# Learning biases may prevent lexicalization of pragmatic inferences

## A case study combining iterated (Bayesian) learning and functional selection

Thomas Brochhagen[1] Michael Franke [2] Robert van Rooij [1]

CogSci 2016, Philadelphia, USA, 2016.08.13

[1]ILLC, University of Amsterdam

[2]Department of Linguistics, University of Tübingen

## Main methodological contribution

- interaction of: (replicator mutator dynamics)
  - fitness-relative replication (replicator dynamics)
  - iterated learning (mutator dynamics)
- learners perform joint-inference:
  - type of pragmatic behavior
    RSA-style probabilistic types (Frank & Goodman 2012)
  - lexical meaning
    LOT-style learning biases (Piantadosi et al. under review)

## Case study on scalar implicatures

- model shows prevalence of:
  - Gricean pragmatic use
  - non-lexicalized upper-bounds

· M. C. Frank and N. D. Goodman. Predicting pragmatic reasoning in language games. ***Science*, 336(6084):998–998, 2012**

· Steven T. Piantadosi, Joshua B. Tenenbaum, and Noah D. Goodman. Modeling the acquisition of quantifier semantics: a case study in function word learnability, under review

# The semantics-pragmatics distinction

Semantics

Literal meaning (truth-conditional)

Pragmatics

Information beyond literal meaning (e.g. defeasible inferences)

# Scalar inferences

(1) $\langle$*some*, *many*, *most*, *all*$\rangle$

a. All students came to class
$\rightarrow$ Some students came to class

b. Some students came to class
$\rightsquigarrow$ Not all students came to class

(2) $\langle$*may*, *should*, *must*$\rangle$

(3) $\langle$*one*, *two*, *three*, ...$\rangle$

(4) $\langle$*or*, *and*$\rangle$

(5) ...

The use of a less informative expression when a more informative one could have been used* can license a defeasible inference that stronger alternatives do not hold

*The hearer assumes the speaker to be knowledgeable and cooperative


· Laurence R. Horn. *On the Semantic Properties of Logical Operators in English*.
**Indiana University Linguistics Club, Bloomington, IN, 1972**

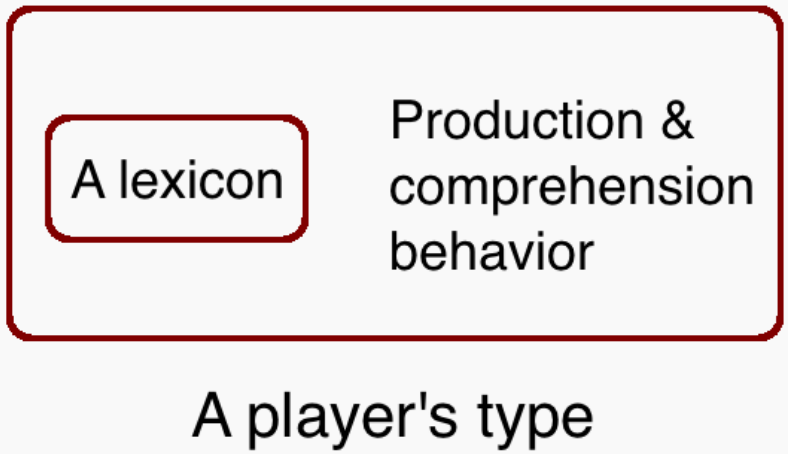
· Gerald Gazdar. *Pragmatics, Implicature, Presuposition and Logical Form*.
**Academic Press, New York, 1979**

· Paul Grice. Logic and conversation.
**In *Studies in the Ways of Words*, chapter 2, pages 22–40. Harvard University Press, Cambridge, MA, 1975**

1. Why are (pragmatically inferred) upper-bounds of weak(er) alternatives not part of semantics?
2. What justifies semantic structure in light of pragmatic enrichment?

# Model

# Components I: Probabilistic (pragmatic) language users




· Anton Benz, Gerhard Jäger, Robert Van Rooij, and Robert Van Rooij, editors. *Game theory and pragmatics*. **Springer, 2005**

· Leon Bergen, Roger Levy, and Noah D Goodman. Pragmatic reasoning through semantic inference. ***Semantics and Pragmatics*, 2016**

· M. C. Frank and N. D. Goodman. Predicting pragmatic reasoning in language games. ***Science*, 336(6084):998–998, 2012**

· Michael Franke and Gerhard Jäger. Pragmatic back-and-forth reasoning. ***Semantics, Pragmatics and the Case of Scalar Implicatures.*, pages 170–200, 2014**

