# The interaction of usage and function in the emergence of constructions

## A novel construction learning study

Florent Perek

Universität Basel

Adele Goldberg

Princeton University

- Learners want to understand messages given forms (comprehension).

and

- They need to choose forms to express the messages they want to convey (production).

Need to categorize

form ~ function pairings: *constructions*.

**Constructions involve:**

(Abstract) surface form

The X-*er* ___, the Y-*er* ___

Subj V Obj1 Obj2

<verb>-*able*, e.g., *reachable*, *dialable*

*think out of the box*, *pull* <oneself> *together*

Function: semantics and/or information structure

Register, genre, dialect

Relationships to other constructions

Remembered exemplars

Subtle semantic differences between constructions

a. Jo baked Sam a cake.

b. Jo baked a cake for Sam.

She gave *him* a book. >
She gave *a man* a book.

Strong statistical skewing toward pronominal/topical recipients. (Bresnan 2010; Dryer 1986; Givon 1979; Langacker 1987; Arnold et al. 2000; Bresnan and Nikitina 2008; Wasow 2002; Levin and Rappaport Hovav 2004; Goldberg 2006)

# Usage-based constructionist approach

- Grammar emerges from usage
- Language acquisition is input-driven
- Speakers are sensitive to statistical information

# Novel construction learning studies

- "Made-up" languages involving scene-sentence pairs
- The statistical structure of the input is manipulated
- The role of statistics in language learning is investigated

(e.g., Goldberg et al. 2004; Casenhiser & Goldberg 2005; Hudson Kam & Newport 2005; Wonnacott, Tanenhaus, & Newport 2008; Wonnacott, Boyd, & Goldberg 2012)

# Wonnacott, Newport, & Tanenhaus (2008)

- Two constructions with same meaning
  - "Verb Agent Patient " (VSO)
  - "Verb Patient Agent *ka* " (VOS-*ka*)
- Distribution varied across conditions
  - Some verbs occurred *only* in either VSO or VOS-ka
  - Some verbs alternated, i.e., they occurred in both

## Wonnacott, Newport & Tanenhaus (2008)

- Learners depended on the statistics in the input:
  - "Lexicalist" input condition:
    No verbs alternated

## Wonnacott, Newport & Tanenhaus (2008)

- Learners depended on the statistics in the input:
  - "Lexicalist" input condition:
    No verbs alternated → completely conservative behavior

## Wonnacott, Newport & Tanenhaus (2008)

- Learners depended on the statistics in the input:
  - "Lexicalist" input condition:
    No verbs alternated → completely conservative behavior
  - 33% of verbs alternated in input condition:

## Wonnacott, Newport & Tanenhaus (2008)

- Learners depended on the statistics in the input:
  - "Lexicalist" input condition:
    No verbs alternated → completely conservative behavior
  - 33% of verbs alternated in input condition:
    Partially general and partially lexically specific behavior.

## Wonnacott, Newport & Tanenhaus (2008)

- Learners depended on the statistics in the input:
  - "Lexicalist" input condition:
    No verbs alternated → completely conservative behavior
  - 33% of verbs alternated in input condition:
    Partially general and partially lexically specific behavior.
  - Alternating input condition:
    All verbs alternated

## Wonnacott, Newport & Tanenhaus (2008)

- Learners depended on the statistics in the input:
  - "Lexicalist" input condition:
    No verbs alternated → completely conservative behavior
  - 33% of verbs alternated in input condition:
    Partially general and partially lexically specific behavior.
  - Alternating input condition:
    All verbs alternated → fully productive behavior

- Does language learning *only* consist of gleaning statistical regularities in the input?
- There are also learning biases & constraints
  - Communicative (e.g., Levy 2008; Gibson et al. 2011; Piantadosi et al. 2012)
  - Cognitive: memory (e.g., Gathercole & Baddeley 1993; Gibson 1994), inductive categorization processes (e.g., Griffiths et al. 2010; Suttle & Goldberg 2011 ), associative learning (e.g., Yu & Smith 2008)
- What about the function of constructions themselves?

