



CrossMark
click for updates

Review

Cite this article: Collier K, Bickel B, van Schaik CP, Manser MB, Townsend SW. 2014 Language evolution: syntax before phonology? *Proc. R. Soc. B* **281**: 20140263. <http://dx.doi.org/10.1098/rspb.2014.0263>

Received: 1 February 2014

Accepted: 29 May 2014

Subject Areas:

behaviour, cognition, evolution

Keywords:

human language, animal communication, evolution, syntax, phonology, comparative approach

Author for correspondence:

Katie Collier

e-mail: katie.collier@ieu.uzh.ch

Language evolution: syntax before phonology?

Katie Collier¹, Balthasar Bickel², Carel P. van Schaik³, Marta B. Manser¹ and Simon W. Townsend¹

¹Institute of Evolutionary Biology and Environmental Studies, ²Department of Comparative Linguistics, and

³Anthropological Institute and Museum, University of Zurich, Zurich, Switzerland

Phonology and syntax represent two layers of sound combination central to language's expressive power. Comparative animal studies represent one approach to understand the origins of these combinatorial layers. Traditionally, phonology, where meaningless sounds form words, has been considered a simpler combination than syntax, and thus should be more common in animals. A linguistically informed review of animal call sequences demonstrates that phonology in animal vocal systems is rare, whereas syntax is more widespread. In the light of this and the absence of phonology in some languages, we hypothesize that syntax, present in all languages, evolved before phonology.

1. Introduction

Human language and its origins have intrigued philosophers and scientists since early antiquity [1]. This is unsurprising, as language is responsible for much that distinguishes humans from other species and makes us so successful, including the transmission of knowledge [2–5]. Unfortunately, the search for the origins of language is complicated by the fact that language, unlike other biological traits, does not fossilize or leave any traces to study its cumulative evolution. Empirical studies must therefore circumvent this problem and various different approaches have been undertaken to attempt to unpack the evolution of language [6–8]. These include, among others, the study of child language acquisition [9], hominid morphology [10–12], genetics [13] and the use of computer simulations to test specific hypotheses [14–16].

One method that has received particular attention is the exploration of similarities and differences between human language and animal communication systems [6,17]. If similarities are found between humans and a closely related species, then it is possible that they are derived from the same feature present in their common ancestor, representing homologues [18]. If, on the other hand, similarities are found between humans and more distantly related species, these features represent analogues and hence do not give any information on the phylogenetic origins of the feature, but can help elucidate the environmental or social conditions favourable to its convergent evolution [4].

One particular feature of human language that has received considerable attention by both linguists and animal communication researchers, and been highlighted as a 'fundamental universal structural characteristic' [19], is duality of patterning [20,21]. Also known as double articulation [22], duality of patterning is a property of language that allows a combinatorial structure on two levels: (i) phonology, where meaningless sounds called phonemes (i.e. the smallest meaning-differentiating elements of a language that do not themselves have meaning) can be combined into morphemes (i.e. the smallest meaningful elements) and words; and (ii) syntax, in which these morphemes and words can be combined into larger structures [23]. Critically, duality of patterning is the property that allows human languages to create a large lexicon from a few distinct signals [21,24–26]. Unpacking the evolutionary route that led to duality of patterning is thus considered central to a more holistic understanding of language evolution.

Table 1. Terms and definitions of different types of sound combinations used in animal communication research (non-bold type) and in linguistics (bold). In the visual representation, the circles of different colours on the left represent the different sounds to be combined, they can either have a meaning (represented by a letter as in the case of lexical syntax) or they can have no meaning. On the right, the series of circles represent call combinations that can have a meaning that is function of the meaning of its parts (e.g. A + B), no meaning or a new meaning (e.g. X).

| animal communication | human language | definitions | visual representation |
|---|------------------------------------|---|-----------------------|
| lexical syntax lexicoding compositional | grammar (syntax and morphology) | the way meaningful parts (morphemes, words) go together to form sentences a sequence of meaningful elements whose meaning is a function of the meaning of the individual elements that compose it and the way they are structured together | |
| phonological syntax phonocoding combinatorial | idioms (lexicon) | an expression whose meaning is not predictable from the parts that compose it | |
| | phonetics | the physical properties of sounds (phones) meaningless sounds are combined into sequences, the sequences obtained having no conventional meaning | |
| | phonology phonemics | minimal meaning-differentiating units (phonemes) that do not themselves bear meaning recombine to create meaningful expressions meaningful elements combine into a meaningful sequence whose meaning is not a function of the meaning of the parts. | |

Researchers of vocal communication in animals have emphasized the fact that animals are also capable of forming different types of sound combinations that could potentially be analogous or homologous to one or both levels of duality of patterning found in human languages [27–31]. Peter Marler played an important role in establishing the link between the levels of patterning found in human language and the different types of call combinations found in animal communication by introducing the terms phonological and lexical syntax, loosely based on the two levels of duality of patterning [32]. Marler defined phonological syntax (or phonocoding) as the level at which meaningless sounds are combined into sequences, and lexical syntax (or lexicoding) as the higher level at which the meaningful elements are combined. More recently, Hurford has used the terms combinatorial syntax (or combinatoriality) and compositional syntax (or compositionality) to designate the same phenomena as phonological and lexical syntax, respectively [26] (see table 1 for the terms and definitions of sound combinations used in animal communication research and their linguistic equivalents). Our goal here is to examine several examples of animal call combinations from a linguistic perspective and determine which level of duality of patterning they most resemble.

