Representational Systems in Zoosemiotics and Anthroposemiotics Part I: What Have the So-Called “Talking Animals” Taught Us about Human Language?

Vilém Uhlíř*

Theoretical and Evolutionary Biology, Department of Philosophy and History of Sciences. Charles University. Viničná 7, 12843 Praha 2, Czech Republic

Received ???, 2018; Accepted ???, 2018

Abstract: This paper offers a brief critical review of some of the so-called “Talking Animals” projects. The findings from the projects are compared with linguistic data from Homo sapiens and with newer evidence gleaned from experiments on animal syntactic skills. The question concerning what had the so-called “Talking Animals” really done is broken down into two categories – words and (recursive) syntax. The (relative) failure of the animal projects in both categories points mainly to the fact that the core feature of language – hierarchical recursive syntax – is missing in the pseudo-linguistic feats of the animals.

Keywords: language • syntax • representation • meta-representation • zoosemiotics • anthroposemiotics • talking animals • general cognition • representational systems • evolutionary discontinuity • biosemiotics

1. The “Talking Animals” Projects

For the sake of brevity, I offer a greatly selective review of some of the more important “Talking Animals” projects. Please note that many omissions were necessary for reasons of space.

The “thought climate” of the 1960s and 1970s was formed largely by the Skinnerian zeitgeist, in which it seemed possible to teach any animal to master any, or almost any, skill, including language. Perhaps riding on an ideological wave, following the surprising claims of Fossey [1] and Goodall [2] concerning primates, as well as the claims of Lilly [3] and Batteau and Markey [4] concerning dolphins, many scientists and researchers focussed on the continuities between humans and other species, while largely ignoring the discontinuities and differences. Fuelled by the focus on inter-species continuities and using the then-contemporary Skinnerian behaviourism as a platform, several scientists independently started the so-called “Talking Animals” projects. The goal of these projects was to show that an animal could learn at least certain features of language, which would prove (latent) linguistic capabilities in non-humans.

In the late 1960s, two pioneering projects were seeking to demonstrate that a chimpanzee could acquire (at least) rudiments of grammar [5–33]. Both used the visual channel of communication, after previous attempts to teach chimpanzees spoken words had failed (the Hayeses [34–37] taught a chimpanzee named Viki in the 1950s, the Kellogs [38, 39] pursued similar research with the Gua Project in 1931–33), because a prevailing contemporary view claimed that it was the position of the larynx that forbade the chimpanzees from articulating phones and phonemes similar to humans [40]. From 1966 to 1970, the Gardners trained a chimpanzee infant named Washoe (1965–2007) to use American Sign Language (ASL), which is a natural human language using the visual channel. She was reared in a home-like environment, where the prevalent medium of communication was the ASL, thus
applying a holistic approach. From 1964 until 1987, the Premacks trained Sarah, a 2-year-old chimpanzee (*1962), to use artificial language, which used non-iconic plastic chips of varying sizes, shapes and colours. The Premacks chose to teach language to Sarah in components. Early reports and texts (claiming that Washoe used combinations of two or three signs) from the Gardners generated a lot of interest – for instance, a prominent linguist at the time, Roger Brown [41], later compared Washoe’s strings of signs to the sentences of a child and claimed to have found similarities in the structures of utterances of both [42], which was, in turn, cited in nearly every following paper by the Gardners.

The central question of the “first generation” of researchers (the Gardners and the Premacks mentioned earlier) had been whether other species, apart from humans, can acquire language. The answer depended on how one defined what language is. The early researchers defined language as the transmission of information via non-iconic arbitrary signs [5, 16, 21–24, 43–49]. Such definitions allowed them to come with a positive answer: yes, animals showed behaviour resembling language behaviour. Had they defined language as the active usage of all the structures that humans use, their trained animals would not pass the test [27–29, 50–63]. Later, even animals that had not been considered much intelligent at the time (the 1970s) in comparison with primates – such as parrots [64] and sea lions [65] – showed behaviour very similar to that of the trained primates.

More projects followed soon after the exposure that Washoe received and secured the Gardners a tighter financing of their follow-up projects. The “second generation” enjoyed similar media attention and generous funding. From 1973 until 1975, Duane Rumbaugh and his colleagues [44, 45, 66–78] trained a juvenile chimpanzee Lana (*1970) to use an artificial visual language using “lexigrams”. Every lexigram was a combination of particular geometric shapes, colour of background, at least in the Lana Project. The Lana Project was the only one using a real grammar – the artificially created “Yerkish” [69, 79, 80]. Sometimes, the term “Yerkish” is applied to the lexigram system, which is incorrect. Yerkish was explicitly a grammar. The same mistake appears again in follow-up projects by the Rumbaughs – the only chimpanzee who was exposed to Yerkish grammar was Lana. Directly inspired by the Gardners, Francine “Penny” [48, 81–85] has been working with a young female gorilla Koko (1971–2018) since 1972; and [49, 86–88] was working with a young orangutan Chantek (1977–2017) from 1979 until 1986. Both Koko and Chantek were trained to use ASL signs. Foutses took care of the Gardners’ chimps and continued the experiments in more or less the same vein as the Gardners a decade earlier [43, 89–100]. All the researchers have claimed that their subjects produced a great number of combinations of two or three signs.

started [162], as such publicity meant better financing and better chances for grants and donations. The early success magnetised more researchers, who started proposing similar projects en masse. For a short while, these “Talking Animals” projects became quite literally a profitable “research industry” with pop appeal [55–58, 162–164]. The researchers quickly produced publications of wildly varying quality, popular books, plethora of articles in magazines and scientific papers, as well as making appearances in television and radio; some of them toured both universities and popular venues.

