A CONSTRAINT-BASED APPROACH TO THE MEANING AND USE OF QUANTIFIED EXPRESSIONS

Chris Cummins[1], Uli Sauerland[2], Stephanie Solt[2]

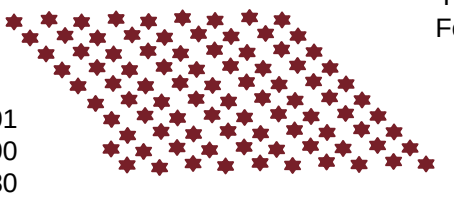
[1] Research Centre for English and Applied Linguistics, Cambridge
[2] Zentrum für Allgemeine Sprachwissenschaft, Berlin

EURO-XPRAG Workshop, June 9-11 2010

Motivation

Between 91 and 93
Between 90 and 100
…

Fewer than 93
Fewer than 94
Fewer than 100
Fewer than 1000
…



More than 91
More than 90
More than 80
More than 50
…

92
About 90
About 100

Constraint-based model

- Idea: quantified expression as output of multiple constraint satisfaction problem
- Potential constraints:
  - Informativeness
  - Granularity/numeral salience
  - Quantifier simplicity
  - Contextual activation of numeral/quantifier

Interpretation in constraints model

- Hearer uses speaker's output to draw inferences about the situation
  - Hearer knows that speaker chose to use utterance U_1 instead of utterance U_2, U_3, etc.
  - Therefore speaker considers U_1 more appropriate for situation than any given alternative
  - Example:
    - If some utterance U_n is more informative than U_1, *and just as good in every other way*, then hearer can infer than speaker does not think U_n is valid (classic SI)

Granularity

*The distance between Amsterdam and Vienna is **one thousand kilometers***

*The distance between Amsterdam and Vienna is **nine hundred sixty-five kilometers***

- Krifka (2009): Granularity modelled via scales that differ in density of scale points:

```
------------------40----------------------------50-------
----35------------40------------45------------50-------
-34-35-36-37-38-39-40-41-42-43-44-45-46-47-48-49-50-51-52-
```

SIs with modified numerals

- Has been observed that modified numerals do not give rise to scalar implicatures

  John has 3 children

  → 'not more than 3'

  John has more than 3 children

  ↛ 'not more than 4', i.e. exactly 4

- Prediction of Constraints Model: Expressions with modified numerals (e.g."more than n") should give rise to SIs at the appropriate granularity level

Experiment 1
Stimuli

A newspaper reported the following:

> **More than 100** people attended the public meeting about the new highway construction project'

Based on reading this, how many people do you think attended the meeting?

A) Between _____ and _____ people attended.

B) _____ people attended.

More than 100
More than 110
More than 93

At least 100
At least 110
At least 93

- Online survey via Amazon MTurk
- n= 100/condition

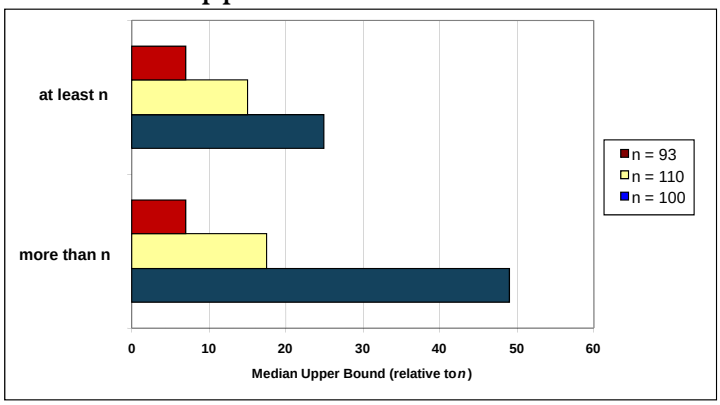
Experiment 1
Results

Median Estimates

	More than *n*		At least *n*	
***n* =**	**Upper bound**	**Most Likely**	**Upper Bound**	**Most Likely**
100	149	110	125	100
110	127,5	112	125	110
93	100	94	100	93

Experiment 1
Results

Median Upper Bound – Relative to *n*



Experiment 1
Subject Comments

- 'More than 100' condition:

  I feel that if there was more than 150, the newspaper would say more than 150.

