AGAINST THE EMERGENT VIEW OF LANGUAGE EVOLUTION

MAGGIE TALLERMAN

Linguistics Section, Newcastle University,
Newcastle upon Tyne, UK
maggie.tallerman@ncl.ac.uk

Under the emergent view of language evolution (Nóbrega & Miyagawa 2015), it is claimed that the language faculty appeared quite recently in human evolution and had no earlier pre-syntactic stages. Language is said to be an amalgamation of two pre-existing systems that also occur in animal communication: an 'E' system, for 'expressive', which is likened to systems of learned birdsong, and an 'L' system, for 'lexical', which is likened to monkey alarm calls. What occurred solely in the case of humans was the advent of a Merge operation, which integrated the two systems. Here I argue against this 'integration hypothesis'. None of the proposed analogues in animal communication have the critical properties occurring in human language. The syntax of birdsong is unlike the syntax of language in all relevant respects. There are no analogues to the functional elements of the E-system in animal communication. However, under a gradualist rather than an emergent view of language evolution, we can see how functional elements and linguistic morphology evolved.

1. Introduction: the Integration Hypothesis

In a recent paper, Nóbrega & Miyagawa (2015) propose what they term an 'emergent' view of language evolution, which 'assumes that the language faculty emerged late in historical development without any prior pre-syntactic stage'. Under this view, language is the result of Merge operating on two pre-existing systems that also occur elsewhere in animal communication, but are only integrated in humans. Nóbrega & Miyagawa specifically argue against the 'gradualist' view of language evolution, which proposes an asyntactic protolanguage initially consisting of wordlike entities without combinatorial properties, as outlined in the work of Bickerton (1990, 1995 et seq.), Hurford (2012) and Jackendoff (2002), among many others. According to Nóbrega & Miyagawa, and contrary to the cited work on protolanguage, words do not precede syntax, but rather, are derived by it. Here, I outline a number of problems for the Integration Hypothesis (Miyagawa et al. 2013, 2014; Nóbrega & Miyagawa 2015), and argue that in fact there are no evident analogues in animal communication for the two main components of language that the authors propose.

The Integration Hypothesis states that language comprises two main components: sentences are constructed of an inner L-layer, for 'lexical', providing basic lexical meaning, and an outer E-layer, for 'expressive', which

has properties such as Tense, or Question, or Focus (Miyagawa et al. 2013, 2014); the latter are functional elements that lack independent status. Nóbrega & Miyagawa (2015) go further, suggesting that not only phrases and sentences but also simple words consist of these two layers, so that there is an inner L-structure at the core of all expressions, and an outer E-layer, the 'expression structure', comprising 'function elements that give shape to the expression' (Nóbrega & Miyagawa 2015): these are categorial and inflectional features. For words, the lexical layer is said to be a ROOT, which lacks syntactic category, so is not yet a noun, verb or adjective etc.; it also lacks inflectional information, so has no number, case or tense etc. A root cannot be or become a word until it combines with categorial and grammatical information, provided by the E-layer. Nóbrega & Miyagawa propose that, crucially, this stage in language evolution does not occur until the late appearance of a Merge operation in the grammar (thus, presumably, in human cognition), since Merge integrates the two pre-existing systems, L and E. At word-level, the proposal is that Merge then produces units with syntactic categories and other grammatical features: words as we know them.

Leaving aside the controversial question of whether or not the Merge stage was a recent development, occurring abruptly within the past 100,000 years (though see Dediu & Levinson 2013, Johansson 2013, Tallerman 2014 for alternative arguments), one major problem for the Integration Hypothesis is that it cannot satisfactorily answer the question of where the categorial and grammatical features themselves come from. Miyagawa and colleagues propose that each of the two layers has an antecedent or an analogue elsewhere in animal communication; it is only Merge that is novel. The E-type is claimed to be similar to birdsong, which is said to have 'syntax without meaning' (Nóbrega & Miyagawa 2015), while the L-type, the lexical component, is likened to monkey calls, considered to be isolated units with real-world reference. An immediate problem here is that we do not and cannot know what monkey alarm calls refer to, if indeed they *refer* directly to anything in the real world at all, rather than being purely affective, so relating to the caller's internal state on perceiving a specific predator.

