Recursion: what is it, who has it, and how did it evolve?

Frederick L. Coolidge,1* Karenleigh A. Overmann1 and Thomas Wynn2

Recursion is a topic of considerable controversy in linguistics, which stems from its varying definitions and its key features, such as its universality, uniqueness to human language, and evolution. Currently, there appear to be at least two common senses of recursion: (1) embeddedness of phrases within other phrases, which entails keeping track of long-distance dependencies among phrases and (2) the specification of the computed output string itself, including meta-recursion, where recursion is both the recipe for an utterance and the overarching process that creates and executes the recipes. There are also at least two evolutionary scenarios for the adaptive value of recursion in human language. The gradualist position posits precursors, such as animal communication and protolanguages, and holds that the selective purpose of recursion was for communication. The saltationist position assumes no gradual development of recursion and posits that it evolved for reasons other than communication. In the latter view, some heritable event associated with a cognitive prerequisite of language, such as Theory of Mind or working memory capacity, allowed recursive utterances. Evolutionary adaptive reasons for recursive thoughts were also proffered, including diplomatic speech, perlocutionary acts, and prospective cognitions.

INTRODUCTION

It was a dark and stormy night. The crew said to the captain, ‘Captain, tell us a story.’ The captain said, ‘It was a dark and stormy night . . .’

—A story in popular culture

Recursion has often been defined as embedding a phrase within another phrase, although there is great debate about this definition or any other single standard definition. When the term is applied to language, there is a definitional problem from recursion’s original use in the fields of mathematics, formal logic, and computer science.1 In these fields, recursion is any instruction that calls another instruction of like type and which is subject to a terminating condition. The introductory story to this article is recursive in at least one sense: it acts as an instruction that calls another instance of itself; it lacks, however, the terminating condition that prevents an infinite loop. The first two sentences of the story are not recursive in the sense of embedding a phrase or sentence within another phrase or sentence; the third sentence, however, repeats the beginning of the story and is therefore recursive in this sense. Thus, at the outset of this discussion, it is important to reiterate: there is no single, universally accepted definition of recursion. Its definition varies across disciplines (e.g., mathematics, logic, computer science, and linguistics), and it varies within these disciplines, particularly within linguistics. In fact, much of the current interest and controversy among scholars about recursion stems from its varying definitions and debate about its key features, such as its universality, uniqueness to human language, and evolution.

A BRIEF HISTORY OF THE DEFINITIONAL DEBATE

Although recursive theory in mathematics and formal logic has been described and employed since the...
1930s, it was not until the 1950s that linguists began to discuss potential recursive aspects of language. In 1953, Bar-Hillel wrote a short paper entitled *On Recursive Definitions in Empirical Science* in which he argued not only that pure mathematics and formal logic should employ recursive definitions but that recursive definitions might greatly facilitate analytical methods in linguistics. In 1956, Chomsky incorporated recursive components in his theory of Generative Grammar, although his use of the term seemed more directed toward the idea that recursion (more recently, the operation labeled Merge) is what creates discrete infinity or productivity in language.

The current debate about the nature of recursion may be traced to a controversial and provocative article by Hauser et al. in which they (1) made a distinction between the faculty of language in the broad (FLB: all communication, human or otherwise) and narrow (FLN: only human language) senses, (2) theorized that computational mechanisms for recursion provided the ability to generate an infinite range of expressions from a finite lexicon, (3) asserted that FLN as a set might uniquely and only include recursion as a member, and (4) speculated that FLN may have evolved for reasons other than communication. They offered the possibility that FLN or particular aspects of it were simply spandrels, that is, byproducts of other naturally selected for, adaptive capabilities. The subsequent commentaries and rebuttals appear, in part, to be reacting to the counterintuitive Hauser et al.’s hypothesis that the faculty of language evolved for purposes other than communication. Hauser et al. provided no discussion or explanation of alternatives as to why language happened if not for communication, other than a brief mention of spandrels, spatial navigation, foraging, number quantification, and social relationships. They also failed to define recursion explicitly, other than stating it was part of a computational system that generated internal representations that mapped onto a sensory–motor interface by the phonological system and onto a conceptual–intentional interface by a formal semantic system, enabling ‘us to communicate an endless variety of thoughts, …’ (p. 1574). It is important to point out that their faculty of language argument for ‘purposes other than communication’ referred only to FLN and not FLB.

