Speaking of quantifiers

Bob van Tiel

Joint work with:

Michael Franke
Philippine Geelhand
Uli Sauerland

Procedural and computational models of semantic and pragmatic processes

Quantity and probability

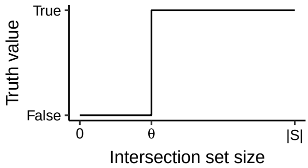
The expression of quantity and probability plays an important role in language and thought. (Kent, 1964; Kratzer, 1991; Shinagare *et al.*, 2019)

What do these expressions mean?

Study 1: Quantity

van Tiel, Franke, & Sauerland (2021). *PNAS*, *118*, e2005453118.

Truth



It has been argued that the meaning of a declarative sentence resides in its truth conditions. Accordingly, knowing what a sentence means entails knowing when it is true or false. (e.g., Frege, 1892; Montague, 1968)

Quantity and probability expressions denote thresholds on their scales: generalised quantifier theory (GQT). (e.g., Gärdenfors, 1987)

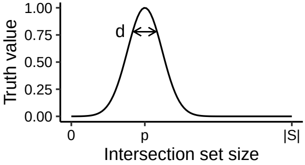
Use

There are often discrepancies between meaning and use:

(1) Several students passed the exam.
$\leadsto$ Not all students passed the exam.

Such discrepancies are to be explained by a theory of pragmatics. One of the central insights of modern pragmatic theory is that communication is a (boundedly) rational activity. (Frank & Goodman, 2012; Grice, 1975; Horn, 1972)

Prototypes

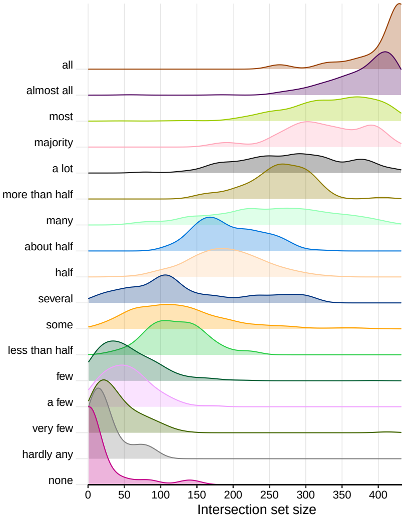


An alternative view holds that meaning should be construed in terms of prototypes that satisfy the most characteristic properties.

> "Cognitive grammar considers discrete categorization to be cognitively unrealistic, and emphasizes instead a prototype or 'central tendencies' model." (Langacker, 1987, p. 5)

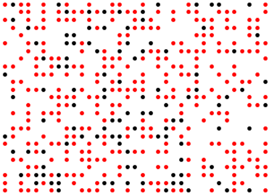
Quantity and probability expressions denote gradient and focalised ranges on their scales. (e.g., Zadeh, 1983)

Production



	Mean	5 – 95% Percentile
all	406	325 – 432
almost all	379	293 – 427
most	335	229 – 417
majority	324	214 – 404
a lot	283	169 – 392
more than half	262	182 – 319
many	254	114 – 387
about half	194	134 – 271
half	191	103 – 279
several	134	28 – 288
some	125	34 – 259
less than half	121	61 – 186
few	56	8 – 133
a few	53	6 – 112
very few	41	5 – 98
hardly any	25	5 – 75
none	15	0 – 83

— of the circles are red.



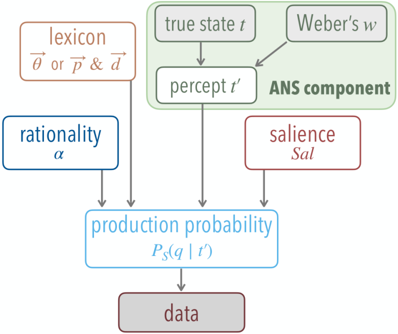
Research question

Can we explain the way people express quantity using a threshold-based semantics? (Or should we encode patterns of gradience and focality at the level of semantics?)

Computational model

The model has the following ingredients:

- Lexicon: the linguistic meaning of QEs.
- Rationality: the rationality of the speaker.
- ANS: the perceived state of the world.
- Salience: the salience of the QE.

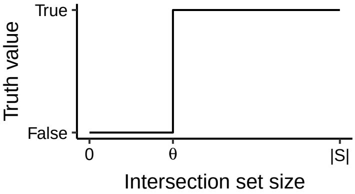


Two types of lexicon

Threshold-based

QEs place thresholds on the intersection set size, based on their monotonicity.

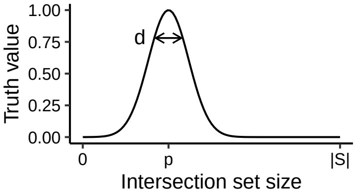
$$\mathfrak{L}_{TB}(m,t) = \begin{cases} 1 \text{ if } m \text{ is mon-}\uparrow \text{ \& } t \geq \theta_m \\ 1 \text{ if } m \text{ is mon-}\downarrow \text{ \& } t \leq \theta_m \\ 0 \text{ otherwise} \end{cases}$$



Prototype-based

QEs denote fuzzy and focalised ranges on the space of intersection set sizes. (e.g., Bocklisch et al., 2012)

$$\mathfrak{L}_{PT}(m,t) = exp\left(-\left(\frac{t-p_m}{d_m}\right)^2\right)$$



Two types of speaker

Literal speaker

Produces messages with probabilities proportional to their truth value.