– Differences in grammatical form typically correspond to differences in function (Bolinger 1968; Goldberg 1995)

→ Use constructions with distinguishable functions.

# Our experiment

- Two word order constructions: SOV and OSV
- OSV order used exclusively with pronoun patients

| | |
|---|---|
| the panda the pig mooped | SOV |
| him the panda mooped | ProSV |

meaning: 'the panda$_{agent}$ pushed the pig$_{patient}$'

- Six novel verbs (e.g., *glim*, *moop*, *wub*) referring to transitive actions (e.g., 'punch', 'push', 'head-butt')

# Our experiment

- Two experimental conditions
  - Lexicalist condition: 3 SOV-only, 3 ProSV-only verbs
  - (Partially) alternating condition: 2 SOV-only, 2 ProSV-only, 2 alternating verbs
- Control condition (same-meaning condition)
  - 3 SOV-only; 3 OSV-only

  to replicate Wonnacott et al. and check that speakers are able to learn verb-specific behavior

Perek & Goldberg
(to appear, *Journal of Memory and Language*)

# Example of exposure pair

the rabbit the panda norped 

# Procedure

- Exposure (2 days)
  - 36 sentence-scene pairs, each verb used 6 times
  - Participants asked to repeat each sentence
- Sentence production task
  - Participants asked to describe new scenes with learned novel verbs.
  - Interspersed with distractor tasks (vocabulary questions, forced-choice sentence comprehension)
- Sentence rating task (not reported here; consistent with production)

# Production task

- Different question contexts:
  - "What happened here?"
  - "What happened to the <patient>?"

Two trials per verb, one in each context

# Example of production trial

what happened here?

# Example of production trial

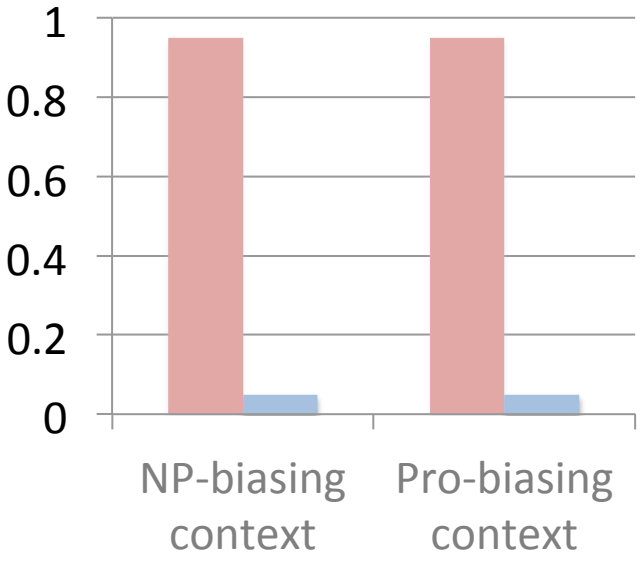
what happened to the panda?