In 2005, Pinker and Jackendoff provided a detailed critique of the arguments in Hauser et al., but unlike them, injected a clear definition of recursion as ‘a procedure that calls itself, or to a constituent that contains a constituent of the same kind’ (p. 203) because in their view Hauser et al. had not offered one. This new definition of recursion—as embeddedness—was similar to some definitions of recursion used in mathematics, formal logic, and computer science (for a review of the latter definitions, see Ref 5). The characteristic of recursive embeddedness possibly changes the intent of Hauser et al., who had emphasized Chomsky’s original notion of recursion as discrete infinity and their mutual idea of recursion enabling the production of an infinite variety of thoughts. Thus, although Pinker and Jackendoff acknowledged that Hauser et al. envisioned recursion as that which produces phrases and sentences from words (syntax), they highlighted an important aspect of recursion, its embeddedness.

Pinker and Jackendoff also noted the various lines of evidence that suggest ancestral humans had a primary auditory system adapted to perform ever increasingly complex auditory analyses of their environment. They found inconceivable Hauser et al.’s claims that there have been no evolutionary changes to the mammalian auditory system of humans for speech perception and that speech perception in modern humans began *de novo*. Chomsky bases his claim that language and its recursive elements may have evolved for reasons other than communication on the parallel argument made by Alfred Russell Wallace (codiscoverer of natural selection) over 125 years ago about the origins of mathematical abilities: ‘[the] gigantic development of mathematical capacity is wholly unexplained by the theory of natural selection, and must be due to some altogether distinct cause’ (Chomsky, Ref 8, p. 53). Chomsky noted that the core element of mathematical capacity, arithmetic, may be a derivative of language. There is, however, a plethora of recent neuroimaging, genetic, and human and animal experimental studies that clearly delineate unique neurological substrates for two core mathematical abilities (subitization and magnitude appreciation), as well as overwhelming evidence for their independence from language functions and their very ancient evolutionary foundations (well over 40 million years; see Ref 9, for an excellent overview of the evolution of numerical cognition). It is also important to note that higher level mathematical concepts (e.g., natural and irrational numbers) are a recent historical and cultural development, and wholly dependent upon language, and that it is not clear whether Chomsky’s recent claim that arithmetic derives from language references innate mathematical abilities or higher level mathematical theory and knowledge (for more details of the Jackendoff and Pinker, Fitch, Hauser, and Chomsky debates, see Refs 5, 6, 8, and 10–14).
RECURSIVE THINKING: HOW AND WHY DID IT EVOLVE?

One of Pinker and Jackendoff’s more important insights into this debate is the idea that ‘the only reason language needs to be recursive is because its function is to express recursive thoughts. If there were not any recursive thoughts, the means of expression would not need recursion either’ (p. 230). At this point in their argument, they offer as plausible candidates for recursive thinking: music, social cognition, decomposition of objects into parts, and the formulation of complex action sequences. Indeed, they state ‘the problem is not a paucity of candidates for evolutionary antecedents but a surfeit’ (p. 230). Ironically, none of their examples directly engage language. The question remains: just what are recursive thoughts that are expressed in language, and why might they be evolutionarily adaptive? If recursion is identified as the cognitive act of designating and subsuming elements, and the elements in question include thoughts as well as phrases and sentences, an important aspect of recursion becomes clearer: the cognitive act of recursion entails holding in mind the relation between the two elements or thoughts, as well as realizing the modification that one element or thought has upon the other. The recursive syntactic structures in language ‘express complex thoughts which themselves display [the] kind of combinatorial capacity’ that is another important characteristic of recursion (Van der Hulst, Refs 15 and 16, p. xxiii). Thus, Pinker and Jackendoff have identified a critical role for recursion—to express recursive thoughts in language—that has been neglected to date, in part because of the lack of precision in the term’s definition and because mathematical- and computer science-based recursivity does not encompass the social, emotional, and communicative intent of utterances. If recursion was an important adaptive feature in the evolution of language, then it may be fruitful to understand the reasons for recursion in speech acts (i.e., the reasons why utterances are made). Three plausible candidates come to mind: diplomatic speech, perlocutory acts, and prospective memory and cognition.

