

# **Processing polarity: Some experimental investigations**

Shravan Vasishth

Cogeti workshop, Tübingen, Feb. 5, 2006

## Acknowledgements

Thanks to Kathleen Raphael and Kai Sippel for running the experiments, and to Larry Horn and the audiences at CUNY 2005 (Tucson, Arizona), and the Polarity meets Psycholinguistics for comments and criticism.

This is collaborative work with Richard Lewis (Michigan), Heiner Drenhaus, Douglas Saddy, Tessa Warren (Pittsburgh), and Masako Hirotani (Leipzig).

## **What's a licensing context for NPIs? (And how does the NPI access this information?)**

In general, syntax, semantics, *and* pragmatics come together to determine if an NPI is licensed or not.

1. Semantic-logical properties: Horn (1997), Giannakidou (1998), Ladusaw (1980), Van der Wouden (1994).
2. Pragmatic properties: Chierchia (2001); Fauconnier (1980); Krifka (1995).
3. A combination of semantic and pragmatic properties: Baker (1970); Linebarger (1987).
4. . . .

In addition, these properties (whatever they are) of the licensing context must be **ACCESSIBLE** to the NPI. This accessibility is determined by hierarchical constituency (Haegeman 1995; Laka 1994; Progovac 2000).

## Syntactic/semantic constraints on German *jemals*, 'ever'

- (1) a. **Kein** Mann, [der einen Bart hatte,] war **jemals** glücklich  
No man who a beard had was ever happy  
'No man who had a beard was ever happy.'
- b. \***Ein** Mann, [der einen Bart hatte,] war **jemals** glücklich  
A man who a beard had was ever happy  
'A man who had a beard was ever happy.'
- c. \*Ein Mann, [der **keinen** Bart hatte,] war **jemals** glücklich  
A man who no beard had, was ever happy  
'A man who had no beard was ever happy.'

## A real-time processing investigation

In a speeded grammaticality judgement task, 24 subjects were shown sentences like (2), 8 sentences per condition and intermixed with 80 unrelated fillers.

(2) a. Accessible licensor

**Kein** Mann, [der einen Bart hatte,] war **jemals** glücklich  
No man who a beard had was ever happy  
'No man who had a beard was ever happy.'

b. No licensor

\***Ein** Mann, [der einen Bart hatte,] war **jemals** glücklich  
A man who a beard had was ever happy  
'A man who had a beard was ever happy.'

c. Inaccessible licensor

\*Ein Mann, [der **keinen** Bart hatte,] war **jemals** glücklich  
A man who no beard had, was ever happy  
'A man who had no beard was ever happy.'

## The intrusion effect

Condition	Accuracy (% correct)	Speed (msecs)
(2a) Accessible licensor	85	540
(2b) No licensor	83	554
(2c) Inaccessible licensor	70	712

1. (2c) was accuracy worse than in other conditions:

(2c) vs. (2a):  $F1(1,23) = 5.11, p < .05$ ;  $F2(1,23) = 8.89, p < .01$ .

(2c) vs. (2b):  $F1(1,23) = 6.11, p < .05$ ;  $F2(1,23) = 10.80, p < .01$ .

2. (2c) responses slower than in other conditions:

(2c) vs. (2a):  $F1(1,23) = 10.25, p < .01$ ;  $F2(1,23) = 8.35, p < .05$ .

(2c) vs. (2b):  $F1(1,23) = 26.68, p < .001$ ;  $F2(1,23) = 11.95, p < .01$ .

In sum, a linearly preceding but structurally inaccessible licensor sometimes ends up getting accessed; let's call it the INTRUSION EFFECT.

## A semantic integration problem appears to cause the intrusion effect

NPI licensing violations are known to trigger an N400, suggesting semantic integration problems (Saddy et al., in press).

In an ERP version of the speeded acceptability study, we replicated the preceding experiment's results and also found an N400 in both the no-licensor and inaccessible-licensor conditions:

(3) b. No licensor

\***Ein** Mann, [der einen Bart hatte,] war **jemals** glücklich

A man who a beard had was ever happy

'A man who had a beard was ever happy.'

c. Inaccessible licensor

\***Ein** Mann, [der **keinen** Bart hatte,] war **jemals** glücklich

A man who no beard had, was ever happy

'A man who had no beard was ever happy.'

