Exemplars in Syntax: Evidence from Priming

Neal Snider, Dept. of Linguistics, Stanford University

ESSLLI Workshop on Exemplar-based Models

August 15, 2007

Hypothesis

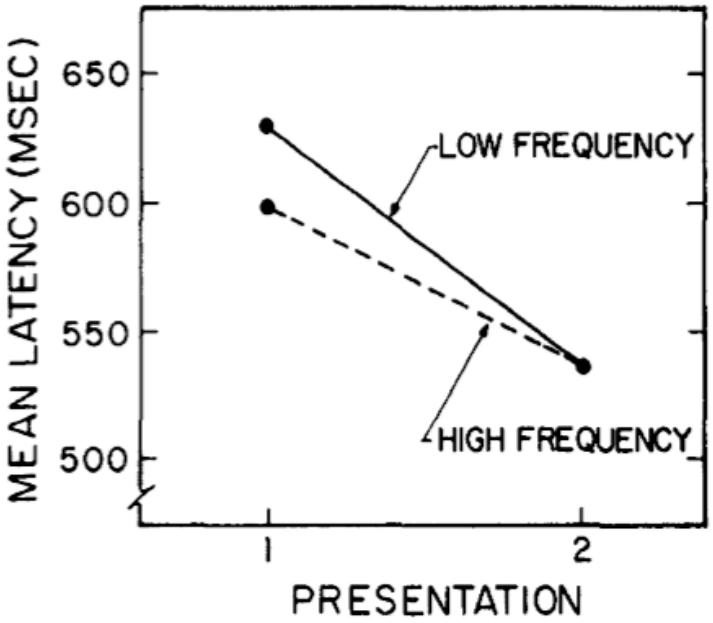
- What do exemplar models predict about language processing?
- No explicit distinction between levels of word and structure

→Lexical access and structural access should be the same process

  - Evidence: priming

Lexical priming

- Words are recognized faster when they have been encountered before
- BUT, high frequency words get less of a boost from priming (Scarborough et al 1977; Versace and Nevers 2003; etc)



"Inverse frequency effect"

Lexical priming

- Also, more similar words prime more (Ratcliff and McKoon 1981; 1988; Schreuder et al 1983)

- Do multi-word structures show priming?

Structural priming (Bock 1986)

TRANSITIVE **DATIVE**

PRIMING SENTENCES

ACTIVE:
ONE OF THE FANS PUNCHED THE REFEREE.

PREPOSITIONAL:
A ROCK STAR SOLD SOME COCAINE TO AN UNDERCOVER AGENT.

PASSIVE:
THE REFEREE WAS PUNCHED BY ONE OF THE FANS.

DOUBLE OBJECT:
A ROCK STAR SOLD AN UNDERCOVER AGENT SOME COCAINE.

TARGET PICTURES

- passive description more likely following passive sentence
- prepositional dative more likely following prepositional dative
- double object dative more likely following double object

Structural priming

- Tendency for syntactic structures to persist in dialogue: (Szmrecsanyi 2005; Gries 2005)

"… I don't feel we should **loan [them] [money].** … I wish our leaders were really seeking the Lord on these things, and if we feel led to **give [a country] [money]** to help them, fine"
(Switchboard corpus)

Prediction

- If lexical and structural priming access the same representations:

1) Less frequent structures should prime more strongly
2) More similar structures should prime more strongly

Exemplar frequency in ditransitive priming

Jaeger and Snider 2007

Data

- Modeling syntactic choice in the ditransitive alternation:

NP***PP***	give [money] [to a country]
NP***NP***	give [a country] [money]

- Bresnan et al (2007) model has 95% accuracy

(baseline is 80%)

Data

- 1,153 tokens from Switchboard corpus (all primed turns from Bresnan et al 2007)
- Naturalistic, spontaneous speech, conversational data
- Analyzed with logistic regression
  - positive response is NPPP

Controls (Bresnan et al 2007 + more)

- **Recipient**
  - *Pronominality*
  - *Givenness*
  - *Definiteness*
  - *Log length*
  - *Animacy*
- **Theme**
  - *Pronominality*
  - *Givenness*
  - *Definiteness*
  - *Log length*
  - *Number*
- **Verb**
  - *Semantic class*
  - *Target verb bias*