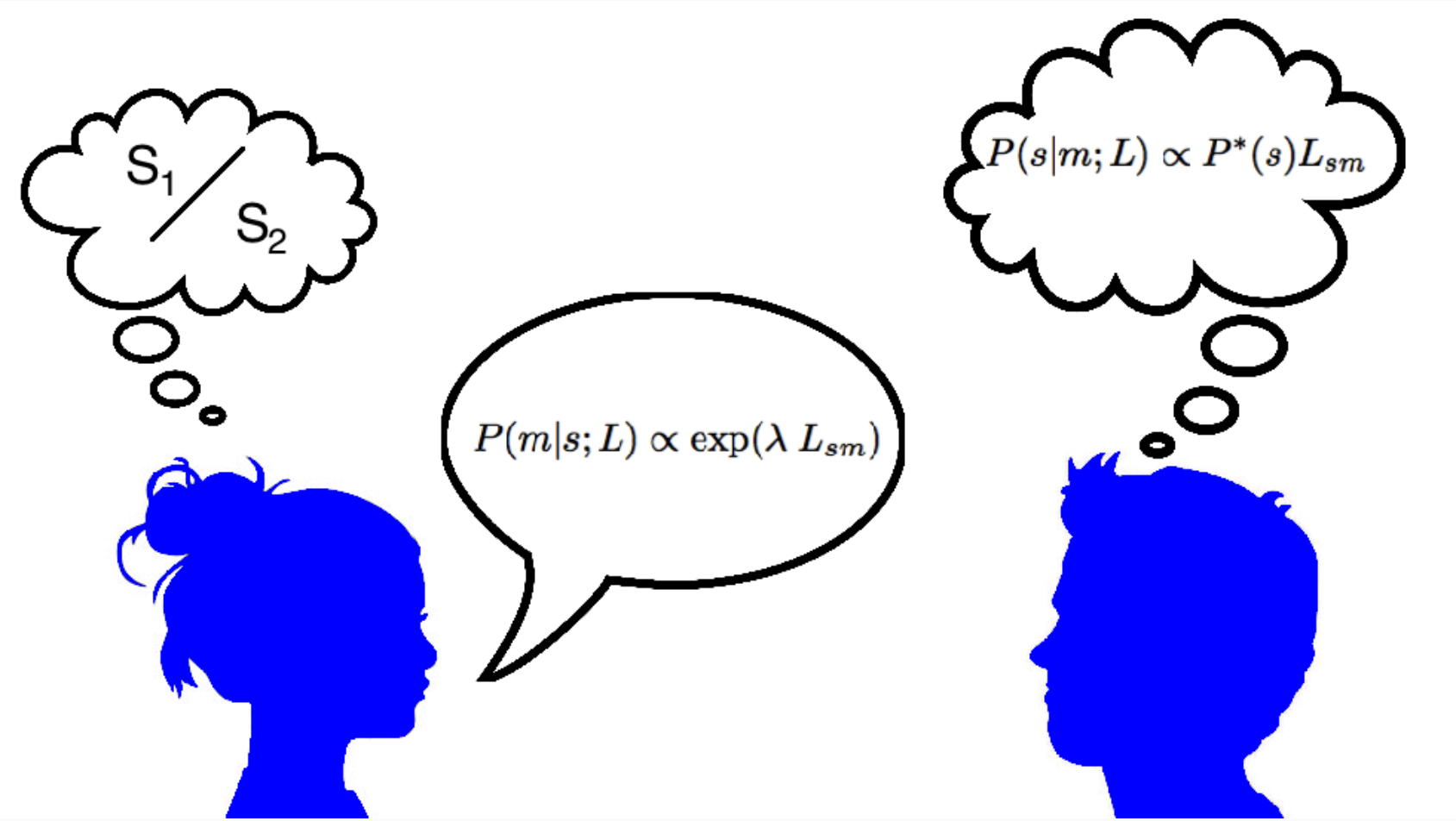
# Lexica

- $s_1$: Bill read some but not all books
- $s_2$: Bill read all books

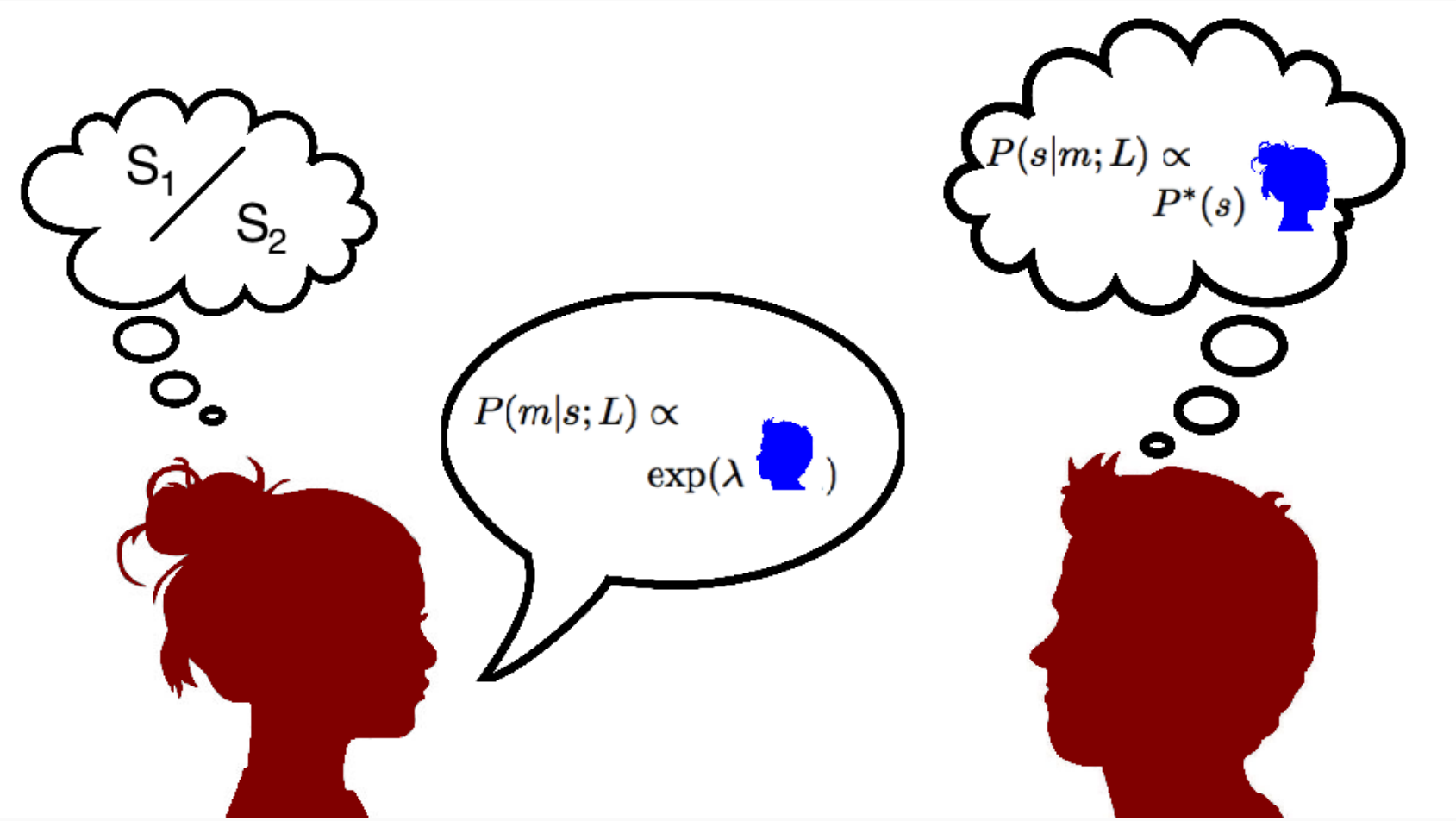
$$L_{\text{lack}} = \begin{array}{c} \\ s_1 \\ s_2 \end{array} \begin{array}{c} \begin{array}{cc} m_{\text{all}} & m_{\text{some}} \end{array} \\ \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \end{array}$$

$$L_{\text{bound}} = \begin{array}{c} \\ s_1 \\ s_2 \end{array} \begin{array}{c} \begin{array}{cc} m_{\text{all}} & m_{\text{some}} \end{array} \\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{array}$$

# Literal behavior



# Pragmatic behavior



# Components II: Cultural transmission

Two competing pressures:

1. Communicative efficiency

   ... as replicator dynamics; $\dot{x}$

2. Learnability

   ... iterated Bayesian learning as mutator dynamics; $Q$

} Replicator-mutator dynamics

$$\hat{x} = \dot{x} \cdot Q$$


· Thomas L. Griffiths and Michael L. Kalish. Language evolution by iterated learning with bayesian agents. ***Cognitive Science*, 31(3):441–480, 2007**

· M. A. Nowak and D. C. Krakauer. The evolution of language. ***Proceedings of the National Academy of Sciences*, 96(14):8028–8033, 1999**

