Review of *Embedded Scalars, Preferred Readings & Intonation: An Experimental Revisit*

For *Journal of Semantics*

November 16, 2014

1   General assessment

This is a well-written paper exploring the role of intonation on the availability of embedded implicatures and, more generally, on preferences between different readings of complex sentences containing scalar items. The paper is clear and thoughtful, it is well time to explore the processing of aurally presented sentences containing embedded scalar items, and the data are important and are discussed with care and expertise. The incremental verification task is also quite interesting, and might lead to new ways of exploring issues in semantics/pragmatics. I believe the paper should be published.

Nevertheless, I believe the paper needs major revisions before it can be published. My main concern, which is what I'll largely focus on in this review, is with the assumed theoretical background and therefore also with the theoretical consequences the authors draw from their data. In particular, the main claim – that the data are problematic for both 'traditionalism' (an inapproriate name, I think – more on this in (3) in section 3 below) and 'grammaticalism' is exaggerated, at best. So far as I can see, most of the data are easily accommodated in both theories without much ado. Of course there remain many open questions, and these should be clarified, but the suggestion that the data are as problematic for the two 'isms' as they are made out to be is objectionable.

I will try to substantiate this claim in section 2 below, but let me say briefly that I think the source of this problem is that the authors assume that certain ways

of resolving choice points within both traditionalism and grammaticalism are representative of those frameworks, and this strikes me as unreasonable when there are live debates about how these choice points should be resolved.

2 Theoretical background and consequences

In this section I will say a little bit about what I see as problematic with the authors' characterization of both grammaticalism and traditionalism, and I will try to speak to how this might affect the main claims of the paper.

2.1 Grammaticalism

There is no justification for the authors' decision to use the Strongest Meaning Hypothesis (SMH) as representative of grammaticalism. The authors say that it is 'the most prominent' of selection principles, but this is not a good justification. First I don't think this is true, at least not among proponents of the grammatical approach. There are live debates about whether an EXH-specific parsing principle is even needed and, if it is, there are differing opinions as to what that might be. Second, even if it were true that the SMH is somehow most prominent, how would that be relevant to whether it has more of a claim to truth compared to competing ideas? The authors are right to point out that their data are *prima facie* problematic for the SMH, but this is not particularly damaging for grammaticalism itself, is it?

The force of this objection gets its bite when we see what happens if we consider what is predicted under a different approach, say Magri (2009, 2011). Unless I'm mistaken, the approach actually makes quite a lot of sense of the authors' data once we assume with Noveck and Posada (2003) (among others) that participants can often be classified into (i) those participants that generally select the literal meaning, and (ii) those participants that generally strengthen. How would these interpretation strategies be implemented in Magri's proposal? Under his approach – which is motivated by data, not *a priori* parsing principles – a complex sentence will have an EXH at each scope site. The literal meaning follows by treating alternatives as irrelevant, and hence by pruning them so that EXH is vacuous. So one strategy, call it the pruning strategy, is to prune alternatives from the domain of EXH. A second strategy, the strengthening strategy, is one in which you don't prune alternatives at all.

Now under this assumption there is no parse with EXH only at the root, nor is there a parse with EXH only embedded. A sentence like *All of these letters are*

connected to some of their triangles is always parsed as: $EXH(C_1)$(*All of these letters* $x, EXH(C_0)$(*x are connected to some of their triangles*)). For subjects who follow the pruning strategy, C_0 and C_1 are the empty set, so both occurrences of EXH are vacuous. This gives you the literal meaning. For subjects who follow the strengthening strategy, C_0 = {x is connected to all of its triangles}, and $C_1 = \{$All of these letters $x, EXH(C_0)$(x is connected to all of its triangles)$\}$. I might be mistaken, but this parse is equivalent to what the authors describe as the 'local' reading: the local EXH gives an embedded 'some but not all' meaning, and the global EXH is vacuous (it adds nothing to say that not all letters are connected to all of their triangles, when you know that all letters are connected to only some of their triangles). In the non-monotonic case of *exactly-one*, the meaning of the Magri-parse for subjects who follow the strengthening strategy is not equivalent to what the authors describe as the 'local' reading, but in the authors' experiment it turns out to behave exactly like that reading, coming out 'true' at Step 5. (I think the predicted reading is that exactly one letter is connected to some but not all of its triangles and either none or more than one are connected to all of their triangles.)

