were acquired in the ‘Rule Application’ phase.

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

Data Acquisition

Data acquisition was performed with a 3 Tesla TIM Trio system (Siemens, Erlangen, Germany) using a 32-channel Siemens head coil. Functional magnetic resonance images (fMRI) were acquired using an optimized 2D single-shot echo planar imaging (EPI) sequence which included online EPI distortion correction with PSF mapping [22]. 350 EPI volumes per session were acquired with a square FOV of 220 mm, an in-plane matrix size of 128×128, with 36 slices of 2.7 mm thickness and 20% gap (i.e.

2.3 mm x 2.3 mm x 2.7 mm voxel size) aligned parallel to the AC-PC plane, a repetition time (TR) of 2000 ms, echo time (TE) 32 ms, and a flip angle of 73°. For anatomical registration, high-resolution T1-weighted MR images were acquired using a 3D MPRAGE sequence (TE = 3.02 ms, TR = 2190 ms, inversion time [TI] = 1300 ms) with a matrix size of 250 x 250 x 256, with isometric voxels with a nominal side length of 0.9 mm, flip angle of 9° and GRAPPA acceleration factor 2.

Data Preprocessing

Image preprocessing and statistical analysis at the individual and group level were performed using SPM 8 (http://www.fil.ion.ucl.ac.uk/spm/). Data were first slice-time and then motion corrected. These corrected data were then spatially normalized using New Segment (SPM manual, FIL Group) and finally smoothed using a 5 mm full-width-at-half-maximum Gaussian filter. For single-subject analyses, evoked hemodynamic responses for the different event types were modeled within a general linear model using delta functions corresponding to the stimulus presentation length convolved with a canonical hemodynamic response function. This way, the model captured differences in reaction time. To this design matrix we added 24 nuisance regressors of no interest, corresponding to the motion realignment parameters and their Volterra expansion [23], to regress out residual motion artifacts. In addition, a 356 s cutoff high-pass filter was applied to account for low-frequency drifts and signal fluctuations. Block regressors were used to correct for session-related mean and scaling effects (added as confounds). Responses corresponding to the rule application phase of the three stimuli types were then summarized across the four sessions and entered into a second-level GLM.

Statistical Analysis

On the group level, a repeated-measures GLM with partitioned error variances (rm-GLMFlex: between- and within-subject error terms are modeled separately) was used to model activity during the application phase. This one-way rm-GLMFlex (with each task's application phase being one level) allowed us to identify hemodynamic responses solely related to the tasks of interest by constructing planned contrasts to answer the different research questions within one model. Rule-based related activation was obtained by contrasting VRT and EIT with the control condition task (PSVT). The differences between recursion-related processes and those resulting from embedded iteration were assessed by directly comparing VRT and EIT (implicitly a comparison of VRT-PSVT versus EIT-PSVT) within the rm-GLMFlex model. Additionally, to test for regions commonly activated during the application phase in the VRT and the EIT, a conjunction analysis across the contrasts VRT-PSVT and EIT-PSVT using the 'conjunction null' hypothesis was performed [24]. All comparisons were masked with the main effect of the one-way rm-GLMFlex and subsequently thresholded at a voxel-wise FDR-adjusted $P < 0.05$ with a 10-voxel extent threshold.

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

Neuro-cognitive Battery

We applied a brief neuropsychological battery to screen our participants for possible cognitive impairments. This battery included computerized versions of Digit Span backwards (DSPAN, a verbal working memory task), Corsi block tapping backwards (CORSI, a spatial working memory task), Tower of Hanoi (ToH, a recursive planning in action sequencing task) [25] (retrieved from http://pebl.sf.net/battery.html) and a paper-and-pencil version of the progressive matrices of RAVEN (a test of non-verbal intelligence). We recorded the maximum number of elements correctly reproduced in DPSAN and CORSI, the maximum length (viz. number of steps) of ToH problems that participants were able to complete without errors, and the number of correct answers in RAVEN.

Results

Behavioral Results

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

In order to prevent participants from using simple heuristic strategies we included different foil categories ("ODD foil" and "POSITIONAL foil") in both VRT and EIT (see Supplementary Methods). Participants performed adequately (> 90%) in all foil categories (see Supplementary Table S1).

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

Rule-based Iterative Processes (within and across hierarchical levels) versus Similarity Assessment

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

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

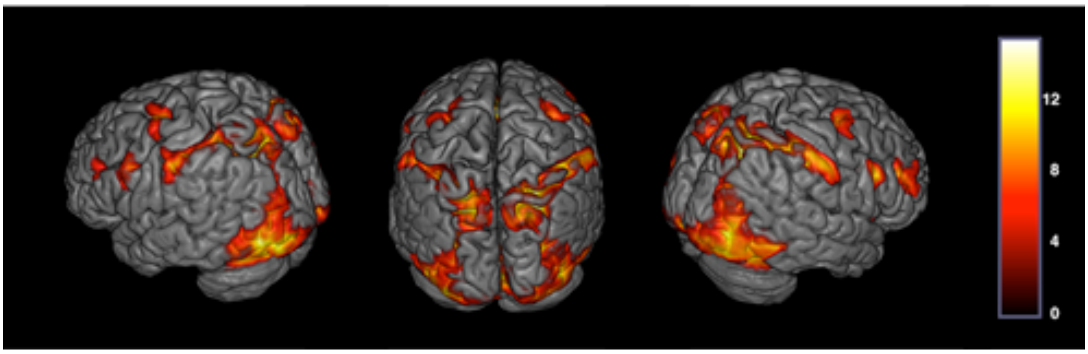


Fig.5. Brain activations specific to both visual recursion and embedded iteration tasks (VRT and EIT), in comparison with a simple similarity task. Both recursive and within-level iterative tasks showed activations within the visual 'dorsal stream', a bilateral network including regions from the occipital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

Visuo-spatial Hierarchy Differences: Within-level Transformations versus Recursion

