The Rational Speech Act framework: an integrative theory of the interaction between literal meaning, world knowledge, and context

Judith Degen

June 21, 2019
XPrag 2019, Edinburgh

Early modern XPrag

focus on implicature cost question

Bott & Noveck 2004; Breheny et al 2006

Default theory

Levinson 2000

Relevance theory

Sperber & Wilson 1995; Carston 1998

Early modern XPrag

focus on implicature cost question

Bott & Noveck 2004; Breheny et al 2006

Default theory

Levinson 2000

Relevance theory

Sperber & Wilson 1995;
Carston 1998

Literal-first hypothesis

Huang & Snedeker 2009

Early modern XPrag

focus on implicature cost question

Bott & Noveck 2004; Breheny et al 2006

Default theory

Levinson 2000

Relevance theory

Sperber & Wilson 1995; Carston 1998

Literal-first hypothesis

Huang & Snedeker 2009

Constraint-based account

Degen & Tanenhaus 2015

Informational privilege

Default theory
Levinson 2000

Relevance theory
Sperber & Wilson 1995;
Carston 1998

extreme informational privilege

extreme parallelism

Literal-first hypothesis
Huang & Snedeker 2009

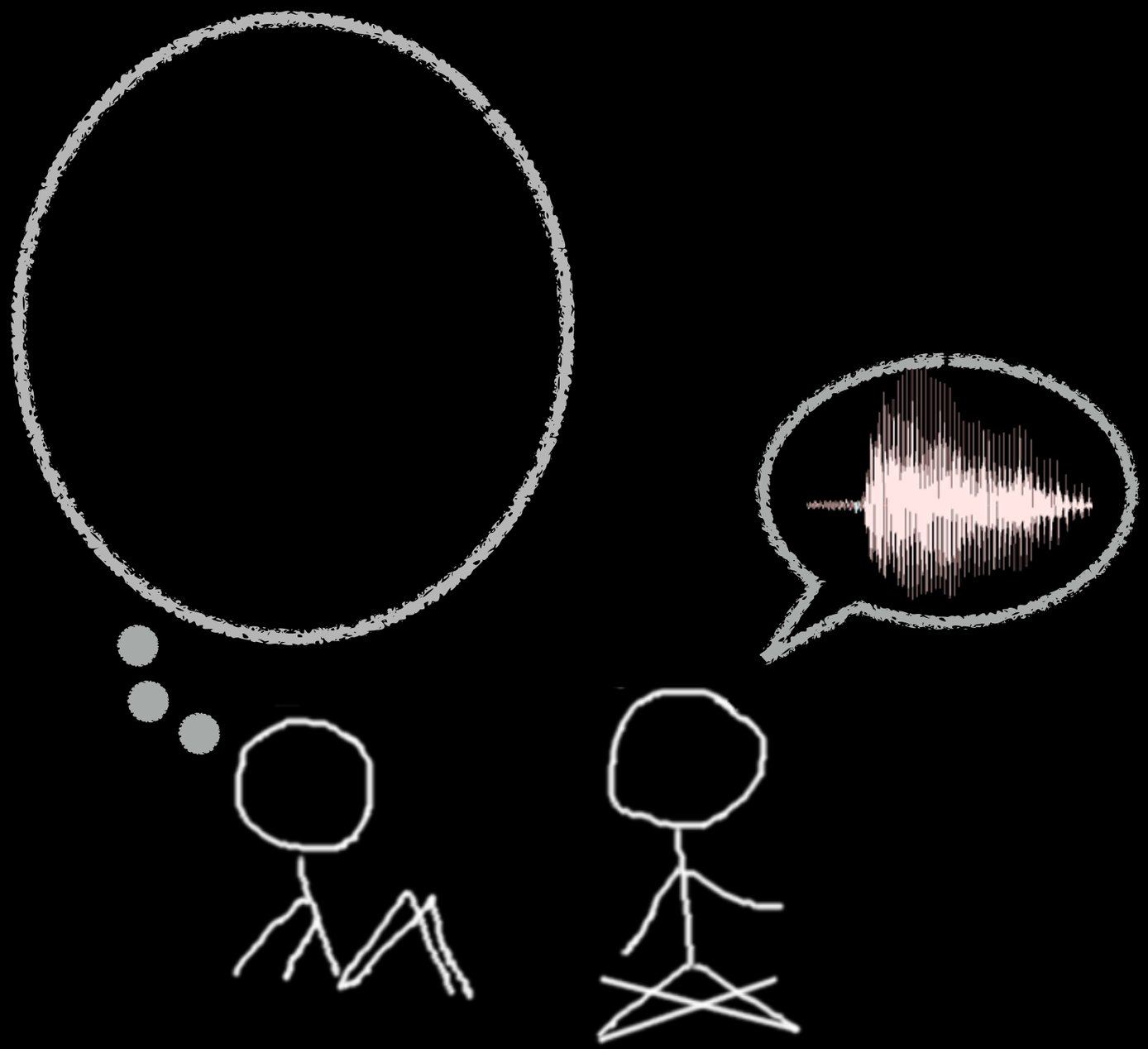
certain information
- **processed earlier**
- **weighted more heavily** in resulting interpretation

Constraint-based account
Degen & Tanenhaus 2015

all information
- **processed in parallel**
- **weighted equally** in resulting interpretation

