Adults, but not children, use spatial deixis to predict a speaker's referent

Tracy Reuter & Casey Lew-Williams (Princeton University)
treuter@princeton.edu

A number of theories claim that prediction supports language processing and learning.[1,2] Supporting this idea are findings that listeners can use informative verbs[3] and articles' number markings[4] to predict upcoming referents. However, precisely *how* prediction occurs remains largely unknown. Pickering and Garrod (2013) propose that prediction occurs via simulation: The listener uses language production mechanisms to simulate the speaker's likely production. Prediction via simulation thus requires taking the speaker's perspective. Moreover, the authors propose that children may lack requisite production mechanisms to predict via simulation.

We tested this claim by evaluating how adults and children comprehend spatial deixis. Spatial deixis (e.g., *this, that, these,* and *those*) could allow listeners to predict both the plurality and proximity of the speaker's referent. First, *'this'* and *'that'* are singular, whereas *'these'* and *'those'* are plural, so deixis number marking may support prediction.[4] Second, *'this'* and *'these'* typically indicate referents proximal to the speaker, whereas *'that'* and *'those'* typically indicate distal referents.[5] Critically, using spatial deixis to predict the referent's proximity would require taking the speaker's perspective, because *'this'* and *'that'* are typically defined by the speaker's perspective – the opposite of the listener's perspective in a face-to-face dialogue.[5] Thus, deictic words provide an optimal test case for evaluating prediction via simulation, allowing a comparison of number-based prediction[4] alongside the novel question of whether their spatial informativity can be used to predict a speaker's referent.

We hypothesized that adults, but not children, can predict via simulation.[2] To evaluate this hypothesis, we tested adults (n=28) and 5-year-olds (n=28) in an eye-tracking task. Participants viewed images of a speaker and four referents which differed in plurality and in proximity to the speaker (Fig. 1). Participants heard deictic sentences (e.g., *Look at that beautiful baby*) and neutral sentences (e.g., *Look at the beautiful baby*). We analyzed participants' looks to referents within 100-ms time-bins, correcting p-values with FDR.[6]

To evaluate whether participants can use deixis to predict the speaker's likely referent, we first analyzed looks to the target referent for deictic and neutral sentences (Fig. 2). We found that adults and children have greater target looks for deictic sentences (adults: 0–1000 ms, ps<0.05; children: 600–900 ms, ps<0.05). However, this analysis does not determine whether listeners predict the plurality of the referent,[4] the proximity of the referent,[5] or both.

To evaluate whether participants can use deixis to predict the *plurality* of the referent,[4] we analyzed looks to plural referents for deictic (i.e., comparing this/that vs. these/those) and neutral sentences. We found, for both adults and children, that the plurality effect began earlier for deictic (adults: -600–1500 ms, ps<0.05, children: -100–1500 ms, ps<0.05) than for neutral sentences (adults: 300–1500, ps<0.05, children: 500–1500, ps<0.05), suggesting that adults and children can use deixis number marking to anticipate the plurality of the upcoming referent.

In a third and critical analysis, we evaluated whether participants can take the speaker's perspective and use deixis to predict the *proximity* of the referent (e.g., use *'this'* to look at a referent close to the speaker). We analyzed looks to proximal referents for deictic sentences (i.e., comparing this/these vs. that/those) and for neutral sentences. For adults, the proximity effect began earlier for deictic sentences (200–1500 ms, ps<0.001) than for neutral sentences (400–1500 ms, ps<0.001). In contrast, for children, the proximity effect began later for deictic sentences (400–1500 ms, ps<0.05) than for neutral sentences (200–1500 ms, ps<0.05).

Together, findings are consistent with the hypothesis that adults, but not children, can predict via simulation.[2] These findings contribute to current theoretical debates about how prediction occurs during language processing. In ongoing experiments, we aim to further examine how language production experience, visual and linguistic context, and paralinguistic cues may jointly influence listeners' predictions during language processing.

