Do animals have Universal Grammar? A case study in phonology
Bridget Samuels, Marc Hauser, & Cedric Boeckx
To appear in *The Oxford Handbook of Universal Grammar*, ed. Ian Roberts

1.   Introduction

Do animals have Universal Grammar? The short answer must be "no." Otherwise, why (as Noam Chomsky has repeatedly asked) do human children learn language with strikingly little conscious effort, while no other animal has even come close to approximating human language, even with extensive training (e.g., apes, dolphins) or exposure (e.g., dogs) to massive linguistic input? But we must qualify this answer, particularly in light of the undeniable fact that a great many of the cognitive capacities which clearly serve our linguistic ability—rich conceptual systems, vocal imitation, categorical perception, and so on—are shared with other species, including some of our closest living relatives. This suggests that the question is more complicated than it might first appear.

We can cast some light on this issue by maintaining the distinction that Hauser et al. (2002a) make between the "Faculty of Language - Broad" (FLB) and "Faculty of Language - Narrow" (FLN). FLB includes all the systems that are recruited for language but need not be unique to language, or to humans, whereas FLN is the subset of FLB that is unique to our species and to language. At present, a dominant hypothesis among proponents of this view is that FLN is very small, perhaps consisting only of some recursive operations (i.e., Merge) and/or lexicalization[1] plus the mappings from narrow syntax to the interfaces (see Hauser at al. (2002a) and Fitch et al. (2005)). Thus, the question of whether animals have UG must be re-cast in terms of whether UG is equated with FLN, or whether we might instead admit that various species have parts of UG which are, by definition, part of FLB.

In the present work, we use phonology as a case study to show what type of cross-species evidence bears on the issue of whether animals have (various components of) UG, broadly construed. Phonology is a particularly interesting area to study with an eye towards determining how much can be attributed to mechanisms which are present in other cognitive areas and in other species. Pinker & Jackendoff (2005:212) claim that phonology constitutes a problematic counterexample to the idea that FLN is nearly empty, because "major characteristics of phonology are specific to language (or to language & music), [and] uniquely human," while Chomsky argues in recent work that phonology is an afterthought, an externalization system applied to an already fully-functional internal language system; phonological systems are "doing the best they can to satisfy the problem they face: to map to the [Sensory-Motor system] interface syntactic objects generated by computations that are "well-designed" to satisfy [Conceptual-Intentional system] conditions" but unsuited to communicative purposes (Chomsky 2008:136). Such a view accords with the evolutionary scenario developed by Hauser et al. (2002a) and Fitch et al. (2005), who hypothesize that language may have emerged rather suddenly as a result of minimal genetic changes with far-

---

[i] Hauser et al. focused on the idea that recursion and the mappings (or interfaces) to semantics and phonology might be the crucial component in FLN. However, it has proven difficult to pinpoint what is meant by recursion in the relevant sense, such that it may be unique to humans and to language. Another hypothesis has been proposed by authors such as Spelke (2003) and Boeckx (In press). On their view, it is not recursion and the interfaces but rather lexicalization—the ability to embed any concept in a "lexical envelope" which allows it to be recursively Merged—which arose uniquely in our species.

reaching consequences. If this is correct, phonology might make extensive use of abilities that already found applications in other cognitive domains at the time externalized language emerged.

One way of testing this hypothesis is to see how many of the mechanisms which potentially underlie phonology can be found in other species, though not necessarily in their communication systems. To the extent we can show that other species can do what phonological computations require—that is, the more we can minimize the need for evolving language- or phonology-specific abilities—the evolutionary picture painted by Hauser et al. gains credibility from an evolutionary/biological standpoint (Hornstein & Boeckx 2009; see already Lindblom, MacNeilage, and Studdert-Kennedy 1984:187). We suggest, in line with this view, that the operations and representations which underlie phonology were exapted, or recruited from other cognitive or perceptuo-motor domains for the purpose of externalizing language; they are thus subsumed by FLB, but not FLN.[2] If UG subsumes not only FLN but also FLB, the answer to our original question—do animals have UG?—can be answered largely in the affirmative.

2.   Animal phonology: what do we look for?

Hauser et al. (2002a:1573) list a number of approaches to investigating the Sensory-Motor system's properties (shown below in (1)), all of which are taken to fall outside FLN.

(1)   a.   Vocal imitation and invention
            Tutoring studies of songbirds, analyses of vocal dialects in whales,
            spontaneous imitation of artificially created sounds in dolphins
       b.   Neurophysiology of action-perception systems
            Studies assessing whether mirror neurons, which provide a core
            substrate for the action-perception system, may subserve gestural
            and (possibly) vocal imitation;
       c.   Discriminating the sound patterns of language
            Operant conditioning studies of categorical perception and the prototype magnet
            effect in mammals and birds;
       d.   Constraints imposed by vocal tract anatomy
            Studies of vocal tract length and formant dispersion in birds and
            primates;
       e.   Biomechanics of sound production
            Studies of primate vocal production, including the role of mandibular
            oscillations;
       f.   Modalities of language production and perception
            Cross-modal perception and sign language in humans versus
            unimodal communication in animals

While all of these undoubtedly deserve attention, they address two areas—how auditory categories are learned and how speech is produced—which are peripheral to the core of phonological computation and tell us little about how phonological objects are represented or

---

[ii] On language as an exaptation, see among others Piattelli-Palmarini (1989), Uriagereka (1998), Boeckx & Piattelli-Palmarini (2005), Hauser et al. (2002a), Fitch et al. (2005).

manipulated.[3] Yip (2006a,b) outlines a more directly relevant set of research aims, suggesting that we should investigate whether other species are capable of the following:[4]

(2) a. Grouping by natural classes
     b. Grouping sounds into syllables, feet, words, phrases
     c. Calculating statistical distributions from transitional probabilities
     d. Learning arbitrary patterns of distribution
     e. Learning/producing rule-governed alternations
     f. Computing identity (total, partial, adjacent, non-adjacent)

This list can be divided roughly into three parts (with some overlap): (2a,b) are concerned with how representations are organized, (2c,d) are concerned with how we arrive at generalizations about the representations, and (2e,f) are concerned with the operations that are used to manipulate the representations. There are at least three more areas to investigate in non-linguistic domains and in other species:

(3) g. Exhibiting preferences for contrast/rhythmicity
     h. Performing numerical calculations (parallel individuation and ratio comparison)
     i. Using computational operations relevant to phonology

In the sections to follow, we present evidence that a wide range of animal species are capable of the solving experimental tasks that provide evidence for features a-i, though it may be the case that there is no single species except ours in which all these abilities cluster in exactly this configuration. In other words, it may be that what underlies human phonology is a unique *combination* of abilities, but the individual abilities themselves may be found in many other species — a point made long ago by Charles Hockett, though with attention focused on different features. In §3, we focus on the mechanisms underlying capacities a,b and h — that is, how phonological material is grouped. Next, in §4, we turn to the abilities c-g, that is, the capacity to identify and produce patterns. In §5, we discuss capacities e and i, focusing on symbolic computation. Finally, we briefly discuss other linguistic modules and conclude in §6.

3.   Grouping

Since the hypothesis put forward by Hauser et al. (2002a) takes recursion and the interfaces to be the central property of FLN, much attention has been paid to groupings, particularly recursive ones, in language. While phonology is widely considered to be free of recursion,[5]

[iii] See Samuels (2009a, §3.2.1) for discussion of (c) and (f), which are relevant to questions of phonological acquisition and the building of phonological categories.

[iv] Yip mentions two additional items which also appear on Hauser et al.'s list: categorical perception/perceptual magnet effects and accurate production of sounds (mimicry).

[v] Some authors have argued for recursion in the higher levels of the prosodic hierarchy (e.g., at the Prosodic Word level or above). See Truckenbrodt (1995) for a representative proposal concerning recursion at the Phonological Phrase level. Even if this is correct (though see Samuels 2009a, Ch. 5), the recursive groupings in question are mapped from syntactic

grouping (of features, of segments, and of larger strings) is nonetheless an integral part of phonology, and there is evidence that infants perform grouping or "chunking" in non-linguistic domains as well (see, e.g., Feigenson & Halberda, 2004). Additionally, segmenting the speech stream into words, morphemes, and syllables depends on the converse of grouping, namely edge detection. We will discuss edge detection and pattern extraction in §4.

Human beings are masters at grouping and making inductive generalizations. In setting up the comparative approach to this problem, Cheney & Seyfarth (2007:118) make the point eloquently: "the tendency to chunk is so pervasive that human subjects will work to discover an underlying rule even when the experimenter has—perversely—made sure there is none." This point holds true across the board, not just for linguistic patterns.

