Effect of surprisal on the form and prosody of referring expressions

Introduction. The production and comprehension of referring expressions involve a complex interplay of syntactic, semantic, and pragmatic factors, including principles such as Gricean maxims (Grice, 1975) and constraints on structural accessibility (Kamp et al., 2010). An important influence on how referring expressions are processed is the interlocutors' background knowledge (Winograd, 1972), especially their prior expectations regarding how plausible certain events are (Achimova et al., 2022, 2024; Stegemann-Philipps et al., 2021).

In an online experiment (Prolific, n = 139, native speakers of American English), we explored how people choose referring expressions to describe participants in familiar vs. surprising events. We used an artificial world depicted through short cartoons featuring three types of monsters: red, yellow, and blue. The monsters engaged in four interactions (attacking, rock throwing, waving, and jumping over) that could result in either the agent or the patient of the action falling over. We engineered situation-specific priors associated with the relative strength of different monsters. During the first phase of the experiment, participants learned that the red monster was stronger than the yellow, and the yellow was stronger than the blue. Regardless of who initiated the interaction, the weaker monster always fell to the ground. This established participants' expectations for the second phase, where they observed an interaction between two monsters and were asked to describe what they saw aloud. The first two trials involved familiar events, where the outcome conformed to the relative monster strength learned during the training phase. The next two trials presented surprising events, where the outcome contradicted the learned monster strength. The final two (post-surprising) trials showed conforming events again. We were interested in which form of anaphoric reference (noun phrase, pronoun, or zero anaphor) participants would use to refer to the falling monster:

(1) *The yellow monster attacked the red monster and (∅ / it / the yellow/red monster) fell down.*

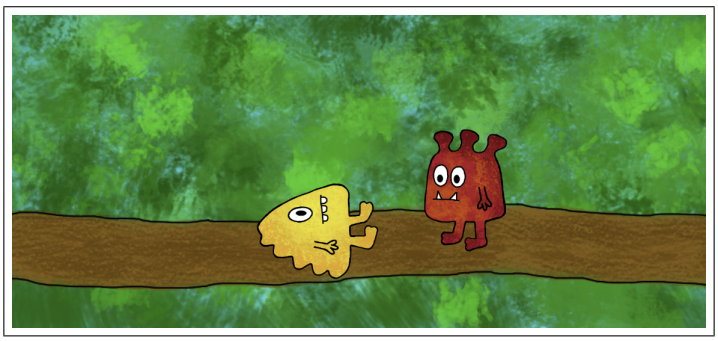


Figure 1: Sample trial.

Choice of referring expressions. When describing surprising events, participants used definite noun phrases (NPs) more frequently than when describing familiar events (Figure 2). A cumulative-link mixed-effects regression model revealed that this effect is present for both agent and patient antecedents ($\beta = 1.48, SE = 0.31, z = 4.73, p < 0.001$). In surprising trials, participants used more overt, unambiguous expressions to prevent potential misinterpretation on the listener side. They also tended to avoid zero anaphors, possibly anticipating that in an imperfect communication channel, listeners can reconstruct a pronoun in their place if this would allow them to arrive at a more plausible interpretation.

Even though the last two (post-surprising) trials featured the same events as the first two (familiar ones), the production rates were different from familiar events ($\beta = 1.23, SE = 0.30, z = 4.04, p < 0.001$) and not significantly different from surprising ones ($\beta = -0.25, SE = 0.30, z =$

$-0.83, p = 0.407$). This reflects the process of belief revision: after observing surprising events, participants' prior expectations about the plausibility of events changed.

Prosodic analysis. In the previous analysis, we established that participants are more likely to use definite NPs to refer to patient antecedents compared to agent antecedents and to mark surprising events compared to familiar ones. However, it is possible that event predictability shapes not only the lexical choice of reference but also how it is expressed prosodically. For prosodic analysis, we selected data from 54 participants (324 utterances in total) based on recording quality and dialect (standard American English) to perform tonal analysis. Each target utterance was tonally annotated following the ToBI conventions (Beckman & Hirschberg, 1994) in PRAAT (Boersma & Weenink, 2008), in which nuclear and prenuclear accents (T*) and boundary tones at the phonological phrasal (T-) and intonational phrasal level (T%) were annotated, as well as the word level parsing. The intonational phrase boundaries were especially relevant because they were used as a reference point in designating nuclear accents.

The results of the pitch accent analysis revealed a robust and previously unreported pattern linking surprisal to nuclear accent placement (Figure 3). A cumulative link mixed-effects model showed that when speakers use an overt form of reference (a pronoun or a definite NP to refer to a previously mentioned character), they are significantly more likely to place a nuclear accent to mark actors of surprising ($\beta = 3.501, SE = 0.529, z = 6.628, p < 0.001$) and post-surprising events ($\beta = 2.113, SE = 0.391, z = 5.400, p < 0.001$) compared to familiar ones. The analysis revealed no difference between patient and agent antecedents in the placement of the nuclear accent and no interaction between antecedent position and event type.

