



HAL
open science

Comparing Human and Nonhuman Animal Performance on Domain- General Functions: Towards a Multiple Bottleneck Scenario of Language Evolution

Joël R Fagot, Raphaëlle Malassis, Tiphaine Medam, Marie Montant

► To cite this version:

Joël R Fagot, Raphaëlle Malassis, Tiphaine Medam, Marie Montant. Comparing Human and Nonhuman Animal Performance on Domain- General Functions: Towards a Multiple Bottleneck Scenario of Language Evolution. Boe, L.J., Fagot, J. F., Perrier, P. & Schwartz, J.L. Origins of Human Language: Continuities and Discontinuities with Nonhuman Primates, Peter Lang D, 2017, 978-3-631-73808-5. 10.3726/b12405 . hal-02435853

HAL Id: hal-02435853

<https://hal.science/hal-02435853v1>

Submitted on 16 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Joël Fagot, Raphaëlle Malassis,
Tiphaine Medam and Marie Montant
*Laboratory of Cognitive Psychology, Institute of Language,
Communication and the Brain, Aix-Marseille University, and CNRS*

Comparing Human and Nonhuman Animal Performance on Domain-General Functions: Towards a Multiple Bottleneck Scenario of Language Evolution

Abstract: How come that *Homo sapiens* is the sole animal species communicating with a language (i.e., a human language)? Theorists of language evolution have mostly adopted a human-centered approach to address this question. This chapter discusses the limits of this approach and proposes an alternative that consists in studying the domain general functions that serve language comprehension and production from a comparative and evolutionary perspective. Special attention is given to domain general processes which allow humans and animals to integrate information in space and time, and thus develop perceptual and more conceptual abstract categories. This chapter presents illustrative studies that reveal the various aspects in which these integration processes differ in human and nonhuman animals. Finally, we discuss the source of these species differences and their potential implications for our understanding of language evolution.

Keywords: Animal communication, integration processes, categorization, language evolution, baboons

1. Introduction

1.1 Limits of strictly human-centered approaches

In the various scientific disciplines interested in language evolution – anatomy, physiology, paleoanthropology, linguistics, psychology, neurosciences, computer sciences – the precisely defined architecture of human language is classically used as the basic reference, the *canon*, to which all nonhuman communication systems are compared. Said differently, the communication system of *Homo sapiens* -an approximately 200,000 year-old isolated species (there are no other living species in the genus *Homo*) with almost

1.5kg brain tissue, living in complex societies with a long history of cultural evolution—is used as the unique reference system in most comparative studies on language evolution (for a review, Hauser et al., 2002). Obviously, this makes sense given that the purpose of these comparative studies was not to build a descriptive catalogue of other animal communication systems but to better understand human language evolution. However, when pushed too far, such an anthropomorphic approach can be misleading because, by definition, nonhuman animals cannot equal human animals in performance when it comes to human language. A current caveat in human-centered comparisons consists in assuming non-explicitly that nonhuman cognitive architectures must *resemble* human cognitive architecture, in parts or as a whole. However, such an assumption can hold only if: 1) human and nonhuman cognitive architectures had followed similar evolutionary paths and were adapted to comparable environmental, social and biological constraints, and 2) if the cognitive architecture of each species was a construction made of independent (non-interacting) cognitive components that are not sensitive to developmental and phylogenetic factors. Given that every species has a unique cognitive architecture, it seems like a vain enterprise to search for strictly identical components in humans and nonhuman animals.

Consider one example: syntax. Different types of syntaxes have been formalized in Chomsky's hierarchy, from simple (finite state) grammars to complex (supra-regular) grammars. In an attempt to better understand human cognitive "uniqueness", major efforts have been put in the investigation of nonhuman species' ability to process supra-regular grammars (e.g., Fitch and Hauser, 2004; Gentner et al, 2006; Abe and Watanabe, 2011). Unsurprisingly, nonhuman animals do not equal *Homo sapiens* in that particular type of linguistic computation. The very attempt to search for strict human-like syntax in nonhuman animals implies that there is only one way to compute information in a complex communication system: the human way. This view does not consider the possibility that each species, even phylogenetically close to *Homo sapiens*, might have developed its own original and complex –possibly multimodal- cognitive architecture, which does not include human-like syntax. It also neglects the existence of interactions between the various components of a cognitive architecture whose effects increase over phylogenetic and developmental time scales (syntax might not exist as an independent computational subsystem, not even in

humans, see Seidenberg and MacDonald, 1999). These interactions make it very unlikely that complex integrated levels of computation (syntax-like processes) in the cognitive architectures of two different species resemble each other.

We must acknowledge, however, that more and more studies are looking for simpler forms of syntax in nonhuman animals. This approach has recently gained interest in the field of ethology and to some extent in linguistics (for a review, Schlenker et al., 2016; see also Petkov and Wilson, 2012). It refers to basic learning principles that allow the extraction of combinatorial semantics, statistical regularities, or adjacent and nonadjacent dependencies according to which qualitatively comparable basic learning principles are supposed to hold for human as well as nonhuman communication sequence learning (e.g., Seidenberg et al., 2002).

The narrow anthropomorphic approach of language evolution that we question here is not only found in linguistics but has dominated other domains as well, such as comparative anatomy. Rapid progress of imaging techniques over the past 20 years has spurred a frantic search for the anatomical landmark of language in *Homo sapiens*. After a couple of unsuccessful attempts to teach human vocal language to nonhuman primates (e.g., Hayes, 1951), it has become clear that our closest cousins – chimpanzees – are not able to pronounce human phonemes. The first proposed explanation for the incapacity to produce speech was that their larynx was too high (e.g., Lieberman, 1968, 1975). The existence of a physical limitation for speech in chimps has been admitted in the research community for many years. However, recent studies suggest that vocal tract anatomy cannot suffice to explain the absence of speech in nonhuman mammals (for a review, Fitch 2010). Fitch et al. (2016) inferred from X-ray videos and a modeling approach that macaca's tract has the potential to produce a broad range of speech sounds. This was confirmed in one of our recent study in which we recorded the spontaneous vocalization of baboons (*Papio papio*, Boë et al., 2017). We found that the baboons produce sounds that share the acoustic F1/F2 formant properties of human [i æ a o u] vowels, and those baboon sounds were produced by movements of the tongue in a human-like articulatory space defined by two axes (anterior-posterior and superior-inferior, see Boë et al., 2017 for more details). Therefore, the inability of nonhuman primates to produce human phonemes is more likely

due differences in the neural circuits that command oro-facial muscles, or to other cognitive differences, than to differences in the anatomy of the vocal tracts.

Other anatomical factors have been proposed to explain the uniqueness of language in the human species. Most of them concern human-centered brain features. Neuroscientists have been looking for the homologue structures of Broca's area and Wernicke's area in the brain of various apes. Broca and Wernicke are important regions of the human brain whose injury provokes severe language disorders, the so-called Broca's and Wernicke's aphasias. Because of interspecific anatomical differences, direct comparisons of brain regions across species are difficult to make. Instead, researchers have compared brain asymmetries across species. In the human brain, the language function is mostly hosted in the left hemisphere. Broca's area, Wernicke's area and the *planum temporale* are bigger in the left than in the right hemisphere (Geschwind and Levitsky, 1968). Those asymmetries, that were first thought to be unique to humans, have been observed as well in nonhuman primates (e.g., Cantalupo and Hopkins, 2001; Cantalupo et al., 2003). In a recent paper, Gomez-Robles and collaborators (2013) propose that, at a whole brain scale, there is continuity in asymmetric variation between humans and chimpanzees: similar brain asymmetries exist in both species, even though the human brain tends to be a little more asymmetric and more sensitive to developmental constraints.

More recently, tremendous efforts have been put into the identification of precise brain regions or neural structures that might be unique to humans and explain the emergence of language in our species. For example, Leroy et al. (2014) have proposed that the superior temporal sulcus (STS) critically differs in human and nonhuman species: human STS is deeper in the right hemisphere compared to the left, and this depth asymmetry is not found in chimps. Given that STS is central in the perisylvian language region, this particular landmark could be a promising "human-only" candidate. However, as the authors acknowledge, the link between this anatomical feature and the language ability in humans remains loose: MRI anatomical measurements made on various groups of human subjects show that the STS depth asymmetry is bigger in men than in women, it persists in children with impaired language development, and is unchanged in adults with reversed language lateralization (*situs inversus*). To drive the point home, an even

more recent study shows that the very same STS brain asymmetry exists in baboons (Meguerditchian et al., 2016). Further investigations on the relationship between variation in human brain anatomy (at a brain network scale) and variation in human language function are needed before one can make between-species comparisons and draw convincing conclusions about the role of precise human brain features in the evolution of human language.

The literature on human brain landmarks for language also raises a crucial question: what can gross anatomy tell us about fine-grained and complex interacting cognitive functions, in particular in a comparative perspective? Without a well-defined theoretical model of the anatomy-to-function relationship in each species that would make the comparison possible, the explanatory power of anatomy brain differences remains very limited. This would not be the case if we knew precisely the functional significance of those particular brain regions in their neural networks, in both species. For example, recent studies have shown that the homologue of Broca's area in chimps is involved in communicative behaviors (Tagliabue et al., 2008; 2011), however little is known about the type of computation it makes, or about the way it deals with a combination of communicative gestures and vocalizations, or about the way this region interacts with deeper structures involved in emotional vocalizations (Jürgens, 1979). Just like the linguistic/cognitive human-centered components, the anatomical landmarks of language -when strictly human-inspired (and functionally underspecified)- seem of very little explanatory value to understand language evolution in the human lineage, at least in the current state of our scientific knowledge.

