The paper was boring and long: A reply to Michel and Löhr's theory of predicate ordering in copredication

Elliot Murphy[1,2]

1.   Vivian L. Smith Department of Neurosurgery, McGovern Medical School, University of Texas Health Science Center at Houston, Houston, TX, 77030, USA
2.   Texas Institute for Restorative Neurotechnologies, University of Texas Health Science Center at Houston, Houston, TX, 77030, USA
Correspondence: elliot.murphy@uth.tmc.edu
Article Type: Discussion

Abstract: In a recent publication, Michel and Löhr (2024) propose a predictive processing (PP) theory of copredication. They do so in an attempt to account for previously documented predicate ordering effects, whereby copredication acceptability is significantly impacted by the order of predicates. I review how their PP model is unfit for purpose, does not account for the empirically documented effects, does not provide novel experimental directions or conceptual innovations, and ultimately falls back on assumptions already presented in the copredication literature. Hence, the PP model fails empirically and theoretically.
Keywords: copredication; polysemy; coherence relations; predicate order

Natural language provides us with the ability to pick out multiple meanings and attribute them to a single nominal, even if these distinct meanings invoke divergent and incompatible categories. For example, (1a) uses two adjectives within the same category (roughly, INFORMATION), while (1b), exhibiting 'copredication', picks out both INFORMATION ('boring', pertaining to the content) and EVENT/PROCESS ('long', pertaining to some aspect of reading the paper) senses (see Michel and Löhr 2024, example (11)).

   (1) a. The paper was boring and dull.
       b. The paper was boring and long.

Copredication remains highly relevant to major debates in linguistics and philosophy, since it cuts to the core of how the lexicon is formatted, and how distinct lexico-semantic representations relate to one another.

Michel and Löhr (2024) (henceforth, M&L) propose a model of copredication acceptability dynamics which is driven by context-sensitive expectations and predications of upcoming linguistic material. I will briefly outline in this article some major obstacles for M&L, in addition to clear cases of misrepresentation of my published work (Murphy 2021a, 2021b).

1. Much Ado About Nothing: The absence of frequency effects in copredication

Perhaps the most fundamental obstacle for M&L is the extensive evidence in Murphy (2021a) that the frequency of the polysemous sense in copredication has no statistically significant relation with sense order acceptability; i.e., there are no Dominant-Subordinate or Subordinate-Dominant predicate order preferences. Frequency has no statistically significant relationship with predicate order acceptability. This was already pointed out in Murphy (2023a), a previous reply to M&L's research, but this has not been accommodated into the present paper. While of course not equivalent, mechanisms of predictability plainly relate in various ways to lexical statistics pertaining to frequency, including within polysemous senses of single nominals.

Going beyond this, in Murphy (2021c) it was shown that what I called the "persistence conditions" for copredication (the conditions under which certain complex polysemous nominals can sustain coherent sense relations in a brief narrative which "takes away" certain senses until the object referred to is reduced in sense number) are also correlated significantly with semantic complexity, *not* sense frequency. This provides further evidence for the centrality not of prediction, but of semantics (see Murphy 2023b, 2024a).

M&L describe the model I put forth in Murphy (2021a), in which Simple-Complex orderings of copredication sentences are typically preferred over sentences in which the more semantically complex sense is placed first, and with the following hierarchy obtaining: PHYSICAL < INFORMATION < EVENT < INSTITUTION. This is termed the *Incremental Semantic Complexity* (ISC) model. M&L write: "One problem with reducing order effects

to a concreteness effect is that the assumption that abstract concepts are generally more difficult to process is far from obvious". Yet, my model does not reduce ordering effects only to a "concreteness effect"; in fact, I do not entertain a concreteness effect at all, but rather a semantic complexity effect, which of course often correlates with concreteness, but not necessarily (for examples, empirically tested, see Murphy 2021a). In addition, it ultimately does matter too much whether concrete concepts are more or less difficult to process than abstract concepts, since in Murphy (2021a) a *priming mechanism* is assumed for the implementation of Simple-Complex orderings and ISC, whereby the reason why ordering effects obtain is not strictly because of the simple sense being simpler, and the complex sense being more complex, but rather because simpler senses prime other senses more strongly (i.e., they pre-activate other senses to a greater extent). Certainly, priming is only one candidate mechanism to implement ISC, and other possible alternatives exist, yet to be empirically tested.

