DE GRUYTER MOUTON Theoretical Linguistics 2016; 42(1-2): 159–171

Shane Steinert-Threlkeld*

Compositionality and competition in monkey alert calls

DOI 10.1515/tl-2016-0009

In their paper, Schlenker et al. (2016b) analyse a wealth of data on monkey alert calls using the tools of formal semantics and pragmatics. Such a pursuit promises to be extremely valuable. By figuring out exactly how such animal communication systems work, we may be able to shed light on evolutionary precursors of human language. While the field is too nascent to say anything concrete on this front, the paper hopes to show the promise of using these methods. In my comments, I want to carefully examine whether the methods are in fact appropriate and what justifies their use. First, I will briefly review scalar implicatures because a similar mechanism plays a critical role in all of their case studies. Then, I will closely look at two of the case studies and raise some questions in each case about the theoretical roles that sentences and compositional interpretation thereof play. Finally, I will conclude with some general worries about the use of competition among alert calls to account for the data at hand.

1 Informativity and scalar implicatures

In all of the main case studies, Schlenker et al. make crucial use of competition among calls on an informativity scale. This mechanism of competition takes its inspiration from the standard treatment of scalar implicatures in human language. Consider, for instance, a friend who utters:

(1) Sue or Felix will be there.

Typically, the hearer of (1) will infer that it is not the case that *both* Sue and Felix will be there. The reasoning might look something like:

***Corresponding author: Shane Steinert-Threlkeld,** Department of Philosophy, Stanford University, 450 Serra Mall, Bulding 90, Stanford, CA 94305, USA, E-mail: shanest@stanford.edu

(2) Enrichment of ‘or’ to its exclusive meaning:
  a. ‘or’ competes with ‘and’ in the sense that the following sentence was an *alternative* to (1):
     (i) Sue and Felix will be there.
  b. Sentence (2-a-i) is strictly more informative than (1).
  c. If the speaker were being cooperative and knew that (2-a-i) was true, she would have uttered it.
  d. Thus, assuming that she is cooperative, she must not know that (2-a-i) is true.
  e. Assuming that the speaker is in a position to know whether or not (2-a-i),[1] the hearer can infer that it is false.

The fact in (2-b) can be captured by saying that the alternatives lie on a scale of informativity:[2]

(3) ‘or’ < ‘and’

where the ‘<’ captures that the ‘and’ sentence is more informative in that it asymmetrically entails the ‘or’ sentence. In general, given an informativity scale, generalized Gricean reasoning can be used to generate the implicatures that all of the more informative alternatives on the scale are false (or at least not known to the speaker).

I will briefly pause over a *prima facie* worry about appealing to competition among calls in monkeys: the reasoning by which competition leads to enriched meaning – as exemplified in (2) – appears to rely on facts about the mental states of the speaker. That is to say, the calculation of the enriched meaning appears to require that the hearer be able to reason about other minds and, in particular, that of the speaker. It is unclear whether the various monkey species under discussion in Schlenker et al. (2016b) can carry out such reasoning: the literature is both inconclusive and has focused mostly on the great apes.[3] Even if true, however, this need not be alarming. Schlenker et al. are careful in claiming that the logic of competition among calls only depends on a strict informativity relationship and some mechanism for selecting the most informative call applicable in a given situation. This mechanism need not involve reasoning about other minds. The authors capture this minimal core with the following principle.

1 This is sometimes called the competence assumption, following van Rooij and Schulz (2004).
2 Such scales are often called Horn scales, following their introduction in Horn (1972).
3 See Premack and Woodruff (1978), Call and Tomasello (2008).

(4) Informativity Principle:
If a sentence S was uttered and if S' is (i) an alternative to S and (ii) strictly more informative than S, infer that S' is false.

Further reason not to be worried about the role of theory of mind comes from recognizing that Grice (1975) himself did not require that reasoning like (2) describe the actual process by which hearers determine what is implicated. Rather, such a piece of reasoning must be able to be provided 'from the outside' as it were. Grandy (1989) puts it thus: "In other words, it suffices if the hearer can intuitively grasp the implicature so long as we conversational theoreticians can supply the rigorous argument" (p. 519).[4] With this *prima facie* worry assuaged, I will proceed to discuss two of the case studies, each of which makes prominent use of (4), in detail. Following that, I will return to raise more fundamental worries about the use of competition among calls to enrich literal meaning.