Thus, the only two form-meaning pairs predicted under Magri (2009, 2011), once a single auxiliary assumption is made about interpretation strategies, gives rise to precisely the two readings found in the authors' data. What is needed, of course, is a statement about why the pruning strategy should be preferred to the strengthening strategy, and there are obvious suggestions worth pursuing, some of which the authors discuss (e.g., literal meanings require less computation, which might be preferred in a demanding task such as the one discussed in the authors' paper). There is also the question of the differential availability of the local reading between *every* and *exactly-one*; this just means there's work to do, but I don't see how this is especially problematic for the grammatical theory. If anything, the data seem to support the approach, given the existence of a ready explanation that combines independently motivated assumptions about grammar (Magri, 2011) and independently motivated assumptions about interpretation strategies (Noveck and Posada, 2003).

I may of course have miscalculated, and the authors are encouraged to double-check everything I've said above. The main point nevertheless stands: after (rightly) rejecting the SMH selection method, the authors are too quick to dismiss grammaticalism itself; I'd encourage them to clarify the nature of the challenge that remains for grammaticalism if it is to account for their data.

2.2 Traditionalism

As with grammaticalism, I feel that the authors are too quick to select from a set of certain choice points made available by traditionalism and take these choices as representative. For example, the authors estabish a 'weak traditionalism' and a 'strong traditionalism' based on whether, by default, crucial contextual features (e.g., speaker-opinionatedness, etc.) are either assumed to hold or not hold. But are non-default strategies not conceivable? For example, suppose that the contextual features under consideration are essential to implicature computation. For purposes of discussion, let us follow Gamut (1991) and take this to mean that you must assume that the stronger alternative is relevant, and that the speaker is opinionated about it, in order to squeeze a scalar implicature (secondary implicature, let's say; Sauerland, 2004) out of the utterance. Need there be a 'default' strategy about this? It might be that unless these features are clearly specified in the context, the listener will need to make a guess about whether the speaker is opinionated, say, and different people might make different guesses based on all sorts of factors. This does not mean that traditionalism makes no predictions, but rather that these predictions will be parameterized by the kinds of answers to these questions different subjects might give. Indeed, one could manipulate these features experimentally and see whether implicature rates change (e.g., Goodman and Stuhlmuller, 2013).

More problematic, however, is the authors' suggestion that a challenge for traditionalism is to explain their finding that accent does not increase the availability of local readings. But isn't an explanation at hand? Surely prosodic prominence itself is not sufficient to trigger the kinds of enrichments the authors have in mind. The typical example used to motivate the connection between accent and local enrichments is rarely an isolated sentence, but rather is a sentence with an explicitly mentioned contrasting alternative. E.g.,'The chili isn't spicy; it's SPICY.' 'The boy who saw SOME of the pictures is happier than the boy who saw ALL of them.' And so on. This crucial feature – the explicit mention of a contrasting alternative – was entirely absent in the authors' materials; the mere presence of accent in absence of salient alternatives might not be enough to trigger the kind of local enrichment envisioned. If it is right that the conditions for triggering local enrichment were not met, the authors' data might not be as problematic as they appear to be. (We might imagine that without a salient alternative focus projection could project focus from embedded 'some' all the way to the root; but with a salient alternative, focus would likely stay narrow.)

The only potentially problematic data point that I see for traditionalism is the

availability of local readings without any accent at all, especially with *exactly-one*, which cannot mimic local strengthening by playing with global alternatives. (In connection with this, I don't see how 'unrestricted traditionalism' (p.8) is a kind of traditionalism. The alternative in (8) does not even entail AS. Why, under a traditional Gricean view, should it even be considered as an alternative?)

3   Minor comments

(1)   p.9: 'The grammatical approach, as desribed so far, is not yet a fully artic-
       ulated theory.'

Theory of what? Theory of grammar, or theory of human behaviour? Would anyone say that a syntactic theory that posited PP-attachment ambiguities was not a 'fully articulated theory'? Of course, these ambiguities need to be resolved somehow, and auxiliary assumptions about processing, memory, planning, etc. are invoked to say how. But this is commonplace, no? I somehow fail to see what the substance of this claim is. (The discussion in Chemla and Singh (2014a,b) might be relevant.)

(2)   A potential difference between EC/LC and literal/strengthened ambigui-
       ties

The authors found with the EC/LC ambiguities that there was a slight preference for delaying the response. This seems to be a rational strategy for a participant to follow – when there is an ambiguity it might make sense to wait for some time for relevant information to come in before deciding how to disambiguate, especially if there is little cost to doing so (here resolving the ambiguity merely requires an answer to the question: where should the PP be attached?). With scalar sentences, on the other hand, we might naturally expect a preference for literal readings when the task imposes high demands on memory, because the literal reading requires strictly less computation: no need to generate alternatives, no need to exhaustify, etc. A lazy participant might wish to avoid all that computation. I wonder if this consideration has any merit, and whether it bears on the interpretation of the results.