2. Examples of combinations in animal communication systems

(a) Winter wrens: phonological syntax?

Some of the best-studied examples of animal sound combinations come from birdsong [33]. One classic example of phonological syntax noted by Marler is the song of the winter wren (*Troglodytes troglodytes*) [32]. Kroodsma &

Momose [34] describe the songs of a Japanese population of winter wrens whose song types consist of a highly predictable sequence of notes or syllable types (a note being a continuous trace on a sonogram and a syllable being a repeated unit of identical notes or groups of notes). In their study population, the typical repertoire for a male includes six or seven song types. These different song types are obtained by reusing many of the same syllables or syllable sequences in a different order. However, as Marler noted, these syllables do not differentiate the song types from one another. In fact, all six or seven song types in a male wren's repertoire convey the same 'message' and none of them have any referential meaning [32]. Therefore, while superficially there seem to be structural similarities between bird song and human phonology, there are important differences when it comes to meaning differentiation. For the wren's song to have phonology in the linguistic sense, the different order of syllables in the different song types would have to bring about a change in meaning between the song types, just as in English *pat* and *tap* differ in meaning but are made up of the same sounds in a different order. Because of this, the structure of the wren's song (and that of most other bird and whale songs) would be better described not as phonological syntax but as *phonetic patterning*. Phonetics describes the physical properties of sound and, unlike phonology, it does not presuppose that sound patterns carry any function that serves to differentiate meanings.

Despite these critical differences, the search for comparative examples of phonology in animal communication has, in a similar way to Marler, continued to focus on bird [35,36] or whale song [37,38]. However, a more phoneme-focused approach could be taken by searching for the use and comprehension of minimal pairs (pairs of meaningful signs or words distinguished by only one element drawn from a

finite list; such as *tap* versus *lap* in English) in animal communication systems [39].

(b) Campbell monkeys: lexical syntax

Both Marler and Hurford argue that lexical syntax is only found in human language [26,32]. However, at least one example of call combination in an animal communication system could correspond to its definition. This is the use of an affixation system by Campbell monkeys (*Cercopithecus campbelli campbelli*) [40]. Campbell monkeys have two main predators: leopards (*Panthera pardus*) and crowned eagles (*Stephanoaetus coronatus*). The Campbell monkeys give a 'krak' call when they detect a leopard and a 'hok' call when they detect a crowned eagle. They can also add an affix '-oo' to both of these calls to produce two new calls: 'krak-oo' and 'hok-oo'. The 'krak-oo' call is given to any general disturbance and the 'hok-oo' call is given to any disturbance in the canopy. The critical aspect here is that the same '-oo' is affixed to both calls ('krak' and 'hok'). It is this use of the same elements, with the same meanings, in different sequences, that makes them compositional rather than combinatorial. The affixation modifies the meaning of the stem calls in a predictable way: changing a call designating a specific predator into a call designating a less specific disturbance in the same general physical space. Perhaps the closest language analogy would be the suffix '-like', changing the meaning of the call from 'leopard' to 'leopard-like (disturbance)'. The meaning of this suffix is fairly abstract: it does not refer to a concrete entity of its own, but directs the hearer to imagine a general situation that is disturbing in a similar way to the presence of a predator yet is not as dangerous as a real appearance of the predator. Abstract meaning operators of this kind are ubiquitous in human languages. Here, Campbell monkeys put together elements that conserve their meaning no matter what sequence they are part of, and obtain assemblies whose meaning reflects the meaning of their parts. This fits Hurford's definition of compositional syntax [26] and so deserves the name *syntax*, even if it is only a very rudimentary one.

(c) Putty-nosed monkeys: a less clear-cut example

The putty-nosed monkey's (*Cercopithecus nictitans*) combinatorial system is not so easy to categorize. In their communication system described by Arnold & Zuberbühler, putty-nosed monkeys produce two different loud calls: 'pyows' and 'hacks' [41]. These calls can be used as alarm calls when a predator is detected. If the predator is a leopard, the putty-nosed monkeys use 'pyows', and if it is a crowned eagle, they use 'hacks'. In addition to this, the monkeys can combine these two calls into another structure, the 'pyow-hack sequence'. This sequence normally consists of two to three 'pyows' followed by up to four 'hacks'. The 'pyow-hack sequences' elicit the movement of the group. While the components of this sequence bear meaning individually, the meaning of the sequence does not appear to derive from the meaning of these components, and so this combination does not conform to Marler's definition of lexical syntax [32]. There do, however, exist three alternative analyses that can be invoked to linguistically categorize and understand this call combination in relation to human language.

