Incremental Truth Value Judgments – A New Method To Measure Semantic Preferences During Comprehension

Oliver Bott & Fabian Schlotterbeck

SFB 833 – Construction of Meaning

12/02/2010
Outline

1. Measuring Preferences Online
2. Experiment 1: ITVJ Of Relative Quantifier Scope
3. Experiment 2: Task Demands of the ITVJ Method
4. Summary
We are interested in...

- Which interpretations are available when processing semantically ambiguous sentences?
- Do the readings differ in preference?
- What are the intermediate representations constructed during incremental interpretation?
- What is the final interpretation?
Offline Methods

**PRO**
- Are well suited to find out about the final interpretation(s)

**CON**
- Cannot tell us whether an interpretation was immediately available
- Cannot tell us whether comprehenders come up with this reading automatically
Online Methods

**PRO**
- Online data tell us which interpretation is harder to process

**CON**
- It’s hard to distinguish the dispreferred from the unavailable
A Novel Method – Incremental Truth Value Judgments

Goal

Combine the benefits of the *Picture Verification Task and online measures*

A sample Trial

1. A disambiguating picture is presented

   ▶ This adds a factor *disambiguating picture* with one level for each reading
A Novel Method – Incremental Truth Value Judgments

Goal

Combine the benefits of the *Picture Verification Task and online measures*

A sample Trial

- A disambiguating picture is presented

- This adds a factor *disambiguating picture* with one level for each reading
Incremental Truth Value Judgments – A Sample Trial

...like...

professors

students
Incremental Truth Value Judgments – A Sample Trial

The sentence is presented in a self-paced fashion.

For each segment, readers have to decide whether the sentence so far is compatible with the picture.

- **Dependent variables**
  - Rejection rates: is this a possible reading
  - RTs of ("yes") judgments: how difficult is this interpretation
Incremental Truth Value Judgments – A Sample Trial

2. The sentence is presented in a self-paced fashion.

3. For each segment, readers have to decide whether the sentence so far is compatible with the picture.

Dependent variables:
- Rejection rates: is this a possible reading?
- RTs of ("yes") judgments: how difficult is this interpretation?
Incremental Truth Value Judgments – A Sample Trial

2. The sentence is presented in a self-paced fashion.
3. For each segment, readers have to decide whether the sentence so far is compatible with the picture.

Dependent variables:
- Rejection rates: is this a possible reading?
- RTs of (“yes”) judgments: how difficult is this interpretation?
Incremental Truth Value Judgments – A Sample Trial

1. The sentence is presented in a self-paced fashion.

2. For each segment, readers have to decide whether the sentence so far is compatible with the picture.

- Dependent variables:
  - Rejection rates: is this a possible reading?
  - RTs of (“yes”) judgments: how difficult is this interpretation?
Each of the professors —— —— - — ——- —— ——.
Incremental Truth Value Judgments – A Sample Trial

- - - - - - - - - likes - - - - - - - - - - -.
Incremental Truth Value Judgments – A Sample Trial

none of the students
Remark

It has been shown that set diagrams are both a reliable and a valid mode of disambiguation to investigate scope preferences using a customary picture verification task (Bott & Radó, 2007).

Does this also carry over to its incremental version?

Exp.1) Using the ITVJ to measure preferences while interpreting relative scope of quantifiers

Exp.2) Potential problems of the ITVJ and how they can be avoided
Outline

1. Measuring Preferences Online
2. Experiment 1: ITVJ Of Relative Quantifier Scope
3. Experiment 2: Task Demands of the ITVJ Method
4. Summary
The Test Case: Scope Ambiguities

Scope preferences can be quite subtle, making the phenomenon an ideal test case for the ITVJ method!
Doubly quantified sentences like:

\[(2) \quad \text{Jeden dieser Schüler lobte genau ein Lehrer.} \]
\[\text{Each of these pupils}^{acc} \text{ praised exactly one teacher}^{nom}. \]
\[\text{Each of these pupils was praised by exactly one teacher.} \]

are ambiguous between two relative scope assignments:

\[(\forall \exists !) \quad \text{For each student holds: exactly one teacher praised him.} \]
\[(\exists ! \forall) \quad \text{For exactly one teacher holds: he praised every pupil.} \]
For each student holds: exactly one teacher praised him.
\((\exists! \forall)\) For exactly one teacher holds: he praised every pupil.
"Ambiguous" [+a] Constructions

Linear order and grammatical function of quantifiers influence the degree of ambiguity in the following German sentences.

(os/+a) Jeden dieser Schüler lobte genau ein Lehrer.  
  Each of these pupils$_O$ praised$_V$ exactly one teacher$_S$.

(so/+a) Genau ein Lehrer lobte jeden dieser Schüler.  
  Exactly one teacher$_S$ praised$_V$ each$_O$ of these pupils.

