Cancellation, underspecification, and experimental pragmatics

Matthijs Westera

Institute for Logic, Language and Computation
University of Amsterdam

SFB colloquium
Düsseldorf, July 2nd 2014
1.1. Aim of this talk

To give a uniform account of three phenomena:

1. Ignorance inferences with scalar modifiers; data from: Geurts et al. 2010; Coppock & Brochhagen SALT 2013; Westera & Brasoveanu SALT 2014. (Joint work with Adrian Brasoveanu, UCSC.)

2. Diversity in exhaustivity inferences; data from: Van Tiel et al. (submitted)


General line of explanation

(i) each phenomenon is highly context-dependent;
(ii) experiments leave the context underspecified;
(iii) participants fill in the gaps based on typical use.
1.1. Aim of this talk

To give a uniform account of three phenomena:

1. **Ignorance inferences** with scalar modifiers; data from:
   - Geurts et al. 2010;
   - Coppock & Brochhagen SALT 2013;
   - Westera & Brasoveanu SALT 2014.

(Joint work with Adrian Brasoveanu, UCSC.)
1.1. Aim of this talk

To give a uniform account of three phenomena:

1. **Ignorance inferences** with scalar modifiers; data from:
   - Geurts et al. 2010;
   - Coppock & Brochhagen SALT 2013;
   - Westera & Brasoveanu SALT 2014.
   (Joint work with Adrian Brasoveanu, UCSC.)

2. Diversity in **exhaustivity inferences**; data from:
   - Van Tiel et al. (submitted)
1.1. Aim of this talk

To give a uniform account of three phenomena:

1. **Ignorance inferences** with scalar modifiers; data from:
   - Geurts et al. 2010;
   - Coppock & Brochhagen SALT 2013;
   - Westera & Brasoveanu SALT 2014.
   (Joint work with Adrian Brasoveanu, UCSC.)

2. **Diversity in exhaustivity inferences**; data from:
   - Van Tiel et al. (submitted)

3. **Diversity in “yes”/“no” licencing**; data from:
   - Brasoveanu, Farkas, and Roelofsen, 2013.
1.1. Aim of this talk

To give a uniform account of three phenomena:

1. Ignorance inferences with scalar modifiers; data from:
   - Geurts et al. 2010;
   - Coppock & Brochhagen SALT 2013;
   - Westera & Brasoveanu SALT 2014.
   (Joint work with Adrian Brasoveanu, UCSC.)

2. Diversity in exhaustivity inferences; data from:
   - Van Tiel et al. (submitted)

3. Diversity in “yes” / “no” licencing; data from:
   - Brasoveanu, Farkas, and Roelofsen, 2013.

General line of explanation

(i) each phenomenon is highly context-dependent;
1.1. Aim of this talk

To give a uniform account of three phenomena:

1. Ignorance inferences with scalar modifiers; data from:
   - Geurts et al. 2010;
   - Coppock & Brochhagen SALT 2013;
   - Westera & Brasoveanu SALT 2014.
   (Joint work with Adrian Brasoveanu, UCSC.)

2. Diversity in exhaustivity inferences; data from:
   - Van Tiel et al. (submitted)

3. Diversity in “yes”/“no” licencing; data from:
   - Brasoveanu, Farkas, and Roelofsen, 2013.

General line of explanation

(i) each phenomenon is highly context-dependent;
(ii) experiments leave the context underspecified;
1.1. Aim of this talk

To give a uniform account of three phenomena:

1. **Ignorance inferences** with scalar modifiers; data from:
   - Geurts et al. 2010;
   - Coppock & Brochhagen SALT 2013;
   - Westera & Brasoveanu SALT 2014.
   (Joint work with Adrian Brasoveanu, UCSC.)

2. Diversity in **exhaustivity inferences**; data from:
   - Van Tiel et al. (submitted)

3. Diversity in “yes”/“no” licencing; data from:
   - Brasoveanu, Farkas, and Roelofsen, 2013.

General line of explanation

(i) each phenomenon is highly context-dependent;
(ii) experiments leave the context underspecified;
(iii) participants fill in the gaps based on typical use.
1.1. Aim of this talk

To give a uniform account of three phenomena:

1. **Ignorance inferences** with scalar modifiers; data from:
   - Geurts et al. 2010;
   - Coppock & Brochhagen SALT 2013;
   - Westera & Brasoveanu SALT 2014.
   (Joint work with Adrian Brasoveanu, UCSC.)

2. Diversity in **exhaustivity inferences**; data from:
   - Van Tiel et al. (submitted)

3. Diversity in “yes”/“no” licencing; data from:
   - Brasoveanu, Farkas, and Roelofsen, 2013.

**General line of explanation**

(i) each phenomenon is highly context-dependent;
(ii) experiments leave the context underspecified;
(iii) participants fill in the gaps based on typical use.
1.2. The bigger picture

The dominant view: conversational implicatures are unreliable.
1.2. The bigger picture

The dominant view: conversational implicatures are *unreliable*.

Yet, *implicatures in plain cases are well known to be flimsy and context-dependent.* (Magri, 2011, p.13)

*But the effect is correspondingly weak: [it] produces cancelable implicatures rather than indefeasible truth conditional effects.* (Beaver & Clark, 2008, p.41)
1.2. The bigger picture

The dominant view: conversational implicatures are *unreliable*.

Yet, *implicatures in plain cases are well known to be flimsy and context-dependent.* (Magri, 2011, p.13)

*But the effect is correspondingly weak: [it] produces cancelable implicatures rather than indefeasible truth conditional effects.* (Beaver & Clark, 2008, p.41)

Phenomena 1 and 2 illustrate *why* this has become the dominant view, and my account suggests that this is unjustified.
1.2. The bigger picture

The dominant view: conversational implicatures are *unreliable*.