# **Theoretical background: A computational model of sentence processing**

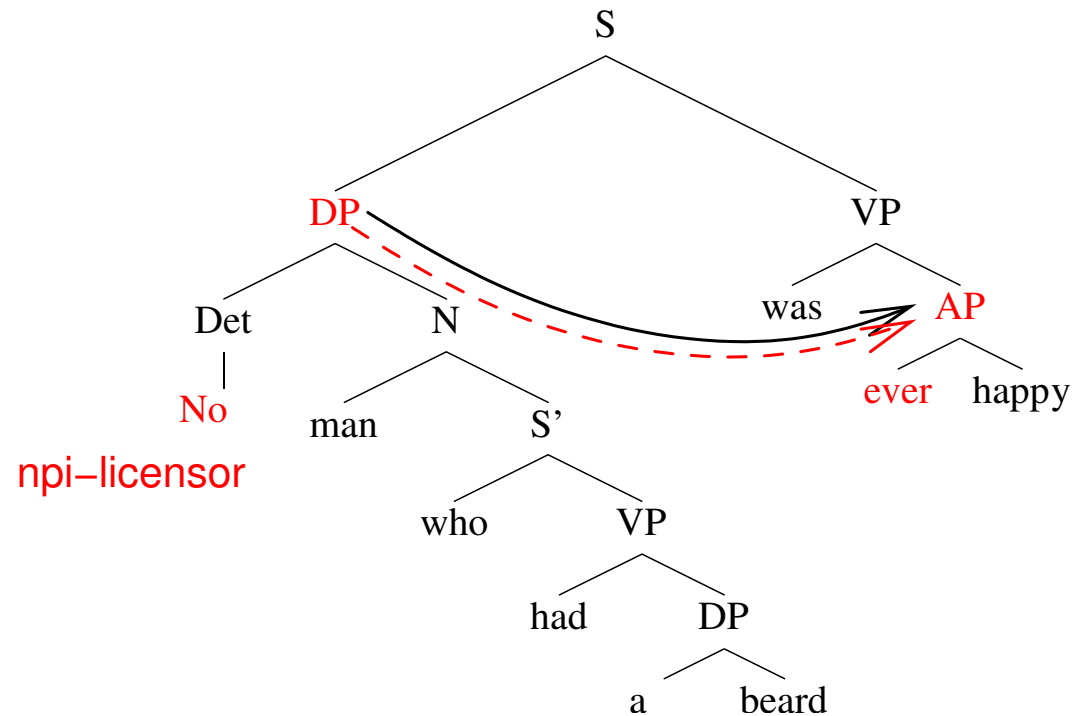
Basic assumptions (elevator version):

- Cue-based retrieval
- Interference
- Decay and reactivation

The model is fully implemented and the associated papers are available from my web page.



**When licenser is present and is in correct location:  
An additional semantic constraint boosts activation of subject DP**



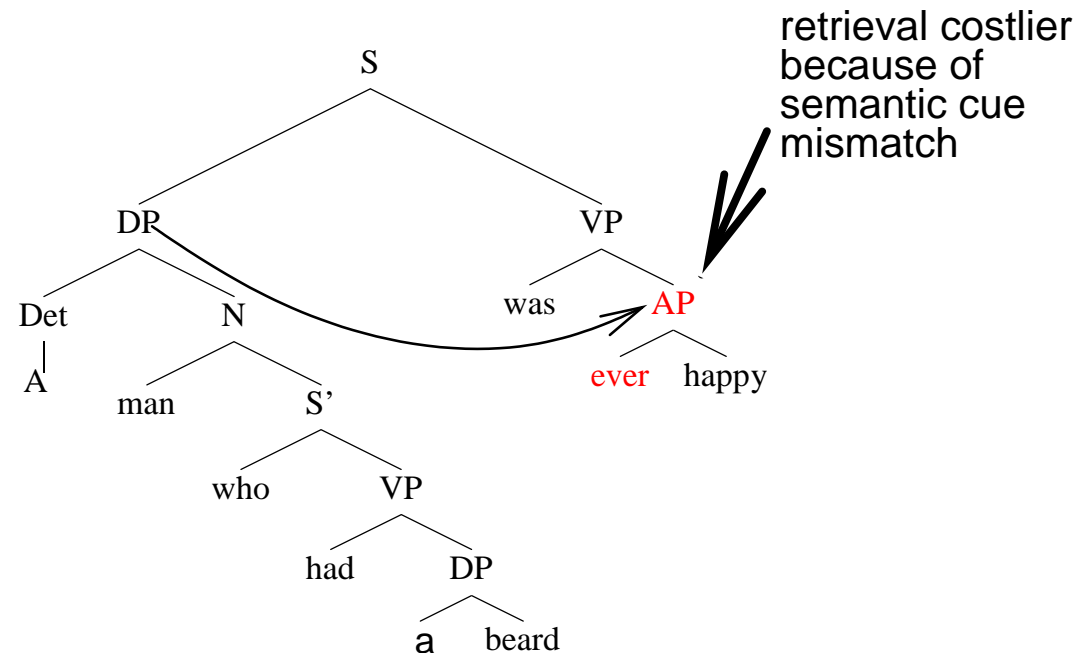
(Syntactic-semantic) retrieval cue # 1: retrieve subject of main predicate