In the late 1970s, a decade into the inquiry of animal language research, a substantial amount of evidence (sometimes anecdotal) had been amassed, which the researchers themselves interpreted as a proof that primates could not only use symbols for reference but also combine words (symbols) according to (what was claimed as) grammatical rules, and also that they could create new meanings. These collective claims were challenged from (at least) three independent directions.

First, from a project, intended as a replication of the Washoe Project by Herbert Terrace and his colleagues [46]. Nim (1973–2000) was a young male chimpanzee reared in an ASL-rich environment. The goal of the Nim Project (running from 1973 until September 1977) was to get a substantial amount of evidence of grammatical structures in the chimp’s sign productions. Initially, the first superficial look at Nim’s productions of sign combinations led the researchers to think that they were generated by simple Finite State Grammar. In the late 1970s, the Nim Project seemed to be the strongest evidence of the linguistic skills of a non-human animal [46]. A closer look at the videotapes of Nim’s signings, however, revealed that Nim’s combinations were long requests for rewards, and Terrace and colleagues decided that there was no basis for interpreting his signings as sentences. Moreover, most of his longest productions occurred when his trainers did withhold the rewards, which meant that none of his signings were truly spontaneous. This led Terrace to notice that the trainers were so focussed on Nim’s signs that they did not notice that Nim only imitated their previous signings and that he used the signings simply to get a reward.

The debate came to a head with an article in Science by Terrace et al. [50], in which Terrace stated that Nim was unable to acquire even fundamental grammar.¹⁵ The science community [163, 164, 172, 173] accepted Terrace’s claim as a final proof that primates cannot produce real sentences. This claim holds very well even today – not even the much-celebrated and acclaimed Kanzi could produce sentences [57–61]. Terrace claimed that his analyses of the available footage of other signing primates (Washoe, Koko and Chantek) proved the same in the context of other signing apes [170, 171]. The other researchers loudly protested, and the Gardners even threatened with legal action if he would ever use photos of Washoe, their chimp.

Second, from the analysis of the Lana Project by Thompson and Church [174]. The authors analysed an extensive corpus of Lana’s productions and concluded that the majority of these combinations were generated by conditional discrimination learning. Lana learned stock sentence templates, which she used according to simple rules – whether the reward was in view or not. They concluded that while Lana did understand the meaning of the lexigrams that stood for particular rewards, there was no evidence that Lana understood the meaning of the remainder of the stock sentence lexigrams. Similarly, Terrace and Premack [27, 29, 168, 170] later noted that while a primate can learn an arbitrary string of symbols, there is no certainty that the primate uses the purported meanings of these signs that the researchers glossed it as. A sequence glossed as “please machine give cola” could just as well have been glossed with randomly chosen letters “K X C V” (or other non-iconic symbols). For example, similar experiments without the language glossing of the strings, were carried out on pigeons and rhesus monkeys. In an experiment by Straub and Terrace [175], pigeons were required to learn lists of arbitrary photographs. A later experiment by Terrace et al. [176] with rhesus monkeys showed that they have remarkable skills in learning lists of nonsense syllables used in tests of human memory. The lists that Terrace’s monkeys had memorised were considerably longer than the sequences memorised by language-trained apes [176]. These skills of monkeys seriously undermine the claims that the sequences produced by language-trained apes might be meaningful or grammatical.

Third, the debate, initiated by Terrace, peaked in 1980 when psychologist Robert Rosenthal and biosemiotician and linguist Thomas Sebeok organised a conference named “Conference on the Clever Hans Phenomenon: Communication with Horses, Whales, Apes, and People” (May 6–7, 1980, at the New York Academy of Sciences [163, 164]). Linguists, psychologists and animal trainers were invited to give speeches (among the speakers were Herbert Terrace, Laura Petitto [both from the Nim project], Duane Rumbaugh and Sue Savage-Rumbaugh, Thomas Sebeok and Heini Hediger). The conference came to a conclusion almost unanimously – skills in serial ordering are

¹⁵ Moreover, for evidence, see Refs. [50–54, 165–171].
not sufficient for language and trivial ordered concatenation does not prove grammar; animals cannot acquire human language (especially syntax) and their behaviour was not genuinely linguistic. As far as I know, there is no record of what exactly transpired there, but literature mentions reported gossip of ad hominem verbal attacks between the Rumbaugh and Terrace with Petitto. It is worth noting that similar attacks apparently transpired already between Terrace and the Gardners and between Terrace and Patterson at that point (hinted at in various texts by Terrace). The conference was seen by the scientific community as the last nail in the coffin of the “Talking Animals” projects. This, however, did not happen. It was an end of what we can retrospectively claim to be the “golden era”, certainly the most medialised, notorious, layman friendly and (in)famous era of such research.

Consequently, the “Talking Animal” projects came under scrutiny, and several critical studies were carried out.16 The studies differed in politeness but were remarkably similar in conclusions [57, 58]. For the sake of brevity, I will present only the recurring and most prevalent arguments.

The projects were criticised [163, 164, 172, 173] as being focussed on the wrong questions: they would have helped with our understanding of human language far more had the researchers been asking not whether animals can acquire human language, but instead what they can reveal about the origin of human language [55, 172, 173].