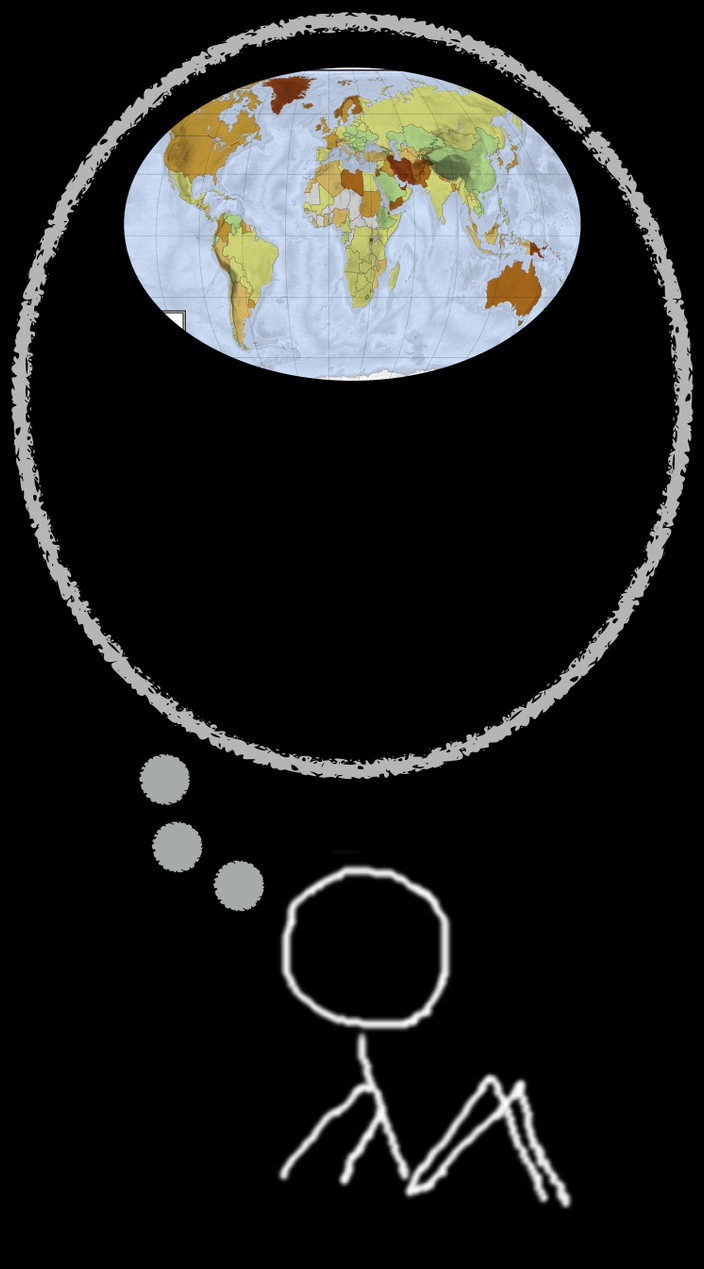
Degen & Tanenhaus 2019

linguistic signal



linguistic signal

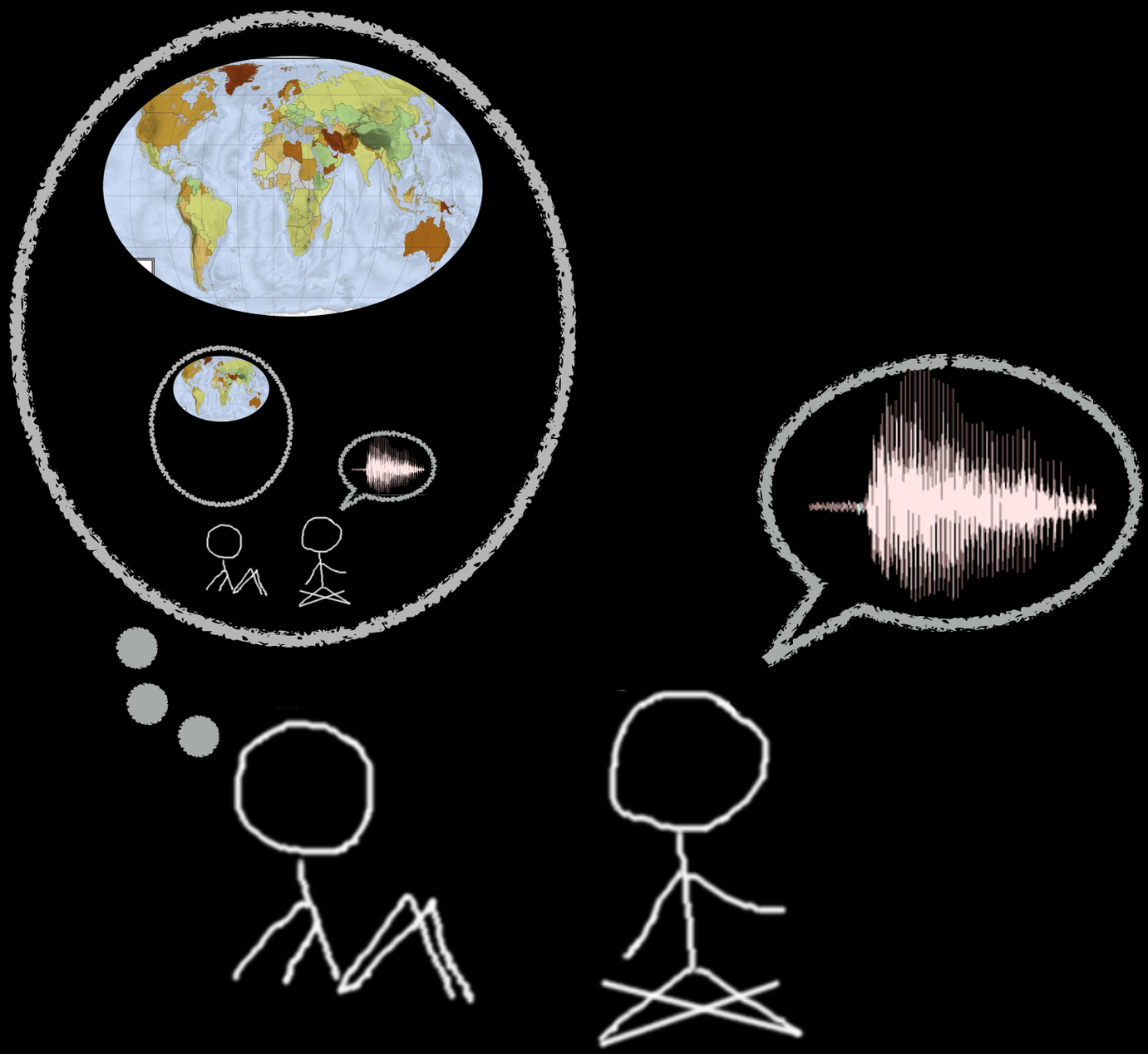
world knowledge



linguistic signal

world knowledge

context

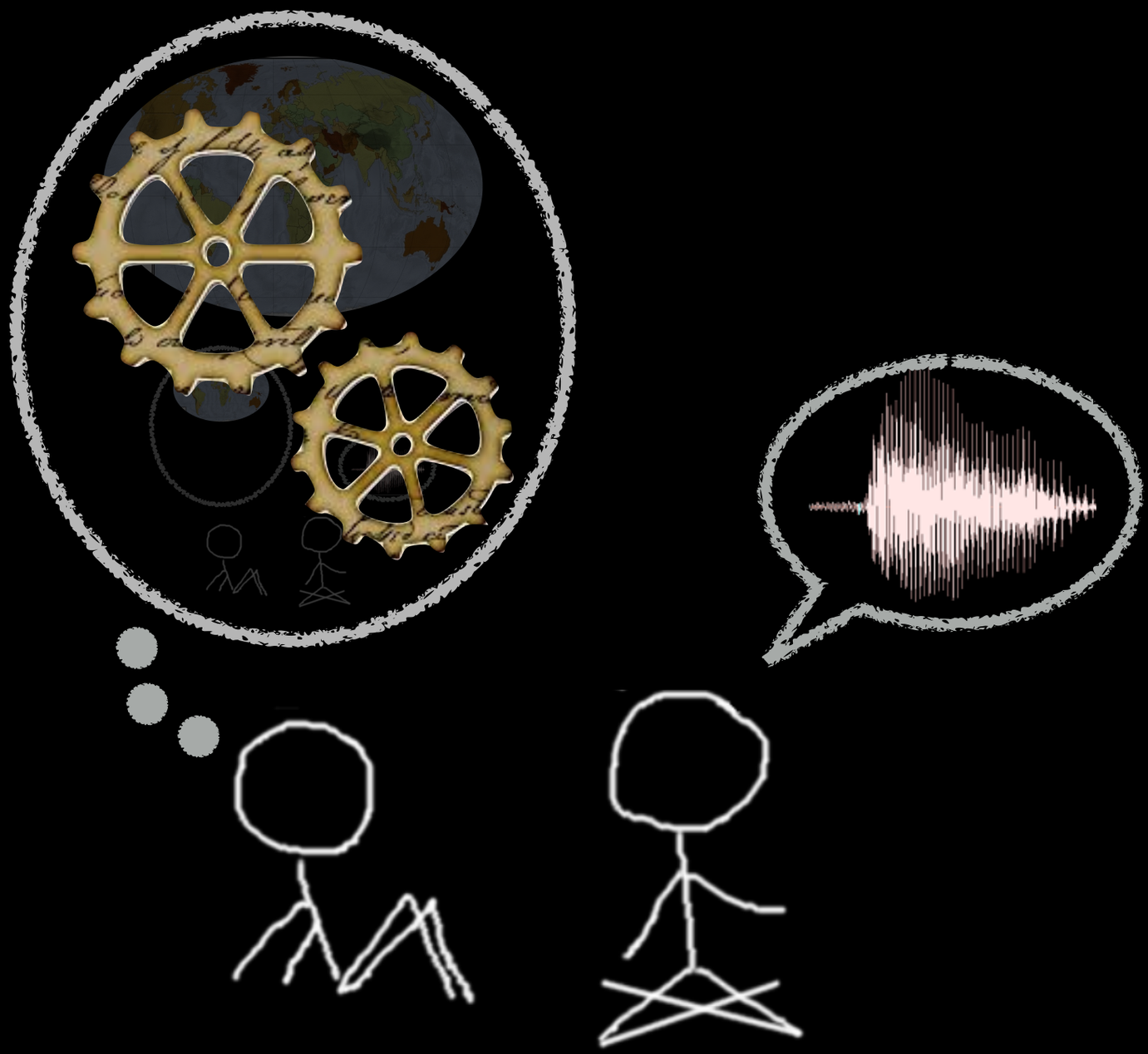


linguistic signal

world knowledge

reasoning

context

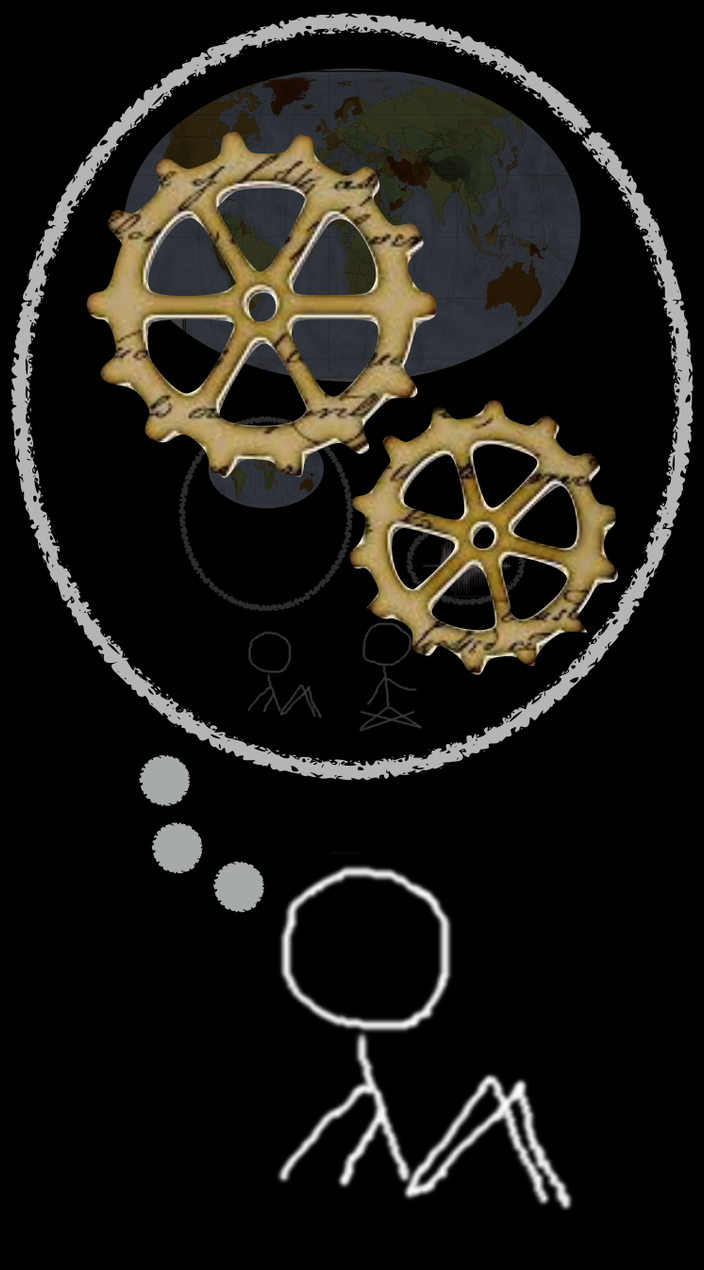


linguistic signal

world knowledge

reasoning

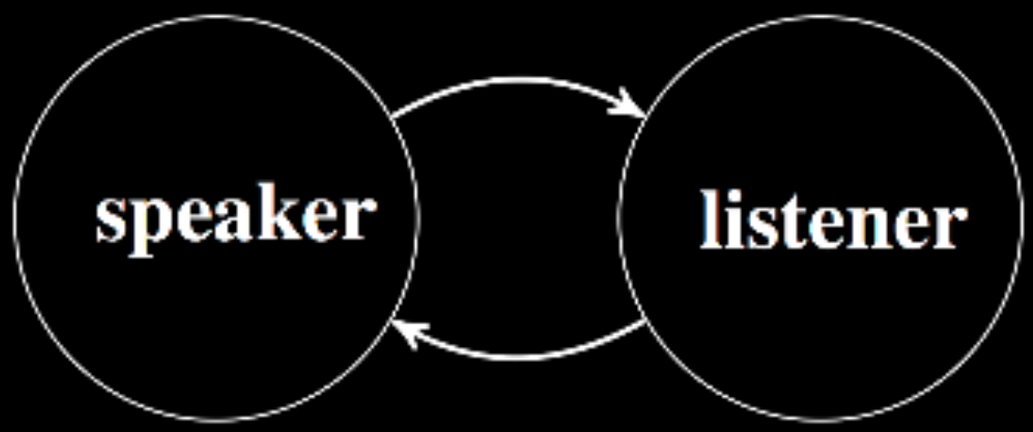
context



linguistic signal

PRAGMATICS

The Rational Speech Act Framework



RSA

Probabilistic pragmatics

Franke & Jäger, 2016; Goodman & Frank, 2016; Scontras, Tessler, & Franke 2018

Reference

Frank & Goodman, 2012; Qing & Franke, 2015; Degen & Franke, 2012; Stiller et al., 2015; Franke & Degen, 2015; Degen et al, under review

Cost-based Quantity implicatures

Degen et al., 2013; Rohde et al., 2012

Scalar implicatures

Goodman & Stuhlmüller, 2013; Degen et al., 2015

Embedded implicatures

Potts et al., 2016; Bergen et al., 2016

M-implicatures

Bergen et al., 2012

Figurative meaning

Kao et al., 2013; 2014; 2015; Cohn-Gordon & Bergen, under review

Gradable adjectives

Lassiter & Goodman, 2013; 2015; Qing & Franke, 2014

Adjective ordering

Hahn et al 2018; Scontras et al 2019

Other

plural predication Scontras & Goodman 2017
I-implicatures Poppels & Levy, 2016
generics Tessler & Goodman, 2019
modals Herbstritt & Franke, 2017
vague quantifiers Schöller & Franke, 2017
convention formation Hawkins et al 2018; 2019
questions Hawkins et al 2015
pragmatic adaptation Schuster & Degen, in prep
exhaustivity inferences Javangula & Degen in prep
atypicality inferences Kratvchenko & Demberg
social meaning Burnett 2017; 2019

Bayesian models in other areas

RSA

Bayesian models in other areas

language processing:

- speech perception
- syntactic adaptation
- reading (surprisal theory)

RSA

Bayesian models in other areas

higher-level cognition:

- reasoning
- categorization
- social reasoning
- intuitive physics

language processing:

- speech perception
- syntactic adaptation
- reading (surprisal theory)

RSA

Bayesian models in other areas

cognitive science more broadly:

- visual perception
- auditory perception
- multi-modal integration

higher-level cognition:

- reasoning
- categorization
- social reasoning
- intuitive physics

language processing:

- speech perception
- syntactic adaptation
- reading (surprisal theory)

RSA

Outline

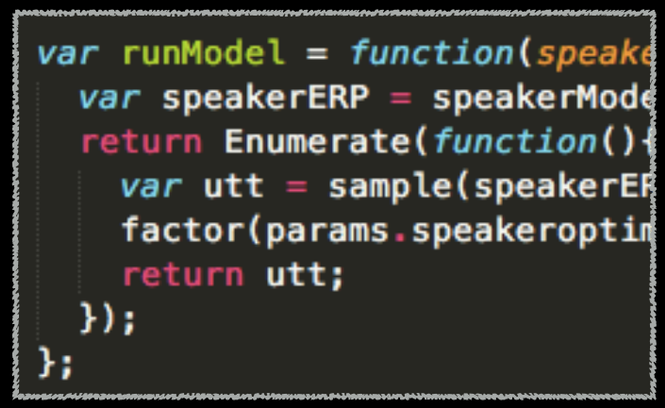
I. “**Overinformativeness**” in production
— why do we do it?

II. “**Underinformativeness**” in production
— how do we deal with it in comprehension?

Outline

I. "**Overinformativeness**" in production
— why do we do it?

II. "**Underinformativeness**" in production
— how do we deal with it in comprehension?



models

corpora

experiments

Part I

Production of referring expressions

Degen, Graf, Hawkins, & Goodman, under review

CONTENT SELECTION

Which features of an object should/do speakers mention?

Degen et al under review

Overinformative referring expressions — color/size asymmetry

color sufficient

the green lightbulb

Deutsch 1976; Pechmann 1989; Sedivy 2003; Gatt et al. 2011; many others

Overinformative referring expressions — color/size asymmetry

color sufficient

the green lightbulb

the big green lightbulb **8-10%**

Deutsch 1976; Pechmann 1989; Sedivy 2003; Gatt et al. 2011; many others

Overinformative referring expressions — color/size asymmetry

size sufficient

the big lightbulb

color sufficient

the green lightbulb

75-80% ***the big green lightbulb*** **8-10%**

Deutsch 1976; Pechmann 1989; Sedivy 2003; Gatt et al. 2011; many others

Overinformative referring expressions — color/size asymmetry

size sufficient

the big lightbulb

color sufficient

the green lightbulb

75-80% ***the big green lightbulb*** **8-10%**

1. speakers produce overinformative referring expressions
2. more overinformative color than size mentions

Deutsch 1976; Pechmann 1989; Sedivy 2003; Gatt et al. 2011; many others

Overinformative referring expressions — color/size asymmetry

size sufficient

the big lightbulb

75-80%

color sufficient

the green lightbulb

the big green lightbulb **8-10%**

1. speakers produce overinformative referring expressions
2. more overinformative color than size mentions

“OVERINFORMATIVENESS”

Deutsch 1976; Pechmann 1989; Sedivy 2003; Gatt et al. 2011; many others

Computational models of REs

- Greedy Algorithm
  Dale 1989
- Incremental Algorithm
  Dale & Reiter 1995
- PRO
  Gatt et al 2013; van Gompel et al 2019

Computational models of REs

Informativeness

- Greedy Algorithm
  Dale 1989
- Incremental Algorithm
  Dale & Reiter 1995
- PRO
  Gatt et al 2013; van Gompel et al 2019

Computational models of REs

- Greedy Algorithm
  Dale 1989

Informativeness

- Incremental Algorithm
  Dale & Reiter 1995
- PRO
  Gatt et al 2013; van Gompel et al 2019

