Speaking of quantifiers

Bob van Tiel, Michael Franke, Manfred Krifka, & Uli Sauerland

The study of words that express quantification has occupied philosophers, linguists, logicians, and psychologists since the work of Aristotle. In formal semantics, the consensus is that quantifying expressions denote generalised quantifiers, i.e., relations between sets. According to this idea, e.g., a quantified statement of the form 'every A is B' is true iff $A \subseteq B$. The theory of generalised quantifiers has been immensely succesful in explaining both empirical observations and experimental data. Nonetheless, psychological research has shown that it offers a poor approximation to the actual usage of quantifying expressions (e.g., Newstead et al., 1987). To illustrate, we conducted two experiments on the production and interpretation of quantifying expressions.

For Exp. 1, we drafted 250 participants (mean age: 33, 151 females) on Amazon's Mechanical Turk. Trials in Exp. 1 showed a display consisting of 432 circles randomly scattered across a 25×36 grid. The circles were either black or red, with the precise distribution of red and black circles varying across trials. Participants were instructed to describe these displays by completing a sentence of the form '— of the circles are red'. Participants were explicitly told not to use numeric descriptions, such as 52 or 251. Every participant was thus presented with 10 trials.

For Exp. 2, we drafted 50 participants (mean age: 32, 17 females) on Amazon's Mechanical Turk. Trials in Exp. 2 showed a statement of the form 'Q of the circles are red' along with one of three questions: (i) What is the lowest number of circles that might be red? (ii) What is the highest number of circles that might be red? (iii) What is your best guess about the number of circles that might be red? Q was varied between the 16 most frequently producted quantifying expressions in Exp. 1. To answer the questions, participants were presented with the same type of grid as used in Exp. 1, this time containing 432 black circles. Participants were instructed to add and remove red circles until they felt the display provided an answer to the relevant question.

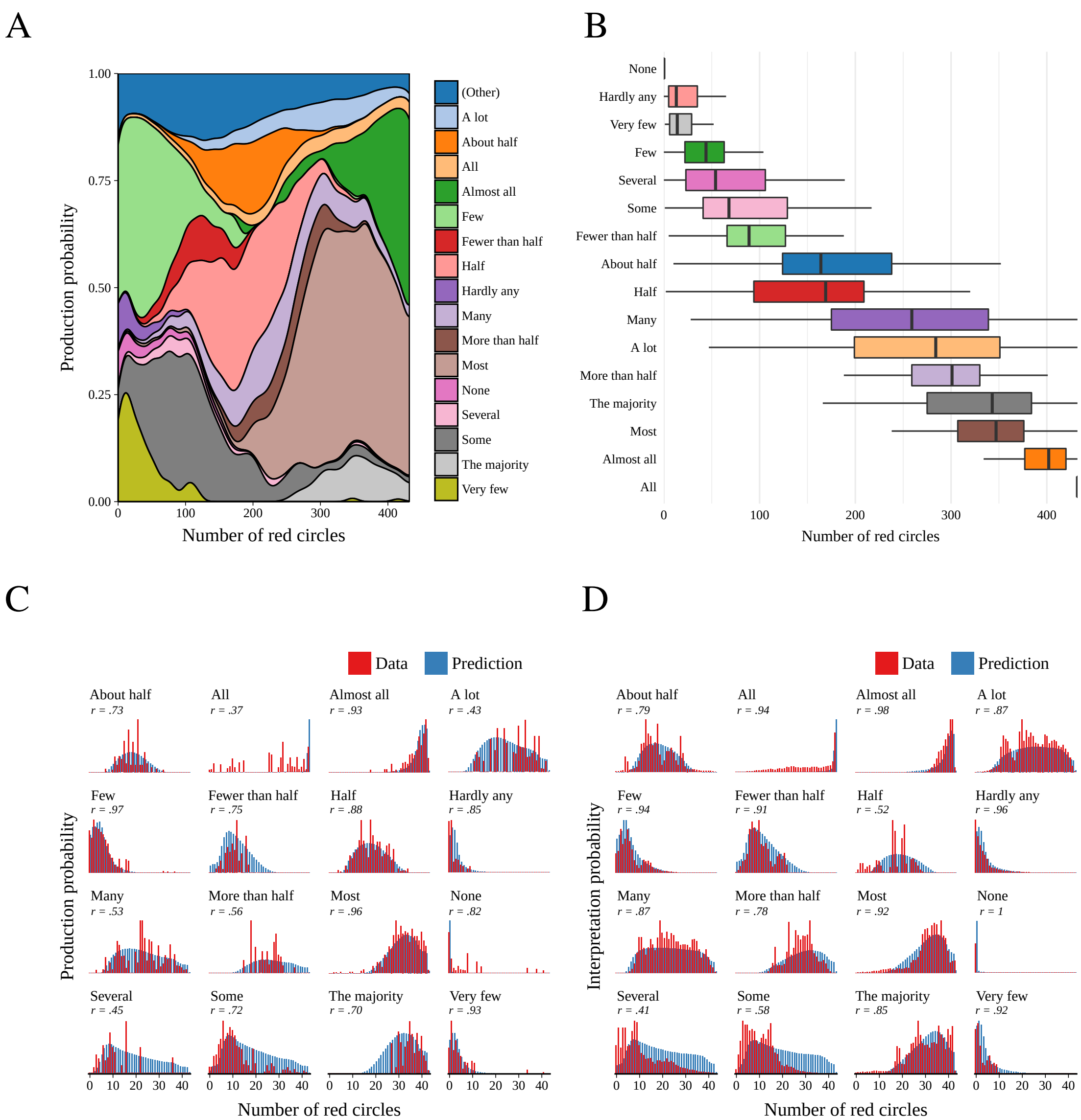
Fig. A shows production probabilities for the 16 most frequently produced quantifying expressions in Exp. 1. Fig. B shows a normalised measure of the results of Exp. 2. This measure assumes that every participant represents quantifying expressions as normal distributions with their answer to (iii) as mean and half of the average distance between this mean and their answers to (i) and (ii) as its standard deviation. For each value in the range between the answers to (i) and (ii), its probability was thus determined and these probabilities were normalised so that they summed to 1.

