Title
Is There Such a Thing as Animal Phonology?

Permalink
https://escholarship.org/uc/item/5f90t5p7

Author
Yip, Moira

Publication Date
2006-06-20

Is there such a thing as animal phonology?
Moira Yip
University College London

The issue of whether language is the result of mechanisms that are specifically human, and specific to language, has been publicly discussed in a recent series of papers in *Science* and in *Cognition*. Hauser, Chomsky and Fitch in their original 2002 paper (henceforth HCF) and again in Fitch, Hauser and Chomsky (2005) (henceforth FHC) argue that recursion is the only mechanism that qualifies on both counts, and they call this the Faculty of Language in the Narrow sense (FLN), in contrast to the Faculty of Language in the Broad sense (FLB). HCF spend little time on phonology, but they define it as a "mapping from narrow syntax to the SM interface". In the later 2005 paper, they say (p.200): "much of phonology is likely part of FLB, not FLN, either because phonological mechanisms are shared with other cognitive domains (notably music and dance), or because the relevant phenomena appear in other species, particularly bird and whale "song"." They further speculate that three aspects of phonology might be in FLN: recursion (if present at all, likely to be at the level of the phonological phrase or above), mapping to the SM interface, and some unknown residue, which must be 'tested for, not assumed'. Their default position appears to be that this residue is a null set. They also take a strong view of what it takes to be included in FLN. Any similarities whatsoever to mechanisms needed in non-language spheres, such as music or dance, or to non-human spheres, such as birdsong, are sufficient to disqualify something from FLN. Rhythm, "clearly shared between language and music", is thus excluded, even if its computation in language calls on additional mechanisms. They also suggest that any signs of recursion in phonology could be just a reflex of recursion in syntax, and might also be the same as recursion in music. Under their view, then, phonology contains little or nothing that belongs in FLN.

Pinker and Jackendoff in two 2005 papers disagree, arguing that there are aspects of speech and language other than recursion that must be included in FLN. The original arguments of Hauser, Chomsky and Fitch barely touched on phonology (as opposed to speech production and perception), but Pinker and Jackendoff (2005; henceforth PJ) bring it into the discussion. They argue that although phonology is not recursive, it does display the property of discrete infinity, something which, when it occurs in syntax, HCF call the "hallmark of language". The suggestion then is that there are some aspects of phonology other than recursion that must be in FLN. They also take a less restrictive view of the qualifications for inclusion. If something has non-language analogues, perhaps in music, and then language has adapted and changed this for specialized use, PJ would consider it to form part of FLN. So Jackendoff and Pinker (2005: 215; henceforth JP) would allow into FLN "[a]ny subsystem that can be shown to have been adapted for language from some evolutionary precursor".

This debate raises a host of interesting questions, and I shall not attempt to deal with them all here. My goal in this squib is to ponder on what it is, as a phonologist, I would take to be core properties of phonological systems, and then ask which, if any, are known to be found outside humans, and which are known to be absent in at least some non-human animals. Not surprisingly, it turns out we have very few answers to these questions, so I then turn to how we might go about designing experiments to test

for phonological skills. There I stop, in the hope that someone might find these ideas sufficiently interesting to actually carry out these or similar experiments, and give us some answers.

There are two things I shall not discuss. First, I will not look systematically at speech production and speech perception. Second, I will not investigate the question of to what extent the skills needed for phonology are language-specific in humans, as opposed to also being found in other domains, such as music.

1. Some interesting phonological skills

It is useful to divide our phonological knowledge into matters relating to the sounds themselves, and to the structures in which they occur. I start with the sounds.

1.1 Sounds

Early in acquisition we acquire the **distinctions** found in our ambient language, and our ability to distinguish other sounds fades away. We also exhibit an awareness of prototypes for these linguistic categories, in what Kuhl (1991) calls the **perceptual magnet effect**. At the completion of this process we have acquired the **inventory of sounds** of our language.

With the sounds in place, phonology proper begins to emerge. One of its central organizing principles is of course that sounds are grouped into **natural classes,** and phonological processes act on these classes, in environments that are also framed in terms of these classes.

Using these classes of sounds as raw material, we also learn a set of **regular patterns of distribution,** and of **alternations** that are **systematic, governed by laws** that may, depending on one's theory of phonology, be stated in terms of rules, parameters, or constraints.