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

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

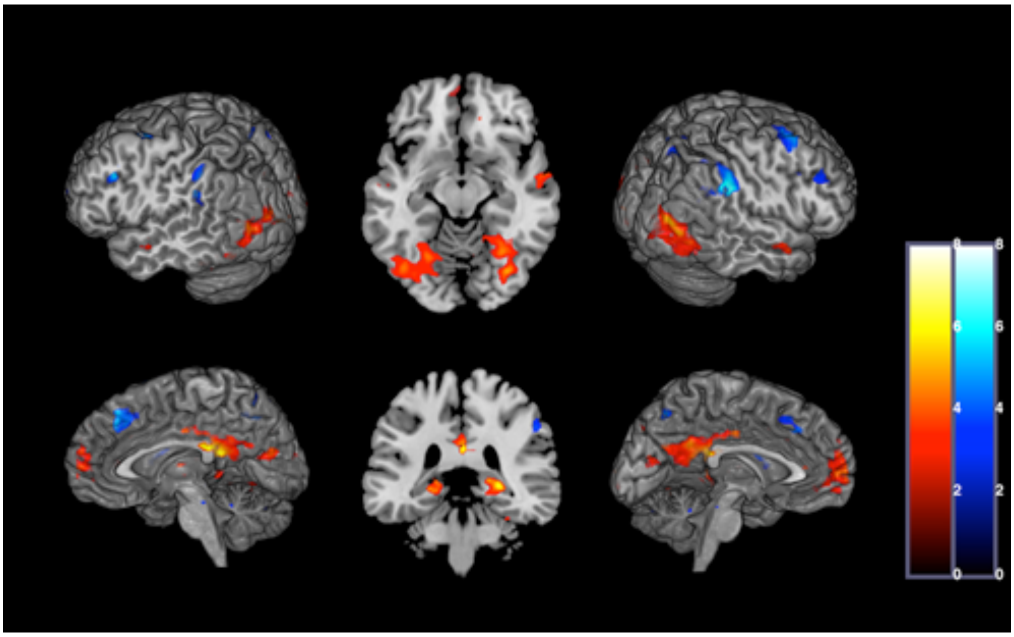


Fig.6. Brain activation contrast between visual recursion task (VRT) and embedded iteration task (EIT). For VRT (shown in red and summarized in Supplementary Table S4) larger responses were found in regions related with (i) the visual 'ventral stream' (including fusiform, lingual, and middle temporal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

Discussion

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

Our main findings were the following: 1) Both rule-based processes (within and between levels) activated a bilateral network (the dorsal stream) which includes visual association areas and fronto-parietal circuits associated with spatial reasoning [13]. Additionally, both rule-based tasks activate the inferior frontal gyrus (IFG, including parts of Broca's area), insula, cingulate cortex and right hippocampus; 2) Compared to within-level transformations, the representation of recursive processes generating new hierarchical levels (i.e. fractals) recruited regions within the parieto-medial temporal pathway (PMT; Fig.2) - including the posterior cingulate cortex (PCC) and retrosplenial cortex (RSC) - and their projections to the medial temporal cortex (MTL), which have been associated with the integration of spatial and semantic information [13]. We also found activations in the anterior portions of superior and middle temporal gyri (STG and MTG, respectively); 3) In contrast, within-level iterative rules activated the following regions more strongly; the dorsal stream, the dorsal fronto-parietal network (FPN), IFG, and basal ganglia. We now elaborate on these three basic findings.

Iterative Processes Generating Hierarchies Activate the Dorsal Stream and IFG

Compared to simple assessment of visual similarity, the cognitive processes involved in the representation of iterative rules correlate with greater activation of visual association areas, including bilateral activations in the intra-parietal sulcus (extending to portions of superior and inferior parietal cortex). These areas comprise the so-called 'dorsal stream', and are involved in the processing of information relating to the location of objects in visual-spatial structures ('where' information) [13]. Furthermore, we found activations in the supplementary motor area (SMA), pre-motor cortex (PMC), and prefrontal cortex (PFC). These areas have been described as projections of the dorsal stream and have been implicated in the control of eye movements, spatial working memory and executive control of visual-spatial processing [13]. We also found activations in the insula and anterior/middle cingulate gyrus, often described as part of a 'salience network' [26]. This network allows switching between external and internal modes of representation (correlated, respectively, with the activation of central executive and default-mode networks) and plays a crucial role in maintenance and update of predictions and expectations [26].

Finally, conjunction analysis revealed activations within Broca's area (and its right hemisphere homologue) for both within-level and cross-level transformations. Broca's area has been shown to be active in the processing of sequential hierarchies in natural language [31, 32], artificial grammars with sound sequences [33, 34], artificial grammars with image sequences [35], music patterns [36, 37] and during the processing of action sequences [28]. However the precise role of this structure is still uncertain, hypotheses ranging from it supporting sequential working memory, to participating in hierarchical or structure unification [29, 30, 38-46]. Our results support the hypothesis that Broca's area may be generally involved in maintaining online information or rules supporting iterative/sequential processes [27-29], rather than in the integration of multiple hierarchical-levels *per se* [30].

Representation of Self-similar Hierarchies (Fractals) Requires Integration of Spatial and Categorical Information

In addition to requiring the participation of the 'dorsal stream', the representation of processes generating new hierarchical levels recruited a bilateral network involving the ventral occipito-temporal cortex, including fusiform and lingual gyri - all parts of the visual 'ventral stream' [14]. This network has been associated with the representation of categorical or semantic information. Furthermore, these rules recruited the anterior regions of STG and MTG, which appear to correlate with the retrieval of abstract categories [47, 48]. Interestingly, VRT also specifically activated areas within the PMT pathway (RSC and PCC), which have been described as intermediary projections of the dorsal stream to the MTL. These areas are involved in the integration of objects in contextual frames [13] and in the integration of spatial and categorical/semantic information. Lesions in these areas are associated with spatial navigation deficits, in particular with an inability to use spatial landmarks (despite an intact ability to retrieve landmark location) [16, 19].

The generation of novel self-similar hierarchical levels also bilaterally activates the PMT projections into the MTL (hippocampus and parahippocampal cortex). These areas have been associated with episodic memory and with the formation of unified representations of items and contexts [49, 50]. The recruitment of the MTL has previously been reported as being crucial in the processing of spatial and social hierarchies [13, 51-53], and in studies investigating the processing of novel (vs. well-trained) linguistic hierarchies [54].

Taken together, these results suggest that episodic memory and the integration of items in contexts are crucial mechanisms in the processing of rule-based generation of novel hierarchical levels using recursive principles. Furthermore, this process requires the integration of spatial and categorical information. This finding is particularly intriguing since the visuo-spatial hierarchies employed in this study do not convey "semantic" information *per se*. We hypothesize that the representation of hierarchical dependencies may require the retrieval of "semantic" information of a rather abstract sort. In order to utilize a spatial landmark one needs both to know its location (where), and to know what it is a landmark of (what) - a type of referential relationship. Processing this abstract relationship between reference and referent may require the activation of traditional 'semantic' networks which would therefore be necessary for the integration of multiple hierarchical levels. Consistent with this supposition, in other domains, such as language, the processing of hierarchies is also associated with the activation of areas related with semantic retrieval (e.g. STG) [32].

Finally, our behavioral results suggest a specific correlation between VRT (but not EIT) and Tower of Hanoi, which requires hierarchical planning of actions and invites a recursive solution [55](Supplementary Results). Crucially, we used a score of Tower of Hanoi (longest sequence performed without mistakes) that cannot easily be explained by simple iterative mechanisms. This behavioral correlation lends support to the hypothesis that our visual recursion task may tap into cognitive resources associated with the processing of recursion.

Within-level Transformations are More Specifically Spatial

Compared with VRT, the representation of iterative processes transforming hierarchies within a fixed level correlated more strongly with the activation of areas in the visual dorsal stream [13]. This suggests, in agreement with previous research [10, 21], that these within-level transformations may rely on specific spatial resources ('where' information), to a greater extent than recursive transformations. Our behavioral results (see Supplementary Results) confirm that both the acquisition and application of within level rules correlate more strongly with working memory abilities than do rules generating novel hierarchical levels. Interestingly, small foci within Broca's area and its right homologue seemed to be more active in within-level transformations than in recursive transformations. These findings suggest that Broca's is not specifically active for the processing of cross-level hierarchical integration [30, 33, 35, 38, 56], but may be more generally involved in the storage and maintenance of rule-based iterative information, or in working memory processes [27-29, 46, 57]. These findings also suggest that recursive embedding is a more memory-efficient method to generate complex hierarchies.

