



Review

# From Fractal Geometry to Fractal Cognition: Experimental Tools and Future Directions for Studying Recursive Hierarchical Embedding

Mauricio J. D. Martins

SCAN-Unit, Department of Cognition, Emotion, and Methods in Psychology, Faculty of Psychology, University of Vienna, 1090 Vienna, Austria; mauricio.martins@univie.ac.at

## Abstract

The study of fractals has a long history in mathematics and signal analysis, providing formal tools to describe self-similar structures and scale-invariant phenomena. In recent years, cognitive science has developed a set of powerful theoretical and experimental tools capable of probing the representations that enable humans to extend hierarchical structures beyond given input and to generate fractal-like patterns across multiple domains, including language, music, vision, and action. These paradigms target *recursive hierarchical embedding* (RHE), a generative capacity that supports the production and recognition of self-similar structures at multiple scales. This article reviews the theoretical framework of RHE, surveys empirical methods for measuring it across behavioral and neural domains, and highlights their potential for cross-domain comparisons and developmental research. It also examines applications in linguistic, musical, visual, and motor domains, summarizing key findings and their theoretical implications. Despite these advances, the computational and biological mechanisms underlying RHE remain poorly understood. Addressing this gap will require linking cognitive models with algorithmic architectures and leveraging the large-scale behavioral and neuroimaging datasets generated by these paradigms for fractal analyses. Integrating theory, empirical tools, and computational modelling offers a roadmap for uncovering the mechanisms that give rise to recursive generativity in the human mind.

**Keywords:** recursive hierarchical embedding; fractal cognition; generativity; cross-domain cognition; cognitive neuroscience; human cognition

Academic Editors: Mingshu Peng and Trung V. Phan

Received: 16 August 2025

Revised: 4 October 2025

Accepted: 7 October 2025

Published: 10 October 2025

**Citation:** Martins, M.J.D. From Fractal Geometry to Fractal Cognition: Experimental Tools and Future Directions for Studying Recursive Hierarchical Embedding. *Fractal Fract.* **2025**, *9*, 654. <https://doi.org/10.3390/fractalfract9100654>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Fractals have long been a central topic in mathematics and signal analysis, providing formal tools to describe self-similar structures and scale-invariant processes across a wide range of natural and artificial systems [1–4]. Classical work in geometry, complex systems theory, and physics has established rigorous methods for characterizing such patterns—whether through fractal dimension, scaling laws, or recursive generation rules—and these concepts have been instrumental in disciplines ranging from geophysics to image compression [5,6].

In the past decade, cognitive science has begun to adapt these principles to investigate a fundamentally different question: how does the human mind represent and manipulate the generative rules that give rise to fractal-like structures [7,8]? This

emerging field, which we term *fractal cognition*, focuses not on measuring the fractal properties of external stimuli but on probing the internal cognitive representations that allow humans to extend hierarchical structures beyond given input and generate self-similar patterns in multiple domains.

At the core of this capacity lies recursive hierarchical embedding (RHE)—the ability to generate or recognize structures in which constituents are embedded within constituents of the same type, allowing complexity to scale with each repetition of the rule [8]. RHE is a hallmark of human generativity, supporting complex syntax in language [9,10], nested harmonic and rhythmic patterns in music [11,12], hierarchical segmentation in vision [13], and structured sequencing in motor actions [14–16]. Theoretical and empirical work has shown that RHE enables the construction of patterns with fractal-like properties [7,8], but unlike mathematical fractals, these cognitive structures are bounded by memory, attention, and learning constraints [17–19].

Recent experimental advances have provided powerful behavioral and neuroimaging paradigms for studying RHE [13,15,20–22]. These methods can disentangle cross-level recursive from within-level iterative representations, track developmental acquisition [23], compare domain-specific and domain-general mechanisms [8,24], and link behavioral performance with neural substrates in both healthy individuals and clinical populations [25,26]. Together, these cross-domain studies have begun to reveal shared computational principles, as well as dissociations that speak to the specialization of cognitive and neural systems [8].

The unique contribution of this paper is to present the research program we have developed to bridge fractal geometry and the cognitive science of fractal cognition. We clarify the conceptual foundations and methodology of RHE, demonstrating how fractal geometry provides a principled framework for studying hierarchical cognition.

A secondary contribution is to exemplify how this framework can be used to address the theoretical question of whether RHE is a domain-general or domain-specific capacity. To this end, we review findings from our own research program and contrast them with results from more established paradigms to study hierarchical cognition, such as Artificial Grammar Learning (AGL), which have enabled precise modeling of hierarchical representations using symbolic, Bayesian, and neural network approaches.

The paper is organized as follows: Section 2 defines RHE and its relationship to recursion, iteration, hierarchical embedding, and fractal self-similarity; Section 3 reviews behavioral and neuroimaging paradigms across language, music, vision, and motor action; Section 4 summarizes key empirical findings and their theoretical implications for domain-generality versus domain-specificity while situating RHE alongside related approaches; and Section 5 proposes that future research can integrate behavioral and neural signatures with symbolic, neural, and Bayesian models to uncover the algorithmic and biological bases of recursive generativity.

By integrating theoretical frameworks, empirical findings, and computational approaches, we aim to provide a roadmap for understanding how the mind instantiates fractal cognition.

## 2. Conceptual Foundations

### 2.1. Defining Recursive Hierarchical Embedding (RHE)

There are two central challenges in the study of fractal cognition. First is the multiplicity of meanings across domains [27–29], and second is the difficulty of distinguishing it from superficially similar generative processes [7,30].

To reduce this ambiguity, our discussion within the cognitive sciences centers on recursive hierarchical embedding (RHE), a construct that integrates two distinct components: hierarchical embedding and recursion. Hierarchical embedding refers to the

nesting of one element (or set of elements) within another, dominant element. For instance, in English, embedding the noun “chocolate” within “cake” to form [[chocolate] cake] changes the meaning to a specific type of cake rather than a kind of chocolate [31–33]. Recursion, by contrast, describes a process in which the output of a function becomes the input for the same function [34]. This can be observed in the generation of natural numbers by the recursive function  $N_i = N_{i-1} + 1$ , which produces the infinite sequence {1, 2, 3, ...}.

When combined, recursion and hierarchical embedding enable the construction of hierarchies of potentially unbounded depth. For example, applying the recursive embedding rule  $NP \rightarrow [[NP] NP]$  allows “birthday” to be added to “chocolate cake,” producing [[[birthday] chocolate] cake], and continuing indefinitely.

Equally important is specifying the level of analysis most relevant for empirical work. Recursion can be discussed at three levels [7,29]: (1) the generative process, (2) the structure of the stimuli, and (3) the cognitive representations involved. The first two are often problematic—generative processes may be inaccessible, and hierarchical patterns can arise in nature without complex cognitive mechanisms, as in the branching of trees or bacterial colony growth. This paper, therefore, focuses on the third level: assessing whether participants can identify the regularities in hierarchical stimuli and use them to produce new embedding levels beyond the structures they have been given.

Defined in this way, RHE can be clearly distinguished from several superficially similar phenomena: non-recursive *iterative embedding* and *self-similarity*.

Non-recursive hierarchical embedding can create hierarchies, but without the property of self-application that allows indefinite extension with a single rule. For instance, with two distinct rules  $A \rightarrow B$  and  $B \rightarrow C$ , we could create the 3-level hierarchy  $A[B[C]]$ , but not generate further levels. *Iteration* allows applying the same operation repeatedly in a linear fashion, where each step is added at the same structural level. For example, repeating the operation  $B \rightarrow C$  adds an infinite number of Cs without generating new hierarchical levels:  $B[CCCCCCC]$ . Recursion, in contrast, applies a rule to its output in a way that can increase hierarchical depth. For example, the repeated application of the rule  $\alpha \rightarrow \alpha[\alpha]$  can generate the 3-level structure  $\alpha[\alpha[\alpha]]$  but also further levels.

*Self-similarity* refers to a structural property rather than a process: a coastline or snowflake may exhibit similar patterns at multiple scales [35–38], but such patterns can emerge through physical or biological growth mechanisms without any cognitive representation of recursive rules.