- All controls replicate

Exemplar frequency

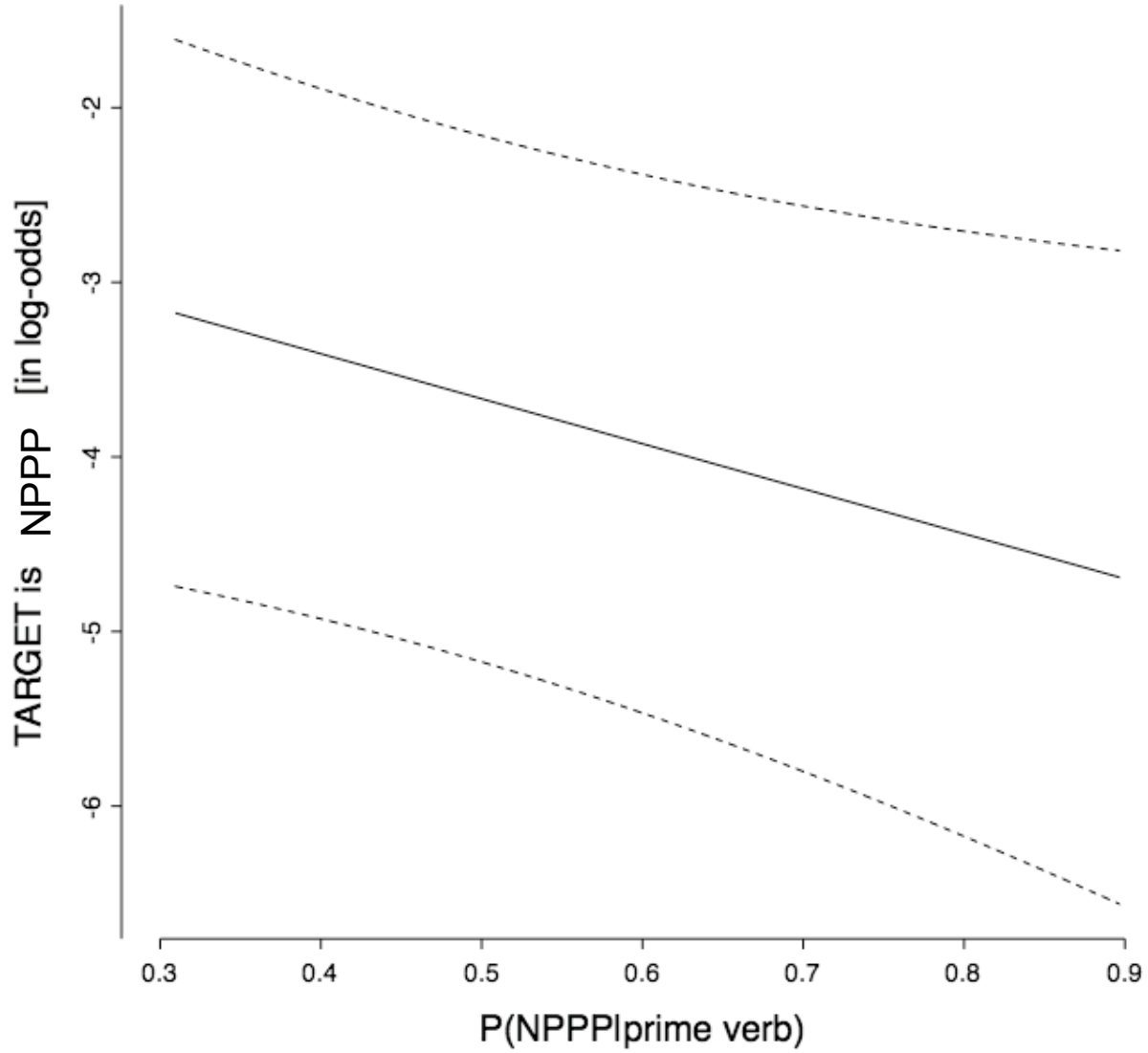
- In this study, exemplar frequency was taken to be the verb bias: P(NPPP|verb)
  - estimated from Switchboard data set
- Because of inverse frequency effect in priming, verb bias is expected to have a **negative** effect on priming
- Thus we called it a surprisal effect

Results of interest

- Split data into NPNP primes (896) and NPPP primes (257)
- *Prime Frequency*
  - **If prime is NPPP** ($p < 0.03$)
  - **If prime is NPNP** (n.s.)
- *Target Frequency* ($p < .01$)

Effect of *Prime Frequency*

Primes least likely to be NPPP make NPPP **4.5** times more likely than primes most expected to be NPPP



Summary so far

- Effect of exemplar frequency on structural priming
- Less frequent exemplars prime more strongly
  - Like in lexical priming
- Can this effect be found in other structures?