(3)   Traditionalism

On p.6, the authors write that they use the word 'traditionalism' because it takes a 'conservative stance' with respect to Grice's work. This totally ignores arguments

by Fox (2007, 2014) that the most conservative understanding of the Maxim of Quantity only yields ignorance inferences, not SIs. It is only with a radical stipulation, grammatically stipulated formal alternatives, that a quantity maxim can be used to derive SIs. Furthermore, as Larry Horn has pointed out in many papers, the debate over the source of the apparent ambiguity is very old, much older than the modern era. In light of these considerations, calling the neo-Gricean approach 'traditional' does not entirely do justice to tradition.

Why not use labels that more accurately characterize the debate? Isn't the issue one about the domain-specificity of the strengthening mechanism?

4   Typos/minor errors

- In the abstract, the authors begin by saying that 'some' receives a 'pragmatic enrichment'; calling it a 'pragmatic enrichment' begs the main question of the paper.

- p.1, last line onto p.2: 'the under universal' → 'under the universal'

- p.3, second paragraph: 'this paper presents results form' → from

- p.3 same paragraph: 'to cater for'...? (I'm not sure this use of 'cater' is felicitous.)

- p.10: Calling something a 'pragmatic enrichment' is not the same as giving an account of such an enrichment; in these cases, how does the enrichment come about, and what about the mechanism makes it 'pragmatic?'

- I'm wondering if the authors might say a little bit more about how the incremental verification task suppresses typicality reasoning. Are the authors assuming that participants will not categorize certain configurations along the way as (a)typical? (Of course, I agree with the authors that 'typicality' is not an explanation but something in need of explanation; nevertheless, perhaps some comment might help further motivate the use of incremental verification.)

- 5.2.3, p.28: 'none of which' → none of whom

- p.29, middle of page: 'as if the upward-monotone quantifier exactly one' → non-monotonic quantifier

6

- Not a particularly relevant question, but just wondering: did the authors by any chance collect RTs at critical regions? (might be informative)

- I fail to see the prior plausibility of the disambiguation criterion the authors state at the bottom of p.39.

- p.39, last paragraph: The QUD criterion in Fox (2007) does make clear predictions. The problem is not with the predictions, but with the unfortunate fact that QUDs are hard to pin down. If we could find a way to overcome this obstacle, the predictions of the theory are clear (e.g., one might try fixing the QUD and see what happens; without an explicitly provided QUD, listeners have to guess one, cf. also a similar point raised in connection with traditionalism in section 2.2).

References

Emmanuel Chemla and Raj Singh. Remarks on the experimental turn in the study of scalar implicature, Part I. *Language and Linguistics Compass*, 8(9):373–386, 2014a.

Emmanuel Chemla and Raj Singh. Remarks on the experimental turn in the study of scalar implicature, Part II. *Language and Linguistics Compass*, 8(9):387–399, 2014b.

Danny Fox. Free choice disjunction and the theory of scalar implicature. In Uli Sauerland and Penka Stateva, editors, *Presufofopposition and Implicature in Compositional Semantics*, pages 71–120. Palgrave Macmillan, New York, NY, 2007.

Danny Fox. Cancelling the Maxim of Quantity: another challenge for a Gricean theory of Scalar Implicatures, 2014.

L. T. F. Gamut. *Logic, Language, and Meaning*, volume 1. University of Chicago Press, Chicago, IL, 1991.

Noah D. Goodman and Andreas Stuhlmuller. Knowledge and implicature: Modeling language understanding as social cognition. *Topics in Cognitive Science*, 5:173–184, 2013.

Giorgio Magri. A theory of individual level predicates based on blind mandatory scalar implicatures. *Natural Language Semantics*, 17:245–297, 2009.

Giorgio Magri. Another argument for embedded scalar implicatures based on oddness in downward entailing environments. *Semantics and Pragmatics*, 4(6): 1–51, 2011.

Ira Noveck and Andres Posada. Characterizing the time course of an implicature: An evoked potentials study. *Brain and Language*, 85(2):203–210, 2003.

Uli Sauerland. Scalar implicatures in complex sentences. *Linguistics and Philosophy*, 27:367–391, 2004.