Among the many mechanisms needed to explain these processes is the ability to **compute identity**, whether total or partial, and over non-adjacent as well as adjacent sounds. The Obligatory Contour Principle (Leben 1973, McCarthy 1986, Yip 1988, Frisch and Zawaydeh 2001, Frisch, Pierrehumbert and Broe 2004), for example, requires this power to ban identity, and so, in a rather different way, does vowel harmony, to enforce it.

1.2 Structures

We know that sounds group themselves into larger units, including **syllable**, **foot**, and **prosodic word**, and that phonological processes may pay attention to these units in a wide variety of ways, including but not limited to (i) phonotactics and (ii) as environments for rules that apply within them but not across them, or at their edges only.

Although not directly related to phonological structures, an important part of our phonological knowledge seems to be based on an ability to pay attention to **statistical distributions**. Recent work has shown this to be true of adults (Hay 2001, Zuraw 2002, Pierrehumbert 2003, and many others). It is also apparently one of the tools we use in acquisition: Saffron et al. (1996) have shown that infants are capable of extracting 'words' from the speech stream purely on the basis of **statistical probability**.

This ability to make use of transitional probabilities extends in human infants **to non-adjacent consonants,** or **non-adjacent vowels** (Newport and Aslin 2004), and

again this is used in the phonology and morphology of languages such as the Semitic languages.

Ramus et al. (2000) have shown that infants can distinguish between **two prosodically distinct languages** even when the segmental content is drastically simplified (so that all stops are [t], all fricatives are [s], all nasals are [n], all glides are [j] and all vowels are [a]). One might suppose that this leaves only pitch and duration cues different between the two languages, but this is a slight over-statement, since manner of articulation differences were also retained, meaning that overall C/V ratios remained distinguishable (and thus the distinction between the CV syllables of Japanese and the much more complex clusters of Dutch), as did stop/fricative/nasal ratios among consonants. Still, if by 'prosodically distinct' we mean languages that differ in syllable structure, foot structure, and intonation structure, it does seem that this is sufficient information for infants to tell them apart.

Having identified some of our key phonological abilities, I now turn to the question of whether they are specific to humans or also found in other animals.

2. Are these abilities found in other animals?

We need to distinguish between evidence that animals have analogues of these abilities in their own communication systems, and evidence that they are able to recognize the sounds and structures of human language. Within their own systems, there has been almost no investigation of what I would call phonology. There has been work looking at internal structure, but much of it has investigated whether an animal call is decomposable into semantically discrete units, or 'words', rather than work on the structure of the sounds themselves. The main exceptions to this have come in work on those species that have 'songs', such as birds, whales, crickets, and gibbons.

As far as their abilities upon exposure to human language are concerned, the bulk of the recent careful work has been done with cotton-top tamarins, a small primate easily studied in the lab. Unlike earlier work, much of this work does not use operant conditioning, but simply surrounds the animals with an ambient 'language', then tests for whether they react to deviations from this. One advantage of this technique is that it can be replicated in human adults and human infants, allowing for a more direct comparison between human and non-human abilities.

2.1 Sounds

If we start with animals' own communications, do they have the equivalents of our phoneme inventories? Starlings seem to have a rough equivalent. They have four song 'types' (also known as motifs): whistle, variable, rattle, and high-frequency; each is generally less than one second long. A motif is an acoustically complex event composed of several notes in a highly stereotyped pattern, where a note is 'classified by the presence of continuous energy in the spectrogram' (Gentner and Hulse 2000). The notes are first grouped into what Doupe and Kuhl (1999) call syllables, defined as 'units of sounds separated by silent intervals'. These in turn are grouped into motifs, which Doupe and Kuhl identify with the phrase, but a phonologist might be more inclined to call a prosodic word. Among starlings, each bird has 11-90 of these motifs, adding new ones throughout life. They are unique to each individual, and they are used to identify the individual. They also have a significant effect on a bird's ability to recognize another individual's song.(Gentner and Hulse 1998). Drawing parallels

between the building blocks of a starling song and the building blocks of human phonology is not an exact science, but. Doupe and Kuhl suggest that the 'notes of the songs are analogous to the 'smallest units of speech, or phonetic units'. I return to this below.