Preferences

Computational models of REs

- Greedy Algorithm
  Dale 1989

Informativeness

- Incremental Algorithm
  Dale & Reiter 1995

Preferences

- PRO
  Gatt et al 2013; van Gompel et al 2019

Probabilities

Computational models of REs

- Greedy Algorithm
  Dale 1989

Informativeness

- Incremental Algorithm
  Dale & Reiter 1995

Preferences

- PRO
  Gatt et al 2013; van Gompel et al 2019

Probabilities

- Rational Speech Act (RSA)
  Frank & Goodman 2012

Computational models of REs

- Greedy Algorithm
  Dale 1989

Informativeness

- Incremental Algorithm
  Dale & Reiter 1995

Preferences

- PRO
  Gatt et al 2013; van Gompel et al 2019

Probabilities

- Rational Speech Act (RSA)
  Frank & Goodman 2012

The RSA framework

Frank & Goodman 2012

$O = \{$  ,  ,  $\}$

$U = \{\text{big}, \text{small}, \text{green}, \text{black}\}$

The RSA framework

Frank & Goodman 2012

$O = \{$  ,  ,  $\}$

$U = \{\text{big}, \text{small}, \text{green}, \text{black}\}$

Literal listener

$P_{L_0}(o|u) = \mathcal{U}(o|\{u \text{ is true of } o\})$

$[[u]] : O \rightarrow \{\text{true}, \text{false}\}$

The RSA framework

Frank & Goodman 2012

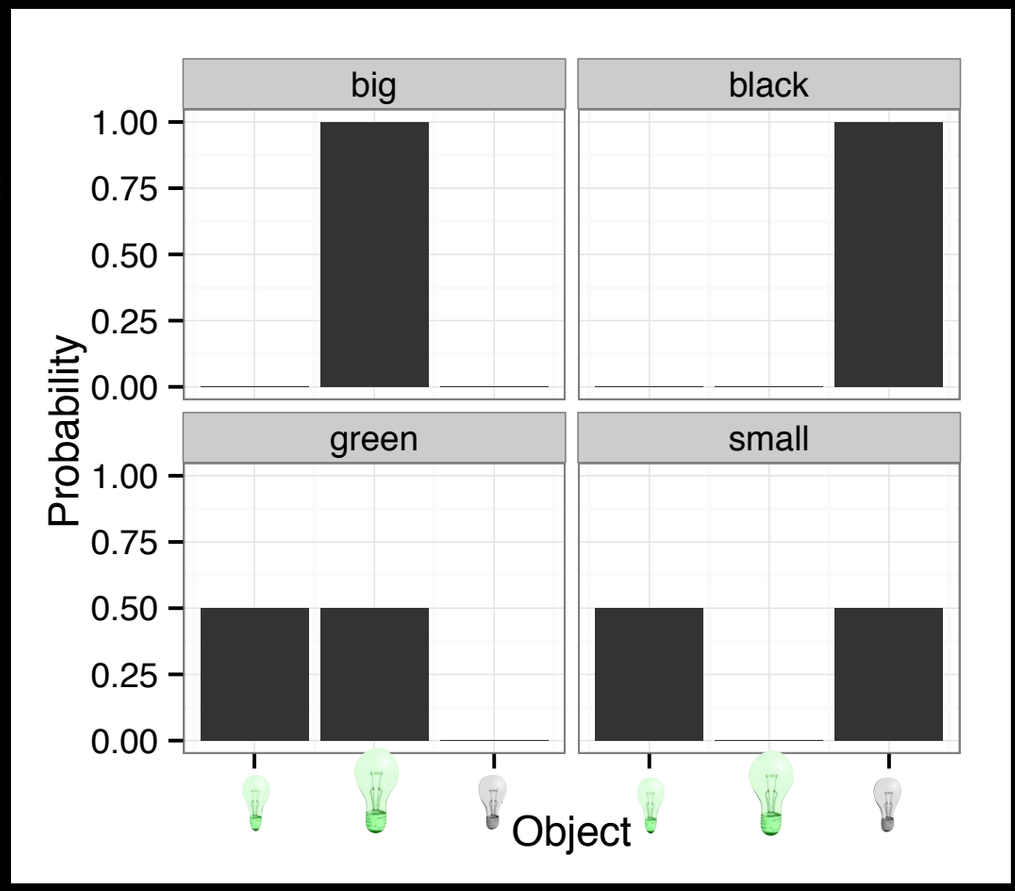
$O = \{$  ,  ,  $\}$

$U = \{\text{big}, \text{small}, \text{green}, \text{black}\}$

Literal listener

$$P_{L_0}(o|u) = \mathcal{U}(o|\{u \text{ is true of } o\})$$

$$[[u]] : O \rightarrow \{\text{true}, \text{false}\}$$



The RSA framework

Frank & Goodman 2012

$O = \{$  ,  ,  $\}$

$U = \{\text{big}, \text{small}, \text{green}, \text{black}\}$

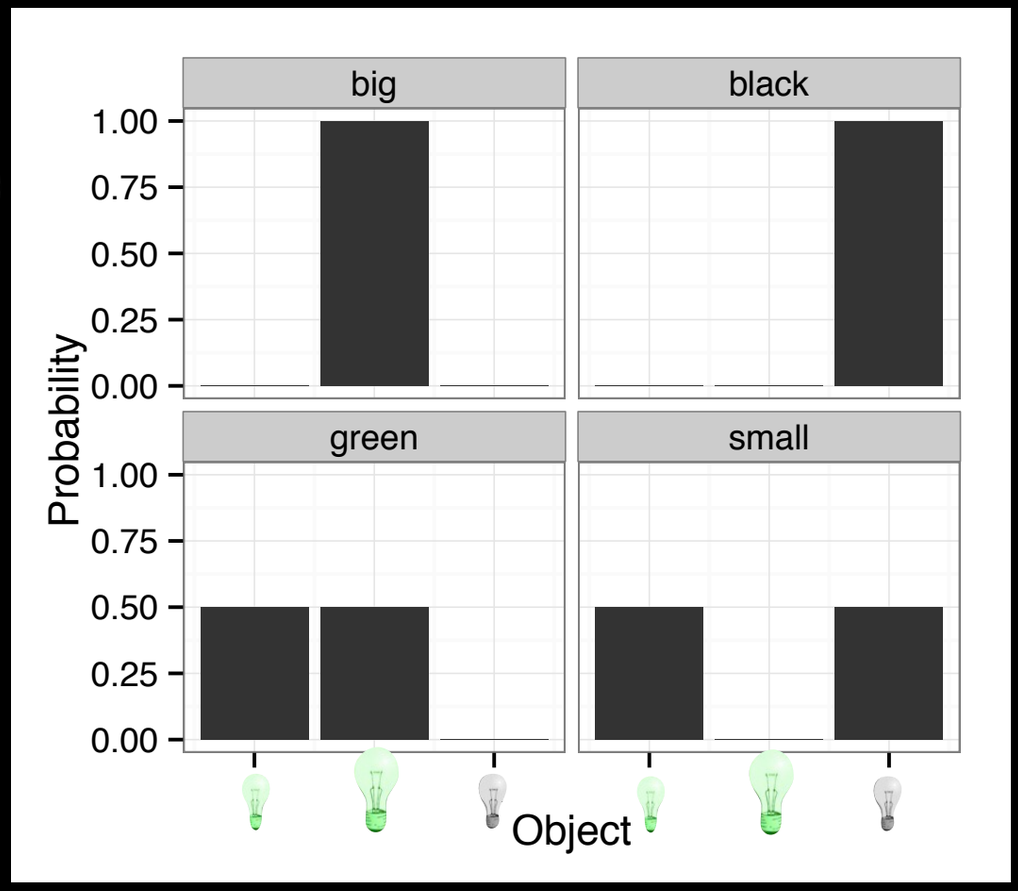
Literal listener

$$P_{L_0}(o|u) = \mathcal{U}(o|\{u \text{ is true of } o\})$$

$$[[u]] : O \rightarrow \{\text{true}, \text{false}\}$$

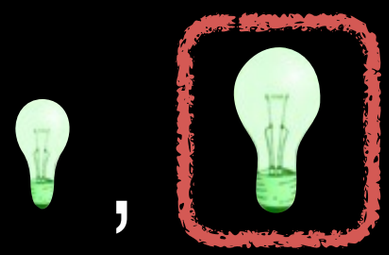
Pragmatic speaker

$$P_{S_1}(u|o) \propto e^{\lambda \cdot (\ln P_{L_0}(o|u) - C(u))}$$



The RSA framework

Frank & Goodman 2012

$O =$ {  }

$U = \{\text{big}, \text{small}, \text{green}, \text{black}\}$

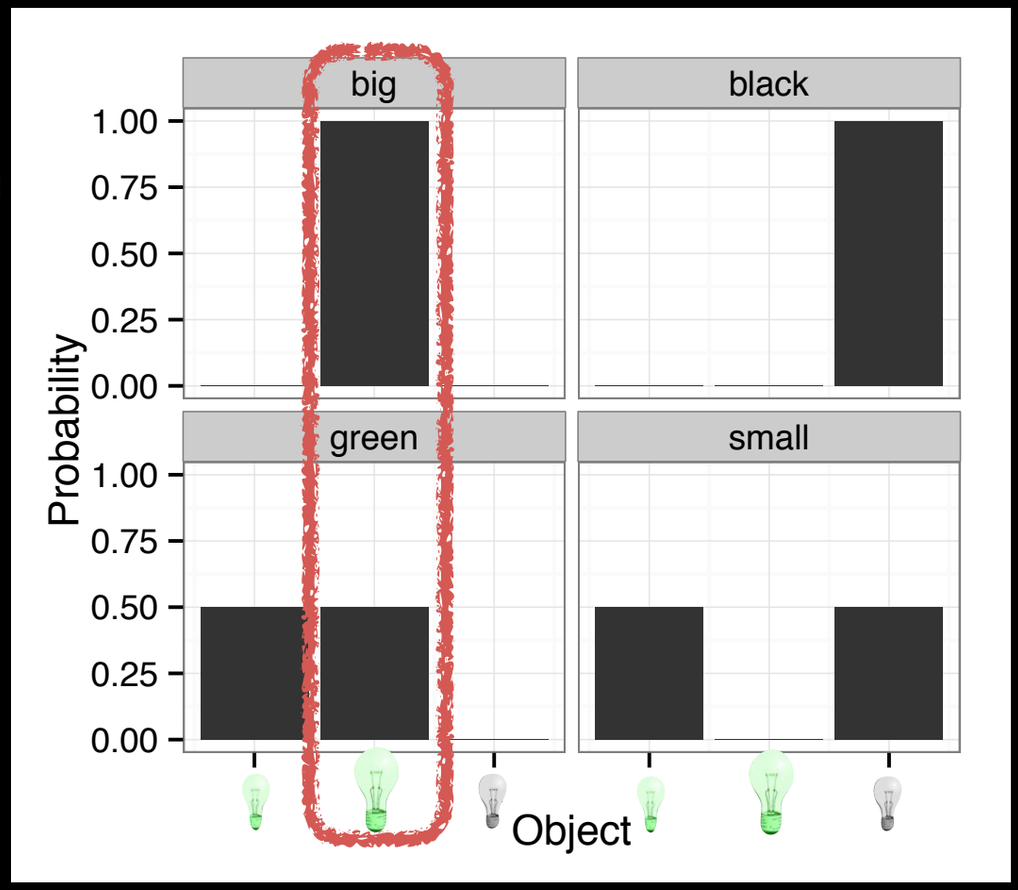
Literal listener

$$P_{L_0}(o|u) = \mathcal{U}(o|\{u \text{ is true of } o\})$$

$$[[u]] : O \rightarrow \{\text{true}, \text{false}\}$$

Pragmatic speaker

$$P_{S_1}(u|o) \propto e^{\lambda \cdot (\ln P_{L_0}(o|u) - C(u))}$$



The RSA framework

Frank & Goodman 2012

$O = \{$  ,  ,  $\}$

$U = \{\text{big}, \text{small}, \text{green}, \text{black}\}$

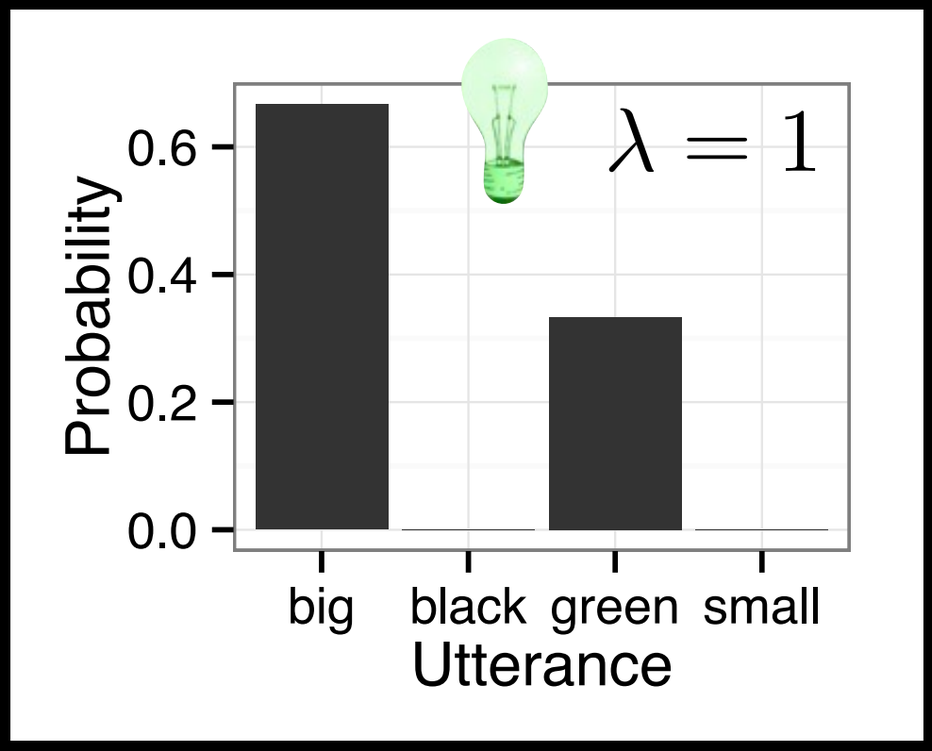
Literal listener

$$P_{L_0}(o|u) = \mathcal{U}(o|\{u \text{ is true of } o\})$$

$$[[u]] : O \rightarrow \{\text{true}, \text{false}\}$$

Pragmatic speaker

$$P_{S_1}(u|o) \propto e^{\lambda \cdot (\ln P_{L_0}(o|u) - C(u))}$$



The RSA framework

Frank & Goodman 2012

$O = \{$  ,  ,  $\}$

$U = \{\text{big}, \text{small}, \text{green}, \text{black}\}$

Literal listener

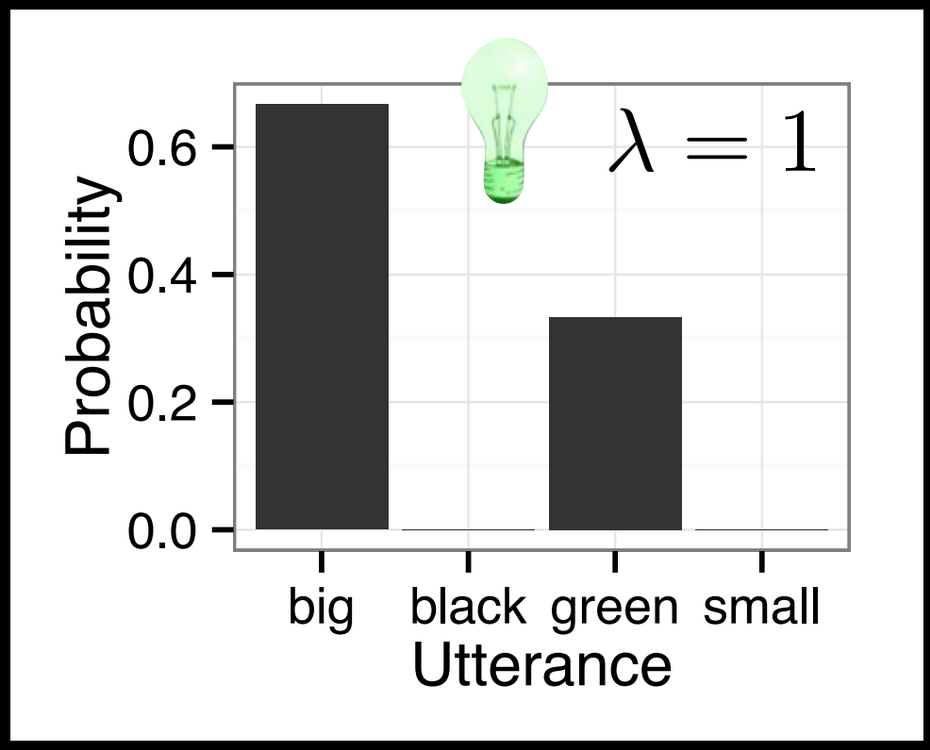
$$P_{L_0}(o|u) = \mathcal{U}(o|\{u \text{ is true of } o\})$$