The search for a unique (cognitive or anatomical) key factor at the origin of human language will inevitably lead to a stalemate. Brains, just like the cognitive function they host, are shaped by species-specific phylogenetical and developmental trajectories. They are complex systems and the differences between human and nonhuman brains and functions cannot be easily reduced to single "keys" components, especially when a strict human-centered definition of these components is applied.

1.2 An alternative ...

What could be the alternative to this narrow human-centered approach? In this chapter, we propose to step backward to enlarge our view of language evolution. Rather than focusing on the presence or absence of strictly human-inspired language features in nonhuman animals, we believe that an alternative is to examine the background of the language function, namely the inherited domain-general elements of “the machinery required to master human language” (Saffran and Thiessen, 2008), that might be used in non-human species for communicative and/or other purposes. Domain-general mechanisms correspond to learning devices that apply to a variety of different cognitive functions, as opposed to domain-specific mechanisms that are dedicated to specific cognitive functions. The opposition between domain-general *versus* domain-specific mechanisms was first proposed in a cognitive development framework where Skinner’s view (1957) was opposed to that of Chomsky (1959). In the present chapter, we use the notion of domain-general mechanism in a more evolutionary perspective. This idea we defend here is that the human language-device is mainly made of domain-general elements, some of which are shared with other species (either because they were present in a common ancestor or they result from convergence processes). These inherited cognitive components are very likely to take different forms in different species, as a function of the cognitive domain they are involved in, because different domains show different regularities and constraints that shape these components. However, we expect that ~~the~~ close comparison between human and nonhuman performance will reveal what aspects of those domain-general cognitive components are shared across species. The underlying hypothesis we uphold here is that complex and phylogenetically recent cognitive functions, ~~including~~ language, are probably the product of intense re-use and re-combination of subsets of inherited anatomical, cognitive, behavioral components (Anderson 2010). Phylogenetically close species might share some (but not all) of these components, as a support of communication and/or other cognitive functions. For example, the serial organization and structuration of elements that we find in the processing of syntax might as well serve the planning of complex motor sequences in humans (Koechlin and Jubault, 2006), in other primates or in birds, including the sequences of bird’s songs (Suzuki et al., 2016).

In the remaining part of this chapter, we will firstly describe some of the multilevel integrative processes that are critically involved in human language, which likely evolved from domain-general inherited functions. We will then consider and discuss main findings of the literature on comparative cognition regarding these domain general mechanisms. In this context, we will pay special attention to the mechanisms by which animals integrate stimulus information at various levels, from low level perceptual grouping mechanisms that lead to global percepts, to more cognitively complex processing mechanisms that make it possible to associate meanings to objects or categories for example. Importantly, this chapter is not aimed at making an exhaustive review of the literature regarding these cognitive processes. Our goal is to document potential *species* differences and similarities concerning these processes, and to illustrate these differences and similarities by a selection of suggestive findings. In a final section of this paper, we will discuss the potential impact of these results on our understanding of language evolution.

2. Integration processes in nonhuman animals

2.1 Integration in time and space

Language production and comprehension involve the processing of a continuous flow of information, and a temporal integration of the different linguistic elements. There is an obvious connection between sequential learning and language, because these two cognitive processes require the extraction and further handling of elements occurring in temporal sequences. However, because most behaviors have a temporal structure, the capacity to relate events in the time dimension might derive from a domain-general function that is involved in many different contexts, beyond language. We will consider below four different aspects of sequence processing, which are the ability to (1) remember sequence of events, and to process that sequence as temporally ordered, (2) to learn the transitional probabilities of items occurring in sequence, (3) to learn and process nonadjacent dependencies, and therefore to know that event A is followed by event B with intervening events between them, and (4) to gain information regarding the general structure of the sequence (i.e., syntax as concerns human language). For all these problems, we document below which aspects of sequence processing seem shared by animals and humans, and which aspects seem more restricted to humans.

2.1.1 *Serial list learning*

List learning corresponds to the capacity to ~~both~~ remember items in a list, ~~but also~~ to remember their temporal ordering. Nonhuman animals can learn and remember lists of items, (e.g. Terrace et al., 2003). Comparative investigations of serial list learning in animals, especially pigeons and monkeys, suggest that the memory of lists of items is not qualitatively different from what is found in humans. When taught that $A < B$, $B < C$, $C < D$, $D < E$, various nonhuman species have been able to properly order previously un-trained test pairs such as $C < E$. Pigeons and monkeys show serial position effects (Wright et al., 1985) sharing similarities with those of humans. Monkeys also show the symbolic distance effects found in humans: their latency to compare two items is shorter when the items are far apart in the list (e.g., AC vs CD ; Colombo and Frost, 2001). All these discoveries show that animals are capable of developing representations of series of items based on their ordinal position. Although the succession of words in a sentence is not strictly equivalent to a succession of unrelated items, serial list learning capacity is probably necessary in both cases. These results encourage the view that the underlying features of the learned representations are shared by human and nonhuman animals, suggesting quite old evolutionary mechanisms for item list learning.

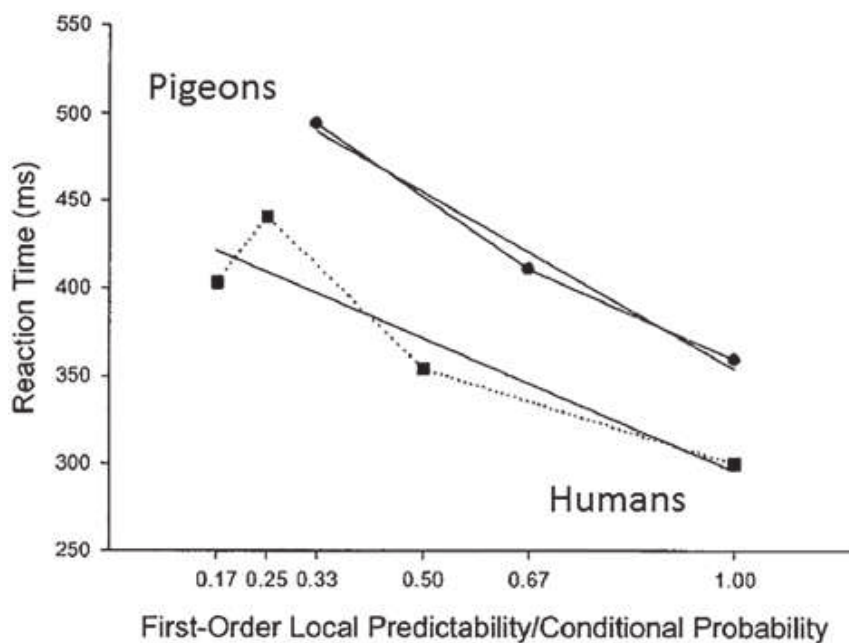
2.1.2 *Processing of adjacent dependencies and chunking*

Adjacent dependencies refer to a predictive relationship between one event and the event immediately following it in the sequence. Consider two three-item sequences, the sequences $A-B-C$ and $A-B-D$, which are presented an equal number of times. In this very small corpus composed of two sequences, A is always followed by B and B is followed half time by C and half time by D . The transitional probability between A and B is thus equal to 1, while the transitional probability between either B and C or B and D is equal to .5. Consideration of the transitional probabilities is one of the mechanisms promoting the learning of auditory and visual sequences in humans (e.g., Hunt and Aslin, 2001). It is one of the mechanisms by which children learn word boundaries and segment speech streams into words: in a sentence, the transitional probability between the last syllable of a word

and the first syllable of the following word is lower than the probability between two successive syllables within one word (Saffran et al., 1996).

Comparative studies have shown that a high transitional probability between two AB items facilitates the processing of the second B item in pigeons and monkeys. Froehlich et al. (2004), for instance, tested pigeons in a serial response time task requiring to peck a stimulus appearing sequentially at three possible locations and in a predefined order. The transitional probabilities between the stimulus locations were controlled in this task, and the authors report that response time to peck is a direct function of these probabilities. Thus, high transitional probabilities gave rise to short response times at the second location of the considered pair, while low transitional probabilities gave rise to longer response times. These authors directly compared their results on pigeons to Hunt and Aslin's study conducted on humans (2001), and report that although slower, pigeons processed information at roughly the same rate as humans, as reflected in similar overall regression slopes (Figure 1).

Figure 1: Average response times of pigeons and humans depending on transitional predictability of the items in a sequence. Results on pigeons are from Froehlich et al. (2004), those on humans are from Hunt and Aslin (2001). Also shown are linear regression lines for the two sets of data. Pigeons responded slower but processed information at roughly the same rate as reflected in similar overall slopes. Figure adapted from Froehlich et al. (2004).



In humans, the processing of transitional probabilities between adjacent elements comes with another type of processing known as chunking that consists in clustering information within sequences (e.g., the syllables in a word). Chunking in humans occurs during the memorization of verbal items, but chunking is not restricted to verbal material, and can be found as well in the visual domain (e.g., Orbán et al., 2008). Experimental evidence suggests that the capacity to organize sequences in chunks is also in the scope of numerous animal species (rats: Fountain, 1990; pigeons: Terrace, 1991; tamarins: Hauser et al., 2001; baboons: Minier et al., 2016). In a recent experiment involving a serial response time task, ~~it was~~ found for example that baboons organized 9-items sequences in three chunks of three items each, and that these chunks precisely included the items sharing the highest transitional probabilities (~~Minier, Rey and Fagot, 2016~~). Therefore, it seems that human and nonhuman animals are prone to statistical learning making use of transitional probabilities to both segment streams of (visual or auditory) information and organize the elements composing these streams into chunks.

