Effects of Domain Complexity on Verification Procedures for *Most* and *More Than Half*

Hadas Kotek, Yasutada Sudo, Edwin Howard, Micha Y. Breakstone and Martin Hackl
MIT

Summary

[1] and [3] propose two competing views of how statements of the form *Most As are Bs* are verified. This study shows that the complexity of the complement set (A–B) affects the verification strategies for *most* and *more than half* in different ways. We argue that these results accord with the superlative semantics for *most* proposed by [1], but are not captured by the alternative put forward by [3].

Background

Most vs. *More than Half*

[1] proposes a superlative meaning for *most* and relates differences in verification procedures between *most* and *more than half* as revealed in Self-Paced Counting (SPC) experiments to differences in their semantics.

⟦ most ⟧(A)(B) = 1 iff ∃X [A(X) ∧ B(X) ∧ ∀Y∈C [Y≠X→|X|>|Y|]]
where C = A (closed under i-sum formation)

⟦ more than half ⟧(A)(B) = 1 iff |A∩B| > ½|A|

Prediction

[1]'s account predicts that the verification of *Most of the As are Bs* is sensitive to the complexity of A, because all non-overlapping subsets of A are compared with A∩B.

Alternative Account

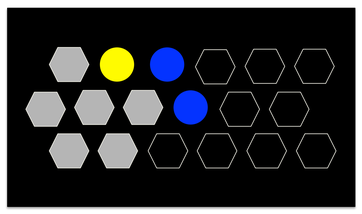
[3] conducted a flash presentation experiment where participants judged the truth of *Most of the dots are blue* against pictures of dots differing in the number of colors, and report no effect of the number of colors on the accuracy rate. Based on these results, [3] suggests a different representation for *most*:

⟦ most ⟧(A)(B) = 1 iff |A∩B| > |A|–|A∩B|

SPC Experiment

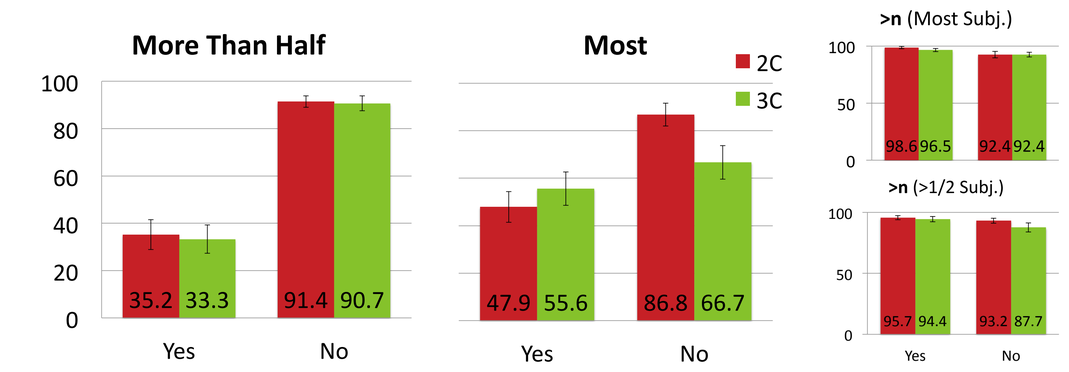
Procedure

The Self-Paced Counting (SPC) method of [1]: Participants answered questions such as *Are most of the dots blue?* relative to dot arrays of varying sizes (14-18). The dots are uncovered in groups of 2 or 3 as the spacebar is pressed. In our study, the number of colors varied between 2 and 3. A third color appears in the 4th frame and on.



