Frequency and distribution of *some* (but not *all*) implicatures

Judith Degen  
Departments of Brain & Cognitive Sciences and Linguistics  
University of Rochester

January 7, 2013
Overview

1. The Frequency Assumption

2. Why should we care?

3. The experiment
   - Collecting implicature ratings
   - Correlating ratings with contextual cues

4. Conclusions & future directions
(1) John: Was the exam easy?
   Mary: Some of the students failed.
   \[\sim\] Some, but not all of the students failed.
Scalar implicature

(1) John: Was the exam easy?
Mary: Some of the students failed.
\[ \leadsto \text{Some, but not all of the students failed.} \]
The assumption

The Frequency Assumption
Scalar inferences are generated more frequently than not when a scalar item like *some* is encountered.
“some typically gives rise to the implicature not all” (Gundel, 2009)

“the lower-bounded interpretation may be vanishingly rare in real-world communication” (Huang & Snedeker, 2009)

“Using a weak expression from a set of stronger alternatives often implies that the stronger alternatives are not applicable” (Bott, Bailey, & Grodner, 2012)

“as a generalized implicatum, the aforementioned [scalar] inference goes through in unmarked contexts, but it may be cancelled” (Horn, 1984)

“[scalar implicatures] show a degree of regularity and have the intuitive feel of components of conventional meaning” (Breheny, Katsos, & Williams, 2006)
Particularized Conversational Implicature (PCI)
Implicature arises in virtue of special features of the context.

Generalized Conversational Implicature (GCI)
Implicature arises unless context blocks it.
(2) John: Was the exam easy?
Mary: Some of the students failed.

\[ \text{GCI} \rightarrow \text{Some, but not all of the students failed.} \]
\[ \text{PCI} \rightarrow \text{The exam was not easy.} \]
(2) John: Was the exam easy?
   Mary: Some of the students failed.
   \[\text{GCI} \rightsquigarrow \text{Some, but not all of the students failed.}\]
   \[\text{PCI} \rightsquigarrow \text{The exam was not easy.}\]

(3) John: Is the teacher doing a good job?
   Mary: Some of the students failed.
   \[\text{GCI} \rightsquigarrow \text{Some, but not all of the students failed.}\]
   \[\text{PCI} \rightarrow \text{The exam was not easy.}\]
   \[\text{PCI} \rightsquigarrow \text{The teacher is not doing a good job.}\]
GCI/PCI and the Frequency Assumption

**Particularized Conversational Implicature (PCI)**
Implicature arises in virtue of special features of the context.

**Generalized Conversational Implicature (GCI)**
Implicature arises unless context blocks it.

- difference between GCI and PCI as difference in frequency of implicature (type? token?)
- but: high frequency of GCI type (like scalar implicature) does not entail the Frequency Assumption
The articulatory bottleneck problem

An argument for efficient communication (Levinson, 2000)

1. There is a significant bottleneck in the rate of information that can be transmitted via human speech.
2. Nevertheless, linguistic communication proceeds at a miraculous speed.
3. Thus, the communicative system must have evolved a solution to the problem of the articulatory bottleneck.

What is the solution?
The GCI-PCI distinction

Particularized Conversational Implicature (PCI)

Implicature arises in virtue of special features of the context.

Generalized Conversational Implicature (GCI)

Implicature arises unless context blocks it.