If we look at the projects, we can see that the methods used differed considerably: all animals were trained to associate arbitrary signs (manual gestures in ASL, plastic chips, lexigrams, and spoken words) with (mental) concepts. All animals were supposed to concatenate the signs into strings that the researchers glossed over as sentences – and subsequently most researchers claimed that their trained animals produced copious amounts of two- and three-sign combinations. There were projects with both an excessive amount and turnover of staff and personnel (Nim), as well as too small a number of caretakers for the ape to socialise (Koko). There were studies with both inadequate knowledge of ASL among teachers and volunteers (Washoe), and studies with the inclusion of native ASL signers (Moja, Tatu and Dar). There was a wide scale between the rigidity in teaching procedures (routinised drilling in classrooms – Nim and Sarah) on the one end and the informal casual training and flexible approach (Chantek and Kanzi) on the other end. Some apes were taught to use plastic chips (Sarah), some to use lexigrams (Kanzi and Lana) and yet others to use ASL signing and gestures (Washoe).

Another often-appearing argument was that the researchers were not linguists themselves – hence, they introduced no grammatical items. The vocabulary of the apes was limited to lexical items and omitted functional words and grammatical items. The fairness of this argument depends on the reader’s consideration. With no knowledge in linguistics, the researchers did not see the key function of the grammatical items in sentence production and possibly thought that they were insignificant. Or perhaps the lack but illustrates the near-impossibility to teach grammar to non-humans (for more details, see Section 2.3 on syntax).

The dispute over whether animals do possess language capacity, as well as whether they are able to acquire a language-like system, is a textbook example of a scientific controversy. There are two mutually contradicting and incompatible views on the subject. The scientists were emotionally invested in their research, because something significant to the ideology they adhered to was at stake. Personal incentives were at stake too – the first scientist to teach non-humans language would become scientifically “immortal” and famous. The disputes were brought to the boil once the accusations of fraud and data forgery appeared [52, 170].

What is at stake in this controversy is the linguistic capability, which has traditionally been seen as the core skill for the self-definition of humans, the skill that sets human essence apart from that of other animals. Language has been seen as the source of human transcendence, as well as the divide from the remainder of nature. The importance of language grew proportionally to the discovery of animal skills that were seen, not long ago, as exclusive to humans: tool use [2] and even tool manufacture [178, 179]. Language, however, differs somewhat from the other claims of human exclusivity – there is no universally accepted definition of language. There is no doubt that the “Talking Animals” experiments were motivated by the desire to tear down the divide between humans and animals and show it as an artificial divide created by humans [95, 117, 124, 157].

There are two different approaches to language that are relevant for this paper – the nativist stance [180–182], which sees language as a complex, unified adaptation for communication, one that is species-specific to humans. Proponents of this approach argue for a holistic approach to language and argue that human language, as a whole, is a qualitative evolutionary discontinuity. The second is the mosaic approach used by Chomsky, Fitch and Hauser [183–187], wherein language is fractioned into sets of sub-skills for easier inter-species evolutionary comparison. The mosaic approach focusses on evolutionary continuities and largely ignores the discontinuities.

16 For sympathetic summaries of the research, see Refs. [77, 78, 162, 177]. For less than sympathetic summaries, see Refs.[57, 58].
solution Chomsky, Hauser and Fitch put forward in their papers suggests that the pseudo-linguistic feats of trained animals could very well be due to either homologies of the relevant organic neural structures found in humans or convergent evolution of analogous domain-specific systems. Due to the mosaical approach to cognitive skills, the experiments gleaned information about the syntactic capacities of animals (discussed in Chapter 2.3), which led Chomsky, Fitch and Hauser to formulate the “Recursion-only” hypothesis, which stated that linguistic syntax is, in fact, co-extensive with cognitive recursion (“narrow language” in their terminology) – which might or might not be what differentiates humans from animals.

The general conclusion, or at least a conclusion consistent with the conventional analyses of these projects, seems to indicate the following: despite this manifold diversity in the “Talking Animal” projects, there were no apparent and remarkable differences in both the responses to language training and the eventual accomplishments of these animals. The considerable and substantial variability in the projects brings no considerable differences in terms of their outcomes. This probably signifies that the problem lies with the animal students, not their teachers. The aim was to show the continuity of language across species. What was shown instead was that the linguistic capacities of other species were limited – but limited in a very interesting way, as the more-recent experiments, which will be discussed later, show. In order to see how limited the achievements of the trained animals were, the question of “Language Acquisition” must be fractioned into two sub-questions.

2. Acquiring Language: Signs and Syntax

The question of Language Acquisition can be divided into two sub-questions: 1) Can animals acquire signs (as words)? and 2) Do animals have a capacity for creating syntactic units made of such signs (sentences)? In other words, the question of language is a question of semanticity and syntax. This dichotomy curiously reflects the divide in linguistics itself [188].