Finally, other paradigms in cognitive science, such as AGL, have tested for the ability to represent relationships that extend beyond adjacent transitions in serial stimuli [39]. These paradigms highlight two additional constructs that are worth distinguishing from RHE: *tree-like structures* and *long-distance dependencies*. *Tree-like structures* capture supra-linear organization in which an object can branch into multiple children, reflecting hierarchy but not necessarily RHE. One of the empirical signals of multiple branching superimposed onto linear structures (as in natural and artificial languages) is the capacity to represent dependencies between non-adjacent objects in a sequence (*long-distance dependencies*), for example,  $A_1-A_2$  in  $[A_1[B_1B_2]A_2]$ . These *long-distance dependencies* imply above-linear representations but not necessarily RHE.

By separating RHE from these related concepts, we can more precisely target the cognitive operations that allow humans to extend hierarchical structures beyond the given input. This distinction is critical for experimental design.

## 2.2. From Fractals to Bounded Cognition

In mathematics, fractals are generated by recursively applying a transformation to its own output, yielding self-similar patterns that, in theory, extend infinitely [4]. RHE in cognition follows the same generative principle—rules applied to their own output to create deeper hierarchical levels [7]—but crucially operates under cognitive constraints. Human recursive production and comprehension are bounded by working memory, attentional focus, and representational fidelity [17,18]. These limits cap the depth of recursion that can be realized, resulting in bounded generativity: the capacity to extend hierarchical structures across multiple levels, but only within the limits imposed by cognitive and perceptual systems.

Recognizing RHE as a bounded, fractal-like process clarifies its theoretical status. It captures the generative principle of self-similar expansion, while emphasizing that, unlike in mathematics, this expansion is inherently resource-limited in humans. This perspective offers a unifying lens for comparing recursive structures across language, music, vision, and action, while maintaining the focus on the mental operations that produce them, rather than on idealized, infinite structures. Testing the extent and limits of these operations requires carefully controlled experimental paradigms, to which we now turn.

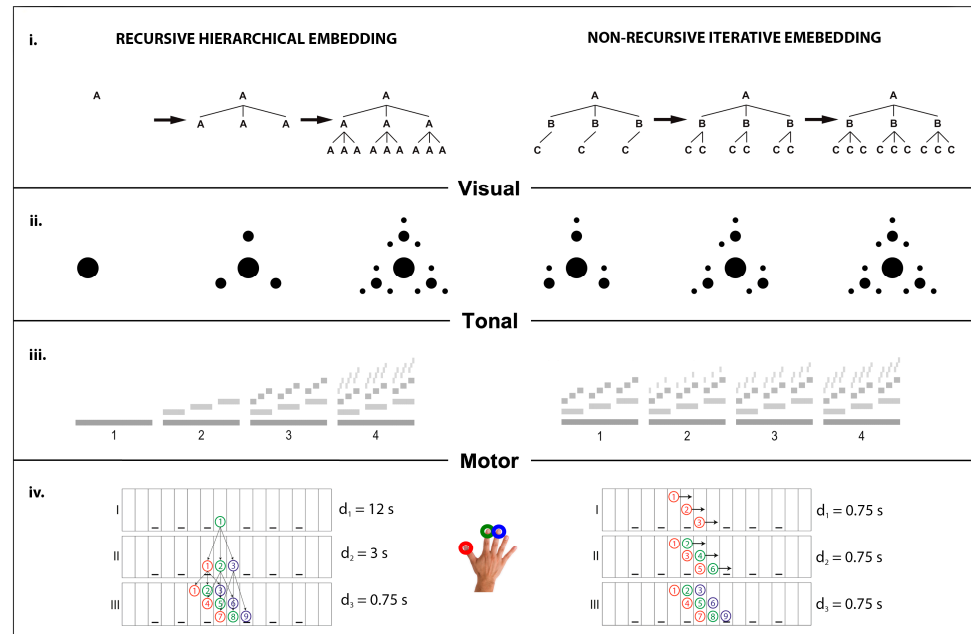
## 3. Experimental Paradigms for Studying RHE

The empirical investigation of recursive hierarchical embedding (RHE) encompasses a range of task formats and cognitive domains. Still, the underlying logic remains the same: to determine whether participants can infer and apply a generative rule that increases hierarchical depth beyond what is provided (Figure 1).

Broadly, RHE paradigms address two complementary questions. The first is, can participants do it? That is, can they recognize or generate recursive structures that extend beyond the provided input? This has been tested through discrimination and production tasks.

The second is, how do they do it? Which mechanisms, cognitive resources, and neural systems support recursive processing? This is probed with interference paradigms, lesion studies, neuropsychological evidence, and neuroimaging designs.

Figure 1 illustrates the general logic: recursive and iterative rules generate perceptually similar outcomes across vision, music, and motor domains, making it possible to compare conditions while controlling for low-level perceptual features. By combining behavioral, neuropsychological, and neuroimaging approaches, researchers can map both the presence of RHE and the mechanisms that underlie it.



**Figure 1.** General distinction between recursive hierarchical embedding and non-recursive iteration across domains. Abstract representation of the generative processes (i) and its application in the visual (ii), tonal (iii), and motor (iv) domains. Regardless of the generative process, the final structure is perceptually and motorically similar between recursive and iterative conditions, controlling for lower-level perceptual features.

### 3.1. Behavioral Methods

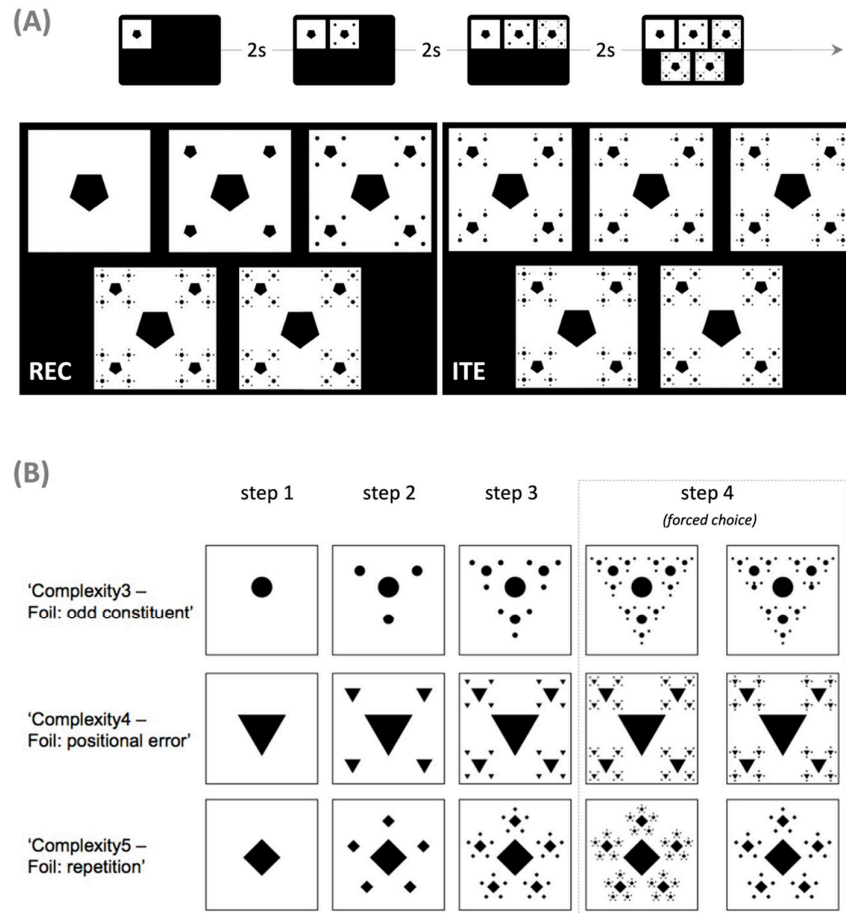
Behavioral investigations of RHE rely on two complementary strategies [40,41]: discrimination tasks, which test recognition of the correct recursive continuation [13,21], and production tasks, which assess the ability to generate it without external guidance [15,42]. Both formats operationalize the same underlying logic—determining whether participants can apply a learned generative rule to produce an additional hierarchical level beyond the given examples—but differ in the modality of response and the degree of external support.

In discrimination tasks, participants are shown the initial steps of a recursive process that generates successive hierarchical levels. Then, in a forced choice paradigm, they must choose between two continuations: one generated by recursive embedding and a foil generated by a different rule. Performance in RHE is then compared to that of non-recursive iteration, which generates the same pair of choice images. Because recursive and iterative rules can produce perceptually similar outputs [7,29] (Figure 1), this format is particularly well-suited for vision [13,20,25,26,43] and music [21,22], where foil stimuli can be carefully controlled to avoid reliance on low-level heuristics. Performance across different foil types (e.g., odd constituent, positional error, repetition) helps infer whether participants are genuinely applying recursive rules (Figure 2B). Using a large set of foils can help in inferring the internal representation from the choice behavioral pattern [44].