$$[[u]] : O \rightarrow \{\text{true}, \text{false}\}$$

Pragmatic speaker

$$P_{S_1}(u|o) \propto e^{\lambda \cdot (\ln P_{L_0}(o|u) - C(u))}$$

Quantity



The RSA framework

Frank & Goodman 2012

$O = \{$  ,  ,  $\}$

$U = \{\text{big}, \text{small}, \text{green}, \text{black}\}$

Literal listener

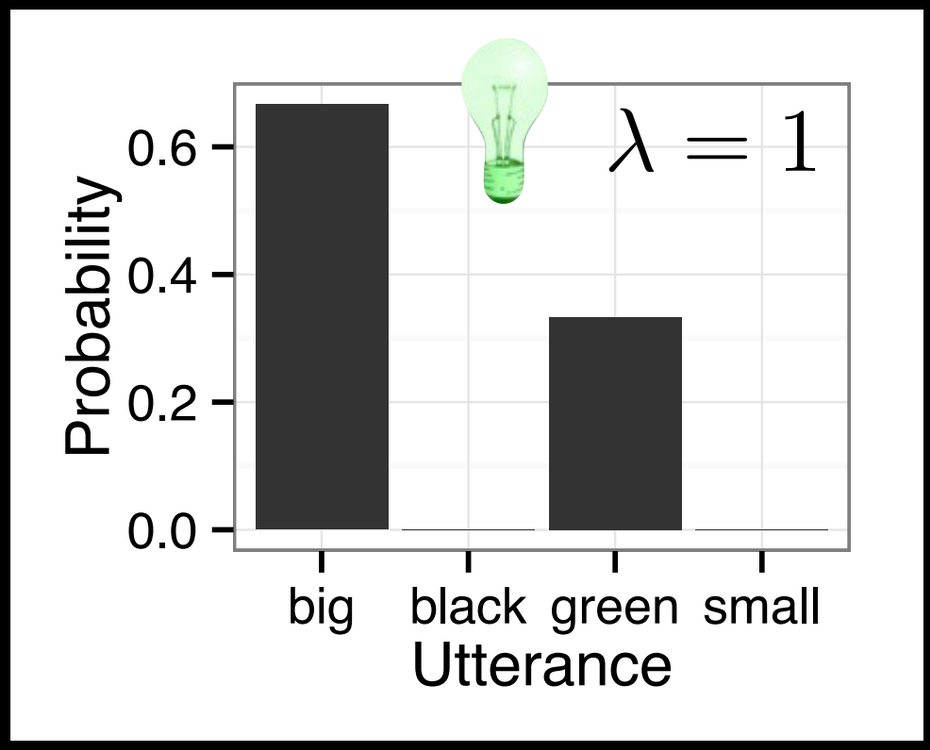
$$P_{L_0}(o|u) = \mathcal{U}(o|\{u \text{ is true of } o\})$$

$$[[u]] : O \rightarrow \{\text{true}, \text{false}\}$$

Pragmatic speaker

$$P_{S_1}(u|o) \propto e^{\lambda \cdot (\ln P_{L_0}(o|u) - C(u))}$$

Manner



The RSA framework

Frank & Goodman 2012

$O = \{$  ,  ,  $\}$

$U = \{\text{big}, \text{small}, \text{green}, \text{black}\}$

obvious problem:
no complex utterances

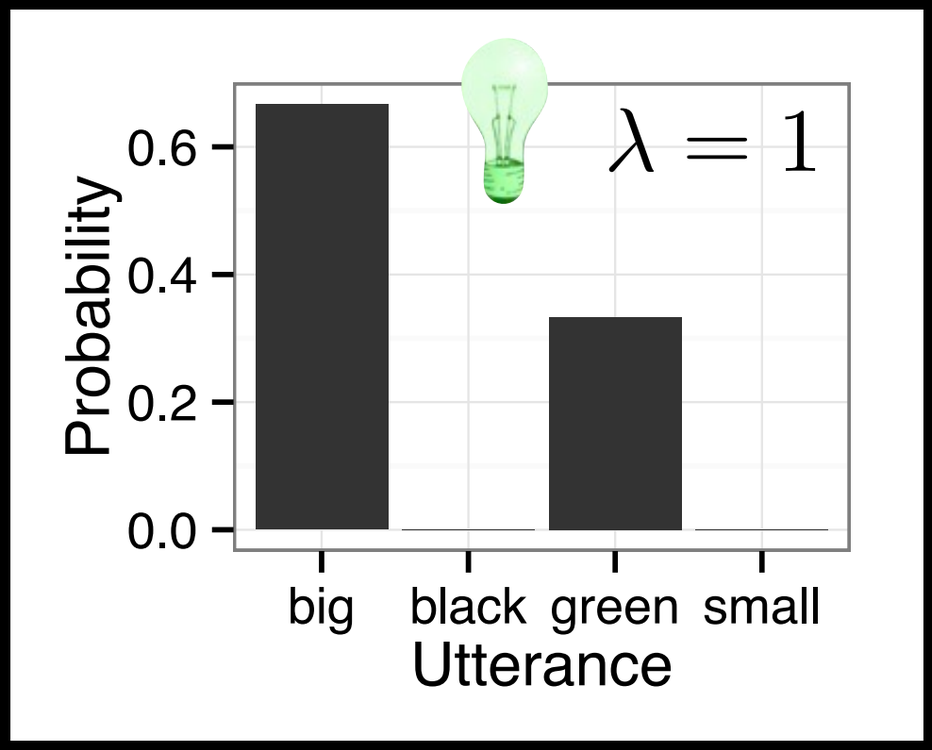
Literal listener

$$P_{L_0}(o|u) = \mathcal{U}(o|\{u \text{ is true of } o\})$$

$$[[u]] : O \rightarrow \{\text{true}, \text{false}\}$$

Pragmatic speaker

$$P_{S_1}(u|o) \propto e^{\lambda \cdot (\ln P_{L_0}(o|u) - C(u))}$$



The RSA framework

Frank & Goodman 2012

$O = \{$  ,  ,  $\}$

$U = \{\text{big}, \text{small}, \text{green}, \text{black}$
$\text{big green}, \text{small green}, \text{small black}\}$

