

## Effects of predictability and optionality on pronominalization

Jet Hoek & Hannah Rohde (University of Edinburgh)

jet.hoek@ed.ac.uk

The idea that predictability influences production has been observed at multiple levels of linguistic structure – e.g., phonetics [1], morphology [2], and syntax [3]. Evidence for the influence of predictability on reduction in speakers' choice of reference, however, is mixed. Work in this domain typically targets the 3<sup>rd</sup> person pronoun and typically uses a referent's rate of re-mention in a subsequent sentence as a measure of predictability. Some studies show more pronominalization for more predictable referents [4,5], while others find no influence of predictability and instead link pronominalization to other factors exclusively related to referent topicality [6,7]. Disentangling these factors is a challenge as it is difficult to manipulate one factor while holding all others constant. Here we test pronominalization in contexts that permit manipulations of predictability while addressing prior confounds of referent optionality.

The primary evidence for the impact of predictability on pronominalization [4,5] shows that a subject referent is more likely to be pronominalized when it corresponds to a thematic role that is favored for subsequent re-mention. However, in such cases, the thematic role manipulation creates systematic differences in the obligatory-vs-optional status of the competing referent.<sup>1</sup> If obligatory arguments are more integral to the event denoted by the verb, they may compete more with the referent who corresponds to the favored thematic role. In this study, we explicitly manipulate the optionality of the competing referent, enabling us to better test how referents' predictability influences subsequent rates of pronominalization.

Participants (N=65) wrote story continuations for prompts that varied in the presence and status of a competing referent (see (1)). We held constant thematic and grammatical role of the target referent (subject *Patrick*) and the order in which the referents were mentioned. To do this, we used two forms of subject-biased implicit causality verbs [8]: predicate adjectives (1a-b) and transitive verbs (1c). A norming study (N=21) allowed us to select the 18 items that were judged to be most interchangeable in use and meaning between (1b-c) out of a larger set of 30 verbs.

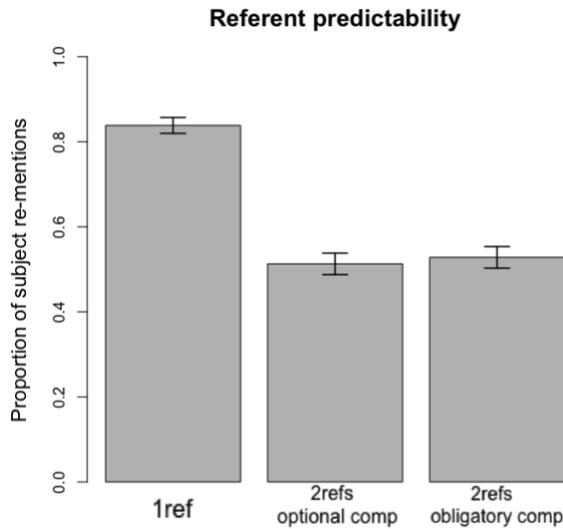
- |                                |                                    |
|--------------------------------|------------------------------------|
| (1) a. Patrick is boring.      | 1 referent                         |
| b. Patrick is boring to Emily. | 2 referents, optional competitor   |
| c. Patrick bores Emily.        | 2 referents, obligatory competitor |

The binary outcomes for re-mention (subject or not) and referential form (pronoun or not) were analyzed using LMER models (maximal RE structure permitted by the data). A main effect of condition was found for both outcomes, and we used pairwise comparisons to follow these up. Figure 1 shows the proportion of continuations that started with a re-mention of the subject referent and Figure 2 shows the rates of pronominalization of the subject. Crucially, the difference in predictability between the first two conditions (1a-b,  $p < .001$ ) does not yield a difference in pronominalization between those conditions ( $p = .61$ ). Furthermore, pronominalization rates between the latter two conditions (1b-c) differ ( $p < .05$ ) despite their matched predictability rates ( $p = .70$ ). Pronominalization is thus no higher for a more predictable referent (1a vs 1b) but it is higher when the competing referent is optional rather than obligatory (1b vs 1c).

