On the *-oo* suffix of Campbell's monkeys

Jeremy Kuhn[1], Sumir Keenan[2,3], Kate Arnold[3], Alban Lemasson[4]

1. *Institut Jean Nicod, Département d'études cognitives, ENS, EHESS, CNRS, PSL Research University, 75005 Paris, France*
2. *Université de Lyon/Saint-Etienne, Equipe Neuro-Ethologie Sensorielle, ENES/Neuro-PSI, CNRS UMR 9197, Saint-Etienne, France*
3. *School of Psychology and Neuroscience, University of St Andrews, Scotland, UK*
4. *Université de Rennes 1, Ethologie animale et humaine (UMR 6552 – CNRS – Université Caen Normandie), Station Biologique, Paimpont, France*

1 Introduction

1.1 Primate morphology?

Ouattara et al. (2009a,b) make the novel claim that Campbell's monkey alarm calls demonstrate a simple pattern of linguistic morphology. The authors observe that there are at least two distinct alarm calls (called *krak* and *hok*) that are used in two different predatory contexts, and that each may be followed by a low frequency sound (called *-oo*) that alters the meaning of both calls in predictable ways, allowing contexts with reduced level of threat. In light of these facts, *-oo* is analyzed as a meaning-bearing, combinatorial unit.

However, the claim that a non-human communication system has a combinatorial system (however primitive) is rare in the literature (see §5 for related patterns), and, indeed, is antithetical to certain claims that structural hierarchy is unique to human language (e.g., Bolhuis et al. 2014). Moreover, it has been noted (Schlenker et al. 2014) that there is redundancy between the apparent semantic contribution of *-oo* and the semantic contribution of a variety of other signal manipulations (e.g. calling rate) that are easiest to explain via non-compositional mechanisms. These facts warrant particular caution when evaluating the pattern as a possible counterexample to generalizations about human language.

Thus, in this squib, we examine the compositional hypothesis further. As counterpoint, we consider a class of more conservative hypotheses in which *-oo* does not itself bear meaning, but instead arises as the side effect of other articulatory processes that non-compositionally affect call meaning. Key to such hypotheses is the premise that *-oo* is articulatorily parasitic on another phonetic process. A major contribution of this squib is thus phonetic: considering the acoustic properties of *-oo*, we conclude that complex calls (*krakoo* and *hokoo*) are produced with a two pulses of a single breath-group. Critically, the production of these complex calls requires an additional articulatory gesture and thus an increase in articulatory effort. An increase in articulatory effort would not be expected on an analysis in which *-oo* arises as a phonetic side effect; we accordingly reject these alternate hypotheses, thus strengthening the robustness of the combinatorial analysis.

1.2 ‘*Merge*’ as the putative defining feature of human language

Bolhuis et al. 2014, following Chomsky (2000), defend the strong hypothesis that the distinguishing feature of human language is the presence of hierarchical syntactic structure. In their words, “human language syntax can be characterized via a single operation that takes exactly two (syntactic) elements a and b and puts them together to form the set $\{a, b\}$.” This operation, called *merge* in the Minimalist tradition (Chomsky 2000), allows two elements that are themselves syntactic units to be combined into a complex unit that can serve as the input to another combinatory operation. In human language, this second operation might be a further application of *merge*, thus recursively generating structures of arbitrary length.

Of course, the presence of *merge* does not guarantee the existence of arbitrarily long sequences; note, for example, that the phrase structure grammar with the terminals $\{$D, N, V$\}$ and the rules $\{$S→NP VP, NP→D N, VP→V NP$\}$ produces sentences with hierarchical structure, but only generates five-word strings. Relatedly, Rizzi (2016) observes that recursive applications of *merge* depend on the presence of a ‘temporary work-space,’ short-term memory storage for non-lexical inputs to *merge*. Without this workspace, a system can produce binary strings of lexical elements, but cannot store these units for further applications of *merge*. For Bolhuis et al. (2014), all non-human animal communication systems disallow hierarchy of any depth. As indicated above, the alarm calls of Campbell’s monkeys pose a potential counterexample; this is thus the question that we address here.[1]