Limitations of the Current Study

It could be argued that participants may have used simple heuristic strategies, comparing items according to their similarity, to solve our tasks. We tried to minimize this problem in three ways: 1) All neuroimaging analyses and comparisons were implicitly performed against a 'similarity task'; 2) We included different 'foil item categories' to block any specific heuristic strategies; and 3) We explicitly instructed and trained participants in the usage of within-level and recursive rules while solving EIT and VRT. Furthermore, VRT performance both inside and outside of the scanner was specifically correlated with Tower of Hanoi, which is considered a recursive planning task [55] and processing of VRT stimuli was more efficient than processing of EIT stimuli, despite both tasks using identical targets. Taken together, this suggests that our experiment design and analysis tapped into the representation of recursive principles rather than the application of simple heuristic strategies.

Conclusion

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

Acknowledgments

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

References

1 Chomsky, N. 1957 *Syntactic structures*. The Hague: Mouton.

2 Nardini, M., Jones, P., Bedford, R.,Braddick, O. 2008 Development of Cue Integration in Human Navigation. *Current Biology*. **18**, 689-693. (10.1016/j.cub.2008.04.021)

3 Unterrainer, J. M.,Owen, A. M. 2006 Planning and problem solving: From neuropsychology to functional neuroimaging. *Journal of Physiology-Paris*. **99**, 308-317. (10.1016/j.jphysparis.2006.03.014)

4 Badre, D. 2008 Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes. *Trends in Cognitive Sciences*. **12**, 193-200. (10.1016/j.tics.2008.02.004)

5 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**, 501-515. (10.1098/rstb.2002.1257)

6 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.

7 Jackendoff, R. 2002 *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford University Press.

8 Badre, D.,D'Esposito, M. 2009 Is the rostro-caudal axis of the frontal lobe hierarchical? *Nature Reviews Neuroscience*. **10**, 659-669. (10.1038/nrn2667)

9 Conway, C. M.,Christiansen, M. H. 2001 Sequential learning in non-human primates. *Trends in Cognitive Sciences*. **5**, 539-546.

10 Martins, M. D. 2012 Distinctive signatures of recursion. *Philosophical Transactions of the Royal Society B: Biological Sciences*. **367**, 2055-2064. (10.1098/rstb.2012.0097)

11 Fitch, W. T. 2010 Three meanings of "recursion": key distinctions for biolinguistics. In The evolution of human language. (ed. R. K. Larson, V. M. Déprez,H. Yamakido), pp. 73-90. New York: Cambridge University Press.

12 de Haan, E. H. F.,Cowey, A. 2011 On the usefulness of 'what' and 'where' pathways in vision. *Trends in Cognitive Sciences*. **15**, 460-466. (10.1016/j.tics.2011.08.005)

13 Kravitz, D. J., Saleem, K. S., Baker, C. I.,Mishkin, M. 2011 A new neural framework for visuospatial processing. *Nature Reviews Neuroscience*. **12**, 217-230. (10.1038/nrn3008)

14 Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G.,Mishkin, M. 2013 The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends in Cognitive Sciences*. **17**, 26-49. (10.1016/j.tics.2012.10.011)

15 Margulies, D. S., Vincent, J. L., Kelly, C., Lohmann, G., Uddin, L. Q., Biswal, B. B., Villringer, A., Castellanos, F. X., Milham, M. P.,Petrides, M. 2009 Precuneus shares intrinsic functional architecture in humans and monkeys. *Proceedings of the National Academy of Sciences*. **106**, 20069–20074.

16 Aguirre, G.,D'Esposito, M. 1999 Topographical disorientation: a synthesis and taxonomy. *Brain*. **122**, 1613-1628.

17 Ranganath, C.,Ritchey, M. 2012 Two cortical systems for memory-guided behaviour. *Nature Reviews Neuroscience*. **13**, 713-726. (10.1038/nrn3338)

18 Sato, N., Sakata, H., Tanaka, Y. L.,Taira, M. 2010 Context-Dependent Place-Selective Responses of the Neurons in the Medial Parietal Region of Macaque Monkeys. *Cerebral Cortex*. **20**, 846-858.

19 Ino, T., Doi, T., Hirose, S., Kimura, T., Ito, J.,Fukuyama, H. 2007 Directional Disorientation Following Left Retrosplenial Hemorrhage: a Case Report with FMRI Studies. *Cortex*. **43**, 248-245.

20 Buzsáki, G.,Moser, E. I. 2013 Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nature Neuroscience*. **16**, 130-138. (10.1038/nn.3304)

21 Martins, M. D.,Fitch, W. T. 2012 Empirical Approaches to Recursion. In The Evolution of Language - Proceedings of the 9th International conference. (ed. T. C.

Scott-Phillips, M. Tamariz, E. A. Cartmill,J. R. Hurford), pp. 219-225. Kyoto: Singapore: World Scientific.

22 Zaitsev, M., Hennig, J.,Speck, O. 2004 Point spread function mapping with parallel imaging techniques and high acceleration factors: fast, robust, and flexible method for echo-planar imaging distortion correction. *Society of Magnetic Resonance in Medicine*. **52**, 1156–1166. (doi:10.1002/mrm.20261)

23 Lund, T. E., Nørgaard, M. D., Rostrup, E., Rowe, J. B.,Paulson, O. B. 2005 Motion or activity: their role in intra- and inter-subject variation in fMRI. *NeuroImage*. **26**, 960–964. (10.1016/j.neuroimage.2005.02.021)

24 Nichols, T., Brett, M., Andersson, J., Wager, T.,Poline, J.-B. 2005 Valid conjunction inference with the minimum statistic. *NeuroImage*. **25**, 653–660. (doi:10.1016/j.neuroimage.2004.12.005)

25 Chumbley, J., Worsley, K., Flandin, G.,Friston, K. 2010 Topological FDR for neuroimaging. *NeuroImage*. **49**, 3057–3064. (doi:10.1016/j.neuroimage.2009.10.090)

26 Mueller, S. T. The Psychology Experiment Building Language. 2011.

27 Sridharan, D., Levitin, D. J.,Menon, V. 2008 A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proceedings of the National Academy of Sciences*. **105**, 12569-12574. (10.1073/pnas.0800005105)

28 Rajah, M. N., Ames, B.,D'Esposito, M. 2008 Prefrontal contributions to domain-general executive control processes during temporal context retrieval. *Neuropsychologia*. **46**, 1088-1103. (10.1016/j.neuropsychologia.2007.10.023)

29 Fazio, P., Cantagallo, A., Craighero, L., D'Ausilio, A., Roy, A. C., Pozzo, T., Calzolari, F., Granieri, E.,Fadiga, L. 2009 Encoding of human action in Broca's area. *Brain*. **132**, 1980-1988. (10.1093/brain/awp118)

30 de Vries, M. H., Christiansen, M. H.,Petersson, K. M. 2011 Learning Recursion: Multiple Nested and Crossed Dependencies. *Biolinguistics*. **5**, 10-35.