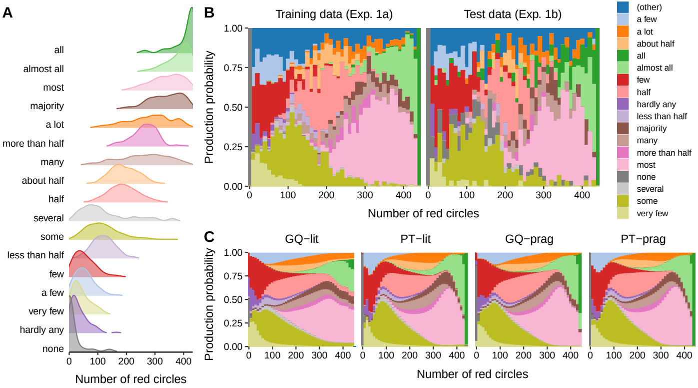
$$P_{S_{\text{lit}}}(m \mid t, \mathfrak{L}) \propto \mathfrak{L}(m, t)$$

Pragmatic speaker

Selects messages that optimise the probability of coordination.

$$P_{S_{\text{prag}}}(m \mid t, \mathfrak{L}) \propto P_{L_{\text{lit}}}(t \mid m, \mathfrak{L})^{\lambda}, \text{where } P_{L_{\text{lit}}}(t \mid m, \mathfrak{L}) \propto \mathfrak{L}(t, m)$$

Model predictions



Free model parameters were estimated using STAN. Model comparison (training vs. test data) indicate that the pragmatic threshold-based model offered the best explanation of the test data.

Interim conclusions

- Quantity expressions have crisp meanings, but gradience may emerge as a consequence of competition between viable messages—as well as cognitive and perceptual biases.
- Tools and techniques from probabilistic modelling in the cognitive sciences can be used as link functions for the study of abstract theories of meaning from linguistics.

Study 2: Probability and autism

van Tiel, Sauerland, & Franke (2022). *Open Mind*, *6*, 250–263.

Autism Spectrum Disorder

ASD is a lifelong neurobiological disorder characterised in part by important pragmatic difficulties.

> "Difficulties understanding what is not explicitly stated (e.g., making inferences) and nonliteral or ambiguous meanings of language (e.g., idioms, humor, metaphors, multiple meanings that depend on the context for interpretation)." (DSM-V, American Psychiatric Association, 2013, p. 48)

These communicative difficulties are commonly connected to difficulties with mental state attribution. (e.g., Baron-Cohen, 1995; Happé, 1993)

Scalar inferences

Surprisingly, autistic people are equally likely to derive scalar inferences as neurotypicals. (e.g., Chevallier *et al.*, 2010; Pijnacker *et al.*, 2009; Su & Su, 2015)

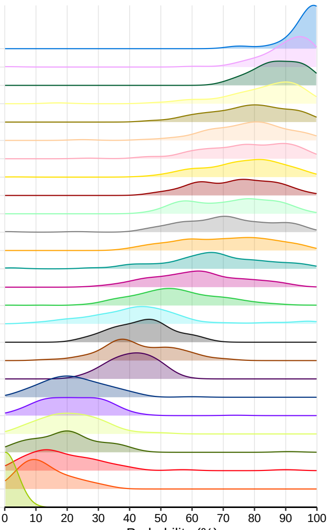
(2) a. Some dogs are mammals.
    b. Zebras have black or white stripes.

However, there is a considerable gap between the experimental task and actual language use.

Research goal

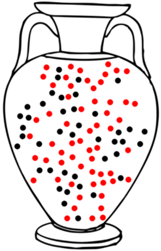
Can we leverage recent probabilistic pragmatic models to quantitatively estimate effects of autism on language use?

Production



	Mean	5 – 95% Percentile
certain	96	86 – 100
almost certain	90	77 – 99
extremely likely	87	75 – 97
very high chance	81	58 – 95
highly likely	78	57 – 95
great chance	76	54 – 96
most likely	76	53 – 95
very likely	76	55 – 94
very good chance	73	53 – 92
high probability	72	53 – 91
high chance	70	47 – 93
more than likely	70	46 – 93
probable	66	38 – 94
likely	61	39 – 87
good chance	54	33 – 79
possible	44	18 – 89
even chance	44	28 – 60
decent chance	43	25 – 61
equal chance	40	29 – 51
low chance	22	7 – 39
unlikely	21	6 – 36
small chance	21	6 – 36
not likely	20	3 – 37
not very likely	20	4 – 39
very unlikely	13	4 – 28
impossible	0	0 – 1

If you randomly take a marble from the urn — that it is red.



