

Overview of Grodner et al. (2010)

Chris Potts, Ling 236: Context dependence in language and communication, Spring 2012

April 16

1 Background

- Central question of Grodner et al. (2010) (henceforth G10): Is there a processing cost to interpreting *some* as *some but not all*?
- Like Huang and Snedeker (2009) (HS09), G10 put subjects in an experimental situation in which the relevant scalar implicatures are highly relevant, in that they are crucial for fast object identification.
 - Thus, neither paper addresses the question of whether generalized implicatures arise as defaults in other kinds of context.
 - Rather, they address how quickly such inferences are made where they are relevant.
 - (Compare with Breheny et al. (2006), who sought to manipulate the contextual relevance of the target implicatures.)

2 Materials

G10 adapt the method of HS09. Using the visual-world eye-tracking paradigm, they analyze how quickly subjects fixate on the target item while listening to prerecorded audio sentences like (3).

(3) Click on the girl who has

a. ...summa the balls.

(Early-summa in fig A, Late-summa in fig. B)

b. ...alla the ballons

(Alla)

c. ...nunna the items.

(Nunna)

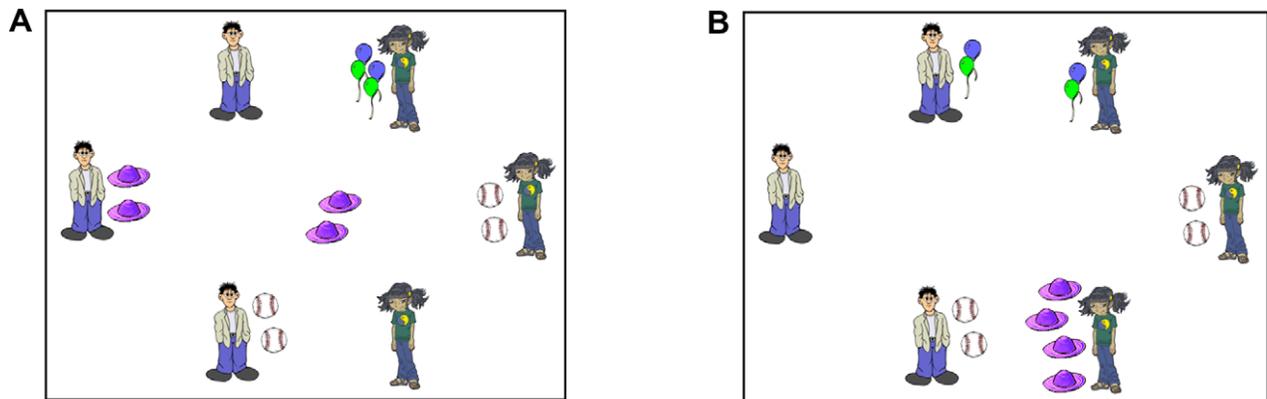


Figure 1: G10's visual scenes.

3 Central differences from Huang and Snedeker (2009)

- (4) G10 use more complex displays with more figures and more objects.
- (5) G10 had a Nunna (none of) condition as a literal control in addition to Alla.
- (6) G10's had 40 fillers (32 test items), whereas HS09 used no fillers.
- (7) G10 use reduced forms of the quantifiers to provide an earlier cue that the partitive is coming. (See p. 46, right, on what this did to the vowels and the length of the nominal.)

Rationale:

- a. The implicatures of *some* are robust only for the partitive. See Degen et al.'s (2009) gumball paradigm, in which gumballs drop from a higher chamber into a lower one. *You got some gumballs* was accepted in situations where all of the gumballs dropped, whereas *You got some of the gumballs* was rejected (p. 44, left).
 - b. Appendix A study: aligns with Degen et al.'s finding that the partitive is less natural than the bare form when picking out all (p. 51, right).
 - c. Thus, it could be that people waited to compute implicatures until they heard *of*, whereas there is no need to wait for the other quantifiers.
- (8) G10's instructions included a description of the domain of discourse that provided the counts of all the items in it (*There are four balls, four planets, and four balloons*; p. 45, right).

Rationale:

"This was to enhance the salience of the full set of each object type as a means of identifying a referential candidate. This potentially makes the contrast between full sets and sub-sets more prominent and could thus facilitate the comparison of alternatives that leads to the scalar inference." (p. 45, left).

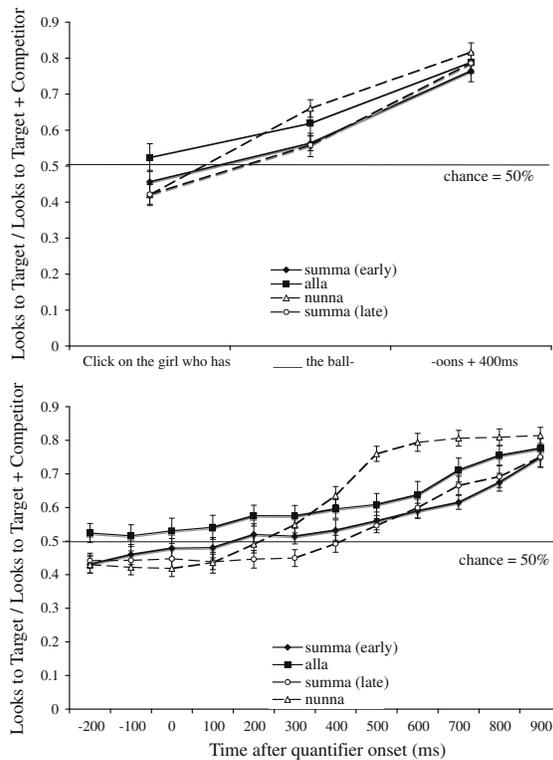
- (9) G10 had no numerical conditions; the only numerical quantifiers people heard were in preliminary descriptions of the domain of discourse.

