Distinct Processing Demands of Production and Comprehension Shape Memory for Language and Language Use

Sarah Brown-Schmidt & Yu Zhou (Vanderbilt University)

Language is interpreted incrementally, with listeners considering multiple candidate referents over time[1-2]. When interpreting an expression like *the striped dress*, candidate referents that are temporarily consistent with the expression (e.g., striped *skirt*; *checkered* dress) tend to be remembered better than entities that never match the expression (e.g., *checkered dress*), a finding attributed to temporary activation of the candidate referents in memory[3]. In contrast, during language production speakers must consider how the referent (*striped dress*) can be distinguished from competitors (*checkered dress*) in order to design an appropriately modified expression (e.g., *the striped dress*)[4]. We speculate these distinct processing demands lead to differences in memory for linguistic contexts. Given that language use is guided by memory representations of the discourse history, we further predict that these memory differences will, in turn, shape *future* language processing. **E1-E2** (pre-registered) test these hypotheses. We interpret the findings in the context of a recent theoretical proposal[3] that different conversational roles (e.g., *speaker, listener, overhearer*) pose distinct cognitive demands, limiting the degree to which a common memory representation of the discourse history is possible.

E1: Eye-tracked participants (Ps, *N*=50) sat at separate computers, taking turns describing images in English. Conversations were unscripted and the addressee's task was to click it (**Fig1**). Scenes had a target (e.g., *striped dress*), an adjective competitor matching the *initial* part of the phrase (A-comp, *striped skirt*), a noun competitor matching the latter (N-comp, *checkered dress*), and an image that did not match (e.g., *checkered skirt*), and two fillers. After the conversation, Ps completed a 2AFC memory test for the images (**Fig2**). Results: Analysis of gaze using a dynamic *IRTree* model[5] examines gaze at three nodes; here we focus on Node 3 which contrasts fixations to A-comp and N-comp as a function of role (**Fig3**): At Node 3, an overall preference to fixate A-comp more than N-comp ($b = 0.29$, $p < .0001$) was mediated by a significant effect of role, with more N-comp fixations in production and more A-comp fixations in comprehension ($b = -0.14$, $p = .009$). These findings show that participant role (as speaker or addressee) differentially drives attention to candidate referents in the scene, with speakers looking more at production-relevant objects (e.g., the checkered dress when saying "*the striped dress*") and addressees looking more at comprehension-relevant items (e.g., the striped skirt when hearing "the striped dress"). Analysis of memory for candidate referents tests if processing differences shape enduring memory for context. In addition to an expected target boost to memory ($b = -.90$, $p<.0001$) and an overall benefit for candidate referents that temporarily matched the referring expression ($b = .32$, $p=.002$), the analysis revealed that former-speakers remembered N-comps better ($b = -0.42$, $p = .014$), whereas former-listeners remembered A-comps better ($b = 0.18$, $p = 0.11$). The findings are consistent with the hypothesis that distinct demands of speaking and listening result in different *memories*.

E2 (*N*=50) was identical except Ps played the game with an RA who described multiple images, e.g., "*the striped dress*", then 3 trials later *"the other dress"*, or "*the other striped one*". We predict the memory asymmetries observed in E1 will guide online interpretation of the term "other" as well as memory for the scene. Results**:** Data collection is complete, and the analysis of eye-tracking data is underway and will be complete in time for the conference. Initial analysis of the memory data reveals, as before, a memory advantage for named targets and candidate referents temporarily consistent with the unfolding speech stream.

Conclusion: The distinct demands of production and comprehension in conversation lead to distinct memories for the contexts of language use. We hypothesize and will test in Experiment 2 whether these distinct memories impact the subsequent language use. These findings are consistent with recent theorization[3] that conversational participants form distinct representations of the discourse history, placing processing-driven limitations on the potential for forming a common memory for past interactions.

***Figure1**. Example stimulus grid for testing the target* striped skirt. In Exp2 repeated trials, the target is either the adj competitor (*striped dress)* or noun competitor (*checked dress*).

***Figure 2.** Example 2AFC memory test trial.*

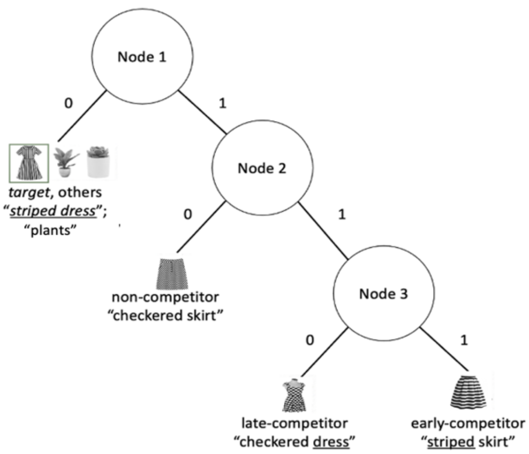

***Figure 3**. Tree diagram for 3-Node IRTree model. The model captures polytomous gaze data using three binary nodes.*

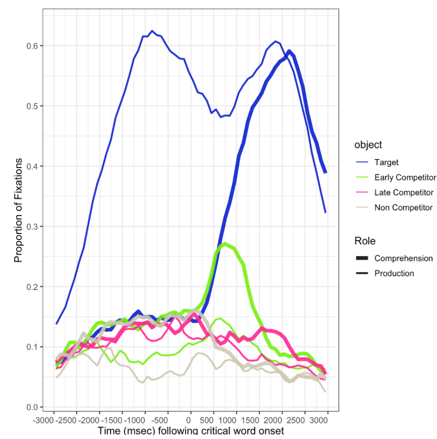

***Figure 4**. Average proportion of fixations to the critical images during production (thin lines) and comprehension (thick lines), aligned at critical word onset (*ms*) in **Exp1**.

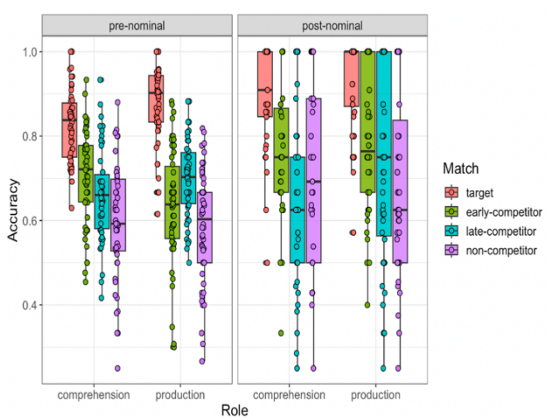

***Figure 5.** Accuracy in 2AFC memory test for images viewed during the communication task in **Exp1**.

References

[1] Eberhard, K.M., Spivey-Knowlton, M.J., Sedivy, J.C. et al. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *J Psycholinguist Res* 24, 409–436.
[2] Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of memory and language*, *38*(4), 419-439.
[3] Lord, K., & Brown-Schmidt, S. (2022). Temporary ambiguity and memory for the context of spoken language. *Psychonomic Bulletin & Review*, *29*(4), 1440-1450.
[4] Pechmann, T. (1989). Incremental speech production and referential overspecification. *Linguistics*, 27, 89–110.
[5] Cho, S.-J., **Brown-Schmidt**, S., De Boeck, P., & Shen, J. (2020). Modeling Intensive Polytomous Time Series Eye Tracking Data: A Dynamic Tree- Based Item Response Model. *Psychometrika, 85,* 154–184.