

**From:** Open Mind em@editorialmanager.com  
**Subject:** Your Submission  
**Date:** April 4, 2016 at 5:55 AM  
**To:** Gregory Scontras scontras@stanford.edu

OM

Ref.: Ms. No. OPMI-D-16-00002  
Subjectivity predicts adjective ordering preferences  
Open Mind: Discoveries in Cognitive Science

Dear Gregory Scontras,

Your manuscript has now been evaluated and the Reviewers' comments are appended to this email. As you can see, the reviewers felt your work addresses an interesting and important question, and is potentially appropriate for this journal. They were positive about the ideas in your work and the design of your studies, but also had some concerns that should be addressed if the work is to be reconsidered for publication in this journal.

In particular, I see the following as the most important points raised by the reviewers (in most cases, mentioned by more than one reviewer):

1) The discussion/framing of other approaches needs some attention: R1 mentions some other hypotheses that deserve discussion, and R2 takes issue with the framing of previous approaches as syntactic vs semantic. This latter point is related to another issue mentioned by R2 and R3: the distinction between online decisions about adjective ordering versus conventionalized ones. Clarifying how your proposal relates to these questions would, I think, help address these concerns.

2) All three reviewers express some dissatisfaction about the lack of direct comparison between your hypothesis and others. As they point out, the fact that your data are well explained by subjectivity tells us little about whether this hypothesis provides a better explanation than any other. Some of the other hypotheses may seem difficult to operationalize, but on the face of it, so is subjectivity. A quantitative comparison to other approaches would considerably strengthen the paper.

3) R1 and R2 both express confusion/skepticism about the explanation for \*why\* subjectivity should determine ordering. I suspect that some readers will be skeptical regardless of the explanation, but it would be good to clarify if possible - or alternatively to admit that the hypothesis is based purely on descriptive adequacy and indeed it is not clear why it holds. (If so, this might suggest that subjectivity is correlated with some other, perhaps as yet undetermined, factor that could provide a more convincing mechanism for ordering.)

The reviewers also mentioned a number of other points which you should take into account if you choose to resubmit. I'm aware that the above changes will likely require adding more material to the paper; you may want to consider whether some of the details (e.g. of model comparisons) could be placed in Supplementary Materials in order to maintain the 4000 word limit.

If you decide to revise the work, please submit a detailed cover letter that addresses each of the points of concern raised by the Reviewers. Please note that there is no guarantee that a revised submission will be accepted, since revision can sometimes uncover new issues that were not clear before.

Your revision is due by 07/03/2016.

To submit a revision, go to <http://opmi.edmgr.com/> and log in as an Author. You will see a menu item call Submission Needing Revision. You will find your submission record there.

Yours sincerely

Sharon Goldwater  
Associate Editor  
Open Mind: Discoveries in Cognitive Science

Reviewers' comments:

Reviewer #1: Review of „Subjectivity predicts adjective ordering preferences"

#### 1. Overview

This short paper explores a cognitive hypothesis to a curious cross-linguistic regularity: The tendency of certain ordering preferences for non-conjoined adjectives within a noun phrase. The hypothesis that the authors explore is that more subjective adjectives occur closer to the noun, while less subjective adjectives occur further away. Using naturalness ratings, corpus data, and a faultless disagreement study, the authors found that subjectivity measures account for about two-thirds to about four-fifths of the variance in adjective orders.

The paper is extremely well written, presents two very elegant sets of experiments, and fits the readership of Open Mind well. However, I think that it needs to be expanded both in breadth and depth in order to make a substantial contribution to the discussion around adjective order.

#### 2. Major points

##### a. Situating the hypothesis within other current approaches:

I really appreciate the historical framing of the literature, but I do miss some very influential semantic or 'psychological' contemporary hypotheses: The hypothesis that not subjectivity, but subsectivity, determines adjective order (Trueswell 2009); the apparentness hypothesis (Sproat & Shih, 1991); even Laenzlinger (2005). Ignoring these literatures make it seem like the authors are uncomfortable with the idea that others had the same intuition they had, which they shouldn't be. It does, however, relativize the present findings: Subiectivity is one

others had the same intuition they had, which they shouldn't do. It does, however, relativize the present findings. Subjectivity is one explanation for a whole range of factors that determine adjective order; and as it stands, I am not convinced that subjectivity explains the data any better than the subjective/intersective distinction, or apparentness. All these concepts overlap significantly, and if I am to buy the explanation of subjectivity, I want at the very least some additional data that allow a comparison.