2.1 Words

In the minimal definition, a word (units in spoken or signed language) is an arbitrary association of phonology (or temporal progression of a gesture in the case of signed language) and a concept (meaning – the privacy of meaning dilemma [189–192] had to be omitted for the sake of brevity), stored in the language user’s vocabulary (the lexicon). The textbook definition is that a word is the smallest element that carries a meaning in isolation, thus being contrasted from a morpheme, a smallest unit that can differentiate a meaning. However, such definitions are incomplete as the meanings of words are not defined just by the relation of word to a meaning but also by the relation to other words in the lexicon (for lexicon, see Refs. [193 pp. 130–31, 194]; for a broader approach, see Refs. [195, 196]). The lexicon, thus, is a hierarchically organised net of sets of meronyms, antonyms, synonyms and so on. This net is organised into hierarchical hyponyms and hypernyms (for hierarchical recursion in lexicon, see Ref. [197]). Moreover, words (with the exception of proper names) are generic and refer to kinds, not specifics. This is true even in children as young as 9 months [198]. Words are not just names – labels for things (not even in small children [199]) – but contain grammatically encoded properties (syntactic categories, irregular morphology, transitivity on verbs, count–mass distinction on nouns, and so on [200, 201]; for evidence in early language acquisition, see Ref. [202]). This (meta)information, distinctive to each word, is stored in the lexicon and determines how the particular word enters into relationships with other words and into interaction with morphology and syntax. This is, however, not identical to the reservoir of the general knowledge of words in a particular sentence [203, 204]. Function morphemes are also part of the lexicon, but the grammatical and semantic information they carry (finiteness, case, agreement, voice and so on) does not substantially alter the semantics of the phrase. Function words (tense markers, prepositions, auxiliary verbs and so on), however, do the most work in human language. Moreover, they are fundamentally tied to syntax and, to be acquired, the language user needs syntactic skills [63, 205, 206].

It seems that a great part of human usage of words consists of the information that is governed by recursive syntax. Therefore, word learning should operate at least partially independently of learning other types of knowledge. The mechanisms of word learning in children seem to be different from fact-learning mechanisms ([207, 208], although this is contested – see Ref. [209]). See Ref. [210] for evidence from neural network modelling, implying possible differences in the neural substrate in word learning and fact learning. For example, both Waxman and Booth [207, 211] and Diesendruck and Markson [208] showed that children do not attribute knowledge of facts to the members of the speech community in the same way that they attribute to them the knowledge of words. It was
also shown that children treat novel words differently than novel facts [212, 213].

Hauser et al. [183 p. 1576] noted that (paraphrasing) the qualitative differences in scale, rate and mode of acquisition of words between human children and non-human primates are so vast that the possibility of an independent mechanism has to be seriously considered. Non-human animals probably use domain-general mechanisms to acquire and recall words (noted by Hauser et al. [183]; for more-nuanced linguistic analysis, see Refs. [57, 58]) – although children in early stages of language acquisition may also use domain-general mechanisms [199]. Bloom and Markson [214] have argued that children’s Theory of Mind (TOM) is responsible for word learning; if that is so, there is evidence that primates do not have TOM [215, 216]. The rate and the way in which children build lexicon are simply qualitatively different from those of non-humans.

As Edward Kako [217] rightly notes, no one has yet tried to teach animals in the “Talking Animals” projects to use function words or function morphemes, as these might be the most important for an objective comparison between the syntactic capabilities of humans and animals. Moreover, the usage of most words of natural language is not conditioned to specific functions in particular conditions as is the use of signals in animal communication systems (alarms calls, food calls) [218]. This is referred to as “Displacement” [57 ch. 5, 58 pp. 29–30]. In other words, the usage of words is detached from the here and now, in addition to being at least partially context dependent. Word usage is referentially intricate in the sense that there is usually not a simple straightforward relationship between the word and the thing it references; this trait has only very weak analogues in animal communication systems [58 chs. 4–9].

The claims that apes can acquire large vocabularies [11, 67, 68, 83] have been generally accepted and less criticised in comparison to the claims that primates can master syntax. Possibly because the claims about vocabularies are less controversial than claims about syntax, these claims have undergone noticeably less rigid scrutiny. It has been shown – especially in the experiments using ASL – that the sizes of the vocabularies have been exaggerated by the researchers: a substantial fraction of the utterances of the signing apes consisted of inborn, species-specific gestures rather than newly learned symbols [6 p. 137, 58].

Likewise, there is no clear evidence that the signs that the trained animals learned have truly functioned as referents for distinct things (or kinds or classes of things) and not as requests for rewards (this is referred to as “Non-instrumentality” [57 pp. 66–67, 58 p. 298]). The fact that animals have learned to use symbols, however, does not yet mean that the researchers have proven that the arbitrary signs had genuine meaning for them. The richness of internal representations of meaning is fundamentally difficult to assess. The apes could have been using absolutely different meanings than the researchers thought they were using, as the experiments in the 1970s were not sophisticated enough to control for the reference. They learned signs not as meaningful linguistic symbols with reference functions, but rather, as effective procedures or performances. Sign strings were instrumental habits, their means to keep getting rewards. This interpretation is firmly anchored in the fact that, even in the most “advanced” apes such as Kanzi, the signing for reference (signing for its own sake, not for immediate gratification nor other behavioural reward) was extremely rare [57–59, 61].

Morphology determines the principles of combining words (and morphemes) into complex compounds 17 (e.g. by affixing inflections to word stems). For simplification, morphology can be thought of as syntax at the level of the word. Morphology is manifested in such phenomena as compounding, derivation, inflection and so on. Proper meaning is constructed on the basis of the context (linguistic and non-linguistic). With one exception (only in the Lana Project), no other chimpanzee was trained to use any explicit grammar; none of the artificial “language” systems taught to the apes had contained anything like morphology. Every natural human language is able to invent new words on the basis of already existing ones (via compounding, derivation and inflection), yet only anecdotal evidence exists in the “Talking Animals” projects [5, 6, 9, 43, 82]. Extremely important, in this sense, are bound morphemes, which by combining with word stems carry a grammatical function or shift meaning (e.g. plural markers, tense markers and case markers).

