

The Crucial Role of Recursion Complexity in the Pedagogical AI Construction



Jun Lai^{1,*}

¹School of International Studies, Sun Yat-sen University, P.R.China

Abstract: Based on psycholinguistic and cognitive theories of recursive structures, the current research investigates the core differences between human recursive thought and large language models' recursive text generation. The present paper focuses on the cognitive processing and acquisition preferences of recursive syntax. Previous studies on education in the artificial intelligence era mostly discussed the application effects or ethical risks. However, there remains a gap in the cognitive mechanisms of processing the underlying syntactical structures, which are crucial for both human languages and computational language. The current paper proposes that recursion is a core psychological tool for mediating high-level cognitive thoughts, instead of merely a syntactical structure. Furthermore, it probes into the cognitive constraints, processing strategies and obstacles to acquisition. In addition, it provides pedagogical suggestions for protecting learners' recursive thinking abilities.

Keywords: recursion, language thoughts, pedagogy, educational strategies, artificial intelligence, syntax

1. Introduction

In the artificial intelligence (AI) era, the acquisition of knowledge in various domains has become easy thanks to the rapid developments of generative AI tools, especially large language models (LLMs). In the AI pedagogical domain, much research has investigated how AI-generated complex material might influence language usage or even the way of thinking. The debate over the content and the method of pedagogy has attracted a lot of attention. Some literacy-based approaches prioritise literacy over knowledge merely, whereas the other theories highlight AI capability, suggesting that the current education should especially focus on training the abilities of collaborating with AI. For instance, precious time and learning opportunities should be reserved for training more irreplaceable capabilities, such as critical thinking, creativity, co-piloting, etc. Although the debate has discussed the content of education, the underlying structure of knowledge itself is still neglected. Yet, the way knowledge is organised, processed and transferred deserves more research attention. AI might overperform human beings in processing vast amounts of information within an extremely short time. Most studies generally recognise that LLMs are efficient in

producing complex sentences with multiple layers of embedding. However, whether LLMs can generate the output in an equivalent way to human reasoning is still questionable. Especially, the crucial linguistic and cognitive ability of human beings is the recursive capability, which is fundamental for language comprehension, production, higher-level reasoning, and the pedagogical AI construction. Much literature has probed into cross-language recursion processing and presented different levels of comprehension challenges. Nevertheless, there still remains a gap in utilising LLMs as parallel processing agents of recursive processing for human-computer comparison. It is worthy to explore whether the models overlap with the cognitive bottlenecks of human parsers, or whether they apply completely different mechanisms.

Based on prior studies, the current paper aims to address three research questions: (1) We analyse recursion at various levels, including syntactical, cognitive, and organisational levels. (2) We compare the performance differences between LLMs and human beings in processing recursion and discuss the fundamental reasons accounting for these differences. (3) We discuss the idea of recursive literacy in the pedagogical AI construction. The goal of the present

Corresponding Author: Jun Lai

School of International Studies, Sun Yat-sen University, P.R.China

©The Author(s) 2024. Published by BONI FUTURE DIGITAL PUBLISHING CO., LIMITED This is an open access article under the CC BY

License(<https://creativecommons.org/licenses/by/4.0/>).

research is to contribute to the central theme of recursion and provide a novice path at the curriculum, teaching, and evaluation levels. The research on it could shed light not only on the psychological mechanisms of complex information processing and language acquisition, but also on the irreplaceability of education.

2. Recursion and its processing constraints.

The concept of recursion originates from the domains of mathematics, computer science and logic (Tomalin, 2007), referring to a self-referential mechanism or “a procedure that calls itself, or to a constituent that contains a constituent of the same kind” (Pinker & Jackendoff, 2005). Since the 1950s, in the field of linguistics, the specific recursive structure has been incorporated into the series of generative grammar theories, which emphasised the discrete infinity which can produce an infinite variety of thoughts (Bar-Hillel, 1953; Chomsky, 1957, 1965, 1995). Furthermore, Hauser et al. (2002) raised a substantial distinction between the faculty of language in the broad sense (FLB) and in the narrow sense (FLN). Generally, FLB contains different systems, such as a sensory-motor system, a conceptual-intentional system, as well as the computational system such as recursion. They conceptualised that recursive mechanisms can generate an infinite set of outputs from a finite set of lexicons. On the contrary, FLN only refers to recursion, which is regarded as the unique component of the human faculty of language (Fitch, 2010; Fitch et al., 2005; Hauser, 2010). Although most of the contemporary human languages possess the recursive structures, it is worth mentioning the controversy that it is absent in some Amazonian tribe language, such as Pirahã (Everett, 2005, 2007; Sakel & Stapert, 2010). Other studies have pointed out that the absence of recursion also exists in animals (Corballis, 2007; Gentner, et al., 2006). Yet, most theories have highlighted the speciality of recursion and admitted it as an essential component of linguistics.

