Speakers sacrifice some (of the) precision in conveyed meaning to accommodate robust communication

Judith Degen and T. Florian Jaeger

University of Rochester

June 3, 2011

Choice of form

What drives speakers' choice of form in encoding an intended meaning?

Outline

Choices

(1) Clearly our understanding of the ability of these organisms to disperse was just underappreciated.

(NOVA science broadcast about Mt. St. Helen, 5/4/2010)

Choices

(1) Clearly our understanding of the ability of these organisms to disperse was just underappreciated.
(NOVA science broadcast about Mt. St. Helen, 5/4/2010)

(2) Choice of referring expression:
   a. Can you pass me the cup?
   b. Can you pass me the teacup?
   c. Can you pass me the black-and-white patterned teacup?

Speaker and addressee

Rational addressee's situation

- infer speaker's most likely intended meaning
- problem: utterances are generally noisy, ambiguous, underspecified

Speaker and addressee

Rational addressee's situation

- infer speaker's most likely intended meaning
- problem: utterances are generally noisy, ambiguous, underspecified
- compute probability distribution over intended meanings given utterance and context

Speaker and addressee

Rational addressee's situation

- infer speaker's most likely intended meaning
- problem: utterances are generally noisy, ambiguous, underspecified
- compute probability distribution over intended meanings given utterance and context

Rational speaker's situation

- maximize probability of conveying intended meaning while balancing other constraints, including cognitive limitations

Robust communication

- language and language use as efficiently organized systems (Zipf, 1949, Genzel & Charniak, 2002, Piantadosi et al., 2011)
- speakers manage tradeoff between redundancy and reduction in production (Aylett & Turk, 2004, Jaeger, 2006, Jaeger, 2010, Levy & Jaeger, 2007, Horn, Levinson, 2000)

Uniform Information Density Jaeger, 2006, Levy & Jaeger, 2007

Uniform Information Density (UID)

Within the bounds defined by grammar, speakers prefer utterances that distribute information uniformly across the signal.

UID prediction for alternations

At a choice point, speakers will choose variant with more uniform information density

Shannon information Shannon, 1948

$$\text{Info}(u) = -\log p(u)$$

Syntactic choice

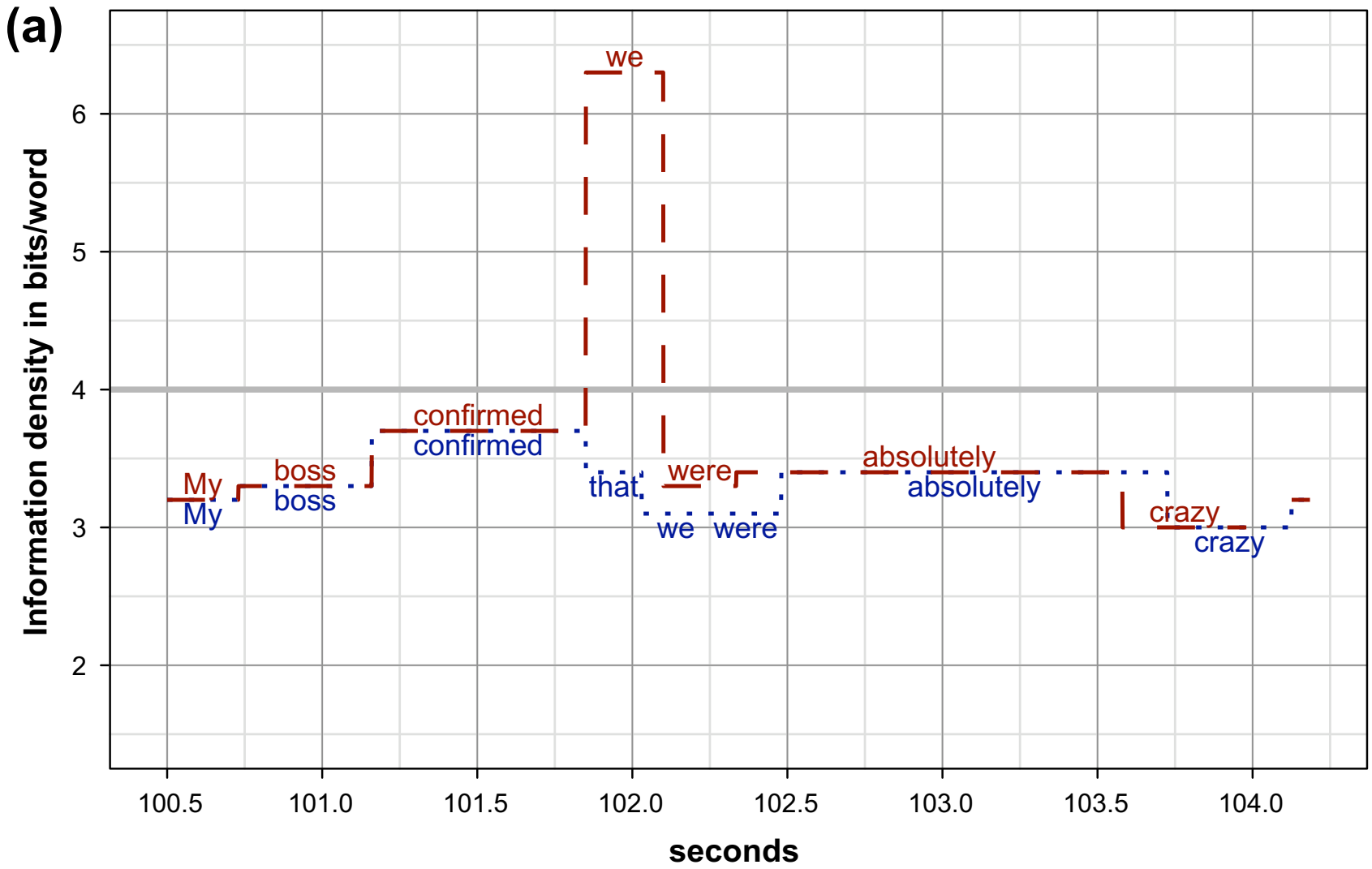
- reduction alternations (e.g. morphosyntactic reduction Frank & Jaeger, 2008, *that*-omission, planning beyond the clause-level Gallo et al., 2008)

(3) that-*omission* (Ferreira & Dell, 2000, Jaeger, 2010)

a. I think that exercise is really beneficial

b. I think ____ exercise is really beneficial

that-omission: avoid peaks and valleys from Jaeger, 2010



Support for UID

- support from alternations on many different levels of linguistic processing:
  - phonetic reduction (Aylett & Turk, 2004)
  - morphosyntactic reduction (Frank & Jaeger, 2008)
  - *that*-omission (Jaeger, 2010)
  - syntactic planning beyond clause level (Gallo et al., 2008)
- speakers prefer the longer form in high-information environments

The question

When given a choice between two forms that are not meaning-equivalent, do speakers sometimes sacrifice subtlety in meaning to robustly communicate the core meaning?

The question

When given a choice between two forms that are not meaning-equivalent, do speakers sometimes sacrifice subtlety in meaning to robustly communicate the core meaning?

The test case: not quite an alternation

(4) Alex ate some cashews. *[simple some]*

(5) Alex ate some of the cashews. *[partitive some]*

Pseudo-alternation

(6) *simple vs partitive* some
Alex ate some (of the) cashews

(7) *that*-omission
I think (that) exercise is really beneficial

Testing robust communication

(8) *core meaning:*
Alex ate SOME cashews

Information estimate

Info(SOME | context) = -log (p(*some* | context) + p(*some of the* | context))

Testing robust communication

(8) *core meaning:*
Alex ate SOME cashews

Information estimate

Info(SOME | context) = -log (p(*some* | context) + p(*some of the* | context))

The prediction

The more information carried by SOME in context, the more likely speakers should be to use the longer form, i.e. the partitive.

Assessing meaning differences

(9) Alex ate some of the cashews

Partitive Constraint
Jackendoff, 1977, Ladusaw, 1982, Reed, 1991

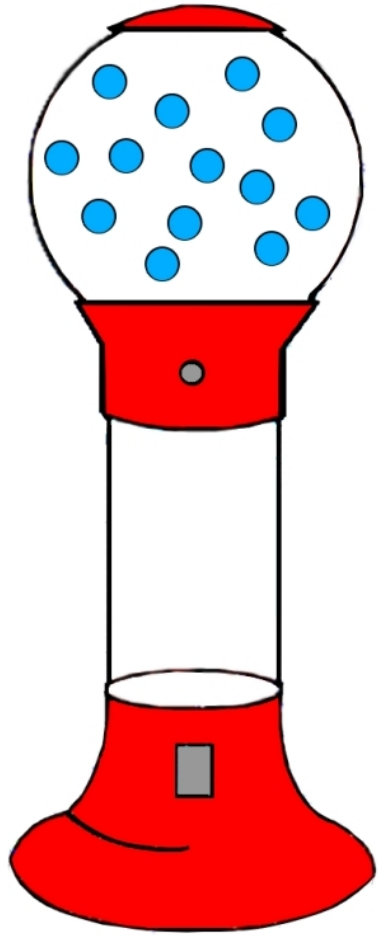
The embedded NP must be definite/specific/given in discourse.

Assessing meaning differences

(9) Alex ate some of the cashews

Partitive Constraint
Jackendoff, 1977, Ladusaw, 1982, Reed, 1991

The embedded NP must be definite/specific/given in discourse.