Ambrose has noted that, approximately 75,000 years ago, Homo sapiens began to transition from restricted landscapes and small group interactions to expanded social landscapes marked by material exchanges, trade, and other forms of reciprocal cooperation. Ambrose believes that one associated innovation may have been a language of diplomacy. He based his language of diplomacy upon the concept of indirect speech by Pinker et al., who noted that people frequently insinuate their intentions indirectly in speech, for example, ‘Gee, officer, is there some way we could take care of the ticket here? [a bribe]’ (p. 833). If it can be determined empirically that diplomatic speech acts require recursion or are best expressed through recursive utterances, then the adaptive nature of recursive thinking becomes more likely.

A second candidate for recursive thinking is perlocutionary speech, which entails persuading, scaring, deceiving, or manipulating someone into doing or realizing something. Perhaps, recursive phrasing makes utterances more persuasive, with or without the speaker’s conscious intentions, for example, ‘I said, I’m sorry!’ What differentiates the first two candidates is that diplomatic speech encompasses the use of elaborated codes (speech that is explicit and ornate because it cannot depend on familiarity and context), whereas in perlocutionary speech acts, the emphasis is upon the pragmatic intent enabled when the speaker and listener are sharing a restricted code (speaker and listener are familiar with each other, the content, and the context).

A third candidate is the use of prospective memory and cognition, which are the abilities to remember to carry out a future act envisioned in the present and to engage in planning for the future, respectively. Gärdenfors and Osvath have argued that the ability to create detached representations for objects or events not currently in the external world would allow and facilitate prospective cognitions (e.g., Suddendorf and Corballis—mental time travel and Tulving—episodic memory and autonoetic thinking), enabling the mental content of an individual to be recalled and shared with others for the purposes of cooperative planning and future action. They maintain that this ability, which they term prospective cognition, is a unique feature of human cognition. Gärdenfors and Osvath suggested that it is the use of symbolic language referring to detached mental representations that truly distinguishes humans from animals. They further argued that symbolic communication was much more efficient than any other method for communicating about nonexistent entities and that this ability to share visions about future goals had increasingly selective advantages over time. It seems possible that prospective cognition, in the form of contingent recursive utterances, might have had an even greater selective advantage. For example, rather than stating ‘tomorrow we hunt’, which is a form of planning and can be shared with others, the ability to say ‘if it rains tomorrow, we will make spears, but if it does not rain, we will hunt’
is a yes–no articulation, suppression of breathing, vocal range, dates, Reuland offers working memory, accuracy of concomitant cognitive and neurophysiological candidates. As to plausible gradual by necessity’ (p. 8), which to him suggests a no gradual increase in recursivity and that it is ‘non-adaptive’ (p. 8) because recursivity is a yes–no proposition. He argues that there can be no gradual increase in recursivity and that it is ‘non-dissociable by necessity’ (p. 8), which to him suggests a possible concomitant genetic change. As to plausible concomitant cognitive and neurophysiological candidates, Reuland offers working memory, accuracy of articulation, suppression of breathing, vocal range, and speed in lexical access.

RECURSION: WHO HAS IT—OR NOT?

Complicating the controversy surrounding recursion are the Pirahã, an Amazonian tribe who apparently lack recursion in their language.29 According to Everett, the Pirahã do string phonemes into words and words into phrases and sentences, but instead of embedding phrases and sentences within other phrases and sentences, they string phrases and sentences together paratactically. Parker,14 Everett (personal communication to Fitch5), Sakel and Stapert,30 and others have noted that the Pirahã may possibly construct and communicate recursive thoughts. The debate is also methodologically questionable, primarily because all observations so far are based on reports by Everett and others and even they disagree.5 Other arguments, such as ‘... Pirahã speakers are perfectly capable of expressing the same underlying conceptual structures as English speakers...’ (Parker, Ref 14, p. 7) are unsubstantiated as Everett never empirically tested any standard range of conceptual structures. Everett also reported that the Pirahã lack origin myths, social hierarchy, autonoesis, words for color, and numbers beyond about three, and their kinship relations and artwork are limited. It is possible that the apparent lack of recursion represents some kind of cultural constraint, as Everett suggests. Certainly, strong critics of Everett’s conclusions31 admit that the Pirahã have some peculiar cultural features. These cultural peculiarities could also be consistent with other explanations, such as an underlying neurophysiological deficiency, like limited working memory capacity, or, as Bickerton32 suggests, an extreme case of acquisitional delay. Another reason for rejecting the Pirahã as relevant to the debate about recursion is that they are related to neighboring groups that do have recursion, but the Pirahã could have lost their recursive abilities for some unexplained reason. The suggestion by Pinker and Jackendoff6 and Parker14 that recursion is not an essential part of language because the Pirahã lack it strikes us as beside the point. Therefore, we agree with Bickerton: the example of the Pirahã as a test case of the universality of recursion in language is flawed. It is important to note that their debatable lack of recursion does not invalidate the arguments that recursion forms an essential part of most languages or that most contemporary human languages possess recursion.