MATCH

(Semantic) retrieval cue # 2: retrieve an NPI-licensor

MATCH

## When no licenser is present



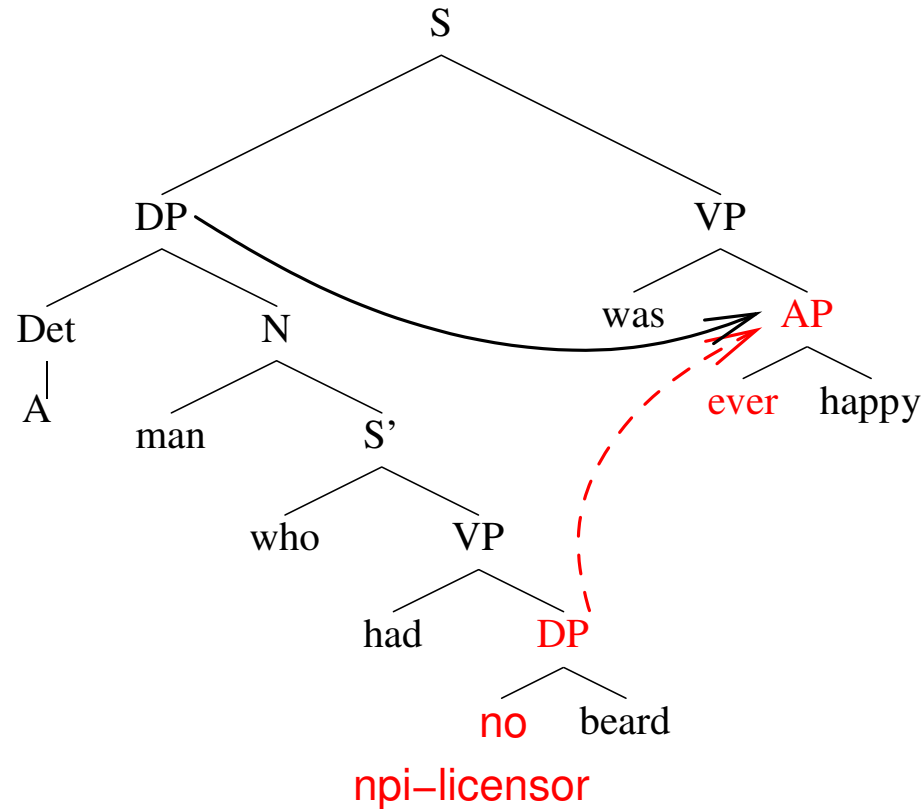
(Syntactic-semantic) retrieval cue # 1: retrieve subject of main predicate

MATCH

(Semantic) retrieval cue # 2: retrieve an NPI-licensor

MISMATCH

## When the licenser is present but in the wrong structural location



(Syntactic-semantic) retrieval cue # 1: retrieve subject of main predicate      MATCH  
(Semantic) retrieval cue # 1: retrieve an NPI-licensor      MATCH WITH EMBEDDED DP

## Modeling percentage of correct judgements: Results of Monte Carlo simulations (50 runs)

Condition	Data	Model
(2a) Accessible licensor	85	96
(2b) No licensor	83	96
(2c) Inaccessible licensor	70	68

## Some open issues

- Perhaps the effects observed are an artefact of the speeded judgement task—a relatively unnatural task for sentence processing. It's important to establish that the cue-based retrieval explanation works in more natural comprehension settings.
- If cue-based retrieval has any validity, it should generalize beyond the NPI data to other phenomena that involve licensors. An example is positive polarity items.

## Positive polarity items or PPIs

These have the curious property that they are allergic to NPI-licensors.

- (4) a. \***Kein** Mann, [der einen Bart hatte,] war **durchaus** glücklich  
No man who a beard had was certainly happy  
'No man who had a beard was certainly happy.'
- b. **Ein** Mann, [der **einen** Bart hatte,] war **durchaus** glücklich  
A man who a beard had was certainly happy  
'A man who had a beard was certainly happy.'
- c. Ein Mann, [der **keinen** Bart hatte,] war **durchaus** glücklich  
A man who no beard had, was certainly happy  
'A man who had no beard was certainly happy.'

## Some assumptions about what a PPI is and does

A simple way to implement the anti-licensing constraint of PPIs is to assume that actually *looks* for an NPI licensor and raises an error signal if there is such a licensor present.

A good reason for taking this approach:

Szabolcsi (2004) has proposed (inter alia) that PPIs have NPI features that “lie dormant” and are “activated” by the NPI licensor.

## Eyetracking study of NPI and PPI processing





## **Eyetracking study of NPI and PPI processing**

### Method:

The three NPI and three PPI conditions were presented in counterbalanced manner to 48 subjects ( $2 \times 3$  factorial design). There were four items per condition (so each subject saw 24 critical items).

Subjects are asked to read sentences on a computer screen and an eyetracker records their eye movements and fixations.

First pass reading time (FPRT): The time spent in a region after it is first entered and before it is exited. Reflects early processing (e.g. lexical retrieval, and immediately following events).

Total reading time (TRT): The sum of all fixations in a region.

## Predictions for NPIs

(5) b. No licenser

\***Ein** Mann, [der einen Bart hatte,] war **jemals** glücklich  
A man who a beard had was ever happy  
'A man who had a beard was ever happy.'

c. Inaccessible licenser

\***Ein** Mann, [der **keinen** Bart hatte,] war **jemals** glücklich  
A man who no beard had, was ever happy  
'A man who had no beard was ever happy.'