31 Berwick, R. C., Friederici, A. D., Chomsky, N.,Bolhuis, J. J. 2013 Evolution, brain, and the nature of language. *Trends in Cognitive Sciences*. (10.1016/j.tics.2012.12.002)

32 Wang, J., Conder, J. A., Blitzer, D. N.,Shinkareva, S. V. 2010 Neural representation of abstract and concrete concepts: A meta-analysis of neuroimaging studies. *Human Brain Mapping*. **31**, 1459-1468. (10.1002/hbm.20950)

33 Lehmann, D., Pascual-Marqui, R. D., Strik, W. K.,Koenig, T. 2010 Core networks for visual-concrete and abstract thought content: A brain electric microstate analysis. *NeuroImage*. **49**, 1073-1079. (10.1016/j.neuroimage.2009.07.054)

34 Eichenbaum, H., Sauvage, M., Fortin, N., Komorowski, R.,Lipton, P. 2012 Towards a functional organization of episodic memory in the medial temporal lobe. *Neuroscience & Biobehavioral Reviews*. **36**, 1597-1608. (10.1016/j.neubiorev.2011.07.006)

35 Ranganath, C. 2010 A unified framework for the functional organization of the medial temporal lobes and the phenomenology of episodic memory. *Hippocampus*. **20**, 1263-1290. (10.1002/hipo.20852)

36 Aminoff, E., Gronau, N.,Bar, M. 2006 The Parahippocampal Cortex Mediates Spatial and Nonspatial Associations. *Cerebral Cortex*. **17**, 1493-1503. (10.1093/cercor/bhl078)

37 Kumaran, D., Melo, H. L.,Duzel, E. 2012 The Emergence and Representation of Knowledge about Social and Nonsocial Hierarchies. *Neuron*. **76**, 653-666. (10.1016/j.neuron.2012.09.035)

38 Zaretskaya, N., Anstis, S.,Bartels, A. 2013 Parietal Cortex Mediates Conscious Perception of Illusory Gestalt. *Journal of Neuroscience*. **33**, 523-531. (10.1523/jneurosci.2905-12.2013)

39 Opitz, B.,Friederici, A. D. 2007 Neural basis of processing sequential and hierarchical syntactic structures. *Human Brain Mapping*. **28**, 585-592. (10.1002/hbm.20287)

40 Friederici, A. D., Bahlmann, J., Heim, S., Schubotz, S. I.,Anwander, A. 2006 The brain differentiates human and non-human grammars: Functional localization and structural connectivity. *Proceedings of the National Academy of Sciences*. **103**, 2458-2463. (10.1073/pnas.0509389103)

41 Goel, V.,Grafman, J. 1995 Are the frontal lobes implicated in "planning" functions? Interpreting data from the Tower of Hanoi. *Neuropsychologia*. **33**, 623-642.

42 Bahlmann, J., Schubotz, R. I., Mueller, J. L., Koester, D.,Friederici, A. D. 2009 Neural circuits of hierarchical visuo-spatial sequence processing. *Brain Research*. **1298**, 161-170. (10.1016/j.brainres.2009.08.017)

43 Bahlmann, J., Schubotz, R.,Friederici, A. 2008 Hierarchical artificial grammar processing engages Broca's area. *NeuroImage*. **42**, 525-534. (10.1016/j.neuroimage.2008.04.249)

44 Fadiga, L., Craighero, L.,D'Ausilio, A. 2009 Broca's Area in Language, Action, and Music. *Annals of the New York Academy of Sciences*. **1169**, 448-458. (10.1111/j.1749-6632.2009.04582.x)

45 Friederici, A. D., Bahlmann, J., Friedrich, R.,Makuuchi, M. 2011 The Neural Basis of Recursion and Complex Syntactic Hierarchy. *Biolinguistics*. **5**, 087–104.

46 Bengtsson, S. L., Haynes, J. D., Sakai, K., Buckley, M. J.,Passingham, R. E. 2008 The Representation of Abstract Task Rules in the Human Prefrontal Cortex. *Cerebral Cortex*. **19**, 1929-1936. (10.1093/cercor/bhn222)

47 Baddeley, A. 2003 Working memory: looking back and looking forward. *Nature Reviews Neuroscience*. **4**, 829-839. (10.1038/nrn1201)

Supplementary Methods 1 - *Stimulus Generation*

Recursion

Stimuli were generated computationally using custom-written Python code. Stimuli were based on the properties of fractal geometry1. Visual fractals can be generated from single constituents like lines, squares or triangles (the initiators) by applying a simple rule of transformation (the generator) a given number of iterative steps. The structures generated by this process are hierarchical and self-similar (Fig.S1).

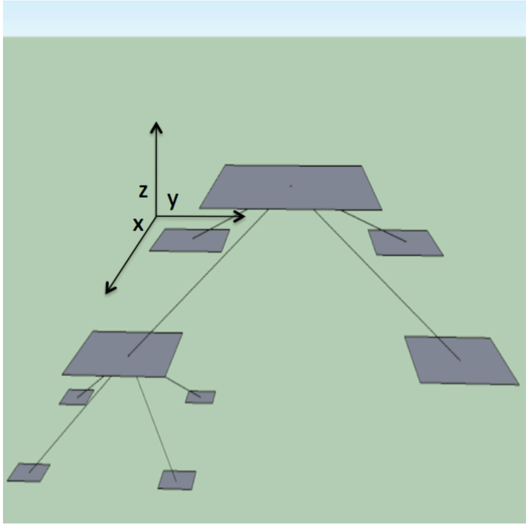


Fig.S1 Visual fractals can be perceived as visuo-spatial hierarchies: a) Different elements (squares) are organized in a 2D space, defined by the xy-axis; b) Different hierarchical levels are organized along the z-axis. An element with a higher value of z is dominant over an element with a lower value of z (if the elements are connected).

The principles underlying the generation of the stimuli were straightforward: There was a pool of 7 geometrical shapes (rectangle, square, circle, star, triangle, pentagon and hexagon), which were used as initiators. Then, different kinds of recursive embedding rules (generators) were applied to these shapes to generate hierarchical structures. Four iterative steps were generated for each fractal (Fig.S2).

There were 2 main categories of generators which determined the visual complexity of the structures that resulted from their application. Here 'visual complexity' refers to the number of elements added to the structure, dependent on the same previously existing element, "dependent" in the sense that the spatial coordinates of the dominant element determine the spatial coordinates of the subordinate elements. For instance, in Fig.S1, the visual complexity is 4, since 4 squares are added in dependency of the same square. In this study, the ‘visual complexity’ utilized was either 3 or 4 (Fig.S2).