However, this is only the first part of the model in Murphy (2021a). The second part concerns coherence relations. In my model, these pertain to the establishment of a *causal* or *temporal connection* between predicates, or some overlap in conceptual features denoted by the predicates (*extensional overlap*). M&L write, incorrectly, that *extensional overlap* is not relevant to the theme of sense order effects in copredication: "While extensional overlap indeed influences acceptability judgements, it does not bear directly on explaining order effects". In Murphy (2021a), one of the first examples given to show that *extensional overlap* does in fact impact sense order preferences is the following:

(2) a. The city has 500,000 inhabitants and outlawed smoking in bars last year.

b. #The city outlawed smoking in bars last year and has 500,000 inhabitants.

c. The city outlawed smoking in bars last year and has 500,000 smokers.

The sense order becomes less salient for acceptability dynamics if the coherence relations are strong enough. Hence, M&L mischaracterize the explanatory scope of coherence relations and their interaction with sense ordering.

Oddly, while they briefly acknowledge that Murphy (2021a) entertains coherence relations as a factor in the ISC, M&L simultaneously insist that Murphy's (2021a) model needs to be replaced by one that considers the importance of "supportive sentence contexts". But this is just another way of formalizing coherence relations, which crucially

3

provide such a supportive context for copredication. If there are other "supportive contexts" that can be shown to impact copredication ordering effects, then this does not undermine my ISC model, nor does it serve as counterevidence, as M&L seem to think. Instead, it simply shows that ISC is implemented alongside various other pragmatic or syntactic processes of interest. ISC is only undermined if an alternative mechanism explains all of the cases covered by ISC more parsimoniously or with greater empirical scope. M&L do not achieve this.

Turning to additional evidence, M&L claim that "[s]ome scholars suggest that children can grasp more general features before specific instances are presented", and thus this acts as evidence, supposedly, against the claim that more abstract concepts are more effortful to process (in adulthood). The logical leap from these studies to the conclusion that copredication ordering effects cannot therefore be grounded in a Simple-Complex bias is thoroughly unmotivated. In addition, consulting the child language processing experiment reported in Murphy (2017), and comparing this with Murphy (2021a), there seems to be no clear relation between the age at which children are first able to articulate a given polysemous sense (likely driven by how frequent that sense is) and the clear Simple-Complex acceptability dynamics documented. There could indeed be such an Early-Late sense ordering preference in terms of age of sense acquisition for polysemous words involved in copredication, but a more systematic experimental investigation is needed to test across different nominal types (e.g., *lunch*-type, *book*-type, *school*-type). Once again, M&L do not attempt to conduct such a survey, and keep to speculation.

Moving into even more speculative domains, M&L provide another piece of "evidence" against the ISC and the model outlined in Murphy (2021a). They say that "for models of cognition following hierarchical Bayesian approaches – which are getting more and more popular and influential (e.g., Tenenbaum et al., 2011) – there is no reason to assume that abstract concepts are more effortful to process", because "inference happens on many levels in the hierarchical [generative] model at the same time". Once again, M&L do not elaborate on why exactly the Bayesian brain hypothesis is incompatible with a psycholinguistic framework in which semantic complexity has parsing implications

– because there is, of course, no such reason. But alas for me, the "popular and influential" Bayesian neuroscientists have spoken. The matter is closed.

M&L cite polysemy processing literature which finds that "the switch from one sense to another (unlike from one meaning to another in the case of homonymy) does not seem to coincide with an increased processing effort – whether or not these senses are abstract or concrete". Crucially, this cited literature does not concern copredication. In addition, the experiments in Murphy (2021a) report a range of different switching contexts, e.g., Pred-Noun-Pred, Noun-Pred-Pred, Pred-Pred-Noun. Empirically documented acceptability dynamics cannot be sidelined simply because one thinks they are incommensurable with research that looks at isolated, single-word sense processing, or child language learning, or Bayesian brain architectures. This does not constitute satisfactory practice in psycholinguistics.