In order to avoid getting my arguments intentionally twisted into easily refutable caricature, it should be noted that this does not mean that animals are not semiotic beings and that they cannot use signs. Quite the opposite 18 – as we will see in the following arguments, their pseudo-linguistic behaviour would be impossible without semiosis. This section shows that as far as we know, no animal in nature and no animal in a laboratory uses a system that involves words in their true linguistic sense, as described and contrasted in the preceding sections.

---

17 Compounds in a general sense, as “Compound” has a specific technical sense in morphology.

18 Very important, in this context, are Ramon Ferrer-i-Carrión’s [219, 220] papers proving that dolphin whistles approximate a Zipfian distribution.
Because the current convention considers language crucially depending on recursion, or even being co-extensive with it (the “Recursion-only hypothesis” of Chomsky, Fitch and Hauser), recursion must be discussed in brief, before syntax.

2.2 Recursion
Recursion is a computational procedure that occurs when a thing is defined in terms of itself or when a constituent of that thing is containing a constituent of the same kind (for recursion in general, logic and mathematics, see Refs. [221–224]; for recursion in language, see Refs. [225–228]. In computing, recursion refers to a situation where a function being defined is applied within its own definition [229, 230] (for a sophisticated computer science-based discussion, see Refs. [231, 232]). In theory, recursion allows an infinite chain of references (in practical computing, this looping is avoided) [233]. In computing theory and formalistic linguistics, various types of recursion are distinguished (see Refs. [227, 234] for more models of recursion in linguistics), of which two are important for the topic of this article:

True recursion is a situation where a constituent has an identical constituent embedded within itself. for example, a computing procedure that has to compute another instance of itself (or an equivalent procedure) in the middle of the computation, and after finishing computing, this embedded instance of itself has to resume the original computation from where it left off. This procedure requires a mechanism of a memory device (stack of pointers, for instance) to indicate where to resume the original computation after the embedded constituent has been completed. A sub-type of True Recursion, relevant to linguistics, is called Centre-Embedded Recursion (in linguistics, it is called “centre embedding” [235]; for a review of linguistic literature up to the early 1980s from a different stance, see Ref. [236]), which is a process of embedding a phrase in the middle of another phrase (of the same type). The most frequently used example involves embedding of a relative clause inside another one (see e.g. the nursery rhyme “This Is the House That Jack Built” in English). Usually more than one level of centre embedding occurs only sparsely in common discourse – Karlsson [237] noted that the maximal centre-embedding recursion is three clauses in written language across various Indo-European languages and two in spoken language – this means that double and triple centre embedding is rare [236, 238]. This constraint, however, concerns only centre embedding, not embedding in general.

The second type is Tail Recursion (sometimes called “Tail Call” in computing), in which a constituent has an identical constituent in its periphery – for instance, a computing procedure that executes the original computation by starting another instance of itself as its final step. For the sake of simplicity, Tail Recursion is a simple iteration, a simple adding of constituents (to a string, phrase or sentence). All the constituents are necessary for full meaning as the constituents are not unrelated. Repeated iteration is a common feature of animal calls for instance.

2.2.1 Recursion and Hierarchical Discrete Infinity in Language
Recursion plays an important role not only in syntax [225–228] but also in semantics [239, 240] and lexicon [197]. Recursion is sometimes described as permeating the structure of the whole language [237, 241, 242]. This might not be fully true – the case of phonology has to be discussed in some length.

While not technically recursive, the phonological structure of language is segmental, discretely infinite and hierarchically structured (although speech is the production of a linear string of phones – for further information on hierarchical structure, see Ref. [243]). Though every human language has an unlimited number of structures built from a finite pool of phonetic segments, languages differ in intonation patterns and in constraints affecting how one sound can influence the pronunciation of other sounds [244]. Recursion does not exist fully in the phonological structure of human language [243]. A full syllable cannot be embedded in another full syllable; language works only with full syllables, which cannot be combined into more syllables, although they can be serially joined in a chain one after another. And there seem to be good adaptive reasons for it – Hockett [245] noticed that the distinct combinatorics in the phonological structure of language (“Duality of patterning” – for discussion, see Refs. [246–248] – one level combines sounds with no meaning into morphemes, and another level combines morphemes into words) allowed the combination of a limited pool of sounds into strings (morphemes into words). This is an “economic” adaptation – it does not require the listeners to pursue increasingly finer discriminations of similar sounds [249–251]. Many phonological rules in fact enhance the ease of discrimination of sounds, thus fulfilling the parity of use of language – to be effectively used by both hearers and speakers in a speech community [252–254].

The principles of combination and holistic adjustment mechanisms in phonology seem to have a different neural substrate than motor programmes (see, especially, evidence from dyslectic children or children with phonology...
problems [255, 256]). Humans learn to articulate speech sounds based on what they hear, whereas the ability to imitate in monkeys and non-human primates is rudimentary at best [257–261]. But there are at least four other species of mammals [262–267] and three species of birds [262, 268–271] that are extremely good at vocal imitation [272]. There is considerable work on the possibility of deep homologies with these animals (e.g. the work on FOXP2 in songbirds [262, 273–278] and these are the types of abilities that are potential candidates as precursor capacities for the Faculty of Broad Language (Chomsky and colleagues) [279, 280]. In humans, the transfer from phonological perception to phonological production is supported [281]. However, since mirror neurons (although this hypothesis is very contentious now), which are activated both in perception and production of motor behaviour, are present in monkeys [282], it seems that transfer from perception to production pre-dates the evolution of human language (see also Whiten’s study [283] on the mirror neurons and imitation in primates and humans). In humans, the brain areas involved in semantic and syntactic processing are mostly the same for perception and production [284], although the motor system does not provide a crucial contribution to perception.