In production tasks, the recursive rule must be internally generated and extended [15,42]. In the motor domain, participants are asked to observe and repeat successive steps of a process that generates a motor fractal and then generate an additional step consistent with the previous hierarchical levels, without visual cues. Then, they deliver the final sequence motorically in a keyboard (examples with cues: Iteration <https://osf.io/gbd8j> (accessed on 30 September 2025), RHE <https://osf.io/6q5jm> (accessed on 30 September 2025); and without cues: Iteration <https://osf.io/7w5ps> (accessed on 30 September 2025)

and RHE <https://osf.io/qh93n> (accessed on 30 September 2025)). Production tasks also allow researchers to analyze error patterns for insights into representational fidelity and to test the limits of recursive processing when external cues are absent. Crucially, production paradigms render the use of perceptual heuristics less likely [40,41].

By combining discrimination and production tasks across domains, researchers can probe both the presence and flexibility of RHE. These behavioral paradigms form the foundation for more advanced designs (interference, neuropsychological, and neuroimaging studies), which shift the focus from whether participants can represent recursive structures to how they do so.



**Figure 2.** Experimental paradigm (adapted from [25]). (A) The presentation of the visual recursion/iteration task (REC, ITE) consisted of four steps, including the successive presentation of steps 1 to 3 at the top of the screen, followed by the two options for a forced choice at the bottom. The location of the correct image was randomized and counterbalanced (e.g., left in the ITE and right in the REC example provided). (B) Examples of fractals used in REC. There were different categories of ‘visual complexity’—3, 4, and 5—and different categories of foils. In ‘odd constituent’ foils, two elements within the whole hierarchy were misplaced; in ‘positional error’ foils, all elements within new hierarchical levels were internally consistent but inconsistent with the previous iterations; in ‘repetition’ foils, no additional iterative step was performed after the third iteration.

### 3.2. Interference, Neuropsychological, and Transfer Paradigms

*Interference paradigms* provide a direct way to test which cognitive resources support RHE by increasing the load on specific subsystems during recursive processing [45–47]. In these designs, participants perform an RHE task while simultaneously engaging in a

secondary task that taxes, e.g., verbal working memory, lexical retrieval, or attentional resources. Performance decrements in the primary task under dual-task conditions reveal which auxiliary systems are recruited to sustain RHE. By targeting the process rather than the outcome, dual-task paradigms help clarify how recursion is implemented in the cognitive system.

Complementary insights come from studies of individuals with focal brain damage or neurodevelopmental differences. *Lesion-to-symptom mapping* [48,49] and *neuropsychological testing* [50,51] can establish whether specific brain regions or cognitive systems are necessary for RHE, moving beyond correlational evidence. For example, if participants with impaired or absent linguistic resources can represent RHE in vision, this would point to partially independent representational pathways. Such cases would underscore that RHE is not a monolithic capacity but one that emerges from the interplay between shared scaffolds and domain-specific systems.

Although less directly investigated in our research program, transfer paradigms can also provide an additional perspective [52–54]. In these designs, participants are trained to apply a generative rule in one domain (e.g., constructing a visual hierarchy) and are then tested on their ability to use the same rule in another domain (e.g., producing a nested rhythmic structure in music). If performance improves due to domain transfer—for example, between visual and music RHE—this could be taken as evidence that the representational codes might overlap [55].

Together, these methods complement neuroimaging by testing whether the observed activations reflect necessary resources rather than mere correlates. However, assessing neural correlates can further constrain the hypothesis space regarding the cognitive and biological systems that instantiate RHE.

### 3.3. Neuroimaging Applications

The principles outlined above can be directly adapted to neuroimaging, with the central question shifting from whether participants can perform RHE behaviorally to whether we can isolate the neural circuits that represent the underlying generative process. To specifically target the generative process, task designs focus on the transition between one recursive step and the next—periods in which no new perceptual input is presented, but participants are required to apply the generative rule internally. This approach minimizes low-level sensory confounds and targets the cognitive operations of rule application, hierarchical updating, and anticipation of the resulting structure.

For example, after hearing an initial sequence and an intermediate recursive transformation (steps 1, 2, and 3 in Figure 1iii), participants might be cued to imagine the next step (step 4), predicting what it would sound like. In motor paradigms, this often involves covert motor imagery [15]—preparing the sequence without overt movement—whereas in sensory domains, it may involve silent rehearsal or mental imagery [22,43]. By capturing neural activity during these internally driven transitions, fMRI and related techniques can probe the planning and representational stages of RHE, rather than the mere perception of its outputs.

Taken together, these behavioral, neuropsychological, and neuroimaging paradigms provide a toolkit for probing both the presence of RHE and the mechanisms that sustain it (Table 1). Having outlined the logic of these methods, we now turn to the findings they have generated. Section 4 reviews how RHE has been tested across language, music, vision, and motor domains, and what these results reveal about the underlying cognitive architecture—most notably, whether recursion reflects a domain-general capacity or emerges from domain-specific systems.

**Table 1.** Experimental paradigms for studying recursive hierarchical embedding (RHE).

Paradigm Type	Domains Applied	Input/Output Structure	Dependent Measures	Advantages/Constraints
Discrimination tasks	Vision, Music, Motor	Participants choose the correct continuation versus foil alternatives	Accuracy, RT, foil-type sensitivity	Requires carefully matched controls (iteration, foils) for low-level heuristics
Production tasks	Motor, Language	Participants extend a sequence beyond the given input without external cues	Accuracy, temporal patterns	Stronger test of internal generativity; reduces reliance on perceptual heuristics
Interference/Dual-task	Vision	RHE combined with secondary task taxing working memory, verbal resources, etc.	$\Delta$ Acc/RT degradation under load	Identifies shared versus independent cognitive resources
Neuropsychological/Lesion	Vision, Language	RHE with brain damage or neurodevelopmental differences	Presence/absence of RHE, domain dissociations	Reveals necessity of systems
Neuroimaging paradigms	Vision	Discrimination/production tasks adapted for fMRI/EEG, focusing on recursive step transitions	BOLD/ERP responses during rule application	Identifies shared/distinct neural resources; Must isolate generative processing from perceptual input

#### 4. Empirical Findings on RHE Across Domains

Recursive hierarchical embedding (RHE) was first formally described in the domain of language [14,28,32,56,57], where syntactic recursion enables clauses to be nested within clauses to potentially unbounded depths. Influential accounts initially proposed that recursion might be a domain-specific property of the language faculty and that, if observed in other domains, it could reflect the use of linguistic mechanisms to implement recursive structure elsewhere [28,56]. As reviewed below, subsequent behavioral, developmental, and neuroimaging research has demonstrated that humans can also generate and manipulate recursively embedded structures in domains such as music, vision, and motor action [8].

The central theoretical question is whether these abilities rely on domain-general mechanisms or on domain-specific representational systems, such as syntactic templates in language or tonal hierarchies in music. Two broad possibilities have been proposed [24,58,59]: (i) that RHE depends on a single recursion mechanism capable of operating across representational formats [28,60], or (ii) that each domain contains its own specialized representational system, with apparent similarities arising from convergent solutions to common structural demands [58].

With the toolkit summarized in Table 1, we were able to test these hypotheses by applying RHE paradigms across multiple domains, task designs, and populations. As the following sections review, the evidence suggests a hybrid account: the acquisition of RHE appears to draw on partially shared cognitive resources, while long-term expertise or reduced access to domain-general symbolic resources increasingly shifts reliance onto domain-specific representational systems.

##### 4.1. Behavioral Evidence—In What Domains is RHE Represented?

Behavioral studies demonstrate that humans can learn and apply recursive hierarchical embedding (RHE) in the visual domain [13], alongside non-recursive iteration (ITE). Adults reliably discriminate RHE continuations well above chance and generalize rules, even in one-forced-choice paradigms. Crucially, visual RHE is not reducible to low-level perceptual heuristics: foils matched for entropy and spatial frequency, as well as aesthetic preference controls, rule out surface-based strategies. Dissociations with working memory further underscore that RHE depends on abstraction rather than storage. Whereas ITE correlates with spatial working memory, RHE correlates with recursive planning tasks such as the Tower of Hanoi, suggesting partially distinct

resources. Adults also show systematic learning curves (reaction times decreasing according to a power-law) and strategy differences: participants who report “analytic” reasoning in RHE are slower but more accurate, consistent with explicit rule extraction.

