Processing Implicatures: a comparison between direct vs. indirect SIs

Experiments in Linguistic Meaning, ELM-2020

P. Marty\textsuperscript{1} J. Romoli\textsuperscript{2} Y. Sudo\textsuperscript{1} B. van Tiel\textsuperscript{3} R. Breheny\textsuperscript{1}

\textsuperscript{1}University College London

\textsuperscript{2}University of Bergen

\textsuperscript{3}Leibniz-Zentrum Allgemeine Sprachwissenschaft
Two effects commonly found in verification tasks for the pragmatic interpretation of scalar sentences with ‘some’ (i.e., ‘some but not all’):
Two effects commonly found in verification tasks for the pragmatic interpretation of scalar sentences with ‘some’ (i.e., ‘some but not all’):

- **Delay effect**
  Pragmatic responses take more time than literal ones
  
  (a.o., Bott et al., 2012; Bott & Noveck, 2004; Cremers & Chemla, 2014)

  (see Chevallier et al., 2008, for similar findings for ‘or’)

Two effects commonly found in verification tasks for the pragmatic interpretation of scalar sentences with ‘some’ (i.e., ‘some but not all’):

- **Delay effect**
  Pragmatic responses take more time than literal ones
  (a.o., Bott et al., 2012; Bott & Noveck, 2004; Cremers & Chemla, 2014)
  (see Chevallier et al., 2008, for similar findings for ‘or’)

- **Memory effect**
  Pragmatic responses decrease under higher memory load
Common interpretation
Common interpretation

1. The literal interpretation of ‘some’ is accessed first and then enriched via scalar reasoning.
Common interpretation

1. The literal interpretation of ‘some’ is accessed first and then enriched via scalar reasoning.

2. This scalar enrichment can be effortful: it may take extra memory resources and/or processing time.
Common interpretation *and further questions*

1. The literal interpretation of ‘*some*’ is accessed first and then enriched via scalar reasoning.

   ✉️ *Do these effects generalize to other scalar terms?*

2. This scalar enrichment *can* be effortful: it *may* take extra memory resources and/or processing time.

   ✉️ *What is the source of these costs?*
Common interpretation *and further questions*

1. The literal interpretation of ‘some’ is accessed first and then enriched via scalar reasoning.
   
   Do these effects generalize to other scalar terms?

2. This scalar enrichment can be effortful: it may take extra memory resources and/or processing time.

   What is the source of these costs?

Compare Direct vs. Indirect SIs to inform both questions and test a recent proposal, the Scallarity Hypothesis.
The Scalarity Hypothesis
Scales and scalability: Processing scalar inferences

Bob van Tiel*, Elizabeth Pankratz, Chao Sun

Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS), Germany

ARTICLE INFO

Keywords:
Scalar inference
Pragmatics
Sentence processing
Working memory
Conversational implicature
Language

ABSTRACT

The scalar word ‘some’ may be interpreted with an upper bound, i.e., as excluding ‘all’. Several studies have found that the computation of this scalar inference may be associated with a processing cost (e.g., Bott & Noveck, 2004; De Neys & Schaeken, 2007), which seems to argue in favour of theories according to which pragmatic inferencing is cognitively demanding (e.g., Sperber & Wilson, 1986). This argument holds on the premise that findings for ‘some’ can be generalised across the entire family of scalar words, which has been called into question by recent work highlighting the diversity within the class of scalar words (e.g., van Tiel, van Miltenburg, Zvakhina, & Geurts, 2016). In order to determine how generalisable the findings for ‘some’ are, we conducted three experiments in which we investigated the cognitive processing of seven scalar words that differ, inter alia, in their scalability, i.e., whether they impose a lower (‘some’, ‘or’, ‘might’, ‘most’, ‘try’) or upper (‘low’, ‘scarcely’) bound on their dimension. We find that the scalar inferences of the negatively scalar words ‘low’ and ‘scarce’ are not associated with a processing cost, unlike the scalar inferences of positively scalar words. We argue that the reported processing cost for scalar inferencing reflects increased cognitive demands associated with the processing of negative information.
An investigation of the cognitive processing of seven scalar words differing, inter alia, in their scalarity:
An investigation of the cognitive processing of seven scalar words differing, inter alia, in their scalarity:

Positively scalar terms
<some, all>, <most, all>, <or, and>, <might, must>, <try, succeed>
i.e., scalar terms with a literal lower-bounded meaning
An investigation of the cognitive processing of seven scalar words differing, inter alia, in their scalarity:

**Positively scalar terms**

<some, all>, <most, all>, <or, and>, <might, must>, <try, succeed>

i.e., scalar terms with a literal lower-bounded meaning

**Negatively scalar terms**

<low, empty>, <scarce, absent>

i.e., scalar terms with a literal upper-bound meaning
An investigation of the cognitive processing of seven scalar words differing, inter alia, in their scalarity:

**Positively scalar terms**
<some, all>, <most, all>, <or, and>, <might, must>, <try, succeed>
i.e., scalar terms with a literal lower-bounded meaning

**Negatively scalar terms**
<low, empty>, <scarce, absent>
i.e., scalar terms with a literal upper-bound meaning

For the positive scalars, the corresponding SI introduces an upper bound while, for the negative ones, it introduces a lower bound.
<table>
<thead>
<tr>
<th>Sentence</th>
<th>Control (T)</th>
<th>Control (F)</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>The battery is low.</td>
<td>![Battery Low]</td>
<td>![Battery Full]</td>
<td>![Battery Low]</td>
</tr>
<tr>
<td>Red flowers are scarce.</td>
<td>![Flower Images]</td>
<td>![Flower Images]</td>
<td>![Flower Images]</td>
</tr>
<tr>
<td>Either the apple or the pepper is red.</td>
<td>![Apple Pepper Images]</td>
<td>![Apple Pepper Images]</td>
<td>![Apple Pepper Images]</td>
</tr>
<tr>
<td>The arrow might land on red.</td>
<td>![Arrow Images]</td>
<td>![Arrow Images]</td>
<td>![Arrow Images]</td>
</tr>
<tr>
<td>Some of the socks are pink.</td>
<td>![Socks Images]</td>
<td>![Socks Images]</td>
<td>![Socks Images]</td>
</tr>
<tr>
<td>Most of the apples are green.</td>
<td>![Apple Images]</td>
<td>![Apple Images]</td>
<td>![Apple Images]</td>
</tr>
</tbody>
</table>
van Tiel et al. (2019): scalarity matters

The MEMORY and DELAY effects generalize to other positively scalar terms, but not to negatively scalar terms (i.e., scarce, low):

- rates of pragmatic responses unaffected by memory load
- pragmatic responses tend be faster than literal responses

---

van Tiel et al. (2019) suggest that these reverse delay effects, present in only some of their experiments, may be incidental rather than systematic.
The Scalarity Hypothesis (SH)
SIAs are cognitively demanding insofar as they introduce an upper bound on the dimension over which the scalar word quantifies.
The Scalarity Hypothesis (SH)
SIs are cognitively demanding insofar as they introduce an upper bound on the dimension over which the scalar word quantifies.

Rationale of the hypothesis

- SIs of positively scalar words introduce negative information
van Tiel et al. (2019): the Scalarity Hypothesis

The Scalarity Hypothesis (SH)
SIs are cognitively demanding insofar as they introduce an upper bound on the dimension over which the scalar word quantifies.

Rationale of the hypothesis

• SIs of positively scalar words introduce negative information
• Processing negative information is cognitively effortful
  (Clark & Chase, 1972; Deschamps et al., 2015; Geurts et al., 2010)
The Scality Hyspothesis (SH)
SIs are cognitively demanding insofar as they introduce an upper bound on the dimension over which the scalar word quantifies.

Rationale of the hypothesis

- SIs of positively scalar words introduce negative information
- Processing negative information is cognitively effortful
  (Clark & Chase, 1972; Deschamps et al., 2015; Geurts et al., 2010)
- MEMORY and DELAY effects induced by difficulties to process the negative information added by upper-bounding SIs
The Scarity Hypothesis (SH)
SIs are cognitively demanding insofar as they introduce an upper bound on the dimension over which the scalar word quantifies.

Rationale of the hypothesis

- SIs of positively scalar words introduce negative information
- Processing negative information is cognitively effortful
  (Clark & Chase, 1972; Deschamps et al., 2015; Geurts et al., 2010)
- MEMORY and DELAY effects induced by difficulties to process the negative information added by upper-bounding SIs

A novel proposal that deserves to be tested more extensively
The Present Study
Contribute to document scalar diversity while offering a more direct test of **SH** by comparing the processing of Direct vs. Indirect SIs.
Contribute to document scalar diversity while offering a more direct test of \textbf{SH} by comparing the processing of Direct vs. Indirect SIs.

**Direct SIs**

John ate \textit{some} of the pie. \hspace{1cm} <\textit{some, all}>

$\sim$ \textit{John ate some of the pie, but not all of it}$\quad$
Contribute to document scalar diversity while offering a more direct test of **SH** by comparing the processing of Direct vs. Indirect SIs.