- Levinson’s Default Model translates the PCI-GCI distinction into processing terms:
  - PCI: implicature not computed by default (slow)
  - GCI: implicature computed by default upon encountering lexical trigger (fast); cancellation in a second step if necessary

- solution to the bottleneck problem: make inference cheap by making GCIs cost-free
The problem

- The problem with the idea that scalar inference is generally cheap/costless: it’s (probably) not true. (Bott & Noveck, 2004; Breheny et al., 2006; De Neys & Schaeken, 2007; Huang & Snedeker, 2009; Bott et al., 2012)
The problem

- The problem with the idea that scalar inference is generally cheap/costless: it’s (probably) not true. (Bott & Noveck, 2004; Breheny et al., 2006; De Neys & Schaeken, 2007; Huang & Snedeker, 2009; Bott et al., 2012)

- And the problem with costly implicature:
The problem

- The problem with the idea that scalar inference is generally cheap/costless: it’s (probably) not true. (Bott & Noveck, 2004; Breheny et al., 2006; De Neys & Schaeken, 2007; Huang & Snedeker, 2009; Bott et al., 2012)

- And the problem with costly implicature:
  - None! – say proponents of the Literal-First hypothesis (Huang & Snedeker, 2009; Bott et al., 2012):
    → implicature cost arises from there being two stages in implicature computation, semantic and pragmatic
The problem with the idea that scalar inference is generally cheap/costless: it’s (probably) not true. (Bott & Noveck, 2004; Breheny et al., 2006; De Neys & Schaeken, 2007; Huang & Snedeker, 2009; Bott et al., 2012)

And the problem with costly implicature:

- None! – say proponents of the Literal-First hypothesis (Huang & Snedeker, 2009; Bott et al., 2012):
  - implicature cost arises from there being two stages in implicature computation, semantic and pragmatic
- Given what we know about frequency effects in other domains of language processing, it’s a mystery if we assume that scalar implicatures are very frequent.
more frequent words are recognized more quickly and more accurately than less frequent words
more frequent and predictable words/structures/collocations are read more rapidly than less frequent ones
predictable inferences are drawn quickly

Why should scalar inferences be the exception?

Maybe because they’re not.

**Hypothesis 1: scalar implicatures are rare overall**

Given a scalar item like *some*, a scalar implicature is unlikely to arise.

\[ p(SI|\text{scalar item}) < p(\neg SI|\text{scalar item}) \]

**Hypothesis 2: scalar implicatures can be predicted from context**

The probability of a scalar implicature contextually depends on a wealth of syntactic / semantic / pragmatic / prosodic cues.

\[
\begin{align*}
& p(SI|\text{scalar item}, c_1, c_2, \ldots, c_n) \leq p(SI|\text{scalar item}) \\
& p(SI|\text{scalar item}, c_1, c_2, \ldots, c_n) \geq p(SI|\text{scalar item})
\end{align*}
\]
Probabilistic, constraint-based account of scalar implicatures (c.f. Russell, 2012)

Hypothesis 1: scalar implicatures are rare overall
Given a scalar item like *some*, support *for* a scalar implicature is lower than support *against* it.

\[ p(\text{SI}|\text{scalar item}) < p(\neg\text{SI}|\text{scalar item}) \]

Hypothesis 2: scalar implicatures can be predicted from context
The support for a scalar implicature contextually depends on a wealth of syntactic / semantic / pragmatic / prosodic cues.

\[ p(\text{SI}|\text{scalar item}, c_1, c_2, \ldots, c_n) \leq p(\text{SI}|\text{scalar item}) \]
\[ p(\text{SI}|\text{scalar item}, c_1, c_2, \ldots, c_n) \geq p(\text{SI}|\text{scalar item}) \]
Three steps: for each utterance with *some* in Switchboard corpus

1. information extraction/annotation
   - extract cue values or
   - annotate for cue values

2. collect implicature ratings on Mechanical Turk

3. correlate cues with implicature ratings
Cues

- **partitive**: $p(SI)$ should increase when *some* in partitive construction
  
  alex ate some (of the) cashews

- **discourse accessibility** of embedded NP: $p(SI)$ greater for accessible referents
  
  - information status (old, new, mediated)
  - topicality (topicalized, subject, other)
  - pre-/post-nominal modification

- **weak vs. strong *some***: $p(SI)$ greater for strong *some*
  
  um, well, some history books are pretty scary
  but my son needed some money

Reed (1991); Horn (1997); Ladusaw (1994)
Three steps:

1. extract database of utterances containing *some*-NPs from a corpus
2. collect implicature ratings for each utterance
3. correlate ratings with cues
Step 1 - getting the dataset

- 1389 items from the Switchboard corpus
- excluded 359 cases with singular count heads:

(4)  
  a. She stuck my name on some list.
  b. *She stuck my name on some, but not all list.