- Legal licensers would be rapidly retrieved
- Intrusive licensers would be harder to process due to the mismatch penalty
- The no-licenser condition should be hardest to retrieve.

## Predictions for PPIs

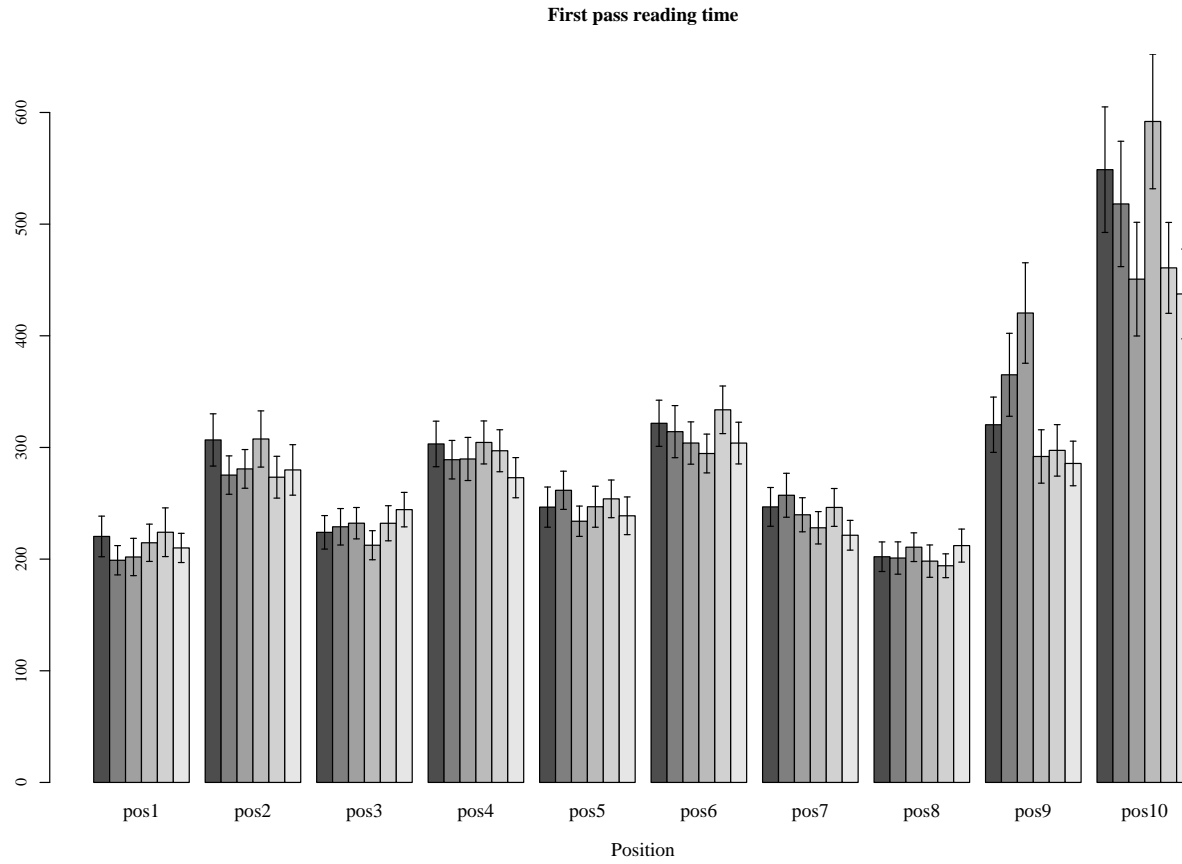
- (6) a. \***Kein** Mann, [der einen Bart hatte,] war **durchaus** glücklich  
No man who a beard had was certainly happy  
'No man who had a beard was certainly happy.'
- b. **Ein** Mann, [der **einen** Bart hatte,] war **durchaus** glücklich  
A man who a beard had was certainly happy  
'A man who had a beard was certainly happy.'
- c. Ein Mann, [der **keinen** Bart hatte,] war **durchaus** glücklich  
A man who no beard had, was certainly happy  
'A man who had no beard was certainly happy.'

- In the legal-NPI-licensor condition processing would be slow at the PPI since an error would immediately be raised.
- In the intrusive-NPI-licensor condition processing should be faster the legal-NPI licensor condition, but slower than the no-licensor condition (due to errorful retrievals).
- In the no-NPI-licensor condition processing would be fastest.

