Perspective-taking behavior as the probabilistic weighing of multiple domains

Daphna Heller (University of Toronto), Christopher Parisien (Nuance Communications), & Suzanne Stevenson (University of Toronto)

daphna.heller@utoronto.ca

Common ground; Perspective taking; Bayesian modeling; Visual world paradigm.

Perspective information distinguishes knowledge shared by both interlocutors (common ground) from knowledge privileged to one interlocutor. Studies ask whether listeners are sensitive to this distinction by examining the referential domain used in interpreting definites, arriving at apparently-contradictory conclusions: listeners use all objects available (including privileged) as the (egocentric) domain [1,2], or listeners use common ground as the domain [3,4]. We propose a novel approach where listeners simultaneously consider **both** domains, weighing them probabilistically. This captures the idea that a listener's perspective-taking behavior arises from considering their knowledge together with the speaker's, which must be inferred from indirect cues.

We model reference resolution as the probability of choosing an object *obj* given a referring expression *RE* and a referential domain *k*: P(*obj*|*RE*,*k*). Using Bayes' rule, this is proportional to: P(*RE*|*obj*,*k*)P(*obj*|*k*). P(*RE*|*obj*,*k*) captures the referential fit: how much is RE expected for each object in *k*; this depends on properties of all objects in *k*. P(*obj*|*k*) captures the prior likelihood that an object will be referred to, with objects privileged to the listener less likely referents. The model weighs the influence of two possible domains, egocentric (*k*=*e*) and common ground (*k*=*c*): αP(*RE*|*obj*,*k*=*e*)P(*obj*|*k*=*e*)+(1-α)P(*RE*|*obj*,*k*=*c*)P(*obj*|*k*=*c*) where α near 1 reflects a listener who weighs *e* more and is thus predicted to exhibit egocentric behavior; α near 0 is a listener who weighs *c* more, showing more adaptation to the speaker's knowledge. α is expected to vary across situations and listeners, depending on how the listener integrates various ground cues.

Our hypothesis is that the difference between [1] (egocentric behavior) and [4] (common ground behavior) is due to referential fit, and not different perspective-taking strategies (cf. [3,4]). Specifically, in our model the referential fit of the RE **in each of the two domains** influences reference resolution. Using elicited production data (A), we model comprehension conditions (B1-2), simulating experiments [1] and [4] respectively.

Production Experiment. Twenty participants instructed a confederate to click on images (interlocutors used different screens). RE data in four conditions (A) that correspond to domains *e* and *c* for (B1-2) were used to estimate the model's referential fit component. The model's response was qualitatively different for (B1) and (B2). In (B2) the target was a better fit to the RE than the competitor in domain *c*, and just as good a fit in domain *e*, so it should be preferred for any $\alpha \neq 1$. In (B1), the target was again a better fit to the RE in domain *c*, but not in domain *e* where the competitor was a better fit, so the target should be preferred for a smaller range of α, and more competition is expected from the (privileged) competitor.

Comprehension Experiment tested these predictions. Participants followed instructions from a confederate to click on images as their eye-movements were recorded. Preliminary results suggest (i) more looks to the competitor in (B1) than (B2) and (ii) response errors only in (B1). This supports our proposal that egocentric and common ground domains are considered simultaneously, allowing us to reconcile the apparently-contradictory results [1] and [4].

Production Experiment: displays from the participant's side (the speaker)
HIDDEN was a grey square with no object; the participant knew the listener saw an object.

(A1e)	Triplet-Domain *e*	big candle, medium candle, small candle, bowl.
(A1c)	Triplet-Domain *c*	HIDDEN, medium candle, small candle, bowl.
(A2e)	Pairs-Domain *e*	big candle, small candle, big bowl, small bowl.
(A2c)	Pairs-Domain *c*	big candle, small candle, big bowl, HIDDEN.

Comprehension Experiment: displays from the participant's side (the listener)
Objects privileged to the listener (**bolded**) had a grey background (not visible to the speaker).

(B1) Privileged-Triplet	**big candle** [competitor], medium candle [target], small candle, bowl.
(B2) Privileged-Pairs	big candle [target], small candle, big bowl [competitor], **small bowl**.
(B3) Control-Triplet	big candle, medium candle, small candle, bowl.
(B4) Control-Pairs	big candle, small candle, big bowl, small bowl.
Spoken instruction:	"Click on the big candle"

[1] Keysar et al. (2000). *Psych. Science*. [2] Keysar et al. (2003) *Cognition* [3] Hanna et al. (2003) *JML*. [4] Heller et al. (2008). *Cognition*.