If we turn to how animals process human language, we find a very confused picture. The process of acquiring a sound inventory clearly rests on the phenomenon of categorical perception. When exposed to synthetic tokens of human speech, this is something that other animals such as chinchillas are known to be capable of (Kuhl and Miller 1975). Like humans from different first-language environments, they differ in where they place the category boundaries. Fitch, Hauser and Chomsky 2005 (henceforth FHC) argue persuasively that so long as the inter-species difference between animals and humans is not greater than the intra-species variation observed in different human language groups, we cannot conclude that the animals' abilities differ from ours. By this criterion, then, categorical perception is not specific to humans.

Turning to the perceptual magnet effect, Kuhl (1991) argues that rhesus monkeys do not show the perceptual magnet effect found in human infants. However, subsequent work on starlings by Kluender, Lotto, Holt and Bloedel (1998) finds remarkable evidence of the effect in starlings. Unlike Kuhl, who worked with monkeys with no experience of human speech sounds, they exposed and trained the starlings by giving them 'experience with approximations to natural distributions of speech sounds.' The starlings then successfully generalized to novel instances of these distributions. What is more, their judgements of goodness of fit as measured by response strength (peck rate) were extremely close to human goodness judgements. They conclude that rather general learning processes may be responsible for the perceptual magnet effect, and that 'constructs such as prototypes are unnecessary.'

Returning to primates, Ramus et al. (2000) replicated their work on infants with cotton-top tamarins, who turn out to be able to be able to distinguish between Dutch and Japanese sentences, like human infants. However, if segment differences are drastically reduced so that all stops are [t], all fricatives are [s], all nasals are [n], all glides are [j] and all vowels are [a], with the result that the remaining differences between the two languages are reduced to F0, segment duration, and the relative frequencies of stops, fricatives, nasals, glides and vowels, the tamarins' performance (still quite good) deteriorates somewhat (whereas for human infants it remains robust). One interpretation of this result is that they are paying more attention to the segments and less to the prosody, while human infants do the reverse. At a minimum, it seems necessary to conclude that even in running speech they can detect segmental information. I return to this interesting paper again below.

I am not aware of any work looking at evidence of natural classes, systematic, lawful phonological processes, or identity computations in non-human animals. One interpretation of the tamarin data in the previous paragraph is that they have no notion of natural classes such as 'stop' or 'fricative', and thus for them the deformed signal bears no systematic relation to the original signals. I return to this in section 3.1.

2.2   Structures
I start with the ability of animals to detect transitional probabilities. Work by Hauser, Newport and Aslin (2001) has shown that cotton-top tamarins have remarkable abilities to do this, both locally and non-locally.

The general methodology of the experiments is roughly the same for all the work cited here. I describe here the work of Hauser et al. (2001). The tamarins were exposed to a language composed of 12 different syllables grouped into 4 trisyllabic 'words'. The words followed each other at random. The transitional probability for syllables within a word was 1.0, and for syllables at word boundaries .33. The materials were synthesized, so there were no prosodic cues to word boundaries, creating a 1-min speech stream which was then looped to form a 21-minute speech exposure stream, played on day one, followed by 2-min re-exposure on day two. They were then tested on 3-syllable items that were either 'words', or 'part-words', with the latter being 3-syllable sequences that spanned 'word' boundaries. The dependent measure was an orienting response to the loud speaker. The tamarins were significantly more likely to orient to part-words (unfamiliar) than to words (familiar). In the later experiments testing non-adjacent transitional probabilities, the basic methodology was the same, but the stimuli were different. The first and third syllables of the 'words' were held constant, and the medial syllables varied. So for example all words were one of *di_tae, po_ga,* or *ke_bu*, but the second syllables could be *ki* or *gu*.