2.   Does the E-system have analogues in animal communication?

2.1 *The E-system and birdsong*

Here, I chiefly consider problems with the proposed antecedents for – and the appearance of – the E-layer in language. A variety of parallels between birdsong and language have been suggested in the literature, including the learning of song by juvenile birds of certain species, the appearance of a critical period in both language and learned song, and so on (e.g Bolhuis et al. 2010). Miyagawa

and colleagues, on the other hand, propose that 'the actual link is between birdsong and the expression structure portion of human language' (Nóbrega & Miyagawa 2015). This layer is claimed to be directly analogous to birdsong (actually the authors go further, claiming a genetic link between birdsong and language). How robust is this analogy? In birdsong, the E-layer is said to have 'expressive' functions such as marking territory or mating availability (Miyagawa et al. 2013, 2014), obviously rather starkly different functions to the linguistic ability to mark a question or focalize an element, also handled by the E-component. More problematically, however, in language these functions are typically signalled by permutations in the syntax, so that for instance focalized constituents move to a special position in a clause – fronted, as in English (*This colour, I really love __*), or an immediately pre-verbal position, as in Basque and Hungarian. Questioned phrases may remain in situ, where they are marked with a *wh*-particle, as in Japanese; but cross-linguistically, *wh*-phrases are often signalled by fronting too. In fact, much of the whole *raison-d'être* of syntax lies in moving elements around to signal different expressive functions, including topicalizing, focalizing, questioning, foregrounding and backgrounding various constituents via passivization (*The window got smashed __ (by the football)*), and so on, all to achieve distinct pragmatic effects. Syntax not only varies word order, but crucially also changes the valency of predicates, altering grammatical relations (for instance, the object of an active sentence is the subject of the corresponding passive), all in order to affect the meaning and pragmatic function of sentences.

In contrast, birdsong specialists note that song variation may be extensive, but specifically does *not* signal different meanings: 'in general the different songs that a bird has convey exactly the same message as each other but do it in a highly varied manner' (Slater 2012: 97). Strikingly, furthermore, even the more complex bird songs are characterized by highly stereotyped ordering sequences, for instance in the song of the chaffinch. Different syllables (in birdsong, a syllable denotes a precisely repeated sequence of one or more notes; Thompson et al. 1994) may be iterated a different number of times, but their sequence remains fixed. Nightingales have a repertoire of around 100–200 distinct song types (Kipper et al. 2006), but again the sequence of these is fixed. In the zebra finch, the song is also stereotyped, though notes may be skipped or added, or the same note repeated, giving an appearance of variation. Interestingly, it seems that the variations in note order may not even be deliberate, but rather are something akin to 'performance' errors in these finches (Hurford 2012: 47f). Even where there are permutations in birdsong syntax, no differentiation occurs in the message: in fact, the song syntax itself seems anything but 'expressive', if the variations carry no distinct meanings. It appears, then, that the proposed analogy between an E-layer in language and the syntax of birdsong is at best tenuous.

2.2 *Are there analogues to functional elements in animal communication?*

Turning now to the role of the E-layer below sentence level, recall that this is claimed by Nóbrega & Miyagawa (2015) to turn L-layer roots into words, via Merge, in the course of language evolution, by adding categorial and inflectional information. Certainly, the relevant literature does not suggest that birdsong displays anything remotely similar to the functional elements and grammatical features of language. Where, then, do these elements come from in language, if there are no obvious antecedents in animal communication? We might expect that an investigation of primate communication could shed light on the origins of functional elements, but there is little evidence for this. Nóbrega & Miyagawa speculate that it is the L-layer, not the E-layer, that derives from primate communication: monkey alarm calls are the 'closest approximation' to linguistic roots in their view. However, alarm calls seem to have more in common with propositions than roots (Bickerton 2009), and might equally indicate the specific escape strategy required rather than essentially naming the predator. Problematically, also, great apes – despite being phylogenetically much closer to humans – lack such specific alarm calls.

In captive great apes, as we see from ape 'language' research, functional elements are generally conspicuous by their absence. Kanzi the bonobo, for instance, is well able to carry out many surprising instructions given in English (e.g. *Can you pour the ice water in the pottty? Go get the balloon that's in the microwave*; Savage-Rumbaugh et al. 1998), to a level of 72% accuracy, but is unable to understand coordination (*Bring me a carrot and a banana*). This may indicate Kanzi's inability to understand hierarchical structure, or a failure to comprehend a critical functional element such as *and* (the authors also report a problem with modal auxiliary *can*). In the examples above, none of the function words are remotely vital to understanding the meanings (Kirby 2000); even in the case of the relative clause, lexical items plus word order suffice: there is no reason to suppose that Kanzi understands subordination. Kanzi is also unable to understand the indefinite article *a*, so that when asked to give *a carrot / a tomato*, he makes an inappropriate response. It seems unlikely, then, that anything in the E-layer, specifically the uniquitous functional elements of language, stems from our shared phylogeny with other apes.