Together with the first four iterations of a fractal, a foil structure was generated. This foil structure corresponded to an “incorrect” fourth iteration, and was generated by applying a rule (generator) to generate the fourth iteration, that was different from the one used to generate the previous three iterations. There were 2 categories of foils, relative to the kind of process used in their generation (Fig.S2). These processes were the following two types: i) Misplacement of one element within each positional scheme (‘Odd constituent foil’); ii) Utilization of a novel positional scheme for all new added elements, different from the scheme used for the previous iterations (‘Positional error foil’). These different foils were used to prevent the development of simple heuristic strategies based simply on comparison of the ‘correct’ and ‘incorrect’ fourth iterations, strategies that could be unrelated to the recursive rule itself.

In overall, the combination of both ‘visual complexity’ and ‘foils’ categories resulted in 4 categories of foil stimuli: Odd constituent foils with complexity 3 and 4; and Positional error foils with complexity 3 and 4. Several stimuli of each category were generated using the programming language Python. For both training and MR sessions, all stimuli categories were equally balanced.

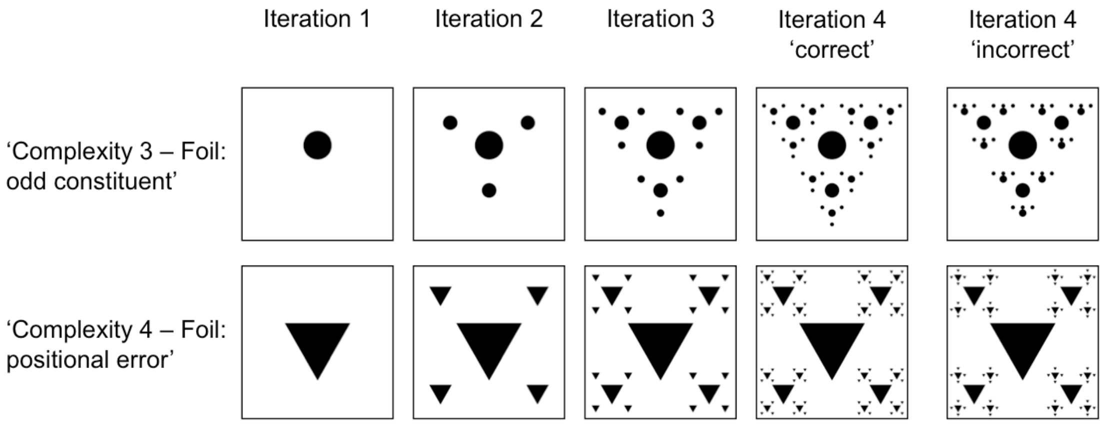


Fig.S2 Examples of fractals used in the Visual Recursion Task: The first four iterations of a fractal generation, as well as one foil (‘incorrect’ fourth iteration), were produced. There were different categories of visual complexity - 3 and 4 – and different categories of foils (see text for details) – ‘Odd constituent’ and ‘Positional error’. Some of the combinations of these categories are depicted.

Iteration

This task was also adapted from Martins et al. (in review). The principle underlying this 'embedded iteration' task (EIT) is similar to the recursion task in the sense that it involves an iterative procedure applied to the same class of hierarchical structures. However, EIT lacks recursive embedding. Instead, in EIT, there is an iterative process which embeds additional elements within a pre-existing hierarchical structure, without producing new hierarchical levels. Stimuli categories were similar to the recursion task. Crucially, in the "Odd constituent foil" category, the element placed in the incorrect position corresponded to the new element being added in the fourth iteration (Fig S3). This means that none of the elements already present in the structure from the third iteration were displaced.

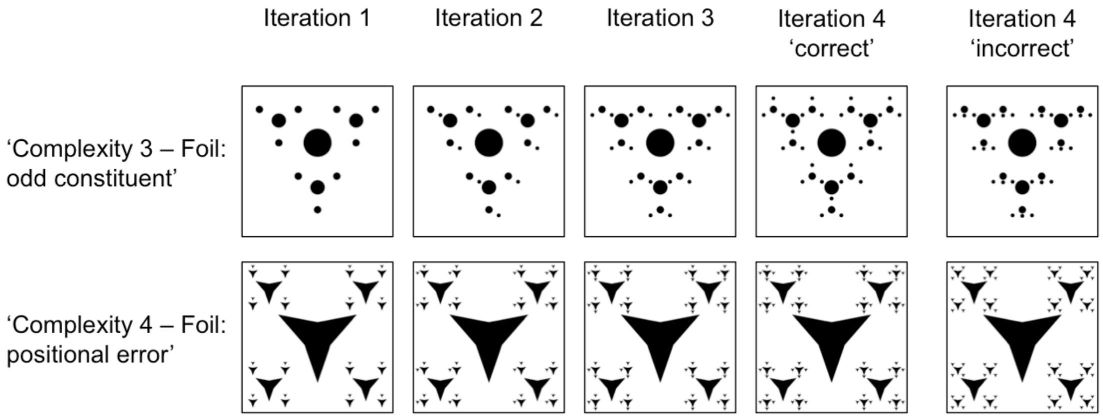


Fig.S3. Examples of fractals used in the Embedded Iteration Task: The first four iterations of a fractal generation, as well as one foil ('incorrect' fourth iteration), were produced. There were different categories of visual complexity - 3 and 4 – and different categories of foils (see text for details) – 'Odd constituent' and 'Positional error'. Some of the combinations of these categories are depicted.

Similarity

The stimuli used in this task were picked randomly from the pool of well-formed fourth iterations generated for the visual recursion task.

Supplementary Methods 2 – Behavioral data

	M	SD
EIT ODD foil	0.90	0.10
EIT POSITIONAL foil	0.93	0.07
VRT ODD foil	0.96	0.06
VRT POSITIONAL foil	0.96	0.06

Table S1 Performance across different foil categories and tasks (see Supplemental materials 1 for details). M, Mean. SD, Standard deviation. VRT, Visual recursion task. EIT, Embedded iteration task.

Supplementary Methods 3 – Neurocognitive assessment

	M	SD	Minimum	Maximum
DSPAN	6.9	1.2	4	9
CORSI	6.5	1.3	3	9
ToH	5.3	1.6	0	7
RAVEN	29.7	2.7	17	32

Table S2 Summary of neuro-cognitive pre-testing results. M, Mean. SD, Standard deviation. DSPAN, Digit Span. CORSI, Corsi block tapping. ToH, Tower of Hanoi. RAVEN, progressive matrices of Raven.