2 Titi monkeys: Whither the sentence?

Titi monkeys exhibit a seemingly complex call system whereby a sequence of calls can encode information about both type and location of various predators. Cäsar et al. (2013) report the results of field experiments in which model predators were placed either in the canopy or on the ground and elicited call sequences were observed. Schlenker et al. summarize the data to be explained as:

(5) Main generalizations (model predator field experiments):[5]
a. Raptor, canopy: A^{++} (average length: 26.8)
b. Raptor, ground: $\underline{A^{++}B^{++}}$ (average number of As: 12.6)
c. Cat, canopy: AB^{++}
d. Cat, ground: B^{++}

Schlenker et al. compare a non-compositional treatment with a compositional treatment of this data. The syntax is given by:

4 See also Saul (2002) for arguments that Grice does not provide a theory of the psychological process of utterance interpretation.

5 Here, A^{+} is the Kleene plus, indicating a sequence of one or more As. A^{++} indicates two or more As.

(6) Sentence types: B^{++}, A^{++}, $A^{++}B^{++}$, AB^{++}

Throughout their paper, a *sentence* is any call sequence that is preceded and followed by a pause that is at least two standard deviations above the mean inter-call pause. The non-compositional semantics can now be specified.

(7) Non-compositional semantics:
a. $[[B^{++}]] = 1$ iff there is a noteworthy event
b. $[[A^{++}]] = 1$ iff there is a non-ground predator
c. $[[A^{++}B^{++}]] = 1$ iff there is a non-ground predator on the ground
d. $[[AB^{++}]] = 1$ iff there is a ground predator in non-ground position

(8) Alternatives:
S' is an alternative to S if it can be obtained by replacing any number of As with the same number of Bs or *vice versa*.

(9) Informativity scale (where $n \geq 2$ and $2 \leq k \leq n-2$):[6]

$$\begin{tikzcd}[column sep=small]
A^k B^{n-k} \arrow[d, no head] & & \\
A^n \arrow[dr, no head] & & AB^{n-1} \arrow[dl, no head] \\
& B^n &
\end{tikzcd}$$

This theory can account for the data in (5) when conjoined with the Informativity Principle (4). For example, if B^n is produced, (4) will generate the inference that A^n and AB^{n-1} are inapplicable.[7] Thus, B^n is enriched to mean there is a noteworthy event and there is no non-ground predator. A sequence with this meaning would only be appropriate in the cat-on-ground contexts.

The main objection to this analysis concerns the time course of the sequences. For instance, the first B in an $A^{++}B^{++}$ sequences appears, on average, after twelve A calls; this, again on average, will be about 16 to 17 seconds after the first A call. Thus, a monkey would have to wait that long to be able to tell whether it's hearing an A^{++} or $A^{++}B^{++}$ sentence. This seems implausible, especially since the two are scalar competitors according to (9). I will return to this point in discussing the next case study.

6 Note that I reformulate their scale in (73) with a *fixed* length of sequences, as context makes it clear is intended.
7 Since A^kB^{n-k} entails A^n, it will also be inapplicable.

To address this time course worry, Schlenker et al. introduce a compositional analysis in which sequences are evaluated relative to times and concatenation shifts the time parameter.

(10) Compositional semantics:
a. $[[B]]^t = 1$ iff there is a noteworthy event at t
b. $[[A]]^t = 1$ iff there is a serious non-ground alert at t
c. $[[wS]]^t = 1$ iff $[[w]]^t = 1$ and $[[S]]^{t+1} = 1$

On this analysis, individual calls – not entire sentences – compete with each other. The same reasoning as above explains why B^{++} is only used in cat-on-ground situations. Simple competition reasoning also explains the data in the two raptor cases, assuming that: (i) even a raptor on the ground is a non-ground threat (so that A is applicable) and (ii) after observing the model raptor being stationary for a long period of time, it is no longer considered a serious alert and so A ceases to be applicable. The former fact explains the canopy situations and the latter the ground situations.