First, this communication system can be interpreted as a simple *phonological system*. Under this analysis, the 'pyows'

and the 'hacks' of the putty-nosed monkeys would be considered as phonemes in the linguistic sense, elements carrying no meaning *per se* but allowing the differentiation between the two single-segment morphemes (i.e. meaningful elements made up of only one sound) 'pyow' ('leopard') and 'hack' ('eagle'), and a morpheme composed of a sequence, 'pyow-hack sequence' ('let's go'). Thus, the element 'pyow' in the single call 'pyow' and in the 'pyow-hack' would be comparable to, say, the sound *s* in the single-segment morpheme *s* (as in *John's*) and in the sequence *so* or *us*—with no meaning in common, but serving as a diacritic for distinguishing meanings.

However, the data also allow alternative analyses that do not assume phonology and duality of patterning. Under one analysis, it would be possible to analyse the 'pyow-hack sequences' as idioms, where the original meanings of 'pyow' and 'hack' have become blurred. A possible etymology is this: the sequence first meant 'leopard and eagle' and then, derived from this by implication, 'danger all over'. This in turn came to mean 'danger all over, therefore let's go' and finally just 'let's go'. The human language analogue would be expressions like *kick the bucket*, the meaning of which is no longer transparently related to the meaning of the components, but has undergone complex etymological developments.

Alternatively, under another analysis, one could ascribe much more abstract meanings to 'pyow' and 'hack', such as 'move-on-ground' and 'move-in-air'. When produced on their own, listeners would seek the contextually most relevant and most suitable interpretation of these calls, possibly using similar heuristic processes such as are well established for human communicators in the theory of implicature inferences [42–44]. A default and common implicature would be, in the case of 'pyow', inference to a prototypical danger on the ground, a leopard; and, in the case of 'hack', a prototypical danger in the air, an eagle. Since under this analysis the calls themselves have very abstract meanings, it is possible to analyse pyow-hack sequences as lexical compositions: meanings like 'move-on-ground' and 'move-in-air' combine to a general meaning like 'we move; let's go' since putty-nosed monkeys themselves move both in the tree canopy and, though more rarely, on the ground [45].

Under either of these last two analyses, putty-nosed monkeys would, contrary to Arnold & Zuberbühler's conclusions [28], have *lexical syntax* in Marler's sense. At first sight, these alternative analyses are perhaps less plausible than positing phonology because they ascribe more complex cognitive processing to the monkeys: language change in the idiom-based analysis or abstract semantics and a well-tuned pragmatic inference machinery in the compositionality-based analysis. However, the communication system of the Campbell monkeys, a species closely related to the putty-nosed monkeys, suggests that their possible use of lexical syntax with abstract semantics is especially worth considering and should not be ruled out *a priori*.

(d) Banded mongooses: a non-primate example of lexical syntax?

Potential examples of lexical syntax are not limited to primate species: there are also examples from species more distantly related to humans, such as the close calls of the banded mongoose (*Mungos mungo*) [29]. Banded mongooses emit close calls while looking for food and these calls differ in structure depending on the exact nature of the behaviour: digging,

searching in the same foraging patch or moving between two patches. In all these contexts, the close call begins with an initial noisy segment that encodes the caller's identity, which is stable across all three contexts. Additionally, in the searching and moving context, there is a second tonal harmonic segment that does not encode identity; however, its length varies consistently with context, the segment being longer when the mongoose is moving rather than searching. These two segments, noisy and harmonic, come together in the call and indicate both the caller's identity and his activity.

As with the putty-nosed monkeys, it is possible to interpret these calls as a simple phonological system, with the noisy segment and short and long harmonic segments being three distinct phonemes. The noisy segment can then be produced alone as a single-segment morpheme when digging, or in combination with one of the other 'phonemes', which allow distinguishing between the different two-segment morphemes for searching or for moving.

In another interpretation, the banded mongoose close calls can act in an analogous way to short sentences: noisy segment + $\emptyset \rightarrow$ 'I (Fred) dig'; noisy segment + short harmonic segment \rightarrow 'I (Fred) search'; noisy segment + long harmonic segment \rightarrow 'I (Fred) move'; with the noisy segment acting as a referential expression that also encodes individual identity (somewhat like the caller's name) and the tonal segment as the 'predicate' that can be compared to simple subject–predicate compositions in human languages. Indeed, some human languages also use individually distinct expressions (i.e. personal names) in lieu of first-person pronouns. This is the case, for example, in Thai, where the use of first-person pronouns equivalent to 'I' is rude. Instead people routinely use their personal name instead of a first-person pronoun, for example saying 'Bill is cooking' while referring to oneself [46]. Under this analysis, the meaning of the assemblies produced by banded mongooses directly reflects the meaning of their different components, making these combinations, in a similar way to the Campbell monkeys', syntactic.

