Luigi Rizzi*

Monkey morpho-syntax and merge-based systems

1. The target article makes a convincing case for the use of the sophisticated analytic tools developed in formal linguistics for the scientific study of monkey languages, with clear implications for the study of other animal communication systems. While the term "language" is construed very broadly, in a way reminiscent of the linguistic notion of e-language (a set of well-formed sentences), the focus really is on (monkey) i-language, the cognitive capacities underlying the call systems of the species under investigation.

In this brief commentary, I will focus on some (morpho-)syntactic aspects of the endeavor. The paper offers various kinds of evidence that at least rudimentary forms of sign combinations are found in monkey languages. Whether or not such forms of combinations can be seen as precursors of the massive combinatorial character of human language is left as a completely open question. The very cautious attitude adopted in the target paper seems to be well justified by the complexity of the issue and by the paucity of relevant evidence; at the same time, the paper forcefully argues for the use of tools developed for the study of human language syntax to highlight properties of monkey call systems. In the spirit of this approach, taking "merge" as the fundamental combinatorial principle operative in human languages, I would like to illustrate a typology of applications of merge based on the nature of the linguistic objects that are combined, and will try to use this typology as an interpretive grid to tease apart combinatorial forms which seem to be accessible to monkey languages, and those which are not.

After an overview of the findings on combinatorial properties in monkey systems, I will ask a few questions raised by the discovered properties with special reference to morpho-syntax. I will then turn to a typology of merge-based systems, and to the status of the observed call combinations with respect to it.

*Corresponding author: Luigi Rizzi, Department of Linguistics, University of Geneva, CH-1211 Genève 4, Switzerland; University of Siena, Siena, Italy, E-mail: luigi.rizzi@unige.ch

2. Campbell monkeys combine the sign -*oo* with two other signs of their repertoire: *krak* (leopard alert) and *hok* (eagle alert), with some noticeable properties:

–   -*oo* does not occur in isolation, but is necessarily combined in a fixed order with the other sign, which must precede it: so it looks like a bound morpheme, a suffix;
–   -*oo* is selective in its combinatorial ability, as it does not combine with another call (*boom*) of the Campbell repertoire;
–   -*oo* seems to have a discernible broadening meaning roughly akin to English – *ish* (*hok-oo*: an eagle-ish alert);
–   There appears to be a regional variation in the interpretation of *krak*, which could suggest a "learned" element in the system; but the variation could also follow from an interplay between pragmatic principles, and properties of the environment.

Female Diana monkeys also have complex calls formed by combinations of two individual calls from the set of calls {L, H, R, A}, which give rise to sequences LA, HA, RA. Here A could be ambiguous between an independent call (following its own occurrence rules), and a suffix attaching to an independent root, much as -*oo* in the male Campbell repertoire; alternatively, LA, HA, RA could be taken as "phrasal" combinations of two words.

A "phrasal" analysis is plausible in the case of male putty-nosed monkeys, whose inventory of alarm calls includes *Pyow* sequences, *Hack* sequences, and also combined *Pyow-Hack* sequences (actually a small and varying number of *Pyows* followed by a small and varying number of *Hacks*); the combined sequences appear to have a discernibly distinct meaning from simple *Pyow* sequences and simple *Hack* sequences, a meaning associated with group movement, rather than with leopard or eagle alerts. The authors discuss a possible non-compositional analysis of the sequence, according to which the combination would be a device to create a new sign whose meaning is not predictable from the meanings of the individual signs, and a compositional analysis, according to which the non-alert meaning of the composite sign follows from pragmatic principles and an inferential mechanism, leaving the issue open.

Black-and-White Colobus monkeys also appear to have combinatorial capacities, producing individual snorts, snort-roar sequences, and roar sequences, with distinct interpretive properties: individual snorts produced in the context of terrestrial animals, sequences of roars produced primarily in the context of aerial predators, and snort-roar sequences produced in every

context. Again, a compositional and a non-compositional analysis are compared and discussed.

Titi monkeys have a system involving A calls, B calls, and combined AB sequences, with distinct interpretations (their system also involves other calls). Again, it is not the case that "anything goes": order modifies interpretation, and not all imaginable orders are actually found.

