Misapplying working-memory tests: A reductio ad absurdum[*]

The preceding essay by Sprouse, Wagers, & Phillips (SWP2) sidesteps the objections we raised in our critique (HSS) of their 2012 article (SWP1). In the process, they attempt to deflect attention away from the fundamental issue at stake: is the experimental method used by SWP1 an effective diagnostic of the role of processing factors in acceptability contrasts? The answer to this question, as far as anyone knows, is "no". Hence, as we argue in HSS, the absence of a correlation between memory measures and acceptability contrasts is unsurprising and in no way undermines processing-based accounts of island effects.

Processing-based accounts of islands centered around working memory (WM) limitations like those of Kluender & Kutas (1993) and Hofmeister & Sag (2010) *do* predict a relationship between WM measures and judgments for islands *if and only if*: (1) the relevant WM measures tap the resources used in processing sentences with island-violating dependencies and (2) the relevant examples allow for individual variation to emerge.[1] But neither concern has been satisfied by SWP. They do not establish that their chosen memory measures (or any other) systematically relate to judgments for *any* hard-to-process but grammatical sentences, nor that this relationship applies universally across the spectrum of sentence processing difficulty.

SWP2 fail to even address point (2), while they attempt to defuse point (1) by a purely "in-principle" argument. In essence, they try to deflect concern about (1) by speculating that other WM tasks would have produced similar results, implying that there should be little concern about their choice of memory measures. Putting aside serious concerns about whether memory tasks are so interchangeable (and the unaddressed fact that the *n*-back and serial recall tasks are not WM tasks; see Kane *et al.* (2004)), this still leaves unanswered the question of how the magnitude of acceptability contrasts tied to resource limitations relates to memory measures. In other words, it is premature to argue about whether all memory tasks relate in the same way to judgments for hard-to-process structures if we don't know that *any* memory measure systematically predicts acceptability judgments for such structures! The relevance of (2) to this issue is not merely a theoretical concern, either. In practice, both HSS and SWP have independently validated objection (2). In HSS, we showed that perhaps the most widely used WM measure, the reading span, fails to predict variation in judgments for extremely hard-to-process sentences (see also Waters & Caplan (1996)), even though they predict variation for "moderately" difficult items. Indeed, both SWP1 & SWP2 fail to mention their finding that judgments for the poster child of grammatical, hard-to-process structures — multiple center-embeddings — were, in fact, higher for individuals with *lower memory scores* on their tests (Sprouse, 2009), contrary to the hypothesis that higher scores should equate to higher acceptability judgments for sentences that stress WM resources. This evidence belies SWP2's claim that "the naturalness or discourse felicity of our materials have little to do with the debate", since in any investigation of individual differences, the nature of the test materials

matters. For example, if the materials universally cause computational breakdown, then individual differences will be inconsequential.

What we can say for certain on the basis of this evidence is that there is no known *universally applicable* way in which memory measures relate to acceptability judgments. Beyond this, we are faced with a stark example of the conclusions we would be forced to draw if we accepted SWP's logic: if a failure to find a correlation between WM measures and acceptability decrements is evidence that grammatical constraints – and not processing constraints – underlie a contrast, then both HSS and SWP have "evidence" that multiple center embeddings are ruled out by grammar and not by processing.

In light of this reductio, we conclude that there is simply nothing new to be learned from either SWP1 or SWP2 about how grammatical constraints and processing factors like WM limitations interact in the processing of island structures.[2] Their conclusions that processing explanations of island effects are to be abandoned in favor of grammatical stipulations (island constraints) are not only unwarranted, they also lead us into a quagmire of non-evidence regarding the underpinnings of acceptability contrasts.

Notes

[*]Comments and suggestions from Herb Clark, Ted Gibson, Robin Melnick, and Tom Wasow helped us considerably in framing the content here. We also thank Jon Sprouse, Matt Wagers, and Colin Phillips for providing access to their materials and stimulating discussion of the issues.

[1]We are open to the possibility that the resource limitations are not strictly memory-related, e.g. processing bottlenecks may play an important role in island effects.

[2]This problem is further compounded by SWP2's claim that "whether the linguistic constraints should be understood as syntactic, semantic, or pragmatic in nature is orthogonal to the current debate," which makes it especially difficult to understand what such accounts predict about the relationship between grammar and WM limitations, e.g. these accounts are compatible with cross-linguistic variation; however, they would be just as compatible with a finding that languages appeared entirely uniform with respect to island effects.

References

HOFMEISTER, PHILIP, & IVAN A. SAG. 2010. Cognitive constraints and island effects. *Language* 86.366–415.

KANE, MICHAEL J., DAVID Z. HAMBRICK, STEPHEN W. TUHOLSKI, OLIVER WILHELM, TABITHA W. PAYNE, & RANDALL W. ENGLE. 2004. The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General; Journal of Experimental Psychology: General* 133.189–217.

KLUENDER, ROBERT, & MARTA KUTAS. 1993. Subjacency as a processing phenomenon. *Language and Cognitive Processes* 8.573–633.

Sprouse, Jon, 2009. Islands and the role of working memory in acceptability judgments. Talk presented at the 32nd Generative Linguistics in the Old World (GLOW 32). University of Nantes, France, April 16-18.

Sprouse, Jon, Matt Wagers, & Colin Phillips. 2012. A test of the relation between working memory and syntactic island effects. *Language* 88.82–123.

Waters, Gloria S., & David Caplan. 1996. Processing resource capacity and the comprehension of garden path sentences. *Memory & Cognition* 24.342–355.