For now, either interpretation is possible, particularly because, in the absence of playback experiments, it is not clear what information listeners extract from these calls. Such experiments are therefore vital in helping shed light on whether banded mongoose close calls represent a syntactic or phonological system.

3. Examples from human languages where phonology is absent

While in animal communication systems sound combinations seem to be the exception, in human language they are the rule: all human languages combine words at the syntactic level and nearly all human languages, spoken or signed, have phonology, or cherology as it is known for sign languages. However, there do exist some languages possessing features without phonology, or lacking phonology altogether. Understanding why this structural feature of language is and can be absent could shed important light on the origins of syntax and phonology in human languages.

(a) Al-Sayyid Bedouin Sign Language

Most sign languages have phonology (cherology). This was first determined by Stokoe [47] in his work on American

Sign Language (ASL). Stokoe specifically demonstrated that ASL has three major categories (hand shape, location and movement) and that they each contain a certain number of features. Replacing one of these features by another causes a change in the meaning of the sign. This allowed Stokoe to conclude that ASL was not made up of holistic signs but of meaningless elements that are recombined into words.

Currently, one sign language is known that does not have phonology, or at least phonology has not fully developed throughout its entire lexicon. This is the Al-Sayyid Bedouin Sign Language (ABSL) described by Sandler *et al.* [48]. ABSL is a relatively new language used in the Al-Sayyid Bedouin group of the Negev region of Israel. The first deaf members of the group were four siblings born around 75 years ago. Over the next two generations, the number of deaf members increased as more were born into the community, most probably due to recessive congenital deafness [48]. There are now around 120–150 deaf members for a total of around 4000 members. ABSL is also used by a significant proportion of hearing members of the community.

Sandler *et al.* [48] looked for phonology in ABSL by searching for minimal pairs. For sign language, these can be distinguished by location, orientation, hand shape or movement. The authors did not find minimal pairs in ABSL [48]. On the contrary, they found a great variety in the signs for single words. For example, the sign for 'tea' can be represented by three different hand shapes and the sign for 'dog' can be made either in front of the mouth or in front of the torso (difference in location), depending on the signer. This lack of minimal pairs lead Sandler *et al.* to conclude that ABSL has no phonology and thus no duality of patterning [48]. Despite its lack of duality of patterning, from a linguistic point of view ABSL is a fully operational language, both in its function, allowing users to have conversations, make plans, tell stories and give instructions, and linguistically, having grammatical regularity at the syntactic, morphological and prosodic levels.

(b) Spoken languages

Of course, it could be that absence of duality of patterning is a peculiarity of an emerging communication system such as ABSL. However, although the spoken languages studied so far undeniably present duality of patterning, it is not implausible to assume, as does Blevins, that duality of patterning is not an absolutely universal property. Blevins discusses segment-sized morphemes in a number of languages [49]. An example is the English morpheme *s*, which can mean 'plural' (*book-s*), 'third-person singular present' (*she look-s*) or 'possessor' (*Rik's*). The reason we analyse *s* as three morphemes with a phoneme /s/ is because the same phoneme recurs in a great number of other morphemes (*soup*, *test*, *miss*, etc.). If this were not the case, one could just as well say that we have a meaning-bearing segment *s* that happens to be three-ways ambiguous. If a language has a large inventory of such meaning-bearing segments and the meanings are sufficiently abstract, this would easily allow a sizeable expressive power without duality of patterning. The two critical requirements for this—abstract meanings and large inventories of segments—are both well established in extant languages.

First, there are languages whose lexicon is composed of words with highly abstract meanings. Consider, for example, words like *st'uswalíc* 'I picked up the rag' in the North

American language Atsugewi, which is composed of a prefix *s'w-* for 'I' followed by the three morphemes *tu* 'do something by hand', *swal* 'for limp (not stiff or resilient) material to move or be located' and *ic* 'upward' [50]. In such a system, a limited number of abstract meanings are strung together and then subjected to a rich machinery of pragmatic inference, deriving concrete meaning effects.

Second, there are languages with impressively large segment inventories. The known maximum is found in !Xõõ in Botswana, with 164 segmental phonemes [51]. Many languages in addition have suprasegmental features like tone (as also found in !Xõõ), vowel and consonant lengthening, nasalization (e.g. *owoku* 'house' versus *õwõngu* 'my house' in the Terena language of Brazil [52]) and holistic sound sequences such as are found in interjections (e.g. *ʔi* 'yes' and *ʔi* 'no' in English). It is easy to imagine that all these possibilities co-occur in a single language, so that inventories quickly reach between 160 and 180 units, each carrying its own abstract meaning.