So what can the tamarins do? Although they can do some of these tasks, they show interesting differences from human infants. Like infants, they can track transitional probabilities for adjacent syllables (Hauser et al. 2001). Also like infants, they can do this for non-adjacent probabilities involving vowels (Newport, Hauser, Spaepen and Aslin 2004). After that, things diverge: infants can do this for non-adjacent consonants, but not for non-adjacent syllables, whereas tamarins can do it for non-adjacent syllables but not for consonants. At this stage we can only speculate as to why. As Newport and Aslin note, in human natural language non-adjacent consonants and vowels play roles in the morphophonology of Semitic and other languages, but non-adjacent syllables are rare. Grammars thus apparently use skills that infants have, and do not use things they are bad at. Importantly, the infants in the experiments were not previously exposed to human languages of the Semitic type, so these abilities cannot be attributed to prior experience. But why should tamarins, who apparently have difficulty with non-adjacent consonants, be able to master non-adjacent syllables? Newport et al (p.109) speculate that the consonant difficulties arise because the acoustic properties of consonants make them harder to perceive and process in running speech. But how can they recognize syllables if consonants are hard? It may be that the particular syllable task can be done without accurate consonant recognition. No syllables differed only in voicing, for example. So vowel recognition and place recognition suffice to identify the syllable, both of which show in formant structure.

What about natural animal communications? Do they show evidence of structures of any kind? The song motifs of the starling occur in sequences about 30secs long, called song-bouts. Within each bout, the transitional probabilities of the song motifs are non-random. They appear to be distinctive for a particular bird, and have a significant effect on a bird's ability to recognize another individual's song.(Gentner and Hulse 1998). If synthetic stimuli are made by splicing, and if the second-order probabilities are maintained the effect is much stronger than a random ordering, with significance P<.05. Third-order probabilities didn't improve things further. It is not at all clear to what aspects of human language, if any, these facts are analogous. Doupe and Kuhl suggest (see section 2.2) that the structure of passerine birdsongs begins with a note, which is the equivalent of a phoneme. A sequence of notes is a syllable, and a

sequence of syllables (a motif) is like a phrase, but for a phonologist this appears to skip the level of organization we call the prosodic word. So perhaps the motif is the prosodic word, and a high-probability sequence of two motifs is roughly like a phrase? The duration of a motif is long compared to a human prosodic word (up to one second long, whereas a typical American English syllable is around 170 ms., with a range from 107-260 ms (Greenberg et al. 1996). The analogy may be imprecise, but the general point that the song can be broken down into smaller internal constituents seems incontestable. Below I diagram the structure of a zebra finch song from Doupe and Kuhl (1999: 571, Fig 2B). It is composed of 4 syllables, labeled a-d, each of which contains from 1-3 notes. The syllables are organized into motifs (my prosodic words) which are repeated:

(1)   song

           motif

           syllable

           notes
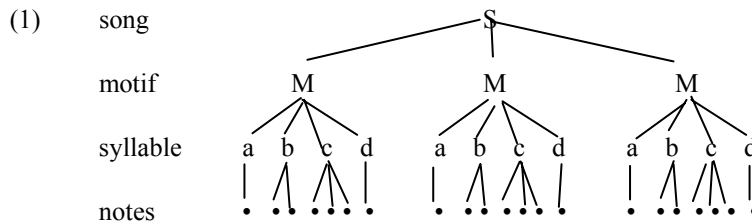
Apart from the obvious fact that the precise parallels are complete guesswork, this differs from human language in other ways. Most strikingly, there is no evidence for any sort of binarity.

In recent work available to me only as a conference poster, Gentner et al. (2005) have argued that starlings can learn new rules for the temporal patterning of song motifs, which is not terribly surprising given that they apparently use temporal patterning to recognize other individuals. More strikingly, they suggest that they can learn center-embedded recursive syntactic structures produced by a context-free grammar, A^nB^n. But even if this is true, is it analogous to syntax or to phonology? The problem is one of interpretation. If a pattern is ABABAB, is this analogous to the CVCVCV patterning required by the phonology of some languages, or is it analogous to the NP VP NP VP pattern imposed by the syntax of some languages? If it is like phonology, recursivity would be surprising since it is apparently not found even in human phonology. If it is like syntax, then it would suggest that even outside humans we find recursiveness in syntax.

Let us grant that the tamarin and starling data suggest some ability to detect structure in both natural birdsong and in human language stimuli. What about using this structure in phonological processes or rules? Here we know nothing at all, and it is not even clear how to test this, a point to which I will return in section 3.2

I now turn to rhythm. To a phonologist, the phrase 'rhythmically distinct languages' would usually suggest languages whose stress systems differed. The little we know about animals' ability to pay attention to rhythm tests for something rather different that is roughly the traditional stress-timed versus syllable-timed distinction, using a particular view of this distinction taken from Ramus et al. (1999).