2 Male Campbell’s monkey alarm calls

2.1 Complex calls

Male Campbell’s monkeys (*Cercopithecus campbelli*) produce at least three distinguishable alarm call stems (i.e. calls not followed by *-oo*), called *krak*, *hok*, and *boom*, classifiable both by ear and automatically (Ouattara et al. 2009a,b, Keenan et al. 2013).[2] The *boom* call is is unique in several respects: it only appears at the beginning of a call sequence, there is visible use of superlaryngeal air sacs, it is never suffixed by *-oo*, and it signals the presence of a non-predatory context; we thus set it aside. Both of the remaining two calls may appear in isolation (‘simple calls’: *krak* and *hok*) or followed by the *-oo* suffix (‘complex calls’: *krakoo*, *hokoo*). The *-oo* particle never appears in isolation.

Critically, Ouattara et al. (2009a) observe that the addition of *-oo* to a base call alters the meaning in a systematic way, acting to attenuate the force of the call. In their data, *hok* only appears in the presence of eagles (predatory disturbances in the canopy); *hokoo*,

---

[1] Although Ouattara et al. (2009b) call *-oo* a ‘suffix’ to characterize the call as a minimal meaning-bearing combinatorial unit, this use of terminology should not be interpreted here as committing to any deeper analogy with spoken language, such as postulating *-oo* as a sub-lexical morpheme versus a sentence-final particle. Certainly, either of these phenomena from human language involve syntactic composition; the question here is whether *-oo* does, too.

[2] Ouattara et al. (2009a,b) additionally identify a stem *wak*, but Keenan et al. (2013) provide evidence that it is a variant of the *hok* call.

too, appears in eagle contexts, but also in contexts of inter-group interaction (non-predatory disturbances in the canopy). *Krak* only appears in the presence of leopards (predatory disturbances on the ground); *krakoo*, too, appears in leopard contexts, but also in reaction to tree falls, inter-group interaction, and eagles. Schlenker et al. (2014) refine these generalizations with further data. While *hok* is associated with eagles and *krak* with leopards, the association is weaker for *krak* than for *hok*. Further, on Tiwai island, which has no leopards, *krak* is used as a general alarm call, including in eagle contexts. For both calls, the complex form is used more widely than the corresponding simple call, including in non-predatory contexts. These observations motivate an analysis in which the meanings of the complex calls *krakoo* and *hokoo* are compositionally derived. The stem communicates locational information (for Schlenker et al., *hok* indicates an 'upwards' disturbance; *krak* is locationally unspecified); the presence of *-oo* adds information regarding the level of threat.

That *-oo* compositionally modulates threat-level is confirmed by the reaction of con- and heterospecifics to natural and artificial stimuli. Ouattara et al. (2009a) report that for Diana monkeys (which associate with Campbell's monkeys), anti-predatory behavior occurs only in response to simple calls. This was confirmed experimentally by Coye et al. (2015), who played back recordings of Campbell's *krak* and *krakoo* calls to groups of Diana monkeys. Both male and female Diana monkeys produced more alarm calls in response to *krak* than to *krakoo* sequences. These results held even for calls that were artificially created by either adding *-oo* to *krak* calls or by removing *-oo* from *krakoo* calls.

2.2 Conjunctive meaning

Even if complex call meanings are compositional, we should ask whether this composition requires anything beyond conjunction. Notably, even in a system without *merge*, if call meanings update an overall information state, the effect is equivalent to the conjunction of the individual calls. On the other hand, any other way of combining meanings requires some kind of function application. Thus, if call combination is found to be non-conjunctive, then syntactic combination is a done deal: the semantic facts alone would be evidence for *merge*. On the other hand, if call combination is conjunctive, the need for *merge* must be decided based on other facts.