Since the 1970's, categorical perception experiments on a wide range of species have provided evidence that animals are sensitive to the phonetic building blocks—features and segments—which are grouped into larger constituents in human phonology. Many studies beginning with Kuhl & Miller's (1975) pioneering work on chinchillas show that mammals (who largely share our auditory system) are sensitive to many of the same acoustic parameters as define phonemic categories in human language (see further discussion in Samuels (2009a), §3.2). We can then begin to ask whether animals might be able to group segments into larger units such as syllables, words, and phrases, all of which are utilized by phonological processes in humans.

One way of approaching the question of whether animals can group sensory stimuli in ways that are relevant to phonology is to see whether their own vocalizations contain internal structure. The organization of bird song is particularly clear, though it is not obvious exactly whether or how analogies to human language should be made. Yip (2006a) discusses how zebra finch songs are structured, building on work by Doupe & Kuhl (1999) and others. The songs of many passerine songbirds consist of a sequence of one to three notes (or "songemes," as Coen (2006) calls them) arranged into a "syllable." The syllables, which can be up to one second in length, are organized into motifs which Yip considers to be equivalent to prosodic words but others equate with phrases, and there are multiple motifs within a single song. The structure can be represented graphically as follows, where *M* stands for motif, σ stands for syllable, and *n* stands for note (modified from Yip (2006a)):

(4)

---

structure, and are therefore not created by the phonological system alone. Furthermore, this type of recursive structure is also quite different from the type found in syntax (for example, sentential embedding) which is limited in its depth only by performance factors.

Song

M_1 M_2 M_3

...

σ_1 σ_2 σ_3 σ_1 σ_2 σ_3

n_1 n_2 n_3 n_4 n_5 n_6 n_1 n_2 n_3 n_4 n_5 n_6
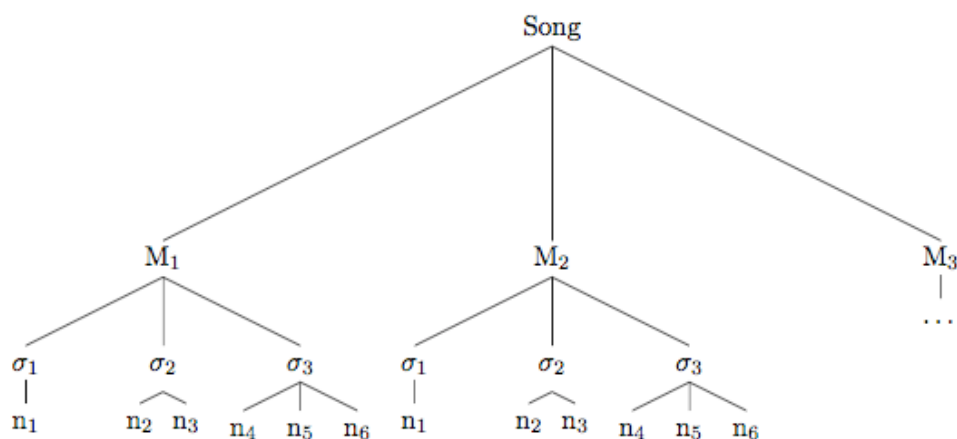
There are a few important differences between this birdsong structure and those found in human phonology, some of which are not apparent from the diagram. First, as Yip points out, there is no evidence for binary branching in this structure, which suggests that the combinatorial mechanism used by birds, whatever it is, cannot be equated with binary Merge. It could, however, be more along the lines of adjunction or concatenation, processes that create a flat structure (see, e.g., Samuels & Boeckx, 2009). Second, the definition of a syllable in birdsong is a series of notes/songemes bordered by silence (Williams & Staples 1992, Coen 2006). This is very unlike syllables, or indeed any other phonological categories, in human language. Third, the examples from numerous species in Slater (2000) show that the motif is typically a domain of repetition (as I have represented it above); the shape of a song is $((a^x)(b^y)(c^z))^w$ with a string of syllables *a, b, c* repeated in order. This is quite reminiscent of reduplication. Payne (2000) shows that virtually the same can be said of humpback whale songs, which take the shape $(a \ldots n)^w$, where the number of repeated components, *n*, can be up to around ten.

We should underscore the fact that the $(a \ldots n)^w$ patterns just discussed are combinatorial rules that define what counts as well-formed song: that is, they describe a type of syntax. Here the distinction made by Anderson (2004) and suggested in earlier work by Peter Marler (1977) is useful: a number of species have a "phonological" syntax to their vocalizations, but only humans have a "semantic" or "lexical" syntax which is compositional and recursive in terms of its meaning. Again, this reiterates Hauser et al.'s hypothesis that what is special about human language is the mapping from syntax to the interfaces (and particularly the semantic interface, as Chomsky emphasizes in recent writings; see, e.g., Chomsky (2004)), not the externalization system.

Concerning the development of phonological syntax, Fehér et al. (2009) conducted a study on zebra finches, which are close-ended learners and can only learn their songs during a critical period of development. They set out to test whether zebra finch isolates—birds who were deprived of the proper tutoring from adults during development—could evolve wild-type song over the course of generations in an isolated community. The experiment succeeded: the community of isolates, after only three or four generations, spontaneously evolved a song which approached the wild-type song of their species. Out of the chaos of stunted song, a well- behaved system emerged, stemming only from the birds themselves and the input they received from other individuals in the isolate colony. Such rapid self-organization of structure in vocalizations is truly remarkable and suggests to us that a similar process could be operative in creating human phonological systems. This is consistent with observations of emerging phonological structure in new sign

languages, such as Nicaraguan Sign Language (Senghas et al. 2005) and Al-Sayyid Bedouin Sign Language (Aronoff et al. 2008).

It should also be noted that both birdsong and whalesong structures are "flat" (in the sense of Neeleman & van de Koot (2006)) or "linearly hierarchical" (in the sense of Cheney & Seyfarth (2007))—they have a depth of embedding which is limited to a one-dimensional string which has been delimited intro groups, as in (5)—exactly what Samuels (2009a) has argued is true of human phonology. It is interesting to note in conjunction with this observation that baboon social knowledge is also of this type, as Cheney and Seyfarth have described. Baboons within a single social group (of up to about eighty individuals) obey a strict, transitive dominance hierarchy. But this hierarchy is divided up into matrilines; individuals from a single matriline occupy adjacent spots in the hierarchy, with mothers, daughters, and sisters from the matriline next to one another. So an abstract representation of their linear dominance hierarchy would look something like this, with each x representing an individual and parentheses defining matrilines:

(5)   (xxx)(xx)(xxxx)(xxx)(xxxxxxx)(xxx)(x)(xxxx)

The difference between the baboon social hierarchy and birdsong, which I translate into this sort of notation below, is merely the repetition which creates a motif (think of baboon individuals as corresponding to songemes and matrilines as corresponding to syllables):

(6)
$$\underbrace{(n_1)(n_2 n_3)(n_4 n_5 n_6)}_{\text{motif}_1} \underbrace{(n_1)(n_2 n_3)(n_4 n_5 n_6)}_{\text{motif}_2}$$

Not only does human language group segments into syllables and larger chunks, there are also melodies (prosodic or intonational contours) that play out over domains larger than the word, as well as languages in which each syllable bears a particular tonal pattern. Processing such melodies requires sensitivity to both relative and absolute pitch. There is evidence that rhesus monkeys, like us, treat a melody which is transposed by one or two octaves as more similar to the original than one which is transposed by a different interval (Wright et al. 2000). Rhesus monkeys can also distinguish rising pitch contours from falling ones, which is also required to perceive pitch accent, lexical tone, and intonational patterns in human speech (Brosch et al. 2004). However, most nonhuman animals are generally more sensitive to absolute pitch than they are to relative pitch; the opposite is true for humans (see Patel 2008). We discuss additional evidence that tamarins can discriminate cross-linguistic prosodic differences in Section 4.

There is evidence to suggest that, as in phonology (but strikingly unlike narrow syntax), the amount of hierarchy capable of being represented by animals is quite limited. In the wild, apes and monkeys rarely perform actions that are hierarchically structured with sub-goals and sub-routines, and this is true even when attempts are made to train them to do so. Byrne (2007) offers one notable exception, namely the food processing techniques of gorillas. Byrne provides a flow chart detailing a routine, complete with several decision points and optional steps, that mountain gorillas use to harvest and eat nettle leaves. This routine comprises a minimum of five steps, and Byrne reports that the routines used to process other foods are of similar complexity. Byrne further notes that "all genera of great apes acquire feeding skills that are flexible and have *syntax-like* organisation, with hierarchical structure. . . . Perhaps, then, the precursors of linguistic syntax should be sought in primate *manual*

abilities rather than in their vocal skills" (Byrne 2007:12; emphasis his). We agree that manual routines provide an interesting source of comparanda for the syntax of human language, broadly construed (i.e., including the syntax of phonology). Fujita (2007) has suggested along these lines the possibility that Merge evolved from an action grammar of the type which would underlie apes' foraging routines. This said, even if such computations are to be found in the manual routines of animals during foraging, a further puzzle arises: why haven't these computations been deployed in other motor domains, including communicative facial and body gestures, as well as vocal gestures? One possible answer to this question is that unlike humans, the capacities that evolve in animals are locked into particular (adaptive) contexts, lacking any form of generality to new domains or contexts. If this is correct, it would say something about the nature of the interfaces, how they have evolved, perhaps uniquely within the genus Homo.