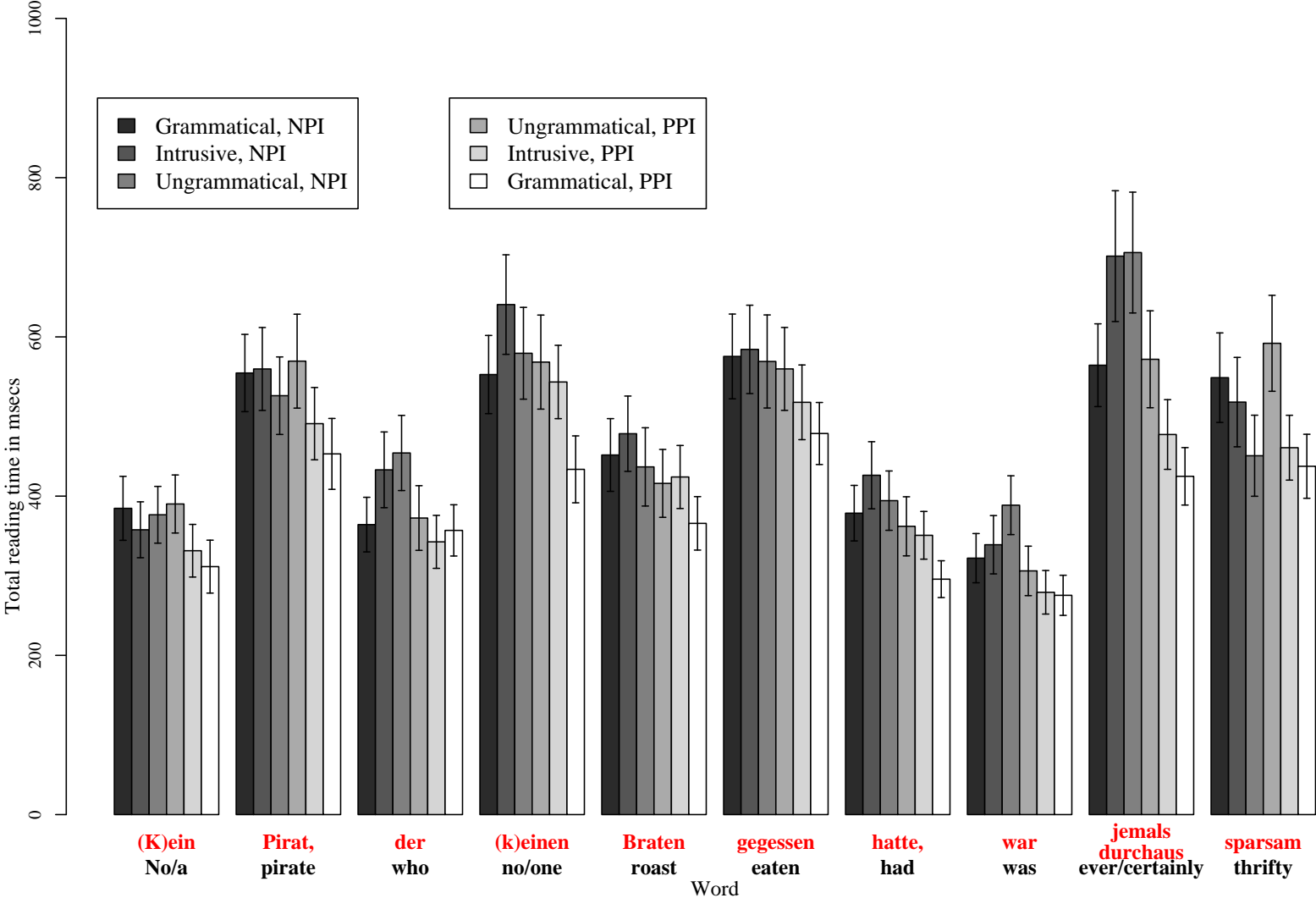
## Analysis and results

- All RTs below 50 milliseconds were removed, on the assumption that they cannot reflect higher-level processes. (Although keeping them does not affect the results.)
- Log transforms were carried out on all reading times before analysis because they were exponentially distributed.
- Linear mixed-effects models (Pinheiro and Bates 2000) were used for computing ANOVA, with subjects (or items) as random effects, and the conditions as fixed effects. Bates' `lme4` package in R was used for analysis.
- The dependent measures reported today are based on FPRT and TRT.

# First pass reading time



Total reading time by word position (in msecs), with 95% CIs



## Results and discussion

- NPIs:
  - Slowest when no licensor present.
  - Faster when intrusive licensor present.
  - Fastest when legal licensor present.
- PPIs:
  - Fastest when no licensor present.
  - Slower when intrusive licensor present.
  - Slowest when legal licensor present.

NPIs and PPIs are perfect mirror images of each other in terms of intrusion effects.

## Summary so far

- Cue-based retrieval has been a robust explanatory mechanism for dependency satisfaction during parsing, and can explain the peculiarities of polarity-driven dependencies as well.
- The cue-based retrieval explanation is validated across three different methodologies: ERPs, speeded grammaticality judgements, and eyetracking.
- The behavior of positive polarity items is best explained by the assumption that they actually look for an NPI licenser and signal an error if one shows up. This is consistent with independently motivated assumptions forwarded by Szabolcsi (2004) and others.



## Broader processing issues

Dependency resolution costs play an important role in determining the development of sentence processing theories:

- (7) a. **Whom**<sub>*i*</sub> did the student standing by the corridor . . . see *t*<sub>*i*</sub>.
- b. The **student** whom I **saw** . . .