M&L also cite some exceptions to the ISC, discussed in Murphy (2021a), in which a Complex-Simple order is preferred. But finding one or two exceptions to a general trend in acceptability dynamics does not a rebuttal make, and certainly cannot provide a statistically meaningful rebuttal. M&L claim: "It seems that if concrete concepts really were preferable in copredication sentences, we would likely not expect these exceptions unless, of course, one could show why the concreteness effect is overwritten in those cases". Quite so. Hence why Murphy (2021a) invokes coherence relations; the cases in question involving a Complex-Simple preference appear to more saliently establish a legitimate coherence relation than the reverse. For example, M&L cite the following from Murphy (2021a):

(3) a. The entrance was being used to deliver food and was in need of repair.

b. The entrain was in need of repair and was being used to deliver food.

In the Simple-Complex case, (3b), an object in need of repair is then claimed to be used for a function, whereas in the reverse case, (3a), it is first established that the object is fit for function, so the subsequent predicate violates to a lesser extent any sense of incongruity. A comprehensive empirical investigation of these claims, which was also called for in Murphy (2021a), would involve manipulating *both* the coherence relations and the sense type, to test for direct interactions between these factors. Murphy (2021a) reports a post-hoc analysis that provides some preliminary evidence for the interaction

between these factors, but a systematic survey would be more conclusive. Instead of conducting these experiments, M&L resort to the unusual practice of casting doubt on empirically documented acceptability dynamics because they can point to one or two exceptions.

When presenting their own model for copredication acceptability dynamics, M&L say that "what drives acceptability intuitions about copredication sentences is the degree to which certain background expectations are fulfilled or violated. Linguistic intuitions arise, so our proposal says, by a process of prediction error minimization in a complex hierarchy of expectations ("priors" as we will call them), which constitutes a generative model of the world, also including knowledge of language".

This leads to an empirical prediction whereby, in cases where there are no differences in predictions for upcoming senses between Simple and Complex (i.e., both senses are 'balanced' in frequency), we will expect to see no effects of predicate ordering on acceptability. Yet, as mentioned, Murphy (2021a) found across a number of experiments that even in cases in which the two predicates have the same frequency, we still see a general Simple-Complex effect.

2. Anyone But You: Attempts to diverge from Murphy (2021a)

M&L propose a "predictive processing (PP) framework, which is increasingly popular in cognitive science". Throughout M&L's paper, one gets the impression that what is deemed trendy and "increasingly popular" in some areas of science, such as predictive processing (justifiably so; see Murphy et al. 2022), must therefore also be useful to solving puzzles about polysemy. But this is not self-evidently so.

For example, M&L cite research within the predictive processing framework that aims to explain syntactic ambiguity resolution purely in terms of prediction satisfaction. But there is much more to reading comprehension than prediction (Huang et al. 2024).

M&L effectively wish to scale up our theories of copredication, skipping over and brushing aside the many important insights that have been provided concerning the lexico-semantic, pragmatic, and syntactic factors that appear to enter into copredication judgments, in favor of a model that invokes our "overall world-model" (M&L) to account for acceptability judgments.

Borrowing another example from Murphy (2021a) without attribution, M&L use their model to explain the following:

(4) a. The newspaper has been attacked by the opposition and publicly burned by demonstrators.

b. #The newspaper was attacked by the opposition and fell off the table.

(4a) is deemed more acceptable, for M&L, because "there is a higher-level prior (expectation), in the form of a familiar situation in which objects are burned in a protest directed against what those objects represent. This situational prior allows us to expect, when reading/hearing "newspaper", both the institution as well as the physical object sense".

This goes no further beyond the explanation already provided in Murphy (2021a), in which a coherence relation (causal connection) is established.

M&L then expand their model in the following way: "The proposal is that in the absence of a rich context and sentence content—as is the case in most experiments by Murphy—by default we prefer an order from more objective to more subjective senses because this is the more reliable way to create common ground (the "Given" information) with the interlocutor."

Here, they inaccurately characterize the "experiments by Murphy", which concern basic structures like "The X was P and Q", or "The P and Q X", but also longer sentences with more complex predicates and prepositional phrases. For even more rich contexts, see Murphy (2021c) and the narratives tested in this experimental report. The important point to stress here is that even though M&L characterize ordering effects as obtaining at the sentence level, with "rich context", even basic phrases exhibit the ISC bias.