Other experiments suggest that non-human primates may be limited in the complexity of their routines in interesting ways. For example, Johnson-Pynn et al. (1999) used bonobos, capuchin monkeys, and chimpanzees in a study similar to one done on human children by Greenfield et al. (1972) (see also discussion of these two studies by Conway & Christiansen (2001)). These experiments investigated how the subjects manipulated a set of three nesting cups (call them A, B, C in increasing order of size). The subjects' actions were categorized as belonging to the 'pairing,' 'pot,' or 'subassembly' strategies, that exhibit varying degrees of embedding:[6]

(7)   a.   *Pairing strategy*: place cup B into cup C. Ignore cup A.
       b.   *Pot strategy*: first, place cup B into cup C. Then place cup A into cup B.
       c.   *Subassembly strategy*: first, place cup A into cup B. Then place cup B
              into cup C.

The pairing strategy is the simplest, requiring only a single step. This was the predominant strategy for human children up to twelve months of age, and for all the other primates—but the capuchins required watching the human model play with the cups before they produced even this kind of combination. The pot strategy requires two steps, but it is simpler than the subassembly strategy in that the latter, but not the former, requires treating the combination of cups A + B as a unit in the second step. (We might consider the construction of the A + B unit as being parallel to how complex specifiers and adjuncts are composed in a separate derivational workspace in the syntax; see Fujita (2007).) Human children use the pot strategy as early as eleven months (the youngest age tested) and begin to incorporate the subassembly strategy at about twenty months. In stark contrast, the non-human primates continued to prefer the pairing strategy, and when they stacked all three cups, they still relied on the pot strategy even though the experimenter demonstrated only the subassembly strategy for them. Though we should be careful not to discount the possibility that different experimental methodologies or the laboratory context is responsible for the non-humans' performance, rather than genuine cognitive limitations, the results are

---

[vi] The situation is actually substantially more complicated than this, because the subjects need not put the cups in the nesting order. To give a couple examples, putting cup A into cup C counts as the pairing strategy; putting cup A into cup C and then placing cup B on top counts as the pot strategy. I refer the reader to the original studies for explanations of each possible scenario. The differences between the strategies as I have described them in the main text suffice for present purposes.

consistent with the hypothesis that humans have the ability to represent deeper hierarchies than other primates. This is, of course, what we predict if only humans are endowed with the recursive engine that allows for infinite syntactic embedding (Hauser et al. 2002a).

Many other types of experimental studies have also been used to investigate how animals group objects. It is well known that a wide variety of animals, including rhesus monkeys, have the ability to perform precise comparisons of small numbers (<4) which fall within the range of the parallel individuation system. They can discriminate between, for instance, groups of two and three objects, and pick the group with more objects in it. As Hauser et al. (2000) note, such tasks require the animal to group the objects into distinct sets, then compare the cardinality of those sets. Further, a wide variety of animals, including birds, rodents, primates, human infants and adults, have the capacity to quantify large numbers approximately using an analog magnitude system (see Gallistel 1990 and Dehaene 1997 for reviews).

Further data concerning mathematical reasoning, and specifically transitivity, comes from Schusterman & Kastak (1993), who taught a California sea lion named Rio to associate arbitrary visual stimuli (cards with silhouettes of various objects printed on them). On the basis of being taught to select card B when presented with card A, and also to select card C when presented with card B, Rio transitively learned the A-C association.[7] Rio also made symmetric associations: when presented with B, she would select A, and so forth. We might consider these groups to be akin to learning arbitrary pairings such as which phonemes participate in a given alternation (A and C bear the same relation to B), or in which contexts a particular process occurs (choose A in the context of B; choose B in the context of C). This is, of course, an inference that would require far more careful testing.

The concept of "natural classes" has also been studied in animals to a certain degree, though not in those terms. We can think of natural classes as multiple ways of grouping the same objects into sets according to their different properties (i.e., features). Alex the parrot had this skill: he could sort objects by color, shape, or material (Pepperberg 1999).

Concerning the ability to group objects, then, we conclude that a wide variety of animals—perhaps especially birds and nonhuman primates—are capable of the basic grouping abilities which phonology requires. They perceive (some) sounds categorically like we do; their vocalizations show linearly hierarchical groupings like ours; they can assign objects arbitrarily to sets like we do; they can categorize objects into overlapping sets according to different attributes like we do. Their main limitations seem to be in the area of higher-degree embedding, but this is (a) at best, a property of phonology which arises because of recursion in syntax, not from a recursive engine within phonology and (b) an expected result if, as Hauser et al. (2002a) hypothesize, recursion is a part of FLN and therefore not shared with other species. We may view the situation in parallel to the difference between humans and animals in the domain of numerical cognition: in both domains, only humans have a recursive engine (Merge), and this mechanism interfaces with others to create novel representations (Hauser et al., 2002a). While many animals (and young human children) seem to be able to represent small numerals precisely and large numbers approximately, only suitably mature (and, perhaps, suitably linguistic) humans go on to learn the inductive principle, which allows them to count infinitely high and with precision. (See

---

[vii] See also Addessi et al. (2008) on transitive symbolic representation in capuchin monkeys, and Cheney & Seyfarth (2007) on transitive inference involving social hierarchy in baboons. Cheney & Seyfarth also discuss both transitive social dominance and learning of symbolic representations in pinyon jays.

§5 for related discussion, and Carey (2009)for a thorough treatise of how the child, but no nonhuman, makes the inductive leap from the evolutionary primitives to a mature system.)

4.   Patterns

The next set of abilities we will consider are those that deal with sequencing, extracting patterns from a data stream, and/or learning arbitrary associations. As mentioned in the previous section, pattern-detection is the flipside of grouping: a pattern is a relation between multiple groups, or different objects within the same group. Thus, the ability to assign objects to a set or an equivalence class is a prerequisite for finding any patterns in which those objects participate; the abilities discussed in the previous section are very much relevant to this one as well.

Several experimental studies on animal cognition more generally bear on the issue of abstract pattern learning. One such study, undertaken by Hauser & Glynn (2009), tested whether wild rhesus monkeys could extract simple algebraic rules like same-different-different (ABB) or same-same-different (AAB) comprised of their species-specific vocalizations. They performed an experiment very similar to one run on infants by Marcus et al. (1999). The habituation material was either in the form of ABB or AAB with novel test material that either matched the habituation stimuli or mismatched. Specifically, after habituating the infants/rhesus to one of these conditions, they tested them on two novel test items: one from the same class to which they had been habituated, and a second from the other class. If subjects extracted the pattern during habituation, then they should respond more to the mismatched test trial than to the novel, but matching test trial. Both infants and rhesus evidenced learning of these simple patterns; they were more likely to dishabituate to the item with the mismatched pattern.

This type of pattern-extraction ability could serve phonology in several ways, such as the learning of phonological rules or phonotactic generalizations. Heinz (2007) showed that phonotactics (restrictions on the co-occurrence of segments, such as at the beginnings or ends of words) can be captured without any exceptions if three segments at a time are taken into account, so it seems on the basis of the rhesus monkey's success in the Hauser & Glynn experiment, together with other work by Murphy et al. (2008) on rats, that learning phonotactics would not be out of their range of capabilities. Furthermore, phonotactics (and all attested phonological rules) can be modeled with finite-state grammars, as has been known since Johnson (1970). In theory, then, if primates or other animals can learn finite-state patterns, we would expect that they could learn any attested phonotactic restriction or phonological rule. The findings of Fitch & Hauser (2004) are also relevant here: at least under one interpretation of the data obtained by Fitch and Hauser, tamarins succeed at learning finite-state grammars but fail to learn more complicated phrase-structure grammars The conclusion that animals are limited to finite-state computation is also supported by van Heijningen et al.'s (2009) nuanced testing of finches on similar patterns, which casts strong suspicion on Gentner et al.'s (2006) claim that starlings can go beyond finite-state patterns to learn a recursive context-free rule; the interpretation of these data are still under debate (ten Cate et al. 2010, Gentner et al. 2010).