3. Some questions are immediately raised by these findings:

(1)   All the examples of combinatorial capacities appear to involve combination of 2 calls, but not of 3, or more (meaningful two-call sequences can be repeated, but the number of repetitions does not seem to matter). Is this just an accident of the sample, or does it tell us something significant of the monkey capacity to combine calls?

(2)   Call order seems to matter. In some cases, this may be a matter of discourse organization and pragmatics, e.g., the effect of an Urgency Principle, determining first occurrence of a call providing critical information on a threat (as in one of the analyses proposed of *Pyow-Hack* sequences in putty-nosed monkeys). In other cases, the order of calls seems to be used in a way analogous to the morpho-syntactic function of word order in a phrase (e.g., AB sequences in Titi monkeys, or snort – roar sequences in Colobus) or morpheme order in a word (the suffixal analysis of *-oo* in Campbell monkeys) in human languages. Are these orderings the effect of articulatory constraints? Or do ordering effects reflect genuine morpho-syntactic choices?

(3)   Should one expect variants based on ordering, i.e. Titi "dialects" with BA sequences, or Colobus dialects with roar – snort sequences? Or variants exhibiting forms of prefixation, rather than suffixation? Or, put in a different way, if the repertoire of calls is innate for a given species, are properties of the sequences also biologically fixed, or do they exhibit elements of the freedom which characterizes human languages?

(4)   The issue of word order in human languages is intimately connected to the head-dependent distinction (whether one assumes a universal base hypothesis with ordering variation determined by movement, à la Kayne (1994), or one assumes head-dependent ordering parameters in syntax or at spell-out). And in fact, much of the morpho-syntax in human languages

is a matter of head-dependent relations. Pursuing the analogy, one important question to ask in monkey formal linguistics is therefore if any analogue of the head-dependent relation can be found. Maybe the evidence on sequences is too limited to meaningfully address the point, but the question is worth asking if the analogy is to be pursued.

(5)   In human languages, the distinction and division of labor between functional and contentive lexicon plays a critical role in syntax, where functional elements trigger major syntactic operations, and are the locus of parameters expressing cross-linguistic variation (Rizzi and Cinque 2016). Is there any trace of this sort of distinction in monkey languages? Perhaps the size of the lexicon is too small to meaningfully investigate this question, even though the proposal that some monkey calls may function as "affixes" is suggestive of a possible division of labor.

(6)   Word order errors and other forms of morpho-syntactic ill-formedness trigger well studied brain responses in humans, e.g. in ERP studies (Grodzinsky and Friederici 2006, and much related work). What should one expect for call sequencing errors in monkey formal linguistics? This raises the broader issue of monkey neurolinguistics: what brain circuits are involved in the calling systems in monkeys?

4. I will assume that a language consists of a lexicon, formed by a finite and possibly small number of items, and a combinatorial device:

(7)   a. Lexicon: a finite list of items A, B, C,...
       b. Combinatorial device: Take two linguistic expressions X and Y and form
          the complex expression [X Y]

where "linguistic expressions" are lexical items and complex entities already created by the combinatorial device. I will use the term "merge" to refer to (7)b, even though other definitions of merge are possible (in the definition in Chomsky 1995 and much subsequent work, merge is a pure set theoretic operation creating the set {X, Y}, and a distinct operation of linearization is needed to stipulate the linear order X>Y).

   Given systems based on (7), we can identify a hierarchy of levels of complexity for such systems, ranked in terms of their generative capacity and of the computational resources they need:

0-merge systems: such systems do not use (7)b at all, so that the only possible linguistic expressions are the individual items taken from the lexicon. The sentence types coincide with the words listed in the lexicon.

Required devices:
– the lexicon

1-merge systems, or word – word merge systems: merge can apply, forming two-word expressions, but then the system stops, i. e., it lacks recursive procedures. Here merge increases the class of linguistic expressions beyond the lexical inventory, but cannot generate any expression consisting of more than two items.