A key factor is locality: the distance between the gap/head and the filler/argument affects processing difficulty.

## (Anti-)locality in German

Konieczny (2000) was the first to show that locality does not hold in German.

Konieczny found a **speedup** at the verb *hingelegt* when a relative clause intervened between the argument *Buch* and the verb.

- (8) a. Er hat das Buch hingelegt, **das Lisa gestern gekauft hatte**  
He has the book laid down, that Lisa yesterday bought had  
“He has laid down the book that Lisa had bought yesterday.”
- b. Er hat das Buch, **das Lisa gestern gekauft hatte**, hingelegt  
He has the book, that Lisa yesterday bought had, laid down  
“He has laid down the book that Lisa had bought yesterday.”

Call this the “anti-locality” effect.

## Polarity licensing: a dependency resolution problem with knobs on

(9) [<sub>NP</sub> **No** man [who had a beard]] was **ever** happy

1. Illegally positioned licensors can mess up the dependency resolution process (the intrusion effect).
2. The licensor and licensee are in a dependency relationship. Therefore, (anti-)locality should affect processing at the NPI.
3. NPIs and their licensors have a very special property—the strength of the dependency can differ:
  - Weak licensors: *Jeder*, every.
  - Stronger licensors: *Kein*, no.

Question: does the **strength** of the licensor affect speed of dependency resolution?

## Locality in German and English NPIs

Warren, Vasishth, Hirotani, and Drenhaus (CUNY 2006) polarity study (self-paced reading):

- (10)
- a. **Kein** Student, der **jemals** Physik studiert hat, kam montags zum Seminar.
  - b. **Kein** Student, von dem der Professor angenommen hatte, dass er **jemals** Physik studiert hat, kam montags zum Seminar.
  - c. **Jeder** Student, der **jemals** Physik studiert hat, kam montags zum Seminar.
  - d. **Jeder** Student, von dem der Professor angenommen hatte, dass er **jemals** Physik studiert hat, kam montags zum Seminar.
  - e. **Der** Student, der **jemals** Physik studiert hat, kam montags zum Seminar.
  - f. **Der** Student, von dem der Professor angenommen hatte, dass er **jemals** Physik studiert hat, kam montags zum Seminar.

## Predictions

English:

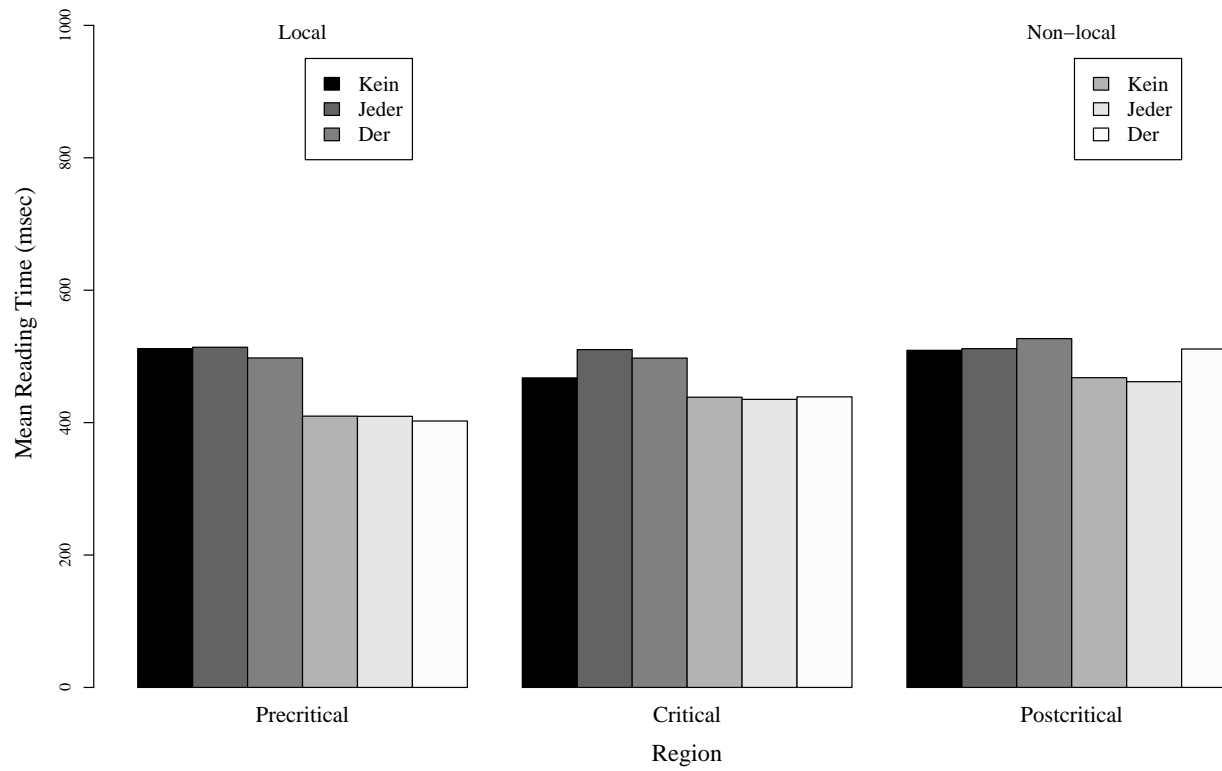
- Locality effect: increasing the licensor-NPI distance should cause a slowdown.
- “Strong” licensors like *No* should be integrated faster than weak licensors like *Every*.
- A slowdown should be seen at NPI when no licensor is present (*The*).