Yet, *implicatures in plain cases are well known to be flimsy and context-dependent.* (Magri, 2011, p.13)

*But the effect is correspondingly weak: [it] produces cancelable implicatures rather than indefeasible truth conditional effects.* (Beaver & Clark, 2008, p.41)

Phenomena 1 and 2 illustrate *why* this has become the dominant view, and my account suggests that this is unjustified.

Indeed, I think the Gricean theory of meaning commits us to conversational implicatures being *as reliable as entailments.*
Outline

Ignorance implicatures and scalar modifiers
   The puzzle
   Experiment design
   Results and discussion

Exhaustivity inferences

Why conversational implicatures may well be strong

“Yes” and “no”

Conclusion
Outline

Ignorance implicatures and scalar modifiers
   The puzzle
   Experiment design
   Results and discussion

Exhaustivity inferences

Why conversational implicatures may well be strong

“Yes” and “no”

Conclusion
2.1. Ignorance implicatures and scalar modifiers

Geurts & Nouwen (2007):

(1) a. I saw at most ten of the coins. \[\sim \text{not sure how many.}\]

b. I saw less than ten of the coins. \[\not\sim \text{not sure how many.}\]
2.1. Ignorance implicatures and scalar modifiers

Geurts & Nouwen (2007):

(1) a. I saw at most ten of the coins. \( \sim not\ sure\ how\ many.\)
    b. I saw less than ten of the coins. \( \nrightarrow not\ sure\ how\ many.\)

- (1a,b) contrast in validity judgment task; (Geurts et al. ’10)
- but not in truth judgment task. (Coppock & Brochhagen ’13)
2.1. Ignorance implicatures and scalar modifiers

Geurts & Nouwen (2007):

(1) a. I saw at most ten of the coins. \(\sim\) not sure how many.
   b. I saw less than ten of the coins. \(\nRightarrow\) not sure how many.

- (1a,b) contrast in validity judgment task; (Geurts et al. ’10)
- but not in truth judgment task. (Coppock & Brochhagen ’13)

Coppock & Brochhagen’s account:

(i) “at most” / “less than” are semantically distinct;
(ii) this yields a difference in ignorance implicature;
(iii) to which truth judgements are insensitive.
2.1. Ignorance implicatures and scalar modifiers

Geurts & Nouwen (2007):

(1) a. I saw at most ten of the coins. \(\sim\) not sure how many.
    b. I saw less than ten of the coins. \(\not\sim\) not sure how many.

- (1a,b) contrast in validity judgment task; (Geurts et al. ’10)
- but not in truth judgment task. (Coppock & Brochhagen ’13)

Coppock & Brochhagen’s account:

(i) “at most” / “less than” are semantically distinct;
(ii) this yields a difference in ignorance implicature;
(iii) to which truth judgements are insensitive.

Problems:

- other implicatures are detected by truth judgement;
  (C&B; see also scalar implicatures literature)
- ignorance implicatures are in fact context-dependent.
2.2. Context-dependence

(2) Exactly how many of the coins did you see?
   I saw \textit{at most} ten of the coins.\(\uparrow\) \(\sim\) ignorance.

(3) Did you see at most ten of the coins?
   (Yes,) I saw at most ten of the coins. \(\not\sim\) ignorance.

(My judgements; actual data to follow.)
2.2. Context-dependence

(2) Exactly how many of the coins did you see?
I saw \textit{at most} ten of the coins.\(^\rightarrow\) \(\sim\) ignorance.

(3) Did you see at most ten of the coins?
(Yes,) I saw at most ten of the coins. \(\nRightarrow\) ignorance.

(My judgements; actual data to follow.)

Ignorance inferences effectively take two steps:
1. \textit{What’s the context like; was a precise answer desired?}
2. \textit{If so, then why didn’t the speaker give one?}
Step 1 relies on an \textit{explicit} QUD or intonation.
2.3. Guessing the QUD

With un(der)specified QUD, participants guess based on:

› their knowledge of the sentence’s typical use;
› the experimental task.

(Because there isn’t anything else.)
2.3. Guessing the QUD

With *(under)*specified QUD, participants guess based on:

- their knowledge of the sentence’s typical use;
- the experimental task.

(Because there isn’t anything else.)

Westera & Brasoveanu’s account

(i) truth judgement task is suggestive of an imprecise context;
(ii) validity judgement task can be precise or imprecise;
(iii) “at most” is used more than “less than” in precise contexts.
2.3. Guessing the QUD

With un(der)specified QUD, participants guess based on:

- their knowledge of the sentence’s typical use;
- the experimental task.

(Because there isn’t anything else.)

Westera & Brasoveanu’s account

(i) truth judgement task is suggestive of an imprecise context;
(ii) validity judgement task can be precise or imprecise;
(iii) “at most” is used more than “less than” in precise contexts.

We take (iii) from Cummins et al.’s (2012) corpus study:

- “less than” occurs relatively more often with round numbers.
2.4. Predictions

We predict for *truth judgement*:

(i) fully specified QUD $\sim$ ignorance depends on QUD;
(ii) underspecified QUD $\sim$ no ignorance implicatures.

Coppock & Brochhagen verified (ii); Geurts et al. verified (iv).

We did two experiments to jointly test (iii) and (iv).
2.4. Predictions

We predict for truth judgement:

(i) fully specified QUD \sim ignorances depends on QUD;
(ii) underspecified QUD \sim no ignorance implicatures.