Another potential challenge to the claim made by Hauser et al.4 that recursion is unique to human language is the finding by Gentner et al.33 that European starlings can learn and ‘accurately recognize acoustic patterns defined by a recursive, self-embedding, context-free grammar’ (p. 1204). Gentner et al. posed their argument that birds understood center-embedding in a sequence in the form A^B^n. However, there are many problems with this research. First, the starlings did not easily learn the recursive A^nB^n phrasings; 9 of their 11 birds required an average of 30,344 trials to reach criterion (two birds
Recursion: what is it?

were unable to reach criterion), and the range was 9400–56,200 trials. The latter is most unlike humans, who acquire recursion early and effortlessly. Second, as Fitch has made clear, A^nB^n grammar is not a litmus test for recursion. A^nB^n is typically viewed as simply canonical grammar beyond finite-state capabilities. As Fitch notes, A^nB^n could be implemented recursively, but A^nB^n can also be implemented without recursion. Third, birds and starlings’ vocal calls, regardless of whether they possess recursion or not, are used for a single purpose: communicating emotional states (e.g., mating, predator warnings, and territoriality), not recursive thought. Fourth, and importantly, Corballis notes that the performance of the birds can be wholly explained by counting strategies and that ‘[t]here appears to be no evidence to date that nonhuman species understand recursion’ (p. 697).

Finally, some of the confusion surrounding recursion derives from the two rather different ways it is used in the literature, a strong version and a weak version [or, in the terms used by Van der Hulst,16 general (generative or Merge) and specific (nested and tail)]. The strong version (i.e., which is the most common in the literature) of linguistic recursion is embeddedness, in which the object embedded is a word, phrase, or sentence, and this creates ‘long-distance dependencies and the need to keep track, or add to memory’ (Parker, Ref 14, p. 3; also see Van der Hulst, Ref 16). Long-distance dependencies are not necessarily a derivative of recursive processes, but they may be observed in nested recursion in particular. The strong version recalls the definition provided by Pinker and Jackendoff—recursion is ‘a procedure that calls itself, or to a constituent that contains a constituent of the same kind’ (p. 203)—although their definition is clearly derived from mathematics and computer science though lacking a terminating condition. Embeddedness can have at least two different forms: tail recursion, seen in possessives (Harry’s father …) and relative clauses (… the wizard who was killed by Lord Voldemort …), and nested recursion (or center-embedded or self-embedded recursion), which leaves material on both sides of the embedded component (… exclaimed, ‘Dumbledore, a wizard whom we who are wizards too admire, will protect Harry!’). It is important to note that embedding phrases within phrases is generally termed hierarchy, and some linguists would argue that hierarchy is recursive only when it embeds phrases of the same type within each other. However, it is debatable whether the procedure must call a constituent that is identical to itself, a nuance that is captured in the definition offered by Corballis: ‘Recursion is a computational procedure that calls itself, or that calls an equivalent kind of procedure’ (p. 698, emphasis ours). Finally, it is also important to note that while embedding can occur independently of recursion, it is not clear that recursion occurs only in syntax and semantics, as Hauser, Chomsky, and Fitch have argued (see Ref 15 for a good explanation of how phonotactics and morphotactics use recursion).

The weak version (i.e., less common) of linguistic recursion is sometimes overlooked in the focus on word, phrase, and sentence embeddedness or skewed by the use of terminology from mathematics and computer science. It is this sense of recursion that specifies the computed output string itself and gives language its power to generate new utterances by enabling the combination of discrete units into larger units (linguists now prefer to label this instruction Merge); it also enables specific recursion.15 Reuland notes that when Chomsky used the term in this sense, he simply meant a recursive operation like the successor function, n + 1, which defines natural numbers. Arguably, the successor function in natural numbers is iteration rather than recursion, but what differentiates the two is significant. Iteration, which is not recursion, does not increase the depth level of embedded structures, while recursion does, within the limits of its constraints.35 Iteration is also counting (n + 1) to infinity, while recursion is counting (n + 1) and stopping at a specific number. Recursion in the sense of Merge—combining two linguistic objects to form a composite available for subsequent merges—still requires a terminating condition. Because whatever mechanism imposes the terminating condition normally caps words, phrases, and sentences within the limits of working memory, it is a mechanism constrained by but perhaps not reducible to working memory (see Ref 28 for an intriguing possibility for the constraints imposed by working memory and the procedural–declarative memory systems interface).

