On the expression of proportion: 
*Most* and *more than half*

Stephanie Solt

Zentrum für Allgemeine Sprachwissenschaft, Berlin
solt@zas.gwz-berlin.de

84th Annual Meeting of the Linguistic Society of America
January 7, 2010

Semantics of quantity/proportion

- Generalized quantifier theory (Barwise & Cooper 1981): Unified analysis of quantificational expressions as ‘quantifying determiners’ that express relationships between two sets

- More recently: Unified analysis fails to capture differences between superficially synonymous expressions:
  - Geurts & Nouwen (2007): *more than 3* vs. *at least 4*
  - Fiengo (2007): *each* vs. *every*

Most vs. *more than half*

- A common intuition: *most* > *more than half*

(4) Unfortunately, the long term maintenance of the reduced weight is poor, and *more than half*, if not *most*, of the persons eventually return to their former obese state

(5) ??...and *most*, if not *more than half*, of the persons eventually return to their former obese state

(6) a. *More than half* of the U.S. population is female
b. *Most* of the U.S. population is female

The facts: female 50.7% vs. male 49.3%

(U.S. Census Bureau 2008)

Distribution of *Most*/*More than Half*

In corpus data, *most* is used for proportions considerably greater than half, while *more than half* is used for proportions close to 50%:

(7) a. The survey showed that *most students* (81.5%) do not use websites for math-related assignments (Education, 129(1), pp. 56-79, 2008)
b. *More than half of respondents* (55%) say that making money is more important now than it was five years ago

(Money, 21(3), p. 72, 1992)
**Distribution of Most / More than Half**

Most is used for proportions considerably greater than half, while more than half is used for proportions close to 50%:

Most is readily followed by a plural noun, yielding a generic-like interpretation; more than half is awkward in similar contexts, and (when acceptable) loses the generic flavor:

(8) a. Most people follow the moral judgments of those around them (Wise, 121(7), pp. 10-11, 2008)
    b. More than half of people follow the moral judgments of those around them

(9) a. Most teens want to fit in with their peers (CNN Health, 11/8/2002)
    b. More than half of teens want to fit in with their peers

More than half — but not most — requires a domain that can be individuated and counted (or otherwise measured)

(13) a. But like most things, obesity is not spread equally across social classes (Wise, 121(7), p. 104, 2008)
    b. But like more than half of things, obesity is not spread equally across social classes

More than half requires the predicate to be precisely defined; most allows a vague predicate

(14) a. ...most of the tourists in the early days were wealthy
    (Nature, 11/7, p. 14)
    b. More than half of the tourists in the early days were wealthy

Use of more than half typically co-occurs with mention of a source of supporting data; this is not the case with most

<table>
<thead>
<tr>
<th>Source of Data Mentioned</th>
<th>More than Half</th>
<th>Most</th>
</tr>
</thead>
<tbody>
<tr>
<td>Americans</td>
<td>9 / 12</td>
<td>13 / 100</td>
</tr>
<tr>
<td>Men</td>
<td>4 / 6</td>
<td>5 / 100</td>
</tr>
<tr>
<td>Women</td>
<td>4 / 5</td>
<td>7 / 100</td>
</tr>
<tr>
<td>Students</td>
<td>5 / 5</td>
<td>36 / 100</td>
</tr>
<tr>
<td>Patients</td>
<td>5 / 5</td>
<td>39 / 100</td>
</tr>
<tr>
<td>Families</td>
<td>1 / 2</td>
<td>11 / 100</td>
</tr>
<tr>
<td>TOTAL</td>
<td>29 / 35</td>
<td>111 / 600</td>
</tr>
</tbody>
</table>

Summary of Corpus Data

- Most and more than half are used to convey distinct ranges of proportions
- Most yields a generic interpretation in contexts where more than half is infelicitous or has a ‘survey results’ interpretation
- More than half (but not most) requires an enumerable domain and a precisely defined predicate
- Use of more than half (but not most) is typically supported by numerical data (count/survey/analysis)
Proposal

The distributional and interpretative differences between most and more than half can be attributed to a fundamental distinction in how proportion is expressed.

Logical form and verification

Hackl (2009): most / more than half require distinct logical forms that are truth-conditionally equivalent but give rise to different verification strategies.

More than half = \frac{\text{many} + \text{–} + \text{than} + \text{half}}{2} \quad (15)

Verification: Count the As that are B, and compare that to half the total number of As.

Most = \text{many} + \text{–est} \quad (16)

Verification: Compare the number of As that are B to the number of As that aren’t B (‘vote counting’)

Support from self-paced counting experiment

More than half – upper bound

* More than half competes with other numerical expressions of proportion.
  * Alternatives can be ordered on scales of varying levels of granularity (Krifka 2009).
  * The choice of more than half implies higher scalar options do not apply; more than half is restricted (by implicature) to values close to half (Horn 1972; Grice 1975).