Developmental findings reveal that the ability to represent RHE emerges gradually [23]. By age 9–10, most children perform reliably on both RHE and ITE tasks (80% accuracy), whereas at 7–8 years, performance is largely at chance. Mastery of ITE appears to scaffold the later acquisition of RHE, as evidenced by task-order effects: completing ITE first improves subsequent RHE performance, but not vice versa. This mimics the developmental path in language in which conjunction ([second] and [red] ball) must be acquired before embedding ([second] red] ball) [30,57]. Younger children are also more sensitive to increased visual complexity and show difficulty rejecting certain foil types, consistent with an early local bias in hierarchical processing. Importantly, both RHE and ITE correlate with grammar comprehension (but only one-level hierarchical nesting [A[B]]) even after controlling for nonverbal intelligence, indicating shared resources for hierarchical processing.

In the auditory domain, both musicians and non-musicians can acquire RHE [21]. In discrimination tasks (Auditory Recursion Task, ART), performance was reliably above chance, though musicians achieved higher accuracy ( $\approx 80\%$ ) than non-musicians ( $\approx 71\%$ ). Importantly, ART performance correlated with visual RHE and Tower of Hanoi (when controlling for ITE), suggesting partially domain-general recursion abilities. Again, learning curves showed that RHE accuracy improved across trials, while ITE remained stable. Moreover, RHE performance was consistent across foil types, whereas ITE performance declined with specific foils. This indicates that RHE relied on abstract rule induction, whereas ITE was often driven by heuristics.

In the motor domain [42], RHE was demonstrated in discrimination and production tasks across two days of training. Adults discriminated RHE from foils above chance as early as Day 1 ( $\sim 71\%$ ), improving to  $\sim 86\%$  by Day 2, which is comparable to ITE ( $\sim 85\%$ ). However, generating novel RHE sequences without cues was significantly more challenging: by Day 2, ITE reached, on average,  $\sim 70\%$ , whereas RHE lagged at  $\sim 59\%$ , with only half of the participants achieving above-chance generation. Learning order mattered: those who practiced ITE first were more likely to succeed in RHE (with a performance boost of 22%). Sleep consolidation also improved RHE by 17%. Latency analyses showed longer cluster-completion times for RHE, consistent with higher hierarchical cluster sensitivity.

Table 2 schematically summarizes these findings. Taken together, the behavioral evidence suggests that humans can acquire and apply RHE across vision, audition, and motor production; however, RHE is consistently more demanding than ITE and emerges only under specific conditions. A recurring theme is that ITE is mastered first and then scaffolds RHE, as evident in the visual, motor, and language domains [23,25,30,42,57]. RHE requires more abstraction, exhibits steeper learning curves, and is less susceptible to surface heuristics [13,21]. At the same time, domain-specific constraints shape performance: auditory RHE is highly sensitive to expertise [21], motor RHE benefits from an additional day of training [42], and visual RHE is constrained by age [23]. Correlations across domains (e.g., with the Tower of Hanoi) highlight shared domain-general resources [13,21], but persistent asymmetries underscore the role of sensory and representational bottlenecks.

**Table 2.** Behavioral evidence for recursive hierarchical embedding (RHE) and iterative embedding (ITE) across domains. Accuracy values are group means as reported in the original studies. RHE generally required more abstraction, showed steeper learning curves, and was less amenable to surface heuristics than ITE. ITE tended to be acquired earlier and more easily, and often scaffolded later RHE performance. However, ITE is also more affected by lesions, atypical development, and interference. Abbreviations: R = RHE; I = ITE; WM = working memory; RT = response time; Disc = discrimination; Prod = production; d1/d2 = Day 1/Day 2; I→R = iteration task performed before recursion (order effect); TD = typically developing controls; Drift = rate of evidence accumulation (decision efficiency); Boundary = amount of evidence required for a decision (cautiousness); ToH = tower of Hanoi—a recursive planning task; SpWM = spatial WM (corsi); VisWM = visual WM (delayed match to sample); ↑ = increase; ↓ = decrease.

Domain-Population	Paradigm	RHE/ITE (Mean Acc)	Performance Patterns
Vision—Adults [13]	Discrimination	R: 82%, I: 92%	RHE: ↑ learning curves; slow strategies more accurate; correlated with ToH; ITE correlated with SpWM.
Vision—Children [23]	Discrimination	7–8 yrs: R: 59%, I: 62% 9–10 yrs: R: 80%, I: 78% Order I→R (ΔR: ↑24%)	RHE: ↑ learning curves; scaffolded by ITE. RHE & ITE correlated with 1-level linguistic embedding (controlling for IQ).
Audition— Non-musicians & Musicians [21]	Discrimination	Non-mus: R: 71%, I: 70% Musicians: R: 84%	RHE: ↑ learning curves; strong expertise effect; correlated with visual RHE and ToH.
Motor—Adults [42]	Discrimination & production (2-day training)	Disc (d2): R: 86%, I: 85% Prod (d2): R: 59%, I: 70% Prod I→R (ΔR: ↑22%) Prod d1→d2 (ΔR: ↑17%)	RHE: ↑ learning curves; Production scaffolded by ITE & boosted by sleep (d2); clustered temporal pattern.
Vision— Lesion patients [25]	Discrimination; Lesion-symptom mapping; Drift diffusion model	R: 80%, I: 82% Order I→R (ΔR: ↑20%)	RHE (v ITE): correlated with 2-level linguistic embedding; Scaffolded by ITE; posterior lesions ↓ drift; frontal lesions ↓ boundary.
Vision—Minimally verbal autistic child [26]	Discrimination	R: 72.2%, I: 64.8% (both pass)	RHE intact despite no linguistic syntax. ITE more fragile.
Vision—Adults [24,61]	Discrimination dual task: - VerbalWM [24] - VisualWM [61] - Lexical retrieval [61] - Serial arithmetic [61]	Baseline [61]: R: 83%, I: 87% verbalWM [24]: R: 86% visWM [61]: R: 85%, I: 83% Lexical [61]: R: 75%, I: 70% Arith. [61]: R: 78%, I: 70%	lexical/arithmic interference ↓ drift and boundary. RHE less affected than ITE; no verbal or visual WM interference.

#### 4.2. Evidence from Lesions, Interference, and Atypical Development

RHE is available across vision, audition, and motor domains, but the key question is how it is implemented: does it rely on a shared resource for hierarchical embedding or on domain-specific implementations? Evidence from lesions, interference, and atypical development helps constrain this debate.

On the one hand, several findings suggest shared resources. Stroke patients with left posterior temporal damage were more impaired in visual RHE than ITE and more impaired in parsing linguistic embeddings with two levels ([A1[B1[C1C2]B2]A2]) than structures with a single level ([A1[B1XXB2]A2]). This indicates a common bottleneck for the processing of nested hierarchical information [25]. Their performance showed reduced evidence accumulation (lower drift rates) for RHE, while frontal lesions instead produced cognitive control problems, consistent with lower boundary separation (reduced cautiousness).

On the other hand, RHE can also operate independently of language and verbal working memory. In minimally verbal autism with hyperlexia, an 11-year-old child performed at typical levels on visual RHE despite lacking recursive syntax, demonstrating that RHE in vision does not require linguistic recursion [26]. Likewise, dual-task paradigms with articulatory suppression (backwards digit span) did not impair visual RHE, ruling out a reliance on phonological rehearsal [24].

More targeted interference studies clarify this asymmetry [13]. When secondary tasks taxed serial symbolic resources—such as lexical retrieval and serial arithmetic—both RHE and ITE performance declined, but iteration was more susceptible. Drift–diffusion modeling revealed that serial symbolic load reduced drift rates (slower evidence accumulation) and lowered thresholds (less cautious responding). Crucially, recursion and iteration diverged in response dynamics: under serial symbolic load, RHE accuracy was less affected and responses became faster, consistent with compensatory strategies, whereas ITE response times slowed down and accuracy degraded. By contrast, a visual delayed match-to-sample interference task did not impair accuracy.

Taken together, these findings suggest a hybrid implementation of RHE. It preferentially engages symbolic and language-like resources (as evidenced by lexical and arithmetic interference, as well as lesion data). Still, it can also be supported by alternative pathways that make RHE more resilient than ITE under load. This dual profile explains why RHE is correlated across domains but can still manifest in atypical populations with impaired recursive language.