(5)  
  a. John kicked some cat off the street.
  b. John kicked some, but not all cat off the street.
Step 2 - collecting ratings

- 1389 items
- 20 items per block
- 10 ratings per item
- ratings collected on Amazon’s Mechanical Turk service (243 participants)
- procedure: rate similarity of original utterance and utterance with implicated content coded explicitly
- ... the more similar, the more the implicated content was arguably part of the speaker’s original meaning (as inferred by participants)
Speaker A: i mean, they just have beautiful, beautiful homes and they have everything. the kids only wear name brand things to school and it's one of these things,

Speaker B: oh me. well that makes it hard for you, doesn't it.

Speaker A: well it does, you know. it really does because i'm a single mom and i have a thirteen year old now and uh, you know, it does.

Speaker B: oh, me.

Speaker A: i mean, we do it to a point but uh, not to where she feels different,

Speaker B: yeah.

Speaker A: but some of them are very rich

but some, but not all of them are very rich

How similar is the statement with 'some, but not all' (green) to the statement with 'some' (red)?

Very different meaning □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ ^{
only 45% of ratings higher than scale midpoint (47% lower) good evidence that some does not usually provide strong support for a scalar implicature
only 45% of ratings higher than scale midpoint (47% lower)
good evidence that *some* does not usually provide strong support for a scalar implicature
(6) 6 – and then some of them are speaking in a foreign language that I don’t even understand.

(7) 5.9 – takes a heck of a long time, I guess, to get there in some states.

(8) 3.4 – what do you feel are some of the main problems.

(9) 2.5 – and I’ve got some broccoli, uh, onions some, some radishes and, uh, uh, beets.

(10) 2 – and he just really missed being able to be by himself, and having some peace and quiet

(11) 1.9 – they were some groups of it, the vicksburg,
Step 3 - investigate effect of cues

- each case in database annotated for cue values
  - quantifier strength: continuous value (experimentally collected ratings)
  - partitive: *no/yes*
  - discourse accessibility:
    - information status: *new/med/old*
    - topicality: *other/subject/topicalized*
    - modification: *unmodified/modified*

- data analysis: linear mixed effects model predicting rating from fixed effects of cues (and log-transformed sentence length) and random by-participant intercepts
(12) Some professors went to the party.

Horn (1997)

<table>
<thead>
<tr>
<th>Strong some</th>
<th>Weak some</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) presuppositional</td>
<td>non-presuppositional</td>
</tr>
<tr>
<td>b) partitive or proportional</td>
<td>cardinal</td>
</tr>
<tr>
<td>c) likely to give rise to SI</td>
<td>unlikely to give rise to SI</td>
</tr>
</tbody>
</table>
Quantifier strength

(12) Some professors went to the party.

Horn (1997)

<table>
<thead>
<tr>
<th>Strong some</th>
<th>Weak some</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) presuppositional</td>
<td>non-presuppositional</td>
</tr>
<tr>
<td>b) partitive or proportional</td>
<td>cardinal</td>
</tr>
<tr>
<td>c) likely to give rise to SI</td>
<td>unlikely to give rise to SI</td>
</tr>
</tbody>
</table>

(13) Weak
   a. but my son needed some money
   b. but my son needed money

(14) Strong
   a. um, well, some history books are pretty scary
   b. um, well, history books are pretty scary
The probability of a scalar inference increases with increasing quantifier strength.
Partitive results

<table>
<thead>
<tr>
<th>partitive</th>
<th>non-partitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>27%</td>
<td>73%</td>
</tr>
</tbody>
</table>