The explanation of cat-in-the-canopy situations, however, gives reason for pause. Crucially, the relevant cats are *ambush* hunters. This means that, once detected, they no longer pose a serious threat. That AB^{++} is always used in such situations then gets explained as follows. The first A reflects that such a cat is a serious non-ground threat. After emitting this first A, however, “it can be assumed that the cat has been detected by conspecifics” (§6.4). Because a cat is an ambush hunter, this means that it no longer poses a serious threat and so A is no longer licensed.

While the general shape of this explanation seems reasonable, the data appears to be too strong to be accounted for by such means. It is rather striking that in cat-in-canopy situations, the Titis always emit a *single* A call before the sequence of Bs. Were the function of such a call to alert conspecifics, one would expect to see (i) more than a single A so as to really ensure that conspecifics are aware and (ii) some variation in the number of As preceding the Bs.

Assuming that an explanation of the same shape, but with lightly modified meanings, can be given, a question immediately arises: in what sense is this semantics really compositional? By analyzing the role of sentences in this theory, it appears that there is no robust sense in which the semantics is compositional. The lone compositional rule says to treat call concatenation as ‘conjunction’ while shifting a time parameter forward. The net effect of these two, however, is that two calls in a row are interpreted as essentially isolated calls according to their lexical meanings. Furthermore, the explanation given for each piece of data in (5) appeals to competition only amongst single calls and not amongst entire sentences. To my mind, then, it appears that the theory preferred by Schlenker et al. in fact needs merely lexical semantics plus informativity reasoning.

Further evidence against the reality of sentences in the case of Titi call sequences comes from closer inspection of the raw data, which has been helpfully made publicly available at http://dx.doi.org/10.5061/dryad.sd1sr. I went through the data and analyzed the call sequences for their sentential structure according to the criterion used in other case studies in the paper: a sentence is a sequence preceded and/or followed by a pause longer than the mean plus two standard deviations. Several observations can be made.

- The eagle-in-canopy and leopard-on-ground situations uniformly elicit *A*s and *B*s, respectively.[8] Thus, while the sentential structure of such sequences is relatively unimportant, the generalizations about these two situations (which most closely resemble naturalistic ones) are the most robust.
- In cat-in-canopy situations, there are exactly four genuine AB^{++} sequences. In *every* one of them, there is an exceptionally long pause (in three cases, more than four standard deviations above the mean) between the *A* and the first *B*. This strongly suggests that AB^{++} sequences do not form a coherent syntactic unit.

The fifth sequence in this situation also shows a striking pattern: it begins *AAABB*. Not only is this not an AB^{++} sequence, Cäsar et al. (2013) report (in their Table 1) that a *second* monkey made its first call after the third call. While the data does not allow us to figure out exactly which calls were made by which monkeys in this particular case, this suggests that one of the monkeys was using A^{++} in this context.

- The raptor-on-ground sequences also reveal interesting patterns, even among the five out of seven observations that Schlenker et al. claim to be $A^{++}B^{++}$ sequences. One monkey[9] begins *ABB*, with a significant pause. Then, what might appear to be a long $A^{++}B^{++}$ sequence follows. Nevertheless, there is a very substantial pause after the *A*s and before the *B*s. Under the existing criterion for sentencehood, the A^{++} and B^{++} would be separate sentences. One monkey[10] by this criterion does emit a genuine $A^{++}B^{++}$ sequence. The remaining three sequences[11] appear to be best analyzed as A^{++} with

8 There are a very few *O*s in one eagle-in-canopy case.
9 The single one emitting calls in Group D.
10 The single one emitting calls in Group A.
11 From Groups M (second responder), P, and R.

various Bs (sometimes even a single one) mixed in: there are no significant pauses following what would be an $A^{++}B^{++}$ sentence and As often follow the B sequences.

Let me clarify the dialectical force of this close inspection of the data. The point is not simply that the data is messy: even in natural language cases where we have strong evidence about the syntax, we will observe plenty of ungrammatical utterances. Rather, the point concerns the role that sentences and compositional interpretations thereof play in this theory. On the semantic side, we saw that each call in a sequence is treated as an independent utterance. On the pragmatic side, competition takes place among individual calls and not among sequences. On the syntactic side, the above analysis shows that the sentence types in (6) posited to handle the data in (5) do not deserve to be called sentences. In total, then, sentences play no role in the final theory of Titi alert calls.

