

# Specifically Human: Going Beyond Perceptual Syntax

Piera Filippi

Received: 8 January 2013 / Accepted: 28 February 2013 / Published online: 11 October 2013  
© Springer Science+Business Media Dordrecht 2013

**Abstract** The aim of this paper is to help refine the definition of humans as “linguistic animals” in light of a comparative approach on nonhuman animals’ cognitive systems. As Uexküll & Kriszat (1934/1992) have theorized, the epistemic access to each species-specific environment (*Umwelt*) is driven by different biocognitive processes. Within this conceptual framework, I identify the salient cognitive process that distinguishes each species typical perception of the world as the faculty of language meant in the following operational definition: the ability to connect different elements according to structural rules. In order to draw some conclusions about humans’ specific faculty of language, I review different empirical studies on nonhuman animals’ ability to recognize formal patterns of tokens. I suggest that what differentiates human language from other animals’ cognitive systems is the ability to categorize the units of a pattern, going beyond its perceptual aspects. In fact, humans are the only species known to be able to combine semantic units within a network of combinatorial logical relationships (Deacon 1997) that can be linked to the state of affairs in the external world (Wittgenstein 1922). I assume that this ability is the core cognitive process underlying a) the capacity to speak (or to reason) in verbal propositions and b) the general human faculty of language expressed, for instance, in the ability to draw visual conceptual maps or to compute mathematical expressions. In light of these considerations, I conclude providing some research questions that could lead to a more detailed comparative exploration of the faculty of language.

**Keywords** Syntax · Faculty of language · Animal cognition · Symbols

## Introduction

Aristotle (≈ 350 b.C./2004) defined man as “zoon logikon”, an expression typically translated as “rational animal” or “linguistic animal”. Currently, the meaning of “logikon”

---

P. Filippi (✉)

Department of Cognitive Biology, University of Vienna, Room 2234, Althanstrasse, 14, 1090 Vienna, Austria

e-mail: piera.filippi@univie.ac.at

is an argument of intense philosophical debate. I believe that data resulting from ethological research into cognition can provide a significant contribution to this topic. Therefore, the goal of this paper is to help refine the definition of humans as “linguistic animals” in light of a comparative approach with nonhuman animals’ cognitive systems.

It is worth highlighting a distinction that not all spoken languages make, namely the semantic distinction between the faculty of *language*, meant in a broad sense, and the actual ability to speak and understand a natural language. The former meaning is used in reference to various cognitive domains such as nonhuman animals’ cognitive and communication systems, as well as to math, art, music, and so forth—while the second refers to the ability to speak (or sign) a natural language such as English, Chinese, or Italian. This terminological distinction is expressed, for example, by the French words *langue-language*.

I will focus my investigation on the sense of the word “language” meant in a broad sense. Specifically, I will use a comparative study of the cognitive skills in nonhuman primates and songbirds in an attempt to demonstrate how the general ability of *language* could be conceived of as a domain of cognitive processes shared among animals. Secondly, in light of this data, I will address a core human-specific cognitive process at the basis of both the ability to speak (or to reason) in verbal propositions and of the general faculty of language expressed, for instance, in the ability to draw visual conceptual maps or to compose musical pieces.

### ***Umwelt* and Syntax: The Species-Typical World**

It is commonly assumed that the epistemic approach to *the* world is reliant on the quantity of information that each different animal species is able to perceive and elaborate. In contrast to this idea, Uexküll & Kriszat (1934/1992) observe that there is no such thing as a single world shared by every animal species. On the contrary, each species has a “biocognitive” apparatus that enables it to access a distinctive, species-typical “world-environment”: the *Umwelt*. This entails that in every living species the system of cognition is strictly intertwined with the sensorial domain in such a way that the sensorial perception of an object consists in the possibility to *recognize* in it something essential for life itself (Uexküll and Kriszat 1992; see Carapezza 2006). In other words, according to Uexküll the perception of an object is made up of the recognition of its effect in terms of functionality for life. However, each species’ *Umwelt* depends precisely on the way in which each species-specific biocognitive apparatus perceives the functions and the effects of those external objects.

This concept is efficiently clarified by Lo Piparo (1998), who compares the biocognitive system of each species to the algorithm implementing a complicated mathematical function that uses facts of the external environment as its set of arguments, and mental states as its set of values generated by applying the algorithm to the arguments. From this perspective, the algorithm used for the pair “state of mind-environment”—i.e. the biological mechanisms underlying the functional perception of the world—has a principal role in the description of animals’ cognitive systems. Following this theoretical paradigm, I assume that the “cognitive algorithm” shared across species consists of the ability to primarily pick up connections or articulations among objects (through the species-specific bio-sensorial modalities), rather than merely isolated single objects in the world.