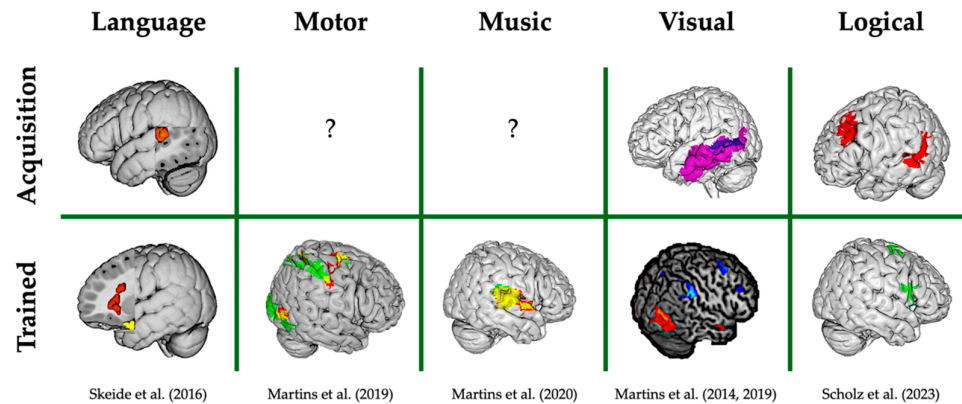
As we will review in the next section, neuroimaging evidence further supports this dual profile.

#### *4.3. Neural Bases of Recursion Across Domains*

Neuroimaging studies suggest a dynamic balance between shared and domain-specific circuitry (Figure 3). When participants are extensively trained in a single domain, RHE vs. ITE contrasts reveal domain-specific temporal and parietal activations: for example, the visual ventral stream in visual recursion [20]; the posterior superior temporal gyrus (pSTG) in musical recursion [22], involved in tone imagery; and parietal–premotor circuits in motor recursion [15], involved in motor planning. This suggests that long-term expertise engages modality-specific representational systems. However, when untrained participants are tested [25,62], different domains elicit a more overlapping network, including the left inferior frontal gyrus (IFG) and posterior superior temporal sulcus (pSTS)—areas classically linked to language syntax [63,64]. The pSTS is particularly relevant when acquiring rules governing hierarchical relations (Figure 3). This pattern suggests a shared “scaffolding” for initial acquisition, followed by domain-specific consolidation with expertise.

Interestingly, the contrast between RHE and ITE also reveals that RHE relies more on the recruitment of the default mode network, associated with internal models of processing, whereas ITE recruits the fronto-parietal network, related to effortful processing of sensorial information [43]. The pattern is consistent with the behavioral evidence that RHE representations are more abstract and less affected by conditions that affect general cognitive resources [13,21,61].

Taken together, these findings indicate that RHE acquisition recruits a partially shared set of cognitive and neural resources across domains [60], which later consolidate into domain-specific circuits as expertise develops [58]. This dual pattern—shared early resources, domain-specific consolidation—provides a framework for understanding how humans flexibly acquire recursive rules in multiple representational formats while also developing specialized neural substrates for their long-term use.



**Figure 3.** Commonalities and differences in the neural processing of recursive hierarchical embedding (RHE) across domains. During acquisition, the left posterior temporal cortex (including pSTS and adjacent regions) is consistently engaged across modalities. With extensive training, RHE processing shifts toward increasingly domain-specialized regions: ventral visual areas for vision, posterior superior temporal regions for music, and premotor–parietal circuits for motor behavior. This pattern suggests a shared scaffolding mechanism during learning, followed by domain-specific consolidation with expertise (figure adapted from [8,15,20,22,25,62,65]).

#### 4.4. Insights from Artificial Grammar Learning and Computational Approaches

AGL paradigms have long been the primary method for probing the human ability to represent above-linear relationships in sequential stimuli [66]. Unlike simple statistical learning of adjacent transitions, AGL tasks assess whether participants can detect and generalize hierarchical dependencies, such as nested or center-embedded structures.

Computationally, these tasks are well formalized through different classes of grammars. Finite-state grammars generate local patterns such as  $(AB)^n$ , whereas context-free grammars generate center-embedded patterns such as  $A^nB^n$ , e.g.,  $[A1[B1[C1C2]B2]A2]$ . This formal distinction allows researchers to infer the nature of participants' internal representations. One influential hypothesis is that learners recruit an external memory device—a push-down stack—to parse nonadjacent dependencies, enabling them to keep “open” elements until closure is encountered [28,56,67,68].

Beyond formal grammars, quantitative models have clarified how hierarchical dependencies are acquired and processed. Bayesian learners—framed as hypothesis testers over generative grammars—capture graded generalization and the preference for recursively embeddable hypotheses when the data support them [69]. Usage-based/Bayesian accounts linked to the “now-or-never” bottleneck [17,70,71] explain why learners rely on chunking and probabilistic abstraction early, yet can still converge on recursive structure as evidence accumulates. Neural and reinforcement-learning models [70], especially recurrent neural networks (RNNs), reproduce human-like performance profiles—including the classic collapse beyond  $\sim 3$ – $4$  embeddings—thus modeling performance limits rather than demonstrating an internal recursive rule [70,72]. Recent Bayesian mixture models [68] further suggest that learners do not universally deploy a push-down stack; instead, they mix recursive and associative heuristics depending on task demands and memory constraints. Together, these approaches illuminate both acquisition dynamics and the mechanistic bottlenecks that shape apparent recursion in behavior.

Armed with this toolkit, AGL has been applied across diverse populations. Adults and older children readily acquire  $A^nB^n$  dependencies and generalize to longer strings [73]. Children's performance, however, is tightly constrained by working memory capacity, with success correlating strongly with digit span. Illiterate Amazonians succeed

despite lacking formal schooling, demonstrating that recursive competence does not depend on literacy. Monkeys and ravens can master AnBn structures [73,74], but only after extensive training, and their acquisition trajectories are best explained by associative or ordinal heuristics [75]. The contrast with humans' relatively rapid and spontaneous generalization highlights the importance of working memory resources as well as the limits of nonhuman cognition [76].

Neuroimaging results converge on a left frontotemporal circuit, with an important evolutionary refinement [64,77–79]. Posterior superior/middle temporal cortex (pSTS/pMTG) plays a central role in structure building—particularly in untrained participants—while BA44/45 in left IFG is increasingly engaged with training and expertise [65,80,81]. Recent evolutionary syntheses [64] argue that BA44 underwent an anterior expansion in humans, enabling language-specific syntactic hierarchy building to occupy anterior BA44, while posterior BA44 supports action. This suggests that these are adjacent but non-overlapping systems embedded in distinct networks. Other accounts indicate that left BA44/45 may be generally involved in controlling domain-specific hierarchical representations from other brain regions, including posterior temporal cortex [82]. Importantly, both accounts are consistent with the idea that similar computational demands across domains might recruit adjacent but non-overlapping circuits [58].

Taken together, AGL, cross-species studies, neuroimaging, and computational modeling converge on a hybrid picture: recursion depends on domain-general resources, but with training, it may consolidate into domain-specific representational systems.

While in some ways similar to AGL, RHE paradigms extend beyond sequential stimuli, as they are also applied to structures extending spatially, such as fractals. RHE also provides stricter experimental controls, holding perceptual outputs constant across recursion and iteration. This reduces the likelihood that apparent recursive performance can be explained by surface heuristics, working memory, or extensive associative training. At the same time, however, RHE research has not yet benefited from the rigorous computational modeling that characterizes AGL, which would allow more precise inferences about the internal strategies participants deploy. Developing such quantitative models represents a crucial step forward.

## 5. Open Questions and Future Directions

Despite substantial progress, important challenges remain for understanding recursive hierarchical embedding (RHE). These challenges arise at both the behavioral and neural levels.

### 5.1. Challenges in Studying RHE

At the behavioral level, RHE is difficult to isolate from confounding factors. In AGL, increasing embedding depth almost inevitably covaries with working memory load, perceptual discrimination difficulty, and attentional switching, making it unclear whether performance drops reflect genuine limits on recursive processing or the exhaustion of auxiliary resources [68,73,76]. While RHE seems more robust against verbal, spatial, and visual WM manipulations (Table 2), these can still play a significant role.

Moreover, discrimination and production tasks likely engage overlapping but non-identical systems: discrimination allows external cues to scaffold performance [40,41], whereas production requires participants to internally generate the recursive step without external guidance. Cross-domain comparability further complicates interpretation, as the same generative rules can differ in representational granularity and familiarity. Visual fractals are relatively easy to recognize, auditory recursion requires tonal imagery, which is more accessible to musicians, and motor recursion remains the most challenging, even after extensive training. This is reflected in the accuracy levels (Table 2).