Supplementary Methods 4 – Activation tables

Brain Region	Label	x	y	z	Cluster size	t-value
R Middle Occipital G	**hIP1**	**30**	**-69**	**27**	**84361**	**18.41**
R Superior Occipital G	hIP1/hIP3	26	-61	42		16.79
Left Cerebellum		-8	-72	-26		16.74
R Superior Parietal Lobule	hIP3	27	-58	52		16.36
Cerebellar Vermis		6	-72	-23		16.34
L Superior Parietal Lobule	SPL(7A)	-21	-64	46		16.01
R Hippocampus		22	-30	-3		16
L Inferior Occipital G	hOC5(V5)	-40	-81	-5		15.64
R Inferior Parietal Lobule	hIP3	32	-49	46		15.32
L SMA	**BA6**	**-2**	**12**	**48**	**6698**	**16.92**
R Middle Cingulate C	BA6	6	17	43		14.5
L Middle Frontal G	BA6	-26	-3	51		14.32
R Inferior Frontal G	**BA44**	**48**	**9**	**28**	**4013**	**17.04**
R Superior Frontal G	**BA6**	**27**	**2**	**52**	**2771**	**13.74**
R Precentral G	BA6	26	-3	49		13.01
R Middle Frontal G	BA6	44	0	57		6.91
R Middle Frontal G	**BA45**	**46**	**33**	**22**	**1759**	**10.11**
R Inferior Frontal G	**BA45**	42	33	16		9.95
L Precentral G	**BA44**	-50	5	28	1668	10.77
L Precentral G	BA6	-57	8	37		9.65
L Insula Lobe	BA13	**-30**	**17**	**7**	**475**	**9.52**
L Inferior Frontal G	**BA45**	**-52**	**29**	**31**	**221**	**6.83**
L Anterior Cingulate C	BA33	**-4**	**10**	**24**	**219**	**7.36**
R Insula Lobe	**BA13**	**39**	**0**	**4**	**27**	**7.1**
R SMA	**BA6**	**10**	**0**	**64**	**6**	**5.73**

Table S3 Anatomical areas of the significant brain activations in the conjunction analysis (VRT and EIT; FDR, $p < 0.05$). Maximum peak of each cluster is depicted in **bold**, other maxima within the same cluster in normal font. MNI coordinates (x, y, z) are depicted. Label (when available) is based on probabilistic maps. L, left hemisphere; R, right hemisphere; G, gyrus; S, sulcus; C, cortex; BA, Brodmann area.

Brain Region	Label	x	y	z	Cluster size	t-value
L Superior Occipital G	**BA17**	**-9**	**-102**	**9**	**6343**	**5.25**
L Lingual G	V4	-26	-64	-9		4.71
L Fusiform G	BA37	-32	-48	-20		4.70
R Middle Occipital G	**BA19**	**40**	**-88**	**13**	**5618**	**5.74**
R Fusiform G	BA37	34	-51	-8		5.54
R Cerebellum		22	-49	-17		4.47
R Posterior Cingulate C	**BA29**	**6**	**-40**	**19**	**3143**	**6.87**
L Precuneus	BA7a	-10	-67	31		6.28
L Cuneus	BA7a	-12	-60	28		5.17
L Superior Medial Frontal G	**BA10**	**-2**	**60**	**4**	**2052**	**4.47**
L Middle Orbital G	**BA10**	-4	66	-3		4.33
L Superior Frontal G	**BA9**	-15	56	30		3.72
R Cuneus	**BA18**	**8**	**-78**	**21**	**510**	**4.74**
R Calcarine G	BA18	3	-67	16		2.93
L Calcarine G	BA17	-6	-82	12		2.69
R Hippocampus		**26**	**-36**	**-2**	**470**	**6.54**
R Lingual G	BA30	10	-40	3		3.93
L Hippocampus		**-20**	**-37**	**3**	**305**	**5.31**
L Lingual G	BA30	-8	-40	-2		4.47
L ParaHippocampal G	**BA35**	-18	-31	-11		4.21
R Middle Temporal G	**BA21**	**60**	**-1**	**-17**	**276**	**4.24**
R Superior Temporal G	**BA21**	63	-10	-9		3.06
R Thalamus		**4**	**-15**	**3**	**196**	**4.64**
L Middle Temporal G	**BA21**	**-54**	**-4**	**-15**	**108**	**3.49**
L Superior Temporal G	**BA22**	-54	-12	-6		3.02
R Middle Orbital G	**BA10**	**9**	**40**	**-8**	**63**	**6.01**
L Middle Cingulate C	**BA32**	**-6**	**18**	**33**	**53**	**3.53**
L Anterior Cingulate C	**BA24**	**0**	**34**	**9**	**31**	**4.37**
L Insula Lobe	**BA13**	**-38**	**-1**	**16**	**23**	**5.49**
R Postcentral G	**BA3**	**40**	**-24**	**43**	**20**	**3.36**
L Cerebellum		**-46**	**-57**	**-26**	**18**	**2.94**
R Anterior Cingulate C	**BA24**	**6**	**34**	**6**	**13**	**3.26**
R Middle Cingulate C	**BA32**	**2**	**24**	**30**	**11**	**3.27**
R Rolandic Operculum		**45**	**-4**	**16**	**10**	**5.70**
L Postcentral G	**BA1**	**-54**	**-28**	**54**	**9**	**2.87**
L Thalamus		**-6**	**-12**	**9**	**9**	**3.19**

Table S4 Anatomical areas of the significant brain activations in the contrast 'Visual Recursion Task > Embedded Iteration Task' (FDR, $p < 0.05$). Maximum peak of each cluster is depicted in **bold**, other maxima within the same cluster in normal font. MNI coordinates (x, y, z) are depicted. Label (when available) is based on probabilistic maps. L, left hemisphere; R, right hemisphere; G, gyrus; S, sulcus; C, cortex; BA, Brodmann area; V4.

Brain Region	Label	x	y	z	Cluster size	t-value
R SupraMarginal G	**PFm**	**52**	**-45**	**43**	**1733**	**7.38**
R Inferior Parietal Lobule	PFm	52	-51	42		7.22
R Angular G	BA7	39	-64	46		4.29
R Middle Frontal G	**BA6**	**28**	**11**	**48**	**1498**	**5.64**
R Superior Frontal G	BA6	32	8	63		4.73
R Precuneus	**BA7a**	**8**	**-58**	**52**	**989**	**5.48**
R Superior Parietal Lobule	BA7a	14	-70	54		4.30
L Middle Frontal G	**BA6**	**-28**	**9**	**57**	**953**	**5.68**
L Precentral Gyrus	BA6	-30	-4	46		3.47
L Superior Medial Frontal G	**BA6**	**-8**	**29**	**37**	**939**	**5.91**
R Superior Medial Frontal G	BA6	6	27	43		5.00
L SMA	BA8	-4	24	49		4.07
R Middle Frontal G	**BA45**	**42**	**30**	**33**	**414**	**4.20**
R Inferior Frontal G	BA44	46	27	30		4.17
R Inferior Frontal G	BA45	51	32	24		3.70
L Inferior Parietal Lobule	**PFm**	**-46**	**-51**	**46**	**323**	**5.62**
L Inferior Parietal Lobule	PF	-48	-48	43		4.84
L Inferior Parietal Lobule	PGa	-42	-55	54		3.53
R Insula Lobe	**BA13**	**32**	**27**	**-2**	**320**	**5.06**
L Inferior Frontal G	**BA44/45**	**-45**	**21**	**34**	**258**	**6.22**
R Caudate Nucleus		**14**	**8**	**9**	**249**	**4.83**
L Caudate Nucleus		**-15**	**2**	**15**	**191**	**4.74**
L Precuneus	**BA7a**	**-8**	**-66**	**54**	**190**	**5.17**
R Pallidum		**20**	**-3**	**7**	**34**	**4.60**
Cerebellar Vermis		**2**	**-49**	**-21**	**33**	**3.65**
R Middle Occipital G	**PGp**	**38**	**-73**	**37**	**31**	**4.04**
L Cerebellum		**-10**	**-72**	**-26**	**24**	**3.99**
R Middle Temporal G	**PGa**	**50**	**-51**	**19**	**5**	**3.45**

Table S5 Anatomical areas of the significant brain activations in the contrast ‘Embedded Iteration Task > Visual Recursion Task (FDR, $p < 0.05$). Maximum peak of each cluster is depicted in **bold**, other maxima within the same cluster in normal font. MNI coordinates (x, y, z) are depicted. Label (when available) is based on probabilistic maps. L, left hemisphere; R, right hemisphere; G, gyrus; S, sulcus; C, cortex; BA, Brodmann area.