Combinatorial properties of phonology (i.e. not True Recursion) have analogues in birdsongs [285, 286] but not in natural primate communication [287, 288].

Language possibly needs to be recursive because its function is to express and communicate recursive human thoughts (were thoughts not recursive, the communication system would not need to be recursive). This view [289] of language as an instrument for expression of thought (Generative Grammar’s interface with conceptual–intentional systems – see “intensions” in the last chapter of Heim and Kratzer [290]; for semiotic interface with grammar, see Ref. [291]), however, vitally depends on whether past hominids and (all) primates did have language “only” for communication. In today’s humans, language certainly has not a single adaptive function [186]. Opinions, however, vary whether language itself derives from prior recursive systems [180, 181] or whether the recursivity is an evolutionarily novel trait [184, 187].

2.2.2 Recursion and Hierarchical Discrete Infinity in Other Cognitive Domains

Some scientists claim that recursion is not unique to human language, and that we can find it dispersed all across human cognition [292–295]. We can find hints of recursion in a visual analysis of objects, especially in the decomposition of a visual representation of an object into parts [296–300], in the representation of complex intentional action patterns and their planning [294, 301–304] and social cognition (in non-humans [305–307]; in humans [308–311]). However, it should be noted that these demonstrations of recursivity in non-linguistic domains are not unambiguous.

A disputed case of recursivity involves numbers – recursive numbering is not an anthropological constant [312 ch. 1, 313–316]. The majority of human cultures (until they were influenced by the Western civilisation) did not have recursive numeral systems [317–319]. It seems that those cultures that have well-developed recursive numeral systems derived them from linguistic recursion [320].

Hierarchical and discretely infinite (though not recursive, despite claims to the contrary in Marcus [321]) organisation of features can be found in other domains (for sophisticated discussion, see Refs. [294, 296]) – in motor control and especially in action sequencing of meticulous manual manipulations [322–324, 193], in solving complex problems [325], in human navigation [326], in processing visual objects into spatial hierarchies [324, 327] and in intentionality and planning [294, 328].

2.3 Syntax

Syntax is the set of principles by which words (and morphemes) are chained into sentences and phrases. Syntax is the capacity to recursively combine and recombine words, phrases and sentences into an infinite range of meanings, thus gaining limitless expressive power. Linguistic syntax (as opposed to the concept of syntax used in computing theory for instance) expresses abstract hierarchical structures (involving true recursion) in rearrangements and permutations involved in particular phrases and sentences. This level of hierarchical recursive organisation goes far beyond the simple (in comparison with human language) mechanisms of concatenation (serial chained iteration) that are sometimes invoked in literature describing animal communication [329–332] and beyond the claims of consistent symbol position, which are found in many of the “Talking Animals” projects. Syntax is “not” consistent linear ordering (concatenation) of symbol positions (which may be a pre-syntactic mechanism of communication used in proto-language [193]). Syntax is the most important mechanism in the system of overall language, as it determines how the meanings of words and morphemes are combined into the meanings of sentences (and phrases), often with changes or shifts of meaning, corresponding to permutations and rearrangements of syntactic structures.
There is a multitude of greatly different descriptions of human syntax. In theory, there are multiple systems of rules, which are capable of generating grammars of different generative powers [333–335]. The most useful and most used distinction is the distinction between Finite State Grammar and Phrase Structure Grammar [333, 336]. Finite State Grammar operates with a finite number of states (simple concatenation of items, which includes superficial parsing), in which the constituents are specified by the transitional probabilities between them. Finite State Grammar is usually used in abstract automata theory and in theoretical computing. This grammar is, however, unable to generate all the structures of any particular human language [333, 336]. All particular human languages work on a level of superior complexity – (on at least the level of) Phrase Structure Grammar [334, 335].

Phrase Structure Grammar allows complex embeddings of phrases within phrases and allows long-distance dependencies, which requires sophisticated parsing abilities, as the components of one string are related to other components some distance away, or even to components of other strings at even greater distances. The consensus is that all human languages operate “at least” on this level of parsing ability [337–340]. However, the generative infinity of grammars makes the presence of the more-sophisticated generative grammars hard to prove, as the limited output of the tested subjects (often non-humans) remains ambiguous to interpretation. Memory and attention limitations and deficits render claims that the subjects have mastered the more-sophisticated grammars inconclusive.

Syntax has several different mechanisms for expressing semantic relations [341 ch. 7, 342 chs. 5 and 7], of which only few are relevant for us in the follow-up paper, as none of them appeared in any of the “Talking Animals” projects:

1. Phrase Structure and function words – which hierarchically chain words into phrases, where the order corresponds to the chaining of the constituents’ meaning (in prototypical cases, where strings of words cannot be grouped ambiguously) [343–345]. A phrase consists of (at least) two levels, the head of a phrase (usually a single word, although it cannot be a phrase) belongs to the same class as the phrase (verb phrase has a verb as the head), which is linked to its complement(s) (it might consist of several phrases embedded within each other). Ambiguities might arise when a head has more than one complement, which is why sentences in human languages include further information relating to the phrase structure via the use of function words (tense markers, prepositions, auxiliary verbs and so on) and functional morphemes (finiteness, case, agreement, voice and so on), which help to make the meanings unambiguous to the language users. Function words and morphemes simply reveal that the phrase structure involves hierarchical embedding and not linear stringing, thus allowing long-distance dependencies (e.g. the X-bar theory describes the vertical relations in a syntactic tree, but not the horizontal relations [345]).

