Quantifying Vagueness

Sverker Sikström[1], Rasmus Bååth[1] & Uli Sauerland[2]

1. Lund University, 2. ZAS

Asking subjects to Quantify Vagueness

Survey
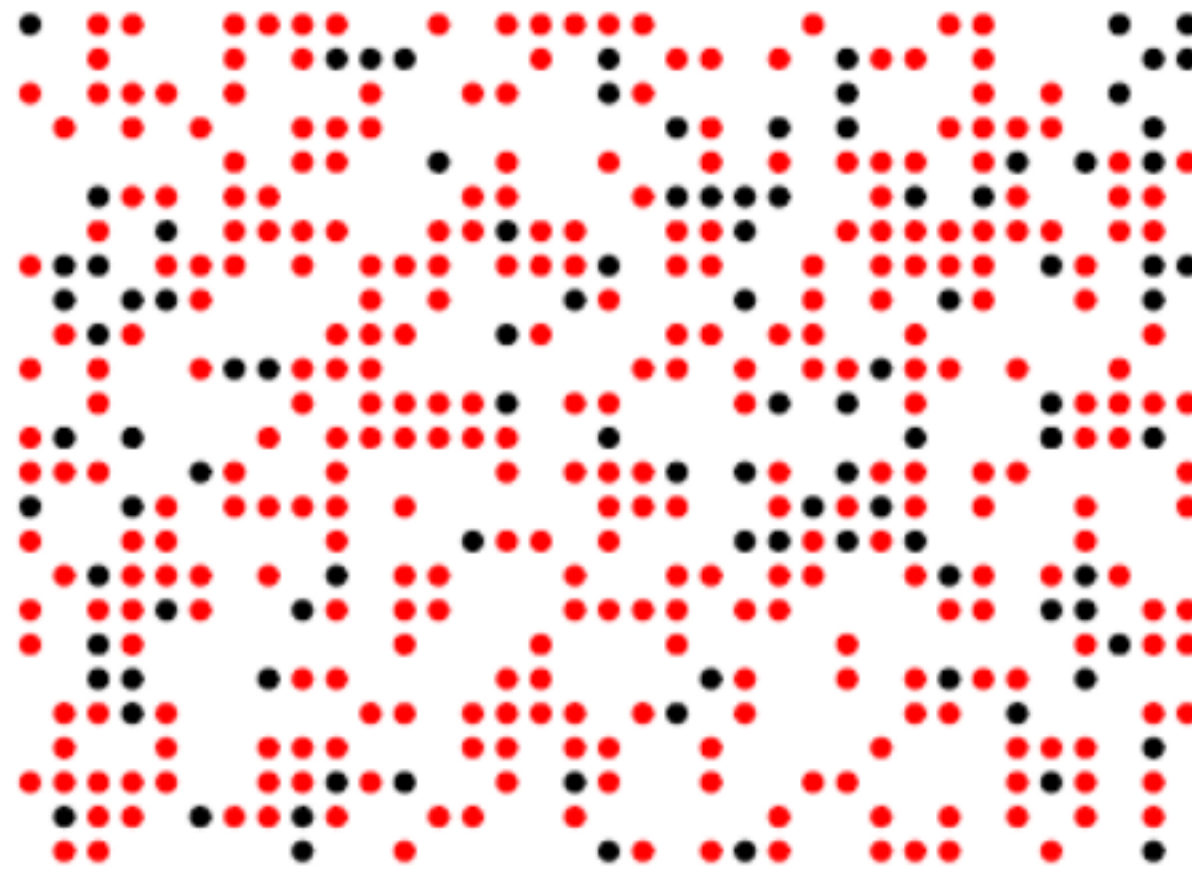
How many of the dots are red?
[   ] are red.