One of the most important obstacles facing a language learner/user falls into the category of pattern-extraction. This difficult task entails parsing the continuous speech stream into discrete units (be they phrases, words, syllables, or segments). This speaks directly to (2b,c). Obviously, segmenting speech requires some mechanism for detecting the edges of these units. Since the 1950's, it has been recognized that one

way to detect the edges of words is to track transitional probabilities, usually between syllables. If Pr(AB) is the probability of syllable B following syllable A, and P(A) is the frequency of A, then the transitional probability between A and B can be represented as:

(8)
$$TP(A \rightarrow B) = \frac{Pr(AB)}{Pr(A)}$$

The transitional probabilities within words are typically greater than those across word boundaries, so the task of finding word boundaries reduces to finding the local minima in the transitional probabilities. Numerous experimental studies suggest that infants do in fact utilize this strategy (among others) to help them parse the speech stream, and that statistical learning is not unique to the linguistic domain but is also utilized in other areas of cognition (see references in Gambell & Yang (2005)). With respect to the availability of this strategy in non-humans, Hauser et al. (2001) found that tamarins are spontaneously able to segment a continuous stream of speech into three-syllable CVCVCV "words" based solely on the transitional probabilities between the syllables. Rats are also sensitive to local minima in transitional probabilities, albeit after a period of operant conditioning (Toro et al. 2005).

While transitional probabilities between syllables are strictly local calculations (i.e., they involve adjacent units), some phonological (and syntactic) dependencies are non-adjacent. This is the case with vowel harmony, for instance, and is also relevant to languages with templatic morphology, such as Arabic, in which a triconsonantal root is meshed with a different group of vowels depending on the part of speech which the root instantiates in a particular context. Comparing the results obtained by Newport & Aslin (2004) and Newport et al. (2004) provides an extremely interesting contrast between human and tamarin learning of such patterns. Newport et al. tested adult humans and cotton-top tamarins on learning artificial languages, all with three-syllable CVCVCV words, involving the three different kinds of non-adjacent dependencies listed below.

(9)   a.   *Non-adjacent syllables*: the third syllable of each word was predictable on the basis of the first, but the second syllable varied.
      b.   *Non-adjacent consonants*: The second and third consonants of each word were predictable on the basis of the first, but the vowels varied.
      c.   *Non-adjacent vowels*: The second and third vowels of each word were predictable on the basis of the first, but the consonants varied.

Both humans and tamarins succeeded in the non-adjacent vowel condition. Humans also succeeded in the non-adjacent consonant condition. These results are expected, at least for humans, because both of these types of dependencies are attested in natural language (in the guises of vowel harmony and templatic morphology). Tamarins failed in the non-adjacent consonant condition, but we suspect that they have the cognitive capability needed to create the appropriate representations, though they might have difficulty distinguishing consonant sounds. In other words, their failure is unlikely to be due to the pattern-detection mechanism, but rather due to the input available to that mechanism. This interpretation is supported by the fact that tamarins succeeded at establishing dependencies between non-adjacent vowels.

From a phonological perspective, perhaps the most intriguing result is that humans failed at this non-adjacent syllable condition, whereas the tamarins succeeded. Newport et al. (2004:111) ask:

"Why should non-adjacency—particularly syllable non-adjacency—be difficult for human listeners and relatively easy for tamarin monkeys? […T]his is not likely to be because tamarins are in general more cognitively capable than adult humans. It must therefore be because human speech is processed in a different way by humans than by tamarins, and particularly in such a way that the computation of non-adjacent syllable regularities becomes more complex for human adults."

They go on to suggest that perhaps the syllable level is only indirectly accessible to humans because we primarily process speech in terms of segments (whereas tamarins process it in more holistic, longer chunks).[8] This is a possible contributor to the observed effect, but other explanations are available. Perhaps, for example, tamarins fail to exhibit a minimality effect that humans exhibit in this circumstance.[9] Let us interpret the tamarins' performance in the non- adjacent consonant condition as suggesting that they either ignore or simply do not perceive consonants. Then for them, the non-adjacent syllable task differs minimally from the non-adjacent vowel task in that the former involves learning a pattern which skips the middle vowel. So rather than paying attention to co-occurrences between adjacent vowels, they have to look at co-occurrences between vowels which are one away from each other. It seems likely, as Newport et al. also suggest, that the adjacent vs. one-away difference represents only a small increase in cognitive demand. But for humans, the non-adjacent syllable condition is crucially different—and this is true no matter whether we are actually paying attention to syllables, consonants, or vowels. These categories have no import for tamarins, but for humans, they are special. The dependency we seek in this condition is between two non-adjacent elements of the same category, which are separated by another instance of the same category. This is a classical minimality effect: if α, β, γ are of the same category and $\alpha > \beta > \gamma$ ($>$ should be read for phonology as "precedes" and for syntax, "c-commands"), then no relationship between α and γ may be established. Perhaps tamarins succeed where humans fail because they do not represent the portions of the stimuli which they track as all belonging to the same abstract category "vowel" which is sufficient to trigger minimality effects for us.

A variety of other studies on primate cognition focus on the ability to learn sequences. Given that sequencing or precedence relationships are extremely important to language, particularly given the Minimalist emphasis on Merge in syntax and Samuels' (2009a) parallel emphasis on concatenate in phonology, these studies are intriguing from a linguist's perspective. As Endress et al. (2009a) emphasize, temporal sequencing underlies everything from foraging routines to structured vocalizations (i.e., the order of segments, morphemes, and words in human speech as well as notes and motifs in animal songs) and requires a specific type of memory which encodes the positions of items within the sequence. Conway & Christiansen (2001) report on a number of studies which compare primates' performances on this kind of task. When presented with an artificial fruit requiring four arbitrary actions to

---

[viii] Newport et al. suggest that tamarins' shorter attention spans reduce the amount of speech that they process at a given time; this would restrict their hypothesis space, making the detection of the syllable pattern easier. It is not obvious how this explains the tamarins' pattern of performance across tasks, however.
[ix] Such effects have been discussed in terms of Relativized Minimality (Rizzi 1990) or the Minimal Link Condition (Chomsky 2000, 2004) in syntax and the No Line-Crossing Constraint (Goldsmith 1976) in autosegmental phonology. Samuels (2009a) argues that minimality in phonology and syntax emerges from the same underlying cause: a directional search mechanism which traverses strings of segments.

open it and thereby reveal a treat, chimpanzees and human preschoolers perform similarly; both succeed at learning the sequence.

However, one apparent cognitive limitation of non-human primates relative to our species in the domain of pattern-learning is that they have extreme difficulty with non-monotonic sequences. This seems to be related to a difference in the way humans and other primates plan and perform sequential actions. One experiment undertaken by Ohshiba (1997) tested human adults, Japanese monkeys, and a chimpanzee on the ability to learn an arbitrary pattern: they were presented with a touch screen with four different-sized colored circles on it and had to touch each one in sequence to receive a reward; the circles disappeared when touched. All the species succeeded in learning a monotonic pattern: touch the circles in order from smallest to largest or largest to smallest. They also all succeeded, but were slower, at learning non- monotonic patterns.[10] But as we will discuss in §5, measurements of reaction times suggest the humans and monkeys used different strategies in planning which circles to touch.

Rhythm, too, is a type of pattern. Rhythmicity, cyclicity, and contrast are pervasive properties of language, particularly in phonology. Everything that has been attributed to the Obligatory Contour Principle (Leben 1973) — a constraint which prohibits the adjacency of two phonological elements which are similar or identical — fits into this category. Walter (2007) argues that these effects should be described not with a constraint against repetition (see also Reiss 2008), but as emerging from two major physical limitations: the difficulty of repeating a particular gesture in rapid succession, and the difficulty of perceiving similar sounds (or other sensory stimuli) distinctly in rapid succession. These are both extremely general properties of articulatory and perceptual systems which we have no reason to expect would be unique to language or to humans.