It is widely reported that Ramus et al. (2000) (discussed earlier) have shown that tamarins, like infants, can distinguish between two prosodically distinct languages like Dutch and Japanese. However, a careful reading of their paper suggests that this is an over-interpretation of their results by others. I have already mentioned that unlike

infants**,** when the segmental content is drastically simplified so that only manner contrasts are retained, leaving little except pitch and duration cues, their performance deteriorates. Surprisingly, Toro et al. (2003) replicated this experiment on rats, and found that unlike the tamarins they did particularly well on the re-synthesized stimuli! In subsequent work, Tincoff et al. (2005) have tested for whether tamarins can distinguish between two rhythmically similar languages, Dutch and English, and between two rhythmically distinct languages, Polish and Japanese, where the criteria for rhythmic similarity are those of Ramus, Nespor and Mehler (1999). They find robust discrimination between Polish and Japanese, and no discrimination between English and Dutch. They used only natural speech, which makes it hard to be certain of the basis for discrimination. For example, the phoneme inventories of Japanese and Polish are highly divergent, whereas those of Dutch and English are more overlapping. Even for human adults, very little work has really tried to test discrimination if *all* information except the consonant and vowel durations that provide rhythmic information are removed. Ramus and Mehler (1999) found that English/Japanese could still be discriminated by adult humans even if F0 was held constant, all vowels were reduced to [a] and all consonants to [s], but as far as I can tell no-one has replicated this work with animals.

What about rhythm in natural animal communication? Of course, birdsongs have a clear musical rhythm, and it would seem reasonable to say that if we see similarities between human language and human music, we must also recognize similarities between human language and animal 'music'. What about outside the avian world? At both extremes of the animal kingdom we can find rhythmic vocalizations: in the great apes, in cetaceans, and in crickets.

A rather simple song is that of the cricket, Acheta domesticus (Stout et al. 1983). (The cricket literature is extensive, and there are many different species. See Olvido and Wagner 2004 for a different view.) The male cricket has a calling song of 'syllables' separated by silences. Females find this most attractive if each sequence is 3 syllables long, rather than 2 or 4, with 1 or 5 still less attractive. The preferred syllable period is 50ms (modal for the natural call at 22C), and the preferred syllable duration is 25ms when the period is 50ms. As the period increases, so does the preferred syllable duration, staying at roughly half the period. So the ideal structure is [(x.)(x.)(x.)], where beat and interval have an equal duration. Note that this entire sequence is within the range of the duration of a human syllable, or 150ms. Absolute duration measures do not seem to be a very useful point of comparison, however, especially when the species in question are as widely distinct as crickets and humans. The auditory and neural systems of the two species are likely to be so divergent that they react to quite different acoustic events. So all we can conclude, I think, is that there is some evidence here for a preference for a preferred rhythmic structure in cricket calls.

In the primates, the clearest example of rhythmic vocalizations are the calls of the gibbons (Geissman 2000). The gibbons produce song bouts lasting around 30 minutes. They are stereotyped and species-specific. Mated pairs also sing duets. The most noticeable aspect of the female duet part is the great call, 'a particularly rhythmic series of long notes uttered with increasing tempo and/or increasing peak frequency'. The male responds with a special reply phrase called a coda. Some species show much more complex duets. The function of these songs is little understood, but it may be

related to strengthening the pair bond. However, Geissman (p.119) points out a major difference between primate vocalizations and human music: 'non-human primates do not seem to be able to keep a steady pulse in their song vocalizations'. Instead, they typically speed up during the call. They are thus not rhythmic in the sense that human language is rhythmic, and it seems likely that the illusion of rhythm is more related to breathing patterns than to cognitive organization. This is clear in the bi-phasic notes of some Old world monkeys and apes, which alternate inhaled and exhaled notes.

We are left with a fragmented set of observations. Some animals show evidence of:

- categorical perception
- perceptual magnet effect
- ability to pay attention to transitional probabilities, including non-adjacent ones
- possible sensitivity to prosody, roughly of the syllable-timed vs. foot-timed variety
- some internal structure in their own systems
- some rhythmic pattering in their own systems

But this leaves unexplored a range of other skills that contribute to human phonology.