# Functional pressure (replicator dynamics); $\dot{x}_i = \frac{x_i f_i}{\Phi}$

- Population of types $x$

  $x_i$ is the proportion of $t_i$ in $x$

- Fitness of type $i$

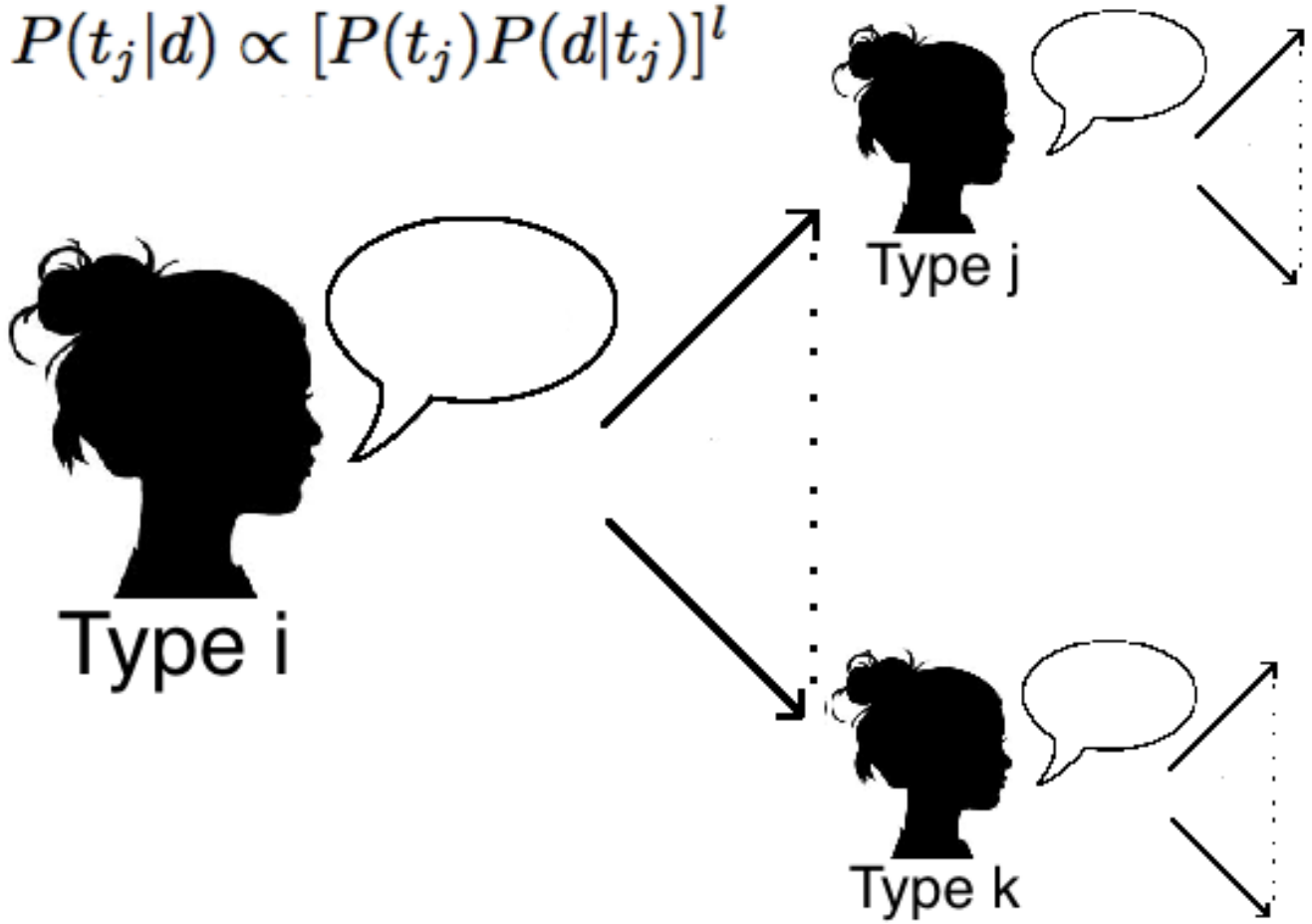
  $f_i = \sum_j x_j U(x_i, x_j)$

- Average fitness in the population

  $\Phi = \sum_i x_i f_i$

# Iterated learning (mutator dynamics)

$$Q_{ij} \propto \sum_d P(d|t_i)\ P(t_j|d)$$
$$P(t_j|d) \propto [P(t_j)P(d|t_j)]^l$$



# Analysis

# Lexica, signaling behavior & types

Lexica subset

$$L_{\text{tautology}} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad L_{\text{bound}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad L_{\text{lack}} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