And for validity judgement:

(iii) fully specified QUD \sim ignorances depends on QUD;
(iv) underspecified QUD \sim ignorance depends on typical usage.
2.4. Predictions

We predict for truth judgement:
(i) fully specified QUD \(\sim\) ignorance depends on QUD;
(ii) underspecified QUD \(\sim\) no ignorance implicatures.

And for validity judgement:
(iii) fully specified QUD \(\sim\) ignorance depends on QUD;
(iv) underspecified QUD \(\sim\) ignorance depends on typical usage.

- Coppock & Brochhagen verified (ii);
- Geurts et al. verified (iv).
2.4. Predictions

We predict for truth judgement:

(i) fully specified QUD $\leadsto$ ignorance depends on QUD;
(ii) underspecified QUD $\leadsto$ no ignorance implicatures.

And for validity judgement:

(iii) fully specified QUD $\leadsto$ ignorance depends on QUD;
(iv) underspecified QUD $\leadsto$ ignorance depends on typical usage.

- Coppock & Brochhagen verified (ii);
- Geurts et al. verified (iv).

We did two experiments to jointly test (iii) and (iv).
2.5. Experiment design

Two experiments with the same design, 3 screens per stimulus:

1. Judge’s question (QUD);
2. Witness’ answer, as self-paced reading task;
3. Judge’s inference, with validity judgement task (5-point scale).
The judge asks:

“What did you find under the bed?”

The witness answers:
_ ______ __ _____ ___ __ the ___________ ______ ___ ___
diamonds
under ___ ___
bed
Based on this, the judge concludes:

“The witness doesn’t know exactly how many of the diamonds she found under the bed.”

How justified is the judge in drawing that conclusion?

(not justified at all) \(1\) \(\circ\) \(2\) \(\circ\) \(3\) \(\circ\) \(4\) \(\circ\) \(5\) \(\circ\) (strongly justified)
2.5. Experiment design

Two experiments with the same design, 3 screens per stimulus:

1. Judge’s question (QUD);
2. Witness’ answer, as self-paced reading task;
3. Judge’s inference, with validity judgement task (5-point scale).
2.5. Experiment design

Two experiments with the same design, 3 screens per stimulus:

1. Judge’s question (QUD);
2. Witness’ answer, as *self-paced reading* task;
3. Judge’s inference, with *validity judgement* task (5-point scale).

- 3 question types × 2 answer types = 6 conditions;
- Latin square design, 108 stimuli (36 items + 72 fillers);
- 35 and 51 participants, respectively (ling. undergrads).
2.6. Items

QUD types experiment I:

- **Polar**: Did you $V$ *Mod* ten of the $N$ *PP*?
  ($V \in \{\text{see, hear, find}\}$; *Mod* as in answer)
- **What**: What did you $V$ *PP*?
- **HowMany**: How many of the $N$ did you $V$ *PP*?
2.6. Items

QUD types experiment I:
- **Polar**: Did you $V$ Mod ten of the $N$ $PP$?
  \[(V \in \{\text{see, hear, find}\}; \text{Mod as in answer})\]
- **What**: What did you $V$ $PP$?
- **HowMany**: How many of the $N$ did you $V$ $PP$?

QUD types experiment II:
- **Approx**: Approximately how many [...]?
- **Exact**: Exactly how many [...]?
- **Disjunct**: Did you $V$ eight, nine, ten or eleven [...]?
2.6. Items

**QUD types experiment I:**

- **Polar:** Did you \( V \) \( Mod \) ten of the \( N \) \( PP \)?
  \((V \in \{\text{see, hear, find}\}; \ Mod \) as in answer\)
- **What:** What did you \( V \) \( PP \)?
- **HowMany:** How many of the \( N \) did you \( V \) \( PP \)?

**QUD types experiment II:**

- **Approx:** Approximately how many [...]?
- **Exact:** Exactly how many [...]?
- **Disjunct:** Did you \( V \) eight, nine, ten or eleven [...]?

**Answer types** (same in both experiments):

- **Sup:** I \( V \) at most ten of the \( Ns \) \( PP \).
- **Comp:** I \( V \) less than ten of the \( Ns \) \( PP \).
2.6. Items

QUD types experiment I:
- **Polar**: Did you $V$ *Mod* ten of the $N$ *PP*? 
  ($V \in \{\text{see, hear, find}\}$; *Mod* as in answer)
- **What**: What did you $V$ *PP*?
- **HowMany**: How many of the $N$ did you $V$ *PP*?

QUD types experiment II:
- **Approx**: Approximately how many [...]?
- **Exact**: Exactly how many [...]?
- **Disjunct**: Did you $V$ eight, nine, ten or eleven [...]?

Answer types (same in both experiments):
- **Sup**: I $V$ at most ten of the $Ns$ *PP*.
- **Comp**: I $V$ less than ten of the $Ns$ *PP*.

Inference (always *ignorance* in items):
The witness doesn’t know exactly how many of the $N$ she $V$ *PP*.
2.7. Results: validity judgements

<table>
<thead>
<tr>
<th>QUD type</th>
<th>Mean judgments and SEs (scale: 1−5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polar</td>
<td>3.2</td>
</tr>
<tr>
<td>What</td>
<td>3.4</td>
</tr>
<tr>
<td>HowMany</td>
<td>3.6</td>
</tr>
<tr>
<td>Approx</td>
<td>3.2</td>
</tr>
<tr>
<td>Exact</td>
<td>3.4</td>
</tr>
<tr>
<td>Disjunct</td>
<td>3.6</td>
</tr>
</tbody>
</table>

The diagrams show the mean judgments and SEs for different QUD types, with markers indicating SUP and COMP conditions.
2.8. Generalizations/discussion: validity judgements

Weaker ignorance in Polar, Approx:
  ▶ Explanation: these do not ask for a precise answer.