At the neuroimaging level, progress is constrained by the difficulty of inferring mechanisms from correlates. Current findings consistently implicate posterior temporal regions in structure building and frontal regions in control and sequencing [25,62,65,82]; however, disentangling representational functions from attentional or decision-related processes remains a challenge. Additionally, the contributions of subcortical [62] and network-level structures [83] are poorly understood. Default mode recruitment in recursion and fronto-parietal engagement in iteration point to large-scale network dynamics, but these patterns remain underexplored [43].

### 5.2. Towards Mechanistic Modeling

A promising way forward is to adopt the modeling toolkit that has proven fruitful in Artificial Grammar Learning (AGL). In AGL, symbolic grammars, Bayesian inference models, and recurrent neural networks (RNNs) provide precise hypotheses about the internal strategies participants deploy. Such models capture not only success but also systematic errors—for example, why both humans and RNNs collapse beyond three to four levels of embedding [18,70], or why Bayesian learners shift between recursive and associative strategies depending on task demands [68].

For RHE, progress has begun with drift diffusion modeling [25,61], which reveals how RHE and ITE differ in evidence accumulation (drift rate) and decision thresholds (boundary separation). Yet this approach remains descriptive. The next step is to develop richer mechanistic models that generate specific predictions about how recursion should be affected by symbolic interference, posterior temporal lesions, or atypical development. These models can then be tested against the behavioral signatures already identified in Table 2: order effects, asymmetric vulnerability of ITE, and resilience of RHE under certain kinds of interference. In this way, computational modeling can bridge fine-grained behavioral results with neural mechanisms, moving beyond correlation to causal inference.

### 5.3. Expanding the Empirical Program

Currently, this modeling agenda has been primarily pursued in the visual domain. Extending it to motor and auditory recursion is a critical next step. For example, motor production tasks provide a stronger test of generativity but remain underexplored in terms of modeling; auditory recursion is especially sensitive to expertise and would provide a natural test case for how domain-specific scaffolding interacts with domain-general mechanisms. Neuroimaging evidence of recursion *acquisition*, summarized in Figure 3, is also disproportionately focused on vision. Systematic extensions to music and motor domains would provide the missing pieces of a full cross-domain profile.

Equally important is broadening the populations tested. Developmental and cross-cultural studies are needed to assess the universality of RHE, and clinical or neuropsychological groups can reveal necessary systems through selective impairments. Transfer designs—training recursion in one modality (e.g., visual fractals) and testing in another (e.g., nested musical rhythms)—would provide a strong test of whether shared scaffolds underlie cross-domain recursion, or whether success depends on domain-specific expertise.

### 5.4. Toward a Unified Science of Fractal Cognition

Taken together, these challenges and opportunities point toward a unified research program. Progress will require integrating behavioral paradigms that isolate recursion from confounds, mechanistic models that formalize the computations involved, and neurobiological evidence that specifies their implementation. The goal is not simply to

show that humans can represent RHE across domains, but to predict when and how recursive competence emerges, collapses, or transfers.

This integrative framework would allow fractal geometry to serve not only as a metaphor but also as a quantitative bridge: linking mathematical self-similarity, cognitive theories of hierarchical generativity, and the neural architectures that implement bounded recursion. By aligning behavioral signatures with algorithmic models and neural dynamics, the field can move toward a mechanistic understanding of how humans instantiate fractal cognition.

## 6. Conclusions

This paper advances a research program in fractal cognition through five core contributions. First, we introduce the theoretical foundations of recursive hierarchical embedding (RHE), clarifying its distinction from iteration, self-similarity, and related constructs. Second, we provide a methodological toolkit—behavioral, neuropsychological, and neuroimaging paradigms—for isolating recursive processing. Third, we illustrate the application of this toolkit across language, music, vision, and motor action to detect the ability to represent RHE and evaluate evidence for domain-general versus domain-specific cognitive resources. Fourth, across domains and populations, the evidence suggests that RHE is more challenging than iteration and emerges later, but is also more resilient to interference and clinical impairment—indicating compensatory mechanisms. We argue that future progress hinges on integrating behavioral and neural signatures across populations and paradigms with precise quantitative models. These contributions provide a roadmap for moving from descriptive demonstrations to a mechanistic science of fractal cognition, where recursive generativity is not only documented but explained and predicted.

**Funding:** Open Access funding provided by the University of Vienna under the KEMÖ consortium.

**Data Availability Statement:** No new data were created or analyzed in this study. Data sharing is not applicable to this article.

**Acknowledgments:** During the preparation of this manuscript, the author used GPT-5 to improve the writing process. The author has reviewed and edited the output and takes full responsibility for the content of this publication.

**Conflicts of Interest:** The author declares no conflicts of interest.

## References

1. Eke, A.; Herman, P.; Kocsis, L.; Kozak, L. Fractal Characterization of Complexity in Temporal Physiological Signals. *Physiol. Meas.* **2002**, *23*, R1.
2. Halsey, T.C.; Jensen, M.H.; Kadanoff, L.P.; Procaccia, I.; Shraiman, B.I. Fractal Measures and Their Singularities: The Characterization of Strange Sets. *Phys. Rev. A* **1986**, *33*, 1141.
3. Mandelbrot, B.B. *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk. Selecta Volume E*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2013; ISBN 1-4757-2763-1.
4. Mandelbrot, B.B. *The Fractal Geometry of Nature*; Freeman: New York, NY, USA; 1982.
5. Fisher, Y. Fractal Image Compression. *Fractals* **1994**, *2*, 347–361.
6. Turcotte, D.L. *Fractals and Chaos in Geology and Geophysics*; Cambridge University Press: Cambridge, UK, 1997; ISBN 0-521-56733-5.
7. Martins, M.D. Distinctive Signatures of Recursion. *Philos. Trans. R. Soc. B Biol. Sci.* **2012**, *367*, 2055–2064. <https://doi.org/10.1098/rstb.2012.0097>.