Signaling behavior

*Literal* or *pragmatic*

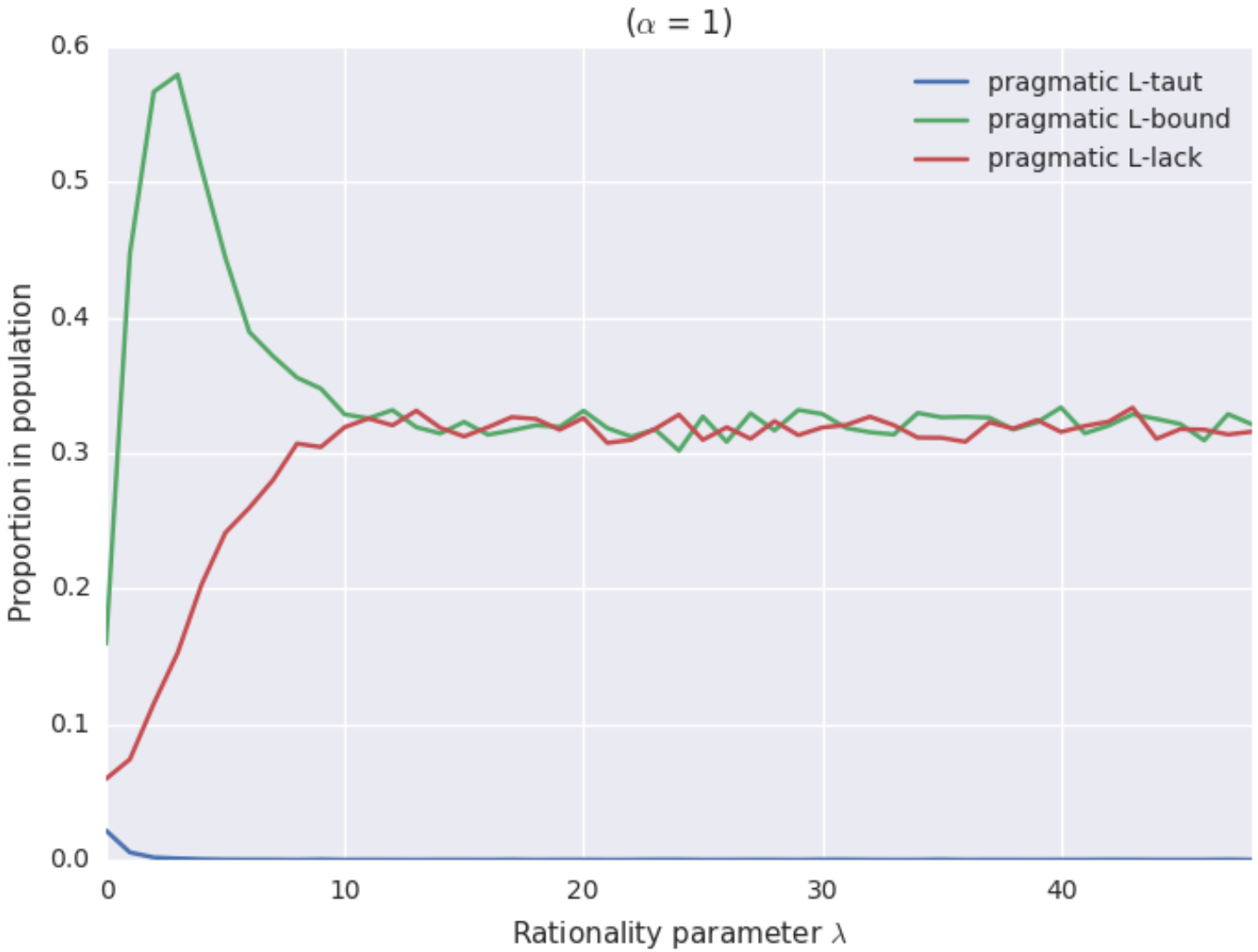
Types

12 types (2 behaviors $\times$ 6 lexica)

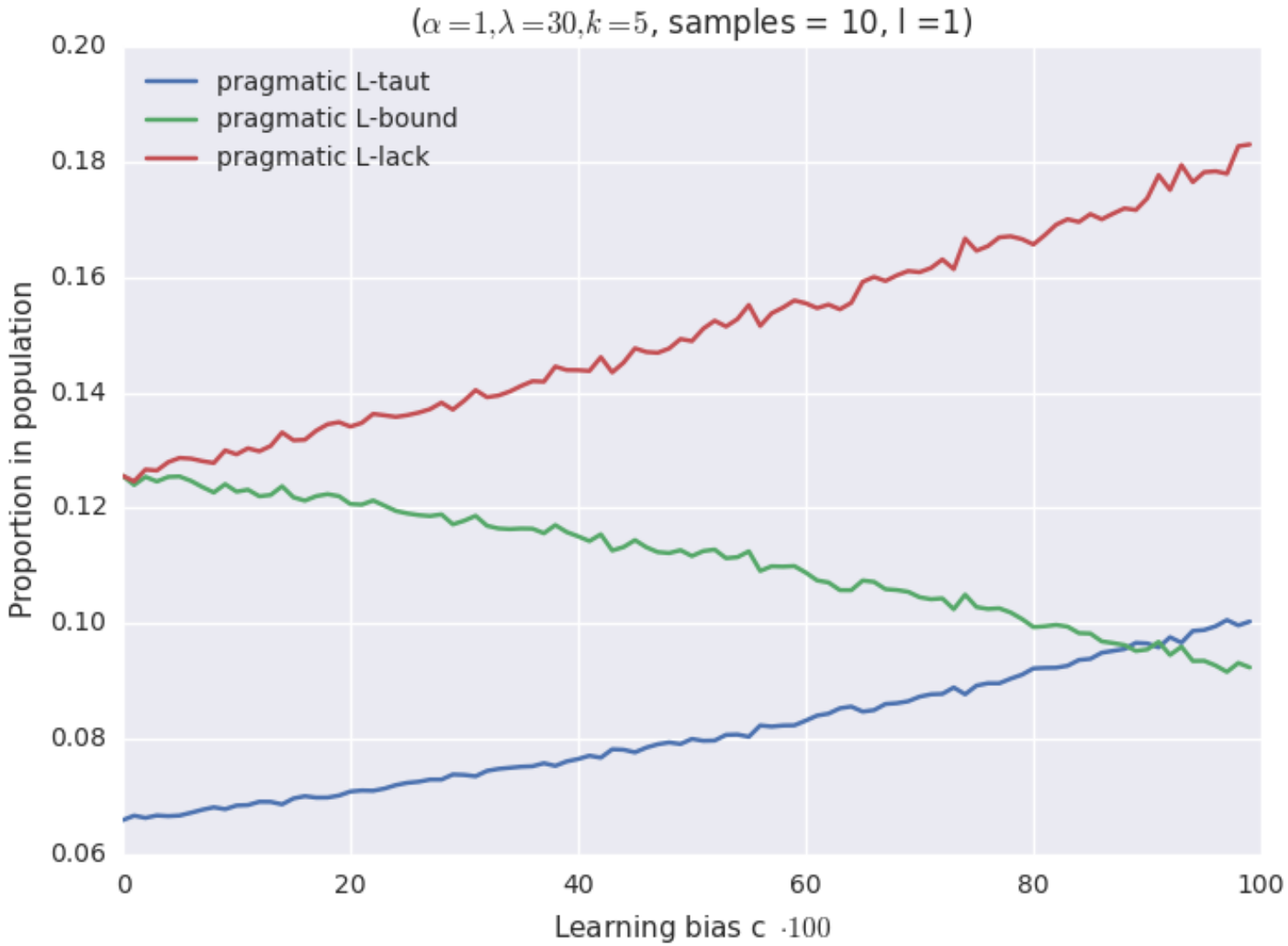
What factors lead to the selection of $L_{\text{lack}}$-like semantics?