Stronger ignorance in What, Exact, Disjunct;
  ▶ Explanation: these ask for a precise answer.

Contrast Sup/Comp only in HowMany:
  ▶ Explanation: this is underspecified for precision...
  ▶ hence the typical use of “at most”/“less than” kicks in.
2.9. Results: reading times experiment 1

Mean reading times (ms) and SEs

- **SUP**
- **COMP**

Regions

<table>
<thead>
<tr>
<th>Mean reading times (ms) and SEs</th>
</tr>
</thead>
<tbody>
<tr>
<td>220</td>
</tr>
<tr>
<td>240</td>
</tr>
<tr>
<td>260</td>
</tr>
<tr>
<td>280</td>
</tr>
</tbody>
</table>

- **HowMany**
- **What**
- **Polar**

*Regions include:*
- ten
- of
- the
- insults
- during
- dinner
2.10. Results: reading times experiment 2

- **Regions**
  - Mean reading times (ms) and SEs

- **Graph**
  - **Sup**
  - **COMP**
  - **Disjunct**
  - **Approx**
  - **Exact**

- **X-axis**: Regions
  - ten
  - of
  - the
  - insults
  - during
  - dinner

- **Y-axis**: Mean reading times (ms) and SEs
  - 200
  - 220
  - 240
  - 260
  - 280

- **Legend**
  - **Sup**
  - **COMP**
2.11. Generalizations/discussion: reading times

**Experiment I**: slower reading $\sim$ stronger ignorance.

**Experiment II**: no effect, probably due to *priming*:
- fillers tested only ignorance inferences (unlike in exp. 1);
- if we ‘correct’ for priming, slower reading $\sim$ stronger ignorance!
2.11. Generalizations/discussion: reading times

**Experiment I**: slower reading $\sim$ stronger ignorance.

**Experiment II**: no effect, probably due to *priming*:
- fillers tested only ignorance inferences (unlike in exp. 1);
- if we ‘correct’ for priming, slower reading $\sim$ stronger ignorance!

**Possible explanations**

Slower reading may be due to:

(A) processing cost of ignorance inference; or

(B) subvocalization with special intonation for ignorance.

(e.g., J.D. Fodor, 2002)
2.11. Generalizations/discussion: reading times

**Experiment I**: slower reading $\sim$ stronger ignorance.

**Experiment II**: no effect, probably due to *priming*:
- fillers tested only ignorance inferences (unlike in exp. 1);
- if we ‘correct’ for priming, slower reading $\sim$ stronger ignorance!

**Possible explanations**

Slower reading may be due to:

(A) *processing cost* of ignorance inference; or

(B) *subvocalization* with special intonation for ignorance.

(e.g., J.D. Fodor, 2002)

If (B), self-paced reading would give us a handle on intonation.
2.12. Conclusion (of this part)

The puzzle(s) are solved in terms of:

- contextual underspecification; and
- typical use;
2.12. Conclusion (of this part)

The puzzle(s) are solved in terms of:

- contextual underspecification; and
- typical use;

And no other type of account seems to be available...

Shouldn’t we also explain typical use? For our purposes, not really, but let’s try:

(i) only “at most” mentions a non-excluded possibility;
(ii) this creates a ‘slight preference’ for use in precise contexts.

Not sure if this is semantics/pragmatics or psychology... Coppock & Brochhagen may assign too much weight to (i).
2.12. Conclusion (of this part)

The puzzle(s) are solved in terms of:

- contextual underspecification; and
- typical use;

And no other type of account seems to be available...

Shouldn’t we also explain typical use?
For our purposes, not really,
2.12. Conclusion (of this part)

The puzzle(s) are solved in terms of:

- contextual underspecification; and
- typical use;

And no other type of account seems to be available...

Shouldn’t we also *explain* typical use?

For our purposes, not really, but let’s try:

(i) only “at most” *mentions* a non-excluded possibility;
(ii) this creates a ‘slight preference’ for use in precise contexts.
The puzzle(s) are solved in terms of:

- contextual underspecification; and
- typical use;

And no other type of account seems to be available...

Shouldn’t we also explain typical use? For our purposes, not really, but let’s try:

(i) only “at most” mentions a non-excluded possibility;
(ii) this creates a ‘slight preference’ for use in precise contexts.

Not sure if this is semantics/pragmatics or psychology...
2.12. Conclusion (of this part)

The puzzle(s) are solved in terms of:

- contextual underspecification; and
- typical use;

And no other type of account seems to be available...

Shouldn’t we also explain typical use?
For our purposes, not really, but let’s try:

(i) only “at most” mentions a non-excluded possibility;
(ii) this creates a ‘slight preference’ for use in precise contexts.

Not sure if this is semantics/pragmatics or psychology...

Coppock & Brochhagen may assign too much weight to (i).
Outline

Ignorance implicatures and scalar modifiers
   The puzzle
   Experiment design
   Results and discussion

Exhaustivity inferences

Why conversational implicatures may well be strong

“Yes” and “no”

Conclusion
3.1. Context-dependence of exhaustivity inferences

Like ignorance, exhaustivity inferences are QUD-dependent:

(4) Is the tea warm?
   (Yeah,) it’s warm.  \(\rightarrow\) It is not hot.