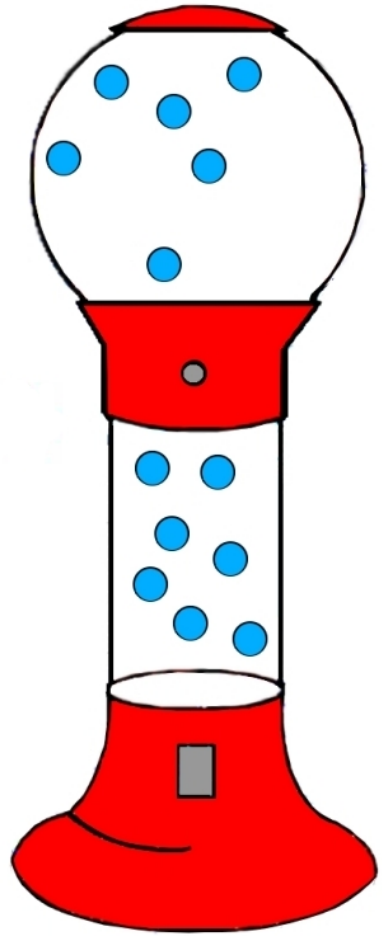
Assessing meaning differences

(9) Alex ate some of the cashews

(10) You got some of the gumballs.

Partitive Constraint
Jackendoff, 1977, Ladusaw, 1982, Reed, 1991

The embedded NP must be definite/specific/given in discourse.



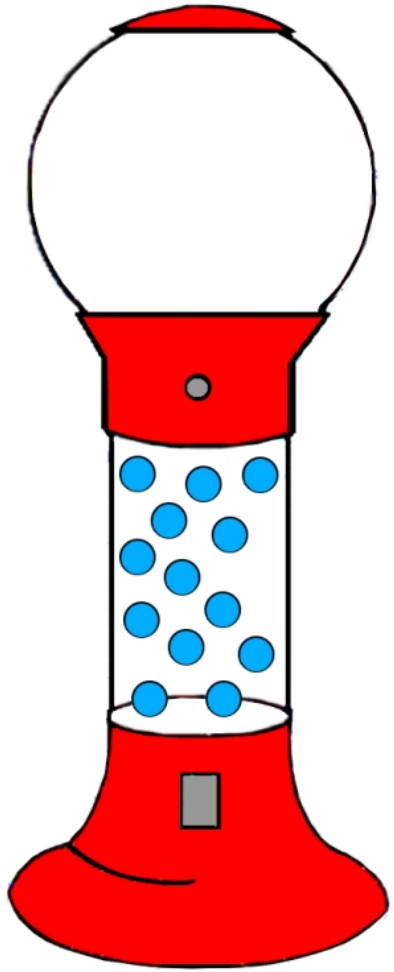
Assessing meaning differences

(9) Alex ate some of the cashews

(10) You got some of the gumballs.

Partitive Constraint
Jackendoff, 1977, Ladusaw, 1982, Reed, 1991

The embedded NP must be definite/specific/given in discourse.



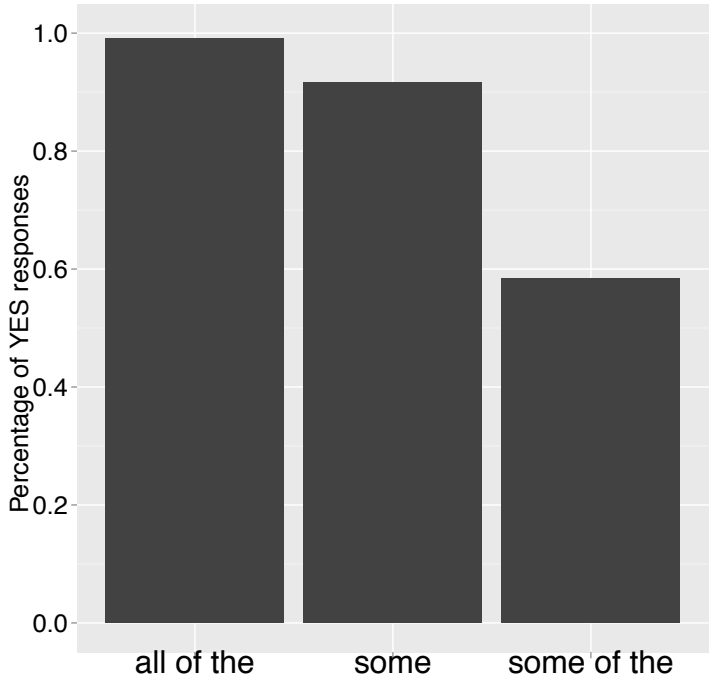
Assessing meaning differences

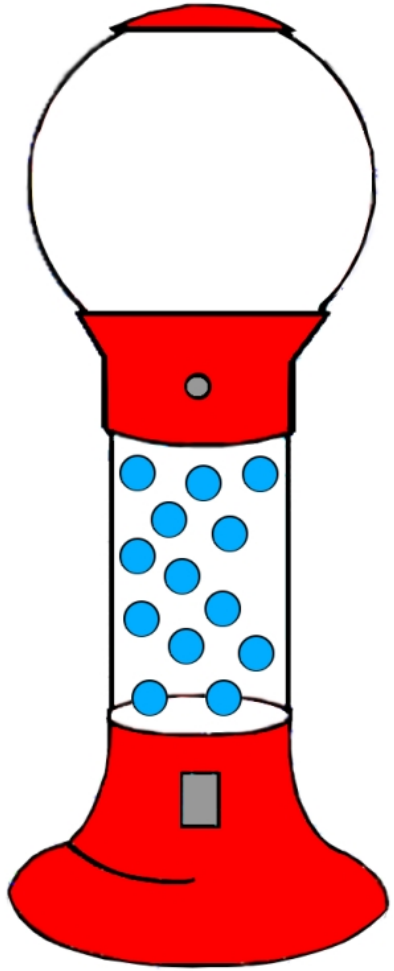
(9) Alex ate some of the cashews

(10) You got some of the gumballs.

Partitive Constraint
Jackendoff, 1977, Ladusaw, 1982, Reed, 1991

The embedded NP must be definite/specific/given in discourse.





Assessing meaning differences

- proxy for predicting meaning differences: discourse givenness

Assessing meaning differences

- proxy for predicting meaning differences: discourse givenness
- 3-level givenness annotation for 25% of Switchboard (Nissim et al., 2004):
  - *new*: NP referent is new to the discourse (not previously mentioned)
  - *mediated*: NP referent has not been explicitly mentioned, but can be inferred from context or is otherwise given
  - *given*: NP referent has been previously mentioned

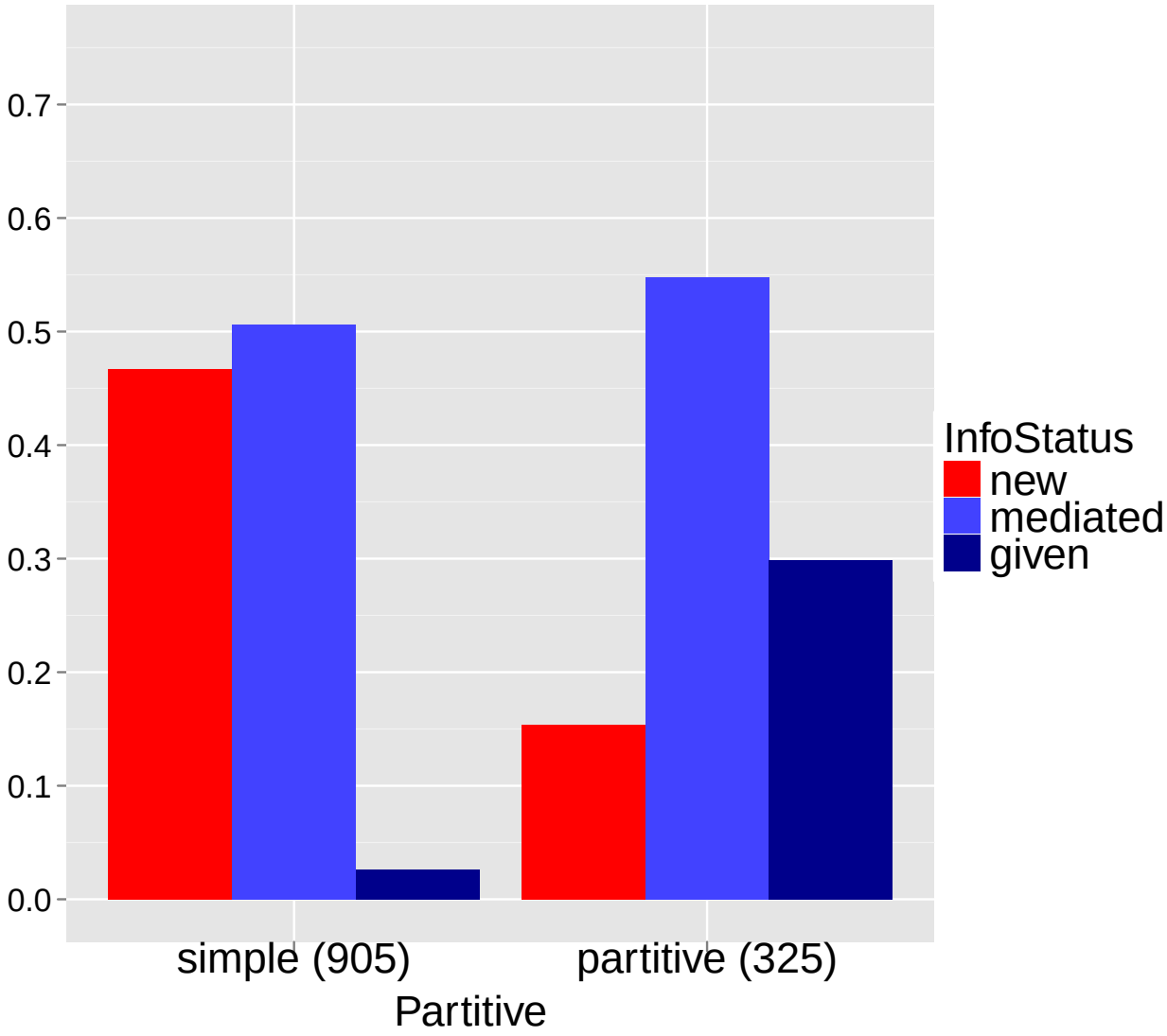
The dataset

- 1951 cases of *some*-NPs from the Switchboard
- givenness annotation for 1389 cases (336 partitives)