One of the most distinctive characteristics of

recursion is its hierarchical feature. The hierarchy enables one level of objects to be inserted into another level, creating embeddings. Furthermore, this special mechanism of embeddedness keeps producing long-distance dependencies, which require keeping track of the associations in the working memory (Corballis, 2007; Pinker & Jackendoff, 2005; Van der Hulst, 2010). Crucially, Pinker and Jackendoff (2005) claimed that the necessity for the presence of recursion in languages was its special function of expressing recursive thoughts.

Nevertheless, from a usage-based point of view, the operational system for processing recursive structure in natural language is the non-linguistic ability of sequence learning (Christiansen & Chater, 2015). The usage-based theoretical approach has explicitly categorised two types of recursion, namely tail recursive structures (see Example 1) and complex recursive structures (see Example 2).

Example 1. The mouse bit the cat that chased the dog that ran away.

Example 2. The dog that the cat that the mouse bit chased ran away.

(examples in Christiansen & Chater, 2015)

Although both types of sentences convey similar semantic meanings, the latter sentence is obviously more challenging for language users to parse, due to the complex syntactic structures with two levels of centre-embeddings inserted. The parsing of the sentence might present a breakdown due to the long-distance dependencies. Human parsers might lose track of the associated components and then misunderstand the meaning. They first need to memorise all the single elements, and then maintain all these elements till the end of processing. Afterwards, they need to extract the memorised items and make the integration. This syntactical complexity poses comparative obstacles for language perception, comprehension, and memorisation (Blaubergs & Braine, 1974; Hamilton & Deese, 1971; Larkin & Burns, 1977; Miller, 1962; Miller & Isard, 1964). A large number of cross-language empirical studies have shown that participants can hardly process the recursive

structures beyond two levels of embedding, such as in English, French, German, Swedish, etc. (Bach et al., 1986; Hoover, 1992; Karlsson, 2007). It might be caused by various cognitive limitations, such as memory span, attention, energy distribution, etc. Moreover, evidence from linguistic corpora also indicated that higher levels of embedding hardly exist in oral speech.

In addition, evidence from developmental psychology research indicated that the powerful ability of generating recursive structures develops in a natural and gradual manner, and independent of other cognitive components. Thus, Christiansen and Chater (2015) argued that the individual differences in comprehending and producing recursive structures could be a reflection of inherent variations in the structural and distributional level of languages. More broadly, the differences in linguistic performance on recursion might be influenced by biological and evolutionary constraints. Hence, the usage-based account claimed that humans' recursive ability was implemented by domain-general mechanisms and evolved for the complex system of sequence learning, instead of domain-specific grammar learning (cf. Christiansen & Chater, 2015).

3. Recursion makes the core distinction between LLMs and human beings

Interestingly, when natural language users have difficulties both in the comprehension and production of higher levels of embedding, LLMs can handle that efficiently. Thus, with the incredibly rapid development of AI, the speciality of recursion stands out due to its powerful generativity and unlimited expressive capability, not only in the field of linguistics, but also at the conceptual level and in the mind. Recursion serves as the dividing line between LLMs and human beings in terms of structural capabilities. Although LLMs are capable of handling a vast amount of information within an extremely short time limit, the underlying mechanism of processing recursive structures demonstrates systematic differences from human beings.

From the first perspective regarding complexity, natural language parsers might exhibit an obvious upper limit, approximately two levels of centre-embedding or four to five levels of right-branching recursive structures based on prior studies. Intuitively, it might be easy to assume that LLMs can handle unlimited levels of hierarchy. However, the latest research with generative AI tools demonstrated a striking result, indicating that LLMs present a similar trend when the complexity increases. In this study, researchers created a CenterBench specifically constructed for testing embedding comprehension (Madhusudan et al., 2026). This research trained LLMs to parse different types of complex recursive sentences in order to distinguish the compression of the underlying syntactical structure from the superficial matching of semantic patterns. They adopted a dataset with 9,720 items testing the plausibility of sentences with centre-embeddings, such as examples in "The cat that the dog chased meowed". Each complete set contains one recursive sentence and its identical counterpart but semantically implausible, as well as six test questions targeting comprehension, syntactical association and logical reasoning. Results testing six models demonstrated a systematic decrease in accuracy when the depth of embeddings increased (Madhusudan et al., 2026). The models showed median gaps up to 26.8 percentage points, indicating that even LLMs would render comparatively deeper syntactical analysis for superficial semantic associations in their parsing performance. It is worth mentioning that semantic plausibility significantly influences performance on tests about resulting actions. Although their reasoning models enhance the general accuracy, their intermediate inference traces reveal traces of semantic shortcuts, as well as unnecessary overthinking, etc. Briefly, this latest research with this CenterBench dataset provided evidence that even LLMs would also prioritise pattern matching over structural analysis when the complexity increases to a certain level.

