

Overspecifying state information in the production of referring expressions

Megan Parker and Daphna Heller (University of Toronto)

megan.parker@mail.utoronto.ca

When speakers choose a referring expression for an object, they may include different amounts of information and types of properties. For example, when referring to a door that is big, brown, and open, a speaker can say *the door*, *the big brown door*, or *the open door*; all could potentially allow an addressee to identify the intended object. It has been observed that speakers usually follow Grice's (1975) maxim of quantity and include the minimal amount of information. But not always: speakers have also been shown to include unnecessary information that does not serve to differentiate the intended referent from other objects in the context, a phenomenon known as overspecification [1]. Some modifiers, such as scalar adjectives (e.g., *long*, *small*), are rarely overspecified, whereas others, most notably color modifiers, are overspecified at a high rate [2,3]. Interestingly, overspecification of color modifiers occurs at a low rate when the color of the referent is predictable (e.g., a yellow banana) [4]. One account suggests that speakers tend to overspecify properties that are atypical of the object (e.g., a blue banana) [5]. Here we ask whether this logic extends to state modifiers (e.g., *dirty/clean*, *open/closed*), examining whether state information will be overspecified for objects in atypical states. A secondary question is whether overspecification is more likely when the context requires producing a (different) modifier (and thus speakers are already planning a modifier), or when no modifier is needed.

Method. In a referential communication task in a lab setting, participants performed the role of director, with the experimenter as the addressee. On each trial, four images appeared and one was highlighted (1); participants were asked to instruct the addressee to click on the highlighted image. We had a 2x2 design. STATE manipulated the state of the target object (e.g., an open vs. a closed door). States were classified as (un)marked by 7 individuals; see Table 1. CONTEXT manipulated whether the target was the only object of this nominal category (singleton), or whether it appeared along a second, identical image contrasting in size (pair). This was done to examine whether the requirement for another modifier raises or reduces the likelihood of overspecification. The experiment consisted of 148 experimental items as well as 74 fillers.

Results ($n=32$). Trials in which the participant was under-informative or provided an incorrect object label were eliminated (6%), resulting in a total of 4434 usable trials. Our first analysis concerns the rate of state modifiers – all cases of overspecification – across the four conditions (Table 2). A mixed-effects logistic regression model with state (marked vs. unmarked) and context (singleton vs. pair) as predictors revealed, first, a significant main effect of state ($z = -9.69$, $p < .001$); speakers produced significantly more overspecifications for marked states (21%) than unmarked states (1%). The main effect of context was also significant ($z = -3.25$, $p = .001$), indicating that the production of unnecessary modifiers is more likely when speakers do *not* need to produce another modifier that is required for referential success (the interaction was not significant, $p = .58$). We also examined markedness by individual object (an example is given in Table 3). For most objects, the marked state was the one overspecified, but for some, it was the state generally classified as unmarked that was mentioned. For example, *open* was mentioned for several objects, but for *laptop* and *eye*, *closed* was mentioned. Similarly, for some objects both states of *empty* and *full* were mentioned (e.g., *jar*, *wine glass*), and for others only one of the two adjectives was ever mentioned (e.g., *empty* for *egg carton*).

Conclusions. We find, first, that objects in atypical states indeed elicit overspecification of state modifiers. While some states were overall marked and thus overspecified more, the pattern seems to depend on the particular object. Second, we demonstrate that overspecification is more likely when another modifier is not being planned. We note that the rate of overspecification of state for marked states (21%) is much lower than previously observed rates for color (Sedivy, 2005: ~40%; Brown-Schmidt & Konopka, 2011: 78%). This may relate to findings from visual memory [6], where an object's color was forgotten more quickly than its state, suggesting that color properties are not initially bound to the object representation,

whereas state information is more unified with the object identity. We propose that the same representations play a role in how likely properties are to be encoded in referring expressions.

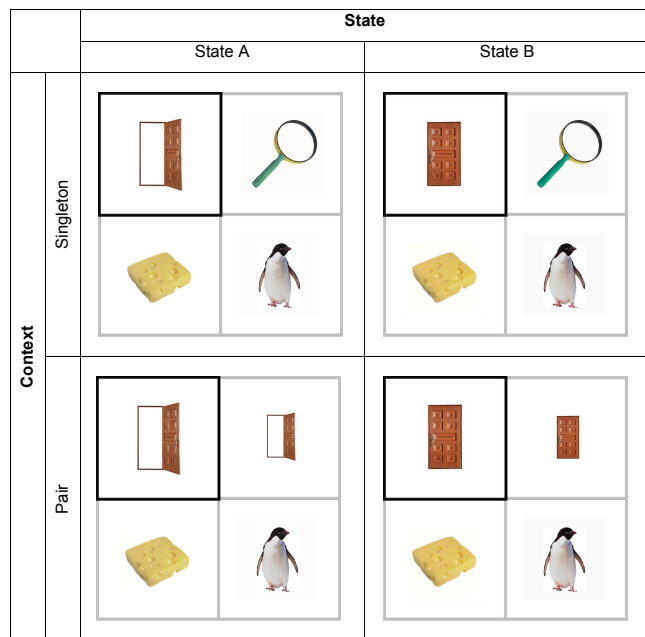


Figure 1. Example experimental displays with an object in two different states (open vs. closed) and in two different contexts (singleton vs. pair).

Marked	Unmarked	Number of items
Open	Closed	50
Broken/ripped	Whole	22
Dirty	Clean	19
Bent/ curved	Straight	18
Sliced	Unsliced	9
Folded	Unfolded	8
Empty	Full	6
Peeled	Unpeeled	6
Rolled	Unrolled	4
Wet	Dry	4
Tied	Untied	2

Table 1. Markedness classification of state pairs

	Marked	Unmarked	
Singleton	24%	1%	13%
Pair	18%	0.5%	9%
	21%	1%	11%

Table 2. Rates of overspecification of state modifiers.

empty	egg carton	jar	wineglass	hamper	garbage	shopping cart
	29% (4/14)	13% (2/15)	6% (1/16)	0	0	0
full	shopping cart	garbage	jar	hamper	wineglass	egg carton
	19% (3/16)	8% (1/12)	7% (1/15)	7% (1/14)	6% (1/16)	0

Table 3. Rates of overspecification for modifiers *empty* and *full*.

References

- [1] Pechmann, T. (1989). Incremental speech production and referential overspecification. *Linguistics*, 27, 89–110.
- [2] Brown-Schmidt, S., & Konopka, A. E. (2011). Experimental approaches to referential domains and the on-line processing of referring expressions in unscripted conversation. *Information*, 2, 302–326.
- [3] Sedivy, J.C. (2005). Evaluating explanations for referential context effects: Evidence for Gricean mechanisms in online language interpretation. In J.C. Trueswell & M.K. Tanenhaus (Eds.), *Approaches to Studying World-situated Language Use: Bridging the Language as product and Language as action Traditions* (153-171). Cambridge: MIT Press.
- [4] Sedivy, J.C. (2003). Pragmatic versus form-based accounts of referential contrast: Evidence for effects of informativity expectations. *Journal of Psycholinguistics*, 32, 3–23.
- [5] Kreiss, E., Degen, J., Hawkins, R. X. D., & Goodman, N. D. (2017) *Mentioning atypical properties of objects is communicatively efficient*. *XPRAG 2017*, Cologne. Poster retrieved from <https://osf.io/8vpj5/>
- [6] Brady, T., Konkle, T., Alvarez, G., & Oliva, A. (2012). Real-world objects are not represented as bound units: Independent forgetting of different object details from visual memory. *Journal of Experimental Psychology*. 142, 791-808