Priming pragmatic reasoning in the verification and evaluation of comparisons

Vishakha Shukla, Madeleine Long, Vrinda Bhatia & Paula Rubio-Fernandez (University of Oslo)
paula.rubio-fernandez@ifikk.uio.no

Most studies on scalar implicature focus on the lexical scale 'some' vs 'all', which tends to elicit high rates of pragmatic responses [1-4]. Here we examined an understudied scale formed by two syntactic constructions: categorizations and comparisons (e.g., 'A robin is a bird' vs 'A robin is like a bird'). Unlike 'some' statements, superordinate comparisons have been found to elicit high rates of logical responses [5], even though they are under-informative when interpreted pragmatically (SI: *A robin is not a bird*). Following recent work on enrichment priming [6-9], we predicted that 'some' and 'all' statements would introduce an *informativity bias* in sentence verification and evaluation, increasing pragmatic responses to under-informative comparisons.

EXP 1 aimed to replicate previous findings by testing whether under-informative comparisons would elicit high rates of logical (vs pragmatic) responses in a sample of 22 UCL students. Replicating prior work [5], high rates of True responses (83%) were observed, in stark contrast to the high rates of True responses previously reported for 'some' and 'all' [1-4].

EXP 2 employed a rating task to test whether 'some' and 'all' statements are more readily perceived as scalemates and elicit scalar implicatures, than categorizations and comparisons. 68 adults from the UK were recruited via Prolific to rate statements on a scale (1=Very bad, 7=Very good). In line with previous work [5], we predicted higher ratings for stronger statements ('all' and categorizations) than weaker ones ('some' and comparisons). Critically, we also predicted comparisons would be rated higher than 'some' sentences (despite both being under-informative). An LMER model of Rating with Sentence Form (Weak vs Strong) and Group (Some & All vs Categorization & Comparison) as FE and maximum RE structure revealed a main effect of Sentence Form (p<.001), with lower ratings for weak forms, and a main effect of Group (p<.001) with higher ratings for categorizations and comparisons. The Sentence Form x Group interaction was also significant (p<.001), driven by a main effect of Group for weak forms (p<.001) (Fig. 1).

EXP 3 tested our main prediction that 'some' and 'all' would prime pragmatic reasoning. 156 adults from the UK were recruited via Prolific and were administered one of two online tasks: sentence verification or evaluation. In both tasks, participants read comparisons and categorizations alone, or randomized with 'some' and 'all' sentences, and had to judge whether the statements were True or False (verification) or Good or Bad (evaluation). An LMER model of Response (True/Good=1, False/Bad=0) with Sentence Type (Categorization, Comparison), Condition (Without Some/All, With Some/All) and Instruction (Verification, Evaluation) as FE and maximum RE structure revealed a marginal Sentence Type x Condition interaction (p=.056), driven by a difference in comparisons (p=.007), but not categorizations (p=.135) across conditions (Fig. 2). Specifically, the rate of True/Good responses was lower for comparisons With Some/All, supporting our hypothesis that participants engaged in pragmatic reasoning when processing 'some' and 'all' statements and as a result responded pragmatically to comparison statements (for full model output, see Table 1). Further support comes from an RT LMER analysis of True/Good responses using the same variables. We found a main effect of Condition (p=.013), with faster RTs in the Without Some/All condition than the With Some/All condition (Fig. 3, Table 2), likely because pragmatic reasoning slows down processing. Along these lines, the B&N effect [3,10] posits that participants will take longer to respond False than True for 'some' statements precisely because deriving scalar implicatures is cognitive costly. We tested this with our own data and replicated these findings with slower RTs for False/Bad than True/Good for 'some' items (p=.040) (Fig. 4). These results extended to comparison statements where RTs were slower for False/Bad than True/Good (p<.001), suggesting the inference *An X is not a Y* is also costly [10].

Our study is the first to show that 'some' and 'all' statements prime pragmatic reasoning in both sentence verification and evaluation tasks. This finding suggests that different scalar terms not only give rise to different rates of scalar implicatures [10-13], but can also affect the interpretation and processing of other types of scalar expressions.

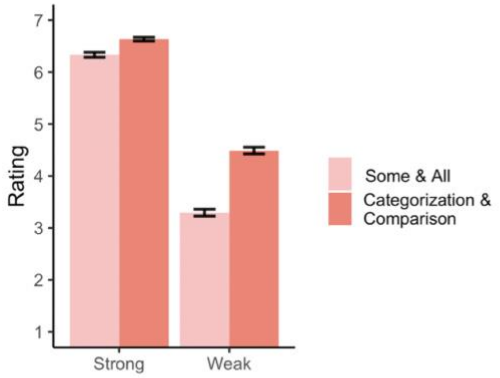

Fig 1. Average sentence ratings in Experiment 2, showing an interaction between Sentence Form and Group.