| parameter | explanation | locus |
|---|---|---|
| $c \in [0, 1]$ | learning bias for upper-bound lack | $P(t_i)$ |
| $l \geq 1$ | sampling to MAP | $[P(t_i)P(d\|t_i)]^l$ |
| $\lambda \geq 1$ | rationality parameter | $\exp(\lambda R_{n-1}(s\|m; L))$ |
| $k = \|d\|$ | datum length | $P(d\|t_j)P(t_i\|d)$ |
| $\|D\|$ | data produced per parent type | $P(d\|t_j)P(t_i\|d)$ |

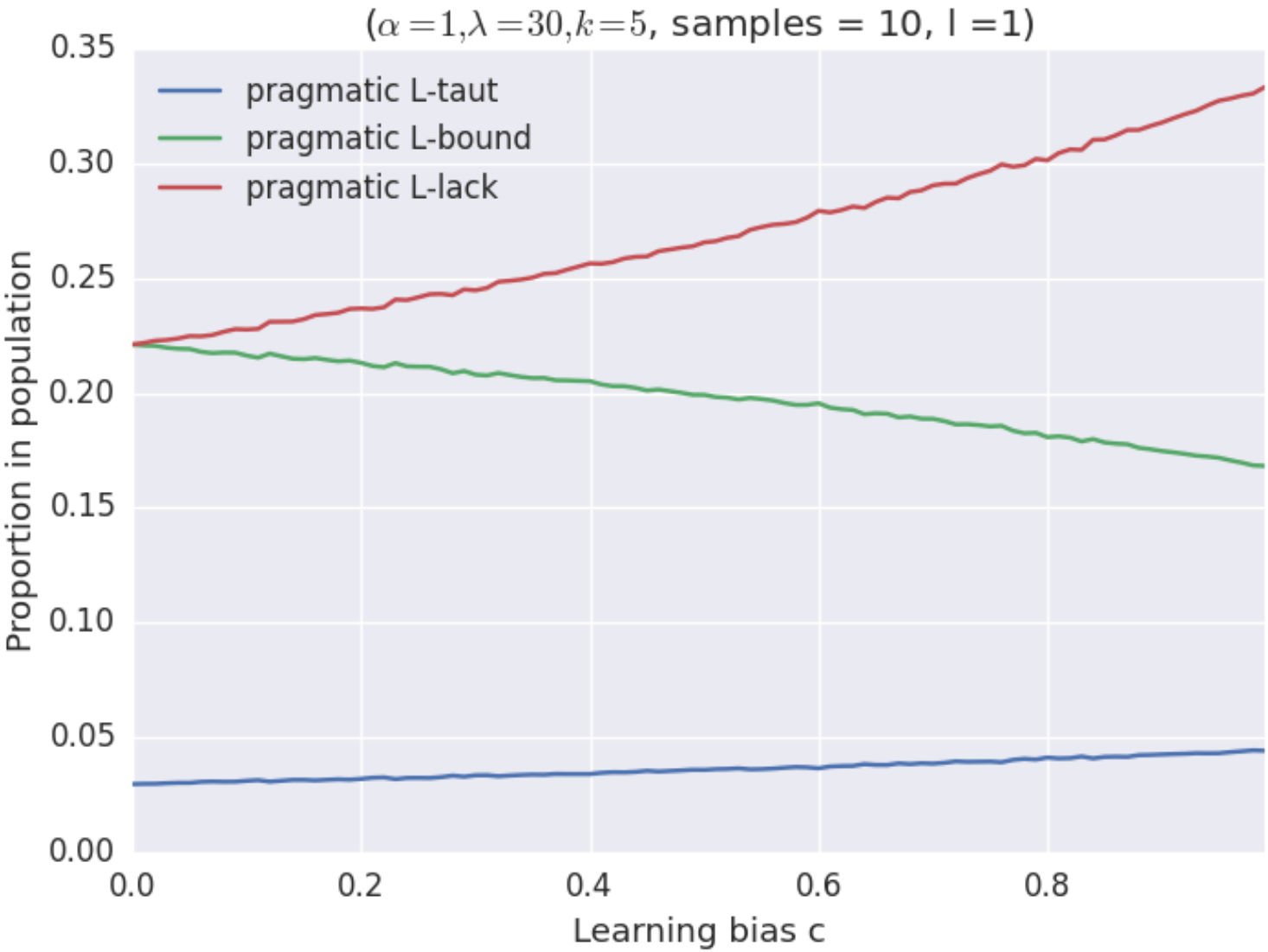
# Expressivity only



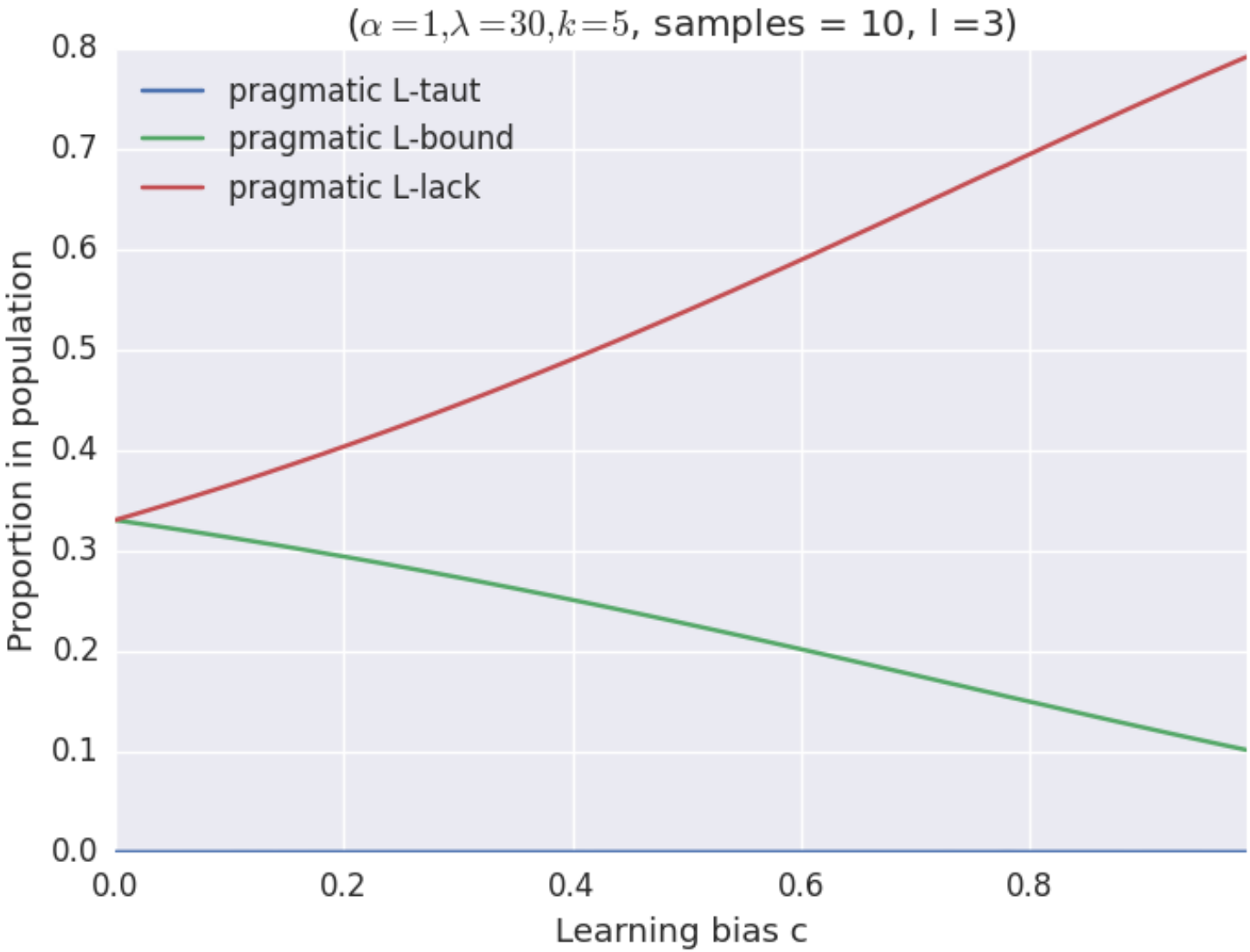
# Learnability only



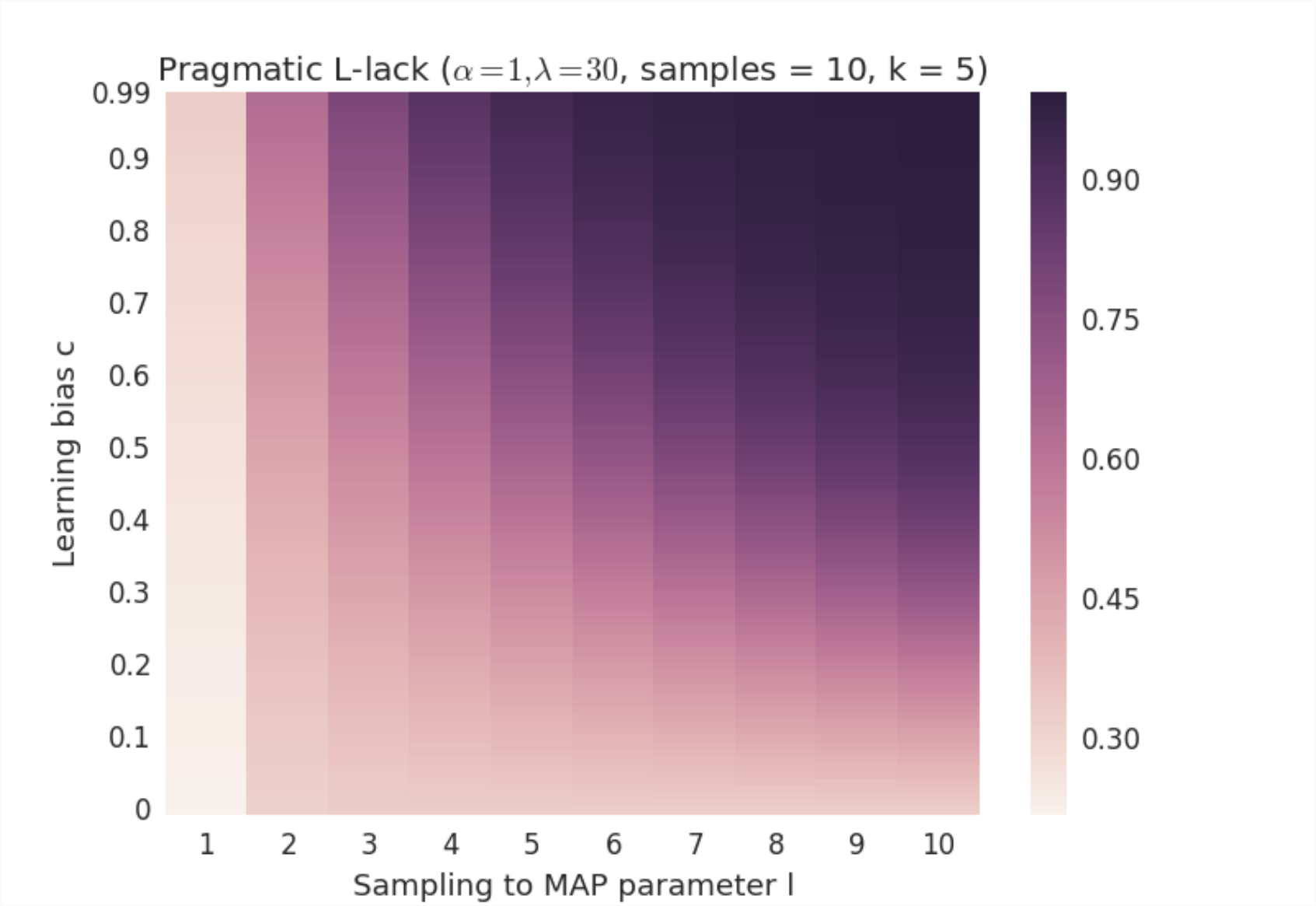
# Expressivity and learnability



# Effect of prior with higher posterior maximization



# Prior and posterior



# Concluding remarks: Application

- Learnability steers language towards simpler semantics
- Pragmatics compensates for potential loss in expressivity

} Lack of semantic upper-bounds

Provided

- Some degree of rationality in learning & choice

# Selection, learning, pragmatic use & lexical meaning

## Main methodological contribution

- interaction of: (replicator mutator dynamics)