Conflicting factors as in (os/+a) should lead to

- Weaker preference for linear scope than in (so/+a)  
  (cf. Kurtzman & MacDonald 1993)
- Slower reading times than in (so/+a) (cf. Filik et al. 2004)
"Ambiguous" [+a] Constructions

Linear order and grammatical function of quantifiers influence the degree of ambiguity in the following German sentences.

(os/+a) Jeden dieser Schüler lobte genau ein Lehrer.
   Each of these pupils praised exactly one teacher.

(so/+a) Genau ein Lehrer lobte jeden dieser Schüler.
   Exactly one teacher praised each of these pupils.

Conflicting factors as in (os/+a) should lead to

- Weaker preference for linear scope than in (so/+a) (cf. Kurtzman & MacDonald 1993)
- Slower reading times than in (so/+a) (cf. Filik et al. 2004)
"Ambiguous" [+a] Constructions

Linear order and grammatical function of quantifiers influence the degree of ambiguity in the following German sentences.

(os/+a) Jeden dieser Schüler lobte genau ein Lehrer.
    Each of these pupils praised exactly one teacher.

(so/+a) Genau ein Lehrer lobte jeden dieser Schüler.
    Exactly one teacher praised each of these pupils.

Conflicting factors as in (os/+a) should lead to

- Weaker preference for linear scope than in (so/+a) (cf. Kurtzman & MacDonald 1993)
- Slower reading times than in (so/+a) (cf. Filik et al. 2004)
Are Both Constructions Really Ambiguous?

Compare OS and SO ambiguous sentences . . .

(os/+a) Jeden dieser Schüler lobte genau ein Lehrer voller Wohlwollen.

Each of these pupils praised exactly one teacher full of goodwill.

(so/+a) Genau ein Lehrer lobte jeden dieser Schüler voller Wohlwollen.

Exactly one teacher praised each of these pupils full of goodwill.
Are Both Constructions Really Ambiguous?

... with their disambiguated counterparts ...

(os/-a) Für jeden Schüler gilt: | ihn | lobte | genau ein Lehrer | ...  
For each pupil holds: | him | praised | exactly one teacher | ...

(so/-a) Für genau einen Lehrer gilt: | er | lobte | jeden Schüler | ...  
For exactly one teacher holds: | he | praised | each pupil | ...
Design

... yielding a $2 \times 2 \times 2$ design:

- Disambiguation picture: $\forall \exists!$ vs. $\exists! \forall$
- Ambiguity: *Ambiguous* vs. *Unambiguous*
- Order: *OS* vs. *SO*
Hypotheses and Predictions

Simultaneous Influence of linear order/ grammatical function

[os/+a] sentences are ”more” ambiguous than [so/+a] sentences from the second quantifier onwards.

- Ambiguous OS sentences differ from their unambiguous counterparts in that inverse scope is easily available
- More difference between ambiguous and unambiguous OS conditions than in SO
Simultaneous Influence of linear order/ grammatical function

[os/+a] sentences are ”more” ambiguous than [so/+a] sentences from the second quantifier onwards.

- Ambiguous OS sentences differ from their unambiguous counterparts in that inverse scope is easily available
- More difference between ambiguous and unambiguous OS conditions than in SO
Hypotheses and Predictions

Simultaneous Influence of linear order/ grammatical function

[os/+a] sentences are "more" ambiguous than [so/+a] sentences from the second quantifier onwards.

- Ambiguous OS sentences differ from their unambiguous counterparts in that inverse scope is easily available
- More difference between ambiguous and unambiguous OS conditions than in SO
Methods

- Incremental Truth Value Judgements
- Latin square design
- 32 items
- 73 fillers (41 true, 32 false)
- 40 participants
Aborted Trials in SO-Sentences

Both [+a] and [-a] sentences only compatible with linear scope
Aborted Trials in OS-Sentences

- Again, [+a] and [-a] show strong preference for linear scope
- Is inverse scope possible in [+a]?
Measuring Preferences Online
Experiment 1
Experiment 2
Summary

RT of “Yes” Responses in Linear Scope Conditions

RT difference only in [os/+a] vs. [os/-a] reflect scope conflict
Discussion

The ITVJ confirms the results of previous studies (eg. Filik et al. 2004)

- It allows us to distinguish a dispreferred from an unavailable reading during incremental interpretation
- It complements existing online paradigms
Discussion

- The ITVJ confirms the results of previous studies (e.g., Filik et al. 2004)
- It allows us to distinguish a dispreferred from an unavailable reading during incremental interpretation
- It complements existing online paradigms
The ITVJ confirms the results of previous studies (eg. Filik et al. 2004)

It allows us to distinguish a dispreferred from an unavailable reading during incremental interpretation

- It complements existing online paradigms
Each of these > exactly one

percent aborted trials (+95 % conf. interv.)