3 Pyow-Hack sequences, revisited again

The second case study that I will discuss concerns the use of *pyow-hack* sequences by male Putty-nosed monkeys. The main idea here is that *pyow*s are used in leopard situations and *hack*s in eagle situations. Short sequences of *pyow*s followed by *hack*s reliably trigger group movement, with length of movement roughly correlated with length of the sequence.

(11) Main generalizations (mostly from field experiments):[12]
  a. *Pyow* discourses: $P^{+}\_\ldots\_P^{+}$ — leopard contexts
  b. *Hack* discourses: $H^{+}\_\ldots\_H^{+}$ — eagle contexts
  c. Transitional discourses: $H^{+}\_\ldots\_H^{+}\_P^{+}\_\cdots\_P^{+}$ — eagle contexts
  d. *Pyow-Hack* sequences: $P^{+}H^{+}$ — trigger group movement

Prior to the present paper,[13] the standard analysis[14] was that such sequences are idiomatic, having syntactic structure but not compositional interpretation.

Schlenker et al. first consider a non-compositional analysis.

12 A discourse is a sequence of sentences; ‘_’ indicates a long pause.
13 See also the companion paper Schlenker et al. (2016a) for more detail.
14 See Arnold and Zuberbühler (2012).

(12) Sentence types: P^+, H^+, P^+H^+

(13) Non-compositional semantics:
a. $[[P^n]] = 1$ iff there is an alert and the alarm level is $\geq n$
b. $[[H^n]] = 1$ iff there is a serious eagle-related alert and the alarm level is $\geq n$
c. $[[P^k H^{n\text{-}k}]] = 1$ iff the group is moving and the alarm level is $\geq n$

(14) Alternatives:
S' is an alternative to S if both are sentences and S' can be obtained from S by replacing some number of Ps with the same number of Hs or *vice versa*

(15) Informativity scale:

$$\begin{tikzcd}[column sep=small] H^n \arrow[dr, dash] & & P^k H^{n-k} \arrow[dl, dash] \\ & P^n & \end{tikzcd}$$

Here, the Informativity Principle (4) explains why P^+ only occurs in leopard contexts: H^+ is more informative and so must not be applicable, so the alert must not be eagle-related.

The compositional analysis differs in a few ways. First, they enrich the syntax:

(16) Sentence types: P^+, H^+, P^+H^+, H^+P^+

(17) Compositional semantics:
a. $[[P]]^a = 1$ iff there is an alert and the alarm level is $\geq a$
b. $[[H]]^a = 1$ iff there is a serious non-ground movement-related alert and the alarm level is $\geq a$
c. $[[wS]]^a = 1$ iff $[[w]]^a = 1$ and $[[S]]^{a+1} = 1$

Now, they also must introduce a second pragmatic principle which provides a ‘first round’ of enrichment before the Informativity Principle provides further enrichment.

(18) Urgency Principle
If S is triggered by a threat and contains calls that convey information about its nature or location, no call that conveys such information should be preceded by any call that does not.

(19) Revised Informativity Principle
As before, but takes into account meanings enriched by the Urgency Principle.

The argument that P^+H^+ can only be used in group movement situations runs as follows. In an eagle or leopard context, P^+H^+ would violate the Urgency Principle (18), since the Hs are preceded by calls that do not contain information about the threat. Thus, P^+H^+ gets enriched to exclude leopard and eagle contexts from its meaning. Secondly, this enriched P^+H^+ meaning is strictly more informative than (the unenriched) H^+ and H^+P^+ (which are both applicable in eagle and movement contexts), so the latter two never get used in group movement situations.

Schlenker et al. note that the compositional analysis can make different predictions than the non-compositional one since, in the former, P^+H^+ has a weaker meaning which requires world knowledge to generate the group-movement reading. While they also claim that it is "arguably more explanatory", the Urgency Principle is, to my mind, arguably just as stipulative as non-compositional semantic clauses.

Before further critical engagement, I want to make a friendly suggestion. For the compositional analysis, Schlenker et al. (2016b) posit H^+P^+ sentences as part of the syntax even though they appear never to be used.[15] They claim that they need to do so because "sentences of type H^+P^+ … serve as alternatives to P^+H^+ when the Urgency Principle is applied to the latter" (§4.3). This, however, appears to be an unnecessary motivation. By the definition of alternatives in (14), the sentence H^n will be an alternative to P^kH^{n-k} since it arises by replacing k Ps with k Hs. But P^kH^{n-k} will still be in violation of Urgency for the same reasons as given above and so will not be used in eagle contexts. A simpler theory that does not posit H^+P^+ in the syntax appears to be easily given.