b. Applicability with a wider range

I am a little bit confused about Experiment 2. If I understand correctly (and I might not!), this study is the same as Experiment 1, but with a wider range of adjectives. Also, the model explains the data much, much less well (88% in Experiment 1 vs. 61% in Experiment 2). Thus, it seems like subjectivity is no better than most models that exist; and certainly worse than Wulff (2003). In addition, the hypothesis that complex concept formation is involved to a significant degree cannot be ruled out, because the set of AAN-combinations in Experiment 1 is limited, and this hypothesis was not tested in Experiment 2.

e. The proposed psychological mechanism for adjective order seems counterintuitive. If it was indeed the case that more subjective adjectives are ruling out miscommunications, then they should come first in post nominal positions. This is indeed the case for Romance languages, for example. However, then one would expect the same linear order in prenominal positions.

3. Minor points

- p.2: The cartographic approach does not seek adjective ordering preferences in their structure (that sounds as if there is something about their morphology that determines their order); it attempts to build a structural model.

- p.2: The idea that miscommunication increases with vague adjectives is a very relevant hypothesis to the author's model, and would be one (interesting) way of expanding the current paper.

- Table 1: Why are there different numbers of adjectives in each category? Isn't there a big danger that the effects are driven by individual items?

- p.3: It would be an even stronger argument for the authors if there were ordering preferences within the same semantic class, based on subjectivity.

p.4: (naturalness) —> (preference), or vice versa everywhere else

Reviewer #2: This paper investigates how the notion of "subjectivity" can explain adjective ordering preferences of English speakers. The question of adjective ordering has been of interest to typologists, syntacticians, and semanticists for a long time, and it's ultimate explanation has interesting implications for the cognitive science of language. I like this study, and I think it represents a contribution to this well-studied topic. Nevertheless I think there are a couple of things the authors could do to strengthen the paper. I outline these below.

(1) The authors suggest a number of alternative accounts of adjective ordering, for example a set of noun classes (as identified by Dixon or used by Cinque), or inherent-ness. They seem to suggest that subjectivity does a better job of operationalizing the relevant semantic notion, and indeed it does explain an impressive amount of the variance in the data. However I would find the paper \*much\* more convincing if the authors actually tested this by doing some model comparison.

(2) I'd like to see a bit more nuance in the presentation of the general question, in particular how previous work has explained adjective ordering based on the interaction of semantics and syntax. The authors set things up as a debate between whether syntax or semantics explains adjective ordering. But actually I think that's a conflation of two distinct (though related) questions: What is the ultimate explanation for why certain (types of) adjectives occur closer to nouns than others (across languages)? And what is the formalization/representation/conventionalization of these ordering preferences in the syntax of a language? Syntacticians like Cinque are largely interested in the latter. But that doesn't mean that they think the universal structure of the DP is unrelated to semantic features of adjectives.

I think this comes up again in the discussion/conclusion, since what the authors end up arguing for is that a particular semantic notion, subjectivity, partially determines order. But of course that leaves open the further question of how this interacts with the syntactic grammar of a language (English or otherwise). Maybe they think that the grammar encodes subjectivity along with frequency, length, etc., together these probabilistically determine order, and that's all there is to it. But it also could be that subjectivity is the ultimate source of a more deterministic set of word order rules, and the latter are encoded in the grammar. The latter seems more in line with how they introduce the issue -- there are these very systematic preferences that people show within and across languages. If subjectivity is really doing so much of the work, then it seems to me that ordering might actually be predicted to be quite flexible. In an individual instance, subjectivity will be determined largely by context, and thus you'd imagine that certain adjectives could vary quite dramatically in their relative ordering. Is that the case? If not, then something still needs to be said about how subjectivity and conventionalized syntax interact.

(3) I'm afraid I find the explanation offered about why decreased subjectivity means closer to the noun very unconvincing. (I'm partial to the "inherentness" explanation myself.) Maybe the authors can elaborate, otherwise, I'm not sure what it adds.