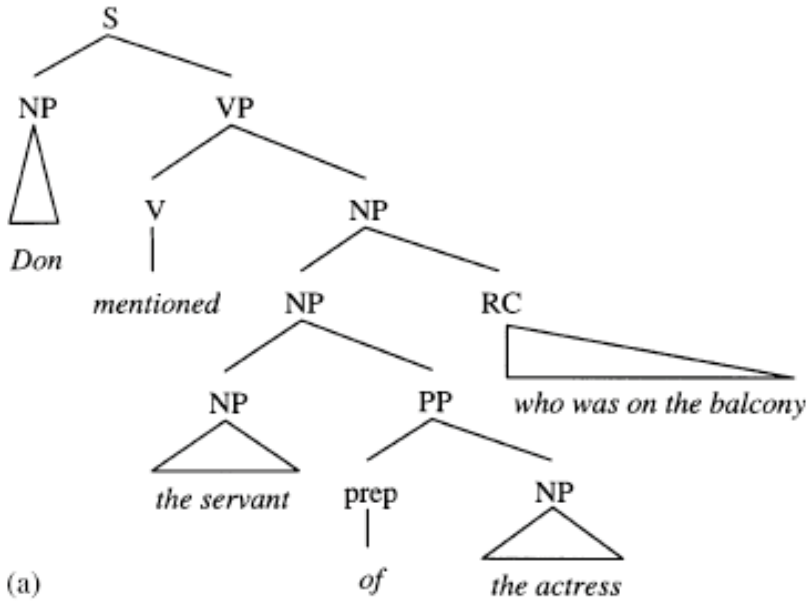
Study 1

Exemplar frequency in RC attachment priming

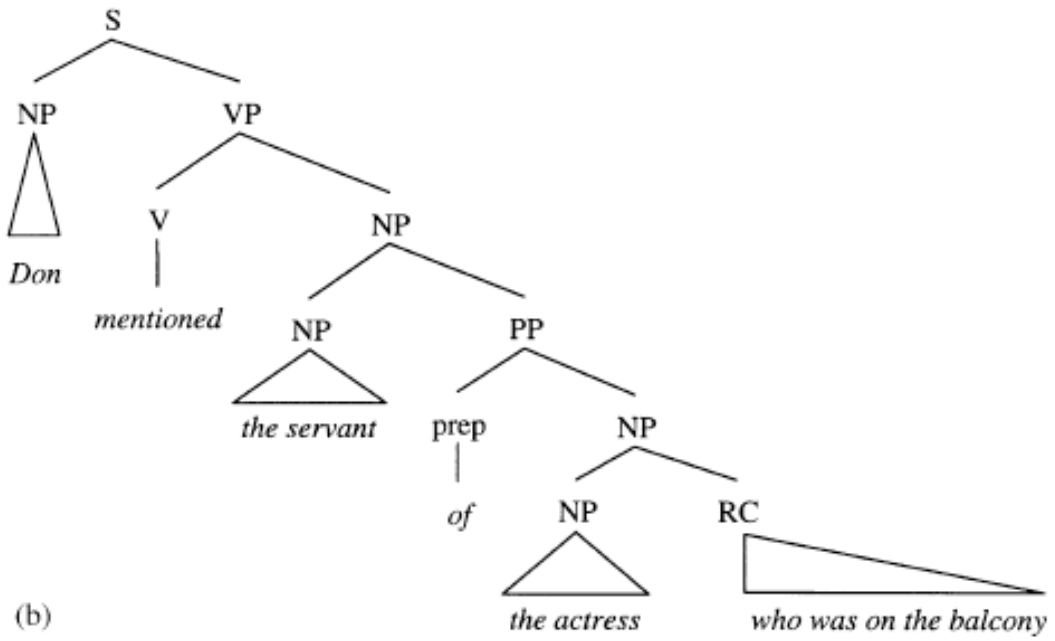
RC Attachment height

- Priming shown in Scheepers (2003), Desmet (2005), etc.

High:



Low:



Corpus study

- Do RC attachments show priming in naturalistic data?
- Are they affected by exemplar frequency?