German:

- Anti-locality effect: increasing the licensor-NPI distance should cause a speedup at the NPI.
- “Strong” licensors like *Kein* should be integrated faster than weak licensors like *Jeder*.
- A slowdown should be seen at NPI when no licensor is present (*Der*).

# German results

German NPIs: The effects of licensor type and locality

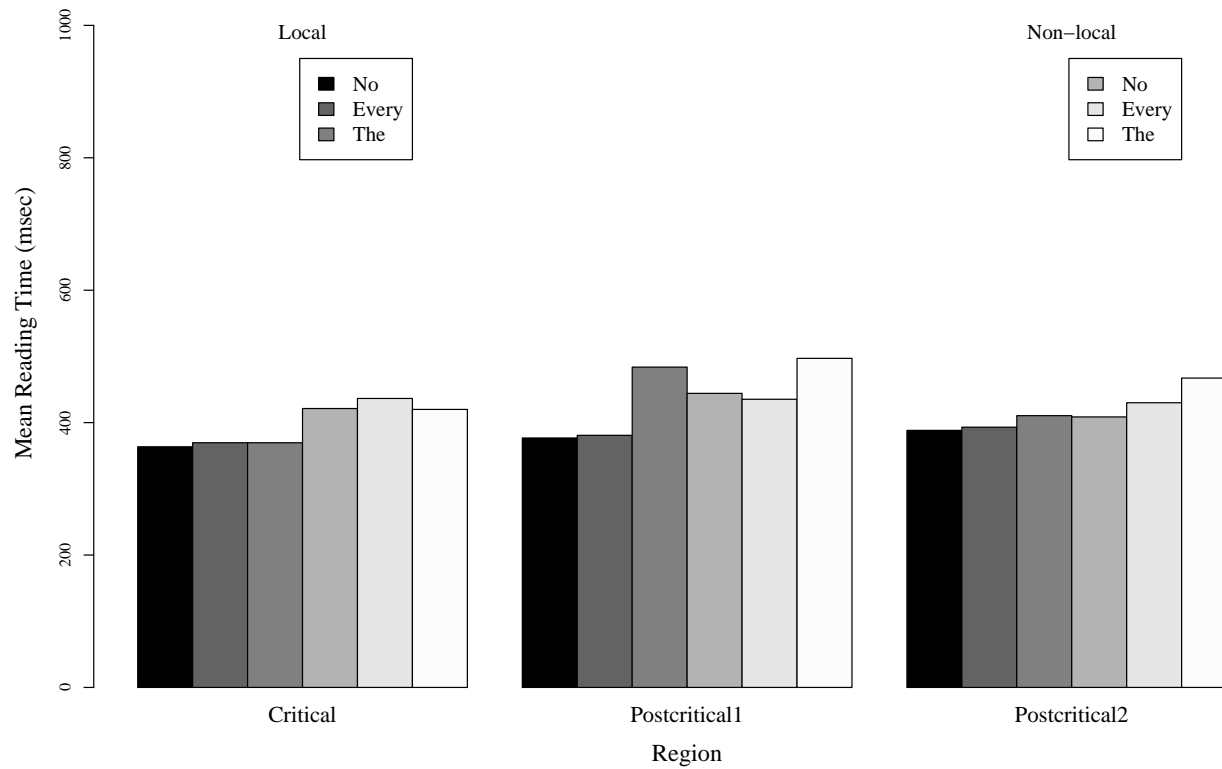


## English version

- (11) a. Every/No/The mailman who ever watched horror movies played baseball for fun.
- b. Every/No/The mailman who the captain claimed ever watched horror movies played baseball for fun.

# English results at NPI and following words

English NPIs: The effects of licenser type and locality





## Summary for English

- At NPI: locality effect
- At post-critical region 1: licenser and locality effect
- At post-critical region 2: licenser effect

## Discussion

- Strong licensors are retrieved faster than weak ones—there is a processing correlate to the independently established assumption about licensor strength.
- The locality effect seen in English dependencies is also seen in NPI dependencies.
- The anti-locality effect seen in German dependencies is also seen in NPI dependencies.