2.1.3 *Nonadjacent dependencies*

Equally important for language is the capacity to detect and learn nonadjacent dependencies. Consider the following sequence structure A-X-B: A is followed by a variable item X, and item X is systematically followed by B. Given this structure, there is a nonadjacent transitional probability of 1 between A and B. Learning nonadjacent dependencies is important for language. For instance, a listener has to detect the relation between the subject and the verb in a sentence despite the presence of intervening words such as adverbs. This capacity can also be useful in very different nonlinguistic contexts, for example when we have to detect a systematic relation between two events separated in time (e.g., the ring of the doorbell signaling that someone is coming), and irrespective of the intervening other events. Experiments have shown that humans and animals can both process nonadjacent dependencies in temporal sequences of events, although nonadjacent probabilities are more difficult to detect and learn than adjacent dependencies (humans: Newport and Aslin, 2004; tamarins: Newport et al., 2004; rats: Fountain and Benson, 2006). Moreover, there are similarities

between species regarding the factors that affect the learning of nonadjacent dependencies (e.g., facilitatory effect of perceptual similarity of the nonadjacent elements in both humans and monkeys; humans: Creel et al., 2004; Gebhart et al., 2009; squirrel monkeys: Ravignani et al., 2013).

However, animals may have more difficulties when presented with complex sequences. Wilson et al. (2015) tested ~~two~~ monkeys and 33 humans using an auditory artificial grammar containing both adjacent and nonadjacent (long-distance) relationships. After an initial exposure to the sequences, the subjects from the two species were exposed to sequences containing violations of either the adjacent or both adjacent and nonadjacent relationships. Both species showed sensitivity to adjacent transitions, but only humans, and even roughly half of them, indicated significant sensitivity to nonadjacent dependencies. Wilson et al. (2015) concluded that in some conditions, nonadjacent probabilities are less salient in macaques than in humans. Although replications and extensions are required, this study suggests that, compared to monkeys, humans have a greater facility to deal with several dependencies of different types (i.e., both adjacent and nonadjacent) at the same time.

2.1.4 *Learning of sequence structure*

Learning the structure of a sequence requires the extraction of the relationships between the constitutive elements of that sequence. This kind of learning probably supports, among others, the encoding of grammatical and syntactic linguistic regularities (e.g., in most German sentences, the verb is at the end). However, such structural regularities also exist in many other (non-linguistic) domains, such as the motor domain. For instance, Byrne et al. (2001) reported that the preparation of food items requiring complex manipulations (thistle leaves) in wild gorillas follows a hierarchical sequential organization.

Marcus et al. (1999) have shown that ~~very young~~ 7-month-old infants can quickly learn that sequences of auditory stimuli follow an ABB or an ABA structure. The processing of such structures was also studied in nonhuman animals, but the results were quite inconclusive: zebra finches (Heijningen et al., 2012), rats (Toro and Trobalón, 2005) and even rhesus macaques (Procyk et al., 2000) do not seem to catch the difference between

the ABA vs ABB structures. One study compared zebra finches and humans using the same experimental procedure and stimuli (Chen et al., 2015): zebra finches did not learn these two types of structures, while humans learned them readily. There is to our knowledge only one study in which an animal species could successfully learn sequence patterns of the ABA/ABB type. This study from Spierings and Ten Cate (2016) compared two avian species, budgerigars and zebra finches, and obtained positive results in the former. From this set of experiments, we can conclude that learning the structure of sequences might very well be in the scope of some nonhuman species, but this ability is clearly not as developed as it is in humans.

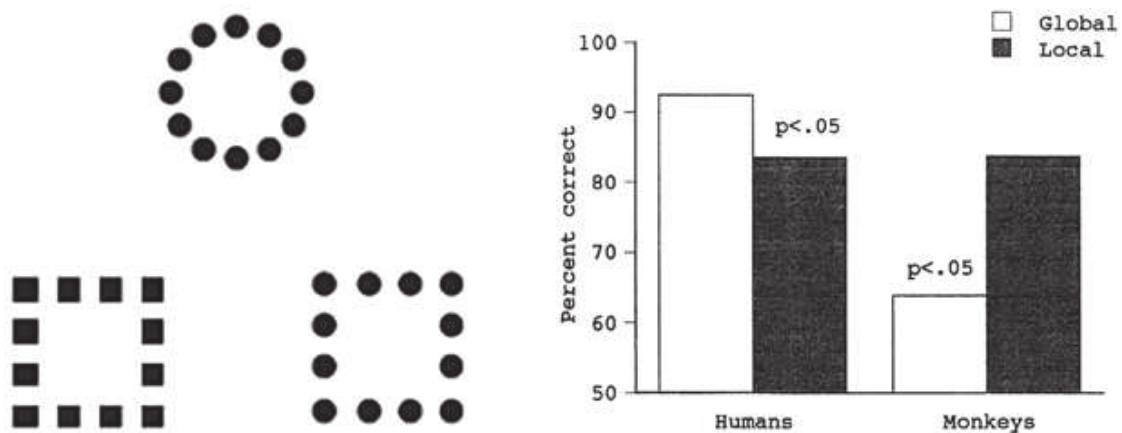
2.1.4 Integration in space

We documented above that human language function makes great use of temporal information, but spatial information is also crucially important for language. Infants at very young age learn that words and communicative gestures refer to entities in their immediate surrounding space. Words also make it possible to refer to objects that are spatially (or temporally) absent (the displacement feature proposed by the linguist Charles Hockett in 1960 to characterize human language). Finally, nonverbal forms of language, such as sign language or writing gestures strongly rely on an encoding of spatial information. No doubt that being capable to integrate and combine information in the spatial domain is also a domain general function important for a complex use of language.

In the late seventies, Navon (1977) has shown that human subjects tend to process the global shapes of visual object before they process their constitutive (local) details. This effect has been named the “global precedence effect” and is often considered to be an attentional phenomenon. Global precedence in humans was demonstrated in experimental research using large letters (global shape) made of smaller letters (local features). The degree to which animals perceive the global properties of the visual input in comparison to more featural ones has been an issue in animal cognition for some time. In our laboratory, we explored this effect in baboons (Deruelle and Fagot, 1998; Fagot and Deruelle 1997) using large shapes (square, circle, cross) made of smaller shapes (again square, circle, and cross) as stimuli. In our tests, the baboons were re-

quired to either match (Fagot and Deruelle, 1997) or identify (Deruelle and Fagot, 1998) these stimuli considering either their global or local structure. For comparative purposes, humans were also tested in the same experimental conditions.

Figure 2: Processing of the global / local stimulus structure in humans and baboons. Left: Illustration of the stimuli used with humans and baboons in Fagot and Deruelle (1997). This experiment required to match hierarchical stimuli considering their global or local structure (in this local trial, they have to match the circle made of circles, with the square made of circles, considering their common local features). Right: percentage correct obtained in humans and baboons in global and local trials. Humans showed an advantage to process the global structure of the stimuli, while baboons showed a local advantage. This local advantage in baboons is accounted for by a general difficulty to “group” the local elements of the hierarchical global/local stimuli into a global whole. Humans are much less sensitive than monkeys to the spatial distance separating the local elements (e.g., Deruelle and Fagot, 1998). Figure adapted from Fagot and Deruelle (1997).



These experiments revealed a striking species differences in global–local processing (Figure 2). Human participants exhibited the global advantage already found by Navon (1977), whereas baboons demonstrated their best performance and fastest response times in the local condition. Several experiments were conducted to understand the cause of this human-baboon difference, which suggested that the performance of the baboons strongly depended on the distance separating the local elements (Deruelle and Fagot, 1998): when the distance was enlarged, the strength of the local bias increased and this effect was amplified in baboons compared to humans. This

effect has been replicated many times in pigeons, capuchin monkeys and chimpanzees (for reviews, see Fagot and Barbet, 2006; Fagot and Parron, 2012).

At first glance, one might consider local precedence in animals as a purely perceptual/attentional phenomenon, but we propose that it is more than this. One of the main properties of language is that it can convey information about things that are not immediately present (spatially or temporally; Hockett, 1960). This displacement feature is crucial in the comparison between human language and other forms of primate communication. We have previously proposed (Fagot and Barbet, 2006) that a strong local bias limits the processing of the relation between and among objects. This effect is for instance demonstrated in Fagot and Parron (2010), showing that an increase in the separation of the distance between two bars of either identical or different colors limits the classification of this stimulus on a same/different relational basis. We suggest that a strong bias in favor of a local processing mode (in either the spatial or temporal domain) in nonhuman animals might place important constraints on their communicative systems. It reduces the possibility to make non (temporally or spatially) adjacent relations between or among the communicative signals, and between or among the communicative signals and the objects in the real world to which they refer, especially when they are far or absent.

2.2 Integration of stimulus dimensions and sensory modalities.

Animals, including humans, live in a rich world of information, and dealing with this complexity is probably critical for the survival of every species. The processing of this complexity may be achieved by a variety of cognitive mechanisms, which include, among others, the integration of the different perceptual dimensions into single entities (i.e., integration of stimulus dimensions), and the grouping of various exemplars of a given object into categories (e.g., object categorization; see a discussion of this issue below). Interestingly, these functions, which are of general adaptive value, are all critical for a multidimensional/multimodal system, such as language. For instance, speech comprehension requires that multiple prosodic (e.g., intonation, stress) and phonemic (e.g., voice onset time, place of articulation) dimensions present in the acoustic signal are processed and integrated.