# Participants

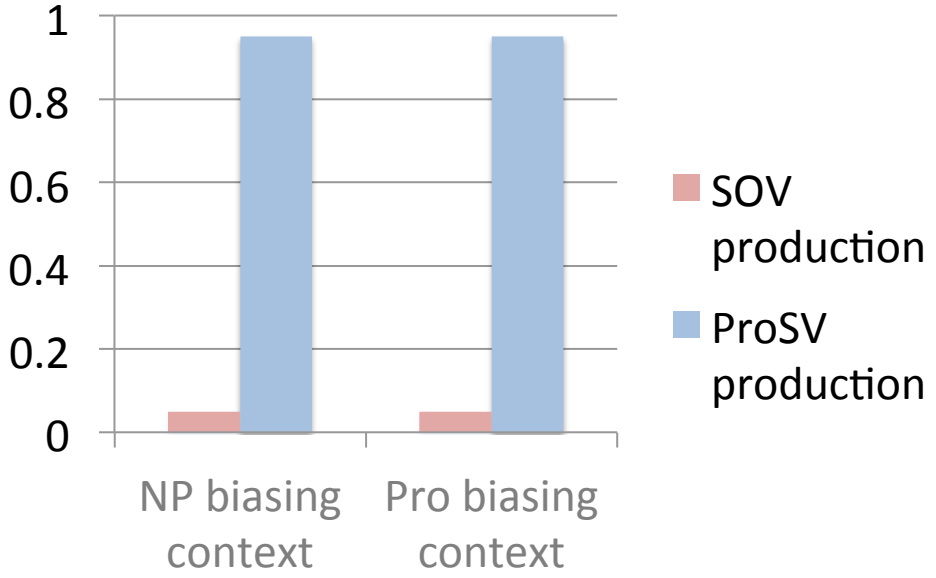
- 64 Princeton undergrads, aged 18-22
  - 24 in lexicalist condition
  - 18 in partially alternating condition (2/6 verbs alternate)
  - 12 in the control, same-meaning. lexicalist condition

To what extent do speakers generalize constructions to unattested verbs?