**Direct SIs**

John ate *some* of the pie.  \(<\text{some, all}>\)

\(\leadsto\) *John ate some of the pie, but not all of it*

**Indirect SIs**

John did *not* eat *all* of the pie.  \(<\text{some, all}>\)

\(\leadsto\) *John didn’t eat all of the pie, but some of it*
Contribute to document scalar diversity while offering a more direct test of **SH** by comparing the processing of Direct vs. Indirect SIs.

**Direct SIs**

John ate *some* of the pie.  

\[ \sim \text{John ate some of the pie, but not all of it} \]

**Indirect SIs**

John did *not* eat *all* of the pie.  

\[ \sim \text{John didn’t eat all of the pie, but some of it} \]

For the weak terms, the corresponding SI introduces an upper bound while, for the (negated) strong ones, it introduces a lower bound.
Direct vs. Indirect SIs: a direct test of SH

*The Scalarity Hypothesis (SH)*

SIs are cognitively demanding insofar as they introduce an upper bound on the dimension over which the scalar word quantifies.
Direct vs. Indirect SIs: a direct test of SH

**The Scalarity Hypothesis (SH)**
SIs are cognitively demanding insofar as they introduce an upper bound on the dimension over which the scalar word quantifies.

**Rationale of our test**

- Pairs of scalar terms that focus on the same dimension and come from the same parts of speech (e.g., <some, all>)}
Direct vs. Indirect SIs: a direct test of SH

*The Scalarity Hypothesis (SH)*

SIs are cognitively demanding insofar as they introduce an upper bound on the dimension over which the scalar word quantifies.

**Rationale of our test**

- Pairs of scalar terms that focus on the same dimension and come from the same parts of speech (e.g., `<some, all>`)  
- Crucially, Indirect SIs introduce a lower bound, unlike Direct SIs
The Scality Hypothesis (SH)

SIs are cognitively demanding insofar as they introduce an upper bound on the dimension over which the scalar word quantifies.

Rationale of our test

• Pairs of scalar terms that focus on the same dimension and come from the same parts of speech (e.g., <some, all>)
• Crucially, Indirect SIs introduce a lower bound, unlike Direct SIs
• According to SH, Indirect SIs should thus pattern with the lower-bounding SIs arising from negatively scalar words
Direct vs. Indirect SIs: a direct test of SH

The Scalarity Hypothesis (SH)
SIs are cognitively demanding insofar as they introduce an upper bound on the dimension over which the scalar word quantifies.

Rationale of our test

- Pairs of scalar terms that focus on the same dimension and come from the same parts of speech (e.g., <some, all>)
- Crucially, Indirect SIs introduce a lower bound, unlike Direct SIs
- According to SH, Indirect SIs should thus pattern with the lower-bounding SIs arising from negatively scalar words

Prediction from SH: no MEMORY OR DELAY effects for Indirect SIs
The Experiment
Materials & Method

Materials and method built upon van Tiel et al. (2019)’s studies$^2$

$^2$van Tiel et al. (2019) JML paper and van Tiel et al. (2019) SuB proceedings
Materials & Method

Materials and method built upon van Tiel et al. (2019)'s studies

- a sentence-picture verification task using the same pictures as in van Tiel et al. (2019) but novel sentences

---

²van Tiel et al. (2019) JML paper and van Tiel et al. (2019) SuB proceedings
Materials and method built upon van Tiel et al. (2019)’s studies\(^2\)

- a sentence-picture verification task using the same pictures as in van Tiel et al. (2019) but novel sentences
- manipulation of the cognitive load on participants’ memory during the verification task

\(^2\)van Tiel et al. (2019) JML paper and van Tiel et al. (2019) SuB proceedings
## Direct SIs: example items

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Crl-True</th>
<th>Ctrl-False</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Some of the apples are green.</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
<tr>
<td>Either the apple or the pepper is red.</td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
<tr>
<td>It is possible that the arrow will land on red.</td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
<td><img src="image9.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Scalar terms: `<some, all>, <or, and>, <possible, certain>`
## Indirect SIs: example items

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Crl-True</th>
<th>Ctrl-False</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not all of the apples are green.</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
<tr>
<td>Not both the apple and the pepper is red.</td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
<tr>
<td>It is not certain that the arrow will land on red.</td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
<td><img src="image9.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Scalar terms: `<some, all>, <or, and>, <possible, certain>`
Further controls: ‘Only some’ sentences

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Ctrl-True</th>
<th>Ctrl-False</th>
</tr>
</thead>
<tbody>
<tr>
<td>Only some of the apples are green.</td>
<td>![Image of apples with some green]</td>
<td>![Image of apples with all green]</td>
</tr>
</tbody>
</table>
Further controls: ‘Only some’ sentences