In the case at hand, a conjunctive analysis initially appears not to be viable: of note, as discussed above, Ouattara et al. (2009a, b) show that simple calls occur in a *subset* of the situations where their corresponding complex calls occur. Conjunction can only restrict a meaning; thus the fact that *-oo broadens* the use of the call suggests that the meaning of *-oo* must be non-restrictive and thus non-conjunctive. However, Schlenker et al. (2014) show that other 'pragmatic' factors complicate the picture. First, they end up concluding that there is an 'alarm parameter' that decreases over time. Thus, the reason why *hokoos* appear in all the same situations where *hok* appears is because—after repeating *hok* for a period of time—the degree of alarm decreases to a sufficiently low level for *hokoo* to be used. Distribution of simple and complex calls supports this hypothesis; in the data from Keenan et al. (2013) (3344 total calls), in sequences that have both *hoks* and *hokoos*, an average of

87.5% of *hoks* appear before the majority of co-sequential *hokoos*. (A similar trend holds for *krak*/*krakoo*.) Second, Schlenker et al. propose that there is competition between call types, akin to scalar implicatures in spoken language. This provides an explanation for why *krak* does not generally appear in situations where *krakoo* would be a more precise call.

In the end, Schlenker et al. (2014) are led to a proposal on which the contribution of *-oo* is restrictive but, for technical reasons, nevertheless not conjunctive (see discussion under their (59)). Their final definition states that for any root R, R*-oo* is used for weak R-type disturbances. Although the definition provided by Schlenker et al. (2014) is not technically conjunctive[3], we consider the situation to be sufficiently unresolved that we cannot conclude based on semantics alone that these calls present a case of syntactic *merge*.

2.3 A single combinatory unit

Regardless whether the semantics is conjunctive or not, facts about timing and distribution nevertheless provide strong evidence that complex calls act as single units that serve as the input for further combinatory processes. In particular, both simple and complex calls are organized into call sequences; in the data from Keenan et al. (2013), sequences have a median of 31 calls, and a maximum of 131 calls. In the same data, an average of 4.60s separate the onset of one call from the onset of the next. Call stems are themselves are an average of 0.13s. In this context, *-oo* shows a strikingly different distribution and timing: *-oo* (average duration 0.093s) always occurs immediately following a call stem, separated only by a short pause averaging 0.060s (Ouattara et al. 2009).

Furthermore, for both simple and complex calls, calls are most commonly found in sequences surrounded by the same call-type. Figure 1 provides the O/E (observed over expected frequencies) for each bigram in the data from Keenan et al. (2013) (total counts: *krak*, 479; *hok*, 421; *krakoo*, 1582; *hokoo*, 862). Values greater than one along the diagonal show that repetition of the same call is more likely than chance for all call types. Naturally, the 'grammar' that derives these sequences of calls will look dramatically different from the grammars of human languages, and there appears to be no motivation to posit sequence generation via *merge*. Nevertheless, even if the system that generates these sequences is a probabilistic model conditioned only on the context of utterance, the difference in co-occurence frequencies between simple and complex calls can only be stated by reference to the complex calls *krakoo* and *hokoo* as themselves combinatorial units.

Thus, both with respect to timing and co-occurence frequency, the complex calls *krakoo* and *hokoo* behave as though they are single calls. In conjunction with the semantic facts motivating decomposition of these calls, we thus have a pattern that appears to exemplify the simplest case of *merge*: two units combining to form one complex unit.

---

[3] The definition of *-oo* is non-conjunctive in the same way that the English adjective *tall* is non-conjunctive, since both must be evaluated with respect to a comparison class; '*tall for a six-year-old*' is different from '*tall for a basketball player*.'