Hypothetical data: conservative, verb-based behavior



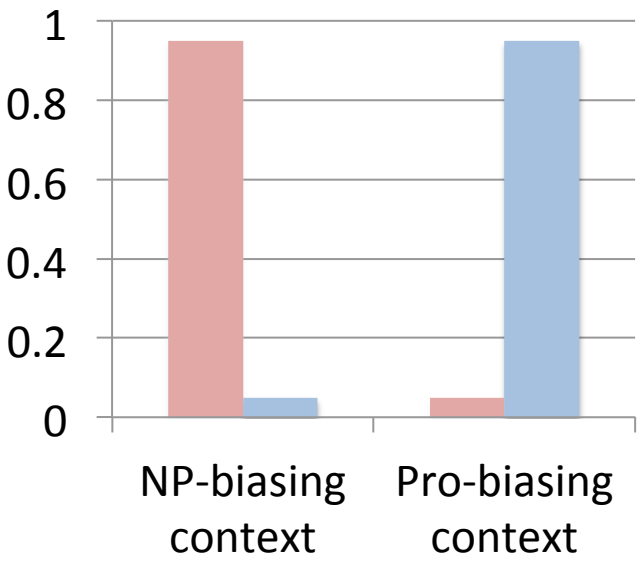

SOV-only verbs



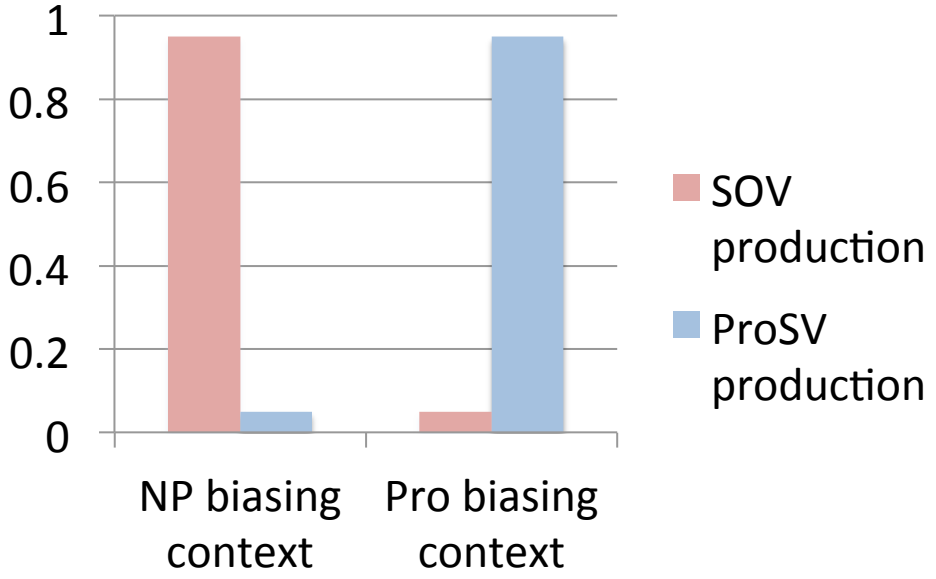

ProSV-only verbs

# To what extent do speakers generalize constructions to unattested verbs?

Hypothetical data: full generalization across verbs

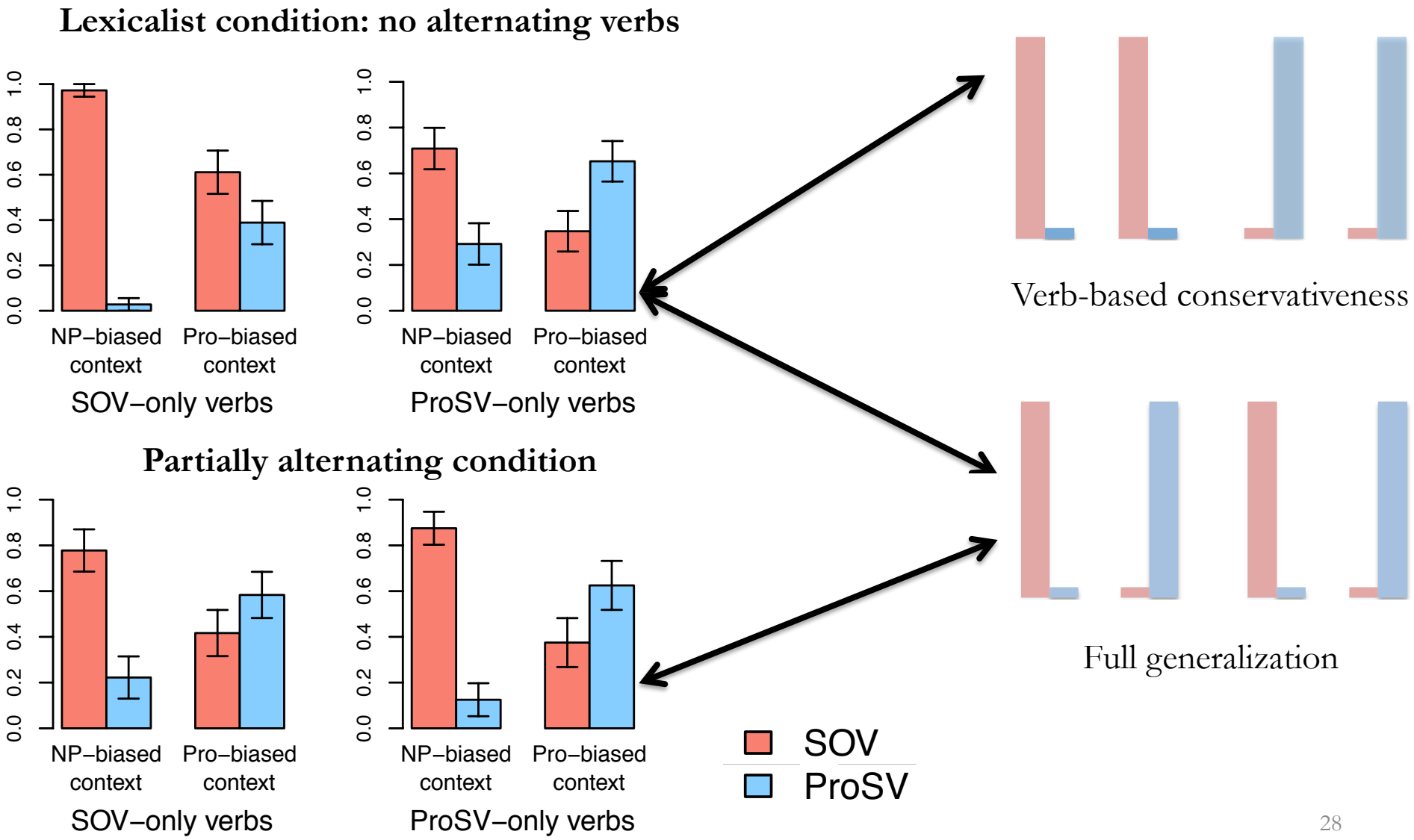



SOV-only in input



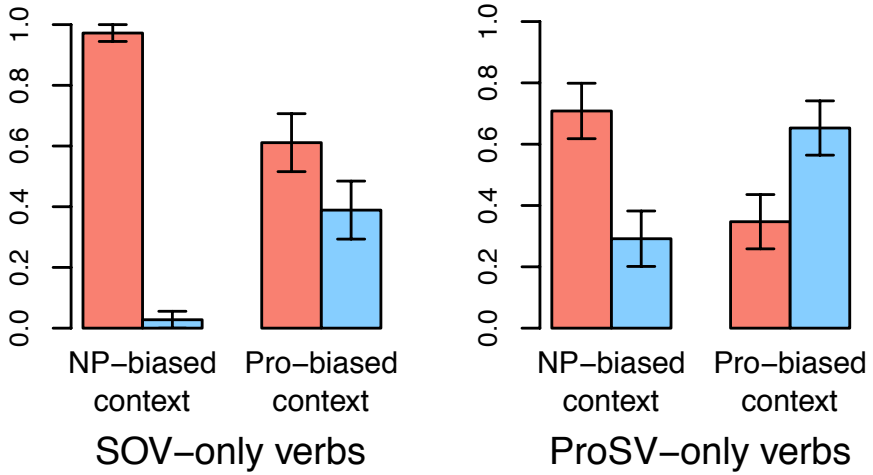

ProSV-only in input