Accordingly, my investigation on the faculty of language (meant in its broad sense) focuses on the ability to connect different elements according to the specific rules allowed by the biocognitive potential of each species. In other words, my investigation of the ability of language coincides with the study of the ability of syntax, used here according to its etymological meaning from the Greek word *syntaxis* (*syn*: together–*taxis*: connect, assemble according to rules).<sup>1</sup> Thus, along this theoretical line, we could state that the cognitive possibility to access each species' *Umwelt* can be referred to as a core ability of "*syntaxis*".

In light of these considerations, I believe that the analysis of exactly this ability in humans could lead to a precise indication of the cognitive processes that make human cognition species-specific. In particular, I assume that one of the most effective strategies to adopt towards this aim would be: (i) to choose one hypothetical constitutive feature specific to humans' ability of *syntaxis* and (ii), to adopt a comparative approach to nonhuman animals' cognitive systems, in order to show if they possess this supposedly unique human cognitive feature. I believe that this specific methodology is the best candidate to provide us with a better grasp of the cognitive similarities among animals and, at the same time, of the mechanisms that underlie the perception of humans' *Umwelt*.

### A Special Kind of *Syntaxis*: Recursion

A distinctive feature of human language is the so called "ruled-governed creativity" (Chomsky 1965): the ability to produce (and understand) an infinite range of various expressions from a finite set of elements and rules. In a highly influential paper, Hauser, Chomsky & Fitch (2002) hypothesize that the only uniquely human component of language mechanism (underlying the property of "ruled-governed creativity") is recursion: the cognitive core responsible for hierarchical processing of self-embedding structures. From this perspective, the kind of *syntaxis* that differentiates human language from other species' system of communication is *at least* the faculty to connect *recursively* the constitutive units of natural languages. Namely, the authors suggest that the ability to produce an infinite range of ever-new expressions relies upon the *syntactic* property of recursion. I will refer to this suggestion as the "recursion distinctiveness" hypothesis.

Linguists within the theoretical framework of generative grammar distinguish two types of recursion: "tail recursion" and "center-embedded recursion". The first one can be mainly identified with a process of iteration, where constituents are just sequentially added at the end (or at the beginning) of a sentence, as in the familiar story of *The House the Jack Built*:

This is the house that Jack built.

This is the malt that lay in the house that Jack built.

This is the rat that ate the malt that lay in the house that Jack built.

This is the cat that worried the rat that ate the malt that lay in the house that Jack built.

<sup>1</sup> Hereafter I will italicize the terms "syntax" and "syntactic" when used according to this specified etymological sense.

As repetitions of sequences are found among several nonhuman species' vocal calls, this level of *syntactic* organization of the signal cannot account for the uniqueness of human language. On the contrary, the center-embedded recursion - defined as a procedure that invokes an instance of itself or of an equivalent procedure in the middle of the computation - is a much more intricate procedure, never encountered among nonhumans' calls. Center-embedded recursion can be explained analyzing the following examples (Corballis 2007; see Christiansen and Chater 1999):

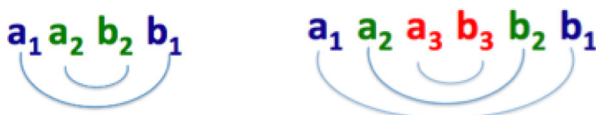
- (i) [The malt (that the rat ate) lay in the house that Jack built.]
- (ii) {The malt [that the rat (that the cat killed) ate] lay in the house that Jack built.}

In (i), the phrase *that the rat ate* is embedded in the sentence *The malt lay in the house that Jack built*, while in (ii), the phrase *that the cat killed* is embedded in the phrase *that the rat ate*. More specifically, the formal structure of these sentences can be illustrated as follows:

- (i)  $a_1$  [the malt]  $a_2$  [that the rat]  $b_2$  [ate]  $b_1$  [lay in the house that Jack built.]
- (ii)  $a_1$ [The malt]  $a_2$  [that the rat]  $a_3$  [that the cat]  $b_3$ [killed]  $b_2$  [ate]  $b_1$  [lay in the house that Jack built.]

In the examples above, the same low indexes depict the connections between the elements. For instance, in (i),  $a_1$  [the malt] is semantically and morpho-syntactically linked to  $b_1$  [lay in the house that Jack built], whereas  $a_2$  [the rat] is linked in the same way to  $b_2$  [ate]. In other words, these sentences respectively have the structure  $a_1 a_2 b_2 b_1$  and  $a_1 a_2 a_3 b_3 b_2 b_1$ , where the 'ab' pairs with higher indexes are hierarchically embedded within 'ab' pairs with lower indexes—namely, a procedure of a–b association invokes an instance of itself in the middle of an equivalent a–b linking procedure (Fig. 1).