Similarly, a series of experiments has also trained neural language models on linguistic big data

with centre-embedding structures and syntactic island constraints on the filler-gap dependencies (Gotlieb Wilcox et al., 2019). With a task of predicting words incrementally, this research examined whether neural language models can learn the hidden structures of stack-like data functionally, and further deploy them in syntactic and sequential processing. In order to complete the task of predicting the correct words in the corresponding structures, the models need to first suppress certain kinds of expectations, maintain the corresponding information and then try to recover these expectations in the end. Hence, the models can push and pop the expectations in a stack-like method. Their findings demonstrated that the trained neural language models were capable of both suppressing and recovering expectations on a stack in the majority of data. However, even the models cannot completely recover the previous grammatical state.

Nevertheless, it is worth mentioning that in comparing the performance of LLMs, it is important to evaluate the standards and the dimensions adopted by the models, especially how they distribute probabilities to legal versus illegal syntactic structures. For instance, some research reminded that necessary linguistic biases might not even be present in the training signal for models (Davis & van Schijndel, 2020). They compared the model output using English and Spanish. The results revealed that models trained on English presented human-like preference syntactically, whereas models trained on did not show a similar human-like syntactical preference. Thus, the findings raise concerns over the relationship between language comprehension and language production.

The similarities of these studies demonstrated that processing complex recursive structures with higher levels of embedding is not only an obstacle for human beings' cognitive system, but also challenging for LLMs when the complex increases to a certain content.

From the second perspective regarding mechanism, human beings and LLMs might differ in the underlying principle of maintaining the

hierarchical representations. When human parsers encounter complex recursive structures, they proceed with comprehending, keeping dependencies in memory, linking associated counterparts and integrating. The performance decreases when the processing cognitive load increases. The inserted long-distance dependencies might cause the overloaded processing system to collapse. Comparatively, LLMs are able to learn these stack-like structures as human beings after being trained with a large number of datasets. However, as the latest research with language models showed, it is also hard for the models to recover the original strength of expectations if the associated dependencies are separated by long-distance embeddings. In other words, the models largely rely on specific lexicons as the boundary signals for structures. The models' structural tracking ability might get loose once the collocation between the associated items deviates from the norm. Unlike human beings are hindered by the inherent cognitive constraints, the models are not affected by cognitive loads but they still collapse at a certain level.

From the third perspective regarding semantics, both human parsers and LLMs utilise semantic information as cues to identify the relationship in order to facilitate processing or alleviate the overloaded burden. For human beings, empirical research has shown semantic matching between words could reinforce the processing of syntactic structure, whereas semantic mismatching might hinder comprehension. Thus, human parsers mainly semantics to recognise the correct hierarchical relationship and further understand the sentential meaning. Semantics serves as a signal for categorical grammatical judgements. However, the models adopt semantics as a substitute way to analyse structures. In the latest study using CenterBench, they set up an artful design, in which each sentence contains a syntactically identical but semantically anomalous counterpart (Madhusudan et al., 2026). Their results testing six models revealed that the performance differences on grammatical versus ungrammatical but semantically anomalous

counterparts became more significant when the embeddings get deeper. The median gap of up to 26.8 percentage points played the role of indicating that models are still functioning well in tracking the correct structures at the level of comparatively shallow embedding; whereas at the level of deeper embedding, the models might render the expected structural analysis for more superficial semantic dependencies (Madhusudan et al., 2026).

As a brief summary, human parsers and LLMs demonstrate similarities in processing complex hierarchical structures with recursive embeddings. However, these two parsers are only similar in the superficial level but differ distinctively in the conceptual representational level and the underlying principles. The differential mechanisms might be caused by distinct methods of resource construction. For human parsers, the completeness of representations is affected since the processing cognitive system is overloaded. For the models, the rigidity of representations is impaired when the associated items become long-distance due to the inserted recursive items. Thus, although both types of syntactic parsers might present similar behaviour results to a certain extent, the influential factors are distinctive. It is the underlying architecture of organising resources that plays one of the most crucial roles.