# Results

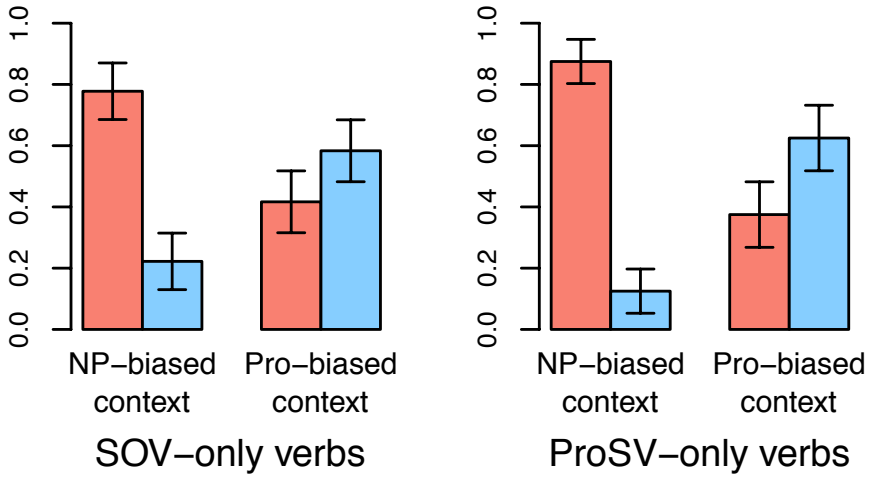

Lexicalist condition: no alternating verbs
0.0 0.2 0.4 0.6 0.8 1.0
NP–biased context
Pro–biased context
SOV–only verbs
NP–biased context
Pro–biased context
ProSV–only verbs
Partially alternating condition
0.0 0.2 0.4 0.6 0.8 1.0
NP–biased context
Pro–biased context
SOV–only verbs
NP–biased context
Pro–biased context
ProSV–only verbs
SOV
ProSV
Verb-based conservativeness
Full generalization