8. Martins, M.D. Cognitive and Neural Representations of Fractals in Vision, Music, and Action. In *The Fractal Geometry of the Brain*; Di Ieva, A., Ed.; Advances in Neurobiology; Springer International Publishing: Cham, Switzerland, 2024; pp. 935–951, ISBN 978-3-031-47606-8.
9. Chomsky, N. *Interfaces + Recursion = Language?: Chomsky's Minimalism and the View from Syntax-Semantics*; De Gruyter Mouton: Berlin, Germany, 2007; ISBN 978-3-11-020755-2.
10. Hulst, H. van der Recursion and Human Language. In *Recursion and Human Language*; De Gruyter Mouton: Berlin, Germany, 2010, ISBN 978-3-11-021925-8.
11. Jackendoff, R.; Lerdahl, F. The Capacity for Music: What Is It, and What's Special about It? *Cognition* **2006**, *100*, 33–72. <https://doi.org/10.1016/j.cognition.2005.11.005>.
12. Lerdahl, F.; Jackendoff, R.S. *A Generative Theory of Tonal Music, Reissue, with a New Preface*; MIT Press: Boston, MA, USA, 1996, ISBN 0-262-26091-3.
13. Martins, M.D.; Martins, I.P.; Fitch, W.T. A Novel Approach to Investigate Recursion and Iteration in Visual Hierarchical Processing. *Behav. Res.* **2016**, *48*, 1421–1442. <https://doi.org/10.3758/s13428-015-0657-1>.
14. Jackendoff, R. Précis of foundations of language: Brain, meaning, grammar, evolution. *Behav. Brain Sci.* **2003**, *26*, 651–665. <https://doi.org/10.1017/S0140525X03000153>.
15. Martins, M.D.; Bianco, R.; Sammler, D.; Villringer, A. Recursion in Action: An fMRI Study on the Generation of New Hierarchical Levels in Motor Sequences. *Hum. Brain Mapp.* **2019**, *40*, 2623–2638. <https://doi.org/10.1002/hbm.24549>.
16. Vicari, G.; Adenzato, M. Is Recursion Language-Specific? Evidence of Recursive Mechanisms in the Structure of Intentional Action. *Conscious. Cogn.* **2014**, *26*, 169–188. <https://doi.org/10.1016/j.concog.2014.03.010>.
17. Christiansen, M.H.; Chater, N. The Now-or-Never Bottleneck: A Fundamental Constraint on Language. *Behav. Brain Sci.* **2016**, *39*, e62.
18. Cowan, N. The Magical Mystery Four: How Is Working Memory Capacity Limited, and Why? *Curr. Dir. Psychol. Sci.* **2010**, *19*, 51–57.
19. Fitch, W.T.; Friederici, A.D. Artificial Grammar Learning Meets Formal Language Theory: An Overview. *Philos. Trans. R. Soc. B Biol. Sci.* **2012**, *367*, 1933–1955.
20. Martins, M.D.; Fischmeister, F.P.; Puig-Waldmüller, E.; Oh, J.; Geißler, A.; Robinson, S.; Fitch, W.T.; Beisteiner, R. Fractal Image Perception Provides Novel Insights into Hierarchical Cognition. *NeuroImage* **2014**, *96*, 300–308. <https://doi.org/10.1016/j.neuroimage.2014.03.064>.
21. Martins, M.D.; Gingras, B.; Puig-Waldmueller, E.; Fitch, W.T. Cognitive Representation of “Musical Fractals”: Processing Hierarchy and Recursion in the Auditory Domain. *Cognition* **2017**, *161*, 31–45. <https://doi.org/10.1016/j.cognition.2017.01.001>.
22. Martins, M.D.; Fischmeister, F.P.S.; Gingras, B.; Bianco, R.; Puig-Waldmueller, E.; Villringer, A.; Fitch, W.T.; Beisteiner, R. Recursive Music Elucidates Neural Mechanisms Supporting the Generation and Detection of Melodic Hierarchies. *Brain Struct. Funct.* **2020**, *225*, 1997–2015. <https://doi.org/10.1007/s00429-020-02105-7>.
23. Martins, M.D.; Laaha, S.; Freiberger, E.M.E.M.; Choi, S.; Fitch, W.T. How Children Perceive Fractals: Hierarchical Self-Similarity and Cognitive Development. *Cognition* **2014**, *133*, 10–24. <https://doi.org/10.1016/j.cognition.2014.05.010>.
24. Martins, M.D.; Muršič, Z.; Oh, J.; Fitch, W.T. Representing Visual Recursion Does Not Require Verbal or Motor Resources. *Cogn. Psychol.* **2015**, *77*, 20–41. <https://doi.org/10.1016/j.cogpsych.2015.01.004>.
25. Martins, M.D.; Krause, C.; Neville, D.A.; Pino, D.; Villringer, A.; Obrig, H. Recursive Hierarchical Embedding in Vision Is Impaired by Posterior Middle Temporal Gyrus Lesions. *Brain* **2019**, *142*, 3217–3229. <https://doi.org/10.1093/brain/awz242>.
26. Rosselló, J.; Celma-Mirallas, A.; Martins, M.D. Visual Recursion without Recursive Language? A Case Study of a Minimally Verbal Autistic Child. *Front. Psychiatry* **2025**, *16*, 1540985.
27. Fitch, W.T. Three Meanings of “Recursion”: Key Distinctions for Biolinguistics. In *The Evolution of Human Language*; Larson, R.K., Déprez, V., Yamakido, H., Eds.; Cambridge University Press: Edinburgh, UK, 2010; pp. 73–90, ISBN 978-0-521-51645-7.
28. Hauser, M.D.; Chomsky, N.; Fitch, W.T. The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science* **2002**, *298*, 1569–1579. <https://doi.org/10.1126/science.298.5598.1569>.
29. Lobina, D.J. Recursion and the Competence/Performance Distinction in AGL Tasks. *Lang. Cogn. Process.* **2011**, *26*, 1563–1586. <https://doi.org/10.1080/01690965.2011.560006>.
30. Roeper, T. The Acquisition of Recursion: How Formalism Articulates the Child's Path. *Biolinguistics* **2011**, *5*, 057–086.
31. Carnie, A. *Syntax: A Generative Introduction*; John Wiley & Sons: Hoboken, NJ, USA, 2021, ISBN 1-119-56931-1.
32. Chomsky, N. *Syntactic Structures*; Mouton de Gruyter: Berlin, Germany, 2002, ISBN 3-11-017279-8.
33. Jackendoff, R. *The Architecture of the Language Faculty*; MIT Press: New York, NY, USA, 1997, ISBN 0-262-60025-0.

34. Odifreddi, P. *Classical Recursion Theory: The Theory of Functions and Sets of Natural Numbers*; Elsevier: Amsterdam, The Netherlands, 1992, Volume 125, ISBN 0-08-088659-0.
35. Koch, H.V. Sur Une Courbe Continue sans Tangente, Obtenue Par Une Construction Géométrique Élémentaire. *Ark. Mat. Astron. Och Fys.* **1904**, *1*, 681–704.
36. Mandelbrot, B.B. How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension. *Science* **1967**, *156*, 636–638.
37. Matsuyama, T.; Matsushita, M. Fractal Morphogenesis by a Bacterial Cell Population. *Crit. Rev. Microbiol.* **1993**, *19*, 117–135. <https://doi.org/10.3109/10408419309113526>.
38. Meinhardt, H. Models of Biological Pattern Formation: From Elementary Steps to the Organization of Embryonic Axes. In *Current Topics in Developmental Biology*; Schnell, S., Maini, P.K., Newman, S.A., Newman, T.J., Eds.; Multiscale Modeling of Developmental Systems; Academic Press: Cambridge, MA, USA, 2008, Volume 81, pp. 1–63.
39. Udden, J.; Martins, M.D.; Zuidema, W.; Fitch, W.T. Hierarchical Structure in Sequence Processing: How Do We Measure It and What's the Neural Implementation? *Top. Cogn. Sci.* **2019**, *12*, 910–924. <https://doi.org/10.1111/tops.12442>.
40. Dedhe, A.M.; Clatterbuck, H.; Piantadosi, S.T.; Cantlon, J.F. Origins of Hierarchical Logical Reasoning. *Cogn. Sci.* **2023**, *47*, 13250. <https://doi.org/10.1111/cogs.13250>.
41. Dedhe, A.M.; Piantadosi, S.T.; Cantlon, J.F. Cognitive Mechanisms Underlying Recursive Pattern Processing in Human Adults. *Cogn. Sci.* **2023**, *47*, e13273. <https://doi.org/10.1111/cogs.13273>.
42. Martins, M.D.; Bergmann, Z.V.; Bianco, R.; Sammler, D.; Villringer, A. Acquisition and Utilization of Recursive Rules in Motor Sequence Generation. *Cogn. Sci.* **2025**, *49*, e70108. <https://doi.org/10.1111/cogs.70108>.
43. Fischmeister, F.P.; Martins, M.D.; Beisteiner, R.; Fitch, W.T. Self-Similarity and Recursion as Default Modes in Human Cognition. *Cortex* **2017**, *97*, 183–201.
44. Ravignani, A.; Westphal-Fitch, G.; Aust, U.; Schlumpp, M.M.; Fitch, W.T. More than One Way to See It: Individual Heuristics in Avian Visual Computation. *Cognition* **2015**, *143*, 13–24. <https://doi.org/10.1016/j.cognition.2015.05.021>.
45. Brown, J. Some Tests of the Decay Theory of Immediate Memory. *Q. J. Exp. Psychol.* **1958**, *10*, 12–21.
46. Frank, M.C.; Fedorenko, E.; Lai, P.; Saxe, R.; Gibson, E. Verbal Interference Suppresses Exact Numerical Representation. *Cogn. Psychol.* **2012**, *64*, 74–92.
47. Peterson, L.; Peterson, M.J. Short-Term Retention of Individual Verbal Items. *J. Exp. Psychol.* **1959**, *58*, 193.
48. Bates, E.; Wilson, S.M.; Saygin, A.P.; Dick, F.; Sereno, M.I.; Knight, R.T.; Dronkers, N.F. Voxel-Based Lesion-Symptom Mapping. *Nat. Neurosci.* **2003**, *6*, 448–450. <https://doi.org/10.1038/nn1050>.
49. Rorden, C.; Karnath, H.-O.; Bonilha, L. Improving Lesion-Symptom Mapping. *J. Cogn. Neurosci.* **2007**, *19*, 1081–1088. <https://doi.org/10.1162/jocn.2007.19.7.1081>.
50. Shallice, T. *From Neuropsychology to Mental Structure*; From Neuropsychology to Mental Structure; Cambridge University Press: New York, NY, USA, 1988; pp. xv, 462, ISBN 978-0-521-30874-8.
51. Caramazza, A. On Drawing Inferences about the Structure of Normal Cognitive Systems from the Analysis of Patterns of Impaired Performance: The Case for Single-Patient Studies. *Brain Cogn.* **1986**, *5*, 41–66.
52. Gentner, D. Bootstrapping the Mind: Analogical Processes and Symbol Systems. *Cogn. Sci.* **2010**, *34*, 752–775.
53. Ruiz, F.J.; Luciano, C. Cross-domain Analogies as Relating Derived Relations among Two Separate Relational Networks. *J. Exp. Anal. Behav.* **2011**, *95*, 369–385.
54. Singley, M.K.; Anderson, J.R. *The Transfer of Cognitive Skill*; Harvard University Press: Cambridge, MA, USA 1989; ISBN 0-674-90340-4.
55. Bouvet, L.; Rousset, S.; Valdois, S.; Donnadieu, S. Global Precedence Effect in Audition and Vision: Evidence for Similar Cognitive Styles across Modalities. *Acta Psychol.* **2011**, *138*, 329–335. <https://doi.org/10.1016/j.actpsy.2011.08.004>.
56. Fitch, W.T.; Hauser, M.D.; Chomsky, N. The Evolution of the Language Faculty: Clarifications and Implications. *Cognition* **2005**, *97*, 179–210. <https://doi.org/10.1016/j.cognition.2005.02.005>.
57. Roeper, T. *The Prism of Grammar: How Child Language Illuminates Humanism*; A Bradford Book: Cambridge, MA, USA, 2007, ISBN 978-0-262-18252-2.
58. Fedorenko, E.; Shain, C. Similarity of Computations Across Domains Does Not Imply Shared Implementation: The Case of Language Comprehension. *Curr. Dir. Psychol. Sci.* **2021**, *30*, 526–534. <https://doi.org/10.1177/096372142111046955>.
59. Fedorenko, E.; Patel, A.; Casasanto, D.; Winawer, J.; Gibson, E. Structural Integration in Language and Music: Evidence for a Shared System. *Mem. Cogn.* **2009**, *37*, 1–9. <https://doi.org/10.3758/MC.37.1.1>.
60. Fadiga, L.; Craighero, L.; D'Ausilio, A. Broca's Area in Language, Action, and Music. *Ann. N. Y. Acad. Sci.* **2009**, *1169*, 448–458.