(5) Is the tea warm or hot?
   It’s warm.  \(\sim\) It is not hot.
3.1. Context-dependence of exhaustivity inferences

Like ignorance, exhaustivity inferences are QUD-dependent:

(4) Is the tea warm?
   (Yeah,) it’s warm.  \[ \rightarrow \text{It is not hot.} \]

(5) Is the tea warm or hot?
   It’s warm.  \[ \rightarrow \text{It is not hot.} \]

As before, with an un(der)specified QUD:

- participants must *guess* based on typical use.
3.2. Van Tiel et al.’s (submitted) results

- their best model still leaves 50% of variance unexplained;
  (based on, e.g., semantic distance)
- might explain it?
3.2. Van Tiel et al.’s (submitted) results

- their best model still leaves 50% of variance unexplained;
  (based on, e.g., semantic distance)
3.2. Van Tiel et al.’s (submitted) results

- their best model still leaves 50% of variance unexplained;
  (based on, e.g., semantic distance)

- might *typical use* explain it?
3.3. A tentative measure of typical use

To explain the variance in terms of typical use:

- we need to *quantify typical use*; in particular:
- the probability that $B$ is relevant given that $A$ is said.
3.3. A tentative measure of typical use

To explain the variance in terms of typical use:

- we need to *quantify typical use*; in particular:
- the probability that $B$ is relevant given that $A$ is said.

**Tentative proposal**

Let’s look in a corpus for:

- $\text{co-relevance}(B,A) \approx \#“A \text{ or even } B” / \#“A \text{ or even”}$; i.e.,
- the probability that, given that there is a relevant, stronger alternative for $A$, it is $B$;
3.3. A tentative measure of typical use

To explain the variance in terms of typical use:

- we need to *quantify typical use*; in particular:
  - the probability that $B$ is relevant given that $A$ is said.

**Tentative proposal**

Let’s look in a corpus for:

- $\text{co-relevance}(B,A) \approx \frac{\#“A \text{ or even } B”}{\#“A \text{ or even}”}$; i.e.,
  - the probability that, given that there is a relevant, stronger alternative for $A$, it is $B$;
  - (taking into account synonyms, polysemy, etc.)
3.4. Explaining Van Tiel et al.’s results

- cheap/free
- sometimes/always
- some/all
- possible/certain
- may/will
- difficult/impossible
- rare/extinct
- may/have to
- warm/hot
- few/none
- low/depleted
- hard/unsolvable
- allowed/obligatory
- scarce/unavailable
- try/succeed
- palatable/delicious
- memorable/unforgettable
- like/love
- good/perfect
- good/excellent
- cool/cold
- hungry/starving
- adequate/good
- unsettling/horrific
- dislike/loathe
- believe/know
- start/finish
- participate/win
- wary/scared
- old/ancient
- big/enormous
- snug/tight
- attractive/stunning
- special/unique
- pretty/beautiful
- intelligent/brilliant
- funny/hilarious
- dark/black
- small/tiny
- ugly/hideous
- silly/ridiculous
- tired/exhausted
- content/happy
3.4. Explaining Van Tiel et al.’s results

- cheap/free
- sometimes/always
- some/all
- possible/certain
- may/will
- difficult/impossible
- rare/extinct
- may/have to
- warm/hot
- few/none
- low/depleted
- hard/unsolvable
- allowed/obligatory
- scarce/unavailable
- try/succeed
- palatable/delicious
- memorable/unforgettable
- like/love
- good/perfect
- good/excellent
- cool/cold
- hungry/starving
- adequate/good
- unsettling/horrific
- dislike/loathe
- believe/know
- start/finish
- participate/win
- wary/scared
- old/ancient
- big/enormous
- snug/tight
- attractive/stunning
- special/unique
- pretty/beautiful
- intelligent/brilliant
- funny/hilarious
- dark/black
- small/tiny
- ugly/hideous
- silly/ridiculous
- tired/exhausted
- content/happy

Google N-grams:

- cheap/free
- sometimes/always
- some/all
- possible/certain
- may/will
- difficult/impossible
- rare/extinct
- may/have to
- warm/hot
- few/none
- low/depleted
- hard/unsolvable
- allowed/obligatory
- scarce/unavailable
- try/succeed
- palatable/delicious
- memorable/unforgettable
- like/love
- good/perfect
- good/excellent
- cool/cold
- hungry/starving
- adequate/good
- unsettling/horrific
- dislike/loathe
- believe/know
- start/finish
- participate/win
- wary/scared
- old/ancient
- big/enormous
- snug/tight
- attractive/stunning
- special/unique
- pretty/beautiful
- intelligent/brilliant
- funny/hilarious
- dark/black
- small/tiny
- ugly/hideous
- silly/ridiculous
- tired/exhausted
- content/happy
3.4. Explaining Van Tiel et al.’s results

- cheap/free
- sometimes/always
- some/all
- possible/certain
- may/will
- difficult/impossible
- rare/extinct
- may/have to
- warm/hot
- few/none
- low/depleted
- hard/unsolvable
- allowed/obligatory
- scarce/unavailable
- try/succeed
- palatable/delicious
- memorable/unforgettable
- like/love
- good/perfect
- good/excellent
- cool/cold
- hungry/starving
- adequate/good
- unsettling/horrific
- dislike/loathe
- believe/know
- start/finish
- participate/win
- wary/scared
- old/ancient
- big/enormous
- snug/tight
- attractive/stunning
- special/unique
- pretty/beautiful
- intelligent/brilliant
- funny/hilarious
- dark/black
- small/tiny
- ugly/hideous
- silly/ridiculous
- tired/exhausted
- content/happy

Google N-grams:
3.5. Conclusion (of this part)

So, variance in exhaustivity might be due to (again):
  - QUD-underspecification; and
  - typical use.

This could be quantified with a suitable corpus measure.
3.5. Conclusion (of this part)

So, variance in exhaustivity might be due to (again):
  ▶ QUD-underspecification; and
  ▶ typical use.

This could be quantified with a suitable corpus measure.

Shouldn’t we also explain typical use?
3.5. Conclusion (of this part)

So, variance in exhaustivity might be due to (again):

- QUD-underspecification; and
- typical use.