	___ Krak	___ Hok	___ Krakoo	___ Hokoo
Krak ___	6.42	0.28	0.32	0.03
Hok ___	0.21	4.84	0.07	1.00
Krakoo ___	0.32	0.07	1.80	0.31
Hokoo ___	0.12	0.88	0.36	2.74

Figure 1: O/E of bigram frequencies

2.4 Regarding the 'holistic' hypothesis

At this point, there is nevertheless another, entirely non-compositional hypothesis that is perfectly compatible with the data: namely, that all four forms (*krak*, *hok*, *krakoo*, and *hokoo*) are holistically memorized as atomic units. These four forms could be given exactly the same meanings as those derived by Schlenker et al. (2014), which, in conjunction with Schlenker et al.'s pragmatic analysis, will generate identical results. On this analysis, there is no need for *merge*; *-oo* would be no more of a syntactic unit than the 'cat' of 'catapult.'

In fact, such an analysis can be posited for any system that generates a finite set of forms, be it the four-form inventory of Campbell's monkeys, or the set of five-word strings of §1.2. In any such case, one cannot falsify a holistic analysis based on form-meaning pairings, as the memorization hypothesis is strictly weaker than the compositional alternative. In the general case, several options can mediate between these hypotheses, but these prove difficult to implement in the case at hand. For example, one can test whether a rule generalizes to a novel form (a 'wug' test). For Campbell's monkeys, though, no such data presently exists, due to the prohibitive difficulty of training a group of monkeys to react to a novel call. More feasibly, one can compare the 'syntactic diversity' of a set of forms to a model on which compositional parts combine independently and interchangeably. Yang (2013) shows that such a model generates a very close fit of the linguistic systems of human adults and children, and notably does not fit the attested productions of the language-trained chimpanzee Nim Chimpsky. In the case at hand, though, this analytic method is confounded by the smallness of the data set (four forms), as well as the fact the contexts that gave rise to calls were often induced by researchers, and thus controlled for frequency.[4]

In light of these challenges, we will not try to put to rest the holistic memorization hypothesis, acknowledging that it is indeed a viable alternative to the compositional theory. On the other hand, we note that what is *lost* on the holistic hypothesis is the semantic connection between *hok* and *hokoo* on one hand and *krak* and *krakoo* on the other; for example, if each form is memorized independently, there is no principled reason why *hok* and *hokoo* should both relate to aerial disturbances. Thus, our goal here will be to consider a second

---

[4] Despite these limitations, we can of course still calculate the relevant values. If stems and *-oo* combine independently and interchangeably, we use the Keenan et al. (2013) data to calculate the expected numbers of calls using the product of the marginal probabilities as *krak*, 555; *hok*, 345; *krakoo*, 1506; *hokoo*, 938. (E.g., the expected count of *krakoo* is $(\text{krak} + \text{krakoo}) \times (\text{krakoo} + \text{hokoo}) / \text{total}$.) These are not far off from the attested values of 479, 421, 1582, 862, which is consistent with the combinatorial story, though the relevance of this result is mitigated by the issues discussed above.

class of non-compositional hypotheses: namely, that *-oo does* systematically modify the call meaning (thus capturing the relation between simple and complex forms), but that it does so via a non-compositional mechanism. This is spelled out in the following sections.

3 Non-compositional modification?

3.1 What does it mean to bear meaning?

The analysis of compositionality in §2.1 is predicated on the assumption that *-oo* itself bears a meaning. For human language, we can say that a morpheme bears a certain meaning if it contributes a stable semantic contribution in all contexts; semantic judgements can be gathered from intuitions of native speakers. For primates, conclusions must be drawn from indirect evidence; as we have already seen, this can include both the context of use and the response of conspecifics or heterospecifics to recordings of the signal in question. Together, these demonstrate that a certain proximate factor is responsible for the signal, and that other animals can interpret the signal in order to react appropriately.

However, these diagnostics cannot determine whether the segment itself bears the meaning or whether the meaning is inferred indirectly. To illustrate this point, we can look to cases of 'paralinguistic' meaning in human speech. Consider, for example, [+excited], a non-concatenative modification of the intensity, pitch range, and speed of an acoustic signal, which combines productively with any utterance and adds the (presupposed) semantic content that the speaker is excited. As with monkey alarm calls, this meaning can be deduced from the context of use (heightened emotional state) and from the reactions of conspecifics to the signal manipulation ("Calm down!"). Intuitively, though, this inference is quite different in origin from the semantic contribution of combinatorial morphemes; whereas morphemes bear meaning themselves, the paralinguistic modification results from the way that the context (the emotional state) directly affects articulation. (In light of §2.2, it bears noting that the meaning of [+excited] is semantically conjunctive.)