In principle, the implementation of a recursive process allows the generation of nested components of sentences within each other, without restriction. This implies the ability to keep in memory a virtually unlimited amount of information about what it has been heard (or seen), while other parts have to be checked first, and resumed from where they are left off. To illustrate this last point, let's return to one of our examples. In (ii), one has to store in memory [the cat], [the malt] as the first constituents that are systematically going to be linked with the correspondent verb, i.e. respectively [killed], [ate], [lay]; thus, while linking  $a_3$  [the cat] to  $b_3$ , i.e. the verb [killed], one has to keep in mind the just heard constituents  $a_1$  [the malt] and  $a_2$  [the rat]; similarly, while connecting  $a_2$  [the rat] to  $b_2$  [ate], one has to resume the previous connection between  $a_3$  and  $b_3$ , and thus, finally process  $a_1 b_1$  {[the malt] [that lay...]}]. Therefore, understanding center-embedded sentences like those above depends on the preservation of all constituents in encoding or decoding the sentence. All the constituents - or more specifically, all the constituents linked according to the



**Fig. 1** Example of center-embedded structure, where 'ab' pairs with higher indexes are embedded in 'ab' pairs with lower indexes

given morpho-syntactic and semantic interconnections—are necessary to grasp the full meaning of the proposition. Importantly, since propositions are *not* mere concatenations of nonsense tokens, but map systematically onto semantic references, a vital indicator of the correct combination among words and phrases is determined by the correct interpretation of the meaning. To show how the semantic interpretation is crucial in determining the syntactic connections in center-embedded sentences, let's take into account once again (i): “**a**<sub>1</sub>[The malt] **a**<sub>2</sub> [that the rat] **a**<sub>3</sub> [that the cat] **b**<sub>3</sub>[killed] **b**<sub>2</sub> [ate] **b**<sub>3</sub> [lay in the house that Jack built.]”. Clearly, the links between **a**<sub>3</sub> [the cat] and **b**<sub>3</sub> [killed], or between **a**<sub>2</sub> [the rat] and **b**<sub>2</sub> [ate the cheese] are not merely perceptual but are determined by the possibility to connect the constitutive items relying on the contextual interpretation of the meaning of the sentence (which can also be driven by the implicit background knowledge that cats eat rats, or that rats love cheese) as well as on morphological rules of patterning, e.g. matching verbs and with nouns.

Hence, the only empirical way we have to assess whether a structure is recursively center-embedded or not is by interpreting its conceptual-semantic value (Fitch 2010). For this reason, testing Hauser et al.'s (2002) “recursive distinctiveness” hypothesis—which implies the use of a comparative approach on nonhuman species' ability to process these type of structures—is highly challenging. A far more empirically tractable hypothesis is the “supra-regular distinctiveness” one (Fitch and Friederici 2012), which, in contrast, concerns patterns of meaningless elements, ruled according to mere perceptual similarities (e.g. matching sounds, colors or shapes). In other words, this hypothesis focuses on a particular type of the faculty of *syntaxis*, which I refer to as “perceptual syntax”.

I assume that the comparative investigation of the ability to process nonsense perceptual patterns can reveal crucial information on both the biocognitive potential of different nonhuman animals and on the evolutionary dynamics of human propositional language. The first comparative study conducted to test the “supra-regular distinctiveness” hypothesis was conducted by Fitch and Hauser (2004). In their research, the authors combined the artificial grammar learning paradigm with the mathematical models used in the formal language theory. In the following paragraph, I will describe this approach in greater detail.