(11) uh, i, i think that in some cases that's considered to prejudice the current case

(12) and i think also some of the, uh, car companies are coming out with, uh, gas powered fleets

Givenness and partitive



partitive

336 *some of the X*

simple

1053 *some X*

Mixed logit model

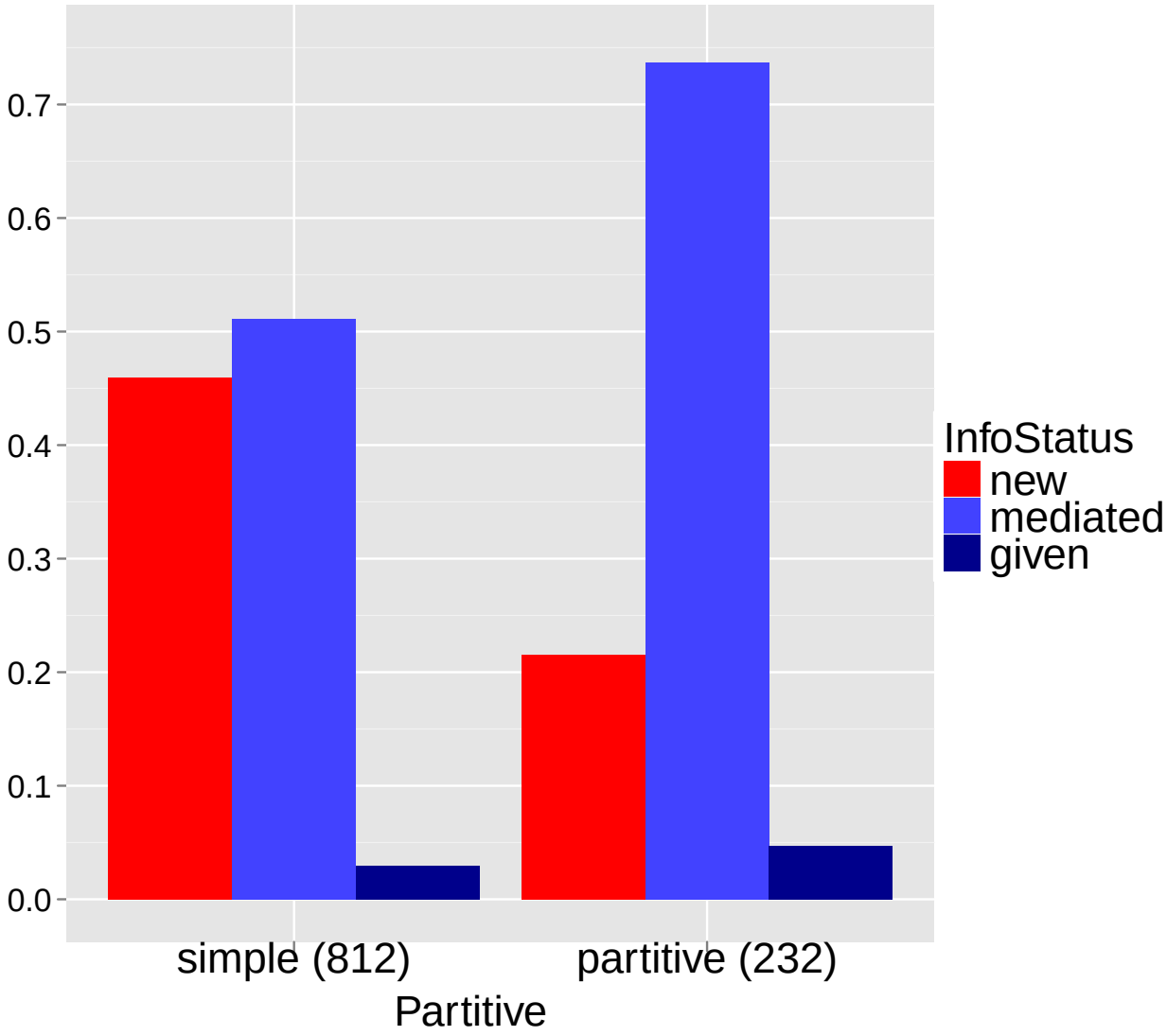
- partitive NPs more likely with more given referents ($p < .0001$)

Partial support for Reed, 1991

Data exclusion

- excluded cases that could only occur in one of the two forms:
  - if NP head was a pronoun (only partitive)
  - if NP head was a singular count noun (only indefinite use)
- leaves 1335 cases (269 partitives)
- givenness annotation for 1139 cases (233 partitives)

Givenness and partitive



After removing pronouns and singular count nouns:

partitive

233 *some of the X*

simple

906 *some X*

Model

partitive NP more likely with more given referents ($p < .0001$)

Testing robust communication

(13) Alex ate [$_{\text{OBJ}}$ SOME cashews]

Information estimates

Info(SOME | context) = -log (p(*some* | context) + p(*some of the* | context))

Testing robust communication

(13) Alex ate [$_{\text{OBJ}}$ SOME cashews]

Information estimates

Info(SOME | context) = -log (p(*some* | context) + p(*some of the* | context))

- information of SOME conditioned on 3 different contexts

Testing robust communication

(13) Alex ate [$_{\text{OBJ}}$ SOME cashews]

Information estimates

Info(SOME | context) = -log (p(*some* | context) + p(*some of the* | context))

- information of SOME conditioned on 3 different contexts

  Info(SOME | previous word)

Testing robust communication

(13) Alex ate [$_{\text{OBJ}}$ SOME cashews]

Information estimates

Info(SOME | context) = -log (p(*some* | context) + p(*some of the* | context))

- information of SOME conditioned on 3 different contexts

  Info(SOME | previous word)
  Info(SOME | NP head)

Testing robust communication

(13) Alex ate [$_{\text{OBJ}}$ SOME cashews]

Information estimates

Info(SOME | context) = -log (p(*some* | context) + p(*some of the* | context))

- information of SOME conditioned on 3 different contexts

  Info(SOME | previous word)
  Info(SOME | NP head)
  Info(SOME | NP grammatical function)

- givenness: *new* / *not new*

Testing robust communication

(13) Alex ate [$_{\text{OBJ}}$ SOME cashews]

Information estimates

Info(SOME | context) = -log (p(*some* | context) + p(*some of the* | context))

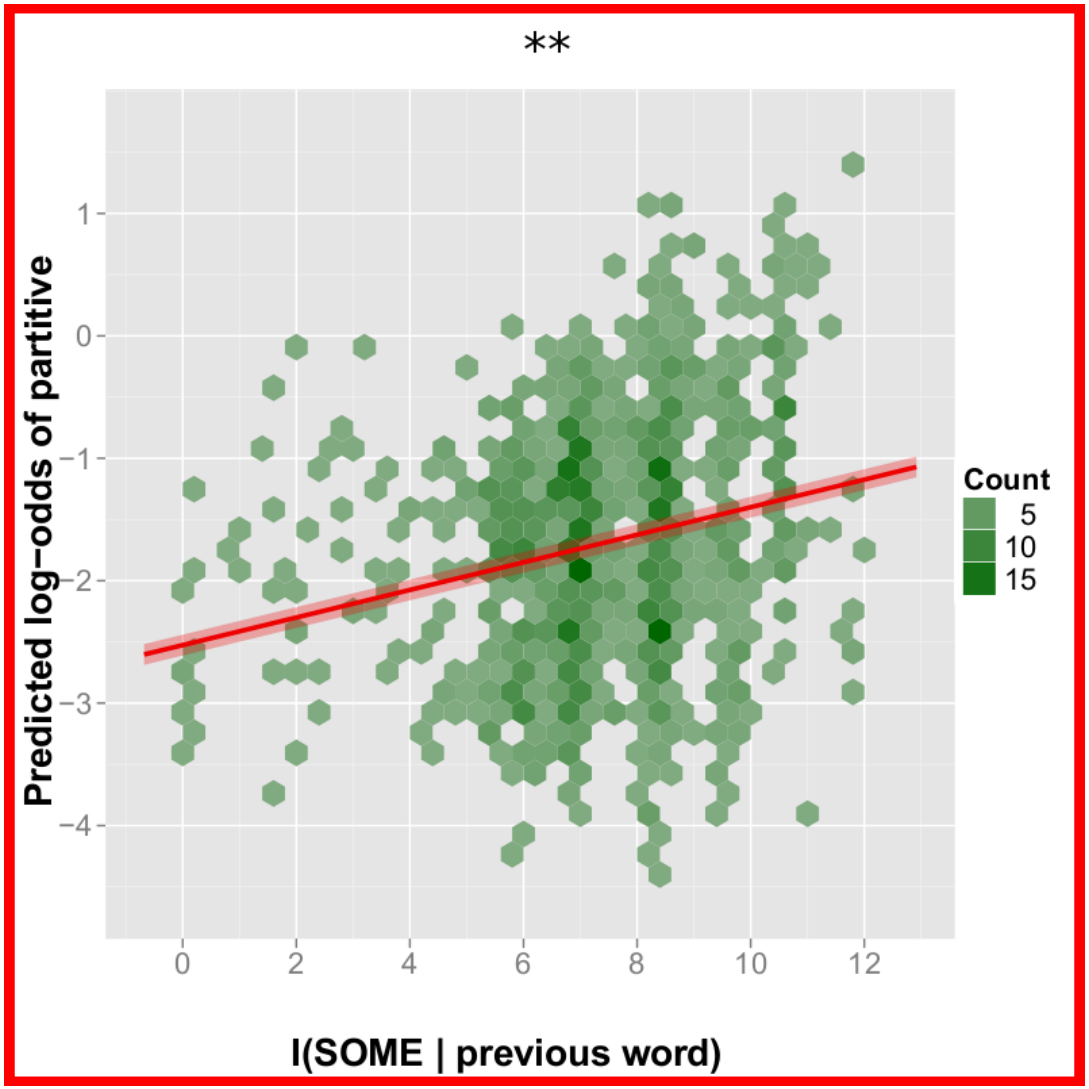
- information of SOME conditioned on 3 different contexts