3.   How can we test these skills further?

The picture I take away from my survey of the recent literature is that we are only beginning to look in non-humans (primates or others) for abilities that verge on what I would call phonology. I have found no literature that looks at such things as:

- an awareness of natural classes of sounds
- the ability to discriminate trochaic from iambic rhythm, or initial from final stress
- the ability to recognize 'words' even if altered by systematic rule-governed alternations not reducible to purely phonetic dictates
- the ability to perform identity calculations

In this final section I suggest questions we might ask, and outline the types of experimental work that might begin to address these issues. They are all focused on asking to what extent animals evidence such abilities when confronted with human language, but of course ideally one would like to search in their natural communications as well. This task seems to me to be beyond our current reach, hence I do not attempt it. One major admission is in order: I am not an experimentalist, and these ideas are sketchy, and might need considerable alteration to be experimentally viable. A final point is that most of the ideas below have not been tried on humans (adults or infants), so in each case the work would need to be carried out on both populations if any human-animal comparison is to be made.

3.1   Natural classes

I mentioned earlier that tamarins perform less well in discriminating Dutch vs. Japanese when segmental contrasts are reduced such that only manner of articulation is preserved. Notice that this particular re-synthesis preserves the natural classes of stop, fricative, nasal etc, and that this information may be used by the human infant in discriminating between the two languages. If tamarins find this harder, could this be interpreted as a lack of natural classes for manner of articulation, such as stop versus

fricative? To test this, we need to contrast performance on a re-synthesized signal that pays attention to natural classes (like the stimuli created by Ramus et al. (2000)), with performance on stimuli which ignore the natural classes, randomly reducing consonants to {t,s,n,l}and deliberately changing their proportions in the output. First we need to test human infants: do they do worse if the segmental reductions are done randomly, not by natural class? If so, human infants use natural classes. Then what about tamarins: we need to take those who do well on the transformed speech that preserves natural classes, and try them on this new transform. The sort of stimuli I envisage might look like this, starting from a CV language:

(2)   Real speech: saki toza nashi loda lapo
        Old transform, preserving natural classes: :   sata tasa nasa lata lata
        New transform, ignoring natural classes: tasa sata tata sasa sata

I can think of a second way to address these questions, based loosely on techniques with human artificial language-learning by Peperkamp and Dupoux (to appear). Suppose we create a pseudo-language composed of both stops and fricatives, and in which all 'words' start with stops, but in which there happens to be a gap in that the language includes no [p]-initial words. We habituate the tamarins to this language, then introduce two new languages: one with [p]-initial words added, and one with [f]-initial words added. If natural classes play a role, the addition of more stop-initial words might go undetected, since all words still begin with stops. The addition of [f]-initial words, however, should not go un-noticed, since it creates an entirely new type of word, beginning with a fricative. A simpler version of this experiment might not pay attention to word position, but just add new sounds to the phoneme inventory of the languages. Do infants and tamarins react differently if the new sound comes from a new natural class?

3.2   Discriminating stress systems

Can animals truly discriminate a pure prominence difference, with syllable structure held constant? Two experiments come to mind. First, we know that tamarins can use transitional probabilities to extract 'words' from the speech stream in an artificial language. Suppose that we expose them to a language in which each of these 'words' carries initial stress. Then we create an otherwise identical language, containing the same 'words', but in which the words carry final stress? Will they discriminate?

Second, we could use natural language stimuli from two CV languages with different stress patterns, such as Warao and Weri, and replicate the work of Tincoff et al. (2005) to see if they can distinguish them reliably. As a control, we could use another pair of CV languages that have the same stress pattern as each other.

3.3   Allophonic rules

Can animals learn that two different sounds are allophones of a single phoneme, each appearing predictably in different environments? The technique that seems promising here is based on one used by Peperkamp and Dupoux (to appear) for adults. Subjects were exposed to an artificial language in which either stop voicing or fricative voicing was allophonic, with the voiced variant occurring intervocalically. The task they used involved picture recognition, and thus cannot be easily adapted for tamarins, so instead one could expose them to the artificial language with its allophonic voicing,

then use a test phase in which the allophonic voicing pattern is violated, with voiceless sounds appearing intervocalically and/or voiced ones appearing in other contexts. Rather surprisingly, in adult humans, this pattern was only learnt easily if the words were linked to semantic information. Otherwise, when the evidence was purely distributional it was only learnt for the actual lexical items in the corpus and not generalized. The goals of any animal experiment would thus have to be modest: do they learn this pattern for the specific 'words' in the exposure phase, and do they show surprise if those same 'words' deviate from this pattern in the test phase.