In human communication more generally, the phonetic properties of speech have been shown to vary with respect to communicational and situational demands (Picheny et al. 1985). Lindblom (1990) describes principles governing these phonetic adaptations in terms of trade-offs: 'hyperarticulated' speech is used to facilitate perception in contexts in which communication is harder or more important (e.g., slow and clear speech in a loud environment); otherwise, when perceptual demands are less severe, speech defaults to an articulatorily easier form. Exactly analogous kinds of patterns have been shown to hold for non-human communication; for example, Candiotti et al. (2012b) show that female Diana monkey contact calls show greater inter-individual acoustic distinctiveness in dark environments (where caller identification relies on sound) than in bright environments.

3.2 Non-compositional modification for Campbell's monkeys

There are independent reasons to think that threat-level impacts the form of Campbell's monkey alarm calls in a non-compositional manner. As we have seen, the presence of *-oo* is associated with contexts with decreased levels of threat. Additionally, though, the level of threat influences Campbell's monkey calls in a variety of other ways. First, Lemasson et al. (2010) show that low-threat contexts are correlated with a slower calling rate. Second, Keenan et al. (2013) show that both *hok* and *krak* calls can be divided into sub-types; the less phonetically stereotyped version of each form is correlated with low-threat contexts.

In both of these cases, compositional analyses are difficult to implement. For call rate, the modification applies to a global property of a call sequence. For call distinctiveness, the modification is most easily stated in terms of the phonetic similarity between multiple lexical items. Neither of these situations is conducive to an explanation in terms of local composition. On the other hand, both the variable calling rate and the acoustic variance can be given a simple non-compositional explanation based on environmental-level factors. On a sequence level, increased calling rate may track emotional state, and increases the redundancy of a signal. On a call level, increased distinctiveness between call types maximizes discriminability, so reduces the chance of communicating the wrong signal in high-threat contexts, where ambiguity can be fatal (Cheney and Seyfarth 1990; see also Arnold and Zuberbühler 2013). Notably, these patterns fit in neatly with the trade-offs discussed in §3.1; high-threat environments, where communication is more important, induce signals that are perceptually clearer, but that are articulatorily harder to produce.

Given that the meaning contributed by *-oo* may also be expressed through non-compositional mechanisms, we may well ask whether *-oo* itself should be analyzed in non-compositional terms.[5] On such a hypothesis, *krakoo* and *hokoo* are simply phonetic variants of *krak* and *hok*. Since *krakoo* and *hokoo* would then be syntactically atomic, there would be no need for *merge*; the hypothesis thus presents a more conservative alternative to the compositional analysis of Ouattara et al. (2009a,b) and Schlenker et al. (2014).

If this is indeed the case, we would then expect *-oo* to adhere to the same principles of communication as the other non-compositional indicators of threat level. In particular, we have seen that *high*-threat contexts induce clearer and faster signals, at the cost of greater articulatory effort. If the presence of *-oo* in *low*-threat environments arises from the same principles, we make two predictions: first, if *-oo* alters the signal perceptually, it should do so in the opposite direction: a slower or less clear signal; second, being the unmarked form, calls with *-oo* should be articulatorily easier to produce than calls without it. In what follows, we will see that these predictions are not borne out: *-oo* has no effect on perceptual properties, and in fact requires *increased* articulatory effort. These results provide grounds to reject the articulatory hypothesis, thus providing support for a morphological analysis.

---

[5] Importantly, though, this redundancy with non-compositional mechanisms does not *necessitate* a non-compositional analysis for *-oo*. After all, even in human language, discourse particles and expressives may express content that can equally well be communicated non-compositionally. For example, the emotive content of *fucking* in the sentence '*I'm going to the fucking store*,' will often be redundant with the semantic content communicated non-compositionally by the tone of voice in which the sentence is uttered.