Data

- Multiply-embedded NPs with RC from Treebank Switchboard
- 272 prime-target pairs
  - 190 LO-attached targets
  - 82 HI-attached targets
- Analyzed with logistic regression
  - positive response is HI attachment

Frequency measure

- P(RC|head N) (from Jaeger 2006)
- But the head N changes depending on the prime
  - And data too sparse to determine bias of N pairs
- How to measure prime surprisal?

Surprisal measure

- What makes a prime surprising?
  - If the RC was attached Hi, but the Lo N has a strong RC bias, and vice versa:

$$\text{prime surprisal} = \begin{cases} P(RC|\text{LO noun}) & \text{if prime = HI} \\ -P(RC|\text{HI noun}) & \text{if prime = LO} \end{cases}$$

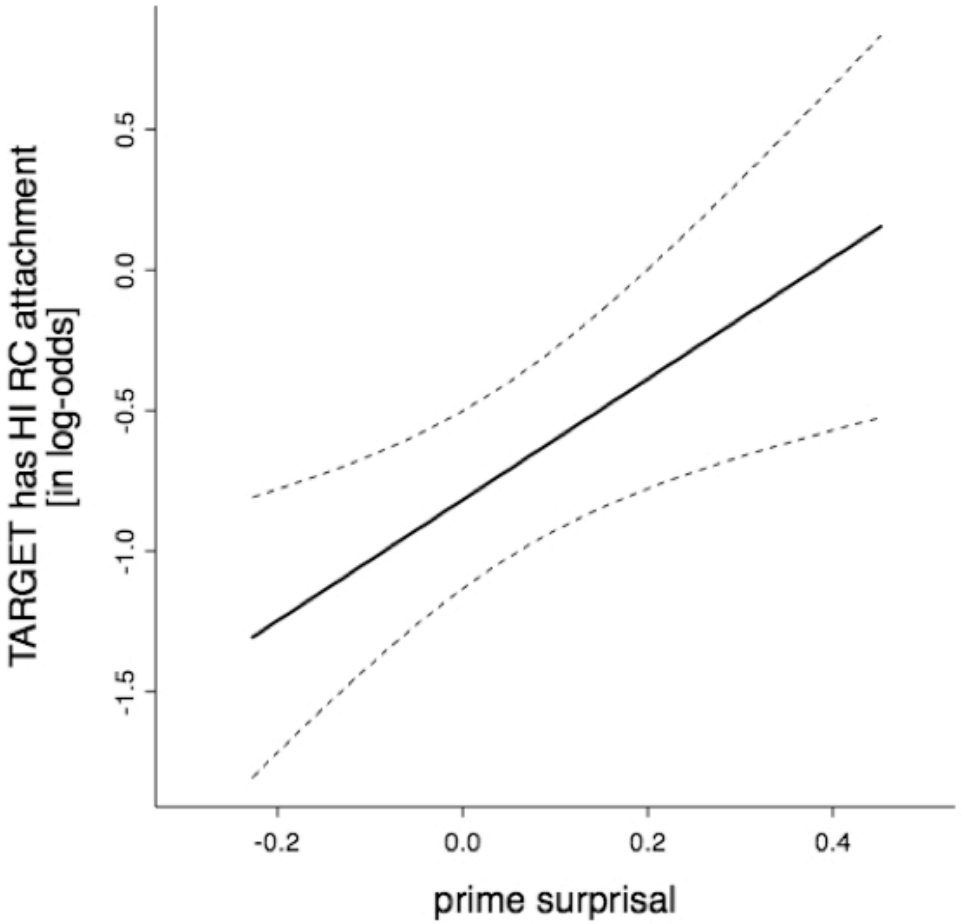
Results

Table 1: Summary of RC attachment analysis

Predictor (independent variable)	*Parameter estimates* Log-odds	S.E.	Odds	*Wald's test* Z	P	$\Delta_x(\Lambda)$-*test* χ^2	P
P(RC | LO target N)	-2.298	0.842	0.10	-2.7	< 0.01	8.0	< 0.005
P(RC | HI target N)	5.297	1.454	200.33	3.7	< 0.005	14.1	< 0.005
prime surprisal	2.293	0.74	9.90	2.8	< 0.005	8.3	< 0.005

- Significant effects of prime bias and target bias

RC prime surprisal



Summary so far

- RC attachment bias affects priming
- Effect of exemplar frequency
  - In ditransitives
  - In relative clause attachment
- Just like in lexical priming, there is an inverse frequency effect in structural priming

- Is there an effect of prime-target similarity?