In any event, M&L's model attempts to diverge from my ISC model (i.e., a Simple-Complex ordering bias, interacting with coherence relations). They instead propose an Objective-Subjective ordering.

There is an immediate problem with this model. The reason why I shifted my model from Concrete-Abstract to the more general schema anchored around semantic complexity is because there are empirically documented cases (in Murphy 2021a) involving Abstract(Simple)-Abstract(Complex) ordering preferences. For example, *lecture*-type (INFORMATION, PROCESS), *newspaper*-type (INFORMATION, INSTITUTION), and

school-type (PROCESS, INSTITUTION) all exhibit clear Simple-Complex ordering preferences, with no meaningful mapping from this dimension to 'subjectivity'. For example, if a school is 'well located' and 'performed highly', surely the performance (the more abstract sense) is more 'objective' here than the quality of the location. If a college is 'composed of three buildings' and 'raising tuition fees', both of these are equally 'objective', yet the Simple-Complex sense order preference obtains (see experiments in Murphy 2021a).

M&L claim that their Objective-Subjective ordering model "seems very attractive from a conceptual point of view". This is debatable, and as is standard practice for M&L, they do not attempt to empirically test these claims, or ground them formally. Instead, there is only speculation. How exactly this model of subjectivity and objectivity relates to predictive processing is totally opaque. There will, for sure, be the occasional example that aligns neatly with Objective-Subjective preferences (ditto Subordinate-Dominant ordering preferences for sense frequency). But the goal of philosophy and psycholinguistics should not be to cherry-pick a handful of examples that align with one's own preconceptions (a grand total of 11 examples are critically discussed by M&L in direct connection with their model).

Another major obstacle for M&L is that their model is fully grounded in work exploring pre-nominal adjective ordering. But copredication can call upon a range of lexical classes, including verbs. Can processes and events denoted in copredication be inherently aligned along dimensions of subjectivity? This seems much less plausible than the model in Murphy (2021a) which purely calls upon semantic complexity and rather elementary pragmatic processes of coherence.

M&L insist that an advantage of their model is that "it can model cross-linguistic diversity" in copredication ordering effects. But Murphy (2021a) experimentally reports English and Italian ordering effects in line with ISC, whereas M&L only present a handful of English examples.

3. Additional Cases of Misdirection

Despite their attempts to introduce a sophisticated predictive processing model, when it comes to directly engaging with copredication M&L ultimately fall back on some

rather generic explanations. For example, when discussing why "The college has five floors and starts at 8am" sounds better than "The college starts at 8am and has five floors" (see Murphy 2021a for experimental data concerning acceptability), they state that "once the physical school building sense has established a physical referent, we can expect a sense of school-activity", and that this ordering gives us a "micro narrative" for our discourse about colleges, and in the Abstract-Concrete ordering "we do not understand the point of the second predicate given the first one". But of course, the same observation *also* obtains for the Concrete-Abstract ordering. And so we are back at square one – although M&L invoke "coherence" to explain the acceptability contrast, which is also the process posited in Murphy (2021a), supposedly the model that M&L are aiming to provide evidence against, not for.

This pattern of rejecting Murphy (2021a) before quietly explaining copredication via the precise same mechanisms as Murphy (2021a) also repeats itself when M&L discuss their last example (their sentence (22)). Again, they say that acceptability is sensitive to "a connection" between predicates established in discourse. This mysterious "connection" in question is, of course, none other than a coherence relation between predicates.

In a footnote, M&L write the following:

> "Given the theoretical nature of this paper, we have not conducted a statistical, empirical study to support claims about acceptability judgments for some example sentences for which no acceptability data is available in the literature. However, for those examples we have carried out an informal and preliminary survey among English speakers that support those intuitions. Rather than closing the empirical file here, one of the main objectives of this paper is to inspire new empirical research that is focused on context effects and the proposed "expectation architecture.""