To date, perhaps the most direct cross-species test of the perception of human speech rhythm (prosody) comes from Ramus et al. (2000) and Tincoff et al. (2005). In Ramus et al.'s experiment, human infants and cotton-top tamarins were tested on their ability to discriminate between Dutch and Japanese sentences under a number of conditions: one in which the sentences were played forward, one in which the sentences were played backward, and one in which the sentences were synthesized such that the phonemic inventory in each language was reduced to /s a l t n j/. The results of these experiments showed that both tamarins and human newborns were able to discriminate between these two unfamiliar and prosodically different languages in the forward-speech condition, but not in the backward-speech condition. A generous interpretation of these results would suggest "at least some aspects of human speech perception may have built upon preexisting sensitivities of the primate auditory system" (Ramus et al. 2000:351). However, Werker & Voloumanos (2000) caution that we cannot conclude much about the processing mechanisms which serve these discrimination abilities; this is of particular concern given that the tamarins' ability to tell Dutch and Japanese apart was reduced in the reduced phonemic inventory condition. This may indicate that tamarins rely more strongly on phonetic cues rather than prosodic ones. Given the apparent importance

---

[x] In some situations, non-human primates fail entirely at learning non-monotonic patterns. For example, Brannon & Terrace (1998, 2000) found that while rhesus macaques taught the first four steps in a monotonic pattern (integers) could spontaneously generalize to later steps, they failed to learn a four-member non-monotonic pattern even with extensive training. It is not clear what to attribute the worse performance in the Brannon & Terrace studies to; there are too many differences between the paradigm they used and the one reported in the main text, including the species tested.

of prosody for syntactic acquisition in human children—specifically, babies seem to use prosodic information to help them set the head parameter—Kitahara (2003:38) put forth the idea that "cotton-top tamarins fail to discriminate languages on the basis of their prosody alone, because syntactic resources that require such prosodic-sensitive system [sic] might not have evolved for them." Though it is unclear how one might either support or disprove such a hypothesis, it is at the very least interesting to consider what prosody might mean for an animal which does not have the syntactic representations from which prosodic representations are built.

Another example of rhythmicity in speech is the wavelike sonority profile of our utterances, which is typically discussed in terms of syllable organization. Syllables range widely in shape across languages. In (10)-(11) I give examples from opposite ends of the spectrum: a series of three CV syllables in (10), and a syllable in (11) that has a branching onset as well as a coda, and additionally appendices on both ends. The relative heights of the segments in (10)-(11) represent an abstract scale of sonority.

(10)
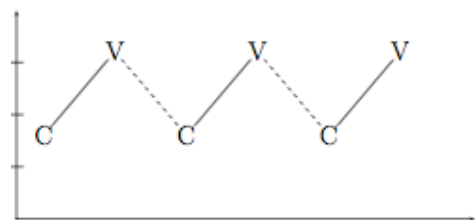
(11)
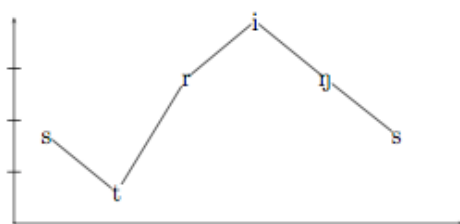
All syllables, from CV (11) to CCCVCC (12), combine to yield a sonority profile roughly as in (12):

(12)

The peaks and troughs may not be so evenly dispersed, and they may not all be of the same amplitudes, but the general shape is always the same. This is hardly a new observation; it is over a century old (e.g., Lepsius & Whitney (1865), de Saussure (1916)). Ohala & Kawasaki-Fukumori (1997:356) point out that it is inevitable:

> "Just by virtue of seeking detectable changes in the acoustic signal one would create as an epiphenomenon, i.e., automatically, a sequence showing local maxima and minima in vocal tract opening or loudness. In a similar way one

could find 'peaks' (local maxima) in a string of random numbers as long as each succeeding number in the sequence was different from the preceding one."

It has long been noted that the ability to break this wave up into periods (based partially on universal and partially on language-specific criteria) aids with the identification of word boundaries: they tend to fall at the local minima or maxima in the wave. And as we saw earlier in this section, we already know that both human infants and tamarins are sensitive to local minima (of transitional probabilities) in speech, which suggests that this is a legitimate possibility.

Animals from a wide variety of clades show preferences for rhythmicity in their vocalizations and other behaviors as well, though it is important to note that our own (non-musical) speech has no regular beat; while language does have a rhythm, it is not a primitive (see discussion in Patel 2008). Yip (2006b) mentions that female crickets exhibit a preference for males who produce rhythmic calls, and Taylor et al. (2008) discovered that female frogs prefer rhythmic vocalizations as well. Rhythmic behaviors, or the ability to keep rhythm, appear to be widespread in the animal kingdom. Gibbons produce very rhythmic 'great calls,' and while Yip (2006b:443) dismisses this, saying that "the illusion of rhythm is probably more related to breathing patterns than cognitive organization," this should hardly disqualify the data. For example, the periodic modulation of sonority in our speech is closely connected to opening and closing cycle of the jaw (Redford 1999, Redford et al. 2001), and it is widely accepted that the gradual downtrend in pitch which human utterances exhibit has to do with our breathing patterns. So for humans, too, there is at least some purely physiological component; however, the fact noted above that females of various species prefer rhythmic calls shows that at the very least, there is also a cognitive component to animals' perception of rhythmicity.

There are also some animals that synchronize the rhythms produced by multiple individuals. For example, frogs, insects, birds, and captive bonobos have all been shown to engage in chorusing during which they synchronize their calls; some fireflies synchronize their flashing, and crabs synchronize their claw-waving (see Merker 2000 and references therein). However, while elephants can be taught to drum with better rhythmic regularity than human adults, they do not synchronize their drumming in an ensemble (Patel & Iversen 2006).

5.   Operations

The final set of abilities which we will discuss are those that pertain to how phonological operations manipulate symbolic objects such as features and phonemes. It is particularly difficult to remain theory-neutral here because this is precisely the area in which phonological theories differ most: there is wide agreement that phonological representations bottom out in phonetic features (though we may debate whether those features are articulatory or acoustic, binary or privative, and so on), but concerning operations, there are a number of quite different proposals, from Autosegmental Phonology to Optimality Theory to Government Phonology. As a compromise, then, we focus here on a small number of basic operations which any theory of phonology (and indeed, of language more broadly) will have to admit: a search mechanism which identifies an object to be manipulated based on some characteristic (feature) of its content, a means by which features may be copied from one object to another and/or deleted from an object, and a mechanism for combining or

concatenating two objects.

Searching[11] is ubiquitous in animal and human cognition. It is an integral part of foraging and hunting for food, to take but one example. The Ohshiba (1997) study of sequence-learning by rhesus monkeys, humans, and a chimpanzee is an excellent probe of searching abilities in primates because it shows that, while various species can perform the multiple sequential searches required to perform the experimental task (touching four symbols in an arbitrary order), they plan out the task in different ways. Human subjects were slow to touch the first circle but then touched the other three in rapid succession, as if they had planned the whole sequence before beginning their actions (the collective search strategy). The monkeys, meanwhile, exhibited a gradual decrease in their reaction times. It was as if they planned only one step before executing it, then planned the next, and so forth (the serial search strategy).

Perhaps most interestingly of all, the one chimpanzee subject appeared to use the collective search strategy on monotonic patterns but the serial search strategy when the sequence was not monotonic. That at least some chimpanzees may employ collective searches is corroborated by the results of a similar experiment by Biro & Matsuzawa (1999). The chimp in this study, Ai, had extensive experience with numerals, and was required to touch three numerals on a touchscreen in monotonic order. Again, her reaction times were consistently fast after the initial step. But when the locations of the two remaining numerals were changed after she touched the first one, her reactions slowed, as if she had initially planned all three steps but her preparation was foiled by the switch.

Another operation which is important to phonological computation, and linguistic computation in general, is copying: for instance, a lexical item must be copied from long-term memory into a numeration each time it is to be used in the derivational workspace (Hornstein 2001). Copying (reduplication) plays a morphological role as well, serving a variety of grammatical functions. We might see the copying of a string of phonological segments, as in reduplication, as parallel to the repetitive patterns of song production found in birds and whales. As we saw in §3, Slater (2000) shows that for many bird species, songs take the shape $((a^x)(b^y)(c^z))^w$: that is, a string of syllables *a, b, c,* each of them repeated, and then the whole string repeated. We also saw that whale songs are similarly structured (Payne 2000). With respect to the copying of a feature from one segment to another (as in assimilatory processes), the relevant ability might be transferring a representation from long-term memory to short-term memory: extracting a feature from a lexical representation and bringing it into the active phonological workspace. This seems like a prerequisite for any task which involves the recall/use of memorized information.

In order to integrate a copied object with other material in the workspace, some sort of concatenative operation is necessary. Concatenation serves not only this grouping function (including assigning objects to sets), but also the ability to chain together sequential actions (recall §3, 4). We have already seen that bird and whale songs have the kind of sequential organization which is indicative of concatenated chunks, and primates can perform multi-step actions with sub-goals. In phonology, concatenation subserves the ability to connect morphemes—a root and an affix, for example—in a manner that creates a linear structure, but not the nested hierarchical structure of Merge.[12]

[xi] For present purposes, we define searching as the cognitive ability to identify the need for a property of an object/mental representation, and a plan for how this need can be met.
[xii] Whereas iterative concatenation yields a flat structure, iterative Merge yields a nested hierarchical structure. Samuels & Boeckx (2009) discuss this issue in greater detail.