Study 2

Exemplar similarity in ditransitive priming

Priming and similarity

- As with lexical priming:
  - More similar exemplars should be more likely to prime!
- How to measure similarity?
- Nearest-neighbor models offer precise metric of exemplar similarity

k-NN similarity metric

- Distance is weighted sum of differences between all the features: (Daelemans 2005)

$$\Delta(X,Y) = \sum_i w_i \delta(x_i, y_i)$$

- Where the difference is defined as:

$$\delta(x_i, y_i) = \begin{cases} \frac{x_i - y_i}{max_i - min_i} & \text{if numeric, otherwise} \\ 0 & \text{if } x_i = y_i \\ 1 & \text{if } x_i \neq y_i \end{cases}$$

Question

- Are exemplars that have less k-NN distance more likely to prime?

Data

- Ditransitive data from Bresnan et al 2007
- Features (and controls):

Recipient

Pronominality
Givenness
Definiteness
Log length
Animacy
Identity

Theme

Pronominality
Givenness
Definiteness
Log length
Number
Identity

Verb

Semantic class
Identity

Examples of exemplar distance

- d=0.66:

  B: He hasn't given it to someone else

  …

  A: I'll take it to him to do it

- d=1.2

  A: I'm there to teach them something

  …

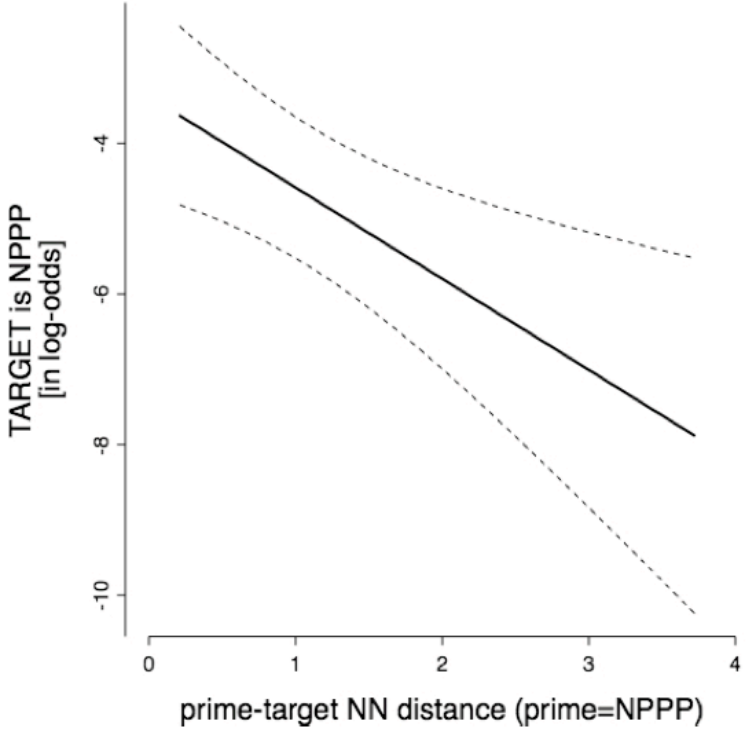
  B: I don't pay much attention to it.

Results

Predictor (independent variable)	*Parameter estimates* Log-odds	S.E.	Odds	*Wald's test* Z	P	$\Delta_x(\Lambda)$-*test* χ^2	P
prime=NPPP	1.665	1.236	11.85	1.4	> 0.15	11.0	< 0.005
prime V = target V * prime=NPPP	-0.024	0.825	0.97	-0.1	< 1	0.03	< 1
NN distance * prime=NPPP	-2.365	0.869	0.05	-2.7	< 0.01	10.8	< 0.005

- Controls replicate
- Distance * prime significant
- Main effect of prime n.s.
- Verb identity * prime n.s.

So effect is not due to known verb repetition boost



Conclusions

- Just like lexical priming:
- Structural priming is affected by
  - (Inverse) exemplar frequency
  - Exemplar similarity

- Evidence that lexical production and structural production access same representations
  - As predicted by exemplar models

Future work

- Experiments
- Improve similarity metric
- Different constructions
- Other phenomena
  - neighborhood effects
  - how to define neighborhood for constructions?

Thanks to:

- Florian Jaeger
- Joan Bresnan