2. Rules for linear ordering of words in particular languages. This means the order of words within sentences (specifying the place of a verb in a sentence) or the order of topic sentences and so on [341, Appendix 2]; for a highly technical summary, see Ref. [346]; for explanation of the phenomenon of free word order, see Ref. [347].

3. Agreement of grammatical categories [348 chs. 3 and 7, 349].

4. Case Marking – the use of inflection depending on the grammatical function that a particular word performs in a phrase [350–352].

These above-mentioned mechanisms are often deployed redundantly [353] – usually, multiple mechanisms working in parallel do reinforce each other, yet in other circumstances, some of the mechanisms dominate over the others. This balance or predominance depends on the context of a sentence and also differs between languages.

Particular natural human languages differ in the extent they rely on these (and other) mechanisms to convey the main semantic relation of the Propositional structure (“who did what to whom (when and where”) (for neurological evidence, see Ref. [354]; for cognitive evidence, see Ref. [355]). Propositional structure and argument hierarchies have been shown to have a neural basis [356]. Though various models of grammar do differ in the description of this basic structure [357], the consensus is that the Argument structure (as the manifestation of Propositional structure in syntax) is the core of language [358–362]. Though various models of grammar do differ in the description of this basic structure [357], the consensus is that the Argument structure (as the manifestation of Propositional structure in syntax) is the core of language [358–362]. Argument structure is distinct from and parallel to Phrase structure, and in sentence production, both structures are mapped together. Propositional and Argument structures are at the heart of every human utterance, even though particular languages differ in the extent of what they need to express explicitly and what remains latent and needs not to be expressed explicitly. In short, no sentence is complete without a subject and a predicate, which is expanded by arguments.

Language with a recursive syntax is an anthropological constant and a human universal, as it emerges spontaneously in ontogeny [339, 340] (although it is possible that syntax is applied to the visual channel instead of the auditory channel in the case of sign languages). Most modern linguists (at least those adhering to Generative
Talking Animals and Human Language

Grammar) agree that recursion is the core of language, an indispensable computational ability unique to humans [226, 228, 241, 243, 363].

For the sake of completeness, a special case of the Amazonian language Pirahã [364, 365] has to be discussed shortly. This language has been claimed (by Everett) as lacking any recursion (as it comprises mono-clausal constructions connected without embedding – for critical reassessment, [366, 367]). Pirahã is still a human language, qualitatively different from anything found in animal communication. Moreover, if languagefaculty [193] is a toolkit, not all languages have to use all the tools all the time. The absence of recursion in one language would not change the fact that the remaining languages do use the tools for embedding clauses. This evidence (had it been correctly assessed by Everett) does not change the fact that recursion is (an innate) part of the language toolkit. This special case also seems to point out that communication is possible without recursion [241, 368, 369], thus strengthening the argument for the case of the recursive syntactical core of language not being “for” communication per se.

2.3.1 Recursion and Syntax

Recursive syntax is not merely co-extensive with hierarchical organisation. Syntactic trees have a characteristic structure (see for example the X-bar theory of phrase structure [345]). In agreement with both Hauser and Fitch and Pinker and Jackendoff [180, 181, 183, 186], one can deduce that syntax is not merely an externalised representational system that happens to be recursive – rather, its multi-directional properties (in the sense that it can move along and between production and perception) point out that language is more likely something akin to meta-mapping between or along other systems. Language is most probably an interface among other representation systems, recursively meta-mapping between (or among) them. This might be the evolutionary novel core of language. Fitch and Hauser claim that language is an externalisation of a single recursive system [183–187], while Pinker and Jackendoff claim that language is a connective tissue between already-recursive systems [180, 181].

There appears to be no unambiguous evidence that non-human species can understand and master recursion. In order to demonstrate True Recursion in an animal, it would be mandatory to unambiguously show that the animal can discriminate the strings according to the rules they were constructed by. Moreover, it would be needed to demonstrate that the constituents in the sentence are seen as being in some relation. For example, in the case of centre-embedding, the subject needs to understand that the elements are “bound” from inside out (or outside inwards) in such a way that the elements at the beginning and at the end of the string are associated and related, and that the following next-to-first and next-to-last are also associated and so on. Until this would be unambiguously proven, there would be no irrefutable and convincing evidence that any non-human species is capable of truly recursive parsing of syntax. So far, all that was proven by the “Talking Animals” projects was that animals can be trained to master rules for linear ordering of words in particular languages. Let us now consider another corpus of existing evidence, which we can judge according to what we have stated so far.

2.3.2 Syntax in Animals

By analogy, in animals, syntax should operate above the level of the single call; yet even after decades of searching for recursion in animal communication systems, hierarchical linguistic recursion (or its equivalent) was not found in animals in the laboratory, in the wild, nor in trained animals. There are, however, claims of exceptions: see Ref. [370] for claims of hierarchy in canary song; see Ref. [332] for combinatorics in Diana and Campbell monkeys; also see vocal improvisation in whale songs [371].