As alluded to in §3, concatenation may underlie the numerical abilities common to humans and many other species as well (for an overview, see Dehaene 1997, Lakoff & Nuñez 2001, Devlin 2005, Shettleworth 2010). This is perhaps clearest in the case of parallel individuation/tracking, or the ability to represent in memory a small number of discrete objects (< 4; see Hauser et al. 2000 and references therein). The connection between parallel individuation and concatenation is suggested by the fact that the speed of recognizing the number of objects in a scene decreases with each additional object that is presented within the range of capability (Saltzman & Garner 1948). This leads us to suspect along with Gelman & Gallistel (1978) that such tasks require paying attention to each object in the array separately, yet they must be represented as a set in order to be manipulated arithmetically. Biro & Matsuzawa (1999) report that the chimpanzee Ai, when given rigorous training over a long period of time, was able to engage in basic counting, addition, and subtraction of natural numbers up to about ten. The ability to subtract has also been shown in pigeons (Brannon et al. 2001). These numerical tasks clearly involve the assignment of objects to sets, which is the fundamental basis of concatenation. Gallistel & Gelman (2005) and Shettleworth (2010) summarize the literature on counting and addition, which have been shown for species as diverse as parrots, rats, and chimpanzees.

6.   Conclusions

The studies of animal cognition and behavior presented here provide evidence that the building blocks of phonology are present in a wide range of species: many animals can group objects, extract patterns from sensory input, perform sequential objects, perform searches, engage in copying behaviors, and manipulate sets through concatenation. These capacities did not evolve for phonology, but rather, for other perceptual, cognitive and motor functions. Looking at the data we currently have, therefore, we conclude contra Pinker & Jackendoff (2005), that phonology provides little challenge to the idea that FLN consists of a small number of computations. With respect to our original question—do animals have phonological Universal Grammar?—we would say "yes" as we find evidence for the abilities that underlie human phonological competence scattered across a wide range of animal species, though no single species besides ours may possess *all* of these abilities. In this sense, what may be unique to humans is our capacity to interface between these different components. Samuels (2009a,b) pursues this idea further.

Although we have focused on phonology here, investigation into other linguistic competences is strongly warranted, though there are several interesting results in morphology and semantics already. For example, Endress et al. (2009b) argue that tamarins share the positional memory mechanism which allows for the encoding of affixation patterns in human language, and that tamarins can learn rules similar to those which produce prefixation and suffixation. It is interesting to juxtapose this with the findings of Ouattara et al. (2009), who report concatenative sequences in Campbell's monkey vocalizations. Such sequences might, under a generous interpretation, be described as something like morphosyntactic combination (i.e., combining meaningful units to create another meaningful unit). We emphasize, however, that it is extremely difficult to offer testable hypotheses about what, if any, type of compositionality beyond the composition of sound sequences (something which is already attested in a huge number of species) might be at work in this case.

On the semantic side, Hurford (2007) delves into the question of how other species construe the world around them and how they conceptualize meaning, with an eye towards understanding the evolutionary path which led to human language semantics. In this domain, there is a somewhat longer tradition of literature, including Cheney & Seyfarth (1990, 2007)

and much work by Zuberbühler (e.g., Arnold & Zuberbühler 2006, Zuberbühler 2005, 2006, et seq.). We hope, however, that this is only the beginning: much research remains to be done in each and every one of the domains discussed here, and we hope that the present work will be taken as an invitation to delve deeper and ask the more sophisticated questions which arise once we identify the basic points of potential consonance and divergence between human and animal cognition as far as phonology is concerned.

On this note, we would like to conclude by addressing one concern about the avenue of research we advance here: how do we know that the animal abilities under consideration are truly comparable to the representations and operations found in human phonology, and what if these abilities are only analogous, not homologous? It is probably premature to answer these questions for most of the abilities we have discussed. But even if we discover that these traits are only analogous, there is nonetheless significant progress to be made in this area of comparative linguistics. Hauser et al. (2002a:1572) underscore this:

> The remarkably similar (but nonhomologous) structures of human and octopus eyes reveal the stringent constraints placed by the laws of optics and the contingencies of development on an organ capable of focusing a sharp image onto a sheet of receptors. [. . .] Furthermore, the discovery that remarkably conservative genetic cascades underlie the development of such analogous structures provides important insights into the ways in which developmental mechanisms can channel evolution [(Gehring 1998)]. Thus, although potentially misleading for taxonomists, analogies provide critical data about adaptation under physical and developmental constraints. Casting the comparative net more broadly, therefore, will most likely reveal larger regularities in evolution, helping to address the role of such constraints in the evolution of language.

Analogies serve to highlight what Chomsky (2005, 2007) calls Third Factor principles—that is, general properties of biological/physical design—which might be at play, and help us to identify the set of constraints which are relevant to the evolutionary history of the processes under investigation. For example, both human infants and young songbirds undergo a babbling phase in the course of the development of their vocalizations. Even though the mechanisms responsible for babbling in the two clades are not homologous, nevertheless

> their core components share a deeply conserved neural and developmental foundation: Most aspects of neurophysiology and development—including regulatory and structural genes, as well as neuron types and neurotransmitters—are shared among vertebrates. That such close parallels have evolved suggests the existence of important constraints on how vertebrate brains can acquire large vocabularies of complex, learned sounds. Such constraints may essentially force natural selection to come up with the same solution repeatedly when confronted with similar problems.
> (Hauser et al. 2002a:1572)

We may not know what those constraints are yet, but until we identify the homologies and analogies between the mechanisms which underlie human and animal cognition, we cannot even begin to tackle the interesting set of questions that arise regarding the constraints on cognitive evolution. The present study, together with the perspective of phonologists like Yip, provides a place for us to begin this investigation in the domain of human phonological computation.

References

Addessi, Elsa, Alessandra Mancini, Lara Crescimbene, Camillo Padoa-Schiopp & Elisabetta Visalberghi. 2008. Preference transitivity & symbolic representation in capuchin monkeys (Cebus apella). *PLoS ONE* 3. E2414.

Anderson, Stephen R. 2004. *Dr. Dolittle's delusion: Animals & the uniqueness of human language*. New Haven: Yale University Press.

Arbib, Michael A. 2005. From monkey-like action recognition to human language. *Behavioral & Brain Sciences* 28. 105–167.

Aronoff, Mark, Irit Meir, Carol A. Padden & Wendy Sandler. 2008. The roots of linguistic organization in a new language. *Interaction Studies* 9. 133–153.

Arnold, Kate & Klaus Zuberbühler. 2006. Language evolution: semantic combinations in primate calls. *Nature* 441. 303.

Biro, Dora & Tetsuro Matsuzawa. 1999. Numerical ordering in a chimpanzee (Pan troglodytes): planning, executing, & monitoring. *Journal of Comparative Psychology* 113. 178–185.

Boeckx, Cedric. 2006. *Linguistic minimalism*. Oxford: Oxford University Press.

Boeckx, Cedric. In press. Some reflections on Darwin's problem in the context of Cartesian biolinguistics. In Anna-Maria Di Sciullo & Calixto Aguero (eds.), *Biolinguistic investigations*. Cambridge, MA: MIT Press.

Boeckx, Cedric & Massimo Piattelli-Palmarini. 2005. Language as a natural object, linguistics as a natural science. *The Linguistic Review* 22. 447–466.

Brannon, Elizabeth M. & Herbert S. Terrace. 1998. Ordering of the numerosities 1 to 9 by monkeys. *Science* 282. 746–749.

Brannon, Elizabeth M. & Herbert S. Terrace. 2000. Representation of the numerosities 1-9 by rhesus macaques (Macaca mulatta*). Journal of Experimental Psychology: Animal Behavior Processes* 26. 31–49.

Brannon, Elizabeth M., Courtney J. Wusthoff, C.R. Gallistel & John Gibbon. 2001. Numerical subtraction in the pigeon: Evidence for a linear subjective number scale. *Psychological Science* 12. 238-243.

Bromberger, Sylvain & Morris Halle. 2000. The ontology of phonology (revised). In Noel Burton-Roberts, Philip Carr & Gerald Docherty (eds.), *Phonological knowledge: Conceptual & empirical issues*, 19–37. Oxford: Oxford University Press.

Brosch, Michael, Elena Selezneva, Cornelia Bucks & Henning Scheich. 2004. Macaque monkeys discriminate pitch relationships. *Cognition* 91. 259–272.

Byrne, Richard W. 2007. Clues to the origin of the human mind from primate observational field data. *The Japanese Journal of Animal Psychology* 57. 1-14.