Recursion in the sense of combinatorics—combining discrete units into larger units—also functions on multiple levels in language: it turns phonemes into morphemes, morphemes into words, and words into phrases and sentences. It might be useful, therefore, to examine (1) whether recursion or Merge is the sole or predominant operation for all levels of recursion. Bickerton suggests the linguistic operation Sequence for the super-sentential (paragraph or discourse) level. Other mechanisms also seem important, like the one that selects a particular phrase or sentence structure from the set of structures in a language and the one that delimits the set of words known to a speaker to those appropriate to
CONCLUSIONS

1. Current interest and controversy among scholars about recursion stems from its varying definitions (e.g., whether it entails embeddedness, calls identical or similar instances) and debate about the features in language (e.g., its universality, uniqueness to human language, and adaptive value in language evolution). However, as Fitch\(^5\) aptly noted, recursion is important and worthy of scientific pursuit because of its generativity or ‘unbounded expressive power’ and its ability to allow ‘language users to express any concept that can be conceived, to whatever degree of accuracy or abstraction … needed’ (p. 89). Thus, it appears to be the computational power of recursion that makes it so important and not its uniqueness to or universality in human languages.

2. Although recursive theory in mathematics and formal logic has been described and employed since the 1930s, it was not until the 1950s that linguists began to discuss potential recursive aspects of language. As a result, mathematical-and computer science-based definitions of recursion have led to some confusion in linguistic recursion research, e.g., whether the A\(^n\)B\(^n\) model of parsing grammar (see Refs 5 and 34 for additional details) is a test for the presence of recursion (the opinion of the latter two authors is ‘no’). Debates about the universality of recursion in human language are interesting, but the absence of recursion in a single human language does not appear to have much value for understanding the ultimate consequences and adaptive value of the computational power of recursion.

3. Recursion in language may have arisen to express recursive thoughts, e.g., diplomatic speech, perlocutionary acts, and prospective memory and cognition. If recursive thoughts had adaptive value, the possibility remains that recursion was an evolutionary spandrel and recursivity evolved because of concomitant beneficent change in a cognitive prerequisite for language. Recursive thoughts and their cognitive prerequisites may provide a fruitful area for future research.

4. There appears to be at least two common senses of linguistic recursion. One is a strong (or commonly used) sense, where it seems to imply embeddedness of phrases/sentences within other phrases/sentences (either tail or nested) and maintenance in working memory of long-distance dependencies. The other is a weak (or less common) sense, where recursion specifies the computed output string itself, giving rise to the enabling of combinations of discrete units into larger units, and meta-recursion, where recursion is both the recipe for an utterance and the overarching process that creates and executes the recipes. Both senses require a terminating condition.

5. The debates about the evolutionary adaptive value of recursion and human language lead to at least two evolutionary scenarios for their emergence, gradual and sudden. The gradualist position posits precursors of modern human language, such as animal communication systems and protolanguages and is consistent with the hypothesis that the primary purpose of language and its evolutionary selective value was communication. The saltationist position (a sudden, large change from one generation to the next), which can also be referred to as discontinuity theory, suggests the technical challenge of discrete infinity is not mitigated by the number of states in a finite-state grammar, leading to an all-or-nothing genetic or epigenetic event scenario. This position most often assumes that there is no gradual development of recursion; it is a yes-or-no proposition. In its most extreme interpretation, it is assumed that modern language is a very recent event (less than 100,000 years ago) and that there were no communications that can
be referred to as ‘language’ before that time. This position sometimes assumes that a neural change such as increased brain size or recursion conferred selective advantages to language but only incidentally. A minority position in the saltationist camp holds that there could have been a delay between the moment the language faculty became possible and the moment when it became fully usable, and this position is consistent with the idea of a protolanguage. In this scenario, a protolanguage exists prior to fully modern language, and some presumably genetic event affected a cognitive prerequisite for language, such as an increase in working memory capacity or its subsidiaries.

Coda

It was a dark and stormy night. The crew said to the captain, ‘Captain, tell us a story.’ The captain said, ‘Stop me if you’ve heard this one . . . ’

REFERENCES