61. Martins, M.; Cook, D.; Villringer, A. Recursion Beyond Language: Lexical and Arithmetic Interference in Visual Hierarchical Embedding. 2025. Available online: [https://osf.io/preprints/psyarxiv/rg9hw\\_v1](https://osf.io/preprints/psyarxiv/rg9hw_v1) (accessed on 6 October 2025).
62. Scholz, R.; Villringer, A.; Martins, M.D. Distinct Hippocampal and Cortical Contributions in the Representation of Hierarchies. *eLife* **2023**, *12*, e87075. <https://doi.org/10.7554/eLife.87075>.
63. Friederici, A.D. Hierarchy Processing in Human Neurobiology: How Specific Is It? *Phil. Trans. R. Soc. B* **2020**, *375*, 20180391. <https://doi.org/10.1098/rstb.2018.0391>.
64. Friederici, A.D. Evolutionary Neuroanatomical Expansion of Broca's Region Serving a Human-Specific Function. *Trends Neurosci.* **2023**, *46*, 786–796. <https://doi.org/10.1016/j.tins.2023.07.004>.
65. Skeide, M.A.; Friederici, A.D. The Ontogeny of the Cortical Language Network. *Nat. Rev. Neurosci.* **2016**, *17*, 323–332. <https://doi.org/10.1038/nrn.2016.23>.
66. Fitch, W.T.; Hauser, M.D. Computational Constraints on Syntactic Processing in a Nonhuman Primate. *Science* **2004**, *303*, 377–380. <https://doi.org/10.1126/science.1089401>.
67. Uddén, J.; Ingvar, M.; Hagoort, P.; Petersson, K.M. Implicit Acquisition of Grammars with Crossed and Nested Non-adjacent Dependencies: Investigating the Push-down Stack Model. *Cogn. Sci.* **2012**, *36*, 1078–1101.
68. Ferrigno, S.; Cheyette, S.J.; Carey, S. Do Humans Use Push-Down Stacks When Learning or Producing Center-Embedded Sequences? *Cogn. Sci.* **2025**, *49*, e70112. <https://doi.org/10.1111/cogs.70112>.
69. Perfors, A.; Tenenbaum, J.B.; Regier, T. The Learnability of Abstract Syntactic Principles. *Cognition* **2011**, *118*, 306–338. <https://doi.org/10.1016/j.cognition.2010.11.001>.
70. Christiansen, M.H.; Chater, N. Toward a Connectionist Model of Recursion in Human Linguistic Performance. *Cogn. Sci.* **1999**, *23*, 157–205. [https://doi.org/10.1207/s15516709cog2302\\_2](https://doi.org/10.1207/s15516709cog2302_2).
71. Christiansen, M.H.; Chater, N. Language as Shaped by the Brain. *Behav. Brain Sci.* **2008**, *31*, 489–509.
72. Lakretz, Y.; Dehaene, S. Recursive Processing of Nested Structures in Monkeys? Two Alternative Accounts. 2021. Available online: [https://osf.io/preprints/psyarxiv/k8vws\\_v1](https://osf.io/preprints/psyarxiv/k8vws_v1) (accessed on 6 October 2025).
73. Ferrigno, S.; Cheyette, S.J.; Piantadosi, S.T.; Cantlon, J.F. Recursive Sequence Generation in Monkeys, Children, U.S. Adults, and Native Amazonians. *Sci. Adv.* **2020**, *6*, eaaz1002. <https://doi.org/10.1126/sciadv.aaz1002>.
74. Liao, D.A.; Brecht, K.F.; Johnston, M.; Nieder, A. Recursive Sequence Generation in Crows. *Sci. Adv.* **2022**, *8*, eabq3356. <https://doi.org/10.1126/sciadv.abq3356>.
75. Rey, A.; Fagot, J. Associative Learning Accounts for Recursive-Structure Generation in Crows. *Learn. Behav.* **2023**, *51*, 347–348. <https://doi.org/10.3758/s13420-022-00564-y>.
76. Rey, A.; Perruchet, P.; Fagot, J. Centre-Embedded Structures Are a by-Product of Associative Learning and Working Memory Constraints: Evidence from Baboons (*Papio Papio*). *Cognition* **2012**, *123*, 180–184. <https://doi.org/10.1016/j.cognition.2011.12.005>.
77. Berwick, R.C.; Friederici, A.D.; Chomsky, N.; Bolhuis, J.J. Evolution, Brain, and the Nature of Language. *Trends Cogn. Sci.* **2013**, *17*, 89–98. <https://doi.org/10.1016/j.tics.2012.12.002>.
78. Hagoort, P. On Broca, Brain, and Binding: A New Framework. *Trends Cogn. Sci.* **2005**, *9*, 416–423.
79. Friederici, A.D.; Chomsky, N.; Berwick, R.C.; Moro, A.; Bolhuis, J.J. Language, Mind and Brain. *Nat. Hum. Behav.* **2017**, *1*, 713–722. <https://doi.org/10.1038/s41562-017-0184-4>.
80. Jeon, H.-A.; Friederici, A.D. Degree of Automaticity and the Prefrontal Cortex. *Trends Cogn. Sci.* **2015**, *19*, 244–250. <https://doi.org/10.1016/j.tics.2015.03.003>.
81. Jeon, H.-A.; Friederici, A.D. Two Principles of Organization in the Prefrontal Cortex Are Cognitive Hierarchy and Degree of Automaticity. *Nat. Commun.* **2013**, *4*, 2041. <https://doi.org/10.1038/ncomms3041>.
82. Matchin, W.; Hammerly, C.; Lau, E. The Role of the IFG and pSTS in Syntactic Prediction: Evidence from a Parametric Study of Hierarchical Structure in fMRI. *Cortex* **2017**, *88*, 106–123. <https://doi.org/10.1016/j.cortex.2016.12.010>.
83. Margulies, D.S.; Ghosh, S.S.; Goulas, A.; Falkiewicz, M.; Huntenburg, J.M.; Langs, G.; Bezgin, G.; Eickhoff, S.B.; Castellanos, F.X.; Petrides, M.; et al. Situating the Default-Mode Network along a Principal Gradient of Macroscale Cortical Organization. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 12574–12579. <https://doi.org/10.1073/pnas.1608282113>.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.