Carey, Susan. 2009. *The origin of concepts*. Oxford: Oxford University Press.

Carstairs-McCarthy, Andrew. 1999. *The origins of complex language: An inquiry into the evolutionary beginnings of sentences, syllables, & truth*. Oxford: Oxford University Press.

ten Cate, Carel, Caroline van Heijningen, & Willem Zuidema. 2010. Reply to Gentner et al.: As simple as possible, but not simpler. *PNAS* 107. E66-E67.

Cheney, Dorothy L. & Robert M. Seyfarth. 1990. *How monkeys see the world: inside the mind of another species*. Chicago: U. of Chicago Press.

Cheney, Dorothy L. & Robert M. Seyfarth. 2007. *Baboon metaphysics*. Chicago: U. of Chicago Press.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: the framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In Adriana Belletti (ed.), *Structures & beyond: The cartography of syntactic structures*, 104–131. Oxford: Oxford University Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 35. 1–22.

Chomsky, Noam. 2007. Approaching UG from below. In Uli Sauerland & Hans-Martin Gärtner (eds.), *Interfaces + recursion = language?*, 1–29. Berlin: Mouton de Gruyter.

Chomsky, Noam. 2008. On phases. In Carlos Otero, Robert Freidin & Maria-Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.

Coen, Michael H. 2006. *Multimodal dynamics: self-supervised learning in perceptual & motor systems*. Cambridge, MA: MIT dissertation.

Conway, Christopher M. & Morten H. Christiansen. 2001. Sequential learning in non-human primates. *Trends in Cognitive Sciences* 5. 539–546.

Dehaene, Stanislas. 1997. *The number sense: how the mind creates mathematics*. Oxford: Oxford University Press.

Devlin, Keith. 2005. *The math instinct*. New York: Thunder's Mouth Press.

Doupe, Allison J. & Patricia K. Kuhl. 1999. Birdsong & human speech: common themes & mechanisms. *Annual Review of Neuroscience* 22. 567–631.

Endress, Ansgar D., Sarah Carden, Elisabetta Versace, & Marc D. Hauser. 2009a. The apes' edge: positional learning in chimpanzees & humans. *Animal Cognition*.

Endress, Ansgar D., Donal Cahill, Stefanie Block, Jeffrey Watumull, & Marc D. Hauser. 2009b. Evidence of an evolutionary precursor to human language affixation in a non-human primate. *Biology Letters* 5. 749-751.

Fehér, Olga, Haibin Wang, Sigal Saar, Partha P. Mitra, & Ofer Tchernichovski. 2009. *De novo* establishment of wild-type song culture in the zebra finch. *Nature* 459. 564-568.

Feigenson, Lisa & Justin Halberda. 2004. Infants chunk object arrays into sets of individuals. *Cognition* 91. 173–190.

Fitch, W. Tecumseh & Marc D. Hauser. 2004. Computational constraints on syntactic processing in a nonhuman primate. *Science* 303. 377–380.

Fitch, W. Tecumseh, Marc D. Hauser & Noam Chomsky. 2005. The evolution of the language faculty: clarifications & implications. *Cognition* 97. 179–210.

Fujita, Koji. 2007. Facing the logical problem of language evolution. Review of Jenkins (2004), *Variation & universals in biolinguistics. English Linguistics* 24. 78–108.

Gallistel, Charles R. 1990. *The organization of learning*. Cambridge, MA: MIT Press.

Gallistel, Charles R. & Rochel Gelman. 2005. Mathematical cognition. In Keith Holyoak & Robert Morrison (eds.), *The Cambridge handbook of thinking & reasoning*, 559-588. Cambridge: Cambridge University Press.

Gambell, Tim & Yang, Charles D. 2005. Word segmentation: quick but not dirty. Ms., Yale University.

Gehring, Walter J. 1998. *Master control genes in development & evolution*. New Haven: Yale University Press.

Gelman, Rochel & Charles R. Gallistel. 1978. *The child's understanding of number*. Cambridge, MA: Harvard University Press.

Gentner, Timothy Q., Kimberly M. Fenn, Daniel Margoliash, & Howard C. Nusbaum. 2006. Recursive syntactic pattern learning by songbirds. *Nature* 440. 1204-1207.

Gentner, Timothy Q., Kimberly M. Fenn, Daniel Margoliash, & Howard C. Nusbaum. 2010. Simple stimuli, simple strategies. *PNAS* 107. E65.

Gil da Costa, Ricardo, Allen Braun, Marco Lopes, Marc D. Hauser, Richard E. Carson, Peter Herscovitch & Alex Martin. 2004. Toward an evolutionary perspective on conceptual representation: Species- specific calls activate visual & affective processing systems in the macaque. *Proceedings of the National Academy of Sciences* 101. 17516–17521.

Goldsmith, John A. 1976. *Autosegmental phonology*. Cambridge, MA: MIT dissertation.

Gould, Stephen J. 1976. In defense of the analog: A commentary to N. Hotton. In R. Bruce Masterson, William Hodos & Harry J. Jerison (eds.), *Evolution, brain & behavior: persistent problems*, 175-179. New York: Wiley.

Greenfield, Patricia M., Karen Nelson & Elliot Saltzmann. 1972. The development of rulebound strategies for manipulating seriated nesting cups: a parallel between action & grammar. *Cognitive Psychology* 3. 291–310.

Hale, Mark & Charles Reiss. 2000a. Phonology as cognition. In Noel Burton-Roberts, Philip Carr & Gerald Docherty (eds.), *Phonological knowledge: conceptual & empirical issues,* 161–184. Oxford: Oxford University Press.

Hale, Mark & Charles Reiss. 2000b. Substance abuse & dysfunctionalism: Current trends in phonology. *Linguistic Inquiry* 31. 157–169.

Hale, Mark & Charles Reiss. 2008. *The phonological enterprise.* Oxford: Oxford University Press.

Hauser, Marc D. 1996. *The evolution of communication*. Cambridge, MA: MIT Press.

Hauser, Marc D. 2000. *Wild minds: what animals really think.* New York: Henry Holt.

Hauser, Marc D. & David D. Glynn. 2009. Can free-ranging rhesus monkeys (Macaca mulatta) extract artificially created rules comprised of natural vocalizations? *Journal of Comparative Psychology* 123. 161-167.

Hauser, Marc D., Susan Carey & Lillian B. Hauser. 2000. Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proceedings of the Royal Society: Biological Sciences* 267. 829–833.

Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002a. The faculty of language: What is it, who has it, and how did it evolve? Science 298:1569–1579.

Hauser, Marc D., Elissa L. Newport & Richard N. Aslin. 2001. Segmentation of the speech stream in a nonhuman primate: Statistical learning in cotton top tamarins. *Cognition* 78. B53–B64.

Hauser, M.D. & Glynn, D. 2009. Can free-ranging rhesus monkeys (Macaca mulatta) extract artificially created rules comprised of natural vocalizations? *Journal of Comparative Psychology* 123: 161-167.

van Heijningen, Caroline A.A., Jos de Visser, Willem Zuidema, & Carel ten Cate. 2009. Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. *PNAS* 106(48). 20538-20543.

Heinz, Jeffrey N. 2007. *Inductive learning of phonotactic patterns*. Los Angeles: University of California dissertation.

Hornstein, Norbert. 2001. *Move! A minimalist theory of construal*. Oxford: Blackwell.

Hurford, James R. 2007. *The origins of meaning: language in the light of evolution.* New York: Oxford University Press.

Ishihara, Shinichiro. 2007. Major phrase, focus intonation, multiple spell-out (MaP, FI, MSO). *The Linguistic Review* 24. 137–167.

Jackendoff, Ray & Steven Pinker. 2005. The nature of the language faculty & its implications for evolution of language (reply to Fitch, Hauser, & Chomsky). *Cognition* 97. 211–225.

Johnson, C. Douglas. 1970. *Formal aspects of phonological representation*. Berkeley, CA: University of California dissertation.

Johnson-Pynn, Julie, Dorothy M. Fragaszy, Elizabeth M. Hirsh, Karen E. Brakke & Patricia M. Greenfield. 1999. Strategies used to combine seriated cups by chimpanzees (Pan troglodytes), bonobos (Pan paniscus), and capuchins (Cebus apella). *Journal of Comparative Psychology* 113. 137–148.

Kahnemuyipour, Arsalan. 2004. *The syntax of sentential stress*. Toronto: University of Toronto dissertation.

Kitahara, Hisatsugu. 2003. Acoustic perception in human newborns & cotton-top tamarins. *Reports of the Keio Institute of Cultural & Linguistic Studies* 35. 31-41.

Kuhl, Patricia K., & James D. Miller. 1975. Speech perception by the chinchilla. *Science* 190. 69–72.