Rationale:

- a. Number terms are more natural than *some* where the number of objects is in the subitizing range.
 - b. Appendix B study: including number terms reduced the naturalness of using *some of*.
- (10) G10 had a Late-summa condition in which even the enriched meaning was ambiguous until the descriptive content of the nominal (*socks vs. soccer balls*).

Rationale: Could further illuminate when people make pragmatic inferences (p. 45, right). Are they willing to do it even when the enriched meaning of the quantifier still under-specifies the referent?

4 Results



(a) G10, fig 2.

(b) HS09, fig. 5.

Figure 2: G10 central timing result compared with those of HS09, experiment 2.

- (11) Prior to the quantifier, there is a bias for looking at the Alla target. This is due to a bias in the visual system for looking at more complex images; HS09 saw these effects as well.
- (12) Thus, the denominator in the y-axis measurement in fig 2, top (given above), combines looks to the target with looks to the Alla target. For Alla, the denominator includes Summa targets. (For discussion of this, see p. 51, right, and fn. 7 on how HS09 addressed this.)
- (13) In the Quantifier interval, looks to the target are above chance for all determiners (p. 47, left). **This is a central contrast with HS09.** Late-summa shows the least gain, due to the fact that it remains ambiguous in this region even if enriched.
- (14) Fig 2, bottom, looks at the 100ms after the quantificational determiner (p. 47, right). The baseline corrects for the bias for Alla targets. Convergence in this region is reliable for all conditions except Late-Summa (which is only fully disambiguated later).
- (15) For Late-summa trials, subjects reliably shifted to one of the two Summa targets within the 100ms interval after the determiner (p. 48).

(16) Is convergence reliably more robust for the Alla condition? If so, that would be in line with HS09's findings. To find support for accepting the null hypothesis that Alla and Summa do not differ, G10 use a Bayesian method (p.49):

- Null hypothesis: the mean target convergence values for All and Early-summa are identical. (Null prior = normal distribution based on the log-odds proportions for Early-summa.)
- Alternative hypothesis: the mean target convergence values for All and Early-summa are maximally dissimilar in that Alla conditions lead to perfect convergence in this 100ms interval and Early-summa conditions lead to 0 convergence in this interval.
- Finding: the null hypothesis is vastly more likely given the likelihood function of the Alla data than is the alternative hypothesis within the range of possible effect sizes.

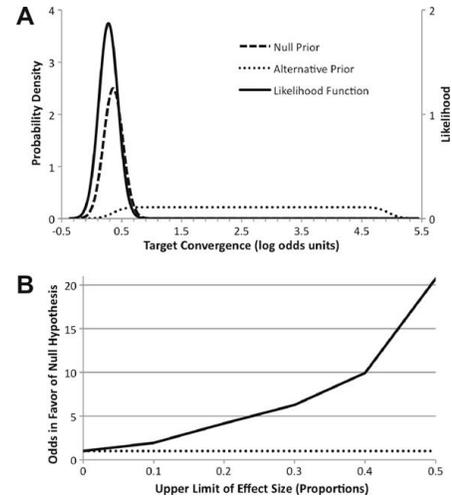


Figure 4: Bayesian modeling for accepting the null hypothesis that Alla and Early-summa do not differ in the 100ms after hearing the determiner.

5 Questions and responses

(17) Did subjects develop a nonce association between Summa utterances and one of the two summa targets? (This would have been detected in HS09's experiment since *two of* also described those images.)

Response:

- Such associations would need to be developed during the experiment, but there was no effect of order on the core results (p. 50, right).
- Such associations would have led people astray during fillers involving *the*.
- Such associations would have needed to be sensitive to the gender of the person in the target picture.
- Summa items were associated with both two- and three-object displays, making encoding without pragmatic comparison even harder.

(18) The Nunna conditions used *items*, which is much more taxonomically abstract than the other objects. Could this have confused subjects?

Response: not impossible, but this won't explain the response patterns (p. 50, right).

References

- Breheny, Richard; Napoleon Katsos; and John Williams. 2006. Are generalised scalar implicatures generated by default? an on-line investigation into the role of context in generating pragmatic inferences. *Cognition* 100:434–463.
- Degen, Judith; Patricia A. Reeder; Katie Carbary; and Michael K. Tanenhaus. 2009. Using a novel experimental paradigm to investigate the processing of scalar implicatures. Slides from Experimental Pragmatics 2009.
- Grodner, Daniel J.; Natalie M. Klein; Kathleen M. Carbary; and Michael K. Tanenhaus. 2010. “some,” and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition* 116(1):42–55.
- Huang, Ti Ting and Jesse Snedeker. 2009. Online interpretation of scalar quantifiers: Insight into the semantics–pragmatics interface. *Cognitive Psychology* 58(3):376–415.