  I chose the above number because I felt had the numbers been higher the paper would have said more than 200.

  I think 125 would be the next increment worthy of mentioning.

- 'More than 110' condition:

  I chose 135 because I felt that a number higher than that would have been described as, "Close to 150 people attended...".

  If it had been >120 it would have described them as such

  If it was more than 115 the writer would have probably said "Almost 120 people attended the meeting"

Prediction from constraints model

- Expressions with modified numerals (e.g."more than *n*") should give rise to SIs at the appropriate granularity level

  - Though assumed granularity level not consistent

However…

- Non-round numbers (e.g. 93) occur only on scale of granularity 1
  (…91…92…**93**…94…95…)
  - Strongest claim: 'more than 93' should implicate 'not more than 94', i.e. 'exactly 94'
  - Instead, respondents typically give range of 94-100
- Why?

Non-round numbers

- When is it felicitous to use *'more than n'* when *n* is not round?
  - Odd in out-of-the-blue context
  - But:

    U.S. coach Bob Bradley will call in more than 23 players when the Americans start their final training camp ahead of the World Cup.

    (http://soccernet.espn.go.com/world-cup/story/_/id/5096975/ce/us/bradley-call-more-23-training&cc=5739?ver=global)
  - Non-round numerals frequently occur in modified form when the numeral is salient in the context

Prediction from constraints model - 2

- Prior activation of the numeral will weaken granularity-based scalar inference
  - Gives speaker a potential additional reason to (re-) use numeral
  - Hearer is aware of this, and moderates strength of inference from this numeral accordingly

Experiment 2
Stimuli

Primed:

A: We need to sell 60 tickets to cover our costs. How are the ticket sales going?
B: So far we've sold fewer than 60 tickets

Unprimed:

A: We need to sell tickets to cover our costs. How are the ticket sales going?
B: So far we've sold fewer than 60 tickets

How many tickets have been sold?
From ____ to _____, most likely _____.

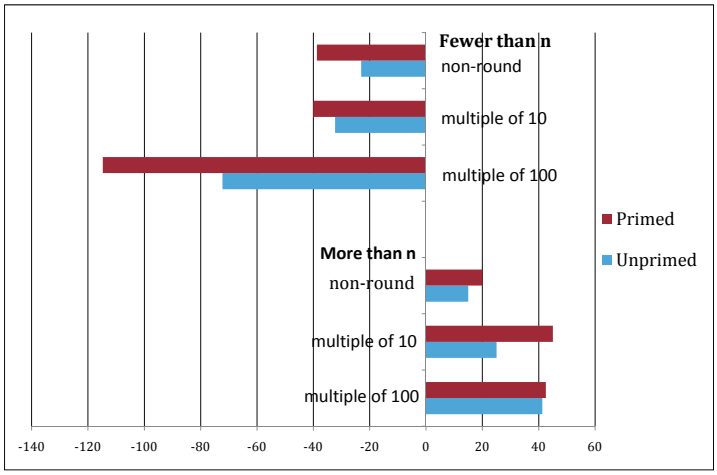
- Paper and pencil questionnaire

Experiment 2
Stimuli

- 3 levels of granularity
  - Multiple of 100
  - Multiple of 10
  - non-round
- 3 numerical expressions:
  - *more than n*
  - *fewer than n*
  - *about n*

