The mechanics of causal discourse expectations

Oliver Bott\textsuperscript{1} & Torgrim Solstad\textsuperscript{2,3}

\textsuperscript{1}: SFB 833, University of Tübingen;
\textsuperscript{2}: NTNU Trondheim; \textsuperscript{3}: SFB 732, University of Stuttgart

DETEC2013
June 20, 2013
Overview

- Implicit causality basics
- Semantic theory of implicit causality
- Discourse continuation experiments
What is implicit causality?

- Implicit causality verbs trigger specific kinds of explanations
- IC bias may be viewed as an epi-phenomenon of explanatory strategies
- Association of explanation with only one of the two participants
- IC properties are well-known from theoretical semantics
- Our theory allows for a systematic manipulation of the bias behaviour of IC verbs
Why is IC interesting?

- Early research: IC dependent on presence of *because*
- Kehler et al.: Bias is identical across *because* and full stop explanation continuations

(1)  
   a. Mary disturbed John because she sang loudly.  
   b. Mary disturbed John. She sang loudly.

- From a lexical semantic point of view: What determines this behaviour?
- From a discourse semantic point of view: Expectations in cases of non-marked discourse relations
- Causality: Are certain explanations preferred over others?
IC bias from the perspective of explanatory discourse

Given a discourse segment $s_n$, which relation may be expected to be established between $s_n$ and the following discourse segment $s_{n+1}$?

(2) a. John disturbed Mary.
   b. John disturbed Mary. He was making lots of noise.

Cognitively plausible model of IC bias.

Insights into the bias

Offers possibility to manipulate not only the bias in terms of “pronominal continuations”

Systematic manipulation of expectations in the prompt.
Questions to be answered in a theory of IC bias

- How does the bias come about?
- Why do some verbs have a clear bias and others not?
- How does the expectation of an explanation come about?
- Why does there exist a correspondence between explanation patterns and coreference?
As far as we are concerned, semantic roles are not as such suited to explain this behaviour.

Causality of verb is not on its own enough: discrepancy – eventive and propositional causality
Basic ingredients

- IC bias is dependent on
  - Causal elaboration possibilities in \( NP1 \) verb-ed \( NP2 \)
  - semantic properties of \textit{because} (clauses)

- Consequently, we need a suitable theory of verb semantics and a typology of explanations
Causes and reasons

- Simple causes are causes of (attitudinal) states or events
  
  (3) **Simple (direct) cause:**
  John disturbed Mary because he sang loudly.

- Reasons are causes of attitudinal states involving intentionality
  
  (4) **Externally anchored reason:**
  John disturbed Mary because she had stolen his textbook.
  
  (5) **Internally anchored reason:**
  John disturbed Mary because he was angry at her.

- Backgrounds are necessary, but not sufficient causes
  
  (6) **Background:**
  John disturbed Mary because she needed silence to concentrate.
Elaboration possibilities I: Stimulus arguments

- *annoy* $\rightsquigarrow f$ CAUSE $s_{att}$
  Some fact brings about the attitudinal state of $y$ being annoyed

(7) **Peter annoyed Mary because he sang loudly.**

$$\lambda f. \lambda y. f \; CAUSE \; s_{att}$$
$$s_{att}: \text{annoyed}(y)$$

- evidence: stimulus arguments may also be realized by *that* clauses

(8) It annoyed Mary *that Peter sang loudly.*
 Ontological constraints

- *because* introduces factive causes, which are propositional in nature.
- *because* cannot specify the causing event of all causative predicates.

\[
\text{murder} \rightsquigarrow e \text{ CAUSE } s
\]

(9) \# *Tom ermordete Louisa, weil er sie in den Rücken stach.*
‘Tom murdered Louisa because he stabbed her in the back’.

- *e* is ‘associated with’ Tom, who is the agent of the causing event.
- Psych verbs: \( f \text{ CAUSE } S_{\text{att}} \)
- Simple causes with psych verbs, **not** with agent-patient verbs.
- Agent–patient verbs should not pattern with stimulus–experiencer or experiencer–stimulus predicates if elaboration of stimulus properties is the decisive source of bias behaviour.
Elaboration possibilities II: Presuppositions

- *because* may elaborate a presupposed preceding event

(10) **NP2 bias:**

\[ \text{Lars bestrafte Melanie weil sie das Geld geklaut hatte.} \]

‘Lars punished Melanie *because she took the money.*’

\[ \{ \begin{array}{c}
\text{e' z} \\
\text{e' <}_t \text{ e} \\
\text{z: AGENT(e')} \\
\text{z=y}
\end{array} \} \quad \begin{array}{c}
\{ \begin{array}{c}
\text{e x y} \\
\text{e: PUNISH(x,y)}
\end{array} \} \]