We further examined the type of accent placed independently of its location, whether prenuclear or nuclear. High peaking accent (H*) was the most used accent in all trials. Overall, 78% of all accents (nuclear and prenuclear) were H*. H* had a rate of 87% among the prenuclear accents and 71% among the nuclear accents. The next most frequently found accent type was the LH*, which involves a steep rise within the accented syllable (12% of all accents, 4% of prenuclear accents, 19% of nuclear accents). Due to its steep rising properties, LH* is often linked to heightened degrees of informativeness in intonational languages (Baumann et al., 2021). We used a binomial mixed-effects regression to evaluate whether the placement of LH* depended on the status of the event (familiar vs. surprising or post-surprising) and the role of the antecedent (agent vs. patient). Indeed, speakers were more likely to use LH* to mark referents of surprising ($\beta = 2.689, SE = 1.104, z = 2.435, p = 0.015$) and post-surprising events ($\beta = 2.560, SE = 1.107, z = 2.312, p = 0.021$). The analysis revealed no difference between agent and patient antecedents with respect to the likelihood of LH* tone placement.

Taken together, these findings demonstrate that prosodic prominence is a key component of how speakers encode and communicate event-related information, and morphosyntactic and prosodic components should be jointly modeled when investigating referring expressions, degrees of familiarity, and ambiguity management. The consistent presence of nuclear accent on subjects in surprising conditions suggests that speakers mark these referents as informationally important, likely because they deviate from listeners' prior expectations. Conversely, the absence of nuclear accent in predictable conditions indicates that no such marking is necessary when the event outcome is consistent with shared background knowledge. Importantly, the fact that these effects are observed regardless of whether the subject is realized as a noun phrase or a pronoun suggests that prosody provides an additional, partially independent channel for signaling familiarity/surprisal scale. Finally, the observed dissociation between referential form and prosodic prominence highlights the need for integrated models that jointly account for morphosyntactic and prosodic choices.

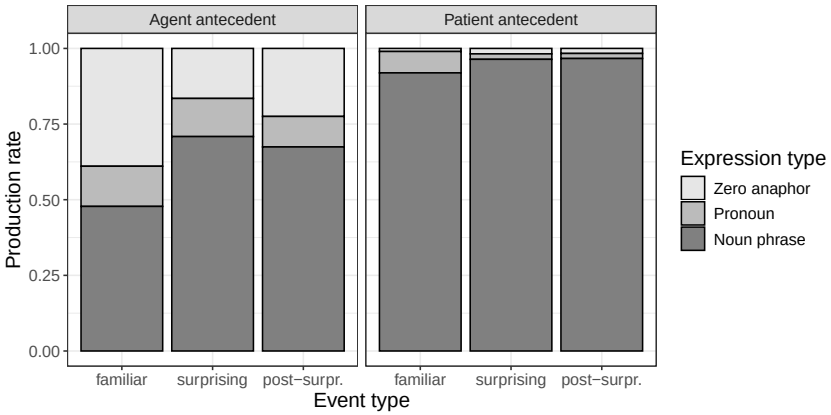


Figure 2: Produced referring expressions.

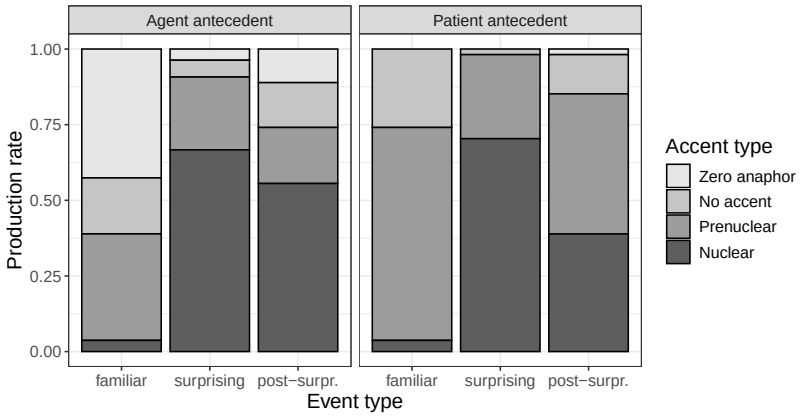


Figure 3: Produced accents (data from 54 participants).

References

Achimova, A., Stegemann-Philipps, C., Winkler, S., & Butz, M. V. (2022). Anaphoric reference in descriptions of surprising events. *Proceedings of Sinn und Bedeutung*, *26*, 20–34.

Achimova, A., van Os, M., Demberg, V., & Butz, M. V. (2024). Interpreting implausible event descriptions under noise. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *46*, 3399–3406.

Baumann, S., Mertens, J., & Kalbertodt, J. (2021). The influence of informativeness on the prosody of sentence topics. *Glossa: a journal of general linguistics*, *6*(1).

Beckman, M. E., & Hirschberg, J. (1994). The tobi annotation conventions. *Ohio State University*.

Boersma, P., & Weenink, D. (2008). Praat: Doing phonetics by computer (version 6.0. 37)[computer program]. *retrieved August*, *12*, 2008.

Grice, H. P. (1975). Logic and conversation. In *Speech acts* (pp. 41–58). Brill.

Kamp, H., Van Genabith, J., & Reyle, U. (2010). Discourse representation theory. In *Handbook of philosophical logic: Volume 15* (pp. 125–394). Springer.

Stegemann-Philipps, C., Butz, M. V., Winkler, S., & Achimova, A. (2021). Speakers use more informative referring expressions to describe surprising events. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *43*(43), 2897–2903.

Winograd, T. (1972). Understanding natural language. *Cognitive psychology*, *3*(1), 1–191.