Required devices:
– the lexicon
– merge, as in (7)b

2-merge systems, permitting word – word merge, and also word – phrase merge. This system requires a lexicon and a temporary work-space, a memory storage; word – phrase merge takes an already formed phrase from the temporary storage and a word from the lexicon, forming a more complex phrase. The system is recursive, hence it can generate a potentially unlimited set of linguistic expressions.

Required devices:
– the lexicon
– merge
– a temporary work-space

3-merge systems, permitting word – word merge, word – phrase merge, and also phrase – phrase merge. The latter requires two separate temporary memory storage spaces, hosting phrases of arbitrary complexity, which can be merged together. This is the kind of merge involved in the formation of structures with complex specifiers: e. g., the merger of a subject and a predicate, both of which can be of arbitrary complexity, is an instance of phrase – phrase merge. For instance, the merger of a subject like *[the men you told me about]* and a predicate like *[know how to address the problem]*: each complex entity is formed by several applications of different types of merge but, if syntactic structures are binary (Kayne 1983), no more than two temporary storage spaces are needed for each application of phrase-phrase merge. I do not discuss here how the temporary work-spaces relate to phases in a phase-based architecture.

Required devices:
– the lexicon
– merge

-   a temporary work-space
-   a second temporary work-space.

Merge is generally considered a unitary phenomenon, and in fact human lan-
guages manifest the full power of 3-merge systems: no human language is
limited to using single words (0-merge), or just two-word sequences (1-merge),
or to disallow complex specifiers (2-merge). I.e., on the latter point, a 2-merge
system would only permit external arguments consisting of one word like
[he [will [meet [the girl]]]], but not of two words like *[[the boy] [will [meet [the
girl]]]]* (a structure which would require the power of a 3-merge system): no
human language appears to have this limitation and disallow complex speci-
fiers. So, known human languages clearly utilize all the cases of the typology.

Nevertheless, the typology of merge clearly yields a hierarchy of cases
involving an increasing complexity, defined by the use of richer computational
resources. Moreover, it is required for theory internal technical issues, such as
the functioning of the labeling algorithm (Chomsky 2013; Rizzi 2016). So, I think
it makes sense to use this hierarchy to evaluate other combinatorial systems,
such as the limited combinatorial options shown in monkeys' call systems.

Where would such systems be placed in the hierarchy? The target article
seems to provide clear evidence that monkey call systems are more powerful
than 0-merge systems, as some combinatorial capacities are attested. If the
limitations to combinations of two calls is strict, and not an accident
of the sample of cases considered, this would correspond to the properties of
1-merge systems.

Such systems, in the simplest form, only permit combinations of two items. It
is conceivable that more 1-merge systems may work in cascades, thus permitting
combinations of more than two items, while only generating a finite (and small)
set of expressions: for instance, one may think of the hierarchical structure of the
syllable as arising from the operation of two such systems: a nucleus is merged
with a coda to determine a rhyme, and an onset is merged with a rhyme, to
determine a syllable, thus giving rise to hierarchically organized structures of
three elements *[$_{Syllable}$ Onset [$_{Rhyme}$ Nucleus Coda]]*. Here, two 1-merge devices
combine to give rise to expressions of three elements, but the overall system is
non-recursive, as it cannot reapply to its own output (i. e., the onset and the coda
of a syllable are not themselves syllables, while in syntax proper, the specifier and
the complement of a phrase are phrases, hence the syntactic system is recursive).
The functioning in cascades can be multiplied, e. g., giving rise to combinations of
four element, etc., but again only a finite number of expressions would be
generated. Going back to the limited combinatorial properties of monkey

languages, they appear to stop at the level of 1-merge systems, without cascades and recursive applications, so that the generated languages are restricted to sequences of two calls.

References

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

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49.

Grodzinsky, Yosef & Angela Friederici. 2006. Neuroimaging of syntax and syntactic processing. *Current Opinion in Neurobiology* 16. 240–246.

Kayne, Richard. 1983. *Connectedness and binary branching*. Dordrecht: Foris Publications.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Rizzi, Luigi. 2016. Labeling, maximality, and the head – phrase distinction. *The Linguistic Review* 33. 103–127.

Rizzi, Luigi & Guglielmo Cinque. 2016. Functional categories and syntactic theory. *Annual Review of Linguistics* 2. 139–163.