## Artificial Grammars for a Comparative Study of Syntax

### Preliminary Notes on Artificial Grammar

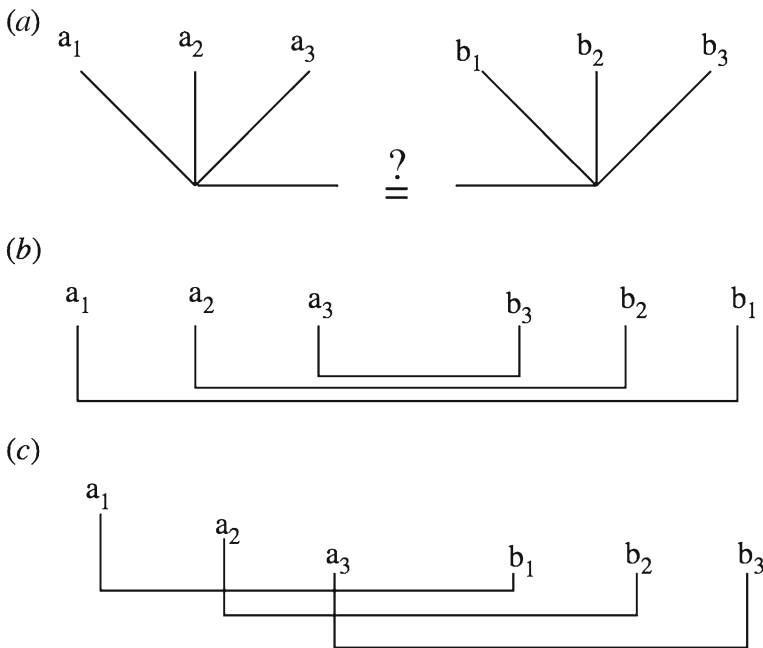
Several studies have been recently comparing the ability of humans and nonhuman animals to process patterns with different levels of complexity. Scientists with this aim were involved in a broad research program, the *artificial grammar learning* paradigm, where an artificial (visual or auditory) mini-language is derived from a set of formal logical rules. In particular, within this paradigm, artificial languages have been created following a system of formal rules (grammar) enunciated by Noam Chomsky (1956) within the generative grammar tradition, in order to distinguish different levels of complexity in language. The linguist has distinguished four classes of formal grammars, which can generate sets of sequences arranged in a mathematical

hierarchy of increasing generative power. For the purposes of this chapter, it is worth focusing on two grammars, namely the so-called “finite state” (or regular) grammars and supra-regular “context-free” grammars (hereafter FSG and CFG). Each of these two formal systems can be generated by a simple mathematical model conceivable as an abstract machine (or automaton) that can be in one of a finite number of states. The machine computing the FSG can be described in terms of the transition probabilities from one state to another and of the triggering condition for each transition. One example of the FSG is the pattern  $(AB)^n$ , where the only rule is to write an occurrence of B right after an occurrence of A for  $n$  times. In order to create such pattern, or to be able to discriminate its formal correctness, all an automaton would need to know is a) the initial input, b) the transition rule to the following state, and c) the number of times after which it has to stop. As only a bounded amount of memory is enough for this kind of automaton to work, this class of grammar is the less complex set of generative rules. In terms of natural language, each state of the machine corresponds to a word, thus the resulting sentence would be a linear sequential concatenation of words or phrases, as in the case of tail recursion.

Notably, propositional language goes beyond the regular grammars’ level of complexity, or in other words, beyond a FSG. This is particularly true, for example, for recursive sentence structures of the example in (i), or for phrase structures with multiple long distance dependencies such as the “if-then” linguistic constructs—where one has to store in memory what comes after “if”, and connect it with the rest of the sentence to form a sense. These latter linguistic structures can be processed by the supra-regular context free grammars, since they require sophisticated parsing capabilities—first, on an abstract level, since one has to be able to recognize so formed patterns - and second, in terms of memory storage. An instance of a supra-regular grammar is the pattern  $A^nB^n$ , where given a certain number of ‘a’s, the exact same number of ‘b’s follows. There are at least three possible supra-regular strategies that could allow to recognize and process this grammar (Fitch and Friederici 2012; Fig. 2): a) a “count and compare” strategy, that is to count all the ‘a’s, store that value in the memory, and compare it to the number of ‘b’s; b) a second strategy is to match each b with the most recently seen ‘a’ (a center-embedded structure); c) match each ‘a’ with a ‘b’ in the order they appear. This last strategy, indeed, yields a so called “crossed” dependency and requires a computational power that goes slightly beyond the one needed to recognize a context free grammar. In all of these cases, as in the case of the propositional language, one has to keep in memory the different parts of the structure, in order to recognize the whole pattern.

### The “Supra-Regular Distinctiveness” Hypothesis

Recently, several studies using an artificial language generated according to the rules described in the formal language theory, have investigated whether nonhumans’ inability to acquire the syntax of a human language could follow from their inability to process patterns that are more complex than the finite state grammars. Within this research framework, the first comparative study was conducted by Fitch and Hauser (2004). The authors compared the ability of a species of nonhuman primate (cotton-top tamarins - *Saguinus oedipus*) and of a group of adult humans to distinguish a regular grammar from a supra-regular grammar. An  $(AB)^n$  structure was adopted as an instance of a finite state grammar, and an  $A^nB^n$  structure as a supra-regular pattern.