3.4   Identity computation

I start with how to discover whether animals can detect identity prohibitions. Suppose we expose them to a language in which all the 'words' observe the OCP-Place. After habituation to this, the test phase then throws in words that violate it (the asterisked cells). Do they show surprise? All the new syllables must already occur in the habituation data, so if a new OCP-violating word is [pabi], [pa] and [bi] must be familiar from words like [padi] and [tabi]. The chart below shows the stimuli for the exposure phase. The test phase would include items from the asterisked cells.

(3)

	CVpV	CVbV	CVtV	CVdV	CVkV	CVgV
pCVC	*	*	pata	pada	paka	paga
bCVC	*	*	bata	bada	paka	baga
tCVC	tapa	taba	*	*	taka	taga
dCVC	dapa	daba	*	*	daka	daga
kCVC	kapa	kaba	kata	kada	*	*
gCVC	gapa	gaba	gata	gada	*	*

In this form, the experiment requires them to recognize p/b as a class. Less ambitiously, one could instead look at the long-distance identity issue by using only voiceless sounds, so that the new stimuli would be papi, tita, kika etc. One cautionary note: if tamarins can't detect non-adjacent consonant transitional probabilities (see above), the task I have just outlined might be beyond them. In that case one would need to use a task where the OCP facts involved vowels, not consonants.

Turning now to languages in which identity is obligatory, one could create an artificial vowel harmony language in which all 'words' have front unrounded vowels i/e or back rounded u/o. After exposure to this harmonic 'language', one could then include test stimuli which mix these two sets, and see if they show surprise at these non-harmonic 'words'. It would also be possible to do this experiment with data from a natural vowel harmony language, like Turkish (eliminating all exceptional non-harmonic words!), and then expose them to a resynthesized pseudo-Turkish in which some percentage of the words are now disharmonic.

4.   Conclusions

This highly speculative squib set out to search for evidence of phonological abilities in non-humans. Some truly interesting experimental work is now being done, but as usual we have far more questions than answers. In section 3 I have outlined ideas for some possible experiments that would keep a large population of tamarins busy for several generations, but it is my hope that some experimental labs may none the less pick up some of these fledgling suggestions and use their expertise to design carefully controlled experiments that address these issues.

Recently a student said to me: "The trouble with Optimality Theory is that it is a theory of anything, not just phonology, and maybe not just humans." I became fascinated by the subtext that any satisfactory theory of phonology *must* be *both* specific to humans *and* specific to phonology. I do not share this common (but rarely voiced) view. It is an empirical question. If the theory that best explains our phonological knowledge also turns out to explain other types of knowledge in humans or other animals, that does not make it wrong. But we are not even close to answering the question at present (at least in so far as it makes claims about humans vs. other animals) because we do not know enough about the abilities of other animals. This squib is an attempt to encourage us to ask these questions.

References

Doupe, A. and P. Kuhl. 1999. Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience* 22, 567-631.

Fitch, W.T., M. Hauser and N. Chomsky. 2005. The evolution of the language faculty: Clarifications and implications. *Cognition* 97, 179-210.

Frisch, S. and B. Zawaydeh. 2001. The psychological reality of OCP-Place in Arabic. *Language* 77.1, 91-106.

Frisch, S., J. Pierrehumbert and M. Broe. 2004. Similarity avoidance and the OCP. *Natural Language and Linguistic Theory* 22, 179-228.

Geissman, T. 2000. Gibbon song and human music from an evolutionary perspective. In N. Wallin, B. Merker and S. Brown (eds.), *The Origins of Music*. 103-123. Cambridge, MA: MIT Press.

Gentner, T., K. Fenn, and D. Margoliash. 2005. Center-embedded syntactic pattern learning by songbirds. Poster presentation given at BU Conference on Language Acquisition.

Gentner, J.Q. and S.H. Hulse 1998. Perceptual mechanisms for individual vocal recognition in European starlings, Sturnus vulgaris. *Animal Behaviour* 56, 579-594.

Gentner, J,Q. And S.H. Hulse 2000. Perceptual classification based on the component structure of song in European starlings. *JASA* 107.6, 3369-3381.