3.3 Perceptual effects of *-oo*

In principle, *-oo* could affect temporal properties of a call sequence; for example, the time it takes to enunciate *-oo* could have the direct effect of slowing down the calling rate. However, this hypothesis is implausible based on the durations involved. The shortest average time between calls reported by Lemasson et al. (2010) is roughly two seconds (in visual eagle scenarios), and ranges up to about six seconds. The smallest significant difference between threats of different levels is approximately one second. In contrast, the average length of the *-oo* suffix is less than a tenth of a second (Keenan et al. 2013). Thus, the amount of time that it takes to enunciate *-oo* is sufficiently small that its addition alone would not alter the rate of call enough to have an effect on the meaning.

Alternatively, *-oo* could affect distinguishability via an acoustic effect on the call stem. Just as co-articulation of an English vowel with a following nasal results in a reduced vowel space (Wright 1986), if Campbell's monkey calls include an *-oo* suffix, then overlap of articulatory gestures could plausibly result in a diminished formant space. However, this possibility, too, is not borne out. Keenan et al. (2013) show that the semantic effect of acoustic sub-type can be dissociated from the presence of the *-oo* suffix: holding stem sub-type constant, both *krak* variants are used more frequently in response to direct observation of a predator; *krakoo* forms are used more frequently in response to another monkey's predator call. This hypothesis is further falsified by the playback experiments of Coye et al. (2015), in which Diana monkeys showed differential behavior to *krak* and *krakoo*, even when these stimuli were artificially constructed from the stems of the other call.

In summary, *-oo* tracks the threat-level of the context, independent of any affect on the call sequence or call stem. There is thus no evidence that the presence of *-oo* affects other perceptual properties of call sequences.

3.4 The articulatory production of *-oo*

Finally, we turn to the articulation of *-oo*. As discussed above, if *-oo* gives rise to meaning indirectly, we would expect it to be associated with a decrease in articulatory effort. This is not borne out. To analyze the articulation of *-oo*, two acoustic facts are of particular import. First, the *-oo* suffix is always separated from the stem by a brief pause, of an average of 0.060s (Ouattara et al. 2009). Second, *-oo* is characterized by a low-pitch band with no higher frequency bands. Figure 2 provides an example spectogram of *krakoo*.

Several sources could account for the pause in phonation: (a) a laryngeal gesture that allows a moment of voicelessness during continued airflow; (b) stopping the airflow by obstruction (as for plosives in human speech); (c) stopping the airflow by a pulmonary gesture. Are any of these explanations compatible with decreased articulatory effort? Possibilities (a) and (b) are not; both require an an additional articulatory gesture that would not naturally occur otherwise. Possibility (c), on the other hand, could result from decreased articulatory effort if the *-oo* suffix is produced via inspiration. Like a human hiccup, the moment at which airflow changes direction would be accompanied by a brief pause of phonation, and continued phonation during inspiration could plausibly result from reduced

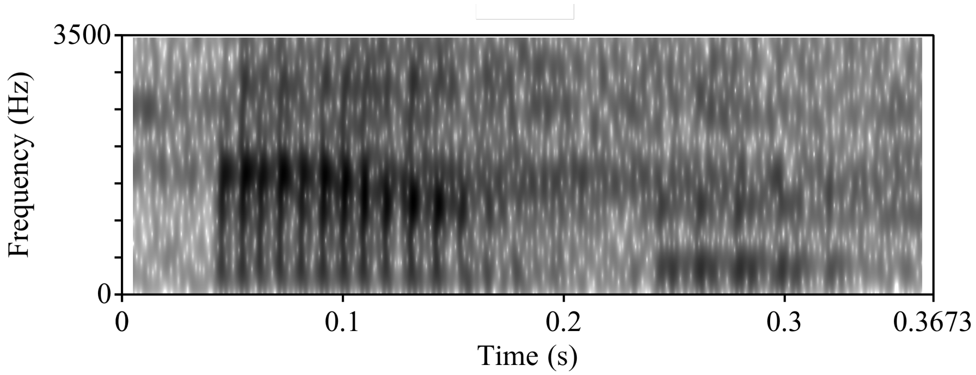

Figure 2: Suspended phonation between *krak* stem and *-oo*. (Recording by KA in Tiwai.)

