

## Referring to Agents in an Artificial World: The Role of Predictability

Asya Achimova, Christian Stegemann-Philipps, Susanne Winkler, Martin V. Butz  
Research Training Group 1808 “Ambiguity: Production and Perception”  
University of Tübingen

When choosing referring expressions (definite NPs, pronouns, or null subjects) to describe agents, speakers face competing pressures. On the one hand, they should avoid using definite NPs to refer to a previously introduced entity—a constraint known as Repeated-name penalty (Gordon et al., 1993). Using a pronoun introduces a different kind of problem: if there are multiple potential antecedents in a scene, a pronoun may be ambiguous. In this paper, we claim that predictability of an event affects the choice of referring expressions in spontaneous descriptions of agents in novel scenes.

Recent corpus (Tily & Piantadosi, 2009) and experimental work (Rosa & Arnold, 2017) suggest that speakers choose more informative expressions to refer to unpredictable content. However, it is difficult to separate the predictability of the event itself from the effect of distributional properties of corresponding event descriptions. To avoid this confound, we introduce an artificial world paradigm where we trained participants to learn about typical interactive event progressions. The participants were then asked to describe similar event progressions verbally.

The scenes featured three distinct types of moving entities: blue cells, red cells, and green viruses (the experiment was conducted in pre-pandemic times). All events consisted of two sub-events: an attack (A attacked B) and fleeing (either A or B fled the scene) (Figure 1). We manipulated the predictability of the event progression that speakers had to describe. In half of the trials, the characters behaved in accordance with previously learned patterns, while in the other half an unexpected progression was shown. We expected the participants to produce more definite NPs when an unusual interaction is observed.

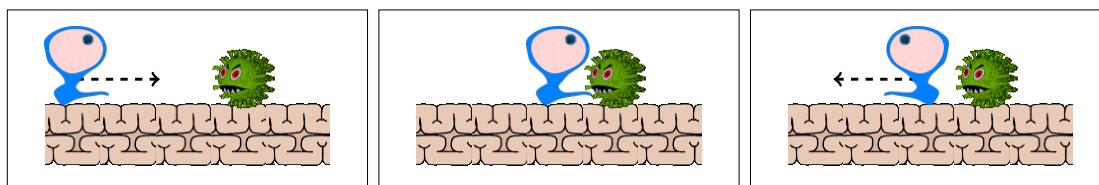


Figure 1: The blue cell attacks the virus and flees.

We anticipated that the participants will produce one of the descriptions in (1) for the event shown in Figure 1.

- (1)
  - a. The blue cell attacks the virus and the blue cell flees.
  - b. The blue cell attacks the virus and it flees.
  - c. The blue cell attacks the virus and flees.

The experiment followed a 2 x 2 design: we varied the thematic role of the fleeing character in the first sub-event (agent or patient) and whether the event proceeded as expected. Participants viewed and described two scenes for each of the four conditions, producing 8 descriptions in total.

We annotated the utterances as to whether the subject of the second verb in the event description was a definite NP or a pronoun. As an alternative, sometimes the speakers

produced a conjoined verb phrase (1c). We assigned such responses to category “null” for null separate subject of the second verb (the whole verb phrase has only one subject). Together null subjects and pronouns formed the “reduced” reference category.

We fitted the data with a generalized linear mixed-effects model, where the type of referring expression served as a binomial dependent variable, surprise status of an event and trial order as the independent variables, and items and subjects as random intercepts. We excluded first trials from the analysis as they featured considerably more reduced references than other trials. As we predicted, speakers produced more definite NPs when they encountered a surprising event ( $\beta = 1.417$ ,  $SE = 0.678$ ,  $z = 2.091$ ,  $p < 0.05$ ), see Figure 2, odds ratios for effects and their interaction are shown in Figure 3. The effect surfaced for events where the same entity acted as an agent in both sub-events. In case when the patient of the first sub-event became the agent of the second sub-event, participants overwhelmingly preferred definite NPs, we relate this effect to the need to clearly mark the change of agent between the two sub-events.

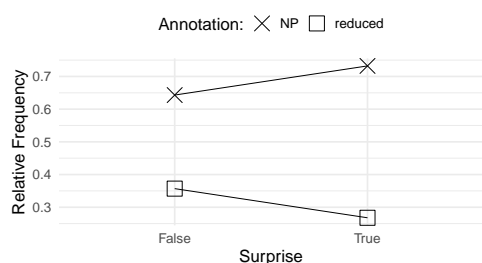


Figure 2: Definite NP vs reduced reference frequencies for surprising and unsurprising events, first trial excluded.

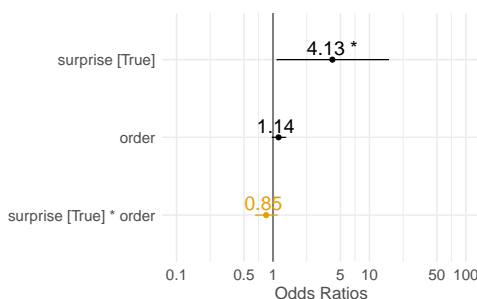


Figure 3: Odds ratios for effects and their interaction. Agents of surprising events are more likely to be referred to with a definite NP.

We conclude that the Repeated-name penalty applies even when there are multiple competing antecedents in the previous clause. The effect is stronger when speakers describe an event where the agent engages in predictable interactions with the patient: in that case speakers are more likely to choose a pronoun or use a conjoined verb phrase to describe the event. However, an unexpected turn in the unfolding of an event progression calls for a referring expression that could boost the salience of the agent to the listener, in accordance with the cooperative view of language production (Grice, 1989). Thus, we show that predictability of an event affects the choice of referring expressions: more informative expressions (definite NPs) are preferred over reduced forms to mark the agent of a surprising event. This effect illustrates the interaction of linguistic constraints on reference and the knowledge of event patterns.

## References

- Gordon, Peter C, Barbara J Grosz & Laura A Gilliom. 1993. Pronouns, names, and the centering of attention in discourse. *Cognitive science* 17(3). 311–347.
- Grice, P. 1989. *Studying in the way of words*. Cambridge, MA: Harvard University Press.
- Rosa, Elise C. & Jennifer E. Arnold. 2017. Predictability affects production: Thematic roles can affect reference form selection. *Journal of Memory and Language* 94. 43–60.
- Tily, Harry & Steven Piantadosi. 2009. Refer efficiently: Use less informative expressions for more predictable meanings. In Kees Van Deemter, Albert Gatt, Roger P. G. van Gompel & Emiel Krahmer (eds.), *Proceedings of the workshop on the production of referring expressions: Bridging the gap between computational and empirical approaches to reference*, Amsterdam.