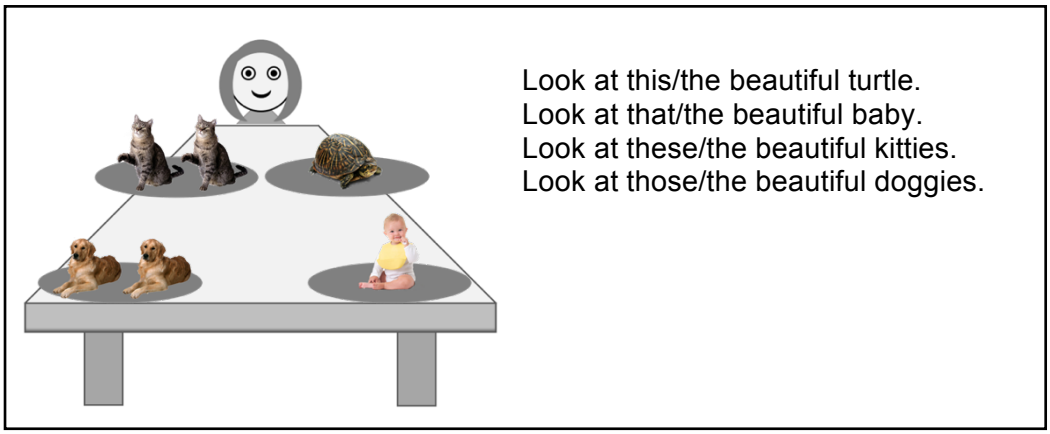

Figure 1: Example visual stimuli (left) and corresponding auditory stimuli (right).

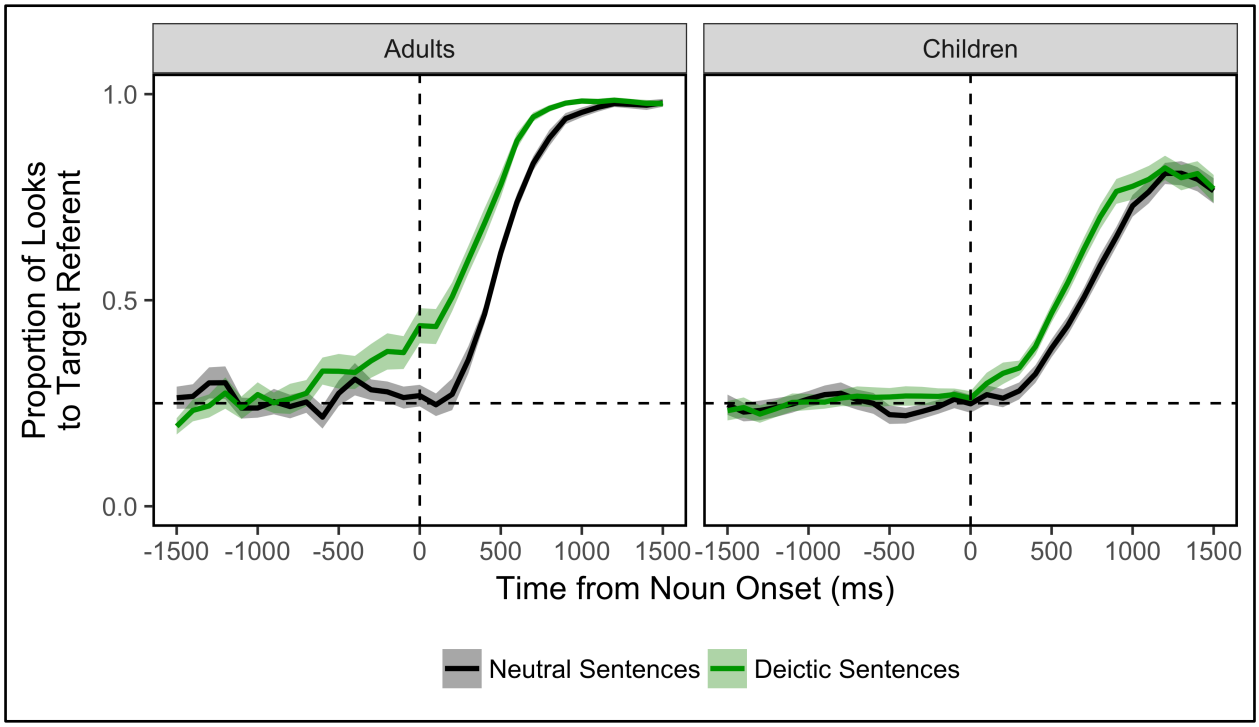

Figure 2: Proportion of target looks for adults (*n*=28) and children (*n*=28) during deictic and neutral sentences. Horizontal dashed line indicates chance performance. Vertical dashed line indicates the onset of the target noun (e.g., baby). Shading represents one standard error from the mean for each condition, averaged by subjects.

References

1. Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psych. Rev.*, *113*(2), 234-72.
2. Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *The Behavioral and Brain Sciences*, *36*(4), 329–347.
3. Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—But only for skilled producers. *J.E.P. H.P.P., 38,* 843-847.
4. Lukyanenko, C., & Fisher, C. (2016). Where are the cookies? Two- and three-year-olds use number-marked verbs to anticipate upcoming nouns. *Cognition*, *146*, 349-370.
5. Fillmore, C. J. (1997). *Lectures on deixis.* Stanford, CA: CSLI
6. Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. of the Royal Statistical Society, 57*(1), 289–300.