The results of both experiments indicate that quantifying expressions are associated with *focal ranges* that are variously wider or more narrow than the corresponding generalised quantifiers. To illustrate, whereas, according to its set-theoretic definition, 'some A are B' is true iff $|A \cap B| \neq \emptyset$, participants in both experiments adopted a much more restrictive meaning. Conversely, 'half of the A are B', according to its set-theoretic definition, is true iff $|A \cap B| = \frac{1}{2}|A|$, but participants in this case appealed to a much more lenient meaning.

We consider two computational models to account for the presence and location of these focal ranges. According to the *typicality-based* model, quantifying expressions exhibit typicality structure (e.g., Novák, 2007). This approach holds that a statement such as 'some A are B' is associated with a prototype in which it is maximally true, e.g., a situation in which 33% of the A are B. The truth value of the statement is a decreasing function of the distance from this prototype.

By contrast, the *rational speech acts* (RSA) model retains the set-theoretic definitions from the theory of generalised quantifiers, but assumes that speakers and hearers engage in pragmatic reasoning about their interlocutor (Frank & Goodman, 2012). In particular, the speaker is assumed to be as efficient and informative as possible given the alternative messages she could have sent.

We implemented both models and calculated the fit with the actual data. For both the production and interpretation data, the RSA model was superior, reaching *r*'s of .91 (Exp. 1) and .87 (Exp. 2). Figs. C and D plot the results of both experiments against the predictions of the RSA model. We thus demonstrate that these data, which appear to undermine the theory of generalised quantifiers, can be accounted for in terms of pragmatic principles of efficiency and informativeness. Moreover, we show that the RSA model, which has hitherto been used in relatively well-controlled settings (e.g., the competition between 'some' and 'all') can be scaled up to provide a precise and tractable account of much more open-ended aspects of language use.



References: Frank, M. & Goodman, N. (2012). In: *Science*, *336* (6084), 998. Newstead, S.E., Pollard, P., & Riezebos, D. (1987). In: *Appl Ergon*, *18* (3), 178–182. Novák, V. (2008). In: *Fuzzy Sets Syst*, *159* (10), 1229–1246.