- ambig. / EA
- unambig. / EA
- ambig. / AE
- unambig. / AE

...verb 2nd qp spillover 1 spillover 2

A Concern
os/-a) Each pupil is such that he was praised by exactly one teacher.
What’s Going On?

- In a follow-up survey employing ordinary picture verification, the participants didn’t make these errors.
  - They must be due to the ITVJ task.
  - Enhanced processing load caused by:
    - Keeping the picture in memory.
    - Dual task during reading.

Is the ITVJ task too hard?
Outline

1. Measuring Preferences Online
2. Experiment 1: ITVJ Of Relative Quantifier Scope
3. Experiment 2: Task Demands of the ITVJ Method
4. Summary
Children’s Extra Object Error with $\forall \exists$

Is every boy riding an elephant?

Meroni (2002)
Pictures in Experiment 1 had two features that might have induced errors:

- extra objects (eo, eg. a non-praising teacher)
- branching lines (bl, eg. teachers praising more than one pupil)
Design – Extra Objects and Branching Lines (1)

os/-a) Each pupil is such that he was praised by exactly one teacher.

![Diagram showing relationships between teachers and pupils, with some pupils not being praised.]

Oliver Bott & Fabian Schlotterbeck
Incremental Truth Value Judgments
Each pupil is such that he was praised by exactly one teacher.
... yielding a $2 \times 2 \times 2$ design

- Disambiguating picture: $\exists! \forall \forall \exists!$
- Extra objects: $+eo$ vs. $-eo$
- Branching lines: $+bl$ vs. $-bl$
Prediction

Both extra objects and branching lines lead to higher error rates with ∀∃! Pictures
Methods

- Incremental Truth Value Judgements
- Latin square design
- 32 items ([os/-a] sentences from Exp. 1)
- 73 fillers (41 true, 32 false)
- 40 participants
∀∃! pictures:

- **Only 1% rejections in [-eo,-bl]!**
- 44% rejections in [+eo,+bl]
- This replicates Exp. 1
- Both eo and bl lead to errors

![Bar graph showing percent rejections for different conditions: -eo,-bl, +eo,-bl, -eo,+bl, +eo,+bl. The graph indicates a higher rejection rate for [+eo,+bl] compared to [-eo,-bl].]
∀∃! pictures:

- Only 1% rejections in [-eo,-bl]!
- 44% rejections in [+eo,+bl]

▷ This replicates Exp. 1
▷ Both eo and bl lead to errors
Results

∀∃! pictures:

- Only 1% rejections in [-eo,-bl]!
- 44% rejections in [+eo,+bl]

▷ This replicates Exp. 1

▷ Both eo and bl lead to errors
∀∃! pictures:

- Only 1% rejections in [-eo,-bl]!
- 44% rejections in [+eo,+bl]

- This replicates Exp. 1
- Both eo and bl lead to errors

![Graph showing percent rejections for different conditions](image-url)
In designing ITVJ experiments, materials have to be carefully controlled for processing demands. When this is done, the ITVJ task can provide completely unbiased data. Errors due to branching lines haven’t been attested in children yet.
In designing ITVJ experiments, materials have to be carefully controlled for processing demands.

When this is done, the ITVJ task can provide completely unbiased data.

Errors due to branching lines haven’t been attested in children yet.
Discussion

- In designing ITVJ experiments, materials have to be carefully controlled for processing demands
- When this is done, the ITVJ task can provide completely unbiased data
- Errors due to branching lines haven’t been attested in children yet
Discussion

Is every boy riding an elephant?
Outline

1. Measuring Preferences Online
2. Experiment 1: ITVJ Of Relative Quantifier Scope
3. Experiment 2: Task Demands of the ITVJ Method
4. Summary
ITVJs...

- measure availability and preferences of an interpretation in an incremental fashion
- provide complementary data for online experiments
- can be easily applied to...
  - anaphora resolution
  - scalar implicatures
  - binding
  - ...

Oliver Bott & Fabian Schlotterbeck
Incremental Truth Value Judgments
ITVJs ...

- reflect processing demands of different interpretations
- allow for controlling these demands (eg. via duration of picture presentation)

- can test resource based accounts in semantics
- can lead to new predictions for language acquisition
Thank you for your attention!

Thanks to Amelie Brinkmann and Anna Pryslopska for drawing hundreds of “bubbles” and running the experiments.
Error Rates Using Classical Picture Verification

**Judgments:**

![Bar chart showing percent correct for different conditions.]

- -eo, -bl
- +eo, -bl
- -eo, +bl
- +eo, +bl
- true fillers
- false fillers

**Judgment times:**

![Bar chart showing judgment time in ms for different conditions.]

- -eo, -bl
- +eo, -bl
- -eo, +bl
- +eo, +bl