Language comprehension is also achieved using a multimodal/multisensory mode of processing when various auditory and visual information are simultaneously integrated (e.g., for the mapping of the lip movements with the auditory signals).

2.2.1 *Combining multiple stimulus dimensions*

Discrimination tasks are often proposed to animals as a mean to assess their perceptual abilities. Typically, stimuli are presented to the subjects, and the subjects' behavioral responses to some perceptual dimensions or combinations of dimensions are reinforced, while responses to the non-relevant dimensions are not. Evidence suggests that animals can discriminate stimuli along various dimensions: color, shape, luminance, orientation of the visual objects, or the pitch of auditory stimuli. Animals can as well base their behavioral responses on combination of two or more stimulus dimensions. For instance, Cook (2001) showed that pigeons can learn to select the computer screen area where horizontal green lines are presented, while avoiding the screen areas showing non-green lines and green lines in a non-horizontal orientation.

In their review article, Lea and Wills (2008) comment on three main trends emerging from the literature on learned discrimination in nonhuman animals. The first one is that unidimensional discrimination is easier to learn than multidimensional discrimination based on combinations of features or conjunctions. Smith et al. (2012) have for instance trained monkeys to sort sine wave gratings depending on their orientations or both the orientation and spatial frequency considered in conjunction. Learning was much faster in the unidimensional than in the bi-dimensional test condition. Another example of this effect comes from research on conceptual discrimination by monkeys (e.g., D'Amato and van Sant, 1988) and the demonstration that discrimination performance relies strongly on an analysis of features, such as color, rather than on configurations of features. Lea and Wills' (2008) second conclusion is that when the stimuli are made of multiple relevant dimensions, nonhuman animals express a tendency to focus their attention on one dimension only, mostly when this dimension has sufficient discriminative values. There are also multiple examples of this trend in the literature. In our laboratory, we found that baboons discriminated

computerized human faces considering exclusively the contour of the face (Martin-Malivel and Fagot, 2001) or pixel luminance information (Martin-Malivel et al., 2006), instead of the multiple levels of information (e.g., configural information, identity, etc.) that the facial stimuli may provide, and that humans process. Finally, Lea and Wills' (2008) report than even with ingenious experimental designs, attempts to force nonhuman animals to process multiple aspects of the stimuli mostly lead to failures. This can be nicely illustrated by Dépy et al. (1997). The baboons in this study were initially trained to discriminate between two categories of stimuli defined by the possession of any combination of two out of three possible binary features. Baboons could sort these two classes of stimuli to a good accuracy level, albeit after a long training process of several thousands of trials, but remained unable to take the three discriminative features into consideration to achieve this performance, two of the three features taking a leading role in the task.

Wang et al. (2015) recently recorded the brain activity (IRMf) of rhesus monkeys and humans in two test phases. In the first phase, subjects from the two species perceived passively sequences of four tones, the last one being either of a lower or higher pitch than the first three. After this habituation procedure, the same subjects perceived sequences violating the general structure used during the habituation phase. Thus, some test sequences contained a number of tones different from the habituation sequences (number deviant), some other sequences contained four tone units with identical pitch (sequence deviant), and a last set of sequences differed from the habituation sequences regarding both the number of items and pitch. In both species, homologous brain areas were particularly responsive to violations in number (intraparietal and dorso premotor areas), and sequence (ventral prefrontal and basal ganglia), but humans were the only primates showing a joint sensitivity to both factors in the perisylvian language region (bilateral inferior frontal and superior temporal gyri). One limitation of this study is that an absence of evidence ~~in baboons~~ is not the evidence of an absence. Although this study does not address directly the relationship between brain and behavior, its results suggest that the perisylvian region is involved in humans only in the integration of various stimulus dimensions contained in auditory sequences of stimuli.

2.2.2 *Combining multiple sensory modalities*

The question of multimodal integration warrants a discussion in this section, due to its relevance to the origin of language. Evidence suggests that the ability to integrate information across sensory modalities is not at all restricted to humans. Cross-modal integration was demonstrated for instance using a task requiring the processing of multi-sensory stimuli (Lanz et al., 2013), cross-modal interference tasks in baboons (Martin-Malivel and Fagot, 2001), and cross-modal matching tasks in chimpanzees (e.g., Davenport and Rogers, 1970). Unfortunately, we are aware of only one study in which the performance of humans and nonhuman animals were directly compared in cross-modal tasks using the same stimulus material. This experiment from Fagot et al. (2000) requested the subjects to categorize pictures of humans and baboons in one condition, and human and baboon vocalizations in another one. In this experiment, the subjects of the two species perceived a prime prior to the presentation of the stimulus to be categorized. Depending on the condition, the prime could be a picture or a vocalization of baboons or humans and three conditions were tested: intra-modal visual-visual priming, intermodal auditory-visual priming, and intramodal visual-auditory priming. Three subjects out of four in each species demonstrated intra-modal priming. Inter-modal priming was demonstrated in the three out of four human subjects in the auditory-visual condition, and all four in the visual-auditory condition, but it was only found in one baboon out of four in each intermodal condition, suggesting that inter-modal integration is more difficult in baboons than in humans. Given the small number of subjects involved in this study, a replication is warranted before drawing any firm conclusion on the evolution of intra-modal integration.

2.3 **Categorization and conceptual integration**

A central aspect of human cognition is our ability to form categories of various kinds of objects or mental entities. Categorization implies that the exemplars of each category are grouped into classes considering physical or more abstract properties. Categorization is a domain-general ability that is fundamental for a variety of more specialized functions (e.g., inference or decision making), including the language function considered in this book.

Historically, Herrnstein and Loveland (1964) were the first to demonstrate categorical abilities in nonhuman animals. They showed that pigeons could sort pictures of humans and pictures devoid of humans in two open-ended categories. Since this study, numerous papers have confirmed that various animal species can efficiently categorize stimuli considering low level perceptual stimulus dimensions, such as the pitch for auditory stimuli, and the color, shape, size, motion, orientation, luminance for visual stimuli (see for instance Berg and Grace, 2011, where pigeons were trained to categorize sine-wave disks considering their spatial frequency and orientation). However, the ability of nonhuman animals to apply categorical processes to more abstract – human-like – stimulus dimensions remains a matter of debate.

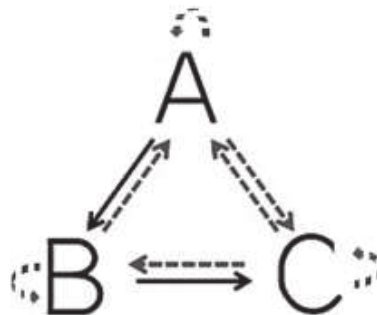
2.3.1 Equivalence classes: grouping arbitrary items within the same category

An important aspect of language is its arbitrariness. Arbitrariness corresponds to the fact that nothing in the physical form (acoustic properties) of most words refers to the objects they designate. For example, the word “car” does not “sound” like the vehicle it refers to. Therefore, words can refer to things in the real world, and things can refer to words, although there is no natural or necessary connection between them. In that case, words and objects are linked by a relation of equivalence, and the many exemplars of a given category of object (e.g., many different tables that vary in shape and color) can be categorized under a unique word label. The ability to form arbitrary connections between words and objects during ontogeny probably comes from the many co-occurrences of words-objects pairs that infants encounter during development. We presume that this ability relies on a more domain-general capacity to make associative (arbitrary) connections between items, or categories of items.

Experiments on “stimulus equivalence” have directly addressed the capacity of human and nonhuman species to group arbitrary items into categories, on the basis of their associative history (Sidman and Tailby, 1982). The prototypical design of stimulus equivalence experiments is shown in Figure 3. In stimulus equivalence experiments, the subject first learns a network of associations (shown in black in Figure 3) with repeated expo-

asures to these associations. Then, probe trials (shown by the red arrows in Figure 3) test if the subject (1) can associate each stimulus to itself (e.g., associate A to A, reflexivity relation), (2) can revert the trained relations (e.g., associate B to A, symmetry), and can associate the stimuli that have a common associate (associate A to C, transitivity). According to Sidman and Tailby (1982), stimulus equivalence is fully shown if the subject demonstrates, without further training, the relations of reflexivity, symmetry and transitivity in post training trials. One may easily imagine the serious limits of a cognitive system that would fail in this task. The formation of equivalence classes has the power to permit to use stimuli interchangeably, and is probably the corner stone of complex symbolic thought.

Figure 3: Typical paradigm for experiments on stimulus equivalence. In this experimental design, the black arrows illustrate the trained associations. The dotted arrows illustrate the untrained associations that emerge in humans after an initial training phase (e.g., Sidman and Tailby, 1982).



Equivalence relations emerge early in human infancy (23-month old, Lipkens et al., 1993). By contrast, the formation of equivalence classes seems especially difficult for nonhuman animals such as pigeons or monkeys, and among the three relations sustaining stimulus equivalence, the relation of symmetry seems the most difficult one to acquire. Lionello-DeNolf (2009) reviewed a total of 24 articles on symmetry testing, and found that the vast majority of these articles reported negative results. Moreover, the handful of articles reporting more positive findings all used a very small number of subjects who received special training procedures, such as long training combining symmetry and reflexivity trials (Frank and Wasserman, 2005) or forms of symbolic training (Pepperberg, 2006).