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Ctrl-True</th>
<th>Ctrl-False</th>
</tr>
</thead>
<tbody>
<tr>
<td>Only some of the apples are green.</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Their literal meaning involves similar steps as the pragmatic meaning of *some* sentences (e.g., generate and negate alternatives).
Further controls: ‘Only some’ sentences

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Ctrl-True</th>
<th>Ctrl-False</th>
</tr>
</thead>
<tbody>
<tr>
<td>Only some of the apples are green.</td>
<td><img src="https://via.placeholder.com/50" alt="Image" /></td>
<td><img src="https://via.placeholder.com/50" alt="Image" /></td>
</tr>
</tbody>
</table>

Proportion of ‘True’ responses: \( \geq 98\% \)
\( \leq 6\% \)

Their literal meaning involves similar steps as the pragmatic meaning of *some* sentences (e.g., generate and negate alternatives).

**Replication of Marty & Chemla (2013):** responses to these sentences unaffected by adding or increasing memory load
NO LOAD condition
Some of the apples are red.

1: Good
0: Not good
LOW LOAD condition
Some of the apples are red.

1: Good
0: Not good
Please recreate the grid that you saw at the start of the trial.

→ Click here to continue
HIGH LOAD condition
Some of the apples are red.

1: Good
0: Not good
Please recreate the grid that you saw at the start of the trial.
The design in a nutshell

- Scales, Pictures and Type of SIs manipulated within subjects
- 3 instances of each condition created by varying the pictures
- Memory load manipulated between subjects: memory condition pseudo-randomly assigned at the start of the study
Participants

- 150 native speakers of English recruited through Prolific
- 50 participants per memory condition
- Data from 3 participants in the HIGH LOAD condition were removed from analyses

---

3Their mean accuracy to the memory task was below the pre-established threshold.
Data analyses

Responses and response times to the verification task

1. Responses across memory conditions as a function of the Scale and the type of SI (Direct vs. Indirect)

2. Response times in the NO LOAD condition
Memory results
Memory results

Proportion of ‘True’ responses (with 95% CIs) for each scale by implicature type as a function of the memory load and picture condition.
Memory results

Proportion of ‘True’ responses (with 95% CIs) for each scale by implicature type as a function of the memory load and picture condition.
Memory results

Proportion of ‘True’ responses (with 95% CIs) for each scale by implicature type as a function of the memory load and picture condition.

Replication: fewer Direct SIs under cognitive load
Memory results

Proportion of ‘True’ responses (with 95% CIs) for each scale by implicature type as a function of the memory load and picture condition.
Memory results

Proportion of ‘True’ responses (with 95% CIs) for each scale by implicature type as a function of the memory load and picture condition.

*Similar patterns of results for Direct and Indirect SIs*
Memory results: discussion

*MEMORY effects for both Direct and Indirect SIs*
Memory results: discussion

**MEMORY effects for both Direct and Indirect SIs**

Challenging for **SH** and for the idea that the polarity of an SI is the only or main explanation of the MEMORY effects.
Memory results: discussion

*MEMORY effects for both Direct and Indirect SIs*

Challenging for **SH** and for the idea that the polarity of an SI is the only or main explanation of the MEMORY effects.

Compatible with Marty & Chemla’s (2013) proposal:

- Memory resources are needed to compare and decide among competing interpretations
- When these resources are impaired, people become in effect more tolerant to under-informative statements
Memory results: discussion

*MEMORY effects for both Direct and Indirect SIs*

Challenging for **SH** and for the idea that the polarity of an SI is the only or main explanation of the MEMORY effects.

Compatible with Marty & Chemla’s (2013) proposal:

- Memory resources are needed to compare and decide among competing interpretations
- When these resources are impaired, people become in effect more tolerant to under-informative statements

*Do the RT results align with the memory results?*
Response time results
Response time results in NO LOAD

Mean logarithmised response times (with 95% CIs) for each scale by SI type as a function of the response type in the Control vs. Target conditions.
Response time results in NO LOAD

Mean logarithmised response times (with 95% CIs) for each scale by SI type as a function of the response type in the Control vs. Target conditions.
Response time results in NO LOAD

Mean logarithmised response times (with 95% CIs) for each scale by SI type as a function of the response type in the Control vs. Target conditions.

**Replication:** target ‘false’ > control ‘false’, target ‘false’ > target ‘true’, interactions between Condition & Response except for some-all
Response time results in NO LOAD

Mean logarithmised response times (with 95% CIs) for each scale by SI type as a function of the response type in the Control vs. Target conditions.
Response time results in NO LOAD

Mean logarithmised response times (with 95% CIs) for each scale by SI type as a function of the response type in the Control vs. Target conditions.