===================
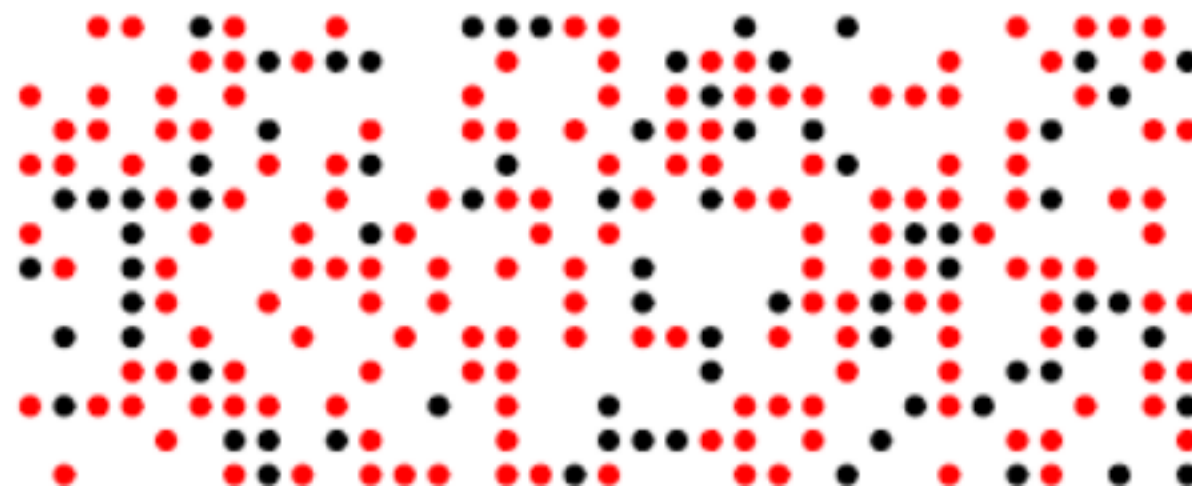
Subjects

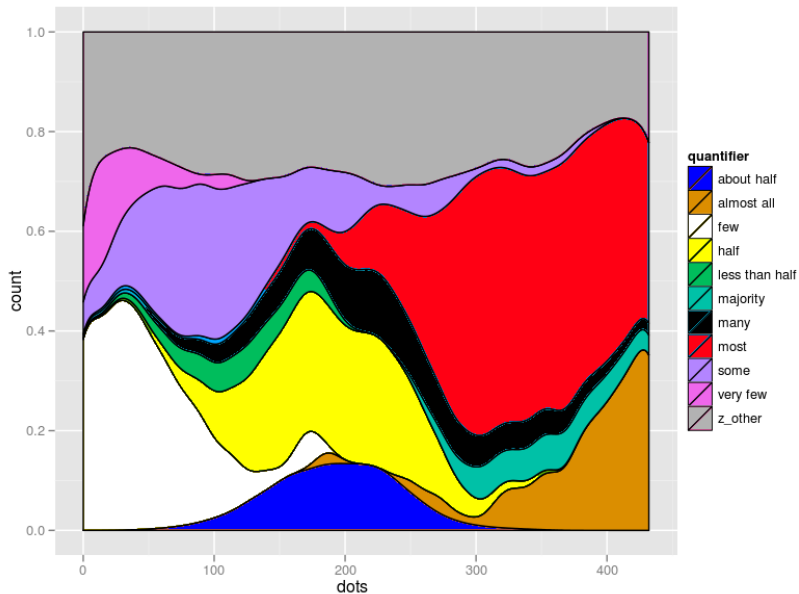|	English	German
n	250	54
age	33.42 (sd=11.8)	32.13 (sd=8.1)
females	62%	28%
pay	$0.20	$0.40

Top English quantifiers

1. most
2. few
3. some
4. almost all
5. half
6. many
7. very few
8. majority
9. about half
10. less than half

Top English quantifiers

Comparison with German: Top ten quantifiers

1. most	1. fast all
2. few	2. meist
3. some	3. einig
4. almost all	4. viel
5. half	5. wenig
6. many	6. hälft
7. very few	7. paar
8. majority	8. sehr viel
9. about half	9. all
10. less than half	10. etwa hälft