Even with this friendly amendment, some issues remain with the analysis. In particular, there are worries about the time course of *pyow-hack* sequences, echoing some I expressed about the Titi case study. The main criticism of the non-compositional analysis, put eloquently in the companion paper Schlenker et al. (2016a), concerns just this: "the *time course* of *pyow-hack* sequences makes it surprising that these should have an idiomatic meaning: they are relatively long, with pauses between calls, which makes it a bit surprising that they should

15 If one modifies the operational definition of 'sentence', some can appear, though they are always significantly rarer than P^+H^+.

be given a lexical meaning as whole units" (p. 12, emphasis in original). It seems, however, that it should be equally surprising in light of the time course data that sentences should be the units of competition for both the Urgency and Informativity Principles. The monkeys would have to wait for many seconds before telling whether they are hearing a P^+ or P^+H^+ sequence, which would trigger different inferences by Urgency.

More generally, it is unclear that the data warrants putting P^+H^+ into competition with the other sentences at all. The distinguishing feature of these sentences is that they trigger group movement. That is to say, what needs to be accounted for is the *effect* these have on the monkeys. While this effect does indeed differ from those of pure P^+ and H^+ sequences,[16] the semantics offered mainly accounts for the *causes* of alarm calls, i. e. the situations in which they are uttered. Schlenker et al. use competition and a world knowledge principle to derive the group movement effect; but it's not clear that one wants to generate the inferences that the other calls are no longer applicable. For instance, P^+H^+ does appear in eagle contexts and so should not be enriched to exclude them.

The data suggests that the alert calls have interesting behavior both in what elicits them and in the effects they produce. Plausibly, they are what Millikan (1995) calls 'pushmi-pullyu' representations: devices that mix descriptive and directive contents in a way that is more primitive than pure descriptive or pure directive content. Pursuing a semantic theory for the alert calls which includes both of these dimensions may provide more satisfying explanations.

4 The data of pragmatics

In the last two sections, I have looked at two of the case studies in some detail in order to examine in what ways the formal semantic and pragmatic machinery is needed. In this last section, I want to raise a more general worry about the use of scalar implicatures to account for monkey alert call data. As mentioned earlier, the worry is not that such reasoning presupposes a theory of mind. Rather, I want to argue that the kind of data that we have in the monkey case studies is too impoverished to justify positing a gap between literal and enriched meanings.

16 In fact, there exists a large lag between the *pyow-hack* sequences and the group movement 'triggered', which leaves open the possibility that other factors trigger the movement.

Schlenker et al. note that informativity scales appear to flip in downward-entailing environments:

(20) I doubt that Sue or Felix will be there.
 '$\not\Rightarrow$' It is not the case that I doubt that Sue and Felix will be there.

The theory that 'or' is inclusive but enriched by competition to exclusive handles the difference between (1) and (20) well: in the latter case, the 'or' sentence is actually more informative and so no enrichment takes place. On the difference between these two sentences, Schlenker et al. write:

> ...we were led *on the basis of truth-conditional data* to posit a division of labor between the literal, or 'semantic' meaning of sentences, and further 'pragmatic' inferences obtained by taking into account *competition* among possible utterances. (§ 2.1.2, emphasis in original)

While there is a sense in which this is true, there are very real differences between the kind of data we have for human language and for monkeys. In our case, the difference between (1) and (20) depends on embedding a sentence in a complex linguistic environment with a known compositional semantics.[17] Moreover, we have linguistic mechanisms for teasing apart literal and enriched meaning. In particular, the latter can be both cancelled and reinforced:

(21) Sue or Felix – maybe even both – will be there.

(22) Sue or Felix – but not both – will be there.

If 'or' had an exclusive meaning, (21) should be a contradiction and (22) a waste of breath. That neither is the case provides some of the strongest evidence for positing a gap between literal and enriched meaning.