Furthermore, Blevins notes that in many languages, meanings depend on position and context [49] (just as the English *-s* means different things depending on whether it follows a noun or a verb stem; cf. above). Even just distinguishing word-initial and word-final positions in two-segment words would thus already yield a potential for more than 300 meanings; adding a noun versus verb distinction could double this number again. Finally, as Blevins also observes, many languages have what are called bi-partite or tri-partite stems, where stems are non-transparently composed of morphemes, like idioms (cf. e.g. in Andi, a language of the Caucasus, *abcho* 'someone washed it', with the bipartite stem *a-ch* 'wash', interrupted by an agreement marker *b-* and followed by a past tense marker *-o* [53]). This quickly adds a few hundred other meanings (in fact, with 180 units that can freely combine with each other in first and second part, a language could potentially have up to 180^2 bipartite stems, which is already beyond an average speaker's lexicon in daily use). Given all these possibilities, it is perfectly possible that there might have been (or will be) a spoken language in the world that lacks duality of patterning. The lexicon of such a language might not (easily) allow growth on the scale of languages with duality of patterning, but if we also allow for borrowing words from other languages, even these limitations are not as detrimental as one might think.

4. Discussion

(a) Syntax before phonology

The examples discussed in this review demonstrate that (i) while phonology in the linguistic sense seems to be rare in animal communication systems, lexical syntax seems to be more widespread than previously thought, and (ii) while there is no human language without syntax, it seems possible for some human languages to lack phonology. This appears to indicate that a single layer of compositional structure (syntax) is less complex to develop than adding to this an extra layer of phonological structure. This leads us to hypothesize that, contrary to the traditional view in both linguistics and animal communication research [54], syntax developed before phonology in human languages.

This hypothesis seems to be further supported by the fact that human languages lacking phonology but possessing syntax, such as ABSL, are emerging languages that do not yet

seem to be fully formed. This suggests that syntax develops first to allow the expression of more concepts with only a few words, while phonology appears later on in the development of a language, when the need for a larger vocabulary makes it a more efficient way to produce an increased number of words. If this is the case, we would expect any new emerging languages to present a similar pattern, with syntax developing before phonology. Preliminary surveys suggest that this may be the case for most spontaneous sign languages [55]. In terms of spoken languages, it is harder to search for similar developmental patterns, as emerging spoken languages such as pidgins and creoles are created when people who speak different languages need to communicate. Therefore, these languages are not created from scratch and their sound system is most often taken from one of the original languages [56].

Why syntax developed before phonology is of course open to discussion, but it could be that, from a cognitive perspective, syntax is simpler to process than phonology. Intuitively, it would seem that syntactical combinations would require less memorizing, as only the meanings of the individual signals would need to be learned and remembered, the meaning of the combination being derived from them. For phonological combinations, on the other hand, it would seem that a new meaning has to be learned for each different sequence of sounds.

(b) Insights into the origins of syntax and phonology

While the examples analysed in this review can give some insight into the order of development of different types of sound combinations, they also allow us to formulate hypotheses regarding the conditions favouring their evolution. One obvious similarity between the species demonstrating combinations of meaningful calls is that they all reside in groups characterized by high sociality. This social dimension may well require such species to express more concepts than would be possible with only the individual calls from their anatomically constrained vocal repertoire. One solution to this constraint is to develop a more open-ended vocal repertoire through learning, as is the case in a number of bird species and social mammal species [57]. Alternatively, as we see here, calls could be flexibly combined to express related (compositional syntax) or even unrelated (combinatorial syntax) meanings [41].

Furthermore, of the three major examples we present, two represent call combinations used in less urgent situations. In the case of the Campbell monkeys signallers use single alarm calls on their own to indicate a predator, whereas they use the affixed call for a more general, less immediately threatening disturbance. In a similar way, banded mongoose call combinations occur while foraging rather than in immediate predation contexts. As a shorter time between the perception of the danger by the emitter and the reaction of the receiver would be more advantageous in urgent situations, one might predict clearer evidence for syntax in more relaxed, social contexts [58]. Indeed, for human language, it is well established that more complex and elaborate kinds of syntax are better represented in written than in spoken language [59] (i.e. in a mode of language use that is removed from the rapid and socially challenging interactions that characterize spoken language).

Given the current absence of unambiguous examples of phonology in the linguistic sense in animal communication systems (i.e. there is no clear evidence of patterns of communication that cannot be explained without assuming phonology), variation

among human languages may provide additional insight into the origins of this feature. First, the examples of human language features lacking phonology, such as segment-sized morphemes or holistic sound sequences, suggest that duality of patterning is an empirically observed correlation and not a logically necessary property of language [48]. New observations are constantly providing additional empirical data to be interpreted. Second, the absence of phonology in certain aspects of languages, or even in whole languages, points towards a non-genetic basis for this feature in human language. Like songbirds [35] and some mammal species (cetaceans [60], pinnipeds [61], elephants [62], bats [63]), humans are vocal learners capable of producing a large number of different sounds. However humans are, as far as we know, the only species that use these sounds phonologically to distinguish between the meanings of two sequences. This suggests that vocal learning and the capacity to produce a large number of different sounds alone are not sufficient to induce the emergence of a phonological level. We therefore argue that the constraints leading to the use of a phonological level are more likely to be cognitive in nature rather than linked to the production capacity of a given species. Specifically, once humans developed the cognitive capacities to memorize phonological combinations and their meanings, phonology itself could become subject to cultural, as opposed to biological, evolutionary processes [23,64]. If this is the case, it might explain why phonology in the linguistic sense is so rare in the communication systems of other species.