Comparison of Literal Translations

quantifiers	p(t-test)	E-mean	G-mean
some vs. einig	0.052	101	81
some vs. paar	0.16	101	72
many vs. viel	0.10	254	290
few vs. wenig	0.79	48	47
almost all vs. fast all	0.60	390	386
most vs. meist	0.47	353	348
majority vs. meist	0.12	361	348
half vs. hälft	0.10	175	200

Quantifying Vagueness Using Signal Detection Theory
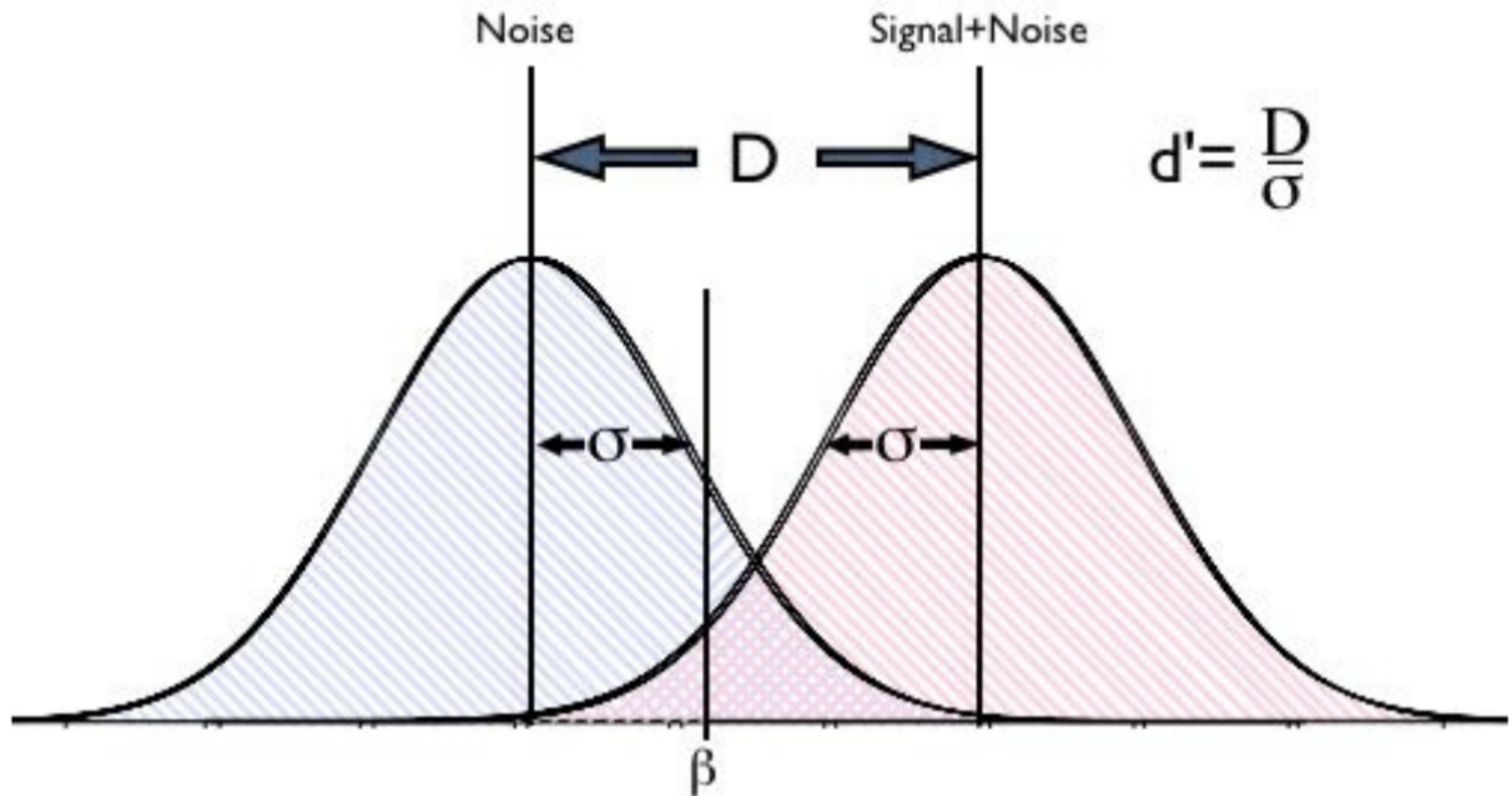
Naive vagueness judgement