On the other hand, the sort of data we have for monkey alert calls is purely distributional. It takes the form: a call (sequence) is used in such-and-such situations.[18] In these simple call systems, there are no mechanisms for explicitly distinguishing the literal and enriched meanings. More generally, when given only correlations between calls and situations, the division between literal meaning and enrichment via competition is massively underdetermined.

Similar points apply to the search for compositionality in these call systems. In the human case, our arguments that a semantic theory must be compositional run as follows: certain properties of our semantic competence – in particular

17 At least, an environment with a known monotonicity profile.
18 And in some cases: a call produces such-and-such behaviors.

that we can learn a system which allows us to understand unboundedly many new sentences in a productive and systematic way[19] – can be had only by positing the grasp of a compositional semantic theory.

These points raise a more general methodological question. Schlenker et al. are careful to repeatedly point out that by using methods from formal semantics like compositional meaning assignments and competition to enrich meanings, "it does *not* follow that the calls under study share non-trivial properties, let alone an evolutionary history, with human language" (abstract). Nevertheless, the reasons that justified the introduction of these two tools in the formal semantics and pragmatics of natural language do depend on features of our linguistic systems that appear not to be shared by these monkey alert call systems. Given that this is so, I would like to invite the authors to expand on how semantic and pragmatic theory choice should take place.

References

Arnold, Kate & Klaus Zuberbühler. 2012. Call combinations in monkeys: Compositional or idiomatic expressions? *Brain and Language* 120(3). 303–309. doi: 10.1016/j.bandl.2011.10.001.

Call, Josep & Michael Tomasello. 2008. Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences* 12(5). 187–192. doi: 10.1016/j.tics.2008.02.010.

Cäsar, Cristiane, Klaus Zuberbühler, Robert J. Young & Richard W. Byrne. 2013. Titi monkey call sequences vary with predator location and type. *Biology Letters* 9(5). 20130535. doi: 10.1098/rsbl.2013.0535.

Davidson, Donald. 1965. Theories of Meaning and Learnable Languages. In Y. Bar-Hillel, (ed.), *Proceedings of the International Congress for Logic, Methodology, and Philosophy of Science*, 3–17. Amsterdam: North-Holland Publishing Company.

Fodor, Jerry. 1987. *Psychosemantics*. Cambridge, MA: The MIT Press.

Grandy, Richard E. 1989. On Grice on Language. *The Journal of Philosophy* 86(10). 514–525.

Grice, Herbert Paul. 1975. Logic and Conversation. In P. Cole and J. Morgan (eds.), *Syntax and Semantics*, 41–58. New York: Academic Press.

Horn, Laurence Robert. 1972. *On the Semantic Properties of Logical Operators in English*. PhD thesis, Los Angeles: University of California.

Millikan, Ruth Garrett. 1995. Pushmi-Pullyu Representations. *Philosophical Perspectives* 9. 185–200.

Premack, David & Guy Woodruff. 1978. Does the chimpanzee have a theory of mind? *The Behavioral and Brain Sciences* 4. 515–526. doi: 10.1017/S0140525X00076512.

Saul, Jennifer M. 2002. What is Said and Psychological Reality; Grice's Project and Relevance Theorists' Criticisms. *Linguistics and Philosophy* 25. 347–372.

19 See Davidson (1965) for the original learnability and Fodor (1987), among others, for systematicity and productivity.

Schlenker, Philippe, Emmanuel Chemla, Kate Arnold, & Klaus Zuberbühler. 2016a. Pyow-Hack Revisited: Two Analyses of Putty-nosed Monkey Alarm Calls. *Lingua* 171. 1–23. doi: 10.1016/j.lingua.2015.10.002.

Schlenker, Philippe, Emmanuel Chemla, Anne M. Schel, James Fuller, Jean-Pierre Gautier, Jeremy Kuhn, Dunja Veselinović, Kate Arnold, Cristiane Cäsar, Sumir Keenan, Alban Lemasson, Karim Ouattara, Robin Ryder, & Klaus Zuberbühler. 2016b. Formal Monkey Linguistics. *Theoretical Linguistics*.

van Rooij, Robert & Katrin Schulz. 2004. Exhaustive Interpretation of Complex Sentences. *Journal of Logic, Language and Information* 13(4). 491–519. doi: 10.1007/s10849-004-2118-6.