Verifying:

\begin{align*}
\text{More than half} & = 1 \iff |A \cap B| > 0.5 |A| \\
\text{Most} & = 1 \iff |A \cap B| > |A| - |B|
\end{align*}

Most

\[ \text{[most A are B]} \equiv 1 \iff A \cap B \text{ larger than } A - B \]

* Logical form involves the comparison of two sets
  * Set sizes can – but need not be – compared by exhaustive counting of set members

... A child can tell which of two plates has more cookies before he can count them (with both readings of before!). Primitive man undoubtedly knows whether an inimical army was larger, smaller, or (roughly) equal to his own ... We may count Peter’s class and Paul’s class and then infer [Peter’s class is larger than Paul’s], and often this is the only way of doing it; but we may also just look at the two classes, and will often ‘see’ which is larger” (Bartsch & Vennemann 1973: 67-68)

Most and comparison

Options for set size comparison:

* Approximately or visually
  \begin{align*}
  \text{Most} & \quad \text{of the students in this class are female} \\
  \text{(17)} & \quad \text{‘Looks like’ more females than males}
  \end{align*}

* By induction or generalization
  \begin{align*}
  \text{Most teens} & \quad \text{want to fit in with their peers} \\
  \text{(18)} & \quad \text{In my experience, teens who want to fit in more common} \\
  & \quad \text{than those that don’t}
  \end{align*}

* Lack of exceptions
  \begin{align*}
  \text{Most} & \quad \text{of the tourists in the early days were wealthy} \\
  \text{(19)} & \quad \text{Few if any exceptions known} \\
  \text{Exhaustive counting of set members not required} \quad \text{Approximate (procedures fail for sets nearly equal in size)}
  \end{align*}
Number Cognition

- Findings from the psychology of number cognition indicate that in addition to a system for the representation of precise numerosity, humans possess a separate system for representing and manipulating approximate number
  - Non-verbal; analog (mental number line)
  - Present in pre-verbal infants, animals, patients with aphasia/acalculia, and societies w/out complex number systems
  - Involved in:
    - Quantity comparison
    - Approximate arithmetic (addition/subtraction – but not multiplication)
    - Ratio-dependent (size and distance effects)

(Dehaene 1997; Dehaene, Dehaene-Lambertz, & Cohen, 1998; Gallistel & Gelman 2000; Pica et al. 2004; Feigenson, Dehaene & Selke 2004)

<table>
<thead>
<tr>
<th>More than half</th>
<th>Most</th>
</tr>
</thead>
<tbody>
<tr>
<td>(</td>
<td>A \cap B</td>
</tr>
<tr>
<td>Requires countable sets</td>
<td>Requires uncountable sets</td>
</tr>
<tr>
<td>Verification via mathematical relation (&gt;) on numbers</td>
<td>Allows verification via approximate algorithms</td>
</tr>
<tr>
<td>Acceptable for proportions close to 50%</td>
<td>Dispreferred for proportions close to 50%</td>
</tr>
<tr>
<td>Competes w/expressions of proportion (e.g. 2/3)</td>
<td>Competes with all as higher alternative</td>
</tr>
<tr>
<td>Survey results interpretation</td>
<td>Generic interpretation</td>
</tr>
</tbody>
</table>

Most and approximate number

- Pietroski et al. (2009): Semantics of *most* based on cardinality comparison, which may be verified via approximate number system
  - Verification displays characteristic size and distance effects
- Further supporting evidence:
  - Children’s acquisition of *most* independent of mastery of counting (Halberda, Taing & Lude 2008)
  - Pirahà – which lacks words for numbers >2 – has an equivalent of *most* (U. Sauerland, p.c.)
- Empirical data explained:
  - Use of *most* where counting precluded
  - Use of *most* when ratio of A \( \cap \) B to A – B is ‘large enough’ to be detected via approximate verification strategy

Most – upper bound

- *Most* does not (directly) encode a proportion
  - Does not invoke a scale of proportions, or compete with items on these scales
  - Alternatives instead are items that express relationships between sets
- (20) a. *Some/many/most/all* teens want to fit in with their peers
  b. *Two thirds/seven-in-ten/70%* of teens want to fit in with their peers
- (21) Some…many….most….all
  - *All* as only higher alternative \( \Rightarrow \) *most* may be used for proportions approaching all

Conclusions

- *Most* and more than half represent two distinct modes of expressing proportion
  - More than half: comparison of numbers
  - *Most*: comparison of sets
- Distinct logical forms allow different possible verification strategies
  - More than half: (precise) mathematical relation
  - *Most*: approximate algorithms/strategies
- Speakers’ choice of lexical items takes into count how the sentence could (in principle) be verified
  - Explains wide range of distributional and interpretive differences between *most/more than half*

...and a possible parallel

Positive vs. comparative forms of gradable adjectives (cf. Kennedy 2007):

(22) a. More than half of the students are female \# female > # male
  b. Most of the students are female \# female >> # male
(23) a. Billy is taller than the average 8-year Height(B) > average
  b. Billy is tall (for an 8-year-old) Height(B) >> average
(24) a. Billy is taller than Freddie Height(B) > Height(F)
  b. Billy is tall compared to Freddie Height(B) >> Height(F)
Acknowledgments

Portions of this work were presented at the 4th Workshop on OT and Interpretation, the 1st VAAG Workshop, Logical Models of Reasoning with Vague Information, and the CUNY Syntax Supper. I would like to express my gratitude to the audiences for their comments and suggestions.

Thanks also to Nicole Gotzner for assistance with data analysis.

This work was funded by the European Science Foundation (ESF) and the Deutsche Forschungsgemeinschaft (DFG) under the EuroCORES call LogICCC.

References