We are interested how precise or vague some words used to quantify a number of items feels to you. If you for example were asked: "How many sheep are black?", "five" would be an precise answer and "a couple of them" would be vague.

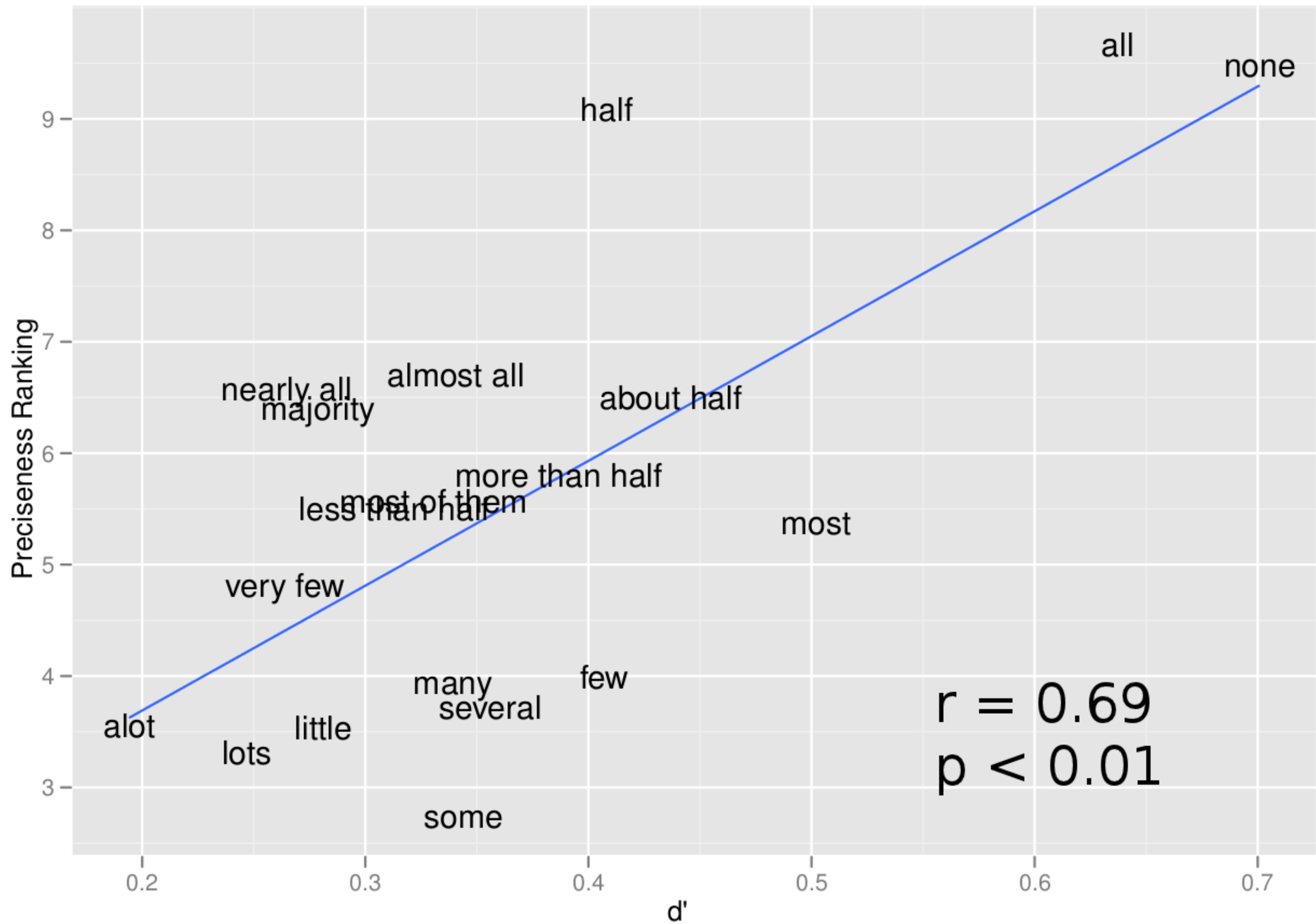Please rate these 20 words on a scale from 1 (very vague) to 10 (very precise).