# Results

**Lexicalist condition: no alternating verbs**




**Partially alternating condition**





**Mixed effects logistic regression**

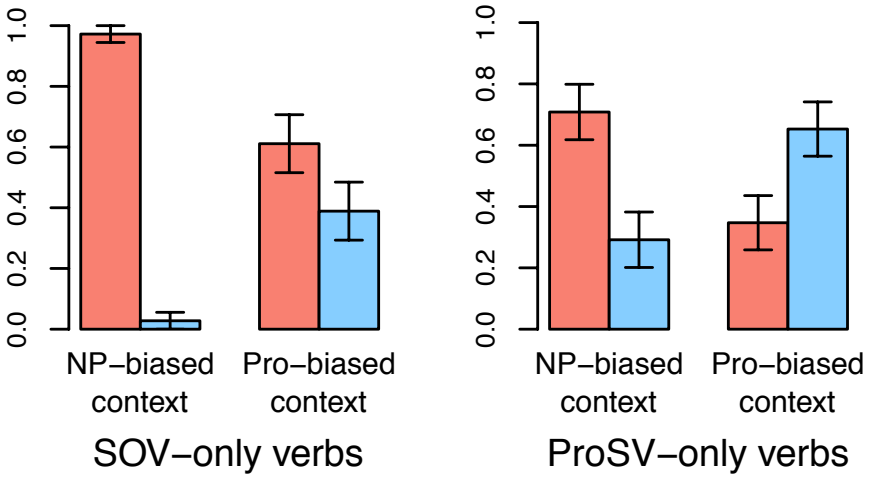
**Main effect of Context:** responses are context-dependent in both conditions ($p < 0.0001$).

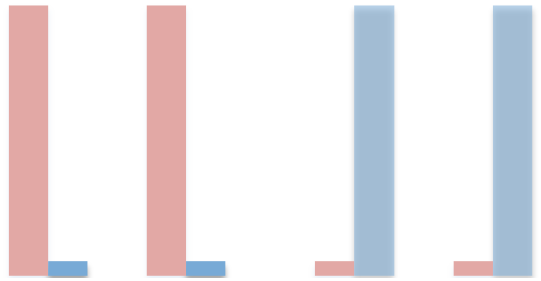
**Interaction between Condition and VerbType** ($p = 0.0001$) : a (conservative) effect of verb type is specific to the lexicalist condition

Subjects, items as random effects.

# Results

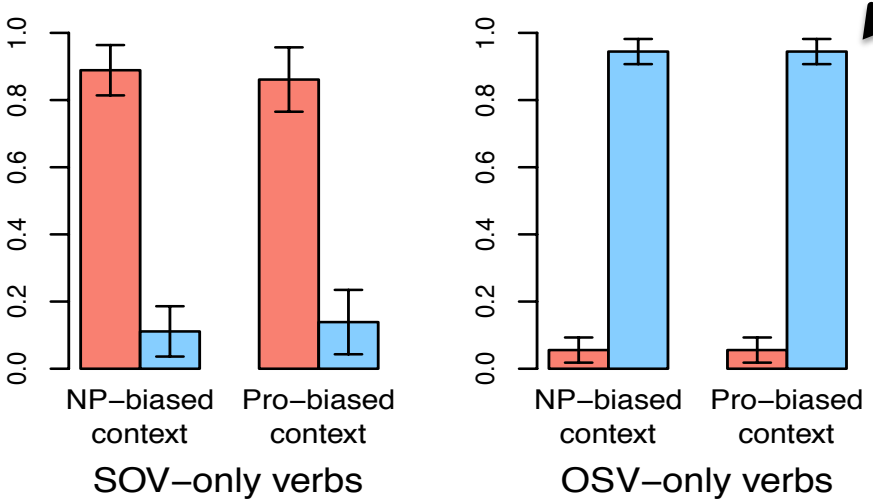
**Lexicalist: no alternating verbs, different functions**