  Info(SOME | previous word)
  Info(SOME | NP head)
  Info(SOME | NP grammatical function)

- givenness: *new* / *not new*
- mixed-effects logistic regression predicting partitive vs. simple *some*

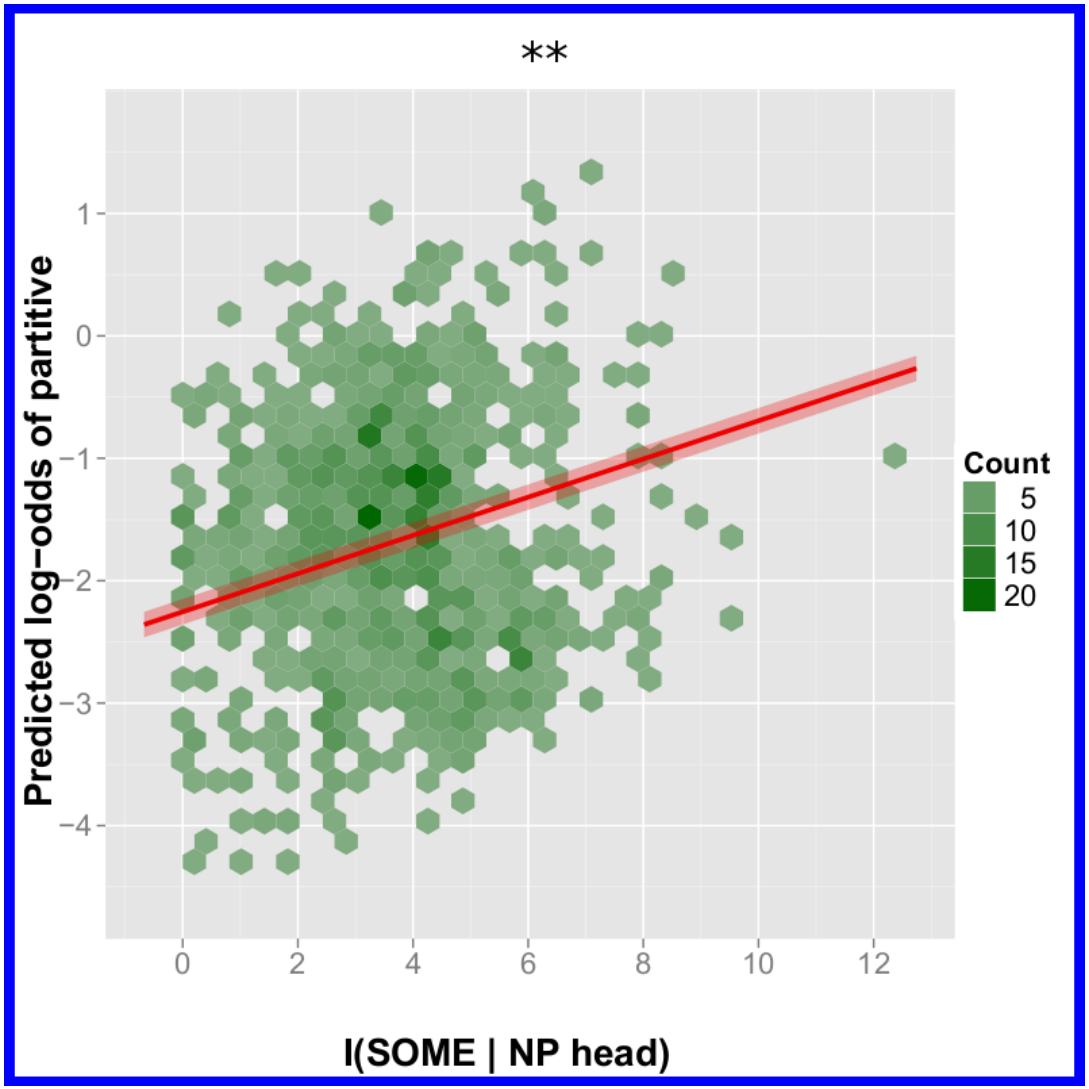
Bigram predictability

Alex ate [$_{\text{OBJ}}$ SOME cashews]



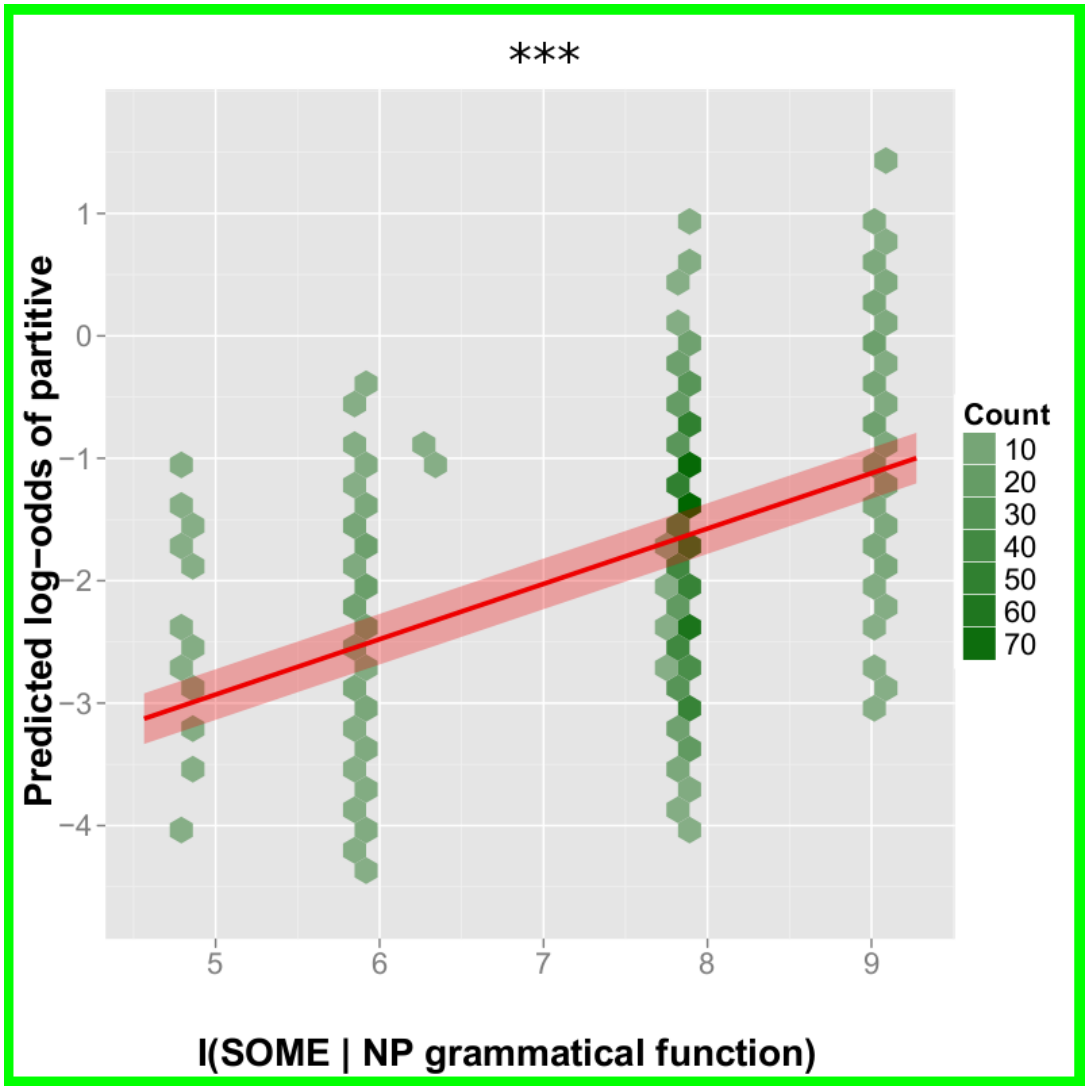
NP head

Alex ate [$_{\text{OBJ}}$ SOME cashews]