most

Vague/Imprecise   Precise/Non-Vague

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

about half

Vague/Imprecise   Precise/Non-Vague

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

almost all

Vague/Imprecise   Precise/Non-Vague
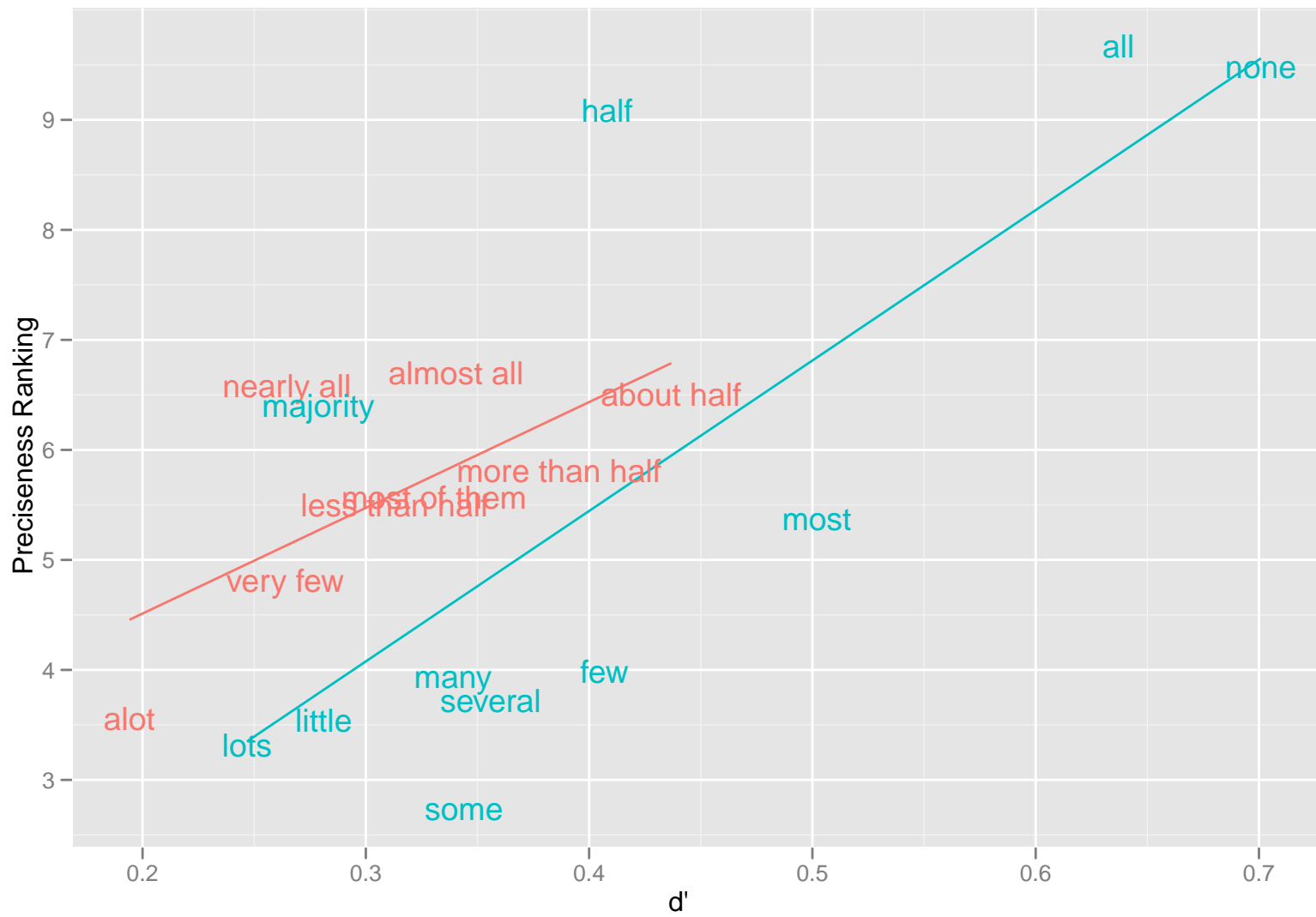
○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