- Each participant was randomly assigned *most* or *more than half*
- 24 target items and 96 filler items
- Half of the target items have the target-to-non-target ratios 9:8 or 8:7; the other half 8:9 or 7:8
- 4 ratios are used in 3C: 9:4+4, 8:3+4, 8:5+4, 7:4+4
- 51 participants (24 for *most*, 27 for *more than half;* all accuracy rate >75%)

Accuracy Rates



Observations:

(i) More correct answers for 'No' for both *most* and *more than half* (Contrast coded mixed logit model with random subject and item effects; p<0.001 for both determiners)

(ii) An interaction of the correct answer type and the number of colors for *most* (Contrast coded mixed logit model with random subject and item effects; p<0.001 for *most*, p=0.884 for *more than half*)

We propose:

(a) Observation (i) is due to **a 'significantly-more'** bias for both determiners

(b) Observation (ii) is due to **a superlative reading** of *most* in subject position

Superlative Reading

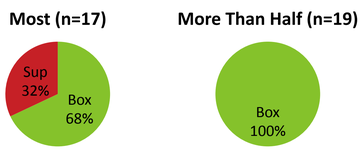
Most of the dots are blue has two readings:

Proportional reading: |blue dots| > |non-blue dots|

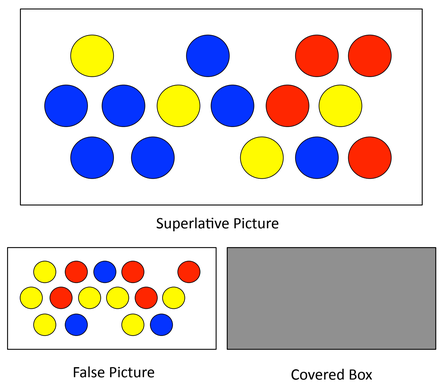
Superlative reading: For all non-blue colors C, |blue dots| > |C dots|

These readings disagree in 3 Color conditions such as 7:4+4 (Prop. → "No"; Sup. → "Yes")

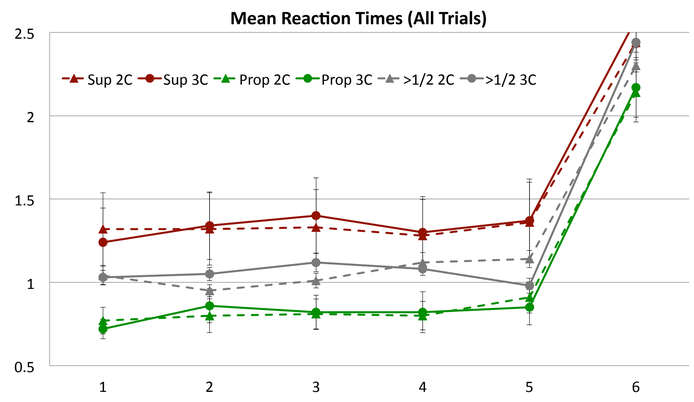
A covered box experiment (cf. [2]) on Amazon Mechanical Turk confirmed that the superlative reading exists for *most* in subject position, but not for *more than half.*



Most of the dots are blue was judged true 32% of the time for superlative pictures (9 subjects judged them true at least once)



Reaction Times Analysis for SPC Experiment



- Participants for *most* were divided into "superlative" and "proportional" speakers (n=12 for each; if accuracy of 2C-no < accuracy of 3C-no, then "superlative")
- We observe a significant difference in RTs (Contrast coded mixed effects model with random subject and item effects; p<0.001 for Frames 1-5)

Discussions

- The superlative reading of *most* is straightforwardly accounted for by [1]'s semantics:
  C=A ⇒ the proportional reading
  C={D_1, D_2, D_3} (C is a salient cover of A) ⇒ the superlative reading
- Why are superlative speakers slower than proportional speakers?
  The superlative reading is hard to verify in 3C; A more accurate, slower strategy is chosen (and generalized to 2C)

References:
[1] Hackl, M. (2009) On the grammar and processing of proportional quantifiers: *Most* versus *more than half. Natural Language Semantics.*
[2] Huang, Y., Spelke, E. & Snedeker, J. "What exactly do numbers mean?" Ms., Harvard University.
[3] Lidz, J., P. Pietroski, T. Hunter & J. Halberda (in press) Interface Transparency and the Psychosemantics of 'Most'. *Natural Language Semantics.*

Massachusetts Institute of Technology

Corresponding author: Hadas Kotek (hkotek@mit.edu)
Support through NSF Grant 0642478: "Experimental Investigations of Verification Procedures for Quantified Statements"