**Fig. 2** Image modified from Fitch and Friederici 2012. Three different strategies can be instantiated to recognize a  $A^n B^n$  pattern: a) count all the ‘a’s and compare this number with the number of ‘b’s; b) match each b with the most recently seen ‘a’; (center-embedded pattern); c) match each ‘a’ with a ‘b’ in the order they appear (cross-dependency pattern)

The occurrences of the class “A” consisted of syllables naturally spoken by a female voice; the class “B”, on the other hand, consisted of syllables spoken by a male voice. The syllables of the two classes consisted of pairs of consonants and vowels; for instance, a pattern defined by the  $A^4 B^4$  grammar would be aaaabbbb; where four instances of ‘B’ follow four instances of ‘A’. On the other hand, the  $(AB)^4$  grammar would define the sequence (ab) (ab) (ab) (ab), where the number of “ab” pairs is specified by the exponent. The authors investigated the ability to learn different levels of *syntactical* complexity in humans and tamarins exposing both of them to the same sets of stimuli so created (see O’Donnell et al. 2005).

This study suggests that parsing procedures at the level of a FSG complexity are spontaneously available to both human and nonhuman primates. On the contrary, only humans mastered the supra-regular  $A^n B^n$  pattern. This outcome explains why nonhuman primates have always failed to acquire natural human languages, whose internal structure goes far beyond the sequential ordering of units. Hence, the data of this research suggest that the ability to process structures that are beyond the complexity of a regular grammar is specific to humans. In this respect, two central considerations need to be strongly stressed. First, since the strategies that could yield a supra-regular  $A^n B^n$  structure—as we have seen in the previous paragraph—are multiple, no conclusion about the “center-embedded recursion distinctiveness” hypothesis can be drawn (see Fitch et al. 2012). Secondly, note that this study focuses on tamarins’ ability to use a ruled pattern on a mere perceptual level (no



meanings, nor morphological/conceptual patterns were present in the stimuli). In fact, this “artificial grammar learning” task provides mere evidence that the ability to process meaningless structures—shown to be present in tamarins—could have grounded the evolution of the more sophisticated capacity to process sentences involving combinatorial semantic values. In other words, the ability to distinguish FSG from CFG investigated adopting nonsense grammars does not address the ability to speak in verbal propositions directly. On the contrary, the comparative investigation of this faculty sheds light on a sort of “cognitive skeleton” that might have been a necessary but not sufficient condition for the evolution of human language. For this reason, one has to be cautious in conflating the capacities to process mere perceptual structures and to process verbal language as the same ability.

Recently, several studies have been applying Fitch and Hauser (2004)’s same paradigm on different nonhuman species, specifically on two species of songbirds (Bengalese finches, and zebra finches). Disregarding these latter crucial observations, the authors of these studies have *ipso facto* blended the ability to process  $A^nB^n$  with the faculty to recognize center-embedded grammars. Furthermore, taking the mistake in an even more misleading direction, they have concluded from their positive results that humans share the faculty to “recognize center-embedded grammars” with nonhuman species. As a matter of clarity, I provide a detailed description of the experimental designs adopted in these studies.

Gentner, Fenn, Margoliash & Nusbaum (2006) claim that European starlings (*Sturnus vulgaris*) are able to acquire information about both  $(AB)^n$  and  $A^nB^n$  syntactic structures. Precisely, the authors trained the birds with only two strings, namely ABAB as an instance of  $(AB)^n$  and AABB as an instance of the grammar  $A^nB^n$ . The starlings were trained to memorize motif patterns constituted by two different song basic motifs termed “rattle” and “warbles”. For example, given the rattles as the units *a* of the class “A”, and the warbles as *bs* of the class “B”, a sequence following an  $(AB)^n$  grammar would be *rattle-warble-rattle-warble*, and a  $A^nB^n$  pattern would be *rattle-rattle-warble-warble*. After acquiring the distinction between the training sets, the starlings were exposed to new sequences constructed from novel song phrases of rattles and warbles, or familiar phrases and elements in a novel sequence. The starlings could discriminate the novel strings. However, although the authors claim that these birds were able to process recursive, center-embedded grammars, Van Heijningen, De Visser, Zuidema & Ten Cate (2009) have clearly shown that in order to solve correctly this task, there is absolutely no need of abstract knowledge of the grammar. In fact, replying the study on zebra finches (*Taeniopygia guttata*), they have demonstrated that success in this task might have derived by attending to shared phonetic characteristics between the familiar and the novel, such as the mere presence of AB vs. AA at the beginning of the sequence, or the presence of BB vs. AB in the last part of the trial. The same “perceptual shortcut strategy” could explain the success of another species of songbirds, i.e. the Bengalese finch (*Lonchura striata* var. *domestica*), in discriminating novel instances of  $(AB)^n$  and  $A^nB^n$  after a training session (Abe and Watanabe 2011). In conclusion, we can’t effectively conclude from these experimental studies that starlings and Bengalese finches are able to process the same center-embedded structures that are involved in human language. This is true for at least three reasons: a) as we have seen above, the correct recognition of a recursive pattern requires as a necessary condition the contextual interpretation of the meaning value of each constitutive unit of the structure.