majority

Vague/Imprecise   Precise/Non-Vague

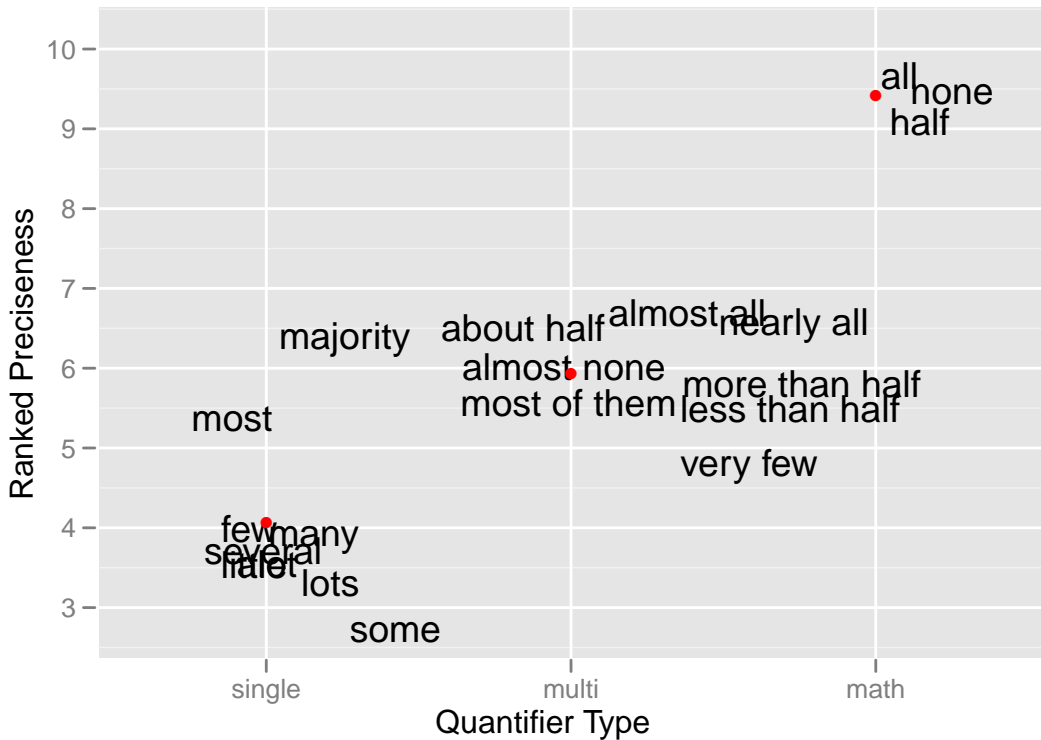○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

few

Vague/Imprecise   Precise/Non-Vague

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

Factors Predicting Vagueness Ranking

- Mathematically well defined quantifiers
  - Well defined (all)
  - Non-well defined (some)
- Number of words in quantifier
  - Single quantifier (most)
  - Multiple word quantifier (more than half)