4. Recursive literacy

In the era of AI co-piloting, it is crucial to realise the similarities and differences between human parsers and LLMs, especially in the case of processing the core structure of human natural languages, namely, recursion. The above comparisons imply that it is worthy to train abilities of managing hierarchical structures at various domains, such as linguistics, way of thinking and architecture of knowledge.

For instance, from the linguistic perspective, when processing long and complex sentences, it is useful first to recognise and identify which type of syntactic structure categories it belongs to. This is often neglected, since learners easily follow the

Good-enough theories, which claim that language parsers tend to construct a good-enough shallow representation, instead of constructing a complete and precise syntactic structure. After differentiating recursion and iteration, it is important to first link the associated items and mark the separate levels. True recognition requires a clear understanding about the structural boundaries for each level, instead of the surficial link between meanings merely. Finally, the comprehension needs an integration utilising both the semantic and syntactic cues. Within this step, it requires maintaining the relevant information sequentially for the later retrieval, as well as reconstruction. Moreover, the criteria of mastering a specific type of linguistic structure do not only concern comprehension, but also production and generation. Thus, this criterion requires learners to fully understand what recursion is and how the recursive rules operate, rather than making simple judgments according to intuition or language habits. When encountering unfamiliar linguistic materials, learners are still able to apply the recursive rules actively, instead of passively getting confused. Only with transfers to novice materials can unlimited generations be produced. It is also a vital ability to clearly describe the operational processes and make logical judgments about the output or the performance.

Although the current LLMs perform well in generating recursive structures, the above analysis has shown that the models still rely heavily on the superficial semantic cues. When the nested embeddings increase to a certain level, the models' performance also drops sharply. Thus, when human parsers utilise LLMs to generate complex linguistic materials, they should be cautious about the output. Although LLMs might produce a large amount of seemingly well-structured utterances, the core remains in a recursive way of thinking. One has to first make the identification about the levels, and try to organize it in the corresponding hierarchy. Only in that way can recursion be realised.

5. Pedagogical methodology and implications

The ability of recursive way of thinking is the necessity of education in this AI era. Even though learners can master the recursive operational system to produce new utterances, they can still fail to understand up to two or three levels of embeddings. Obviously, the depth of human recursive processing has an upper limit, which does not disappear even with more exposure to complex materials. Thus, targeted training, especially for the underlying recursive way of thinking, stands out in the current pedagogical background. Under the convenient AI co-piloting scenario, learners might have easy access to the seemingly perfect complex structures. Yet, learners can truly master recursive thinking when they are actively performing the recursive loops. Afterwards, they are able to make reasonable judgments over the AI-generated output, efficiently realising the true co-piloting. Thus, specific tasks containing recursive language, ways of thinking and knowledge in general need to be designed, arranged and evaluated.

To sum up, the current paper has probed into a specific structure, namely, recursion, which is fundamental for human language. The characteristics of recursion have been discussed, as well as its processing constraints. Recursion has been analysed at various levels, including syntactic, cognitive, and organisational levels. Comparisons have been conducted on the performance differences between LLMs and human beings in processing recursion. Interestingly, two types of parsers present similar performance when the complexity increases. Nevertheless, the underlying principles differ due to various fundamental factors. In the end, recursive literacy stands out as the crucial capacity in the pedagogical AI construction.

Conflict of Interest

The author(s) declare that there are no conflicts of interest to this work.

Acknowledgement

This research was supported by the National Social Science Fund of China (25BYY029)

References

- Bach, E., Brown, C., & Marslen-Wilson, W. (1986). Crossed and nested dependencies in German and Dutch: a psycholinguistic study. *Language and Cognitive Processes*, 1(3), 249–262. <https://doi.org/10.1080/01690968608404677>
- Bar-Hillel Y. (1953). On recursive definitions in empirical science. *Proceedings of the 11th International Congress of Philosophy, Brussels*, 5, 160–165.
- Blaubergs, M. S., & Braine, M. D. S. (1974). Short-term memory limitations on decoding self-embedded sentences. *Journal of Experimental Psychology*, 102(4), 745–748. <https://doi.org/10.1037/h0036091>
- Chomsky, N. (1957). *Syntactic Structures*. Mouton, The Hague.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Christiansen, M. H., & Chater, N. (2015). The language faculty that wasn't: A usage-based account of natural language recursion. *Frontiers in Psychology*, 6, Article 1182. <https://doi.org/10.3389/fpsyg.2015.01182>
- Coolidge, F.L., Overmann, K.A. and Wynn, T. (2011), Recursion: what is it, who has it, and how did it evolve? *WIREs Cognitive Science*, 2(5), 547-554. <https://doi.org/10.1002/wcs.131>
- Corballis, M. C. (2007). Recursion, language, and starlings. *Cognitive Science*, 31(4), 697–704. <https://doi.org/10.1080/15326900701399947>
- Davis, F., & van Schijndel, M. (2020). Recurrent Neural Network Language Models Always Learn English-Like Relative Clause Attachment. *ArXiv*, abs/2005.00165.
- Everett, D. L. (2005). Cultural constraints on grammar and cognition in Pirahã: another look at the design features of human language. *Current Anthropology*, 46(4), 621–646. <https://doi.org/10.1086/431506>
- Everett, D. L. (2007). Challenging Chomskyan linguistics: The case of Pirahã. *Human Development*, 50, 297-299.