The constraints driving the cultural evolution of phonology should be widespread across human cultures, reflecting the distribution of the property itself. These constraints could include the need for distinctiveness and learnability, as well as a tendency to keep meaningful distinctions while trying to make an utterance sound similar to other utterances in a population [23]. As Hockett noted, phonology is most efficient when there is a large set of meanings to be expressed, because the combination of phonemes is generally less constrained than the combination of morphemes: the combination of morphemes must ‘make sense’ [24]. ABSL may lack phonology because it does not currently have these constraints. It is a small community language and its users know each other, potentially making pragmatics and inference an important part of their communicative understanding. However, if the use of ABSL were to spread to a larger population of signers, we could expect a gradual emergence of phonology. In fact,

ABSL already seems to have a blueprint for the development of phonology, with the emergence of categories, the regularization of signs within familylects and young signers using conventionalized signs rather than iconic ones [48].

(c) Conclusion

Duality of patterning is considered an important feature of language. From a comparative perspective, this has led to great interest in animal call combinations and their similarities to the two levels of structure found in duality of patterning: phonology and syntax. In this review, we have shown that there exist no clear examples for phonology in the linguistic sense in animal communication systems, and that, contrary to traditional thought, syntax or compositionality is actually more widespread. When also analysing the structure of human languages, we found that some parts of some languages, and at least one entire language, do not display phonology. From these observations, we alternatively argue that syntax developed before phonology and that the former seems to be a cognitively simpler process, with the latter possibly being the product of cultural evolution. This could be taken into account in future research on meaningful animal call combinations by assuming lexical, and not phonological, syntax as the simplest explanation.

If a certain language property, such as phonology, is not universally present in all human languages, then it is probably unsurprising that it is non-existent in a large number of animal communication systems. However, if the factors leading to the presence (or absence) of this property can be determined, they may allow us to make predictions on which species or social contexts to focus our research effort to find these analogous or homologous properties in animal communication systems if they do exist. This focus fits with recent developments in linguistics that increasingly challenge the idea of a given set of properties defining all and only human languages, and instead probe into the social and biological factors that condition how specific properties of language arise, develop and disappear again in the course of time [65–67].

Acknowledgements. We thank Per Lundberg and two anonymous reviewers for helpful comments on a previous version of the manuscript, and Sabrina Engesser for discussions.

Funding statement. This work was funded by a collaborative University of Zurich Research Priority Program grant (URPP, Evolution in Action).

References

1. Aristotle 350 BC. *The history of animals*. (Reprinted by Kessinger Publishing 2004.)
2. Darwin C. 1871 *The descent of man and selection in relation to sex*. New York, NY: D. Appleton and Company.
3. Tomasello M. 2008 *Origins of human communication*. Cambridge, MA: MIT Press.
4. Fitch WT, Huber L, Bugnyar T. 2010 Social cognition and the evolution of language: constructing cognitive phylogenies. *Neuron* **65**, 795–814. (doi:10.1016/j.neuron.2010.03.011)
5. Hurford JR. 2007 *The origins of meaning: language in the light of evolution*. Oxford, UK: Oxford University Press.
6. Hauser MD, Chomsky N, Fitch WT. 2002 The faculty of language: what is it, who has it, and how did it evolve? *Science* **298**, 1569–1579. (doi:10.1126/science.298.5598.1569)
7. Christiansen MH, Kirby S. 2003 Language evolution: consensus and controversies. *Trends Cogn. Sci.* **7**, 300–307. (doi:10.1016/S1364-6613(03)00136-0)
8. Gong T, Shuai L, Wu Y. 2013 Multidisciplinary approaches in evolutionary linguistics. *Lang. Sci.* **37**, 1–13. (doi:10.1016/j.langsci.2012.09.002)
9. Tomasello M. 2003 *Constructing a language: a usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
10. Lieberman P. 1973 On the evolution of speech. In *Festschrift for Morris Halle* (eds S Anderson, P Kiparsky), pp. 107–127. New York, NY: Holt, Rinehart and Winston.
11. Lieberman P. 1984 *The biology and evolution of language*. Cambridge, MA: Harvard University Press.
12. Lieberman P. 1998 *Eve spoke: human language and human evolution*. New York, NY: W. W. Norton.
13. Enard W, Przeworski M, Fischer SE, Lai CS, Wiebe V, Kitano T, Monaco AP, Pääbo S. 2002 Molecular evolution of FOXP2, a gene involved in speech and language. *Nature* **418**, 869–872. (doi:10.1038/nature01025)