This could be quantified with a suitable corpus measure.

Shouldn’t we also explain typical use?

- This seems to be a sociological, not linguistic, issue;
- it is about what we generally talk about;
- not how we manage to communicate.
3.5. Conclusion (of this part)

So, variance in exhaustivity might be due to (again):
- QUD-underspecification; and
- typical use.

This could be quantified with a suitable corpus measure.

Shouldn’t we also explain typical use?
- This seems to be a sociological, not linguistic, issue;
- it is about what we generally talk about;
- not how we manage to communicate.

(And the same holds for ‘lexical scales’.)
Ignorance implicatures and scalar modifiers
   The puzzle
   Experiment design
   Results and discussion

Exhaustivity inferences

Why conversational implicatures may well be strong

“Yes” and “no”

Conclusion
4.1. Cancelability

The main reason for regarding conversational implicatures as weak is their cancelability. Textbook example:

(6) I saw some of the students. Indeed, I saw all.
4.1. Cancelability

The main reason for regarding conversational implicatures as weak is their cancelability. Textbook example:

(6) I saw some of the students. Indeed, I saw all.

But in fact:

(7) A: Did you see all of the students?
    B: I saw some of the students. # Indeed, I saw all.
4.1. Cancelability

The main reason for regarding conversational implicatures as weak is their cancelability. Textbook example:

(6) I saw some of the students. Indeed, I saw all.

But in fact:

(7) A: Did you see all of the students?  
    B: I saw some of the students. # Indeed, I saw all.

(8) A: Who saw some of the students?  
    B: I saw some of the students. Indeed, I saw all.
4.1. Cancelability

The main reason for regarding conversational implicatures as weak is their cancelability. Textbook example:

(6) I saw some of the students. Indeed, I saw all.

But in fact:

(7) A: Did you see all of the students?  
    B: I saw some of the students.  # Indeed, I saw all.

(8) A: Who saw some of the students?  
    B: I saw some of the students. Indeed, I saw all.

Thus, cancelation in (6) is in fact contextual underspecification, disambiguated by “indeed, I saw all.”. 
4.1. Cancelability

The main reason for regarding conversational implicatures as *weak* is their *cancelability*. Textbook example:

(6) I saw some of the students. Indeed, I saw all.

But in fact:

(7) A: Did you see *all* of the students?
   B: I saw *some* of the students. # Indeed, I saw all.

(8) A: Who saw some of the students?
   B: I saw some of the students. Indeed, I saw all.

Thus, cancelation in (6) is in fact *contextual underspecification*, disambiguated by “indeed, I saw all.”.

Geurts (2010): *to actually make a CI and then contradict it can hardly be cooperative.*
4.2. Cancelability as context-dependence

Context-dependence seems to be what Grice had in mind:

[Conversational implicatures] may be explicitly canceled, by the addition of a clause that states or implies that the speaker has opted out [of the Cooperative Principle], or it may be contextually cancelled, if the form of utterance that usually carries it is used in a context that makes it clear that the speaker is opting out. (Grice, 1989, p.57)
Context-dependence seems to be what Grice had in mind:

[Conversational implicatures] may be explicitly canceled, by the addition of a clause that states or implies that the speaker has opted out [of the Cooperative Principle], or it may be contextually cancelled, if the form of utterance that usually carries it is used in a context that makes it clear that the speaker is opting out. (Grice, 1989, p.57)

“Okay, but doesn’t their context-dependence imply that conversational implicatures are weaker?”
4.3. Calculability

Conversational implicatures are necessarily strong/reliable:

1. they're part of what a speaker intends to convey (Grice '89);
2. a rational speaker will try to ensure that her intention is realized;
3. hence, she will try to ensure that the relevant contextual features are mutually known (e.g., by intonation).

The presence of a conversational implicature must be capable of being worked out; for even if it can in fact be intuitively grasped, unless the intuition is replaceable by an argument, the implicature (if present at all) will not count as a conversational implicature; (Grice, 1989, p.31)

This may be about linguists, but the same holds for language users.
4.3. Calculability

Conversational implicatures are necessarily strong/reliable:

1. they’re part of what a speaker *intends to convey* (Grice ’89);
4.3. Calculability

Conversational implicatures are necessarily strong/reliable:
1. they’re part of what a speaker *intends to convey* (Grice ’89);
2. a rational speaker will try to ensure that her intention is realized;
4.3. Calculability

Conversational implicatures are necessarily strong/reliable:

1. they’re part of what a speaker *intends to convey* (Grice ’89);
2. a rational speaker will try to ensure that her intention is realized;
3. hence, she will try to ensure that the relevant contextual features are mutually known (e.g., by intonation).
4.3. Calculability

Conversational implicatures are necessarily strong/reliable:

1. they’re part of what a speaker *intends to convey* (Grice ’89);
2. a rational speaker will try to ensure that her intention is realized;
3. hence, she will try to ensure that the relevant contextual features are mutually known (e.g., by intonation).

*The presence of a conversational implicature must be capable of being worked out; for even if it can in fact be intuitively grasped, unless the intuition is replaceable by an argument, the implicature (if present at all) will not count as a conversational implicature; [...]*

(Grice, 1989, p.31)
4.3. Calculability

Conversational implicatures are necessarily strong/reliable:

1. they’re part of what a speaker intends to convey (Grice ’89);
2. a rational speaker will try to ensure that her intention is realized;
3. hence, she will try to ensure that the relevant contextual features are mutually known (e.g., by intonation).

The presence of a conversational implicature must be capable of being worked out; for even if it can in fact be intuitively grasped, unless the intuition is replaceable by an argument, the implicature (if present at all) will not count as a conversational implicature; [...] 