- agent of presupposed event identified with ‘bias argument’ of *punish*-event
- “causes precede their effects”
- explanations are external reasons
- we expect strong bias effects: **AVOID ACCOMMODATION**
Elaboration possibilities II: Presuppositions

- *because* may elaborate a presupposed preceding event

\[\begin{align*}
\text{NP2 bias:} & \\
Lars & \text{bestrafte Melanie, weil sie das Geld geklaut hatte.} \\
\text{‘Lars punished Melanie because she took the money.’}
\end{align*}\]

\[
\begin{cases}
\begin{align*}
e' &= e \\
z &= \text{AGENT}(e')
\end{align*}
\end{cases}
\quad \begin{align*}
\begin{array}{c}
\text{e' <}_t e \\
z = y
\end{array}
\end{align*}
\]
Elaboration possibilities II: Presuppositions

- *because* may elaborate a presupposed preceding event

(10) NP2 bias:

*Lars bestrafte Melanie, weil sie das Geld geklaut hatte.*

‘Lars punished Melanie because she took the money.’

\[
\begin{align*}
\left\{ 
\begin{array}{c}
\text{e'} z \\
\text{e'} <_t e \\
\text{z: AGENT(e')} \\
\text{z=y}
\end{array}
\right. \\
\end{align*}
\]

- agent of presupposed event identified with ‘bias argument’ of *punish*-event
- “causes precede their effects”

- explanations are external reasons

▷ we expect strong bias effects: AVOIDACCOMMODATION
Notion of mechanics is too strong
Focus on explanation types
Bias as epiphenomenon
Passives and other constructions may interfere between explanation type and co-reference realization pattern

(11)  
   a. Mary annoyed Peter with ... because
   b. STATIVE PASSIVE (or similar)
Manipulating the bias

- Missing information triggers systematic specification
- By specifying the missing information in the prompt, we should be able to influence the preferences for explanation
- Drop in explanations in full stop.
- Shift in explanation types.
- Correspondingly, the bias should be altered.
Stimulus-Experiencer verbs allow us to specify the stimulus by means of *durch* phrases.

Similar to *by+-ing* or *with* phrases

(12)  
   a.  
   b.  

(13)  
   a.  
   b.  

Specification of causing entity in causal relation
Modifying PSP verbs
Experimental study
Predictions

- In general, we predict:
- Verbs with slots should trigger more explanations than verbs without slots
- Verbs where the slot is filled in the matrix clause should trigger significantly less explanations.
Experiments 1A and 1B
Comparing Stimulus-Experiencer and Agent-Patient verbs
Pretest: *durch* (*by*) vs. *weil* (*because*)

What are the restrictions imposed by *weil* (*because*) and *durch* (*by*)?

(1) Mary (Peter) fascinated Peter (Mary) because…
(2) Mary (Peter) fascinated Peter (Mary) by…
(3) Mary (Peter) killed Peter (Mary) because…
(4) Mary (Peter) killed Peter (Mary) by…

- Continuation study: 48 participants ended German sentences involving 20 SE and 20 causative AP verbs
- $2 \times 2 \times 2$ within design (SE/AP verb $\times$ *durch*/*weil* $\times$ male/female)
- Continuations annotated for explanation type
- *weil* triggers all four explanation types
- *durch* only compatible with simple causes
- *weil* can only target simple causes of SE verbs, not AP verbs
- Evidence for proposed ontological restrictions
Pretest results: factive vs. simple causes

- *weil* triggers all four explanation types
- *durch* only compatible with simple causes
- *weil* can only target simple causes of SE verbs, not AP verbs

Evidence for proposed ontological restrictions

![Graph showing the proportion of explanation type in % for weil vs. durch with various explanation types: simple cause, external reason, internal reason, background. The graph compares the usage of weil and durch with 'because' and 'by' as connectors.]
Methods: Exp. 1A/B

- Between items/within participants:
  - Exp. 1a: 20 SE–verbs
    - SE) Ben (Mia) amused Mia (Ben) (by telling a funny story) a) full stop ... b) because ...
  - Exp. 1b: 20 causative AP–verbs
    - AP) Jasmin (Henry) killed Henry (Jasmin) (by a shot) a) full stop ... b) because ...