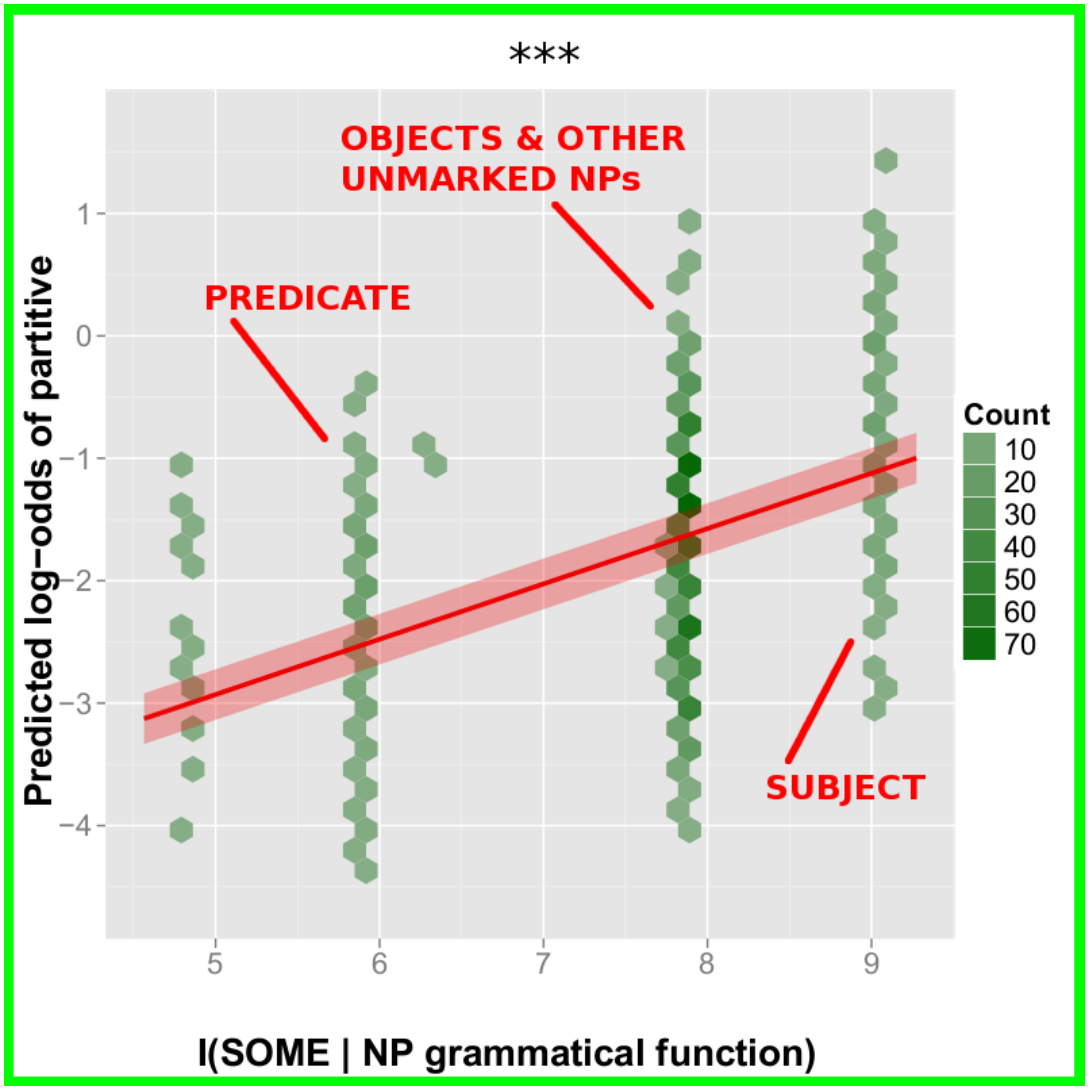
Grammatical function

Alex ate [$_{\text{OBJ}}$ SOME cashews]



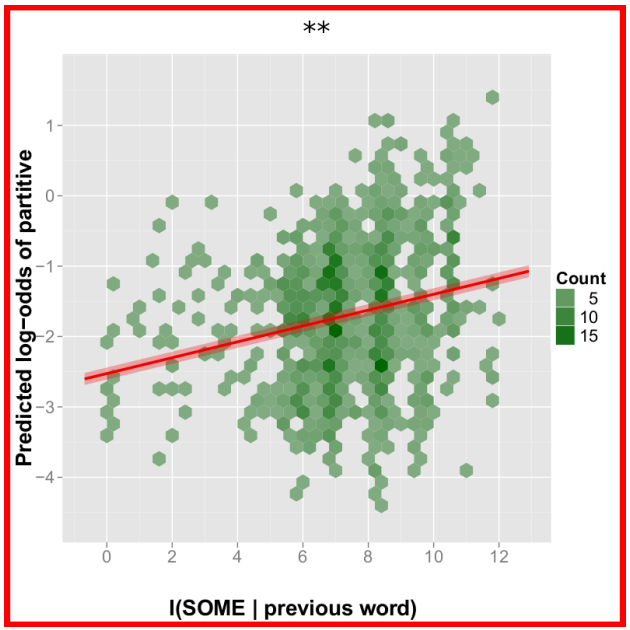
Grammatical function

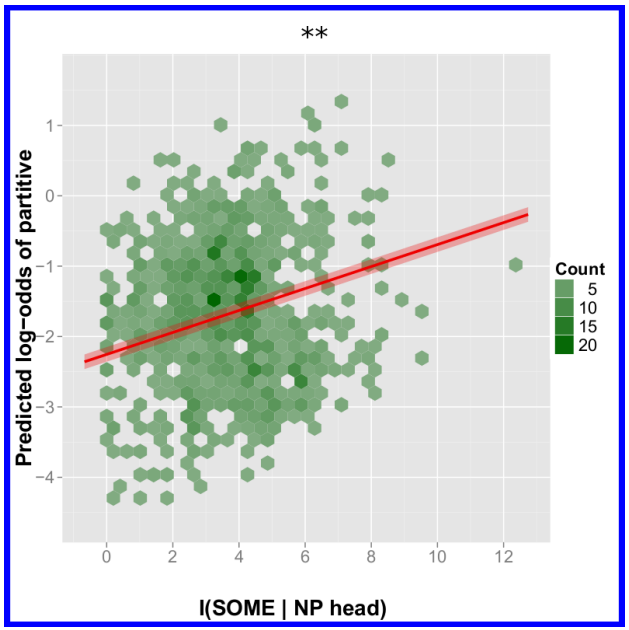
Alex ate [$_{\text{OBJ}}$ SOME cashews]