(Grice, 1989, p.31)

This may be about linguists, but the same holds for language users.
4.4. Uncalculable ‘implicatures’

If an intended inference is *uncalculable*, the hearer will think:

i. either the speaker (experimenter?) didn’t mean to convey it; or

ii. something went wrong in communication (e.g., no intonation).

Either way, the detected inferences may be rather weak... But this doesn’t mean conversational implicatures are weak.
4.4. Uncalculable ‘implicatures’

If an intended inference is *uncalculable*, the hearer will think:

i. either the speaker (experimenter?) didn’t mean to convey it; or

ii. something went wrong in communication (e.g., no intonation).

- In case of (i), a similar *inference* may still surface as a weaker *typicality* inference.
4.4. Uncalculable ‘implicatures’

If an intended inference is uncalculable, the hearer will think:

i. either the speaker (experimenter?) didn’t mean to convey it; or

ii. something went wrong in communication (e.g., no intonation).

- In case of (i), a similar *inference* may still surface as a weaker *typicality* inference.
- In case of (ii), a similar *inference* may still surface depending on *typical use* of the expression.
4.4. Uncalculable ‘implicatures’

If an intended inference is uncalculable, the hearer will think:

i. either the speaker (experimenter?) didn’t mean to convey it; or
ii. something went wrong in communication (e.g., no intonation).

▶ In case of (i), a similar inference may still surface as a weaker typicality inference.
▶ In case of (ii), a similar inference may still surface depending on typical use of the expression.
▶ Either way, the detected inferences may be rather weak...
4.4. Uncalculable ‘implicatures’

If an intended inference is *uncalculable*, the hearer will think:

i. either the speaker (experimenter?) didn’t mean to convey it; or
ii. something went wrong in communication (e.g., no intonation).

- In case of (i), a similar *inference* may still surface as a weaker *typicality* inference.
- In case of (ii), a similar *inference* may still surface depending on *typical use* of the expression.
- Either way, the detected inferences may be rather weak...

But this doesn’t mean conversational implicatures are weak.
Outline

Ignorance implicatures and scalar modifiers
  The puzzle
  Experiment design
  Results and discussion

Exhaustivity inferences

Why conversational implicatures may well be strong

“Yes” and “no”

Conclusion
5.1. “Yes” and “no” licencing

(9) John didn’t come to the party.
a. Yes he *did*. / No, he didn’t.
b. Yes, he didn’t. / No, he *did*.
5.1. “Yes” and “no” licencing

(9) John didn’t come to the party.
   a. Yes he did. / No, he didn’t.
   b. Yes, he didn’t. / No, he did.

Krifka’s (2013) account

1. “yes”/“no” confirm/negate a salient proposition;
2. negative sentences make pos. and neg. proposition salient.
5.1. “Yes” and “no” licencing

(9) John didn’t come to the party.
   a. Yes he *did*. / No, he didn’t. ~ relative to pos. prop.
   b. Yes, he didn’t. / No, he *did*. ~ relative to neg. prop.

Krifka’s (2013) account

1. “yes”/“no” confirm/negate a *salient proposition*;
2. negative sentences make pos. and neg. proposition salient.
5.1. “Yes” and “no” licencing

(9) John didn’t come to the party.
   a. Yes he *did*. / No, he didn’t.  \(\sim\) relative to pos. prop.
   b. Yes, he didn’t. / No, he *did*.  \(\sim\) relative to neg. prop.

Krifka’s (2013) account

1. “yes”/“no” confirm/negate a *salient proposition*;
2. negative sentences make pos. and neg. proposition salient.

Problems:

(i) “yes”/“no”-licensing is very much context-dependent;  
   (my judgements)

(ii) words like “never”, “no one”, DE quantifiers...  
    (Brasoveanu et al., 2013)
5.2. Problem (i): context-dependence

It seems to me that:

- if some proposition is particularly relevant to a speaker...
- she will use that as the reference point for “yes” / “no”.

(Disclaimer: my own judgement only.)
5.2. Problem (i): context-dependence

It seems to me that:

- if some proposition is particularly relevant to a speaker...
- she will use that as the reference point for “yes”/“no”.

(10) A: I want you to tell me who didn’t come to the party.
    B: John didn’t come.
    a. A: ? Yes he did. / ? No, he didn’t. \(\leadsto\) relative to pos. prop.
    b. A: Yes, he didn’t. / No, he did. \(\leadsto\) relative to neg. prop.
5.2. Problem (i): context-dependence

It seems to me that:

- if some proposition is particularly relevant to a speaker...
- she will use that as the reference point for “yes” / “no”.

(10) A: I want you to tell me who didn’t come to the party.
    B: John didn’t come.
    a. A: ? Yes he did. / ? No, he didn’t. \(\sim\) relative to pos. prop.
    b. A: Yes, he didn’t. / No, he did. \(\sim\) relative to neg. prop.

(Disclaimer: my own judgement only.)
5.3. Problem (ii): results by Brasoveanu et al.

Median probabilities of a YES response & 95% CRIs

- Positive
- DE-obj
- DE-subj
- DE-adv
- N-word-obj
- N-word-subj
- N-word-adv
- Negative
5.3. Problem (ii): results by Brasoveanu et al.

Median probabilities of a YES response & 95% CRIs

John slept.
?? No, he did.

No one slept.
No, no one did.

John never slept.
No, he never did.

John didn't sleep.
No, he didn't.
5.3. Problem (ii): results by Brasoveanu et al.

Median probabilities of a YES response & 95% CRIs

John slept.  
?? No, he did.

John rarely slept.  
? No, he rarely did.