There are indeed a few suggestions in the literature that something approaching a functional element – a potential antecedent for the E-layer – occurs in certain monkey calls (e.g. Zuberbühler 2002, Ouattara et al. 2009). In Campbell's monkeys, a 'boom' vocalization preceding (and apparently attenuating) other alarm calls has been interpreted as a 'syntactic rule' (Zuberbühler 2002), with the 'boom' acting as a kind of modifier. Notably, though, the alarm calls occur about 25 seconds later than the boom, indicating that these are not in fact call combinations after all. Ouattara et al. (2009) also

suggest that Campbell's monkeys use a stem + suffix arrangement to broaden the meaning of several distinct calls, so that for instance the eagle alarm is *hok* while *hok-oo* (i.e. purportedly the *hok* call with an *-oo* suffix) is used for 'general arboreal disturbance'. However, while *-oo* may be described as a suffix if the term is used very loosely, it has little resemblance to affixes and other types of morphology in languages.

3.   Functional elements have a purely linguistic development

A wide linguistic literature offers clear views of the sorts of attested pathways that languages follow in the development of independent functional elements and grammatical formatives (e.g. Heine & Kuteva 2007; Carstairs-McCarthy 2010). In essence, inflectional and derivational morphology has a life history, and that history depends exactly on words and morphemes being uttered in closely-adjoining sequences, something that does not occur in primate calls. Functional elements of all kinds in language do not occur spontaneously, but develop from pre-existing lexical items, via various routes of grammaticalization that recur with remarkable frequency cross-linguistically (Heine & Kuteva 2002, 2007). For instance, a 'go (to)' verb often develops a future meaning, or even becomes a future marker, devoid of the semantics of movement, as in English *I'm gonna sit still*; the same process is observed in French, Welsh, Sotho, Zulu etc. Similarly, the Baka adverb *ngili* 'yesterday', is grammaticalized as a verbal tense suffix *-ngi*, marking the near past. Many, though not all, affixes originate as separate words, subsequently morphologized; the history of verbal inflection in the Romance languages is another standard example. New morphological variants (allomorphs) frequently appear as a result of very standard phonological assimilation processes occurring at morpheme and word boundaries yielding distinct forms of the same morpheme; German umlaut illustrates, e.g. *Fuss* [fʊs] 'foot' vs. Füss-e (foot-PLURAL) [fys-ə] 'feet', where the vowel in the noun stem underwent fronting in the plural, under the influence of what was historically a front vowel in the plural suffix. A gradualist view of language evolution is required in order to understand the kinds of development outlined here, since they are by nature gradually-developing phenomena.

In sum, then, we have a very good idea of where grammatical information and functional elements come from, but nothing in the communication systems of other animals seems to provide a promising candidate for antecedents of the E-layer in language. The L-layer seems equally to be a totally distinct system to primate calls. I therefore suggest that we cannot accept the proposal that Merge simply integrates two pre-existing systems, contrary to the Integration Hypothesis. If the E-system and the L-system were there already before Merge existed, it is because both of them evolved – uniquely – in human cognition.

4. A pre-language stage in evolution

Finally, Nóbrega & Miyagawa's concept of language evolution also raises other questions. If the L-layer existed before Merge, then roots existed. Although the authors deny the existence of a 'structureless, protolinguistic stage', they actually do envisage an earlier stage of language containing pre-words: 'The pre-language stage is composed of root-like elements, each occurring in isolation of the others' (Nóbrega & Miyagawa 2015). These, they say, would have reference but lack syntactic category, and thus would be unable to 'participate in any combinatorial systems'. In fact, such a stage seems to be extremely similar to what is proposed in much of the literature on protolanguage, e.g. Bickerton (1990, 1995), Hurford (2012), Jackendoff (2002), Jackendoff & Wittenberg (2014), Tallerman (2007), and many others. In the Nóbrega & Miyagawa system, such pre-words do not join together to form hierarchical structures, since without syntactic categories a pre-word must lack the properties that words have, such as headedness and taking obligatory complements. There could be, then, no syntax. However, nothing in this system seems to prevent roots being strung together in very short sequences, quite arbitrarily, where their meanings seemed to be related, thus forming short, asyntactic sequences of adjoining roots, exactly as outlined in the cited work of Ray Jackendoff. If some of these pre-words are frequently uttered in close proximity – simply abutting each other, without forming hierarchical structures – they are likely to become loosely associated, as Jackendoff suggests. Such a pre-Merge system can in time lead to formalized relationships between pre-words, so that they start, on a case-by-case basis, to take on the properties of argument structure. For instance, a verb-like root could start to co-occur with a noun-like root as its complement, first very informally, then obligatorily. In this way, the verb-like element becomes a head. As supporting evidence, we already know that two adjoining flat structures can subsequently acquire a hierarchical subordination relationship from studies of the many attested historical changes from paratactic to hypotactic syntax (Heine & Kuteva 2007: ch. 5). In such cases, a formerly independent clause becomes a complement or adjunct to a main clause, via the usual processes of grammaticalization. If these speculations are on the right lines, then 'root-like elements' can become syntactic heads over time, and – rather surprisingly – Nóbrega & Miyagawa (2015) have in fact converged on the gradualist concept of a protolanguage.