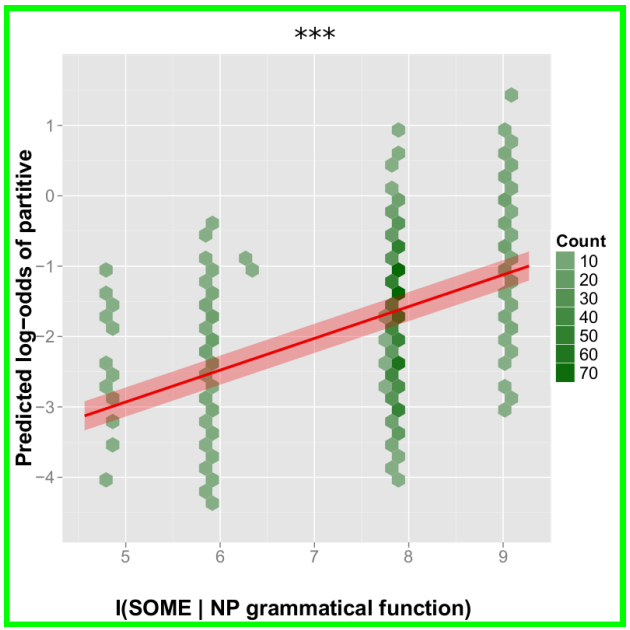


Results

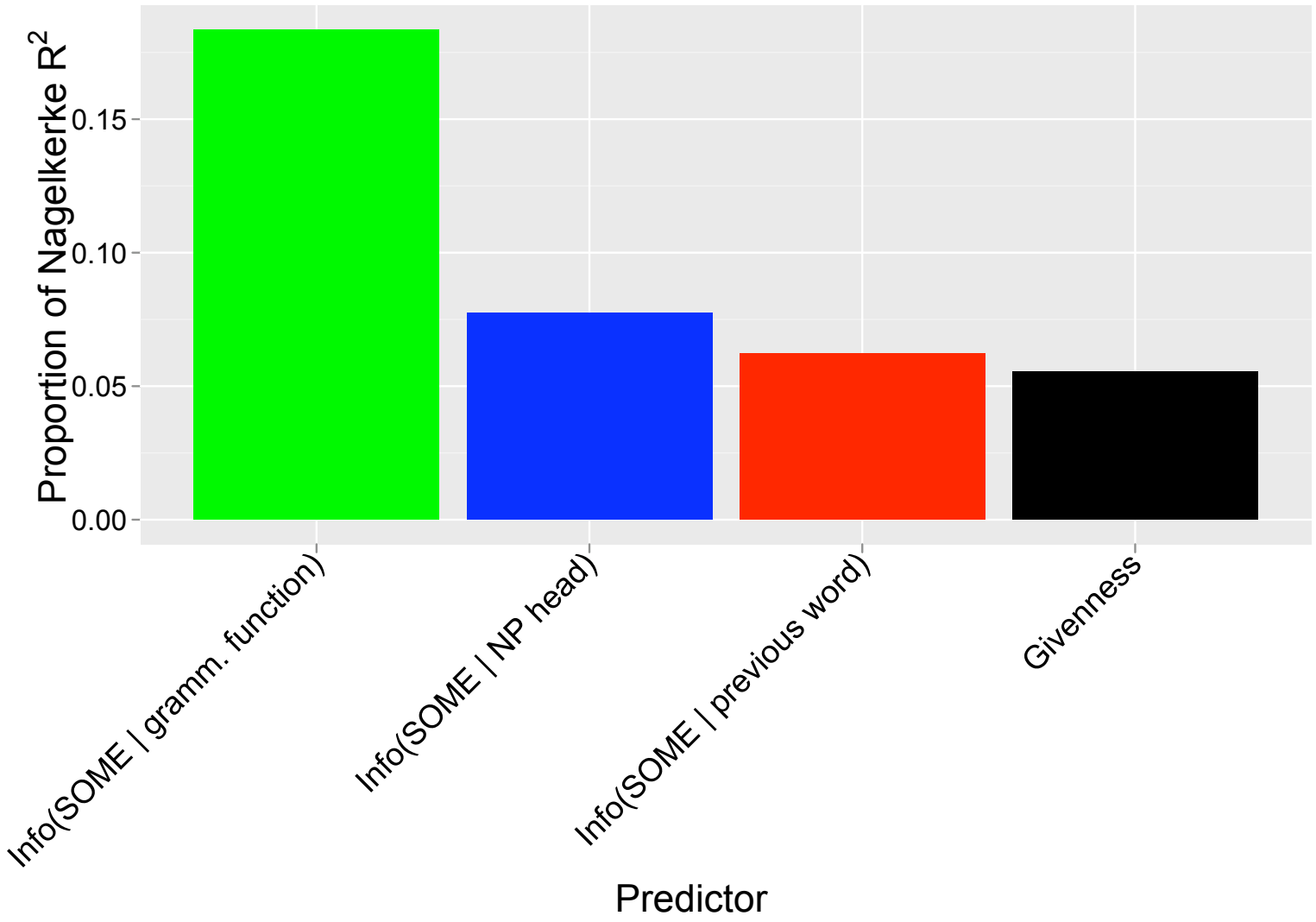
Alex ate [$_{\text{OBJ}}$ SOME cashews]







Strength of predictors



Summary

- the more information carried by SOME the more likely the partitive
- meaning differences captured by givenness contribute less to the choice of partitive vs. simple *some* in our dataset than informational measures

Conclusion

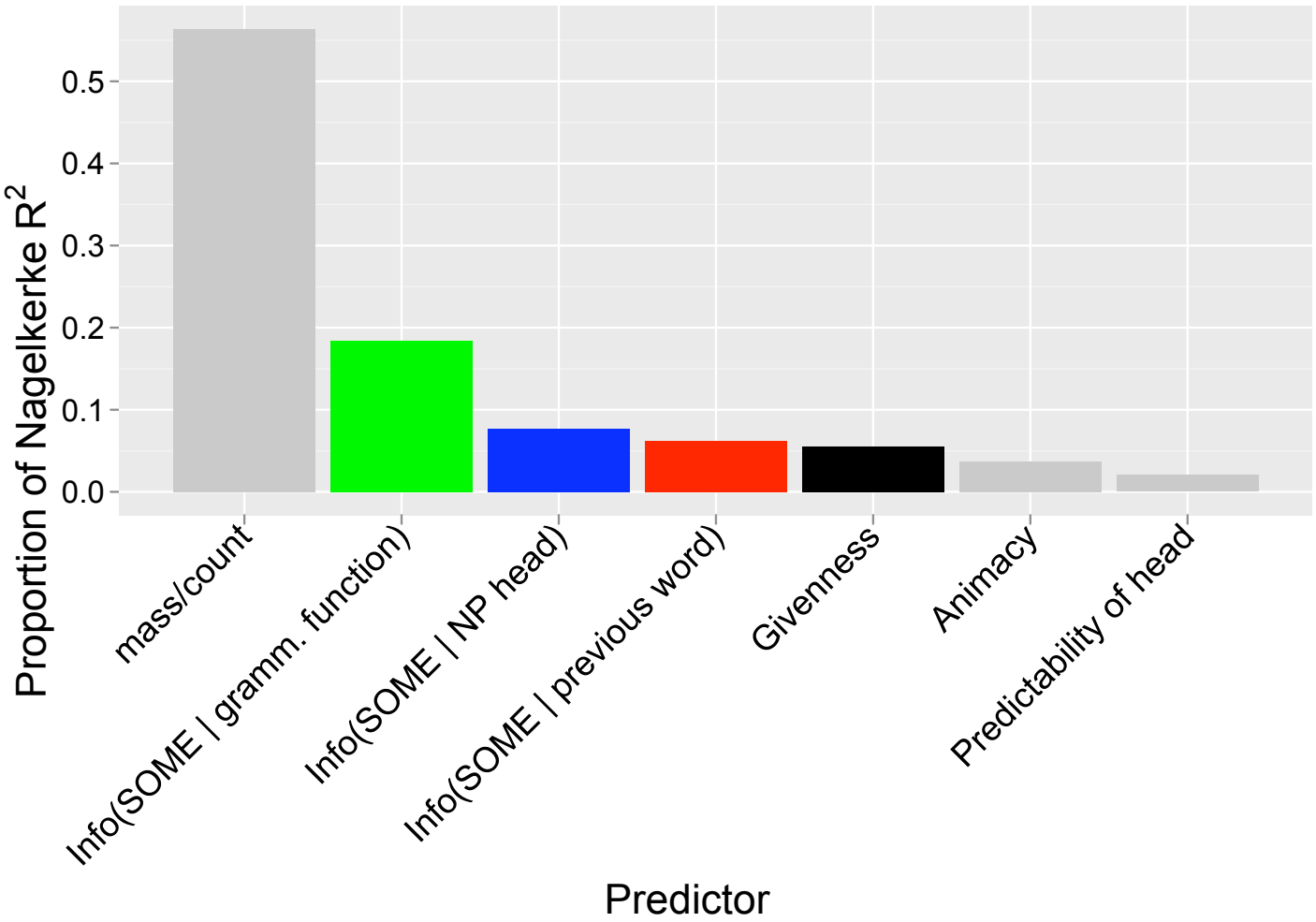
Robust communication and precision in encoding intended meaning

Rather than applying after the restriction of permissible forms by semantics, the pressure to robustly communicate a core meaning applies in parallel with the goal to find the most precise form to encode an intended meaning, and may sometimes override it if the benefit of robust communication outweighs the potential cost of transmitting more information.

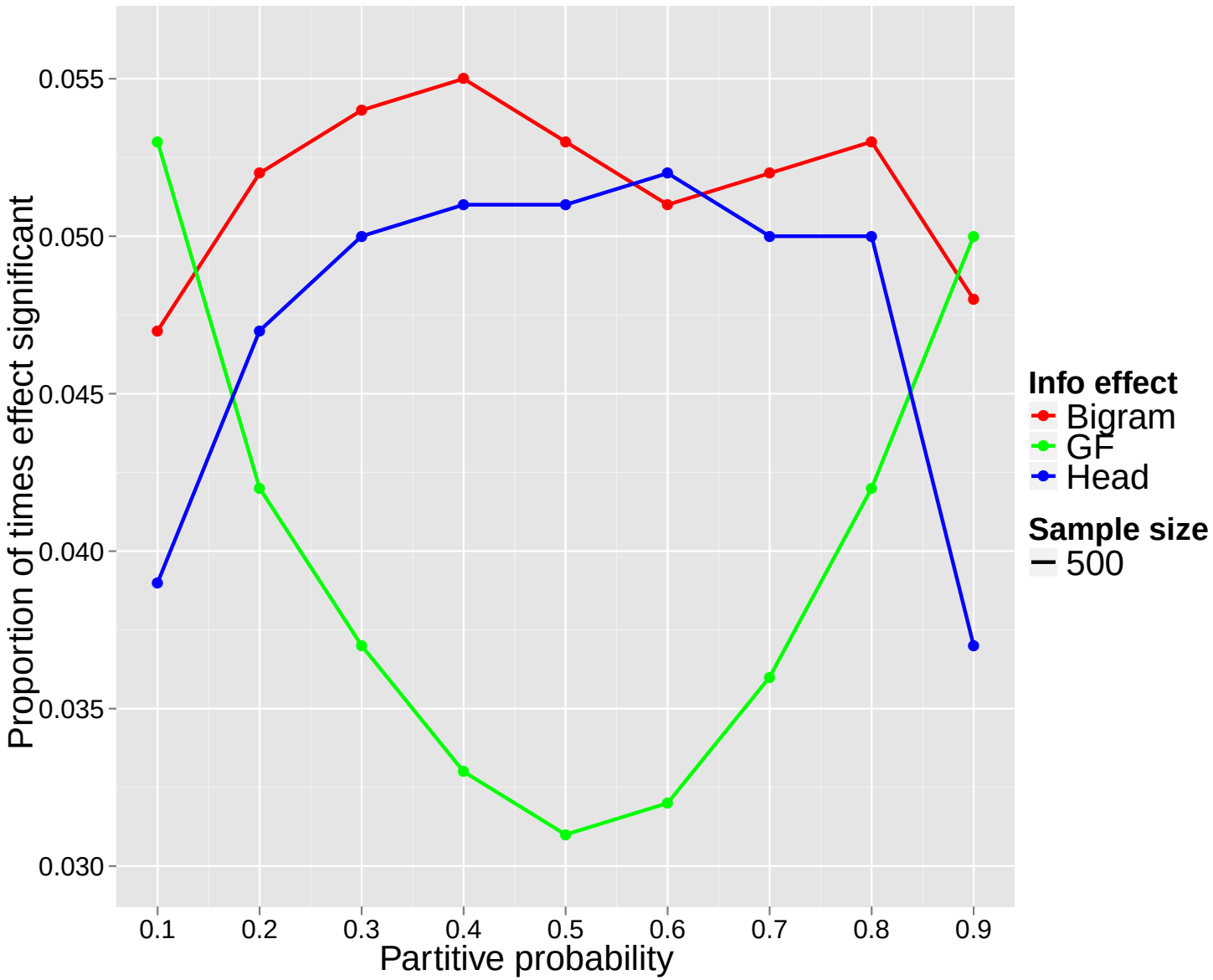
Thanks to

- Chelsea Marsh and Joe Benincasa for givenness annotation
- HLPlab
- Tanenhaus lab