---

cond:  positive  DE-obj  DE-subj  DE-adv  N-word-obj  N-word-subj  N-word-adv  negative

medians: 0.0  0.2  0.4  0.6  0.8  1.0
5.3. Problem (ii): results by Brasoveanu et al.

Median probabilities of a YES response & 95% CRIs

John slept.

?? No, he did.

John rarely slept.

? No, he rarely did.

No one slept.

No, no one did.

No one slept.

No, no one did.
5.3. Problem (ii): results by Brasoveanu et al.

Median probabilities of a YES response & 95% CRIs

- John slept.
  - ?? No, he did.
- John rarely slept.
  - ? No, he rarely did.
- No one slept.
  - No, no one did.
- John never slept.
  - No, he never did.

cond:
- positive
- DE-obj
- DE-subj
- DE-adv
- N-word-obj
- N-word-subj
- N-word-adv
- negative
5.3. Problem (ii): results by Brasoveanu et al.
5.4. Accounting for the data

To explain the data, Krifka might say that constructions vary in:

- *how salient they make the positive proposition.*

But this is *ad hoc.*
5.4. Accounting for the data

To explain the data, Krifka might say that constructions vary in:
  ▶ how salient they make the positive proposition.

But this is ad hoc.

Different proposal (familiar strategy)

  (i) which propositions are salient is primarily contextual;
5.4. Accounting for the data

To explain the data, Krifka might say that constructions vary in:

- *how salient they make the positive proposition.*

But this is *ad hoc.*

Different proposal (familiar strategy)

(i) which propositions are salient is primarily contextual;
(ii) Brasoveanu’s experiment had *underspecified context.*
5.4. Accounting for the data

To explain the data, Krifka might say that constructions vary in:
- *how salient they make the positive proposition.*

But this is *ad hoc.*

Different proposal (familiar strategy)

(i) which propositions are salient is primarily *contextual*;
(ii) Brasoveanu’s experiment had *underspecified context*;
(iii) variance due to... typical use!
5.4. Accounting for the data

To explain the data, Krifka might say that constructions vary in:

- *how salient they make the positive proposition.*

But this is *ad hoc.*

**Different proposal (familiar strategy)**

(i) which propositions are salient is primarily *contextual*;
(ii) Brasoveanu’s experiment had *underspecified context*;
(iii) variance due to... typical use!

In particular, let’s assume the constructions vary in:

- *how often they are used in response to their negation:*
  ‘positive’ sentences < DEQ < N-words < negated sentences
5.4. Accounting for the data

To explain the data, Krifka might say that constructions vary in:

- *how salient they make the positive proposition.*

But this is *ad hoc.*

Different proposal (familiar strategy)

(i) which propositions are salient is primarily *contextual*;

(ii) Brasoveanu’s experiment had *underspecified context*;

(iii) variance due to... typical use!

In particular, let’s assume the constructions vary in:

- *how often they are used in response to their negation:*
  
  ‘positive’ sentences $<$ DEQ $<$ N-words $<$ negated sentences

This might be found in a corpus,
5.4. Accounting for the data

To explain the data, Krifka might say that constructions vary in:

- how salient they make the positive proposition.

But this is *ad hoc*.

Different proposal (familiar strategy)

(i) which propositions are salient is primarily contextual;
(ii) Brasoveanu’s experiment had underspecified context;
(iii) variance due to... typical use!

In particular, let’s assume the constructions vary in:

- how often they are used in response to their negation:
  ‘positive’ sentences < DEQ < N-words < negated sentences

This might be found in a corpus, but for now a conceptual reason:

- we are primarily interested in *what there is*;
- *what there isn’t* is typically only indirectly relevant.
5.4. Accounting for the data

To explain the data, Krifka might say that constructions vary in:

- how salient they make the positive proposition.

But this is *ad hoc*.

Different proposal (familiar strategy)

(i) which propositions are salient is primarily contextual;
(ii) Brasoveanu’s experiment had underspecified context;
(iii) variance due to... typical use!

In particular, let’s assume the constructions vary in:

- how often they are used in response to their negation:
  ‘positive’ sentences < DEQ < N-words < negated sentences

This might be found in a corpus, but for now a conceptual reason:

- we are primarily interested in what there is;
- what there isn’t is typically only indirectly relevant.

(Again, this is more a sociological than a linguistic issue.)
5.5. Conclusion (of this part)

In sum, for “yes”/“no”-licencing:

- underspecification and typical use may be to blame;
- the hypothesized use patterns are conceptually plausible;
- but they should of course be tested, e.g., on a corpus.
Outline

Ignorance implicatures and scalar modifiers
   The puzzle
   Experiment design
   Results and discussion

Exhaustivity inferences

Why conversational implicatures may well be strong

"Yes" and "no"

Conclusion
6.1. General conclusion

I have tried to reduce three puzzles to an interaction between:

- contextual underspecification; and
- typical use.
6.1. General conclusion

I have tried to *reduce* three puzzles to an interaction between:

- contextual underspecification; and
- typical use.

Why is this a ‘reduction’?

- it is a unifying account of three phenomena;
- it potentially simplifies the job left for semantics/pragmatics;
- (leaving typical use for sociology/psychology to explain).
6.1. General conclusion

I have tried to *reduce* three puzzles to an interaction between:

- contextual underspecification; and
- typical use.

**Why is this a ‘reduction’?**

- it is a unifying account of three phenomena;
- it potentially simplifies the job left for semantics/pragmatics;
- (leaving typical use for sociology/psychology to explain).

**Methodological gain**

- typical use can be independently measured (e.g., in a corpus);
- hence *factored out* when interpreting exp. data;
- or, better yet, its influence can be avoided altogether.
Thank you for your attention!