- Within items/participants: 2 × 2 design (connective × modification)
- Counterbalanced between NP1/2 = masc./fem.
- 40 items + 40 fillers, latin square design
- Instructions: “continue the discourse in the most natural way” (or skip it if you can't come up with a continuation)

- Two blocks:
  - 1st block: full stop conditions
  - 2nd block: weil conditions
The resulting corpus of continuations was annotated wrt. the following categories:

- **Bias annotation** (agreement: Cohen’s $\kappa = .86$)
  - NP1 vs. NP2 vs. no anaphor
  - unequivocal vs. ambiguous (*he loved her*)
  - kind of anaphor (pronoun, DP, proper name)

- **Discourse relation** (agreement: Cohen’s $\kappa = .80$)
  - sensible continuation (skips and nonsense $\bowtie$ mis. values)
  - temporal order (event\textsubscript{continuation} after event\textsubscript{prompt}?)
  - explanation? (*weil* test)

- **Causal annotation**: *simple cause vs. ext. reason vs. int. reason vs. background* (agreement: Cohen’s $\kappa = .74$)
The resulting corpus of continuations was annotated wrt. the following categories:

- **Bias annotation** (agreement: Cohen’s $\kappa = .86$)
  - NP1 vs. NP2 vs. no anaphor
  - unequivocal vs. ambiguous (*he loved her*)
  - kind of anaphor (pronoun, DP, proper name)

- **Discourse relation** (agreement: Cohen’s $\kappa = .80$)
  - sensible continuation (skips and nonsense $\triangleright$ mis. values)
  - temporal order (event$_{continuation}$ after event$_{prompt}$?)
  - explanation? (*weil* test)

- **Causal annotation**: simple cause vs. ext. reason vs. int. reason vs. background (agreement: Cohen’s $\kappa = .74$)
The resulting corpus of continuations was annotated wrt. the following categories:

- **Bias annotation** (agreement: Cohen’s $\kappa = .86$)
  - NP1 vs. NP2 vs. no anaphor
  - unequivocal vs. ambiguous (*he loved her*)
  - kind of anaphor (pronoun, DP, proper name)

- **Discourse relation** (agreement: Cohen’s $\kappa = .80$)
  - sensible continuation (skips and nonsense $\triangleright$ mis. values)
  - temporal order (event$_{continuation}$ after event$_{prompt}$?)
  - explanation? (*weil* test)

- **Causal annotation**: *simple cause* vs. *ext. reason* vs. *int. reason* vs. *background* (agreement: Cohen’s $\kappa = .74$)
For SE verbs, *durch*-PPs reduce explanations from 51% to 19%.

For AP verbs without underspecified content, less explanations overall (19% vs. 11%).

Stronger decrease in SE than in AP conditions (interaction: $LRCS_1(1) = 6.9; p < .01$).
Explanation types: predictions for SE verbs and AP verbs

- **SE verbs**
  - Drop in Simple causes
  - Increase in Internal reasons
  - Increase in Backgrounds

- **AP verbs**
  - no significant changes in explanation profile
SE and AP verbs: Explanation types

**SE verbs**

- **durch-PPs reduce simple causes** \((LRCS_{1/2}(1) \leq 266; p_{1/2} < .01)\)
- After *weil*, simple causes are substituted by int. reasons and backgrounds

**AP verbs**

- No sign. effects of *modification*
- Same explanation profiles across conditions
Bias: predictions for SE and AP verbs

SE verbs

- Explanations in unmodified conditions
  - Overwhelmingly simple causes
  - Strong NP1 bias
- Explanations after *durch*
  - no simple causes, internal reasons and backgrounds, instead
  - Bias shifts towards NP2

AP verbs

- Explanations in unmodified conditions
  - External and internal reasons
- Explanations after *durch*
  - Unchanged explanation profile
  - Same bias
Bias: predictions for SE and AP verbs

SE verbs

- Explanations in unmodified conditions
  - Overwhelmingly simple causes
  - Strong NP1 bias
- Explanations after *durch*
  - no simple causes, internal reasons and backgrounds, instead
  - Bias shifts towards NP2

AP verbs

- Explanations in unmodified conditions
  - External and internal reasons
  - Balanced bias
- Explanations after *durch*
  - Unchanged explanation profile
  - Same bias
SE and AP verbs: Bias

**Stimulus-experiencer verbs**

- *durch* phrases decrease the likelihood of rementioning NP1! \( (\text{LRCS}_{1/2}(1) \geq 72; p_{1/2} < .01) \)

**Agent-patient verbs**

- *durch* phrases increase the likelihood of rementioning NP1! \( (\text{LRCS}_{1/2}(1) \geq 3.9; p_{1/2} < .05) \)
SE and AP verbs: Bias

**Stimulus-experiencer verbs**

- *durch* phrases **decrease** the likelihood of rementioning NP1! \((LRCS_{1/2}(1) \geq 72; p_{1/2} < .01)\)

**Agent-patient verbs**

- *durch* phrases **increase** the likelihood of rementioning NP1! \((LRCS_{1/2}(1) \geq 3.9; p_{1/2} < .05)\)
In contrast to ordinary AP verbs, SE verbs in *name verbs name* sentences attract explanations.