Verb-based conservativeness

**Same-meaning: no alternating verbs, same function**

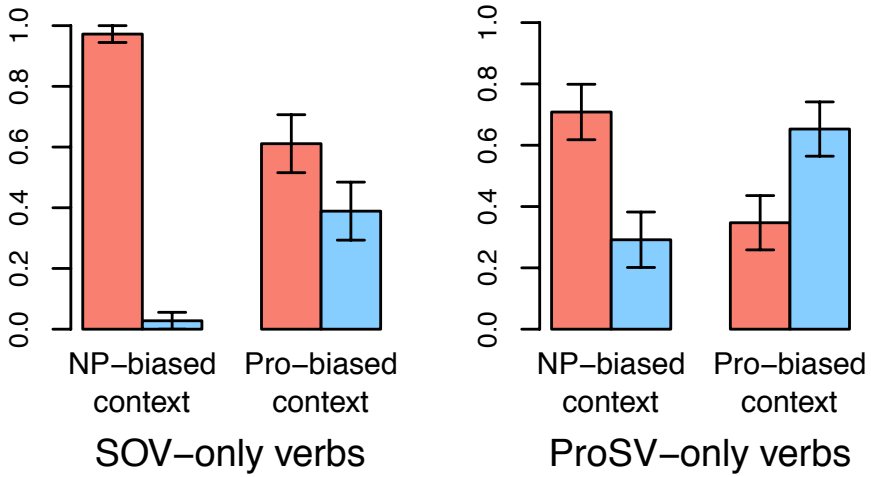



Full generalization

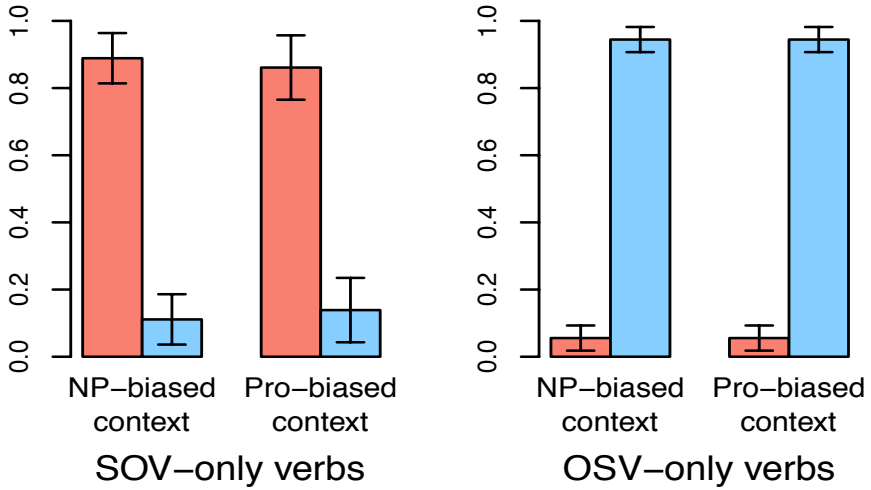
Perek & Goldberg
(to appear, *Journal of Memory and Language*)

# Results

**Lexicalist: no alternating verbs, different functions**




**Same-meaning: no alternating verbs, same function**




**Mixed effects logistic regression**

Condition interacts with both Context ($p < 0.0001$) and VerbType ($p = 0.0029$)

No effect of Context in the same-meaning condition.

Effect of VerbType stronger in same-meaning than in lexicalist condition

Perek & Goldberg
(to appear, *Journal of Memory and Language*)