obvious problem:
no complex utterances

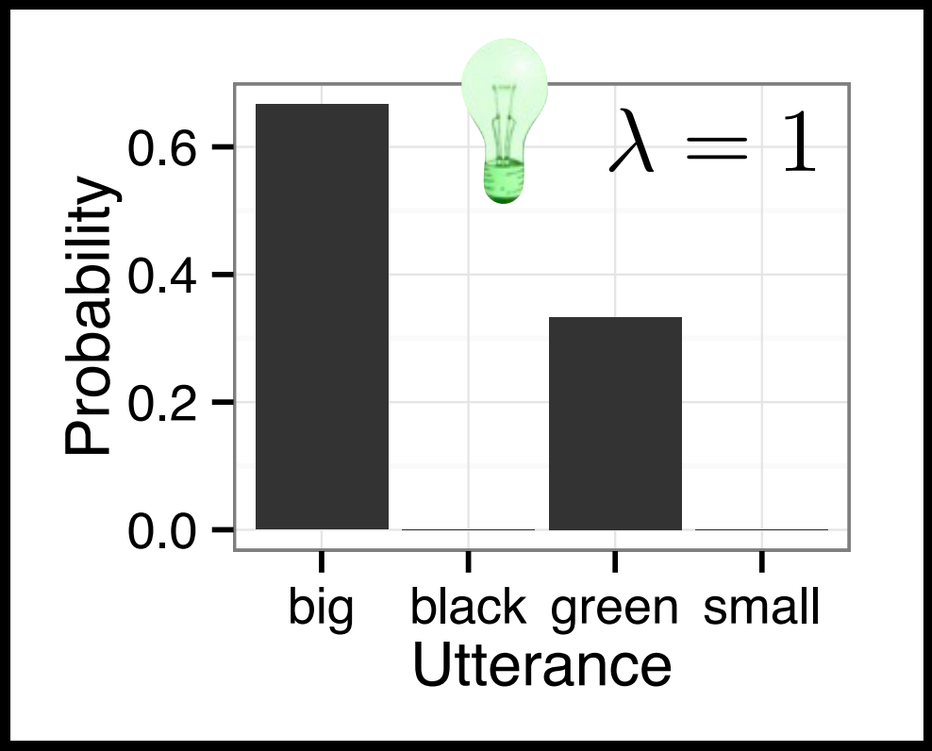
Literal listener

$$P_{L_0}(o|u) = \mathcal{U}(o|\{u \text{ is true of } o\})$$

$$[[u]] : O \rightarrow \{\text{true}, \text{false}\}$$

Pragmatic speaker

$$P_{S_1}(u|o) \propto e^{\lambda \cdot (\ln P_{L_0}(o|u) - C(u))}$$



Utterance semantics & cost

Intersective semantics

$$[[u]] = [[u_1]] \wedge [[u_2]]$$

$$[[\text{big green}]] = [[\text{big}]] \wedge [[\text{green}]]$$

Cost

$$C(u) = C(u_1) + C(u_2)$$

Utterance semantics & cost

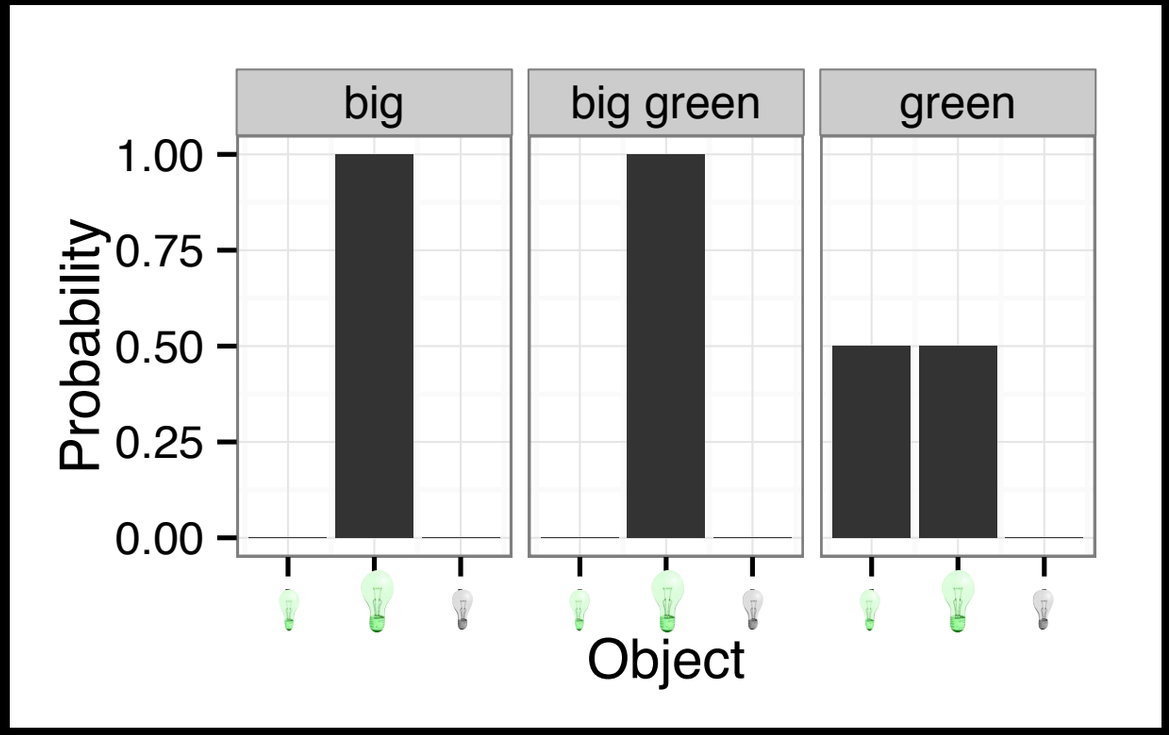
Intersective semantics

$$[[u]] = [[u_1]] \wedge [[u_2]]$$

$$[[\text{big green}]] = [[\text{big}]] \wedge [[\text{green}]]$$

Cost

$$C(u) = C(u_1) + C(u_2)$$



Utterance semantics & cost

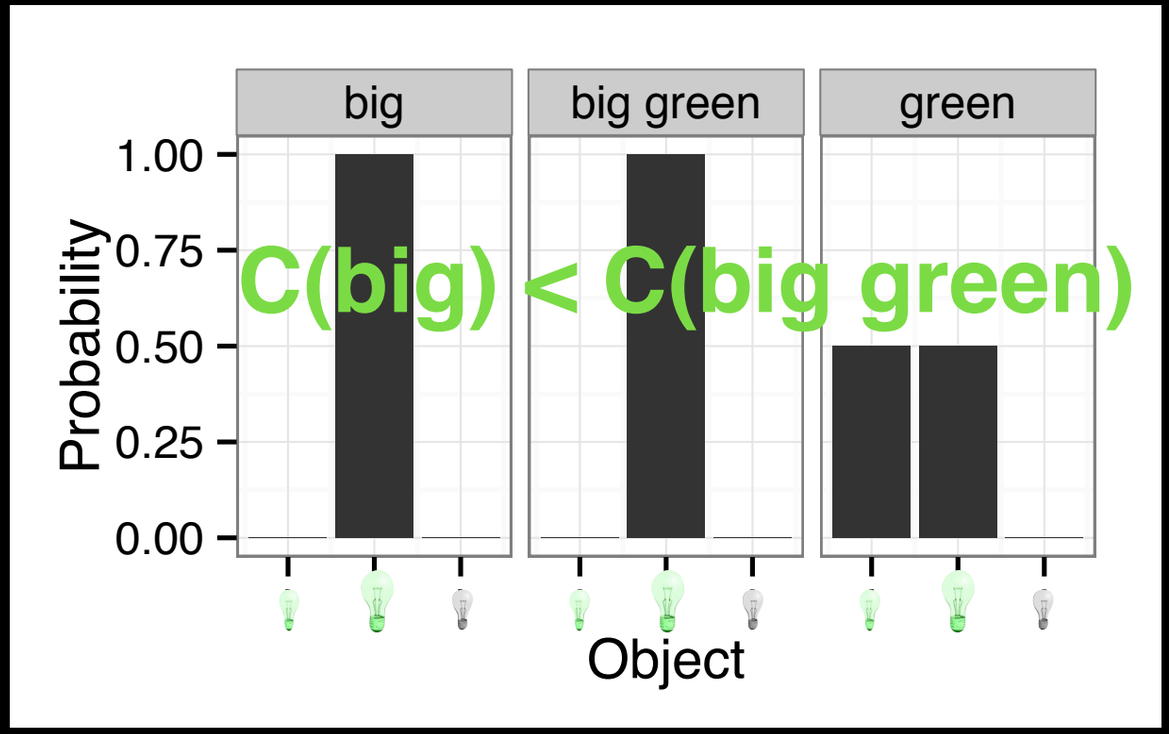
Intersective semantics

$$[[u]] = [[u_1]] \wedge [[u_2]]$$

$$[[\text{big green}]] = [[\text{big}]] \wedge [[\text{green}]]$$

Cost

$$C(u) = C(u_1) + C(u_2)$$



Utterance semantics & cost

Intersective semantics

$$[[u]] = [[u_1]] \wedge [[u_2]]$$

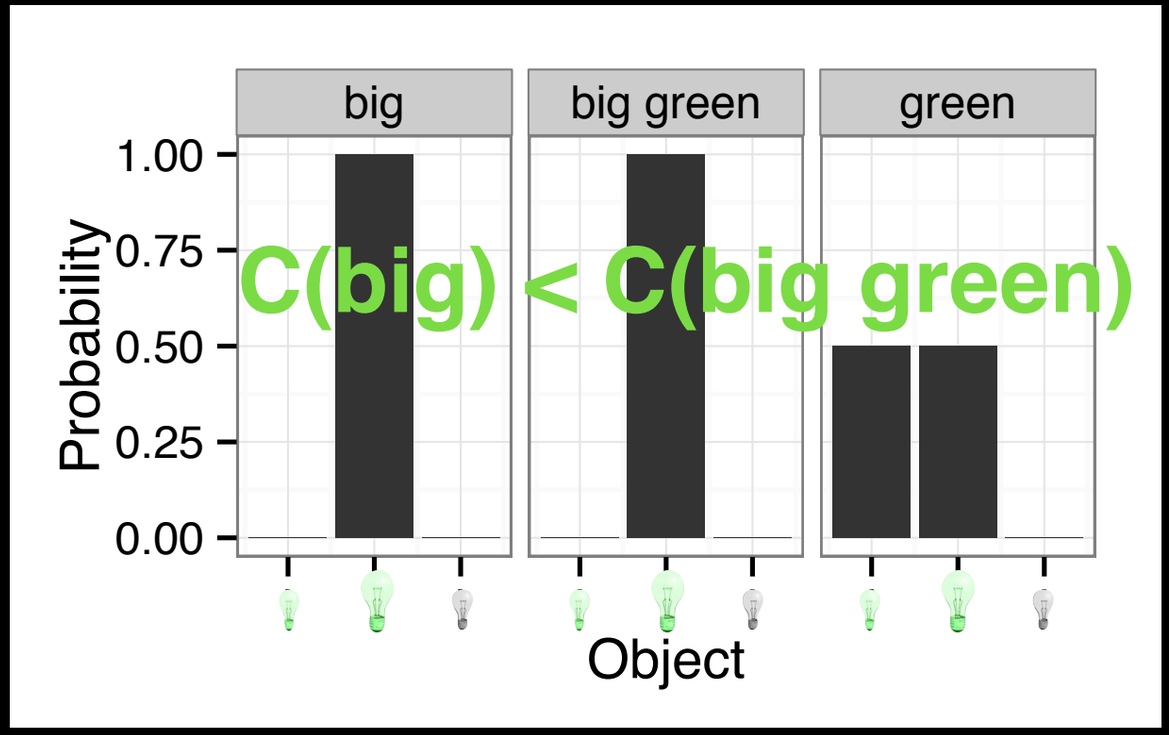
$$[[\text{big green}]] = [[\text{big}]] \wedge [[\text{green}]]$$

Cost

$$C(u) = C(u_1) + C(u_2)$$

RSA does not produce overinformative REs…

Gatt et al 2013; Westerbeek et al 2015



Utterance semantics & cost

Intersective semantics

$$[[u]] = [[u_1]] \wedge [[u_2]]$$

$$[[\text{big green}]] = [[\text{big}]] \wedge [[\text{green}]]$$

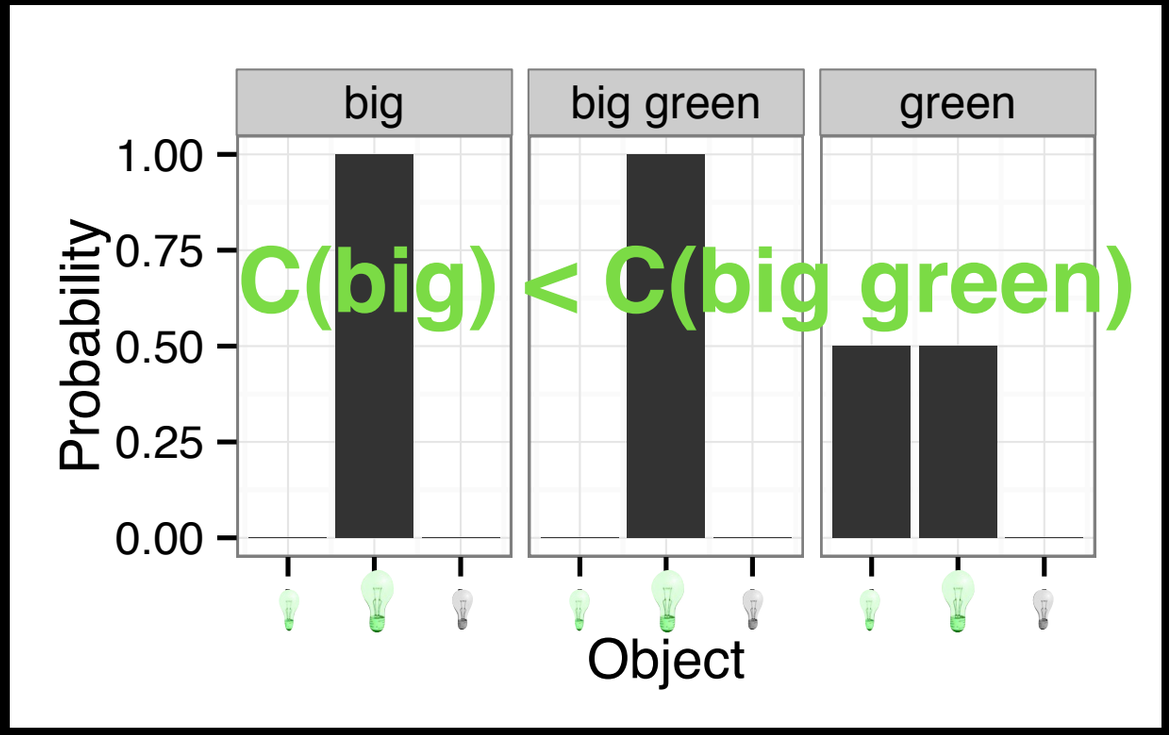
Cost

$$C(u) = C(u_1) + C(u_2)$$

RSA does not produce overinformative REs…

Gatt et al 2013; Westerbeek et al 2015

…with deterministic semantics



Motivation for non-deterministic semantics?

Modifiers differ:

size adjectives are vague and context-dependent
in a way that color adjectives are not

Kennedy & McNally 2005

color is intrinsically salient in a way that size is not

Arts et al 2011; Gatt et al 2013

size adjectives are judged to be more subjective
than color adjectives

Scontras, Degen, & Goodman 2017

Non-deterministic semantics

Literal listener

$$P_{L_0}(o|u) \propto \begin{cases} 1-\epsilon & [[u]](o) = \text{true} \\ \epsilon & \text{otherwise} \end{cases}$$

Non-deterministic semantics

Literal listener

fidelity

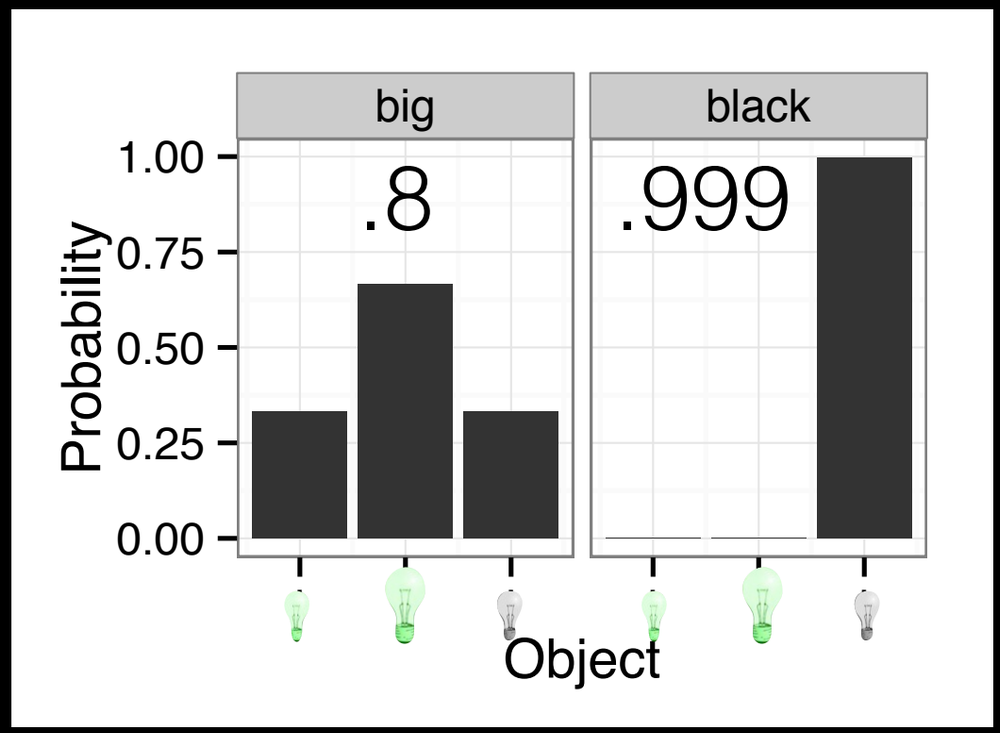
$$P_{L_0}(o|u) \propto \begin{cases} 1-\epsilon & [[u]](o) = \text{true} \\ \epsilon & \text{otherwise} \end{cases}$$

Non-deterministic semantics

Literal listener

fidelity

$$P_{L_0}(o|u) \propto \begin{cases} 1-\epsilon & [[u]](o) = \text{true} \\ \epsilon & \text{otherwise} \end{cases}$$



Non-deterministic semantics

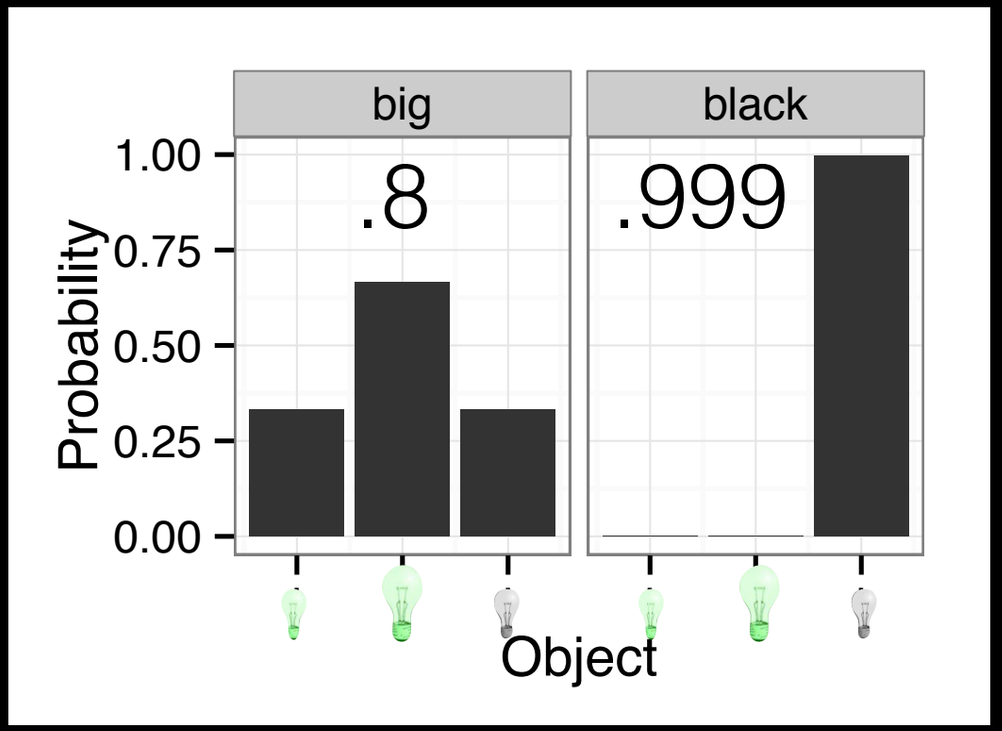
Literal listener

fidelity

$$P_{L_0}(o|u) \propto \begin{cases} 1-\epsilon & [[u]](o) = \text{true} \\ \epsilon & \text{otherwise} \end{cases}$$