14. Deacon TW. 2003 Universal grammar and semiotic constraints. In *Language evolution* (eds MH Christiansen, S Kirby), pp. 111–139. Oxford, UK: Oxford University Press.
15. Kirby S. 2001 Spontaneous evolution of linguistic structure: an iterated learning model of the emergence of regularity and irregularity. *IEEE J. Evol. Comput.* **5**, 102–110. (doi:10.1109/4235.918430)
16. Nowak MA, Komarova NL. 2001 Toward an evolutionary theory of language. *Trends Cogn. Sci.* **5**, 288–295. (doi:10.1016/S1364-6613(00)01683-1)
17. Zuberbühler K. 2005 The phylogenetic roots of language: evidence from primate communication and cognition. *Curr. Dir. Psychol. Sci.* **14**, 126–130. (doi:10.1111/j.0963-7214.2005.00357.x)
18. Fedurek P, Slocumbe KE. 2011 Primate vocal communication: a useful tool for understanding human speech and language evolution? *Hum. Biol.* **83**, 153–173. (doi:10.3378/027.083.0202)
19. Hurford JR. 2002 The roles of expression and representation in language evolution. In *The transition to language* (ed. A Wray), pp. 311–334. Oxford, UK: Oxford University Press.
20. Hockett CF. 1960 Logical considerations in the study of animal communication. In *Animal sounds and communication* (eds WE Lanyon, WN Tavolga), pp. 392–430. Washington, DC: American Institute of Biological Sciences.
21. Hurford J. 2008 The evolution of human communication and language. In *Sociobiology of communication: an interdisciplinary perspective* (eds P D'Etorre, DP Hughes), pp. 249–264. Oxford, UK: Oxford University Press.
22. Martinet A. 1949 La double articulation linguistique. *Cercle Linguistique de Copenhague* **5**, 30–37. (doi:10.1080/01050206.1949.10416289)
23. De Boer B, Sandler W, Kirby S. 2012 New perspectives on duality of patterning: introduction to the special issue. *Lang. Cogn.* **4**, 251–259. (doi:10.1515/langcog-2012-0014)
24. Hockett CF. 1960 The origin of speech. *Sci. Am.* **203**, 88–111. (doi:10.1038/scientificamerican0960-88)
25. von Humbolt W. 1836 *Über die Kawi-Sprache auf der Insel Java*. Berlin, Germany: Druck-erei der Koniglichen Akademie der Wissenschaften.
26. Hurford JR. 2011 *The origins of grammar: language in the light of evolution II*. Oxford, UK: Oxford University Press.
27. Zuberbühler K. 2002 A syntactic rule in forest monkey communication. *Anim. Behav.* **63**, 293–299. (doi:10.1006/anbe.2001.1914)
28. Arnold K, Zuberbühler K. 2012 Call combinations in monkeys: compositional or idiomatic expressions? *Brain Lang.* **120**, 303–309. (doi:10.1016/j.bandl.2011.10.001)
29. Jansen DA, Cant MA, Manser MB. 2012 Segmental concatenation of individual signatures and context cues in banded mongoose (*Mungos mungo*) close calls. *BMC Biol.* **10**, 97. (doi:10.1186/1741-7007-10-97)
30. Hailman JP, Ficken MS, Ficken RW. 1985 The 'chick-a-dee' calls of *Parus atricapillus*: a recombinant system of animal communication compared with written English. *Semiotica* **56**, 191–224. (doi:10.1515/semi.1985.56.3-4.191)
31. Freeberg TM, Lucas JR. 2002 Receivers respond differently to chick-a-dee calls varying in note composition in Carolina chickadees, *Poecile carolinensis*. *Anim. Behav.* **63**, 837–845. (doi:10.1006/anbe.2001.1981)
32. Marler P. 1998 Animal communication and human language. In *The origin and diversification of language* (eds NG Jablonski, Aiello), pp. 1–20. San Francisco, CA: California Academy of Sciences.
33. Berwick RC, Okanoya K, Beckers GJL, Bolhuis JJ. 2011 Songs to syntax: the linguistics of birdsong. *Trends Cogn. Sci.* **15**, 113–121. (doi:10.1016/j.tics.2011.01.002)
34. Kroodsma DE, Momose H. 1991 Songs of the Japanese population of the winter wren (*Troglodytes troglodytes*). *Condor* **93**, 424–432. (doi:10.2307/1368959)
35. Catchpole CK, Slater PJB. 2008 *Bird song: biological themes and variations*, 2nd edn. Cambridge, UK: Cambridge University Press.
36. Gentner TQ, Fenn KM, Margoliash D, Nusbaum HC. 2006 Recursive syntactic pattern learning by songbirds. *Nature* **440**, 1204–1207. (doi:10.1038/nature04675)
37. Payne RS, McVay S. 1971 Songs of humpback whales. *Science* **173**, 587–597. (doi:10.1126/science.173.3997.585)
38. Suzuki R, Buck JR, Tyack PL. 2006 Informational entropy of humpback whale songs. *J. Acoust. Soc. Am.* **119**, 1849–1866. (doi:10.1121/1.2161827)