Fitch and Hauser [185] carried out an ingenious experiment on cotton-top tamarin monkeys (Saguinus oedipus), which relied on behavioural measures in order to find an answer to the question of whether an animal can extract a (grammatical) rule from a set of strings generated according to this rule and learn to use different behavioural responses to each rule. The results have shown that tamarins are unable to process acoustic strings with long-distance dependencies (which are manifestations of Phrase Structure Grammar [343, 344]). It must be mentioned that the rule for proving the presence of Phrase Structure Grammar was “the number of As must be equal to the number of Bs”. Although tamarins were clearly able to process sequential regularities of Finite State Grammar (“A must be followed by B”), they could not handle Phrase Structure Grammar (were not able to differentiate AAAABBBB from AAABBBBB). Moreover, the authors claim that this computational limitation is not a result of constraints in memory or attention. In comparison, humans have mastered and distinguished both grammars easily. It seems that non-humans are stuck trying to interpret strings generated at a higher level of sophistication, with equipment being able to process only strings generated by a lower level of sophistication. The authors warned against an
incorrect assumption regarding their experiment – namely that their methodology concerns recursion in syntax. Their experiment addressed neither syntactic recursion nor embedding [372].

In a follow-up experiment by Gentner et al. [373], the same approach that Hauser and Fitch applied to tamarins was applied on starlings (a species of songbird known for elaborately structured songs). The results suggested that starlings could master a set of rules above Finite State Grammar – the authors have claimed that starlings could work with Phrase Structure Grammar (based on the “the number of As (rattles) must be equal to the number of Bs (warbles)” rule). Starlings could discriminate strings generated by Phrase structure rules up to \( n=4 \) (four rattles or warbles), although high number of trials was required for the starlings to learn the discrimination. It is, however, not clear whether the distinction between (hierarchical) phrase structure and recursive structure was clearly used in these experiments. Considering that the results were based on the previously stated rule (same number of rattles and warbles), the results raise a question whether the birds understood the strings according to a recursive rule (Phrase Structure Grammar) or whether they adopted an alternative rule to differentiate these strings from Finite State Grammar strings. The simpler, alternative strategy would involve iterations, where the birds counted the number of successive rattles and then the number of successive warbles and judge the numbers as equal. This simpler strategy does not involve Phrase Structure Grammar at all, yet it requires memory and an ability to match quantities – if not a straightforward counting mechanism. We know that songbirds do specialise in the production of repetitive sequences [374] and the perception of numerosity in auditory sequences possibly surpasses the same skill in humans [375–378]. Pigeons have comparable numerosity skills [377, 379–381]. The African grey parrot (Psittacus erithacus) is able to count up to six [155, 158, 159] (the starlings were tested only up to four. The data in Ref. [373] suggest that the discrimination by starlings was not perfect (although somewhat better than chance). The birds might not have been using counting, but their matching might have been based on the judgement of duration, tempo or rhythm [382, 383]. Starlings were previously taught to discriminate between two completely artificial sounds looped in a cycle (XXXXOOOO and OXOXOOXX; and later, the sounds were either reversed or replaced by new sounds) regardless of the starting point in the cycle [384]. This supports the possibility that the birds discriminated the strings without knowledge of any grammar used to create them. Further research will be necessary to determine how the songbirds do this and whether this ability truly involves True Recursion and whether the mechanisms they use are similar to (and how similar) or different from human faculties. It is most probable that animal capacity is restricted to parsing symbols in short temporal proximity from each other. Starlings may be using iterative Tail Recursion to arrive at their achievements of matching quantities.

Moreover, it must be noted that analyses (but not theoretical underpinnings) in both experiments are based on formalistic analyses that seem to be following the paradigm of computational linguistics, emphasising the mathematical underpinnings of linguistic computation. The computational linguistics paradigm, however, deals only with superficial sequences of strings, not the structural analyses and interpretations that modern syntax models deploy [385]. Grammars used in these experiments addressed neither recursion nor centre embedding.

The studies on tamarins by Fitch and Hauser [185] and on starlings by Gentner et al. [373] are not sufficient inquiries into the presence of recursive syntax in non-human animals, despite the claims of their authors. Gentner et al. [373] do not necessarily demonstrate the understanding of the recursivity in simple centre-embedded sequences.

3. Conclusions and Discussion

No animal has likely adopted the open, unbounded, hierarchically recursive system that allows humans to, quite literally, express anything (Chomsky, Pinker and others). The animals certainly did achieve something. The fundamental questions are the following: What was it? And how did they achieve it?

Hauser and Fitch suggest that the pseudo-linguistic feats of trained animals could very well be due to either homologies of the relevant organic neural structures found in humans or convergent evolution of analogous domain-specific systems.

The solution, which will be presented in a follow-up paper, differs substantially from the mosaic solution that Hauser and Fitch put forward in their papers. The follow-up paper will suggest that the idea that animals used alternative strategies to achieve their “pseudo-linguistic feats”, possibly relying heavily on semantic processing bound to General Cognition, might be seriously entertained.
Acknowledgements

I would like to thank Anton Markoš for reviewing the thesis on which this text is based and for providing advice, suggestions and guidance. This paper is dedicated to my ex-colleague, collaborator and teacher, Marco Stella. This paper is a shortened version of my master's thesis, with inclusions of parts from my unpublished popular book, named “the Logonauts”, on the same topic.

References


Talking Animals and Human Language


Savage-Rumbaugh, S.E., Pate, J.L., Lawson, J., et al., 1983. Can a chimpanzee make a


[177] Čadková., 2015


[179] Lestel, D., 2011. What capabilities for the animal?


[188] Harris., 1995


Science, 3(6), 555–563.