Pragmatic speaker

$$P_{S_1}(u|o) \propto e^{\lambda \cdot (\ln P_{L_0}(o|u) - C(u))}$$



Non-deterministic semantics

Literal listener

fidelity

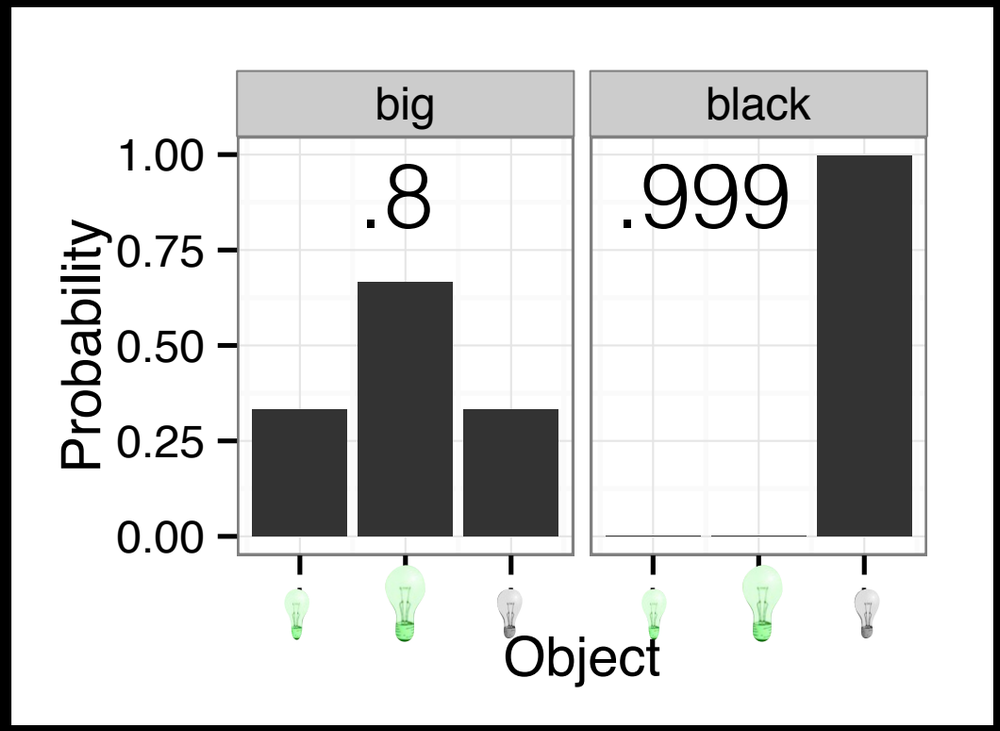
$$P_{L_0}(o|u) \propto \begin{cases} 1-\epsilon & [[u]](o) = \text{true} \\ \epsilon & \text{otherwise} \end{cases}$$

Pragmatic speaker

$$P_{S_1}(u|o) \propto e^{\lambda \cdot (\ln P_{L_0}(o|u) - C(u))}$$

Two free fidelity parameters:

$\text{fid}(size)$ $\text{fid}(color)$



Non-deterministic semantics

Literal listener

fidelity

$$P_{L_0}(o|u) \propto \begin{cases} 1-\epsilon & [[u]](o) = \text{true} \\ \epsilon & \text{otherwise} \end{cases}$$

Pragmatic speaker

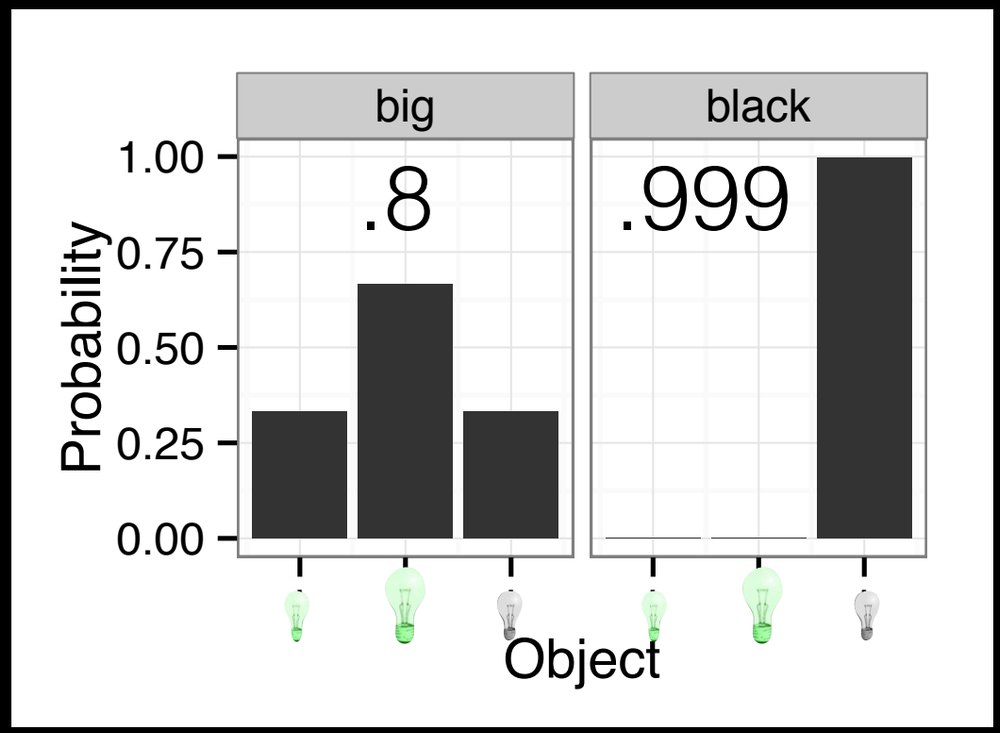
$$P_{S_1}(u|o) \propto e^{\lambda \cdot (\ln P_{L_0}(o|u) - C(u))}$$

Two free fidelity parameters:

$\text{fid}(size)$ $\text{fid}(color)$

Two free cost parameters:

$C(size)$ $C(color)$



Non-deterministic semantics

Literal listener

fidelity

$$P_{L_0}(o|u) \propto \begin{cases} 1-\epsilon & [[u]](o) = \text{true} \\ \epsilon & \text{otherwise} \end{cases}$$

Pragmatic speaker

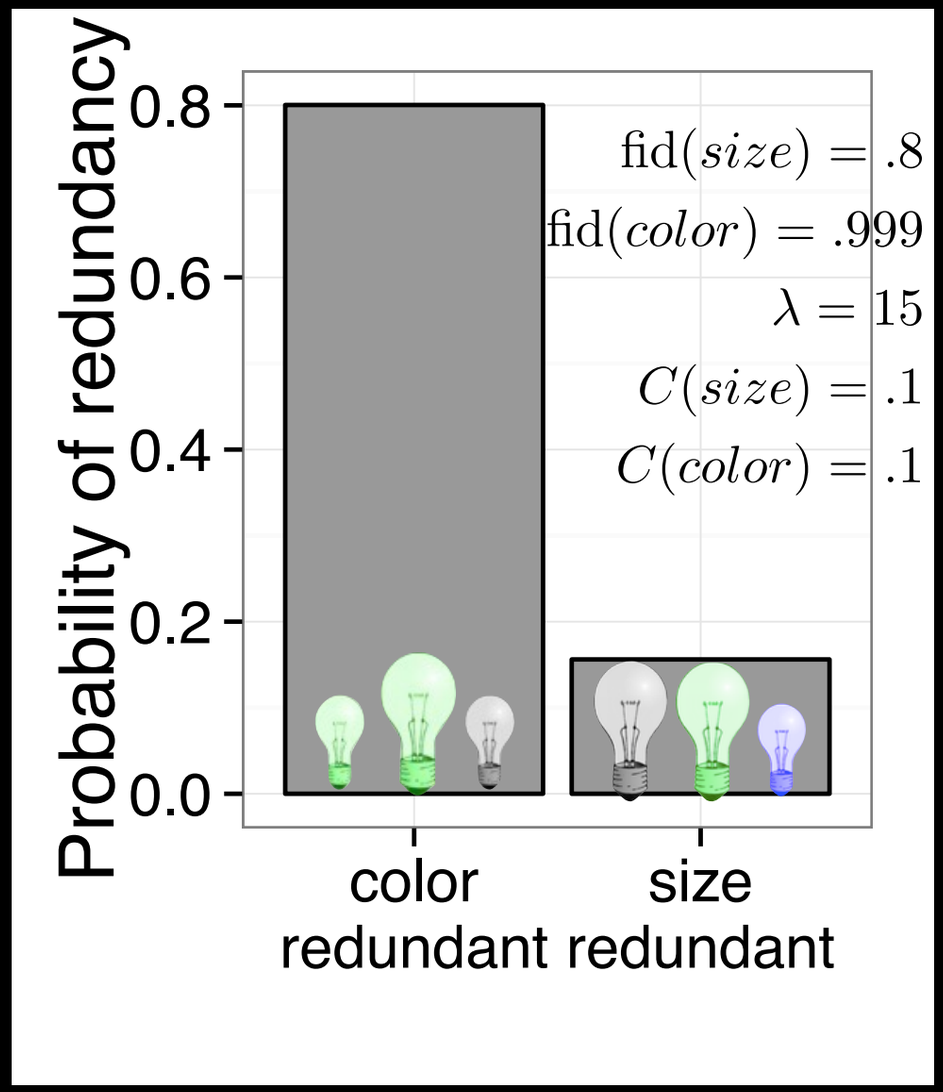
$$P_{S_1}(u|o) \propto e^{\lambda \cdot (\ln P_{L_0}(o|u) - C(u))}$$

Two free fidelity parameters:

$\text{fid}(size)$ $\text{fid}(color)$

Two free cost parameters:

$C(size)$ $C(color)$



Non-deterministic semantics

Literal listener

fidelity

$$P_{L_0}(o|u) \propto \begin{cases} 1-\epsilon & [[u]](o) = \text{true} \\ \epsilon & \text{otherwise} \end{cases}$$

Pragmatic speaker

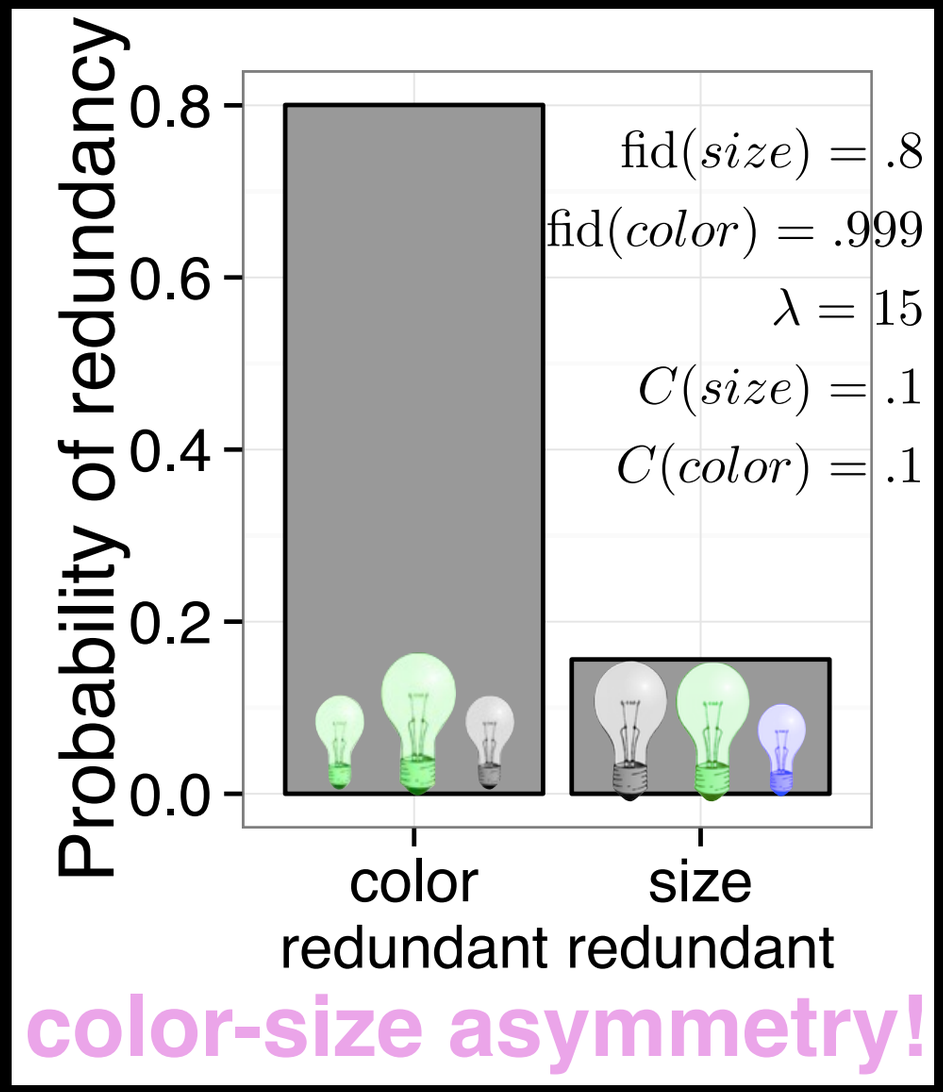
$$P_{S_1}(u|o) \propto e^{\lambda \cdot (\ln P_{L_0}(o|u) - C(u))}$$

Two free fidelity parameters:

$\text{fid}(size)$ $\text{fid}(color)$

Two free cost parameters:

$C(size)$ $C(color)$



color-size asymmetry!