8. General Discussion

Recursion is an interesting theoretical concept and its empirical investigation faces many challenges. In this thesis I presented a new framework, a novel task, and a set of experiments that aimed at characterizing recursion as a testable cognitive end psychological construct.

In the theoretical framework presented here, recursion is defined as the ability to represent hierarchical self-similarity. Behaviorally, it can be detected by the ability to use information shared across hierarchical levels of a structure for generating new levels within that structure. This ability can be useful in domains such as language, social navigation and problem solving, which benefit from the generation of new hierarchical levels using simple and parsimonious rules.

I began by defending the position that we need to investigate the processing of recursion in a variety of domains and search for common principles. I argued that only by taking these empirical steps will it be possible to evaluate whether recursion is a domain-specific module, an umbrella term for a set of different modules, or an epiphenomenon resulting from the interaction of several cognitive abilities.

Working towards this goal, my colleagues and I developed a visual recursion task (VRT) that tests for the ability to detect cross-level hierarchical regularities in the visual-spatial domain. We also developed a control task using embedded iteration (EIT), which tests for the ability to transform hierarchies using a non-recursive rule. We devised a set of experiments to validate these tasks.

First, we found that VRT is cognitively distinct both from general intelligence and simple embedded iteration, and does not correlate with visuo-spatial working memory. However, VRT has good internal reliability and correlates with other "recursive" tasks such as the Tower of Hanoi (Goel & Grafman, 1995). Second, we found that participants were more sensitive to cross-level hierarchical information in VRT than in EIT. These findings give us confidence that VRT measures something specific to hierarchical self-embedding and recursion, and that it can be used to tap into this cognitive construct.

In the next sections I will summarize about the empirical findings on recursion in the visual domain. Overall, our empirical results suggest that recursion is more abstract than simple embedded iteration and that it uses hierarchical integration processes. Furthermore, it is useful for parsing complex structures, and it is not language domain-specific.

8.1. Recursion is abstract

In this thesis, I presented a series of results suggesting that visual recursion is only weakly correlated with resources associated with specific visuo-spatial processing. Instead, it seems to be strongly correlated with systems and tasks that process hierarchical structures and use abstract categories.

First, in comparison with simple embedded iteration, recursion correlates less strongly with visuo-spatial memory and with non-verbal intelligence. Furthermore, it activates to a lesser extent the visuo-spatial brain pathway ‘dorsal stream’, which is involved in the processing of spatial information (Kravitz, Saleem, Baker, & Mishkin, 2011).

Second, in our visual recursion task (VRT) higher reaction times and self-reported analytical strategies (vs. intuitive) correlated with higher scores. Furthermore, the usage of recursive rules correlated with accuracy in a task called Tower of Hanoi (ToH), in which participants were asked to build a sequence of nested movements in their minds, before initiating the physical manipulation of the stimulus material. In other words, in order to solve a typical ToH problem participants had to build a nested representation of the kind ‘[movement A [movement B [movement C]]]’ before manipulating the stimuli. The ability to form these representations correlated with performance in our visual recursion task.

Third, consistent with these behavioral findings, the representation of recursive rules activated the anterior frontal lobe to a greater extent than embedded iteration. Since the anterior frontal lobe is associated with abstract thinking (Badre, 2008; Badre & D'Esposito, 2009), this seems to confirm the theoretical hypothesis that recursive representations are more abstract than simple embedded iteration.

Finally, in our developmental experiment, the acquisition of recursive rules required the previous acquisition of simple iterative rules, but not the other way around. This learning path from iteration to recursion was influenced only by experience, and not by developmental factors. It is interesting that children also need to acquire simple iterative representations in language before they are able to master recursion (Roeper, 2011). This again suggests that recursive representations might be more abstract and may rely on the use of categories that are harder to acquire.

8.2. Recursion is integrative

In the theoretical framework presented here I argued that recursion requires the integration of cross-level hierarchical information. Our experiments provided evidence supporting these assumptions.

First, in our behavioral experiments, participants seemed to be sensitive to variation in the complexity of the rules governing the generation of new hierarchical levels: the more complex the rule the worse the performance. Based on results from EIT, we could also rule out that this was a simple effect of visual complexity. These findings, together with consistent performance across different foil categories, strongly suggest that participants were not using simple visual heuristic strategies to solve VRT. They used the cross-level hierarchical information to make their judgments about whether recursive structures were 'well-formed' or not.

Second, in comparison with EIT, VRT activated a set of brain circuits involved in the processing and integration of spatial and semantic information (Kravitz et al., 2011; Kravitz, Saleem, Baker, Ungerleider, & Mishkin, 2013), including the posterior cingulate cortex and medial temporal cortex. Lesions in these areas are associated with spatial navigation deficits, in particular with an inability to use spatial landmarks (despite an intact ability to retrieve landmark location) (Aguirre & D'Esposito, 1999; Ino et al., 2007). This highlights the integrative role of the posterior cingulate cortex in hierarchical visuo-spatial processing.

Interestingly, even though we used meaningless stimuli as visual fractals, we found activations in the superior temporal sulcus, which has been related with semantic/categorical processing (Lehmann, Pascual-Marqui, Strik, & Koenig, 2010; Opitz & Friederici, 2007; Wang, Conder, Blitzer, & Shinkareva, 2010). This again hints at the central role of categorical processing in recursive representations.

Finally, the ability to acquire both recursive and simple iterative rules correlated with the development of grammar comprehension independently of the effects of intelligence. This means that our visual tasks do tap into cognitive systems shared with language processing, which might be related with the representation of hierarchical structures, or as we discussed above, with categorical processing. Future research will be necessary to elucidate the nature of these shared resources.

8.3. Recursion is useful

In the theoretical chapters, I argued that recursion could be useful in a variety of domains, since it allows the usage of simple rules to represent complex hierarchies. However, as I discussed above, recursion seems to be harder to acquire than other rules that can also be used to represent hierarchical relationships. How do we reconcile these perspectives?

First, even though recursion seems harder to acquire than embedded iteration, once it is available, it seems to enhance the accuracy and speed of processing of hierarchical structures. In our fMRI study, participants were faster and better in VRT than in EIT, but only after training with both tasks.