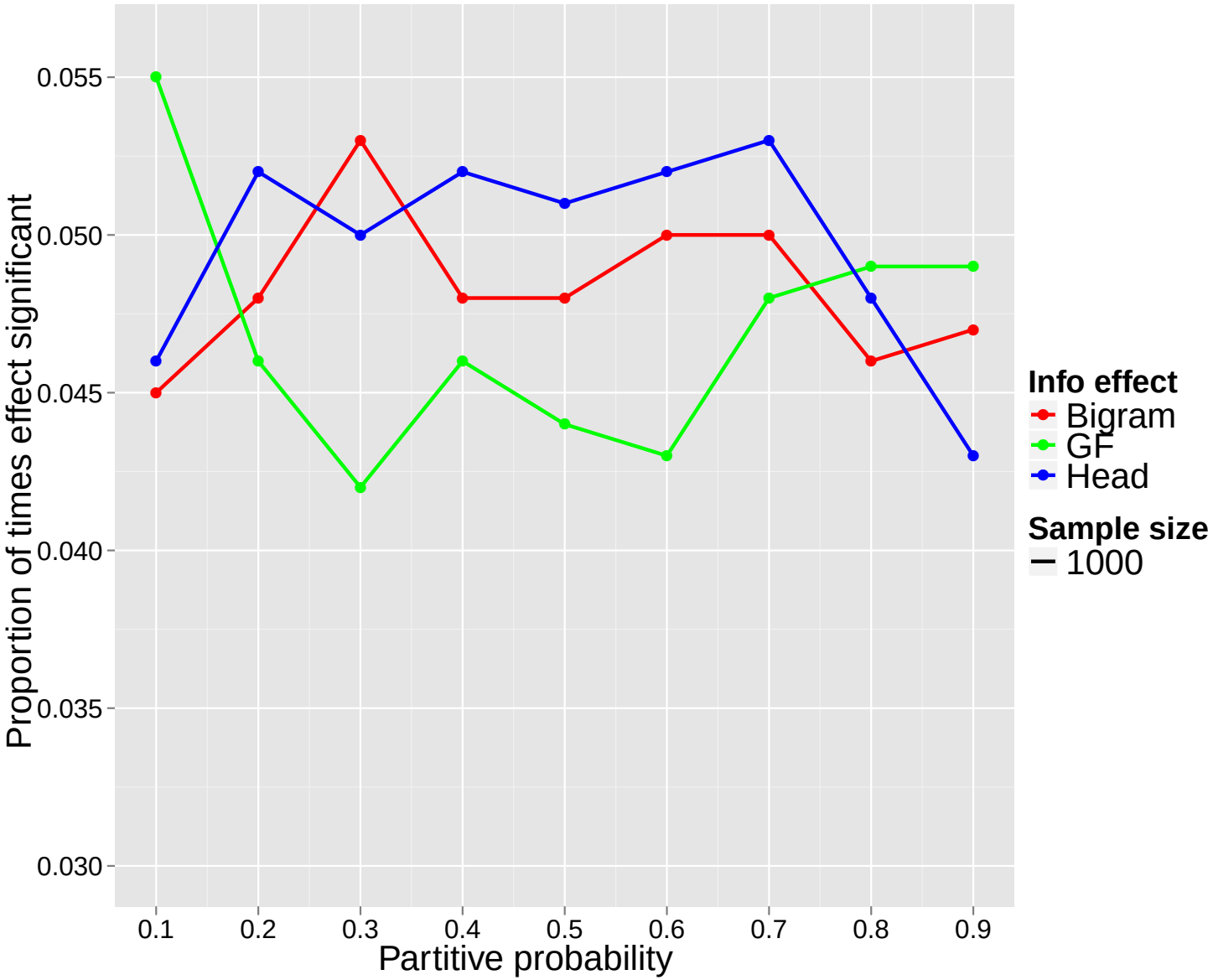
Strength of predictors



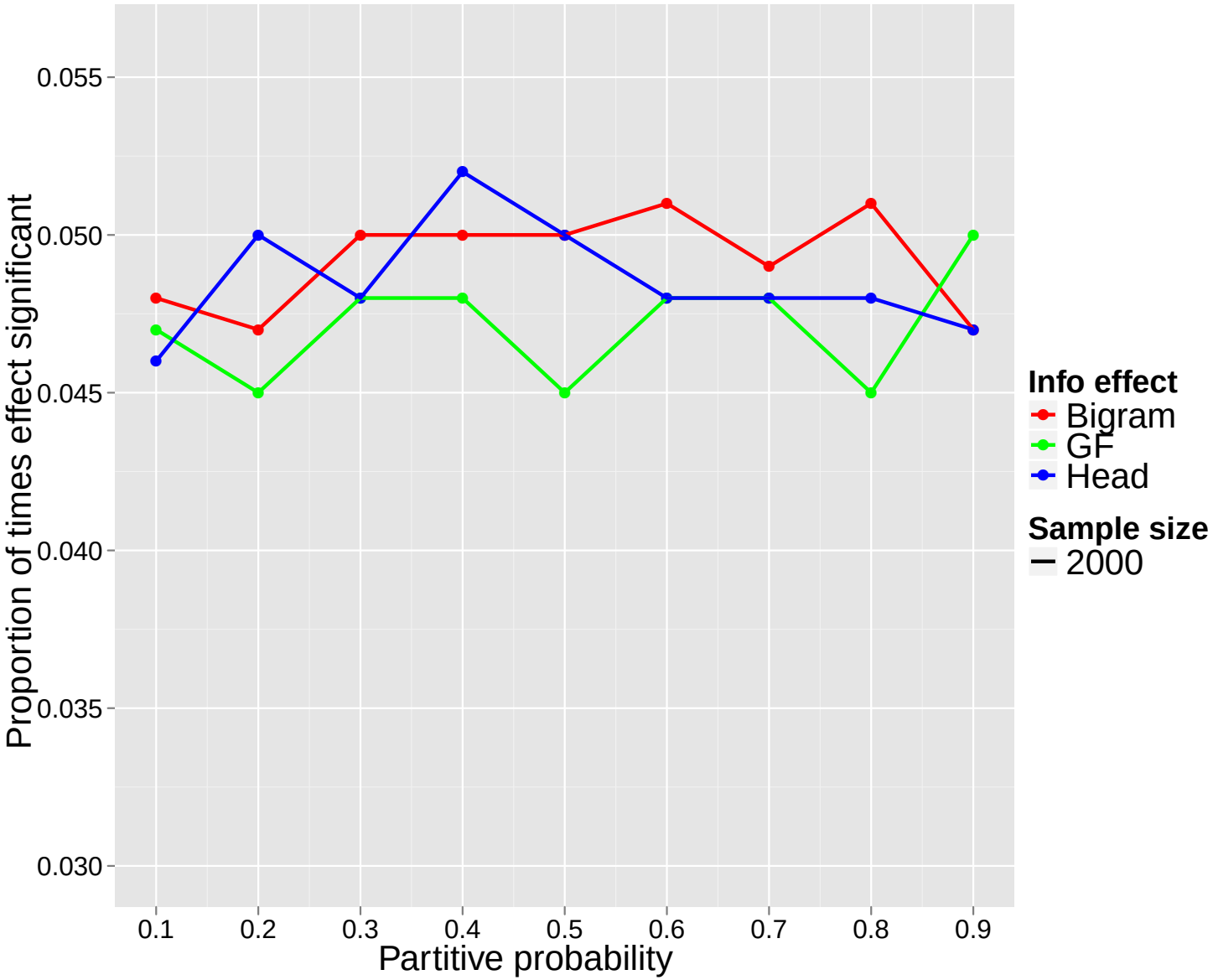
Simulation results



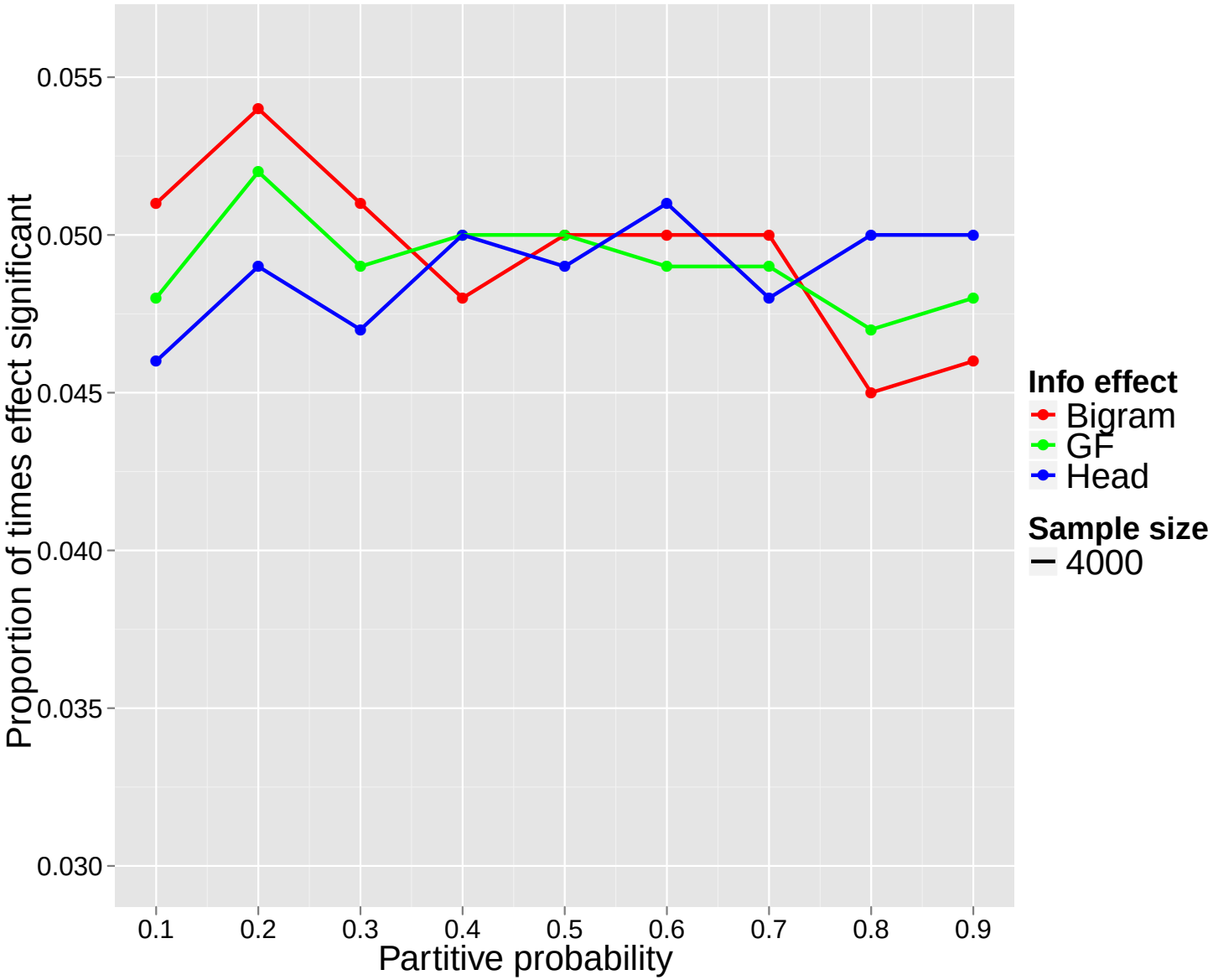
Simulation results



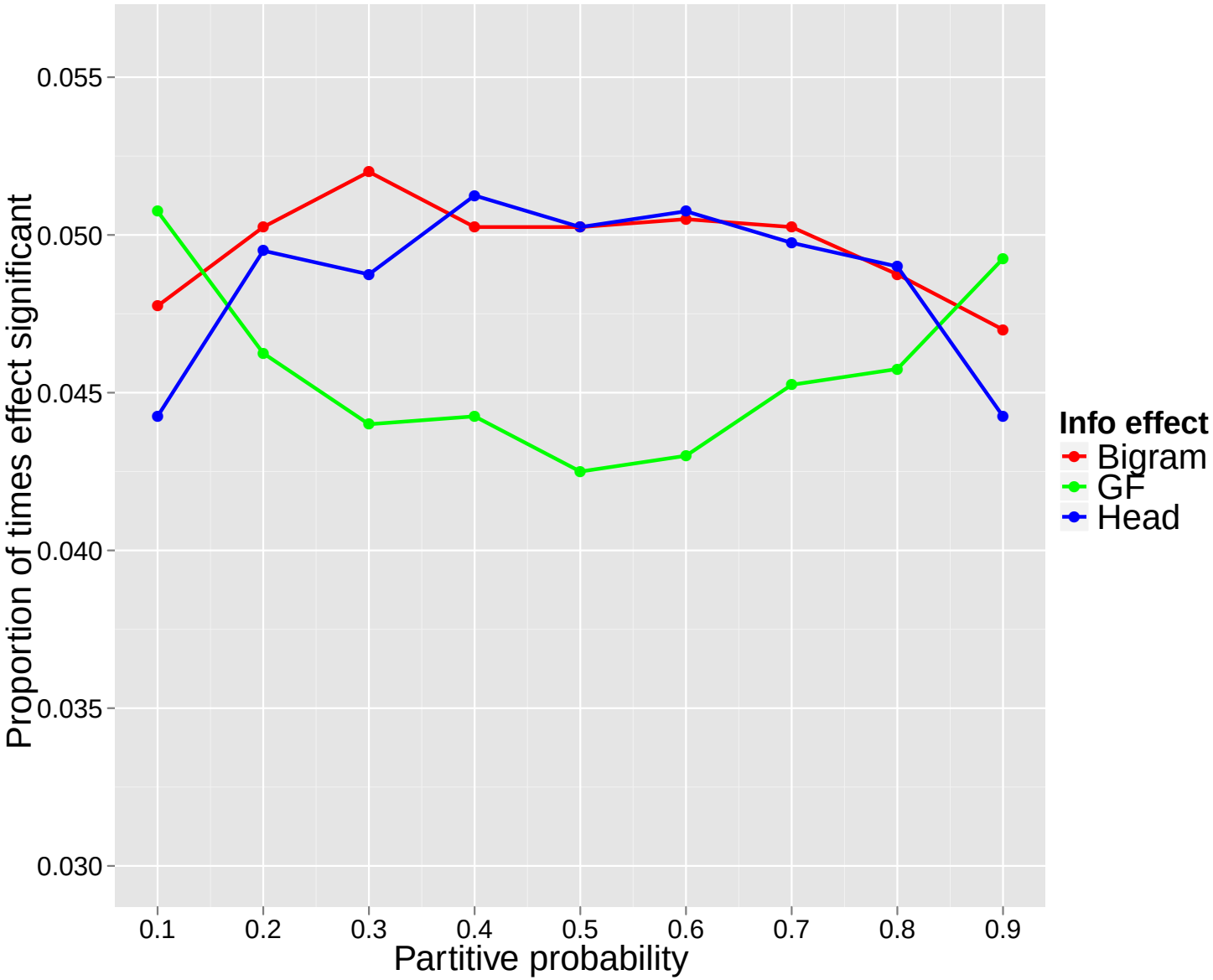
Simulation results



Simulation results



Simulation results



References I

Aylett, M. P., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*, *47*(1), 31 - 56.

Ferreira, V. S., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology*, *40*(4), 296 - 340.

Frank, A., & Jaeger, T. F. (2008). Speaking rationally: Uniform information density as an optimal strategy for language production. In *The 30th annual meeting of the cognitive science society.*

Gallo, C. G., Jaeger, T. F., & Smyth, R. (2008). Incremental syntactic planning across clauses. In *The 30th annual meeting of the cognitive science society* (p. 933 - 938).

References II

Genzel, D., & Charniak, E. (2002). Entropy rate constancy in text. In *Proceedings of the association of computational linguistics* (p. 199 - 206). Philadelphia, PA.

Jackendoff, R. (1977). *X-bar syntax: a study of phrase structure*. Cambridge, MA: MIT Press.

Jaeger, T. F. (2006). *Redundancy and syntactic reduction in spontaneous speech*. Unpublished doctoral dissertation, Stanford University.

Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology*, *61*(1), 23 - 62.