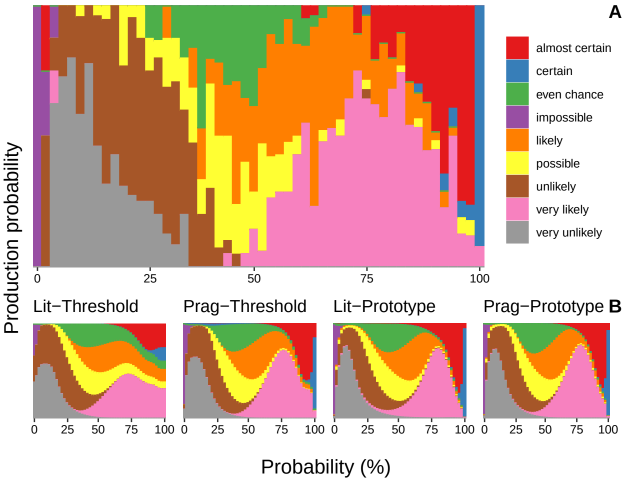
Autism quotient

We measured participants' autism quotient as a measure of the degree to which they exhibited autistic traits. (Baron-Cohen et al., 2001)

(3) a. I find social situations easy.
    b. I find it difficult to work out people's intentions.
    c. It does not upset me if my daily routine is disturbed.

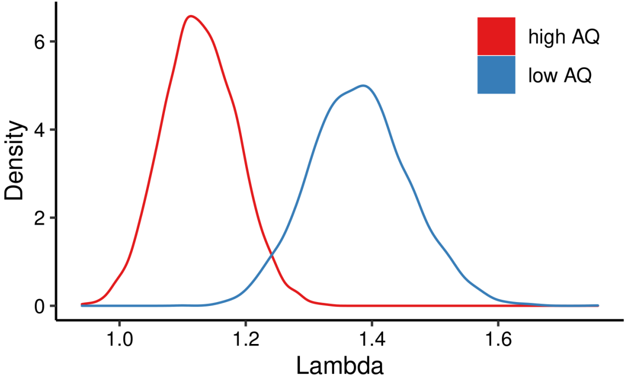
Participants were categorised based on a median split (M = 20.5).

Model predictions



Free model parameters were estimated using STAN. Model comparison (training vs. test data) indicate that both pragmatic models explain the data equally well (r = .89).

Lambda



We fit the model for all participants, but allowed the λ parameter to vary depending on one's AQ. High-AQ participants were estimated to have a significantly lower λ parameter than low-AQ participants.

Conclusions

- The results support a modular view on linguistic meaning:
  - a semantic module that calculates the truth-conditional meaning of an utterance.
  - a pragmatic module that reasons about the probability that the utterance receives the intended interpretation.
- Pragmatic differences between high and low AQ participants emerge in naturalistic settings, thus explaining puzzling findings on scalar inferences. However, the differences are surprisingly subtle.

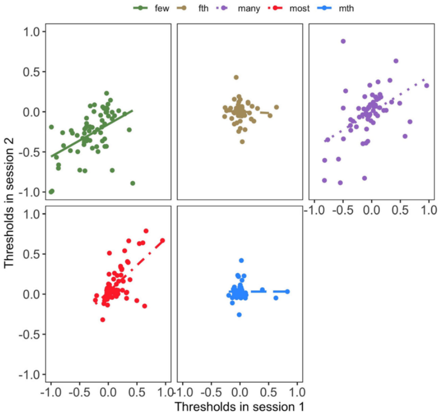
Argumentativity

Participants' choice of quantity expression is partly guided by their communicative goals. (e.g., Anscombre & Ducrot, 1983)

(2) $$\left\{\begin{array}{l}\text{Many people}\\ \text{Not everyone}\end{array}\right\}\text{ enjoyed the workshop.}$$

In our experiment, the overall goal was coordination about the proportion of red circles. It will be interesting to see how to formalise behavioural differences under different goals.

Variability



Our model assumes that each quantity expression is associated with a single threshold; but, Ramotowska *et al.* (2023) observe cross-participant variability.

Sources of vagueness

In our GQT model, the approximate perception of numerosity is the only source of vagueness. However, there may be many other sources.

- Disagreement about the location of the threshold (or about the goal of the conversation or about the appropriate level of granularity or...).
- Uncertainty about the location of the threshold.
- Beliefs about disagreement about the location of the threshold.

In our data, these sources are difficult to disentangle. The results of Ramotowska *et al.* suggest different sources for different quantity words.

Semantics

What is the study of semantics about? (Partee, 1999)

1. Intuitions about the meanings of expressions. (Lakoff, 1987)
2. Generalisations over speaker meanings. (Grice, 1975)
3. Abstractions that underlie compositionality. (Montague, 1973)

Our approach is in line with the third idea. We cannot observe meanings, but we can leverage computational models to make inferences about the meanings that underlie language use.

Semantics

What is the study of semantics about? (Partee, 1999)

1. Intuitions about the meanings of expressions. (Lakoff, 1987)
2. Generalisations over speaker meanings. (Grice, 1975)
3. Abstractions that underlie compositionality. (Montague, 1973)

Our approach is in line with the third idea. We cannot observe meanings, but we can leverage computational models to make inferences about the meanings that underlie language use.

Thank you for your attention!