What is somewhat queer about this is that one would assume that M&L would therefore provide at least some direction or (ironically) a prediction for their model, in terms of acceptability dynamics, or processing/parsing effects, in terms of judgment

ratings or reading times, and so forth. How might one test their "expectation hierarchies" for polysemous senses? This remains unclear.

M&L conclude by discussing "algorithmic" and "(neural) implementation levels", and end their paper by discussing trans-magnetic stimulation techniques, and other "popular and influential" things that unfortunately have no bearing on the matter of the present discussion. To be clear: M&L's model has no relation whatsoever to a "neuromechanical model of copredication"; simply invoking predictive processing does not automatically make one's model neurobiologically plausible and grounded. For more relevant literature from cognitive neuroscience on this matter, see Murphy (2015, 2020, 2024b) and Murphy et al. (2022).

Overall, M&L have a peculiar motivation. They aim to show that every single mechanism in ISC (Murphy 2021a, 2021b) is falsely motivated. Yet every alternative mechanism they put forth is either inadequate or can simply be reduced to a factor already accommodated in prior frameworks (Murphy 2019, 2021a, 2021b, 2021c).

References

Huang, K-J., Arehalli, S., Kugemoto, M., Muxica, C., Prasad, G., Dillon, B., & Linzen, T. (2024). Large-scale benchmark yields no evidence that language model surprisal explains syntactic disambiguation difficulty. *Journal of Memory and Language* 137: 104510. https://doi.org/10.1016/j.jml.2024.104510.

Michel, C., & Löhr, M. (2024). A cognitive psychological model of linguistic intuitions: Polysemy and predicate order effects in copredication sentences. *Lingua* 301: 103694. https://doi.org/10.1016/j.lingua.2024.103694.

Murphy, E. (2015). The brain dynamics of linguistic computation. *Frontiers in Psychology* 6: 1515. doi:10.3389/fpsyg.2015.01515.

Murphy, E. (2017). Acquiring the impossible: developmental stages of copredication. *Frontiers in Psychology* 8: 1072. doi:10.3389/fpsyg.2017.01072.

Murphy, E. (2019). Acceptability properties of abstract senses in copredication. In *Perspectives on Abstract Concepts: From Cognitive Processing to Semantic Representation and Linguistic Expression*, edited by M. Bolognesi, and G. Steen. Human Cognitive Processing Series. Amsterdam: John Benjamins. 145–165.

Murphy, E. (2020). *The Oscillatory Nature of Language*. Cambridge: Cambridge University Press.

Murphy, E. (2021a). *Linguistic Representation and Processing of Copredication*. PhD thesis, University College London.

Murphy, E. (2021b). Predicate order and coherence in copredication. *Inquiry* https://doi.org/10.1080/0020174X.2021.1958054.

Murphy, E. (2021c). Persistence conditions of institutional entities: investigating copredication through a forced-choice experiment. *Frontiers in Psychology* 12: 528862. doi:10.3389/fpsyg.2021.528862.

Murphy, E. (2023a). Copredication and complexity revisited: a reply to Löhr and Michel. *Cognitive Science* 46(10): e13207. doi:10.1111/cogs.13207.

Murphy, E. (2023b). The citadel itself: defending semantic internalism. *Global Philosophy* 33: 7. https://doi.org/10.1007/s10516-023-09669-z.

Murphy, E. (2024a). What is a word? arXiv:2402.12605.

Murphy, E. (2024b). ROSE: A neurocomputational architecture for syntax. *Journal of Neurolinguistics* 70: 101180. https://doi.org/10.1016/j.jneuroling.2023.101180.

Murphy, E., Holmes, E., & Friston, K. (2022). Natural language syntax complies with the free-energy principle. arXiv:2210.15098.

Murphy, E., Woolnough, O., Rollo, R.S., Roccaforte, Z.J., Segaert, K., Hagoort, P., & Tandon, N. (2022). Minimal phrase composition revealed by intracranial recordings. *Journal of Neuroscience* 42(15): 3216–3227. doi:10.1523/JNEUROSCI.1575-21.2022.

Tenenbaum, J.B., Kemp, C., Griffiths, T.L., & Goodman, N.D. (2011). How to grow a mind: statistics, structure, and abstraction. *Science* 331(6022): 1279–1285. doi:10.1126/science.1192788.