**Findings:** target ‘false’ \( \approx \) control ‘false’, target ‘false’ < target ‘true’, interactions between Condition & Response except for or-and
Response time results: discussion

Classical delay effects for Direct SIs, but not for Indirect SIs:

1. Pragmatic ‘false’ ≈ Control ‘false’

2. Literal ‘true’ > Pragmatic ‘false’ & Reverse DELAY effects
1. Pragmatic ‘false’ ≈ Control ‘false’
1. Pragmatic ‘false’ ≈ Control ‘false’

Compatible with SH and with the idea that lower bounding SIs are easier to process and faster to verify than upper bounding SIs.
Response time results: discussion

1. Pragmatic ‘false’ ≈ Control ‘false’

Compatible with **SH** and with the idea that lower bounding SIs are easier to process and faster to verify than upper bounding SIs.

Cremers & Chemla (2014), Bill et al. (2018): SIs may also affect participants’ response times in the control conditions

- Some evidence in our data that people derived more Indirect than Direct SIs (64% vs. 47% in NO LOAD targets)
- Possible confound in our RT controls more damaging for Indirect SIs than Direct SIs (possible follow-up: Training)
2. Literal ‘true’ > Pragmatic ‘false’ & Reverse DELAY effects
2. **Literal ‘true’ > Pragmatic ‘false’ & Reverse DELAY effects**

This is not an uncommon pattern of results:

- *not all* in Cremers & Chemla (2014, Experiment 1)
- *not always* in Romoli and Schwarz (2015)
- *low* and *scarce* in van Tiel et al. (2019)
Response time results: some speculations

“Not all of the apples are green”  Literal ‘true’ > Pragmatic ‘false’
Response time results: some speculations

“Not all of the apples are green”  Literal ‘true’ > Pragmatic ‘false’

Literal responders may have adopted a (time-consuming) two-step strategy to verify the literal meaning of these sentences:

Step 1  Evaluate the prejacent ALL OF THE APPLES ARE GREEN
Step 2  Reverse the response
Response time results: some speculations

“Not all of the apples are green”   Literal ‘true’ > Pragmatic ‘false’

Literal responders may have adopted a (time-consuming) two-step strategy to verify the literal meaning of these sentences:

Step 1  Evaluate the prejacent ALL OF THE APPLES ARE GREEN
Step 2  Reverse the response

Pragmatic responders considered the whole sentence to derive the SI and then looked for a truth-maker for the enriched meaning.
Response time results: some speculations

“Not all of the apples are green”  Literal ‘true’ > Pragmatic ‘false’

Literal responders may have adopted a (time-consuming) two-step strategy to verify the literal meaning of these sentences:

Step 1  Evaluate the prejacent ALL OF THE APPLES ARE GREEN
Step 2  Reverse the response

Pragmatic responders considered the whole sentence to derive the SI and then looked for a truth-maker for the enriched meaning.

*Remaining puzzle: Literal ‘true’ > Control ‘true’?*
Conclusion
The effects of Memory and RTs do not always pattern together: we found memory effects but no delay effects for Indirect SIs.
The effects of Memory and RTs do not always pattern together: we found memory effects but no delay effects for Indirect SIs.

This discrepancy could partly result from a confound in our RT controls (see Cremers & Chemla 2014, Bill et al. 2018).
The effects of Memory and RTs do not always pattern together: we found memory effects but no delay effects for Indirect SIs.

This discrepancy could partly result from a confound in our RT controls (see Cremers & Chemla 2014, Bill et al. 2018).

Yet it could suggest that both types of measures reflect very distinct cognitive effects in verification tasks.
The effects of Memory and RTs do not always pattern together: we found memory effects but no delay effects for Indirect SIs.

This discrepancy could partly result from a confound in our RT controls (see Cremers & Chemla 2014, Bill et al. 2018).

Yet it could suggest that both types of measures reflect very distinct cognitive effects in verification tasks.

<table>
<thead>
<tr>
<th>Memory effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relatively consistent across SIs, reflecting that WM resources are needed to decide to derive an SI or to decide among competing interpretations.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time course effects (delay and reverse delay)</th>
</tr>
</thead>
<tbody>
<tr>
<td>More variability across SIs, suggesting an interplay of different factors, among which the polarity of the base sentence and of its SI seem to appear.</td>
</tr>
</tbody>
</table>
Thank you!
Selected references


Selected references (cont.)