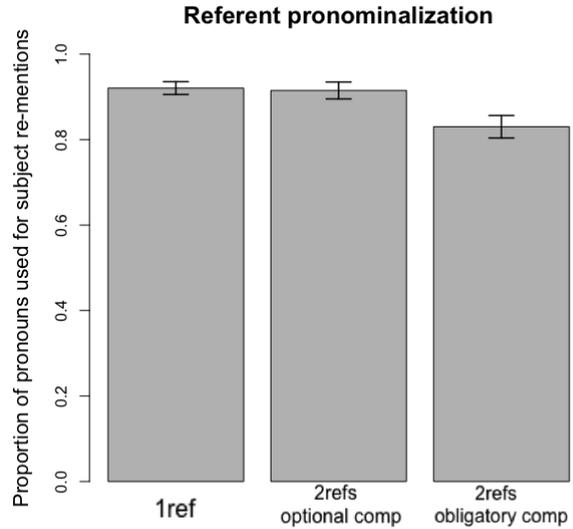
We tried to replicate these findings in a second experiment (N=54). The experiment included only conditions (1b-c), but the rest of the experiment and analysis remained the same. As in Exp. 1, re-mention rates did not differ between conditions (1b-c,  $p = .51$ ). However, this time we did not find a difference in pronominalization rate of the subject referent between conditions (1b-c,  $p = .73$ ). A third replication attempt with all three conditions (N=63) did not find the difference in pronominalization rate between any of the conditions ( $p = .98$ ). Over the whole dataset, the difference in pronominalization rate between conditions (1b-c) is also not significant, see Figure 3. As such, our study has not found any evidence that optionality played a role as a potential confound in [4,5] that could account for the differences between [4,5] and [6,7].

## Footnote

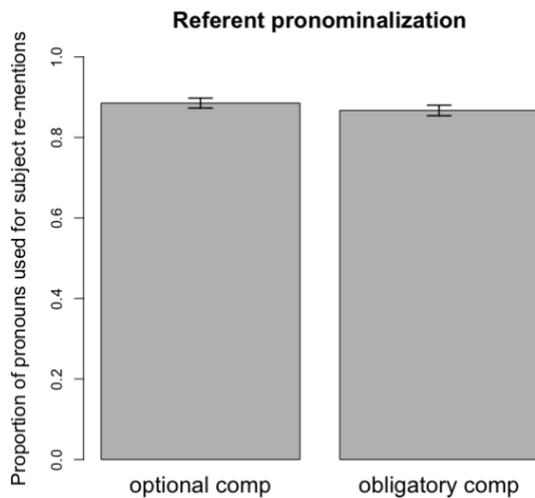
<sup>1</sup> For the well-studied class of transfer verbs, manipulating the thematic role of the subject referent tends to yield non-minimal pairs: Events described with the Source referent in subject position have an obligatory Goal argument (e.g., *Kyle gave a book to Sue*), while events described with the Goal referent in subject position have an optional Source (e.g., *Sue got a book [from Kyle]*).



**Figure 1:** Proportion of subject re-mentions out of all continuations from Experiment 1. Pattern of results is the same with analysis restricted to subject and object continuations. Error bars represent the standard error of the mean.



**Figure 2:** Proportion of pronouns used for subject re-mentions in Experiment 1. Error bars represent the standard error of the mean.



**Figure 3:** Proportion of pronouns used for subject re-mentions for conditions (1b-c) in all data collected in Experiments 1-3.

## References

- [1] Gahl, S. (2008). "Time" and "thyme" are not homophones: Word durations in spontaneous speech. *Language*, 84(3), 474-496.
- [2] Frank, A. F. & Jaeger, T. F. (2008). Speaking rationally: Uniform information density as an optimal strategy for language production. In *Proceedings of the 30th annual conference of the Cognitive Science Society*, 933-938.
- [3] Ferreira, V. S. & Dell, G. S. (2000) Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology* 40, 296-340.
- [4] Arnold, J. E. (2001). The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes* 3, 137-162.
- [5] Rosa, E. C. & Arnold, J. E. (2017). Predictability affects production: Thematic roles can affect reference form selection. *Journal of Memory and Language* 94, 43-60.
- [6] Fukumura, K. & van Gompel, R.P.G. (2010). Choosing anaphoric expression: Do people take into account likelihood of reference? *Journal of Memory and Language* 62, 52-66.
- [7] Rohde, H. & Kehler, A. (2014). Grammatical and information-structural influences on pronoun production. *Language, Cognition, and Neuroscience* 29(8), 1-16.
- [8] Garvey, C., Caramazza, A. & Yates, J. (1976). Factors Underlying Assignment of Pronoun Antecedents. *Cognition* 3, 227-243.