Notably, no semantic values were involved in the patterns adopted in the studies on songbirds; b) it is possible to correctly process these patterns just paying attention to short phonetic segments that are shared between the training and the test stimuli; c) even accepting *per absurdum* that the birds were not using any “perceptual shortcut strategy”, but were truly transferring abstract knowledge about the grammar from the familiarization to the novel test stimuli, it would be more parsimonious to infer that these species can master a supra-regular grammar. Although from these studies we cannot draw any conclusion about the level of pattern complexity that these songbirds can effectively master, it is important to highlight that both Bengalese finches and zebra finches spontaneously *produce* syntactically complex songs; despite lacking a context-free structure, the sequences of syllables (i.e. uninterrupted notes—see Berwick et al. 2011) in their songs are not uttered in a random manner; on the contrary, syllables are concatenated following specific rules that can be conceived as a particular kind of FSG (Okanoya 2004). Therefore, the ability to perceive a supra-regular pattern on the one hand, and to actively produce songs that follow the rules of a finite state grammar on the other - can be interpreted as evolutionary traits shared by humans and songbirds, selected in response to similar external pressures. The naturally following question then is whether what distinguishes humans is their the ability to both produce *and* perceive (or more precisely, to *understand*) well-formed context-free patterns.

Ultimately, I believe that the comparative investigation of the ability of *syntaxein* (on both the perception and the production level) in different species could reveal crucial information on the evolutionary primitives of the faculty to process verbal language. In fact, further empirical and theoretical research is needed in order to solve the following issue: *how* did the ability to elaborate structures on a perceptual domain evolve into the ability to understand propositional syntax? In order to answer this question, it is necessary to pinpoint in further detail the core properties of the syntactic structures determining human language’s specificity. This will be the object of next paragraph, where I will illustrate an important aspect of the nature of human language that in my opinion needs to be integrated in the research paradigm so far described.

### Symbolic *Syntaxis*

In order to answer the last question in the previous paragraph, it’s necessary to return for a moment to the methodological strategy described at the beginning of this chapter: to realize a comparative study on the faculty of language (in its broad sense), one should (i) choose one hypothetical constitutive feature specific to humans’ ability of *syntaxis* and (ii) adopt a comparative approach to nonhuman animals’ cognitive systems to show if they have it. As to (i), the exploration of the “supra-regular distinctiveness” hypothesis should be pushed forward in order to pinpoint in detail how the ability to process combinatorial patterns in humans differs from that in nonhuman animals.

I assume that what makes humans’ *syntaxis* ability specific is the intrinsic possibility that the *interdependencies* between the elements of the linguistic structure have an external referential value; in other words, human language cannot be conceived of as a collection of isolated tokens, with each referring to a specific external meaning. An important clarification of this concept is made in Wittgenstein’s *Tractatus logico-philosophicus* (1922), where the author explains the nature of a verbal sentence referring to it as an image [*Bild*] representing a state of affairs in the World:

The world is the totality of facts, not of things./What is the case, the fact, is the existence of atomic facts [*Sachverhalten*]./An atomic fact is a combination of objects (entities, things)./That the elements of the picture [*Bild*] are combined with one another in a definite way, represents that things are so combined with one another.[...]/The representing relation consists of the coordination of the elements of the picture [*Bild*] and the things./These co-ordinations are as it were the feelers of its elements with which the picture touches reality.  
(§ 1.1, § 2, § 2.01, § 2.15, § 2.1514, § 2.1515)

Here, Wittgenstein claims that what establishes human epistemic access to a fact is the syntactic pattern given in a verbal *Bild*. Indeed, in a verbal sentence the structural morpho-syntactic dependencies [co-ordinations] governing its constitutive units make evident (and at the same time map the perception of) the pattern of connections between objects.<sup>2</sup> Conversely, we cannot perceive single entities of the external world as isolated objects per se. On the contrary, the perception of external objects strictly depends on their possibility to be perceived within a pattern of connections with other objects. In other words, the atoms of reality are not isolated, indivisible entities, but rather cores of connected objects: facts, which are recognized and elaborated through verbal language. For instance, the sentence “Mary invited Paul for dinner” describes a scene whose internal dynamic is expressed and perceived solely according to the logic underlying the sentence itself. Notably, the exact referential value of the verb “invited” results by connecting it to the subject effectively realizing this action, or to a possible subsequent specification of it being, for example, an invitation to a birthday party. Indeed, as Wittgenstein has pointed out, it is exactly these linguistic connections that, as feelers, tap into existent states of affairs in the external world. Moreover, this overall meaning resulting from the “syntactic” structuring of the units (both external objects and linguistic elements) is to be treated as a deriving single, internally complex image (*Build*).