Smaller comments:

pg. 3 "...cartographic approach", one could say... I find this an odd comment. On the one hand, Dixon was less interested in claims about syntactic structure than in understanding the semantics types of adjectives and how they are distributed across languages. On the other hand the syntactic theory that Cinque subscribes to is literally called Cartography...

pg. 6 Any comments about why noun-specific effects are not found? Along the lines of my long-winded comment (2) above, it seems like the subjectivity account might easily predict that particular adjectives might be more subjective when applied to certain nouns than others. But maybe your data just aren't able to actually show that, and with more data per noun (with different adjectives) you'd see it? Are there noun-specific differences in subjectivity ratings? Do you think they are simply ignoring the nouns in the ordering judgment task?

pg. 11 Can you say something more about the analyses that use adjective class configurations? Why are the subjectivity ratings sometimes better at accounting for these?

One last general thing: Can you show that adjectives that are similar in subjectivity have more flexibility in relative order when they occur together? Would be nice to have some examples of this.

Reviewer #3: This paper presents a series of experiments arguing that the way in which subjects prefer to order adjectives ("big blue box" vs. "blue big box") is determined by the subjectivity of the adjective property. Subjects prefer to put more subjectively-determined adjectives (e.g. "big") further from the noun as compared to more objective adjectives ("blue"). This paper is very carefully written and constructed, with detailed analysis and friendly visualizations of the data. The topic is interesting in that there has been a long history of attempts to explain these ordering constraints with different kinds of linguistic and psycholinguistic theories. The results seem very clear that subjectivity influences ordering.

My primary worry is that I don't know how other theories do on this kind of data, so I can't tell how good the author's theory is. In the introduction, the authors outline a number of other features that prior authors thought determined adjective ordering including "inferentness", conceptual combination, specificity, etc. Does subjectivity matter over and above those other factors? Does it explain more variance than they do individually? Or overall? In this kind of domain where there's lots of ideas, it's hard to argue for one idea by showing that it works well; better is to show that it outperforms other prior theories. The authors state that there are difficulties in formalizing these ideas, but I don't see why subjectivity is any different. They could pick a reasonable formalization of any of these other factors (ratings or corpus measures) and look to see how well it predicts the ordering preferences. Doing this is particularly important since I'd think that subjectivity will tend to be correlated with some of these other predictors.

My second (perhaps ignorant) question is why we should think that anything abstract determines ordering preferences. If I just look at the frequency of usage---average distance from an adjective to its noun in a corpus---do I find that predicts ordering preference on new constructions? That could suggest that we just store some ordinal/weighting information about adjective position attached to each adjective. Of course in that analysis the causality could go either way... But in general I'd like to know what kind of evidence rules out those simple frequency/usage-based theories? Here's another version of the question: what if you gave me a new word and varied the subjectivity of its meaning. A fruit is "paxy" if you tend to like its color (subjective) vs. if it has more than 10 seeds (objective). Would you see ordering preferences for "paxy" depending on which meaning you gave people? If people really in their heads have a subjectivity->order rule, it should apply on novel words like that, no? In general, it would be especially nice in this paper to make and test some novel predictions beyond just ratings for ordering preferences and adjective measures.

Additional comments:

In section 3.3 (and perhaps elsewhere), the  $R^2$  values are a little hard to interpret because they include variance from the noise in measuring the subjectivity preferences as well as the naturalness ratings. Because of this measurement error, you could never expect  $R^2=1$ , even for a perfect theory. A useful thing to do might be to do an analysis to see how much of the potentially explainable variance (e.g. subtracting off the measurement error) you actually do explain. Such an analysis would take into account the reliability of the x- and y- values in the correlation and adjust the correlation. Check out Spearman's Prophecy formula.

- I would like to know a little more what these findings tell us in a broader sense beyond adjective ordering. Do the authors expect general subjectivity effects in language? If so, why? I know they avoided saying too much about this due in order to avoid speculating, but I think talking about the bigger picture would help to situate this work for the broader cogsci community.

- How many subjects ran in multiple experiments?

- When I think about the example of "big blue box" vs. "blue big box", these seem to elicit very different contexts in my imagination. In the first, there are many blue boxes, one of which is big; in the latter there are many big boxes, one of which is blue. So it feels like there is some difference in focus/pragmatics/context that is relevant to these effects. With respect to context, wouldn't you expect that different situations (like I just described) would bias people for one preference or the other? Maybe then there is not a single simple factors that matters?