## Experiment 3

Self-paced reading and ERP study:

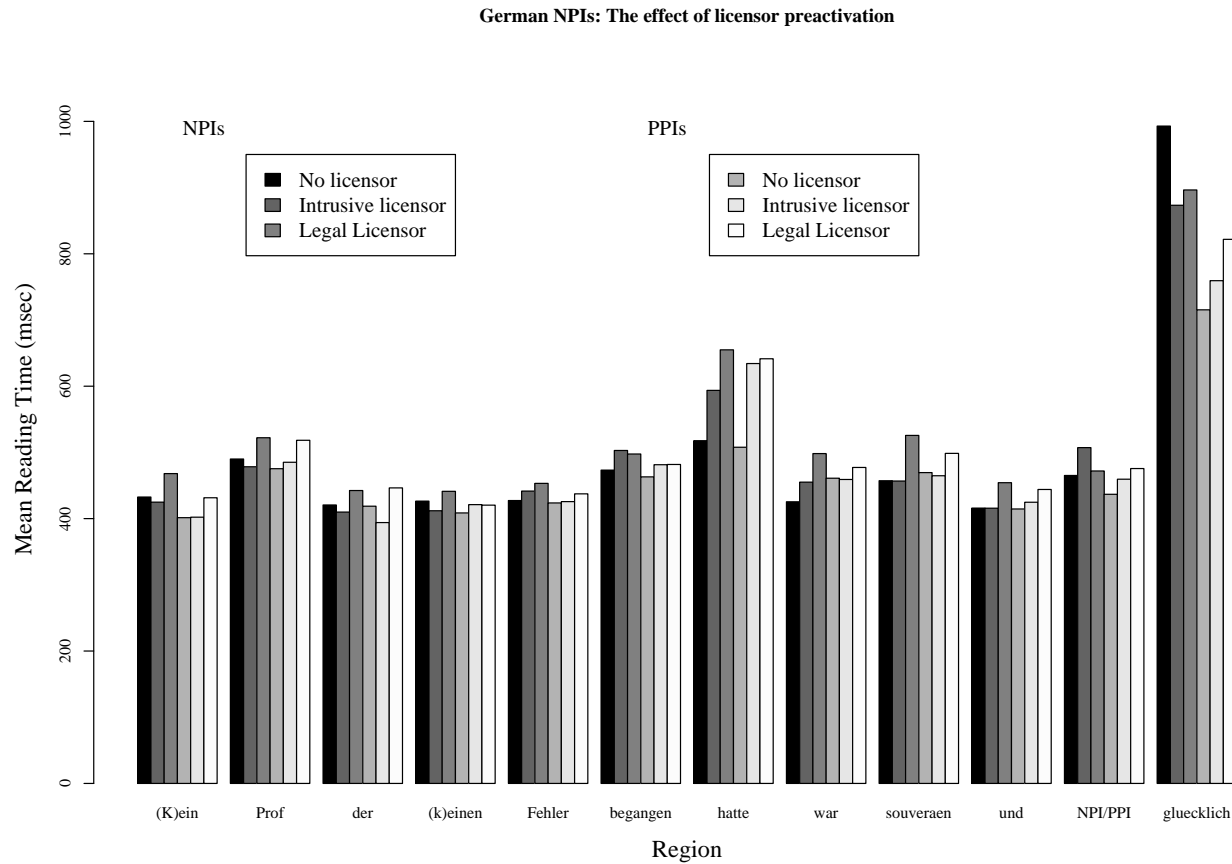
(12) (K)ein Professor, der (k)einen Fehler begangen hatte, war souverän und  
jemals/durchaus glücklich

Motivation:

- What is the cost of retrieving *Kein Professor* versus ein Professor independent of NPI licensing?
- If the two retrieval cues at the NPI (c-commander and NPI-licensor) are of equal strength then:
  - If the NPI licensor is made to have a higher activation, the intrusion effect should disappear.

If the intrusion effect does not disappear, it means that the semantic NPI-licensor cue is stronger than the syntactic c-command cue.

# Intermediate results (n=24) of ongoing SPR experiment



## Discussion

- At “hatte” a negative quantifier is harder to retrieve *per se*.
- At the first adjective “souverän” we see the same pattern.
- Preactivating the matrix NP does not seem to change the intrusion effect for NPIs or PPIs—this suggests that the semantic cue at the NPI (“give me an NPI-licensor”) is actually stronger than the syntactic cue (“give me a c-commander”).

## Concluding remarks and broader implications

- The only explanation for the intrusion effect is stochastic, cue-based retrieval: this rules out all sentence processing theories that rely on rigid metrics like head-dependent distance (Gibson, Hawkins) or any other deterministic measures.
- The locality and anti-locality effects seen in the psycholinguistic literature occur even in NPI licensing—the nature of the dependency between the NPI and the licenser is qualitatively identical to other ones.
- NPIs' lexical entries contain information about the semantic strength of the licenser: strong licensers are retrieved faster than weak licensers.
- The behavior of PPI processing in real time can be explained only by assuming that the PPI sets out to find an NPI licenser and raises an error flag the moment it finds one. AFAIK there is no other dependency quite like this in natural language.

The architecture of the parser **has** to be based on stochastic, cue-based retrieval. All alternatives will fail to explain the available data on dependency resolution.