Consistent with this theoretical framework, Deacon (1997) defines humans as the “symbolic species”. In fact, he observes that in general, nonhuman animals possess the ability to recognize several bilateral associations between one auditory or visual token and a correspondent external object or action (indexical connection). Human interaction with the world differs because it is guided by the perception of the relationships between objects *through* the logical relationships<sup>3</sup> between the tokens (Fig. 3). According to the author, this is exactly what makes human cognition unique.

Keeping with this philosophical paradigm of human language, I assume that what differentiates the latter from other animal communication systems is the ability to categorize the units of a pattern going beyond their isolated perceptual characteristics, i.e. in a “symbolic” dimension. Following Deacon’s theory, I refer the term “symbolic” to the uniquely human ability to combine semantic units within a network of logical combinatorial relationships, which map into (being retroactively mapped by) connections among external objects.

In light of these considerations, it shouldn’t be surprising that some nonhuman species have shown the ability to recognize complex *syntactic* patterns: all they were doing was applying acquired rules on a merely perceptual level of operant association

<sup>2</sup> Here the term “object” is used in a broad sense, referring to abstract or concrete entities, as well as to actions, subjects, etc.

<sup>3</sup> The logical connections are those expressed both through morpho-syntactic links and logical connectors.

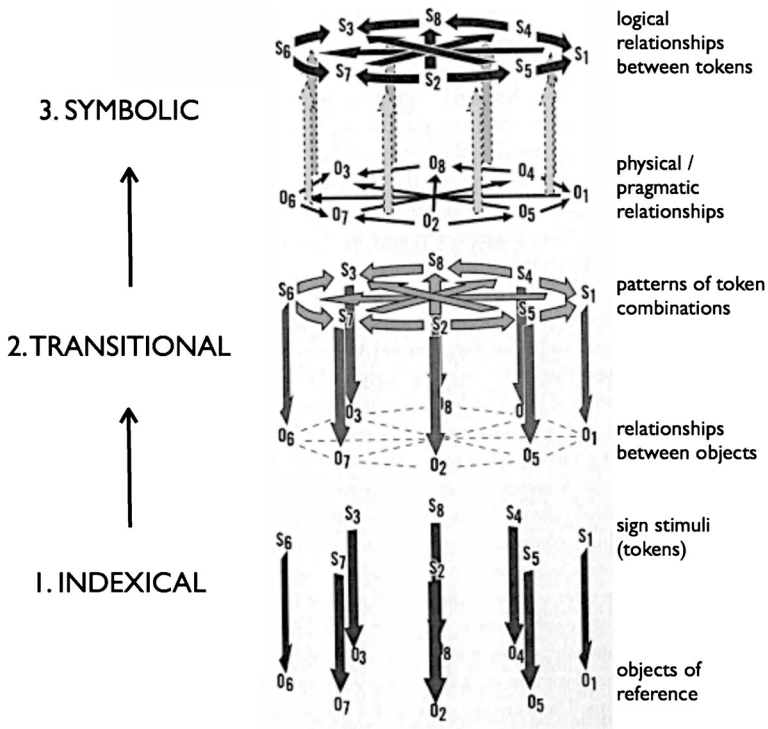


Fig. 3 Image modified from Deacon 1997

between units. The ability to actively process regular patterns on a merely perceptual level—i.e. with no semantic, nor morphological matching values involved—shown in nonhuman animals could, indeed, be identified as a shared cognitive ability, but it is far from the human species-typical ability to process context-free semantic structures. In fact, *only* in humans these patterns have item-based interdependencies, which are linked to connections in the external referential dimension. In human *linguistic* expressions the specific feature distinguishing *syntactic* connections between the constitutive units of a verbal sentence is their intrinsic possibility of a reference (e.g. an object, an action or an emotion). This also applies to other typical human cognitive possibilities such as computing a mathematical expression or drawing a conceptual map.