attention to articulation. The hypothesis that some guenon vocalizations may be produced via inspiration has also been suggested previously by Riede and Zuberbühler (2003) for Diana monkeys. However, this hypothesis doesn't square with the acoustic facts. Eklund (2008), in a review of ingressives in both human and animal sound production, characterizes ingressives as acoustically noisier and less cyclic than their egressive counterparts. Here, *-oo* is not notably noisier than the stem, and has a cyclic period that is visibly distinguishable in both the waveform and spectrogram. Moreover, if the pause is produced entirely with the source of phonation (with no additional articulatory gesture), then the formants produced by inspiration should be identical to those produced by expiration, since the vocal tract filtering the call remains the same. This is not the case; the spectral bands change dramatically from stem to *-oo*. We conclude that *-oo* is not produced via ingression; the production of the call stem and the *-oo* suffix are produced as part of a single breath-group. Importantly, if the stem and *-oo* are two pulses of a single breath-group, the pause in airflow between the two must result from an additional articulatory gesture; this additional gesture requires an increase in articulatory effort. This fact is at odds with any theory in which *-oo* is an articulatory side-effect.

4 Discussion

The question of whether and how animal calls "bear meaning" has been discussed in both the biological literature (Seyfarth and Cheney 1990) and the philosophical literature (Grice 1957, Quine 1973). The present paper extends this discussion to a possible case of hierarchical composition within a non-human primate (Ouattara et al. 2009a,b). The topic is of particular importance to recent claims (e.g. Bolhuis et al. 2014) that hierarchical structure is the defining characteristic of human language. We investigated the hypothesis that *-oo* is a combinatorial, meaning-bearing unit, using as counterpoint the hypothesis that the use of *-oo* arises indirectly from articulatory mechanisms. In this spirit, we discussed both the contextual factors that influence call articulation and the articulation of *-oo* itself.

Ultimately, we were able to reject a class of hypotheses in which *-oo* is a side-effect of articulation, thus bolstering the hypothesis that *-oo* itself carries semantic content.

The argumentation developed here is useful for further investigations into the evolution of hierarchical compositionality in human language, especially as more repertoires of acoustically complex calls are being described for non-human primates in the literature: Bouchet et al. (2010) on the "Uh" unit in mangabeys; Candiotti et al. (2012a) on the "A-calls" of female Diana monkeys; Bouchet et al. (2012a) on the "I" unit in De Brazza's Monkey; Arnold and Zuberbühler (2006) on the "pyow-hack" sequences of putty-nosed monkeys; and Bene et al. (2012) on various calls among Colobus monkeys. While these patterns are a far cry from the complex combinatory processes of human language, detailed examination of them—especially informed by modern linguistic theory—promises to offer insight into the evolution of syntactic and semantic composition in natural language.

References

Arnold, K., & Zuberbühler, K. (2006). Semantic combinations in primate calls. *Nature*, *441*.

Arnold, K., & Zuberbühler, K. (2013). Female putty-nosed monkeys use experimentally altered contextual information to disambiguate the cause of male alarm calls. *PLoS ONE*, *8*.

Bene, J.-C., Ouattara, K., Bitty, E. A., & Kone, I. (2012). Combination calls in Olive Colobus monkeys (Procolobus Verus) in Tai National Park, Côte d'Ivoire. *Journal of Asian Scientific Research*, *2*(9).

Bolhuis, J. J., Tattersall, I., Chomsky, N., & Berwick, R. C. (2014). How could language have evolved? *PLoS Biology*, *12*(8).