A recent paper from our research group nicely illustrates that difficulty (Medam et al., 2016). We trained baboons to associate pictures of bears and

pictures of cars to two different shapes which served as category labels. The baboons demonstrated category learning in this task, but failed to respond correctly when asked to reverse the trained relation and to associate the category labels to the pictures of cars and bears. The immediate conclusion from this study and others (see Lionello-DeNolf, 2009) is that the preferred mode of processing in nonhuman animals is not favorable to the spontaneous emergence of symmetrical relations, which is fundamental for the emergence of equivalences classes. Arbitrariness, as defined by linguists, requires arbitrary, bidirectional associations between the words and their referent (e.g., the written “CAT” refers to the cat animal, as the cat animal refers to the “CAT”). Evidence suggests that nonhuman animals too can form arbitrary associations between various items with no logical connections between them, but they apparently have great difficulties to process these associations as bi-directional.

2.3.2 *First order relations*

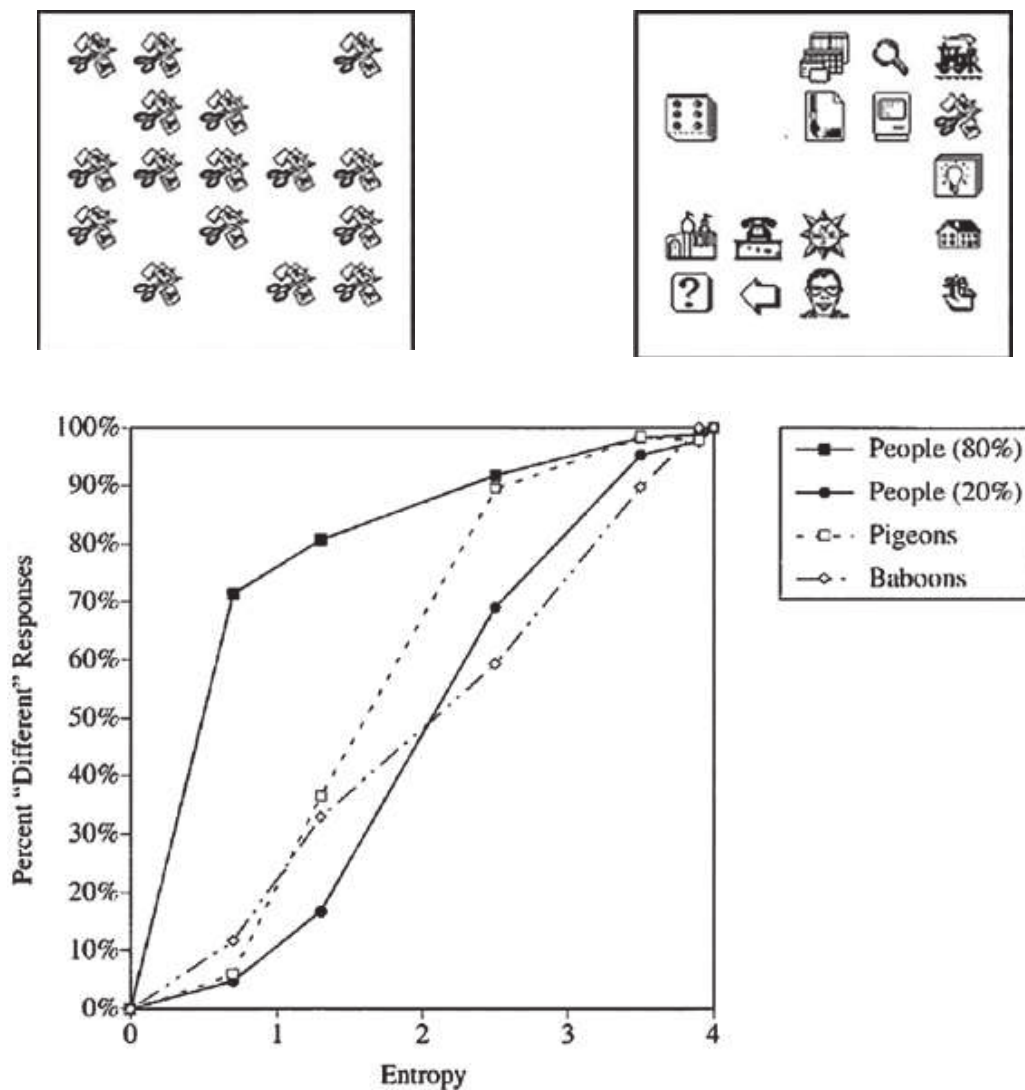
Nonhuman animals can master a broad range of discrimination tasks, and some of them involve the processing of first-order relations. First-order relations refer to spatial or more abstract relations among objects, such as the fact that an object is above or below another one, or that two objects have the same functions. Particularly important for our linguistic system is the first-order relation of sameness/differentness. Our language makes great use of categories, and the abstract concept of sameness is essential for the development of verbal categories. The concept of sameness can provide the basis for the most complex cognitive operations, such as the conservation of volumes or areas, or analogical reasoning (see below). Without this concept, we would be unable to understand sentences such as “This is a cat!”, and to get from this sentence the idea that the animal we see belongs to the cat category. We would as well be unable to understand sentences such as “It is warm again”, suggesting a similarity between the current and past weather. Comparative psychologists have shown that a variety of nonhuman animal species, such as the chimpanzee (Premack, 1983), the baboon (Wasserman et al., 2001) and the rat (Wasserman et al., 2012), succeed in learning same-different relational tasks. However, the nature of the mechanisms supporting this competence in nonhuman animals, and their similarity with humans, remain unclear.

Wright and Katz (2006) asked the following question: how much training nonhuman animals need to form the same/different concept? To answer this question, they tested pigeons, capuchin monkeys and rhesus monkeys with the same test design. Animals from the three species saw two pictures in succession on a touch screen, and two kinds of trials were distinguished. In the “same” trials, the second picture was identical to the first one, while it was different from the first one in the “different” trials. In both kinds of trials, a white key was always displayed on the right of the second picture. When the first and second pictures were identical (identical trials), then the subject was asked to touch/peck the second picture to obtain a reward. If the second picture was different from the first one (different trials), then a touch/peck at the white response key was considered correct. The monkeys needed much less items (about 32) to develop the concept of sameness, than did the pigeons (256). The number of trials children would need in this task is not known, but studies have shown that children can categorize cats as different from dogs with only 12 training exemplars (Quinn et al., 1993), and by 10 months of age, they can form categories with only 7 or 8 training exemplars (Younger and Cohen, 1986). The data therefore suggest an evolutionary trend in this ability: humans would require exposure to a smaller number of items than the other animals, to form categories and develop same/different concepts.

The concept of sameness can be applied to a broad range of attributes, from the most perceptual to the most abstract ones, and another interesting issue in the comparative literature is to know if animals use the same kind of information as humans, when solving similar tasks requiring an abstract concept of sameness/differentness. This kind of questions has been addressed extensively by Wasserman and collaborators (see review in Wasserman et al., in press; Wasserman et al., 2004). These authors trained pigeons, baboons and humans to categorize displays resembling those of the top of Figure 4. They consisted in arrays of 16 icons which were either all same (same relation) or all different (different relation). After they received this category training, the subjects were tested with arrays containing mixtures of icons, in which some icons were duplicated a number of times in the array. The authors reasoned that if the subjects have formed the concept of sameness, then they should classify the arrays containing at least one item different from the others as “different arrays”, irrespective of the fact that some icons are repeated. Figure 4 illustrates the most substantial findings of this set of experiments. This

figure indicates the percentage of different responses with the mixtures, as a function of the entropy of the stimulus. Entropy in this experiment should be understood as a quantification of the perceptual variability of the array: the all-same arrays have an entropy of 0, and the all different-arrays have the highest possible entropy value of 4 (mixtures have intermediate entropy values).

Figure 4: Use of perceptual (entropy) cues by pigeons, monkeys and humans, in a same/different discrimination task. The top panel shows the kind of displays employed in Wasserman and collaborators' experiments (Wasserman et al., 2004, left: same array, right: different array). The bottom figure shows that the same-different response of monkeys and pigeons is controlled in this task by the entropy of the arrays while 80% of the humans humans used more abstract cues in this task. Figure adapted from Wasserman et al. (2004).



The bottom part of Figure 4 shows that the behaviour of the baboons and the pigeons was controlled by the entropy of the arrays, which is a perceptual cue. Response from a subset of humans (about 20% of the group) was also largely controlled by the entropy of the arrays, but this constraint was released in most of the subjects (80%) who treated the arrays containing at least one item different from the others as illustrations of the “different” concept. It can therefore be concluded that humans expressed more abstract judgments than pigeons and baboons in this task. We will not present in this chapter the full series of experiments using this kind of stimuli with animals (for a recent review, see Wasserman et al., in press), but the reader should be aware that clear demonstrations also exist that monkeys can also base their same/different responses on abstract cues independently of the entropy of the stimuli (Flemmig et al., 2013). Nevertheless, although several animal species seem capable of abstract same/different judgements, humans, more readily than other animals, apply qualitative, rule-based frameworks on the Same-Different discrimination task.