In conclusion, I hypothesize that what makes humans' general faculty of language species-typical is the ability to process patterns adopting cross-modal mappings. This includes the faculty of *syntaxis* and of applying a referential value (relying on the internal items' interdependencies) as one indissoluble cognitive phenomenon. Hence, a comparative study of language, which aims to investigate the specificity of humans' faculty of language should, in my opinion, apply the following methodology. First, it is necessary to address the faculty of *syntaxis*, conceived in the following operational definition: the ability to associate a combinatorial pattern in a linguistic dimension to a *syntactic* combination among external objects or categories of objects. Secondly, the comparative exploration of the faculty of *syntaxis* has to be more specific - on the one hand - about the specific computational strategies that could be applied to process

a structures, and on the other hand, on the effective combinatorial value of the items used in the given patterns. Specifically, I assume that this framework would significantly favor progress in the exploration of other animals' ability of *syntaxis* if it will focus on their capacity a) to process simple *syntactic* patterns with internal dependencies between the elements, b) to refer these basic *syntactic* structures to a pattern of external objects. These considerations could provide invaluable information on animal combinatorial abilities, as well as on the combinatorial patterns ruling the sentences of any natural spoken language, and other cognitive domains such as vision, music or math. Ultimately, this could allow the understanding of what makes a species-typical human linguistic expression out of a pattern of perceptual stimuli and consequently, help refining the meaning of "linguistic", when applied to the definition of man as "zoon logikon".

**Acknowledgments** I thank Marco Carapezza, Franco Lo Piparo and W. Tecumseh Fitch for their invaluable help in thinking through the concepts explained in the manuscript, and for encouraging my research. Funding was provided by the University of Palermo and by a European Research Council Advanced Grant SOMACCA (No. 230604) awarded to W. T. Fitch. The funders had no role in the decision to publish or prepare the manuscript.

## References

- Abe, K., & Watanabe, D. (2011). Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nature Neuroscience*, *14*(2011), 1067–74.
- Aristotle (2004). *Metaphysics*. (H. Lawson-Tancred, Trans.) England: Penguin Classics. (Original work ≈ 350 b.C.).
- Berwick, R. C., Okanoya, K., Beckers, G. J. L., & Bolhuis, J. (2011). Songs to syntax: the linguistics of birdsong. *Trends in Cognitive Science*, *159*, 113–21.
- Carapezza, M. (2006). Uexküll, la nozione di *Umwelt* e il parlare di animali. *Annali del dipartimento di Filosofia, Storia e Critica dei Saperi, Università di Palermo*, 47–60.
- Christiansen, M. H., & Chater, N. (1999). Toward a connectionist model of recursion in human linguistic performance. *Cognitive Science*, *23*, 157–205.
- Chomsky, N., Three models for the description of language, *Information Theory*, 1956.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: MIT.
- Corballis, M. C. (2007). Recursion, language, and starlings. *Cognitive Science*, *31*, 697–704.
- Deacon, T. W. (1997). *The symbolic species. The coevolution of language and the brain*. New York: Norton & Company.
- Fitch, W. T., & Hauser, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science*, *303*, 377–80.
- Fitch, W. T. (2010). Three meanings of "recursion": key distinctions for biolinguistics. In R. K. Larson, V. Deprez, & H. Yamakido (Eds.), *The evolution of human language: biolinguistic perspectives* (pp. 73–90). New York: Cambridge University Press.
- Fitch, W. T., & Friederici, A. D. (2012). Artificial grammar learning meets formal language theory: an overview. *Philosophical Transactions of the Royal Society*, *367*, 1933–1955.
- Fitch, W. T., Friederici, A. D., & Hagoort, P. (2012). Pattern Perception and computational complexity: an introduction to the special issue. *Philosophical Transactions of the Royal Society*, *367*, 1925–1932.
- Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature*, *440*, 1204–207.
- Hauser, M. D., Fitch, W. T., & Chomsky, N. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science*, *298*, 1569–79.
- Lo Piparo, F. (1998). Il mondo, le specie animali e il linguaggio. La teoria zoocognitiva di Wittgenstein. In M. Carenini & M. Matteuzzi (Eds.), *Percezione, Linguaggio, Coscienza. Saggi di filosofia della mente* (pp. 183–202). Macerata: Quodlibet.

- O'Donnell, T. J., Hauser, M. D., & Fitch, W. T. (2005). Using mathematical models of language experimentally. *Trends in Cognitive Sciences*, *9*, 284–89.
- Okanoya, K. (2004). Song syntax in Bengalese Finches: proximate and ultimate analyses. *Advances in the Study of Behavior*, *14*, 297–346.
- Uexküll, J. von e G. Kriszat (1992). A stroll through the worlds of animals and men: A picture book of invisible worlds. *Semiotica*, *89*(4). doi:10.1515/semi.1992.89.4.319.
- Van Heijningen, C. A., De Visser, J., Zuidema, W., & Ten Cate, C. (2009). Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. *Proceedings of the National Academy of Sciences*, *106*(48), 20538–43.
- Wittgenstein, L. (1922). *Tractatus logico-philosophicus*. (F. P. Ogden & C. K. Ramsey, Trans.). Oxford, England: Routledge and Kegan Paul. (Original work published 1921).