Lakoff, George & Rafael Nuñez. 2001. Where mathematics comes from: how the embodied mind brings mathematics into being . New York: Basic Books.

Leben, William. 1973. *Suprasegmental phonology*. Ph.D. dissertation, MIT.

Lepsius, Richard & William Dwight Whitney. 1865. On the relation of vowels & consonants. *Journal of the American Oriental Society* 8. 357–373.

Levin, Juliette. 1985. *A metrical theory of syllabicity*. Cambridge, MA: MIT dissertation.

Mailhot, Frédéric & Charles Reiss. 2007. Computing long-distance dependencies in vowel harmony. *Biolinguistics* 1. 28–48.

Marcus, Gary F., S. Vijayan, S. Bandi Rao, & P. Vishton. 1999. Rule learning by seven-month-old infants. *Science* 283. 77–80.

Marler, Peter. 1977. The structure of animal communication sounds. In T.H. Bullock (ed.), *Recognition of complex acoustic signals,* 17-35. Berlin: Dahlem Konferenzen.

Merker, Björn. 2000. Synchronous chorussing & human origins. In Nils Wallin, Björn Merker & Steven Brown (eds.), *The origin of music*, 315-328. Cambridge, MA: MIT Press.

Murphy, Robin A., Esther Mondragon, & Victoria A. Murphy. 2008. Rule learning by rats. *Science* 319. 1849-1851.

Neeleman, Ad & Hans van de Koot. 2006. On syntactic & phonological representations. *Lingua* 116. 1524–1552.

Newport, Elissa L. & Richard N. Aslin. 2004. Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology* 48. 127–162.

Newport, Elissa L., Marc D. Hauser, Geertrui Spaepen & Richard N. Aslin. 2004. Learning at a distance II. Statistical learning of non-adjacent dependencies in a non- human primate. *Cognitive Psychology* 49. 85–117.

Ohala, John J. 1992. Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In *Papers from the CLS Parasession on the Syllable*, 319–338.

Ohala, John J., & Haruko Kawasaki-Fukumori. 1997. Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In Stig Eliasson & Ernst H. Jahr (eds.), *Language & its ecology: Essays in memory of Einar Haugen*, 343–365. Berlin: Mouton de Gruyter.

Ohshiba, Nobuaki. 1997. Memorization of serial items by Japanese monkeys, a chimpanzee, & humans. *Japanese Psychological Research* 39. 236–252.

Ouattara, Karim, Alban Lemasson, & Klaus Zuberbühler. 2009. Campbell's monkeys concatenate vocalizations into context-specific call sequences. *PNAS* 106(51). 22026-

22031.

Patel, Aniruddh. 2008. *Language & music*. Oxford: Oxford University Press.

Patel, Aniruddh & John R. Iversen. 2006. A non-human animal can drum a steady beat on a musical instrument. In *Proceedings of the 9th International Conference on Music Perception & Cognition*, 477.

Payne, Katherine. 2000. The progressively changing songs of humpback whales: a window on the creative process in a wild animal. In Nils Wallin, Björn Merker & Steven Brown (eds.), *The origins of music*, 135–150. Cambridge, MA: MIT Press.

Pepperberg, Irene M. 1999. *The Alex studies: cognitive & communicative abilities of grey parrots*. Cambridge, MA: Harvard University Press.

Piattelli-Palmarini, Massimo. 1989. Evolution, selection, & cognition: from 'learning' to parameter setting in biology & in the study of language. *Cognition* 31. 1–44.

Pinker, Steven & Ray Jackendoff. 2005. The faculty of language: what's special about it? *Cognition* 95. 201–236.

Raimy, Eric. 2000. *The phonology & morphology of reduplication*. Berlin: Mouton de Gruyter.

Raimy, Eric. 2003. Asymmetry & linearization in phonology. In Anna Maria Di Sciullo (ed.), *Asymmetry in grammar*, *vol. 2*, 129–146. Amsterdam: John Benjamins.

Ramus, Franck, Marc D. Hauser, Cory Miller, Dylan Morris & Jacques Mehler. 2000. Language discrimination by human newborns & by cotton-top tamarin monkeys. *Science* 288. 349-351.

Redford, Melissa A. 1999. *An articulatory basis for the syllable*. Austin, TX: University of Texas dissertation.

Redford, Melissa A., Chun Chi Chen & Risto Miikkulainen. 2001. Constrained emergence of universals & variation in syllable systems. *Language & Speech* 44. 27–56.

Reiss, Charles. The OCP and NOBANANA. 2008. In Bert Vaux & Andrew I. Nevins (eds.), *Rules, constraints & phonological phenomena*, 252-301. Oxford: Oxford University Press.

Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, MA: MIT Press.

Saltzman, I.J. & W.R. Garner. 1948. Reaction time as a measure of span of attention. *Journal of Psychology* 25. 227–241.

Samuels, Bridget. 2009a. *The structure of phonological theory*. Cambridge, MA: Harvard University dissertation.

Samuels, Bridget. 2009b. The third factor in phonology. *Biolinguistics* 3(2-3). 355-382.

Samuels, Bridget & Cedric Boeckx. 2009. What emerges from Merge in phonology? Paper presented at OCP 6, Edinburgh.

Saussure, Ferdinand de. 1916. *Cours de linguistique general*. Paris: Payot.

Schusterman, Ronald J. & David Kastak. 1993. A California sea lion (Zalophus californianus) is capable of forming equivalence relations *The Psychological Record* 43. 823–839.

Senghas, Richard J., Ann Senghas & Jennie E. Pyers. 2005. The emergence of Nicaraguan Sign Language: Questions of development, acquisition, and evolution. In J. Langer, S. T. Parker, & C. Milbrath (eds.), *Biology & knowledge revisited: From neurogenesis to psychogenesis*. Mahwah, NJ: Lawrence Erlbaum Associates. 287-306.

Seyfarth, Robert M., Dorothy L. Cheney & Thore J. Bergman. 2005. Primate social cognition & the origins of language. *Trends in Cognitive Sciences* 9. 264– 266.

Shettleworth, Sara J. 2010. *Cognition, evolution, & behavior, second edition*. Oxford: Oxford University Press.

Slater, Peter J.B. 2000. Birdsong repertoires: their origins & use. In Nils Wallin, Björn Merker & Steven Brown (eds.), *The origins of music*, 49–64. Cambridge, MA: MIT Press.

Spelke, Elizabeth. 2003. What makes us smart? Core knowledge & natural language. In Dedre Gentner & Susan Goldin-Meadow (eds.), *Language in mind: advances in the study of language & thought*, 277-311. Cambridge, MA: MIT Press.

Taylor, Ryan C., Barrett A. Klein, Joey Stein & Michael J. Ryan. 2008. Faux frogs: Multimodal signalling & the value of robotics in animal behavior. *Animal Behaviour* 76. 1089–1097.

Tincoff, Ruth, Marc D. Hauser, Fritz Tsao, Geertrui Spaepen, Franck Ramus, & Jacques Mehler. 2005. The role of speech rhythm in language discrimination: further tests with a non-human primate. *Developmental Science* 8. 26-35.

Toro, José M., Josep B. Trobalón & Núria Sebastián-Galles. 2005. The effects of backward speech & speaker variability in language discrimination by rats. Journal of Experimental Psychology: Animal Behavior Processes 31. 95–100.

Uriagereka, Juan. 1998. *Rhyme & reason*. Cambridge, MA: MIT Press.

Walter, Mary Ann. 2007. *Repetition avoidance in human language*. Cambridge, MA: MIT dissertation.

Werker, Janet F. & Athena Voloumanos. 2000. Language: Who's got rhythm? *Science* 288. 280-281.

Williams, Heather & Kirsten Staples. 1992. Syllable chunking in zebra finch (Taeniopygia guttata) song. *Journal of Comparative Psychology* 106. 278–286.

Wright, Anthony A., Jacquelyne Rivera, Sara H. Hulse, Melissa Shyan, & Julie Neiworth. 2000. Music perception & octave generalization in rhesus monkeys. *Journal of Experimental Psychology: General* 129. 291–307.

Yip, Moira. 2006a. Is there such a thing as animal phonology? In *Wondering at the natural fecundity of things: studies in honor of Alan Prince*, http://repositories.cdlib.org/lrc/prince/15.   (8 October 2008.)

Yip, Moira. 2006b. The search for phonology in other species. *Trends in Cognitive Sciences* 10. 442–445.

Zuberbühler, Klaus. 2006. Language evolution: The origin of meaning in primates. *Current Biology* 16. R123-R125.

Zuberbühler, Klaus. 2005. The phylogenetic roots of language: evidence from primate communication and cognition. *Current Directions in Psychological Science* 14. 126-130.