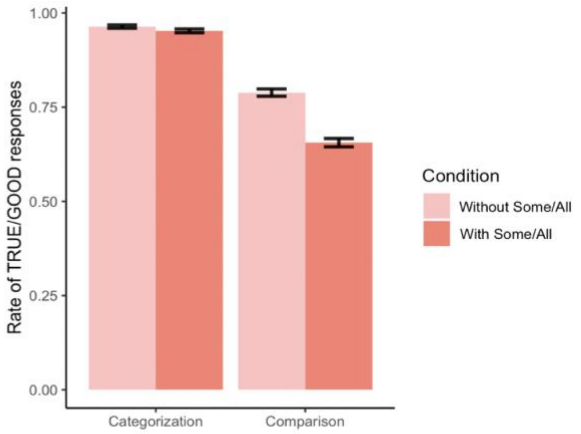

Fig. 2. Average rates of TRUE/GOOD (logical) responses to categorizations and comparisons in Experiment 3, showing the Sentence Type by Condition interaction.

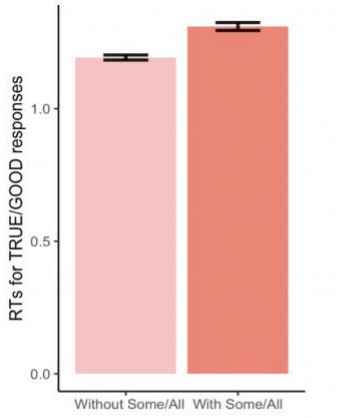

Fig. 3. Average RTs to categorizations and comparisons in Experiment 3, showing the main effect of Condition.

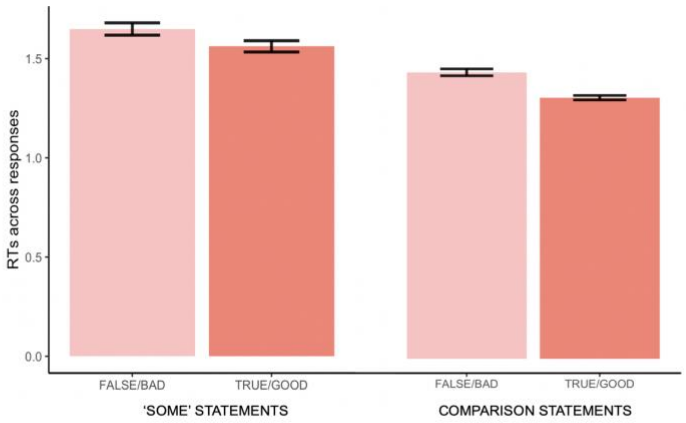

Fig. 4. Average RTs for FALSE/BAD (pragmatic) responses and TRUE/GOOD (logical) responses for 'some' items and comparison items in Experiment 3. Pragmatic responses to both types of under-informative statement were significantly slower than logical responses, confirming that the derivation of the corresponding scalar implicatures is cognitively costly.

Table 1. Full model output for *accuracy* rates.

Fixed effect	Coefficient	SE	*p* value
Sentence Type	-1.435	0.308	<.001
Condition	-0.922	0.316	0.004
Instruction	-1.330	0.338	<.001
Sentence Type x Condition	-1.108	0.580	0.056
Sentence Type x Instruction	2.034	0.566	<.001
Condition x Instruction	-0.078	0.620	0.900
Sentence Type x Condition x Instruction	1.057	1.143	0.355

Note: Significant and marginal main effects and interactions are shaded.

Table 2. Full model output for RTs to True/Good responses.

Fixed effect	Coefficient	SE	*p* value
Sentence Type	0.128	0.019	<.001
Condition	0.110	0.044	0.013
Instruction	-0.022	0.043	0.620
Sentence Type x Condition	-0.008	0.038	0.843
Sentence Type x Instruction	-0.042	0.041	0.313
Condition x Instruction	0.108	0.087	0.216
Sentence Type x Condition x Instruction	-0.013	0.079	0.870

Note: Significant main effects and interactions are shaded.

References

[1] Noveck, 2001. *Cognition*. **[2]** Noveck & Posada, 2003. *Brain and Language* **[3]** Bott & Noveck, 2004. *Journal of Memory and Language* **[4]** Feeney, Scrafton, Duckworth & Handley, 2004. *Canadian Journal of Experimental Psychology* **[5]** Rubio-Fernandez, Geurts & Cummins, 2017. *Review of Philosophy and Psychology* **[6]** Bott & Chemla, 2016. *Journal of Memory and Language* **[7]** de Carvalho, Reboul, der Henst, Cheylus & Nazir, 2016. *Frontiers in Psychology* **[8]** Rees & Bott, 2018. *Cognition* **[9]** Rees, Bott & Schumacher, 2019. *Neuroscience Letters* **[10]** van Tiel, Pankratz & Sun, 2019. *Journal of Memory and Language* **[11]** Doran, Baker, McNabb, Larson & Ward, 2009. *International Review of Pragmatics* **[12]** Doran, Ward, Larson, McNabb & Baker, 2012. *Language* **[13]** van Tiel, van Miltenburg, Zevakhina & Geurts, 2016. *Journal of Semantics*.