- global mean: 3.9
- medians:
  - global: 4
  - partitive: 5
  - non-partitive: 3
### Partitive results

<table>
<thead>
<tr>
<th>partitive</th>
<th>non-partitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>27%</td>
<td>73%</td>
</tr>
</tbody>
</table>

- **global mean**: 3.9
- **medians**:
  - global: 4
  - partitive: 5
  - non-partitive: 3

---

**The partitive increases the probability of a scalar inference.**

Judith Degen
Distribution of scalar implicatures
January 7, 2013 26 / 37
Partitive results

<table>
<thead>
<tr>
<th>partitive</th>
<th>non-partitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>27%</td>
<td>73%</td>
</tr>
</tbody>
</table>

- global mean: 3.9
- medians:
  - global: 4
  - partitive: 5
  - non-partitive: 3

- The partitive increases the probability of a scalar inference.
## Discourse accessibility results I: information status

Increasing discourse accessibility of the embedded NP referent increases the probability of a scalar inference.

<table>
<thead>
<tr>
<th></th>
<th>new</th>
<th>mediated</th>
<th>old</th>
</tr>
</thead>
<tbody>
<tr>
<td>33%</td>
<td>55%</td>
<td>12%</td>
<td></td>
</tr>
</tbody>
</table>
Increasing discourse accessibility of the embedded NP referent increases the probability of a scalar inference.
### Discourse accessibility results II: topicality

<table>
<thead>
<tr>
<th>other</th>
<th>subject</th>
<th>topicalized</th>
</tr>
</thead>
<tbody>
<tr>
<td>84%</td>
<td>15%</td>
<td>2%</td>
</tr>
</tbody>
</table>

*Topic:* and **some of those people**, they don’t deserve to be let loose.

*Subject:* **some of the stuff** is good

*Other:* it was starting to burn **some oil**
### Discourse accessibility results II: topicality

<table>
<thead>
<tr>
<th>other</th>
<th>subject</th>
<th>topicalized</th>
</tr>
</thead>
<tbody>
<tr>
<td>84%</td>
<td>15%</td>
<td>2%</td>
</tr>
</tbody>
</table>

- Topicalized *some*-NPs and *some*-NPs in subject position increase the probability of a scalar inference.

**topic:** and **some of those people**, they don’t deserve to be let loose.

**subject:** **some of the stuff** is good

**other:** it was starting to burn **some oil**
**Discourse accessibility results III: modification**

<table>
<thead>
<tr>
<th></th>
<th>modified</th>
<th>unmodified</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>48%</td>
<td>52%</td>
</tr>
</tbody>
</table>

*unmodified:* uh, but, uh, *some friends* have, uh, gone through this.

*modified:* uh, but, uh, *some friends of mine* have, uh, gone through this.
Modification interacts with information status: modification increases the probability of a scalar inference for new, but not old/mediated NP referents.

**unmodified:** uh, but, uh, **some friends** have, uh, gone through this.

**modified:** uh, but, uh, **some friends of mine** have, uh, gone through this.
• preliminary evidence that the Frequency Assumption does not hold (at least for *some*)

• support for a scalar implicature is increased
  • with increasing quantifier strength
  • when the *some*-NPIs partitive
  • when the embedded NP referent is discourse-accessible

• conclusion: cognitive cost associated with scalar implicatures may be a frequency effect rather than an effect of staged semantics-first processing
Future directions

- investigate further cues: monotonicity properties of the context, prosodic cues (e.g. *some* duration)
- investigate further scalar items
- investigate whether these cues affect the online processing of scalar implicatures
Thank you

- Mike Tanenhaus & the Tanenhaus lab
- Christine Gunlogson
- Greg Carlson
- Polly Jacobson
- Florian Jaeger


Horn, L. (1997). All John’s children are as bald as the King of France: Existential import and the geometry of opposition. In Cls 33 (pp. 155 – 179).