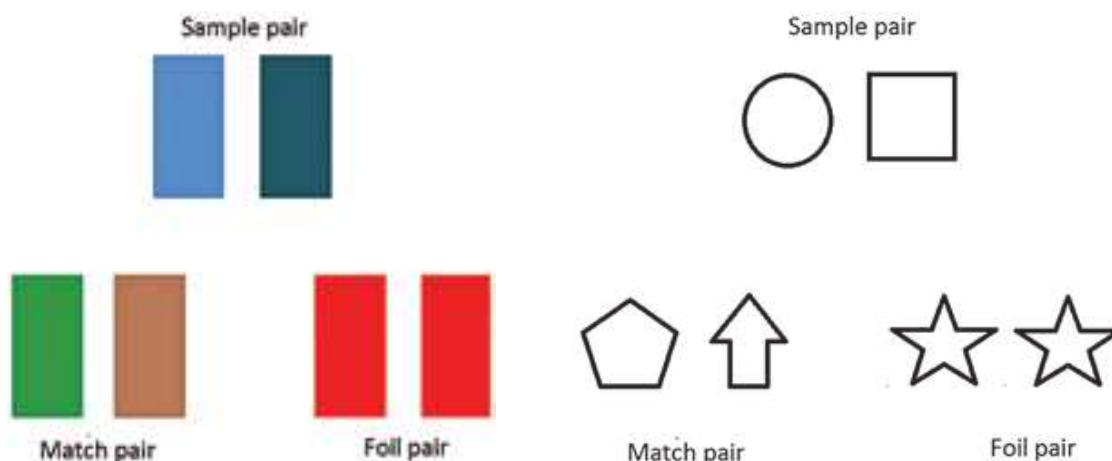
2.3.3 *Analogical (second-order) relational processing*

Our linguistic systems make great use of analogies, and our capacity to produce and understand analogies is considered by many as the “the Fuel and Fire of Thinking” (Hofstater and Sander, 2013). Developmental studies have shown that analogical reasoning is facilitated in children by the capacity to represent abstract relations in symbolic terms via linguistic labels (Christie and Gentner, 2014).

Most research on analogical reasoning in animals has used the Relational Matching-to-Sample task (RMTS: e.g., Fagot and Thompson, 2011) illustrated in Figure 5. In this task, the subject first perceives one pair of objects which are either identical or different. Two comparison pairs are then presented, and the subject must indicate the stimulus pair exemplifying the same (same or different) relation as the sample pair. In other words, the task can be conceptualized as “if AA then BB, and if AB then CD”. Researchers in the domain of comparative cognition have tested several animal species, including pigeons, monkeys and apes using the RMTS task (for a review, see Wasserman, Castro and Fagot, in press). Most of these attempts failed (Thompson and Oden, 2000), but a handful of studies also provide more

positive results, in particular in tests involving thousands of training trials (e.g., Fagot and Thompson, 2011).

Figure 5: Illustration of the relational matching task used, in baboons, with color (left part of the figure) and shape (right) stimuli.



In our laboratory, we could demonstrate that the baboons can successfully solve the RMTS task with pairs of color patches as stimuli (Fagot and Parron, 2010, see Figure 5). In a different study (Fagot and Thompson, 2011), we could further demonstrate that baboons can also solve this RMTS task considering the shape of the items (Figure 5). Again, this cognitive feat also required an extensive training period (from 17 to 30 000 trials per subject). In both studies, the different generalization tests confirmed the real abstract nature of the processes at work in these two tasks. For instance, the baboons could continue to solve the task with a high level of performance when we used novel colors (Fagot and Parron, 2010), and novel shapes (Fagot and Thompson, 2011) as stimuli. However, although cognitive flexibility is suggested by these findings, the data also suggest limits in this processing. In Fagot and Parron's (2010) study, color cues were in fact proposed to 6 baboons, and 4 of these 6 subjects eventually learned the task. In Fagot and Thompson's (2011) study, the same task was given one year later to a larger group of subjects, including the 6 already tested in Fagot and Parron (2010). Six out of 29 baboons learned the RMTS task with shapes, but importantly, none of the baboons who had initially learned the task with colors could also learn it with shapes. In other words, learning the RMTS task with color

cues did not help at all the subjects to learn. Generalization across domains is central to analogical reasoning in humans, and is probably what makes human reasoning so flexible. Generalization across domains allows us, for instance, to understand the meaning of a sentence like “atoms are like tiny solar systems” or “life is a gift, a chocolate box”. Baboons – and probably other nonhuman primates as well – are quite flexible to process items and their relations within given domains, those they have been trained with (e.g., color), but are clearly not as skilled as humans to generalize across domains. This, we believe, is another factor that may greatly affect nonhuman animal’s potential for developing elaborated forms of language.

3. Summary and Conclusions

The main goal of the chapter was to examine the origin of human language from the standpoint of comparative psychology. Language in its various forms (e.g., gesture, writing, speech) is a multi-level integrative process that requires, at the perceptual stage, the segmentation and grouping of perceptual information to extract the general meanings of the communicative signals. We have argued above that many of the integrative processes involved in the language function are in fact domain-general processes that can also be found in non-linguistic functions.

Considering that language uses a multitude of domain-general functions, examination of these functions in animals and especially in nonhuman primates should provide important information on the cognitive background that made it possible for language to emerge in our evolutionary history. Following this reasoning, we comparatively examined in this chapter a number of domain-general cognitive functions, which imply various forms of integration of perceptual/conceptual information. Among the considered functions, we examined the ability of animals to integrate information in time and space, to combine stimulus dimensions, to group objects into categories, and to develop conceptual/relational processes (first- and second-order concepts). Of course, this list is not exhaustive but we believe that it represents a significant selection of basic cognitive domain-general processes that serve language perception and comprehension.

The present overview of the literature allowed us to reach two main conclusions. First, we have identified clear-cut demonstrations that non-

human animals are capable of grouping information in time and space, can combine stimulus dimensions, and can form categories at different levels of abstractness. Thus, we suggest that these functions are shared by humans and nonhumans, at least to some extent, and that they are not language specific (i.e., they have a long phylogenetic history). Second, the literature also reveals important differences in performance among nonhuman animals, and between nonhuman animals and humans. For example, evidence suggest that, compared to humans, in nonhuman animals the integration of information in time and space is more “local”, and the capacity to integrate the information on a larger scale is more restricted. When we come to consider how animals integrate various stimulus dimensions, experimental evidence suggest that they tend to focus on some particular physical dimensions of the stimuli more than humans do, and they hardly combine information from different stimulus dimensions. When we come to consider more general categorization processes, it appears that nonhuman primates form categories but their categories seem to be more strongly tied to the perceptual input than those of humans, and abstract processes, when they emerge, need many more trials to develop, and/or do not generalize to untrained dimensions as readily as for humans. In other words, for most of the integrative functions we have considered in this chapter, nonhuman animals show behaviors that differ at least quantitatively, if not qualitatively from human behavior, and we propose that these differences might be the bottlenecks for the evolution of language. Obviously, evolution has no direction and mastering some human-like language is not and has never been an issue in nonhuman animals. However, it might be that a particular pattern of development of these domain-general functions was a prerequisite to the emergence of human language, and that favorable conjunctions occurred only once in the phylogeny of the primate group, giving rise to the human language. The idea that some domain-general nonlinguistic functions form the bottlenecks of language evolution will be further developed below.

To account for the evolution of language, many theorists have focused their attention on language specific functions, which were considered as key factors for the evolution of human language. For example, many traditional theories claimed that only humans have a low larynx (which was disconfirmed since), or that only humans have the ability to understand

and produce recursive structures allowing an infinite variability in language production. Here, we do not want to discount these explanations, but we think that they do not take the problem at its roots. At this point in the scientific endeavor, we think that it is now necessary to step back a bit to enlarge our view of the problem. Doing so, we can imagine two different scenarios on the origin of language.

The "language-first" scenario would be that our primate/prehominid ancestors had rather limited cognitive resources and it was the appearance of language that boosted their general cognitive capacities. Although there is no doubt that language is a booster for cognitive functions, we think that this scenario does not hold. To illustrate our rebuttal of this theory, imagine an animal cognitive system with a limited working memory span of $N=1$ item. How could a language system develop with such limited memory resources? A clear expansion of this memory system would be required before the animal would have a chance to develop a proto-language system. This very extreme example shows that language evolution requires a cognitive background to support it, and it is only when such a cognitive background has evolved – potentially as a response to ecological pressures – that a form of proto-language can have a chance to emerge.

The second possible scenario, called "prerequisites-first", is that language only emerged in our evolutionary history once critical domain-general functions had gained in cognitive power in our animal/pre-hominid ancestors. This scenario raises one major question: which domain general function(s), or combination of functions, must have evolved at first for the emergence of language? We have no clear answer to this question but can provide several lines of thinking on this issue. First, we note that studies focusing on the so called low level perceptual mechanisms showed important differences between the platyrrhines and catarrhines species, but very little differences are observed in perceptual functions in monkeys, apes and human species (see for instance Fobes and King, 1982 for a review of visual perception). This mere fact suggests that the evolution of these perceptual functions is probably not the factor that made the difference and triggered the evolution of speech in humans. Secondly, comparative experiments suggest more pervasive differences between humans and the other primates in two domains at least. The first domain of importance is the domain of working memory. Working memory in humans seems to depart from that of other animal

species in several important aspects, for instance regarding the ability to process large amount of information in parallel (e.g., Fagot and De Lillo, 2011), or to process long-distance dependencies (e.g., Wilson et al., 2015). Although not discussed in the context of this chapter, working memory in humans may also qualitatively depart from that of the other animals in its use of a phonological loop facilitating memorization in the short term. The second domain for which strong differences emerge between humans and the other animals is the domain of attention. Nonhuman animals seem to focus on single stimulus dimensions more than humans do, and tend to have a more local mode of processing of the perceptual input than humans. There remains a debate on whether the increase in cognitive functions followed the phylogenetic order, from the remote prosimian species to the ape species phylogenetically closest to humans (e.g., Reader et al., 2011), or whether variations in cognitive power among the different primate species have occurred at multiple times in the course of evolution, in independent unrelated primates groups, for instance under the pressure of social factors such as the complexity of the social network (Dunbar, 1998). Discussing these hypotheses is out of the scope of the current chapter. However, whatever the source of this increment in cognitive power is, we propose that the language ability appeared in the evolution of primates at a point in time where domain-general cognitive capacities, especially those pertaining to attention and working memory converged and were sufficiently developed to permit its evolution. This idea is in line with recent usage-based theories suggesting that ~~that~~ language could be acquired in humans by means of domain-general – evolutionary old – processes (Bybee, 2010; Tomasello, 2005).