Bouchet, H., Blois-Heulin, C., & Lemasson, A. (2012). Age- and sex-specific patterns of vocal behavior in De Brazza's Monkeys (Cercopithecus neglectus). *American Journal of Primatology*, *74*(12).

Bouchet, H., Pellier, A.-S., Blois-Heulin, C., & Lemasson, A. (2010). Sex differences in the vocal repertoire of adult red-capped mangabeys (cercocebus torquatus): A multi-level acoustic analysis. *American Journal of Primatology*, *72*, 360–375.

Candiotti, A., Zuberbühler, K., & Lemasson, A. (2012a). Context-related call combinations in female Diana monkeys. *Animal Cognition*, *15*, 327–339.

Candiotti, A., Zuberbühler, K., & Lemasson, A. (2012b). Convergence and divergence in diana monkey vocalizations. *Biology Letters*, *8*.

Cheney, D., & Seyfarth, R. (1990). *How monkeys see the world: Inside the mind of another species*. University of Chicago Press.

Chomsky, N. (2000). Minimalist inquiries: the framework. In R. Martin, D. Michaels, & J. Uriagereka (Eds.) *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, (pp. 89–155). Cambridge, MA: MIT Press.

Coye, C., Ouattara, K., Zuberbühler, K., & Lemasson, A. (2015). Suffixation influences receivers' behaviour in non-human primates. *Proceedings of the Royal Society B*, *282*(1807), 20150265.

Eklund, R. (2008). Pulmonic ingressive phonation: Diachronic and synchronic characteristics, distribution and function in animal and human sound production and in human speech. *Journal of the International Phonetic Association*, *38*(03), 235–324.

Grice, H. P. (1957). Meaning. *The Philosophical Review*, *66*(3).

Keenan, S., Lemasson, A., & Zuberbühler, K. (2013). Graded or discrete? a quantitative analysis of campbell's monkey alarm calls. *Animal Behavior*, *85*, 109–118.

Lemasson, A., Ouattara, K., Bouchet, H., & Zuberbühler, K. (2010). Speed of call delivery is related to context and caller identity in campbell's monkey males. *Naturwissenschaften*, *97*, 1023–1027.

Lindblom, B. (1990). Explaining phonetic variation: a sketch of the h & h theory. In W. J. Hardcastle, & A. Marchal (Eds.) *Speech Production and Speech Modelling.*, (pp. 403–439). Kluwer Academic Publishers.

Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009a). Campbell's monkeys concatenate vocalizations into context-specific call sequences. *Proceedings of the National Academy of Sciences*, *106*(51), 22026–22031.

Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009b). Campbell's monkeys use affixation to alter call meaning. *PLoS ONE*, *4*(11).

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing ii: Acoustic characteristics of clear and conversational speech. *Journal of Speech, Language, and Hearing Research*, *29*, 434–446.

Quine, W. (1973). On the reasons for the indeterminacy of translation. *Journal of Philosophy*, *12*.

Riede, T., & Zuberbühler, K. (2003). The relationship between acoustic structure and semantic information in diana monkey alarm vocalization. *Journal of the Acoustical Society of America*, *114*(2).

Rizzi, L. (2016). Monkey morpho-syntax and merge-based systems. *Theoretical Linguistics*, *42*(1-2), 139–145.

Schlenker, P., Chemla, E., Arnold, K., Lemasson, A., Ouattara, K., Keenan, S., Stephan, C., Ryder, R., & Zuberbühler, K. (2014). Monkey semantics: Two 'dialects' of campbell's monkey alarm calls. *Linguistics and Philosophy*, *37*(6), 439–501.

Wright, J. T. (1986). The behavior of nasalized vowels in the perceptual vowel space. In J. J. Ohala, & J. J. Jaeger (Eds.) *Experimental phonology*, (pp. 45–67). Orlando, FL: Academic Press.

Yang, C. (2013). Ontogeny and phylogeny of language. *Proceedings of the National Academy of Sciences*, *110*(16), 6324–6327.