Quantifying Vagueness using Sematic Spaces

- A computational method to quantify semantics
- Basic assumption: Meaning is co-occurrence
- Words are represented as points in space. The meaning of a word is given by it's distance to other words
- Current applications:
    - Semantic associations
    - Estimating text difficulties
    - Judging quality of essays

Result (translated)

Top vague words

roughly
maybe
rather
about
soon
slightly
pretty
fuzzy
small
large

Top precise words

exactly
absolutely
precisely
now
left
right
two
correct
no
yes

● Leave-one-out cross-validation gives
a correlation of 0.78, p < 0.001

Applications of the vague norm

- Eye-witness testimonies
- Child language
- Specific language impairment (SLI)

Eye-witness testimony

Verbal reports of criminal scene

Ok someone was filming and it was on a bus stop where people were, or there was none to begin with, but then there was someone who went to the mailbox and sent a letter, and then there were cars going by and than a girl came that sat down to wait for the bus and was looking to see when her bus would arrive. And then there was another girl who came by and the girl still sitting asked her what time it was. So that she knew when the buss would arrive. So she were sitting and waiting, and then she rose to get to the bus. Then a car approached and out of it comes two guys, men, out of the car. They took her and well, took her into the car. And then they walked away.

Vagueness and Correctness in Eyewitness Statements

- Theory: Signal detection theory interpretation of vagueness

- Prediction: Vague statements are less correct

- Result: Criminal scene
  - Experiment 1, $p < 0.05$
  - Experiment 2, $p < 0.001$

Vagueness in Child Language

Oral narratives from 108 children aged 4-17
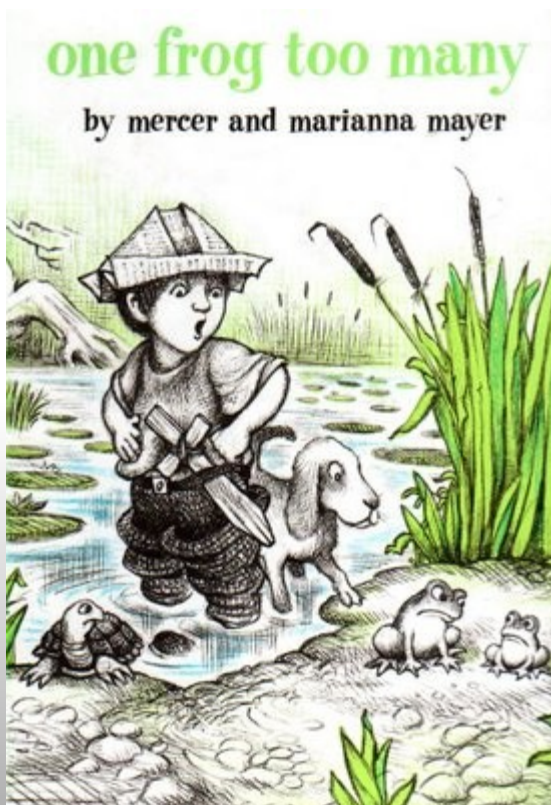
Linguistic Maturity and Vagueness

- Theory: Vagueness is a marker for linguistic elaboration
- Prediction: Vagueness increases with linguistic maturity
- Result: $p < 0.001$

Vagueness and Language Impairment

- Data: 103 oral narratives from children diagnosed with specific language impairment (SLI)
- Theory: SLI children are linguistically impaired
- Prediction: Children with SLI are less vague
- Result: $p < 0.01$

Linguistic Gender and Vagueness

- Theory: Women have a earlier/higher linguistic maturity

- Prediction: Vagueness in girls is higher than in boys of the same age.

- Result: $p < 0.01$

Conclusion quantifying vagueness

- Quantifier factors
    - Not well defined mathematically
    - Single word quantifier.
    - Empirical distributions.
- Subject factors
    - Age
    - Gender
    - Language impairment