SE verbs trigger simple causes that specify a property of the stimulus argument.

*durch* modification affects explanation types and bias in a fully predictable way.

But why should *durch* lead to a bias effect in AP verbs?

(1) John killed Mary by a shot. (= John’s shot)
(2) John amused Mary by telling a joke. (= John’s joke)

Across verb types, *durch*-phrases should enhance salience of NP1.

This influenced the specific form of the explanation sentence, but not its type!
In contrast to ordinary AP verbs, SE verbs in *name verbs name* sentences attract explanations.

SE verbs trigger simple causes that specify a property of the stimulus argument.

*durch* modification affects explanation types and bias in a fully predictable way.

But why should *durch* lead to a bias effect in AP verbs?

1. John killed Mary by a shot.  
   (= John’s shot)
2. John amused Mary by telling a joke.  
   (= John’s joke)

Across verb types, *durch*-phrases should enhance salience of NP1. This influenced the specific form of the explanation sentence, but not its type!
In contrast to ordinary AP verbs, SE verbs in *name verbs name* sentences attract explanations

SE verbs trigger simple causes that specify a property of the stimulus argument

*durch* modification affects explanation types and bias in a fully predictable way

But why should *durch* lead to a bias effect in AP verbs?

(1) John killed Mary by a shot. (= John’s shot)
(2) John amused Mary by telling a joke. (= John’s joke)

Across verb types, *durch*-phrases should enhance salience of NP1

This influenced the specific form of the explanation sentence, but not its type!
Experiment 2
Presupposition verbs
Methods: Exp. 2

- AP–verbs with a presupposition

  PSP) Sebastian (Jaqueline) congratulated Jaqueline (Sebastian) (on winning the first prize in the dancing competition) a) full stop... b) because...

- Within items/participants: $2 \times 2$ design ($connective \times modification$)

- Counterbalanced between NP1/2 = masc./fem.

- Run together with Exp. 1A/B, same methods

- Annotation as in Exp. 1A/B
Predictions for PSP verbs

1. Specifying missing content explicitly in the prompt should lead to a drop in the number of explanations.
2. Unmodified conditions should trigger external reasons.
   - Modified conditions should trigger internal reasons (or backgrounds).
3. This should alter IC bias from NP2 towards NP1.
After a full stop, explanations drop from 56% to 31%

\( LRCS_{1/2}^{(1)} \geq 31; p < .01 \)
PSP verbs: Explanation types

- Unmodified conditions trigger ext. reasons
- Modification leads to a shift towards int. reasons

\[(LRCS_{1/2}(1) \geq 193; p < .01)\]
Unmodified conditions are strongly biased towards NP2
Modification eliminates the bias

\( LRCS_{1/2}(1) \geq 93; p < .01 \)
Cross-linguistic validity
Our account only hinges on verbal semantics and possible explanations. We therefore expected IC bias and the underlying distribution of explanation types to be crosslinguistically stable.

We chose German and Norwegian because...

- verbs can be perfectly matched in meaning
- the two languages differ in discourse segmentation (Fabricius-Hansen et al. 2005, Ramm 2011)
  - German: hypotactic sentences, explicit discourse marking
  - Norwegian: paratactic sentences, less explicit marking
Our account only hinges on verbal semantics and possible explanations. We therefore expected IC bias and the underlying distribution of explanation types to be crosslinguistically stable.

- We chose German and Norwegian because...
  - verbs can be perfectly matched in meaning
  - the two languages differ in discourse segmentation (Fabricius-Hansen et al. 2005, Ramm 2011)
    - German: hypotactic sentences, explicit discourse marking
    - Norwegian: paratactic sentences, less explicit marking
Our account only hinges on verbal semantics and possible explanations. We therefore expected IC bias and the underlying distribution of explanation types to be crosslinguistically stable.

- We chose German and Norwegian because...
  - verbs can be perfectly matched in meaning
  - the two languages differ in discourse segmentation (Fabricius-Hansen et al. 2005, Ramm 2011)
    - German: hypotactic sentences, explicit discourse marking
    - Norwegian: paratactic sentences, less explicit marking
52 Germans and 48 Norwegians participated in a discourse continuation study.

Same methods and annotation as in Exp. 1A/B & Exp. 2

101 verbs: 16 se; 18 es; 10 ap/+psp; 57 ap/–psp

1. block full stop: NP1 verbed NP2. . . .
2. block because: NP1 verbed NP2, because . . .

10,100 productions
Verb-by-verb comparison

(14) NP1 verbed NP2 because he/she...

- German and Norwegian verbs closely corresponded in bias ($r = .92; p < .01$)
The ratio of int. to ext. reasons is a good predictor for IC-bias
(accounted variance: $R^2 = .75$)