References

Bickerton, D. (1990). *Language and species*. Chicago: University of Chicago Press.

Bickerton, D. (1995). *Language and human behavior*. Seattle: University of Washington Press.

Bickerton, D. (2009). *Adam's tongue: how humans made language, how language made humans*. New York: Hill & Wang.

Bolhuis, J. J., Okanoya, K. & Scharff, C. (2010). Twitter evolution: converging mechanisms in birdsong and human speech. *Nature Reviews Neuroscience, 11,* 747–759.

Carstairs-McCarthy, A. (2010). *The evolution of morphology*. Oxford: Oxford University Press.

Dediu, D. & Levinson, S. C. (2013). On the antiquity of language: the reinterpretation of Neandertal linguistic capacities and its consequences. *Frontiers in Psychology, 4*, 1–17.

Heine, B. & Kuteva, T. (2002). *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Heine, B. & Kuteva, T. (2007). *The genesis of grammar: a reconstruction*. Oxford: Oxford University Press.

Hurford, J. R. (2012). *The origins of grammar: language in the light of evolution*. Oxford: Oxford University Press.

Jackendoff, R. (2002). *Foundations of language: brain, meaning, grammar and evolution.* Oxford: Oxford University Press.

Jackendoff, R. & Wittenberg, E. (2014). What you can say without syntax: A hierarchy of grammatical complexity. In F. J. Newmeyer & L. B. Preston (Eds.), *Measuring grammatical complexity*, (pp. 65–82). Oxford: Oxford University Press.

Johansson, S. (2013). The talking Neandertals: what do fossils, genetics, and archaeology say? *Biolinguistics, 7,* 35–74.

Kipper, S., Mundry, R., Sommer, C., Hultsch, H. & Todt, D. (2006). Song repertoire size is correlated with body measures and arrival date in common nightingales, *Luscinia megarhynchos*. *Animal Behaviour, 71*, 211–217.

Miyagawa, S., Berwick, R. C. & Okanoya, K. (2013). The emergence of hierarchical structure in human language. *Frontiers in Psychology, 4*, 1–6.

Miyagawa, S., Ojima, S., Berwick, R. C. & Okanoya, K. (2014). The integration hypothesis of human language evolution and the nature of contemporary languages. *Frontiers in Psychology, 5*, 1–6.

Nóbrega, V. A. & Miyagawa, S. 2015. The precedence of syntax in the rapid emergence of human language in evolution as defined by the integration hypothesis. *Frontiers in Psychology, 6*, 1–8.

Kirby, S. (2000). Review of *Apes, language and the human mind* by S. Savage-Rumbaugh, S. G. Shanker & T. J. Taylor. *Journal of Linguistics, 36*, 190–195.

Ouattara, K., Lemasson, A. & Zuberbühler, K. (2009). Campbell's monkeys use affixation to alter call meaning. *PLOS One.* DOI: 10.1371/journal.pone.0007808

Savage-Rumbaugh, S., Shanker, S. G. & Taylor, T. J. (1998). *Apes, language and the human mind*. Oxford: Oxford University Press.

Slater, P. (2012). Bird song and language. In M. Tallerman & K. R. Gibson (Eds.), *The Oxford handbook of language evolution* (pp. 96–101). Oxford: Oxford University Press.

Tallerman, M. (2007). Did our ancestors speak a holistic protolanguage? *Lingua, 117*, 579–604.

Tallerman, M. (2014). No syntax saltation in language evolution. *Language Sciences, 46*, 207–219.

Thompson, N., LeDoux, K. & Moody, K. (1994). A system for describing bird song units. *Bioacoustics: The Journal of Animal Sound and its Recording, 5*, 267–279.

Zuberbühler, K. (2002). A syntactic rule in forest monkey communication. *Animal Behaviour, 63*, 293–299.