39. Yip MJ. 2006 The search for phonology in other species. *Trends Cogn. Sci.* **10**, 442–446. (doi:10.1016/j.tics.2006.08.001)
40. Ouattara K, Lemasson A, Zuberbühler K. 2009 Campbell's monkeys use affixation to alter call meaning. *PLoS ONE* **4**, e7808. (doi:10.1371/journal.pone.0007808)
41. Arnold K, Zuberbühler K. 2006 Language evolution: semantic combinations in primate calls. *Nature* **441**, 303–303. (doi:10.1038/441303a)
42. Grice HP. 1975 Logic and conversation. In *Syntax and semantics III: Speech acts*. (eds P Cole, JL Morgan), pp. 41–58. New York, NY: Academic Press.
43. Sperber D, Wilson D. 1986 *Relevance: communication and cognition*. Oxford, UK: Blackwell.
44. Levinson SC. 2000 *Presumptive meanings: the theory of generalized conversational implicature*. Cambridge, MA: MIT Press.
45. Bitty EA, McGraw WS. 2007 Locomotion and habitat use of Stampflii's putty-nosed monkey (*Cercopithecus nictitans stampflii*) in the Tai National Park, Ivory Coast. *Am. J. Phys. Anthropol.* **134**, 383–391. (doi:10.1002/ajpa.20679)
46. Iwasaki S, Horie IP. 2005 *A reference grammar of Thai*. Cambridge, UK: Cambridge University Press.
47. Stokoe WC. 1960 *Sign language structure*. Studies in Linguistics, occasional paper 8. Buffalo, NY: Department of Anthropology and Linguistics, University of Buffalo.
48. Sandler W, Aronoff M, Meir I, Padden C. 2011 The gradual emergence of phonological form in a new language. *Nat. Lang. Linguist. Theory* **29**, 503–543. (doi:10.1007/s11049-011-9128-2)
49. Blevins J. 2012 Duality of patterning: absolute universal or statistical tendency? *Lang. Cogn.* **4**, 275–296. (doi:10.1515/langcog-2012-0016)
50. Talmy L. 1972 *Semantic structures in English and Atsugewi*. Linguistics Department, University of California, CA, USA.
51. Traill A. 1985 *Phonetic and phonological studies of !Xäoo Bushman*. Hamburg, Germany: H. Buske.
52. Akinlabi A. 1996 Featural affixation. *J. Ling.* **32**, 239–289. (doi:10.1017/S0022226700015899)
53. Bickel B, Nichols J. 2007 Inflectional morphology. *Lang. Typol. Syntact. Des.* **3**, 169–240. (doi:10.1017/CBO9780511618437.003)
54. Fitch WT. 2010 *The evolution of language*, 1st edn. Cambridge, UK: Cambridge University Press.
55. Meir I, Sandler W, Padden C, Aronoff M. 2010 Emerging sign languages. In *Oxford handbook of deaf studies, language, and education*. Oxford, UK: Oxford University Press.
56. McGregor WB. 2009 *Linguistics: an introduction*. London, UK: Continuum.
57. Slater PJB, Janik VM. 2010 Vocal Learning. In *Encyclopedia of animal behavior* (eds MD Breed, J Moore), pp. 551–557. Oxford, UK: Academic Press.
58. Candiotti A, Zuberbühler K, Lemasson A. 2012 Context-related call combinations in female Diana monkeys. *Anim. Cogn.* **15**, 327–339. (doi:10.1007/s10071-011-0456-8)
59. Karlsson F. 2010 Syntactic recursion and iteration. In *Recursion and human language* (ed. H. van der Hulst), pp. 43–67. Berlin, Germany: De Gruyter Mouton.
60. Tyack PL, Sayigh LS. 1997 Vocal learning in cetaceans. In *Social influences on vocal development*, pp. 208–233. Cambridge, UK: Cambridge University Press.
61. Schusterman RJ. 2008 Vocal learning in mammals with special emphasis on pinnipeds. In *The evolution of communicative flexibility: complexity, creativity, and adaptability in human and animal communication* (eds DK Oller, U Griebel), pp. 41–70. Cambridge, MA: MIT Press.
62. Poole JH, Tyack PL, Stoeger-Horwath AS, Watwood S. 2005 Animal behaviour: elephants are capable of vocal learning. *Nature* **434**, 455–456. (doi:10.1038/434455a)
63. Boughman JW. 1998 Vocal learning by greater spear-nosed bats. *Proc. R. Soc. Lond. B* **265**, 227–233. (doi:10.1098/rspb.1998.0286)
64. Tomasello M. 2003 *Origins of human communication*. Cambridge, MA: MIT Press.
65. Bickel B. 2007 Typology in the 21st century: major current developments. *Linguist. Typol.* **11**, 239–251. (doi:10.1515/LINGTY.2007.018)
66. Bickel B. In press. Linguistic diversity and universals. In *Cambridge handbook of linguistic anthropology* (eds N Enfield, J Sidnell, P Kockelman). Cambridge, UK: Cambridge University Press.
67. Evans N, Levinson SC. 2009 The myth of language universals: language diversity and its importance for cognitive science. *Behav. Brain Sci.* **32**, 429. (doi:10.1017/S0140525X0999094X)