- <https://doi.org/10.1159/000109831>
- Fitch, W. T. (2010). Three meanings of “recursion”: Key distinctions for biolinguistics. In R. K. Larson, V. Déprez, & H. Yamakido (Eds.), *The evolution of human languages: Biolinguistic perspectives* (pp. 73–90). Cambridge University Press.
- Fitch, W. T., Hauser, M. D., & Chomsky, N. (2005). The evolution of the language faculty: Clarification and implications. *Cognition*, 97(2), 179–210.
<https://doi.org/10.1016/j.cognition.2005.05.005>
- Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature*, 440(7088), 1204–1207.
<https://doi.org/10.1038/nature04675>
- Gotlieb Wilcox, E., Levy, R., & Futrell, R. (2019). Hierarchical Representation in Neural Language Models: Suppression and Recovery of Expectations. *ArXiv*, abs/1906.04068.
- Hamilton, H. W., & Deese, J. (1971). Comprehensibility and subject-verb relations in complex sentences. *Journal of Verbal Learning and Verbal Behavior*, 10(2), 163–170.
[https://doi.org/10.1016/S0022-5371\(71\)80008-7](https://doi.org/10.1016/S0022-5371(71)80008-7)
- 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(5598), 1569–1579.
<https://doi.org/10.1126/science.298.5598.1569>
- Hauser, M. D. (2010). On obfuscation, obscurantism, and opacity: Evolving conceptions of the faculty of language. In R. K. Larson, V. Déprez, & H. Yamakido (Eds.), *The evolution of human languages: Biolinguistic perspectives* (pp. 91–99). Cambridge University Press.
- Hoover, M. L. (1992). Sentence processing strategies in Spanish and English. *Journal of Psycholinguistic Research*, 21(4), 275–299.
<https://doi.org/10.1007/BF01067514>
- Karlsson, F. (2007). Constraints on multiple center-embedding of clauses. *Journal of Linguistics*, 43(2), 365–392.
<https://doi.org/10.1017/S0022226707004616>
- Larkin, W., & Burns, D. (1977). Sentence comprehension and memory for embedded structure. *Memory & Cognition*, 5(1), 17–22.
<https://doi.org/10.3758/BF03209186>
- Miller, G. A. (1962). Some psychological studies of grammar. *American Psychologist*, 17(11), 748 – 762. <https://doi.org/10.1037/h0044708>
- Madhusudan, S., Chen, K., Emami, A. (2026). The dog the cat chased stumped the model: measuring when language models abandon structure for shortcuts. *arXiv*.
<https://doi.org/10.48550/arXiv.2510.20543>
- Miller, G. A., & Isard, S. (1964). Free recall of self-embedded English sentences. *Information and Control*, 7(3), 292–303.
[https://doi.org/10.1016/S0019-9958\(64\)90310-9](https://doi.org/10.1016/S0019-9958(64)90310-9)
- Pinker, S., & Jackendoff, R. (2005). The faculty of language: what’s special about it? *Cognition*, 95, 201–236.
<https://doi.org/10.1016/j.cognition.2004.08.007>
- Sakel, J., & Stapert, E. (2010). Pirahã—in need of recursive syntax? In H. Van der Hulst (Ed.), *Recursion and human language* (pp. 3–16). De Gruyter/Mouton.
- Tomalin, M. (2007). Reconsidering recursion in syntactic theory. *Lingua*, 117, 1784–1800.
<https://doi.org/10.1016/j.lingua.2006.11.001>
- van der Hulst, H. (2010). Re recursion. In H. Van der Hulst (Ed.), *Recursion and human language* (pp. xv–liii). De Gruyter/Mouton.

How to Cite: Lai, Jun. (2026). The Crucial Role of Recursion Complexity in the Pedagogical AI Construction. *Contemporary Education and Teaching Research*, 07(10), 357-363.
<https://doi.org/10.61360/BoniCETR262020981005>