Ladusaw, W. (1982). Semantic constraints on the english partitive constructions. In D. Flickinger, M. Macken, & N. Wiegand (Eds.), *Proceedings of wccfl 1* (p. 231 - 242). Stanford: Stanford Linguistics Association.

Levinson, S. C. (2000). *Presumptive meanings - the theory of generalized conversational implicature*. The MIT Press.

References III

Levy, R., & Jaeger, T. F. (2007). Speakers optimize information density through syntactic reduction. In B. Schlökopf, J. Platt, & T. Hoffmann (Eds.), *Advances in neural information processing systems (nips) 19* (p. 849 - 856). Cambridge, MA: MIT Press.

Nissim, M., Dingare, S., Carletta, J., & Steedman, M. (2004). An annotation scheme for information status in dialogue. In *Fourth language resources and evaluation conference.* Lisbon, Portugal.

Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*.

Reed, A. M. (1991). On interpreting partitives. In D. J. Napoli & J. A. Kegl (Eds.), *Bridges between psychology and linguistics: a Swarthmore festschrift for Lila Gleitman.* L. Erlbaum Associates.

Shannon, C. (1948). A mathematical theory of communication. *Bell Systems Technical Journal*, *27*(4), 623 - 656.

References IV

Zipf, G. K. (1949). *Human behavior and the principle of least effort: An introduction to human ecology*. Addison-Wesley.