Non-deterministic semantics

Literal listener

fidelity

$$P_{L_0}(o|u) \propto \begin{cases} 1-\epsilon & [[u]](o) = \text{true} \\ \epsilon & \text{otherwise} \end{cases}$$

Pragmatic speaker

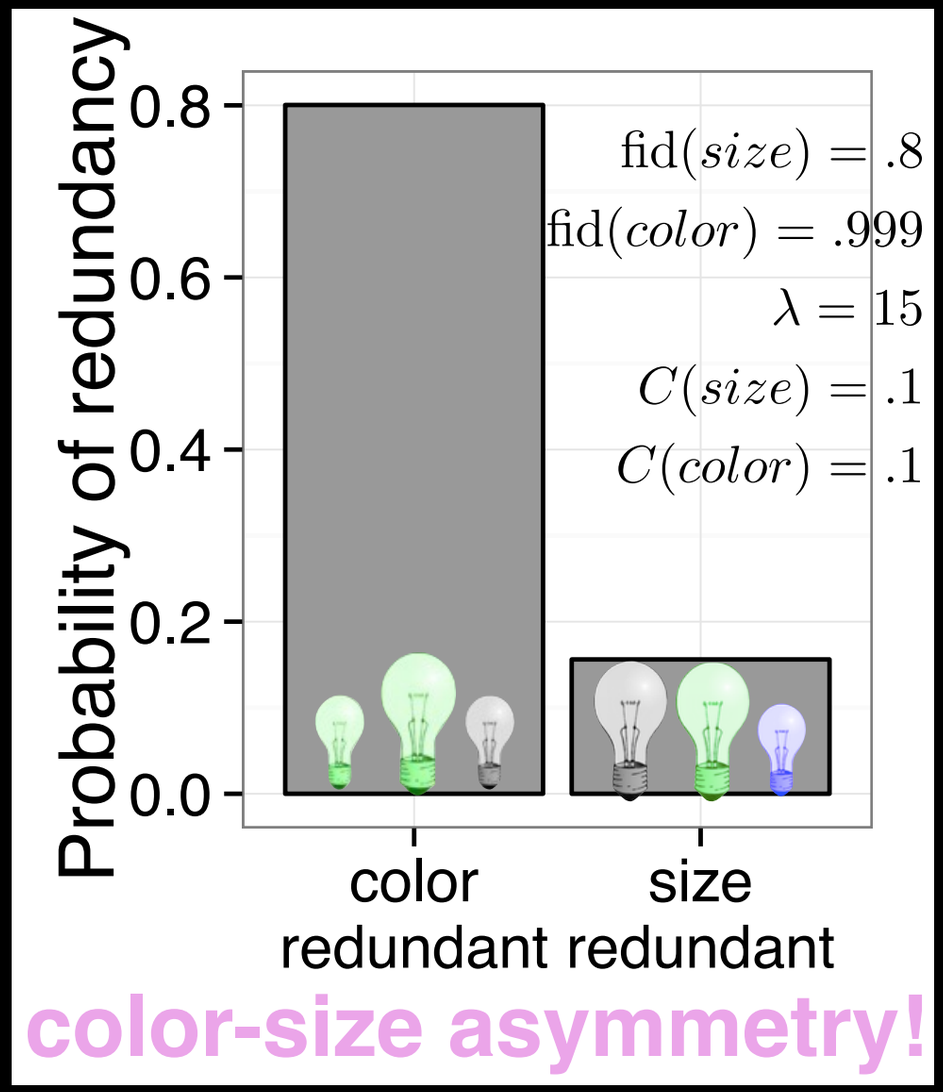
$$P_{S_1}(u|o) \propto e^{\lambda \cdot (\ln P_{L_0}(o|u) - C(u))}$$

Two free fidelity parameters:

$\text{fid}(size)$ $\text{fid}(color)$

Two free cost parameters:

$C(size)$ $C(color)$



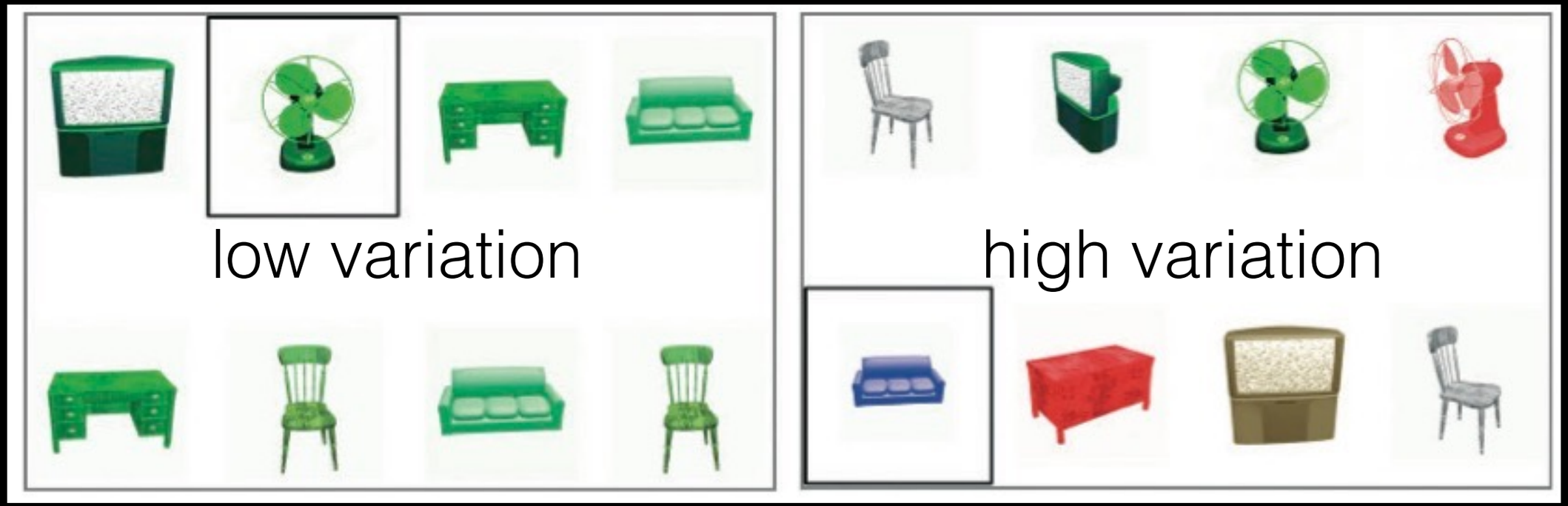
color-size asymmetry!

If modifiers don't "work perfectly",

adding modifiers adds information

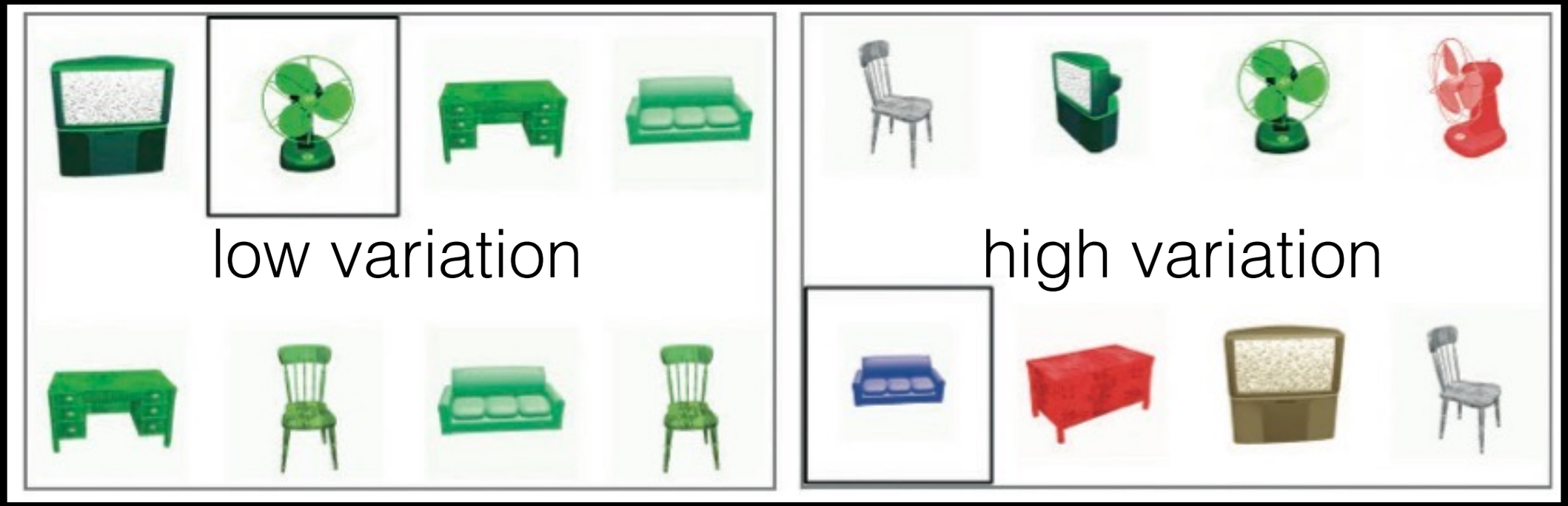
Independent empirical evidence for RSA with non-deterministic semantics?

Scene variation



Koolen et al 2013, Davies & Katsos 2013

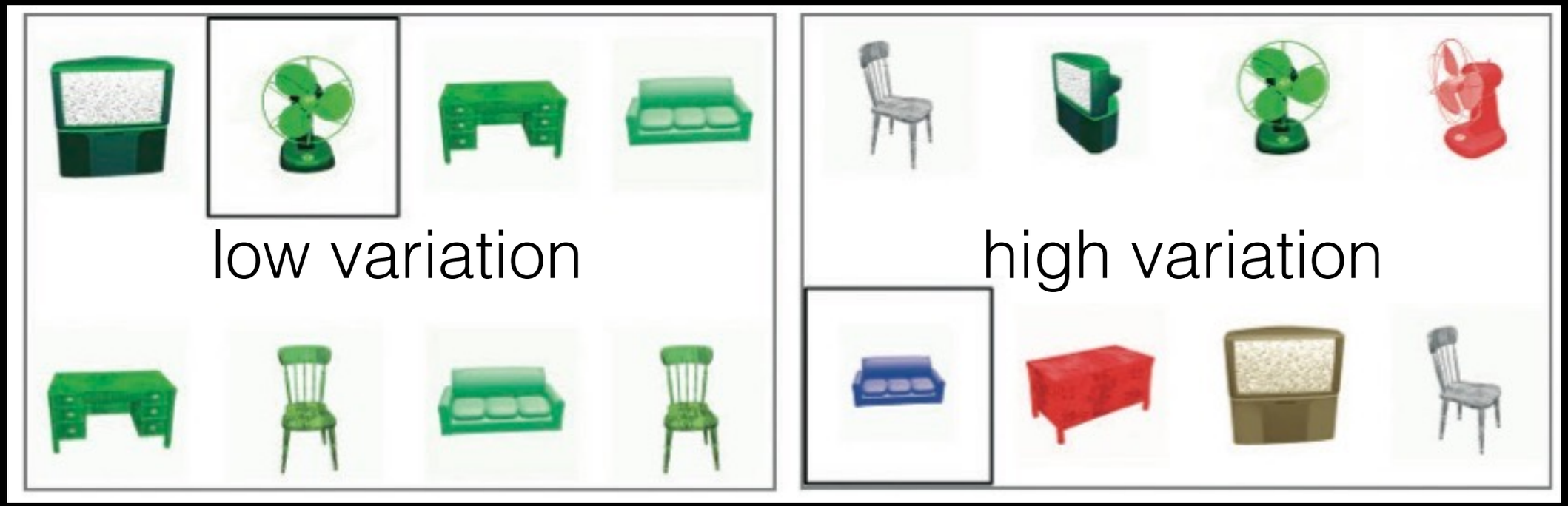
Scene variation



more redundant color use in high-variation scenes

Koolen et al 2013, Davies & Katsos 2013

Scene variation



more redundant color use in high-variation scenes

Koolen et al 2013, Davies & Katsos 2013

non-deterministic RSA predicts this result

Independent
quantitative evidence
for non-deterministic
RSA?

Scene variation

scene variation increases probability of redundancy

Scene variation

scene variation increases probability of redundancy

$\frac{n_{\text{diff}}}{n_{\text{total}}}$ proportion of total distractors that don't share target value on insufficient dimension

Scene variation

scene variation increases probability of redundancy

$\frac{n_{\text{diff}}}{n_{\text{total}}}$ proportion of total distractors that don't share target value on insufficient dimension

sufficient dimension: size

insufficient dimension: color → $\frac{n_{\text{red}}}{n_{\text{total}}} = \frac{2}{4} = .5$

Scene variation

scene variation increases probability of redundancy

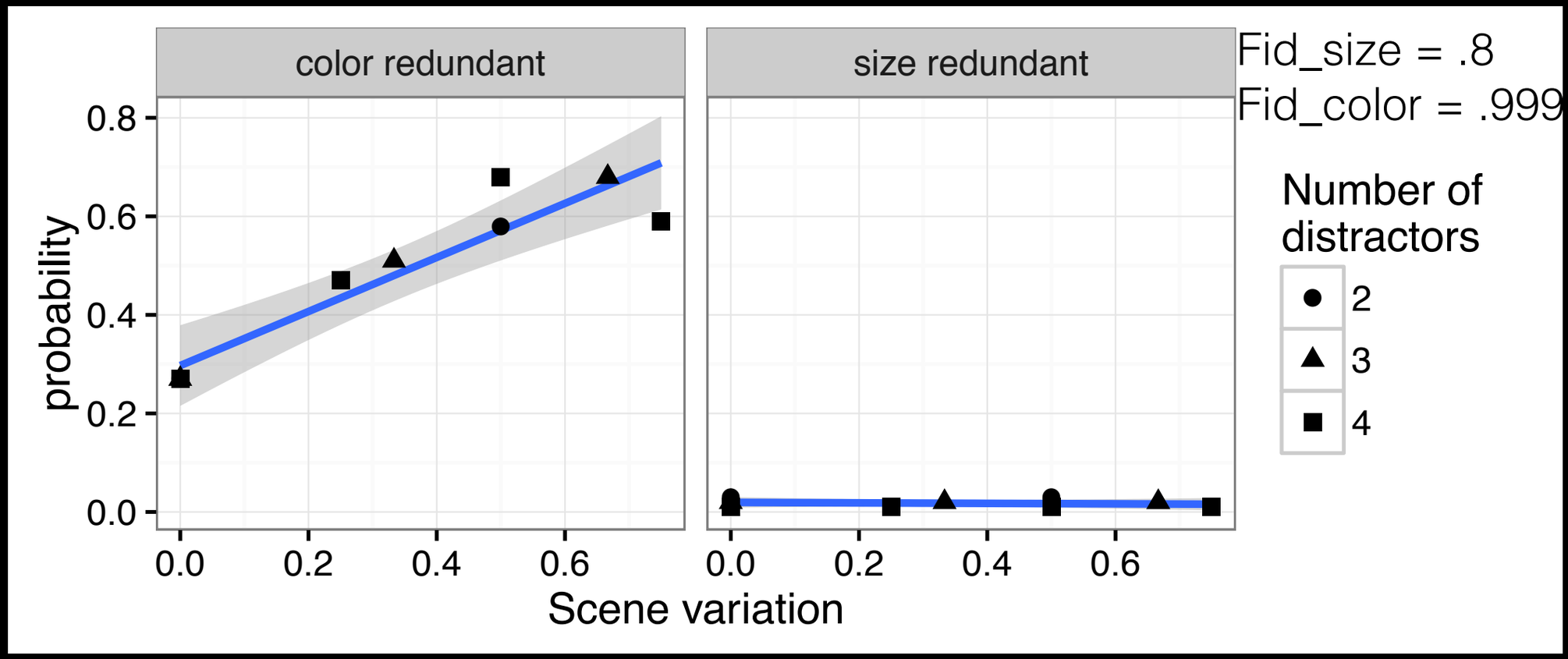
$\frac{n_{\text{diff}}}{n_{\text{total}}}$ proportion of total distractors that don't share target value on insufficient dimension

sufficient dimension: size

insufficient dimension: color → $\frac{n_{\text{red}}}{n_{\text{total}}} = \frac{2}{4} = .5$

greater proportion = more variation

Model predictions



Prediction: increase in redundant adjective use with increasing scene variation for color but not size

Interactive reference game experiment

- 58 pairs of participants on Mechanical Turk
- random assignment to speaker/listener role
- 72 trials (half targets, half fillers)
- 36 object types
- on all target trials, one of size or color was sufficient
- **scene variation manipulation**:
  - total number of distractors (2, 3, 4)
  - number of distractors that shared the insufficient feature value with target

Speaker's perspective

You: the stapler

listener: which one??