Second, in the developmental study, the use of recursive rules seem to lead to increased rejection of structures with errors deeply nested within the visual hierarchies. In agreement with previous research on the visual domain (Alvarez, 2011), compressed representations of hierarchical structures can lead to a better processing of fine-grained details. Together, these findings support the view that recursive representations are harder to acquire, but once available, can lead to enhanced representations of hierarchical structures, perhaps by compressing the amount of information that needs to be temporarily stored in domain-specific working memory buffers.

8.4. Recursion is not language domain-specific

To date, the most influential theory about recursion hypothesizes that this ability is either language domain-specific, or that uses in other domains are parasitic on language resources (Fitch et al., 2005; Hauser et al., 2002).

I divided this problem in two questions:

1. Is recursion specific to the linguistic domain?
2. Is language necessary to use recursion in non-linguistic domains?

The answer to the first question is ‘No’. I have shown that recursion is available in the visuo-spatial domain. Regarding the second question, the results presented here strongly suggest that humans can build recursive representations independently of language. Evidence supporting this is the following: (1) The acquisition of recursive rules in vision does not require on-line access to verbal resources; (2) the representation of recursive rules while generating hierarchical levels does not activate the typical perisylvian brain areas associated with on-line language processing; and (3) grammar comprehension is not specifically correlated with visual recursion.

However, we did find some interesting parallels between syntactic and visuo-spatial recursion: (1) They follow similar learning constraints (prior acquisition of iteration is necessary); (2) both domains activate brain areas associated with categorical/semantic processing; and (3) grammar comprehension correlates with rule-based iterative processes in vision (embedded iteration and recursion).

Taken together, these data suggest that general mechanisms for hierarchical processing may exist, and that recursion is a particular subset of hierarchical processing. Crucially, these mechanisms can operate independently of on-line verbal resources, though they may still depend on the availability of resources shared with language, for example, semantic or categorical knowledge.

9. Future directions

Our results demonstrate that recursion in vision is correlated with other ‘recursive’ tasks such as Tower of Hanoi. To further investigate issues concerning the domain-specificity or generality of recursive representations we have to broaden our research scope. My colleagues and I are currently developing a music recursion task, in which participants have to acquire and apply processes generating music fractals (Martins, Gingras, & Puig-Waldemueller, 2014). Our preliminary results confirm that recursion is less correlated with domain-specific (music) memory resources than embedded iteration, and that recursive representations enhance the detection of errors deeply nested within hierarchical structures. Furthermore, in comparison with embedded iteration, music recursion seems to correlate with brain activation in the anterior frontal lobe. These results are exciting and promising. Together, visual recursion and music recursion tasks will provide conclusive evidence regarding the nature of recursion, and its relationship with other cognitive resources.

Finally, because both these tasks are non-linguistic, they can in principle be used to test non-human animals. We could for example test our closest primate relatives, the chimpanzees, and finally address the question whether recursion evolved specifically in humans, and whether its availability is associated with our outstanding ability to generate complex hierarchical structures.

10. Bibliography (General introduction and conclusion)

Aguirre, G., & D'Esposito, M. (1999). Topographical disorientation: a synthesis and taxonomy. *Brain,* 122(9), 1613-1628.

Alegre, M. A., & Gordon, P. (1996). Red rats eater exposes recursion in children's word formation. *Cognition,* 60, 65-82.

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences,* 15(3), 122-131.

Badre, D. (2008). Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes. *Trends in Cognitive Sciences,* 12(5), 193-200.

Badre, D., & D'Esposito, M. (2009). Is the rostro-caudal axis of the frontal lobe hierarchical? *Nature Reviews Neuroscience,* 10(9), 659-669.

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.

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. (2011). *The Recursive Mind*. Princeton: Princeton University Press.

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.

Fitch, W. T., Hauser, M., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition,* 97(2), 179-210.

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.

Gordon, P. (2004). Numerical Cognition Without Words: Evidence from Amazonia. *Science,* 306(5695), 496-499.

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.

Hulst, H. v. d. (2010). *Recursion and Human Language*. Berlin/New York: de Gruyter Mouton.

Humboldt, W. v. (1972). *Linguistic Variability and Intellectual Development*. Philadephia: University of Pennsylvania Press.

Ino, T., Doi, T., Hirose, S., Kimura, T., Ito, J., & Fukuyama, H. (2007). Directional Disorientation Following Left Retrosplenial Hemorrhage: a Case Report with FMRI Studies. *Cortex,* 43(2), 248-245.

Jackendoff, R. (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*: Oxford University Press.

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.

Kravitz, D. J., Saleem, K. S., Baker, C. I., & Mishkin, M. (2011). A new neural framework for visuospatial processing. *Nature Reviews Neuroscience,* 12(4), 217-230.

Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., & Mishkin, M. (2013). The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends in Cognitive Sciences,* 17(1), 26-49.

Laury, R., & Ono, T. (2010). Recursion in conversation: What speakers of Finnish and Japanese know how to do. In H. v. d. Hulst (Ed.), *Recursion and Human Language*): de Gruyter Mouton.

Lehmann, D., Pascual-Marqui, R. D., Strik, W. K., & Koenig, T. (2010). Core networks for visual-concrete and abstract thought content: A brain electric microstate analysis. *NeuroImage,* 49(1), 1073-1079.

Lobina, D. J. (2011). Recursion and the competence/performance distinction in AGL tasks. *Language and Cognitive Processes,* 26(10), 1563-1586.

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.

Lowenthal, F., & Lefebvre, L. (2014). *Language and Recursion*. New York: Springer.

Luuk, E., & Luuk, H. (2010). The redundancy of recursion and infinity for natural language. *Cognitive Processing,* 12(1), 1-11.

Martins, M. D. (2012). Distinctive signatures of recursion. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 367(1598), 2055-2064.

Martins, M. D., Gingras, B., & Puig-Waldemueller, E. (2014). *Cognitive representation of fractal music: empirical approach to music recursion.* Paper presented at the EvoMus: The evolution of music and language in a comparative perspective, Vienna.

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.

Miller, S. A. (2009). Children's Understanding of the Second-Order Mental States. *Psychological Bulletin,* 135(5), 749-773.

Nardini, M., Jones, P., Bedford, R., & Braddick, O. (2008). Development of Cue Integration in Human Navigation. *Current Biology,* 18(9), 689-693.

Opitz, B., & Friederici, A. D. (2007). Neural basis of processing sequential and hierarchical syntactic structures. *Human Brain Mapping,* 28(7), 585-592.

Pinker, S., & Jackendoff, R. (2005). The faculty of language: what's special about it? *Cognition,* 95(2), 201-236.

Reboul, A. (2012). Language: Between cognition, communication and culture. *Pragmatics & Cognition,* 20(2), 295-316.

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.

ten Cate, C., & Okanoya, K. (2012). Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 367(1598), 1984-1994.

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.

Wang, J., Conder, J. A., Blitzer, D. N., & Shinkareva, S. V. (2010). Neural representation of abstract and concrete concepts: A meta-analysis of neuroimaging studies. *Human Brain Mapping,* 31(10), 1459-1468.

Watumull, J., Hauser, M. D., Roberts, I. G., & Hornstein, N. (2014). On Recursion. *Frontiers in Psychology,* 4.

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.