From a more practical standpoint, we conclude from this chapter that the comparative investigation of non-linguistic, domain general functions should be considered of central importance in the debate on the evolution of language. Unfortunately, real comparative studies, in which humans and other species are tested on the same problems using the same tasks, are relatively rare in this literature, and most of them only concern a very limited number of species. Such studies will become mandatory to further test the hypothesis that the expansion of domain-general functions in nonhuman primates served as a basis for the evolution of language.

Acknowledgements

This project was supported by ANR-16-CONV-0002 (ILCB), ANR-11-LA-BX-0036 (BLRI) and ANR-11-IDEX-0001-02 (A*MIDEX).

References

- Abe, K., and Watanabe, D. (2011). Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nature Neuroscience*, 14, 1067–1074.
- Anderson M. L. (2010). Neural reuse: a fundamental organizational principle of the brain. *Behavioural Brain Science*, 33, 245–266.
- Berg, M.E., and Grace, R.C. (2011). Categorization of multidimensional stimuli by pigeons. *Journal of Experimental Analysis of Behavior*. 95, 305–326.
- Boë, L.J., Berthommier, F., Legou, T., Captier, G., Kemp C., Sawallis, T.R., Becker, Y., Rey, A., and Fagot, J. (2017). Evidence of a vocalic proto-system in the baboon (*Papio papio*) suggests pre-hominin speech precursors. *PLoS ONE*, 12, e0169321.
- Bybee, J. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press.
- Byrne, R.W., Corp, N., and Byrne, J.M. (2001). Estimating the complexity of animal behaviour: How mountain gorillas eat thistles. *Behaviour*, 138, 525–557.
- Cantalupo, C., and Hopkins, W. D. (2001). Asymmetrical Broca’s area in great apes. *Nature*, 414, 505.
- Cantalupo, C., Pilcher, D., and Hopkins, W. D. (2003). Are planum temporale and sylvian fissure asymmetries directly related? A MRI study in great apes. *Neuropsychologia*, 41, 1975–1981.
- Chen, J., van Rossum, D., and ten Cate, C. (2015). Artificial grammar learning in zebra finches and human adults: XYX versus XXY. *Animal Cognition*, 18, 151–164.
- Chomsky, N. (1959). A review of B.F. Skinner’s “Verbal Behavior.” *Language*, 35, 26–58.
- Christie, S., and Gentner, D. (2014). Language helps children succeed on a classic analogy task. *Cognitive Science*, 38, 383–397.

- Colombo, M., and Frost, N. (2001). Representation of serial order in humans: A comparison to the findings with monkeys (*Cebus apella*). *Psychonomic Bulletin and Review*, 8, 262–269.
- Cook R. G. (2001). Hierarchical stimulus processing in pigeons. In R. G. Cook (Ed.), *Avian visual cognition* [On-line]. Available: www.pigeon.psy.tufts.edu/avc/cook/
- Creel, S. C., Newport, E. L., and Aslin, R. N. (2004). Distant melodies: statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 1119–1130.
- D'Amato, M. R., and Van Sant, P. (1988). The person concept in monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes*, 14, 43–55.
- Davenport, R.K., and Rogers, C.M. (1970). Intermodal equivalence of stimuli in apes. *Science*, 1968, 279–280.
- Dépy, D., Fagot, J. and Vauclair, J. (1997). Categorization of three-dimensional stimuli by humans and baboons (*Papio papio*): Search for prototype effects. *Behavioral Processes*, 37, 299–306.
- Deruelle, C., and Fagot, J. (1998). Visual search for global/local stimulus features in humans and baboons. *Psychonomic Bulletin and Review*, 5, 476–481.
- Dunbar, R.I.M (1998). The social brain hypothesis. *Evolutionary Anthropology*, 6, 178–190.
- Fagot, J., and Barbet, I. (2006). Grouping and segmentation of visual objects by baboons (*Papio papio*) and humans (*Homo sapiens*). In E. A. Wasserman and T. R. Zentall (Eds.), *Comparative cognition: experimental explorations of animal intelligence* (pp. 11–24). New York: Oxford University Press.
- Fagot, J., and De Lillo, C. (2011). A comparative study of working memory: immediate serial spatial recall in baboons (*Papio papio*) and humans. *Neuropsychologia*, 49, 3870–3880.
- Fagot, J., and Deruelle, C. (1997). Processing of global and local visual information and hemispheric specialization in humans (*Homo sapiens*) and baboons (*Papio papio*). *Journal of Experimental Psychology: Human Perception and Performance*, 23, 429–442.

- Fagot, J., Martin-Malivel, J., and Dépy, D. (2000). What is the evidence for an equivalence between objects and pictures in birds and nonhuman primates? In J. Fagot (Ed.), *Picture Perception in Animals* (pp 295–320). East Sussex: Psychology Press.
- Fagot, J. and Parron, C. (2010). Relational matching in baboons (*Papio papio*) with reduced grouping requirements. *Journal of Experimental Psychology: Animal Behavior Processes*, 36, 184–193.
- Fagot, J., and Parron, C. (2012). Visual cognition in nonhuman primates. In O. Lazaverra, T. Shimizu and E.A. Wasserman (Eds.), *How animals see the world* (pp. 371–389). Oxford: Psychology Press.
- Fagot, J., and Thompson, R. K. R. (2011). Generalized relational matching by Guinea baboons (*Papio papio*) in two by two-item analogy problems. *Psychological Science*, 22, 1304–1309.
- Fitch, W.T. (2010) *The evolution of language*. Cambridge University Press.
- Fitch, WT, de Boer, B., Mathur, N., and Ghazanfar, A.A. (2016). Monkey vocal tracts are speech-ready. *Science advances*, 2:e1600723.
- Fitch, W. T., and Hauser, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science*, 303 (5656), 377–380.
- Flemming, T. M., Thompson, R. K. R., and Fagot, J. (2013). Baboons, like humans, solve analogy by categorical abstraction of relations. *Animal Cognition*, 16, 519–524.
- Fobes, J.L., and King, J.E. (1982). Vision: The dominant primate modality. In J.L. Fobes and J.E. King (Eds.), *Primate behavior* (pp. 219–243). New York: Academic Press.
- Fountain, S. B. (1990). Rule abstraction, item memory, and chunking in rat serial-pattern tracking. *Journal of Experimental Psychology. Animal Behavior Processes*, 16, 96–105.
- Fountain, S. B., and Benson Jr., D. M. (2006). Chunking, rule learning, and multiple item memory in rat interleaved serial pattern learning. *Learning and Motivation*, 37, 95–112.
- Frank, A.J. and Wasserman, E.A. (2005). Associative symmetry in the pigeon after successive matching-to-sample training. *Journal of Experimental Analysis of Behavior*, 84, 147–165.
- Froehlich, A. L., Herbranson, W. T., Loper, J. D., Wood, D. M., and Shimp, C. P. (2004). Anticipating by pigeons depends on local statistical infor-

- mation in a serial response time task. *Journal of Experimental Psychology: General*, 133, 31–45.
- Gebhart, A. L., Newport, E. L., and Aslin, R. N. (2009). Statistical learning of adjacent and nonadjacent dependencies among nonlinguistic sounds. *Psychonomic Bulletin and Review*, 16, 486–490.
- Gentner, T. Q., Fenn, K. M., Margoliash, D., and Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature*, 440, 1204–1207.
- Geschwind, N., and Levitsky, W. (1968). Human brain: Left-right asymmetries in temporal speech region. *Science*, 161, 186–187.
- Gomez-Robles, A., Hopkins, W.D., and Sherwood, C.C. (2013). Increased morphological asymmetry, evolvability and plasticity in human brain evolution. *Proceedings of the Royal Society: B*, 280, 20130575.
- Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how does it evolve? *Science*, 298, 1569–1579.
- Hauser, M. D., Newport, E. L., and Aslin, R. N. (2001). Segmentation of the speech stream in a nonhuman primate: statistical learning in cotton-top tamarins. *Cognition*, 78, B53–B64.
- Hayes, C. (1951), *The Ape in Our House*. New York: Harper.
- Heijningen, C. A. A. van, Chen, J., Laatum, I. van, Hulst, B. van der, and Cate, C. ten. (2012). Rule learning by zebra finches in an artificial grammar learning task: which rule? *Animal Cognition*, 16, 165–175.
- Herrnstein, R.J., and Loveland, D.H. (1964). Complex visual concept in the pigeon. *Science*, 23, 549–551.
- Hockett, C. (1960). The origin of speech. *Scientific American*, 203, 88–111. Reprinted in Wang, William S-Y. (1982) *Human Communication: language and its psychobiological bases*, Scientific American pp. 4–12.
- Hofstater, D., and Sanders, E. (2013). *Surfaces and essences: analogy as the fuel and fire of thinking*. New-York: Basic Books.
- Hunt, R. H., and Aslin, R. N. (2001). Statistical learning in a serial reaction time task: Access to separable statistical cues by individual learners. *Journal of Experimental Psychology: General*, 130, 658–680.

