DO COMPREHENDERS PREDICT FROM THEIR OWN OR THEIR PARTNER'S PERSPECTIVE?

Ruth Corps, Charlotte Brooke, & Martin Pickering (University of Edinburgh)
Ruth.Corps@ed.ac.uk

Listeners often use the semantics of an utterance to associatively predict what they are going to hear. For example, they will preferentially look at edible objects immediately after hearing *The boy will eat…*, predicting that the speaker is about to mention such an object [1]. These predictions are fairly rapid, and appear to be automatic. But do listeners predict what they themselves would say if they were producing the utterance, or what they believe the speaker is likely to say?

It would be useful for listeners to predict appropriately, from the speaker's perspective, because their predictions will tend to correspond to what the speaker actually ends up saying. But are listeners able to "step into the speaker's shoes" in this way, given that prediction is time sensitive and cognitively demanding [2]? And are these predictions harder to make than associative predictions, because they require cognitive adjustment?

We addressed these questions in three experiments using the visual-world paradigm, in which participants listened to sentences (N=28; e.g., *I would like to wear…*) while viewing four objects on-screen. We created differences in perspective by creating differences in gender identity. Thus, we manipulated the gender of the speaker (as indexed by their voice and their face; [3]), the participants, and the characters in the sentences. In particular, participants heard a male or a female speaker producing sentences about gender-stereotyped objects displayed on-screen (gender stereotypy was assessed using a separate pre-test; N=80). One target (a dress) and one distractor (a hairdryer) were stereotypically female, while the other target (a tie) and distractor (a drill) were stereotypically male. Sentences began with *I* in Experiment 1, *You* in Experiment 2, and the name *James* or *Kate* in Experiment 3. To determine whether participants considered perspective when predicting, we fitted Bayesian generalized linear mixed effects models to binomial fixations in 50 ms time bins from 1000 ms before to 1500 ms after critical verb onset (*wear*).

In Experiment 1, participants (N=24, 12 males) fixated targets more than distractors from 500 to 1500 ms after verb onset (all *p*s < .05; Fig. 2), before the target was even named, suggesting listeners predicted associatively. Participants also fixated appropriate targets, which matched the speaker's gender, more than inappropriate targets, which matched their own gender, from 1000-1500 ms after verb onset (all *p*s < .05). This appropriate effect occurred later than the associative effect, suggesting that it took time to adjust for perspective and that this process is cognitively demanding.

We found similar effects in Experiment 2, in which sentences began with the pronoun *You* rather than *I*, so that appropriate prediction was not tied to the speaker's perspective. Participants (N=32; 16 males) predicted associatively from 450-1500 ms after verb onset (all *p*s < .05) and appropriately from 1250-1500 ms (all *p*s < .05). Note that this appropriate effect was later in Experiment 2 than Experiment 1, perhaps because there is some ambiguity as to who *You* may refer to [4]. But importantly, this appropriate effect again occurred later than the associative effect, providing further evidence that adjusting for perspective is demanding.

In Experiment 3, participants (N=32; 16 males) listened to sentences referring to a male (James) or a female (Kate) character. Participants predicted associatively from 450-1500 ms after verb onset (all *p*s < .05). They also predicted appropriately (looking at the target that matched the character's gender) from 600-1500 ms (all *p*s < .05), later than the associative effect.

We conclude that listeners take perspective into account when predicting, but not from the earliest moments of prediction. The difference in the time-course of associative and appropriate prediction is consistent with a multiple mechanisms account of prediction [5, 6], in which associative predictions are made automatically, while appropriate predictions (based on situational knowledge) are under strategic control.

References

[1] Altmann, G. T. M. & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition, 73,* 247-264.
[2] Ito, A., Corley, M., & Pickering, M. J. (2018). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism: Language and Cognition, 21,* 251-264.
[3] Van Berkum, J. J., Van den Brink, D., Tesink, C. M., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of cognitive neuroscience, 20,* 580-591.
[4] Brunyé, T. T., Ditman, T., Mahoney, C. R., Augustyn, J. S., & Taylor, H. (2009). When you and I share perspectives: Pronouns modulate perspective taking during narrative comprehension. *Psychological Science, 20*, 27-32.
[5] Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research, 1626,* 118-135.
[6] Pickering, M. J. & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin, 144,* 1002-1044.

Figure 1. Examples of stimuli used in the three experiments