Greenberg, S. J. Hollenback, D Ellis. 1996. Insights into spoken language gleaned from phonetic transcription of the Switchboard corpus. International Conference on Spoken Language Processing,

Hauser, M., N. Chomsky, and W.T. Fitch. 2002 The faculty of language: What is it, who has it, and how does it evolve? *Science* 298, 1569-1579.

Hauser, M. D., E. Newport and R. Aslin. 2001. Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. *Cognition* 78, B53-B64.

Hay, J. 2001. Lexical frequency in morphology: is everything relative? *Linguistics* 39.6, 1041-1070.

Hayes, B. 1995. *Metrical Stress Theory: Principles and Case Studies.* University of Chicago Press.

Jackendoff, R. and S. Pinker. 2005. The nature of the language faculty and its implications for the evolution of language (Reply to Fitch, Hauser and Chomsky). *Cognition* 97, 211-225

Kluender, K., A. Lotto, L. Holt, and S. Bloedel. 1998. Role of experience for language-specific functional mappings of vowel sounds. *JASA* 104.6, 3568-3582.

Kuhl, P. 1991. Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics* 50.2, 93-107.

Kuhl, P. And . Miller. 1975. Speech perception by the chinchilla: voiced-voiceless distinction in alveolar plosive consonants. *Science* 190, 69-72.

Leben, W. 1973. Suprasegmental Phonology. PhD Dissertation, M.I.T.

McCarthy, J. 1986. OCP effects: gemination and antigemination. *Linguistic Inquiry* 17, 207-263.

Newport, E. and R. Aslin 2004. Learning at a distance: I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology* 48, 127-162.

Newport, E., M. Hauser, G. Spaepen and R. Aslin. 2004. Learning at a distance: I. Statistical learning of non-adjacent dependencies in a non-human primate. *Cognitive Psychology* 49, 85-117.

Olvido, A, and W. Wagner. 2004. Signal components, acoustic preference functions and sexual selection in a cricket. *Biological Journal of the Linnean Society* 83.4, 461-472.

Peperkamp, S. and E. Dupoux. to appear. Learning the mapping from surface to underlying representations in an artificial language. In J. Cole and J. Hualde (eds.), *Laboratory Phonology 9.*

Pierrehumbert, J. 2003. Probabilistic phonology: discrimination and robustness. In R. Bod, J. Hay, and S. Jannedy (eds), *Probabilistic Linguistics*. 177-228. Cambridge, MA: MIT Press.

Pinker, S. and R. Jackendoff. 2005. The faculty of language: What's special about it? *Cognition* 95, 201-236.

Ramus, F., M. Hauser, C. Miller, D. Morris and J. Mehler. 2000. Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science* 188, 349-351.

Ramus, F. and J. Mehler. 1999. Language identification with suprasegmental cues: A study based on speech re synthesis. *JASA* 105.1, 512-521.

Ramus, F., M. Nespor and J. Mehler. 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 73, 265-292.

Saffran, J., E. Newport and R. Aslin. 1996. Statistical learning by 8-month old infants. S*cience* 274, 1926-1928.

Stout, J., C. DeHaan, and R.McGhee. 1983. Attractiveness of the male Acheta domesticus calling song to females. *Journal of Comparative Physiology A: Sensory, Neural and Behavioral Physiology* (Historical Archive) 153.4, 509-521.

Tincoff, R., N, Hauser, F. Tsao, G. Spaepen, F. Ramus, and J. Mehler. 2005. The role of speech rhythm in language discrimination: further tests with a non-human primate. *Developmental Science* 8.1, 26-35.

Toro, J., J. Trobalon and N. Sebastian-Galles. 2003. The use of prosodic cues in language discrimination tasks by rats. *Animal Cognition* 6, 131-136.

Yip, M. 1988. The Obligatory Contour Principle and Phonological Rules: A Loss of Identity. *Linguistic Inquiry* 19.1, 65-100.

Zuraw, K. 2002. Aggressive reduplication. *Phonology* 19.3, 395-440.

Department of Phonetics and Linguistics
University College London
Gower Street
London WC1E 6BT
ENGLAND
moira@ling.ucl.ac.uk
http://www.phon.ucl.ac.uk/home/moira/home.htm