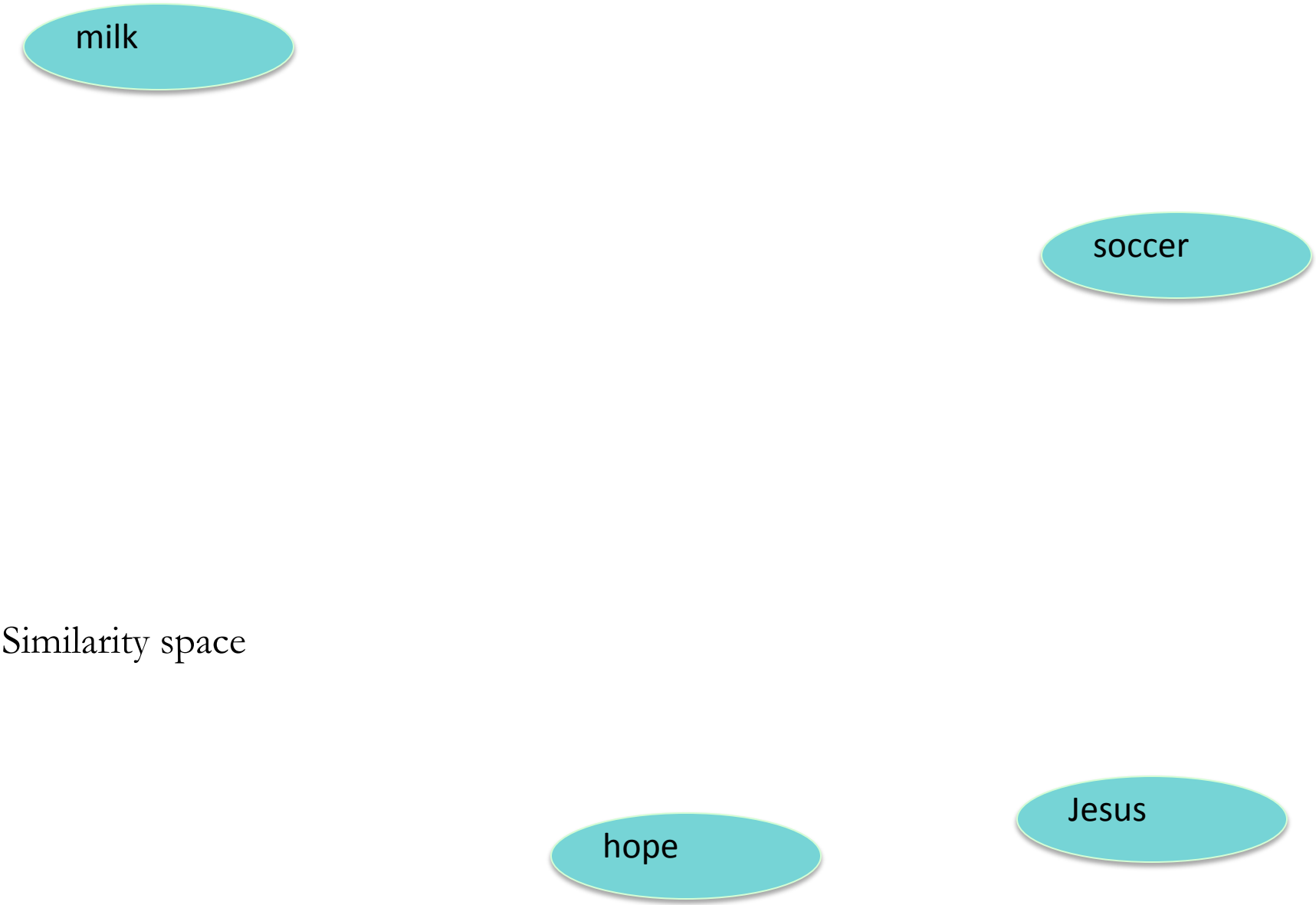
# Summary

- Tendency for participants to generalize, using contextually appropriate construction
  - Evidence of ignoring verb-specific input
- This tendency interacts with the input
  - Alternating verbs promote productivity, as in Wonnacott et al.
  - But here: full generalization when only 1/3 of verbs alternate
- Sentence rating results support production data

Perek & Goldberg
(to appear, *Journal of Memory and Language*)

- SOV and ProSV constructions are distinguished by discourse contexts (not by verb semantics).

- Constructions define relevant dimensions of similarity




Similarity space

milk

<something that is essential for good quality of life.>

soccer

Got _____?
as rhetorical question

got milk?

got soccer?

Constructions determine dimensions of similarity

Jesus

hope

got Jesus?

# Conclusions

Refinement of the usage-based approach

- Statistical information is essential to learn both item-specific patterns and general constructions
- But communicative functions of constructions determine which dimensions of similarity are relevant to generalizations

Thanks for your attention!

florent.perek@gmail.com

*Got constructions?*