- Jürgens, U. (1979). Neural control of vocalization in nonhuman primates. In H.D. Steklis and M. J. Raleigh (Eds.), *Neurobiology of social communication in primates* (pp. 199–206). New-York: Academic Press.
- Koechlin, E, and Jubault, T (2006). Broca's area and the hierarchical organization of human behavior. *Neuron*, 50, 963–74.
- Lanz, F., Moret, V., Rouiller, E.M., and Loquet, G. (2013). multisensory integration in nonhuman primates during a sensory-motor task. *Frontiers in Human Neuroscience*, 7,799.
- Lea, S. G. E., and Wills, A. J. (2008). Use of multiple dimensions in learned discriminations. *Comparative Cognition and Behavior Reviews*, 3, 115–133.
- Leroy, F. et al. (2014). New human-specific brain landmark: The depth asymmetry of superior temporal sulcus. *PNAS*, 112, 1208–1213.
- Lieberman, P. (1968). Primate vocalizations and human linguistic ability. *Journal of the Acoustic Society of America*, 44, 1574–84.
- Lieberman, P. (1975). *On the Origins of Language: an Introduction to the Evolution of Human Speech*. New York: Macmillan.
- Lionello-DeNolf, K.M. (2009). The search for symmetry: 25 years in review. *Learning and Behavior*, 37, 188–203.
- Lipkens, G., Hayes, S. C., and Hayes, L. J. (1993). Longitudinal study of derived stimulus relations in an infant. *Journal of Experimental Child Psychology*, 56, 201–239.
- Marcus, G. F., Vijayan, S., Rao, S. B., and Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, 283, 77–80.
- Malivel, M. and Fagot, J. (2001). Cross-modal interference and conceptual categorization in baboons. *Behavioural Brain Research*, 122, 209–213.
- Martin-Malivel, J. and Fagot, J. (2001). Perception of pictorial human faces by baboons (*Papio papio*): effects of stimulus orientation on discrimination performance. *Animal Learning and Behaviour*, 29, 10–20.
- Martin-Malivel, J. Mangini, M., Fagot, J. and Biederman, I (2006). Do humans and baboons use the same information when categorizing human and baboon faces ? *Psychological Science*, 17, 599–607.
- Medam, T., Marzouki, Y., Montant, M., and Fagot, J. (2016). Categorization does not promote symmetry in Guinea baboons (*Papio papio*). *Animal Cognition*, 19, 987–99.

- Meguerditchian, A., Marie, D., Love, S. A., Margiotoudi, K., Bertello, A., Lacoste, R., Roth, M., Nazarian, B, Anton, J.-L., and Coulon, O. (2016). Human-like brain specialization in baboons: an in vivo anatomical MRI study of language areas homologs in 96 subjects. In S.G. Roberts et al. (Eds.), *The evolution of language: Proceedings of the 11th International Conference (EVLANG11)*. New Orleans, LA, USA.
- Minier, L., Fagot, J., and Rey, A. (2016). The temporal dynamics of regularity extraction in nonhuman primates. *Cognitive Science*, 40, 1019–1030.
- Navon, D. (1977). Forest before the tree: The precedence of global feature in visual perception. *Cognitive Psychology*, 9, 353–383.
- Newport, E. L., and Aslin, R. N. (2004). Learning at a distance I. Statistical learning of nonadjacent dependencies. *Cognitive Psychology*, 48, 127–162.
- Newport, E. L., Hauser, M. D., Spaepen, G., and Aslin, R. N. (2004). Learning at a distance II. Statistical learning of nonadjacent dependencies in a nonhuman primate. *Cognitive Psychology*, 49, 85–117.
- Orbán, G., Fiser, J., Aslin, R.N., and Lengyel, M. (2008). Bayesian learning of visual chunks by human observers. *Proceedings of the National Academy of Sciences USA*, 105, 2745–50.
- Pepperberg, I.M. (2006). Ordinality and inferential abilities of a grey parrot (*Psittacus erithacus*). *Journal of Comparative Psychology*, 120, 205–216.
- Premack, D. (1983). The codes of man and beast. *Behavioral and Brain Sciences*, 6, 125–137.
- Petkov, C. I., and Wilson, B. (2012). On the pursuit of the brain network for proto-syntactic learning in nonhuman primates: conceptual issues and neurobiological hypotheses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367, 2077–2088.
- Procyk, E., Ford Dominey, P., Amiez, C., and Joseph, J. P. (2000). The effects of sequence structure and reward schedule on serial reaction time learning in the monkey. *Brain Research. Cognitive Brain Research*, 9, 239–248.
- Quinn, P.C., Eimas P.D., and Rosenkrantz, S.L. (1993). Evidence for representations of perceptually similar natural categories by 3-month-old and 4-month-old infants. *Perception*, 22, 463–475.

- Ravignani, A., Sonnweber, R.-S., Stobbe, N., and Fitch, W. T. (2013). Action at a distance: dependency sensitivity in a New World primate. *Biology Letters*, 9, 20130852.
- Reader, S.M., Hager, Y. and Laland, K.N. (2011). The evolution of primate general and cultural intelligence. *Philosophical Transactions of the Royal Society, B*, 366, 1077–1027.
- Saffran, J. R., Aslin, R. N., and Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926–1928.
- Saffran, J.R., and Thiessen, E.D. (2008). Domain-general learning capacities. In E. Hoff and M. Shatz (Eds.), *Blackwell Handbook of Language Development* (pp. 68–86). New-York: Blackwell Publishing Ltd.
- Schlenker, P., Chemla, E., and Zuberbühler, K. (2016). What do monkey calls mean? *Trends in Cognitive Sciences*, 20, 894–904.
- Seidenberg M., and MacDonald M. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science*, 23, 569–588.
- Seidenberg, M., MacDonald, M., and Saffran, J. (2002) Does grammar start where statistics stop? *Science*, 295, 553–554.
- Sidman, M., and Tailby, W. (1982). Conditional discrimination vs. matching to sample: an expansion of the testing paradigm. *Journal of Experimental Analysis of Behavior*, 37, 5–22.
- Skinner, B. F. (1957). *Verbal Behavior*. New York: Appleton-Century-Crofts.
- Smith, J. D., Berg, M.E., Cook, R.G., Murphy, M.S., Crossley, M.J., Boomer, J., Spiering, B., Beran, M.J., Church, B., Ashby, G., and Grace, R.C. (2012). Implicit and explicit categorization: A tale of four species. *Neuroscience and Biobehavioral Reviews*, 36, 2355–2369.
- Spierings, M. J., and Ten Cate, C. (2016). Budgerigars and zebra finches differ in how they generalize in an artificial grammar learning experiment. *Proceedings of the National Academy of Sciences of the United States of America*, 113, E3977-3984.
- Suzuki, T. N., Wheatcroft, D., and Griesser, M. (2016). Experimental evidence for compositional syntax in bird calls. *Nature Communication*, 7, 10986.

- Taglialatela, J.P., Russell, J.L., Schaeffer, J.A., and Hopkins, W.D. (2008). Communicative signaling activates “Broca’s” homologue in chimpanzees. *Current Biology*, 18, 343–348.
- Taglialatela, J.P., Russell, J.L., Schaeffer, J.A., and Hopkins, W.D. (2011). Chimpanzee vocal signaling points to a multimodal origin of human language. *PLoS ONE*, 6, e18852.
- Terrace, H. S. (1991). Chunking during serial learning by a pigeon: I. Basic evidence. *Journal of Experimental Psychology: Animal Behavior Processes*, 17, 81–93.
- Terrace, H. S., Son, L. K., and Brannon, E. M. (2003). Serial expertise of rhesus macaques. *Psychological Science*, 14, 66–73.
- Thompson, R. K. R., and Oden, D. L. (2000). Categorical perception and conceptual judgments by nonhuman primates: The paleological monkey and the analogical ape. *Cognitive Science*, 24, 363–396.
- Tomasello, M. (2005). *Constructing a language*. Harvard: Harvard University Press.
- Toro, J. M., and Trobalón, J. B. (2005). Statistical computations over a speech stream in a rodent. *Perception and Psychophysics*, 67, 867–875.
- Wang, L., Uhrig, L., Jarraya, B., and Dehaene, S. (2015). Representation of numerical and sequential patterns in macaque and human brains. *Current Biology*, 25, 1966–1974.
- Wasserman, E.A., Castro, L. and Fagot, J. (in press). Relational Thinking in Animals and Humans. In Call J.P. (Ed). *From Percepts to Concepts”: APA Handbook of Comparative Psychology*. New-York: American Psychological Association.
- Wasserman, E. A., Castro, L., and Freeman, J. H. (2012). Same-different categorization in rats. *Learning and Memory*, 19, 142–145.
- Wasserman, E. A., Fagot, J., and Young, M. E. (2001). Same-different conceptualization by baboons (*Papio papio*): The role of entropy. *Journal of Comparative Psychology*, 115, 42–52.
- Wasserman, E.A., Young, M.E., and Cook, R.G. (2004). Variability discrimination in humans and animals: implications for adaptive action. *American Psychologist*, 59, 879–890.
- Wilson, B., Smith, K., and Petkov, C. I. (2015). Mixed-complexity artificial grammar learning in humans and macaque monkeys: evaluating learning strategies. *The European Journal of Neuroscience*, 41, 568–578.

- Wright, A. A., and Katz, J. S. (2006). Mechanisms of same/different concept learning in primates and avians. *Behavioural Processes*, 72, 234–254.
- Wright, A. A., Santiago, H. C., Sands, S. F., Kendrick, D. F., and Cook, R. G. (1985). Memory processing of serial lists by pigeons, monkeys, and people. *Science*, 229, 287–289.
- Younger, B.A., and Cohen, L.B. (1986). Developmental change in infants' perception of correlations among attributes. *Child Development*, 5, 803–81.