  - fitness-relative replication (replicator dynamics)
  - iterated learning (mutator dynamics)
- learners perform joint-inference:
  - type of pragmatic behavior
  - lexical meaning

## Case study on scalar implicatures

- model shows prevalence of:
  - pragmatic use
  - non-lexicalized upper-bounds

# References I


[BJVRVR05] Anton Benz, Gerhard Jäger, Robert Van Rooij, and Robert Van Rooij, editors.
***Game theory and pragmatics.***
Springer, 2005.

[BLG16] Leon Bergen, Roger Levy, and Noah D Goodman.
**Pragmatic reasoning through semantic inference.**
*Semantics and Pragmatics*, 2016.

[FG12] M. C. Frank and N. D. Goodman.
**Predicting pragmatic reasoning in language games.**
*Science*, 336(6084):998–998, 2012.

[FJ14] Michael Franke and Gerhard Jäger.
**Pragmatic back-and-forth reasoning.**
*Semantics, Pragmatics and the Case of Scalar Implicatures.*, pages 170–200, 2014.

[Gaz79] Gerald Gazdar.
***Pragmatics, Implicature, Presuposition and Logical Form.***
Academic Press, New York, 1979.

[GK07] Thomas L. Griffiths and Michael L. Kalish.
**Language evolution by iterated learning with bayesian agents.**
*Cognitive Science*, 31(3):441–480, 2007.

[Gri75] Paul Grice.
**Logic and conversation.**
In *Studies in the Ways of Words*, chapter 2, pages 22–40. Harvard University Press, Cambridge, MA, 1975.

# References II


[Hor72] Laurence R. Horn.
***On the Semantic Properties of Logical Operators in English*.**
Indiana University Linguistics Club, Bloomington, IN, 1972.

[HS03] Josef Hofbauer and Karl Sigmund.
**Evolutionary game dynamics.**
*Bulletin of the American Mathematical Society*, 40(04):479–520, 2003.

[NK99] M. A. Nowak and D. C. Krakauer.
**The evolution of language.**
*Proceedings of the National Academy of Sciences*, 96(14):8028–8033, 1999.

[PTGew] Steven T. Piantadosi, Joshua B. Tenenbaum, and Noah D. Goodman.
**Modeling the acquisition of quantifier semantics: a case study in function word learnability, under review.**