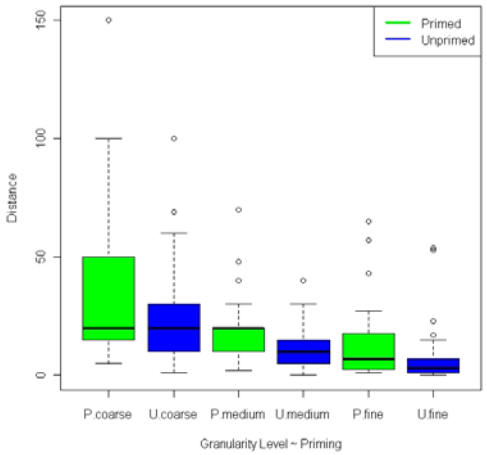
Experiment 2
Results

Median Upper/Lower Bound – Relative to *n*



Experiment 2
Results

Most Likely Value (vs. *n*)



- ANOVA shows significant effects for both granularity and priming ($p<.001$)

Experiment 3

Primed:

Salesperson: This shelf unit is designed to hold 60 CDs. How many CDs do you own?

Customer: I have more than 60 CDs

Unprimed:

Salesperson: This shelf unit is designed to hold CDs. How many CDs do you own?

Customer: I have more than 60 CDs

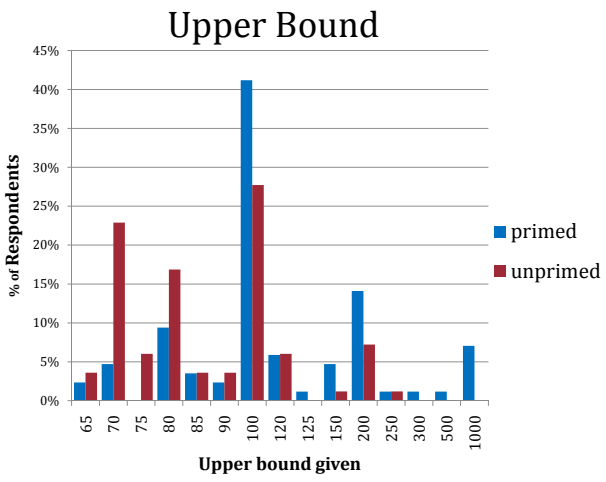
How many CDs do you think the customer owns?

Between ____ and _____; most likely number _____

- 1-question MTurk survey

Experiment 3

Results



Median Responses

Upper bound

Unprimed: 81

Primed: 100

Best estimate

Unprimed: 70

Primed: 80

Prediction from constraints model - 2

- Prior activation of the numeral will weaken granularity-based scalar inference

Implications for theory – Exp 1

- SI predicted by standard Gricean means
  - Now no need to stipulate absence of SI with these structures
- From RT point of view, also reasonable
  - Use of "more than n" indicates that "more than m" is not valid, if m > n, assuming m could be used without additional cognitive costs
- In either case, need to incorporate notion of granularity/numeral salience into model

Implications for theory – Exp 2/3

- Still in the spirit of the Gricean approach
  - Inferences arise from what we choose not to say; no choice, no inference!
- Also agrees with a Relevance-based account
  - Previously-mentioned numerals advantaged in Relevance
  - Therefore more likely to be selected in broader range of contexts
  - Hence less 'cue validity' for SIs

Is this actually RT?

- Selecting optimal output by constraint-based model similar to RT approach
- Could be viewed as a proposal unpacking the notion of Relevance in quantifier case
  - Elaborates Relevance by
    - Identifying its contributory factors
    - Quantifying their contribution
    - Evaluating which option achieves greatest Relevance

Role of constraint-based model

- Model does not appear *so far* to challenge
  - Relevance-theoretic accounts
  - traditional Gricean intuitions
- However, does present a useful means of generating non-obvious predictions
  - So far, restricted to the domain of quantification
  - Possible extension?
    - Issues: constraint set, assessing extent of violations

Summary

- SIs available with "more/fewer than n"
  - contrary to previous literature
  - restricted to appropriate granularity
- Inferences tempered by contextual activation
- Novel predictions from constraint-based account empirically validated
  - Not necessarily evidence for this account versus alternative accounts of pragmatic inference
  - However, do indicate that granularity and context effects *must be incorporated in models*

Thank you!