You: big purple

Round 1 of 72

Send

Listener's perspective

speaker: the stapler

You: which one??

speaker: big purple

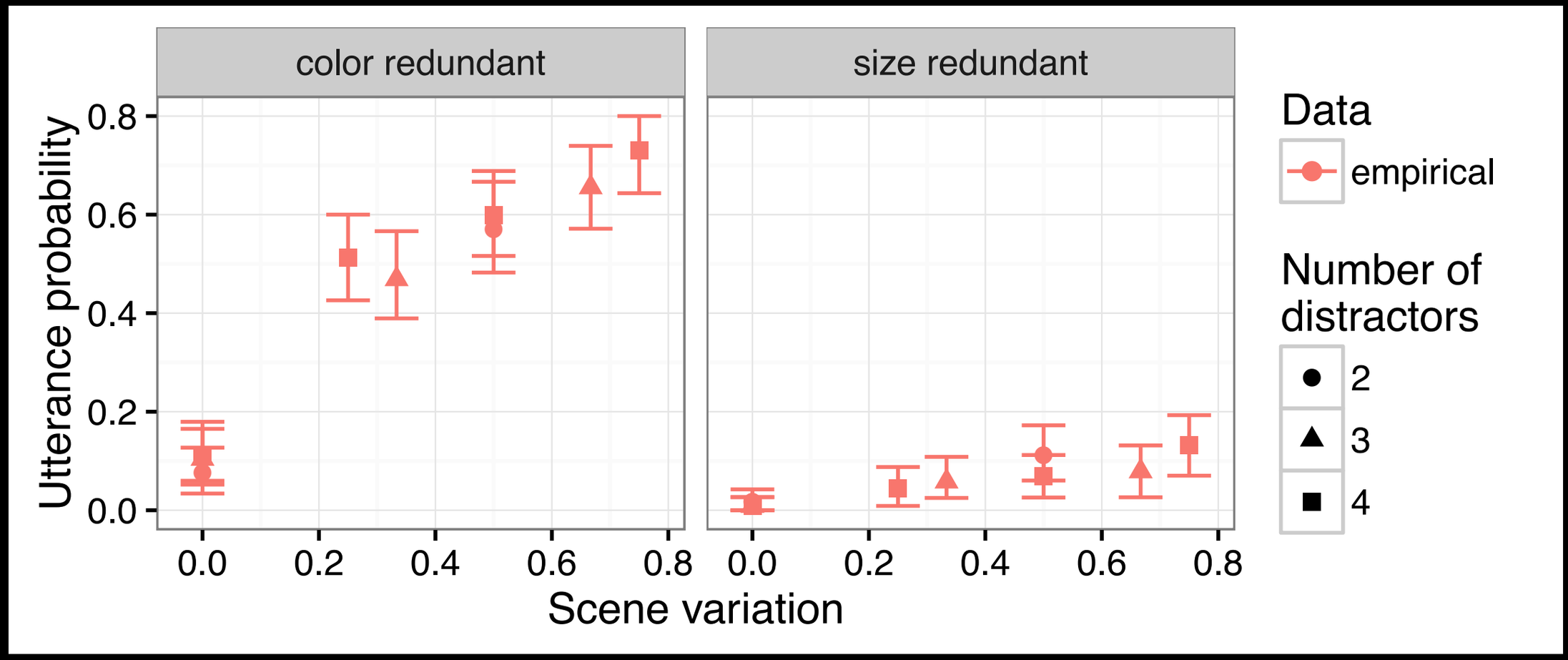
Round 1 of 72

Send



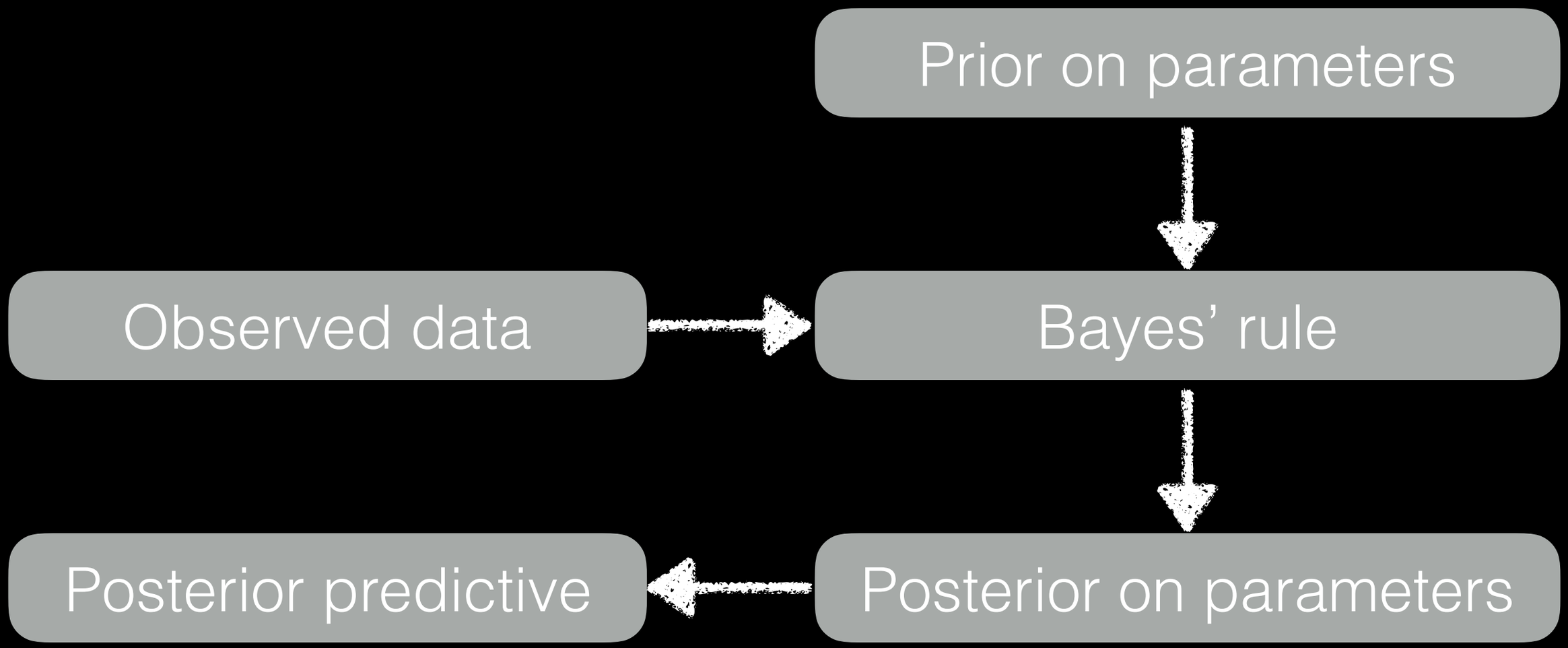


Results

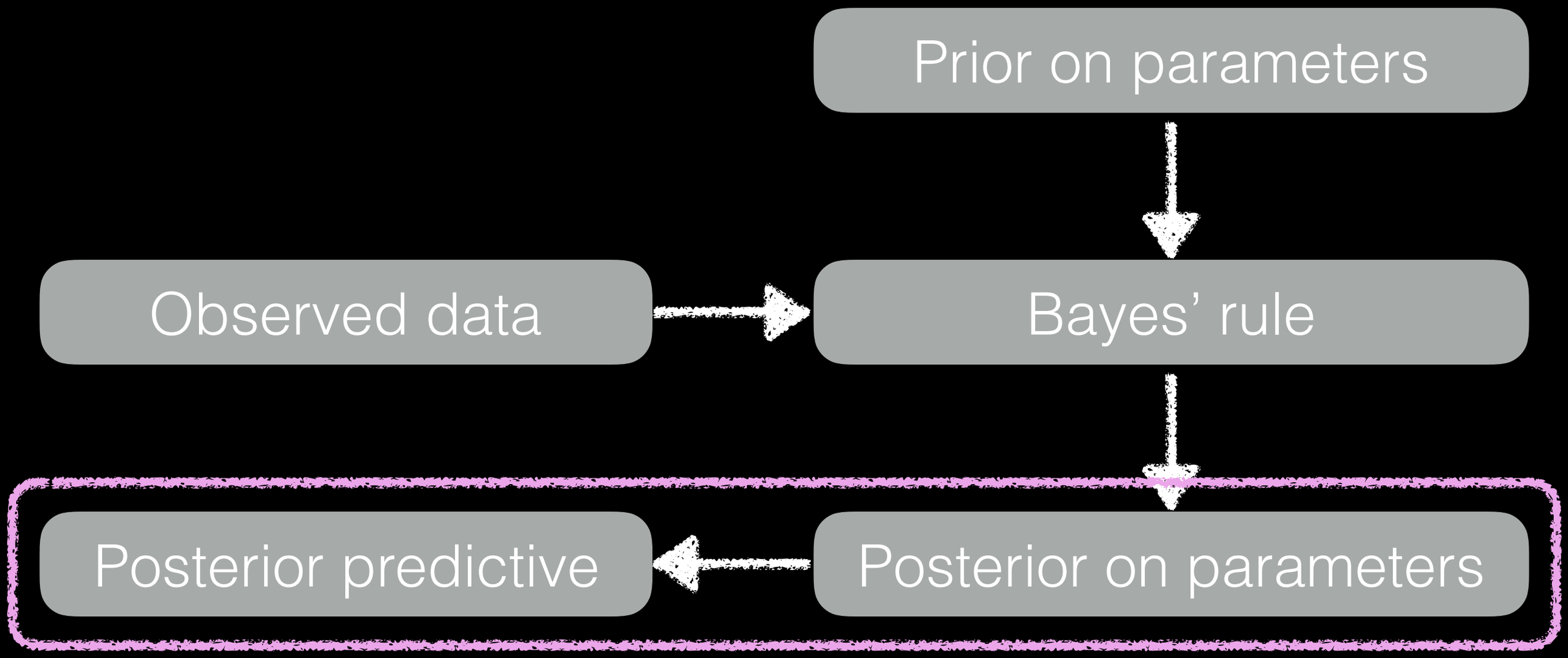


1. more redundant adjective use with greater scene variation
2. greater effect of scene variation for color than size

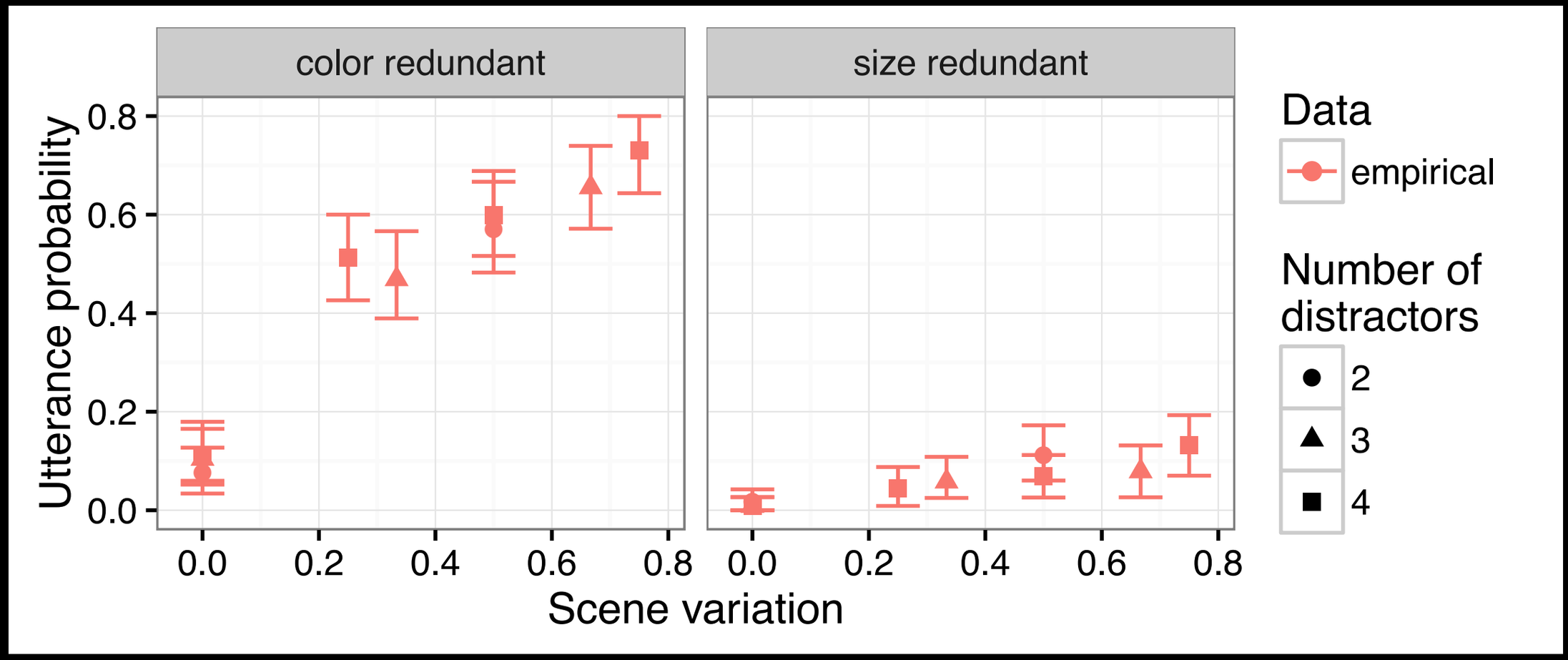
Bayesian data analysis



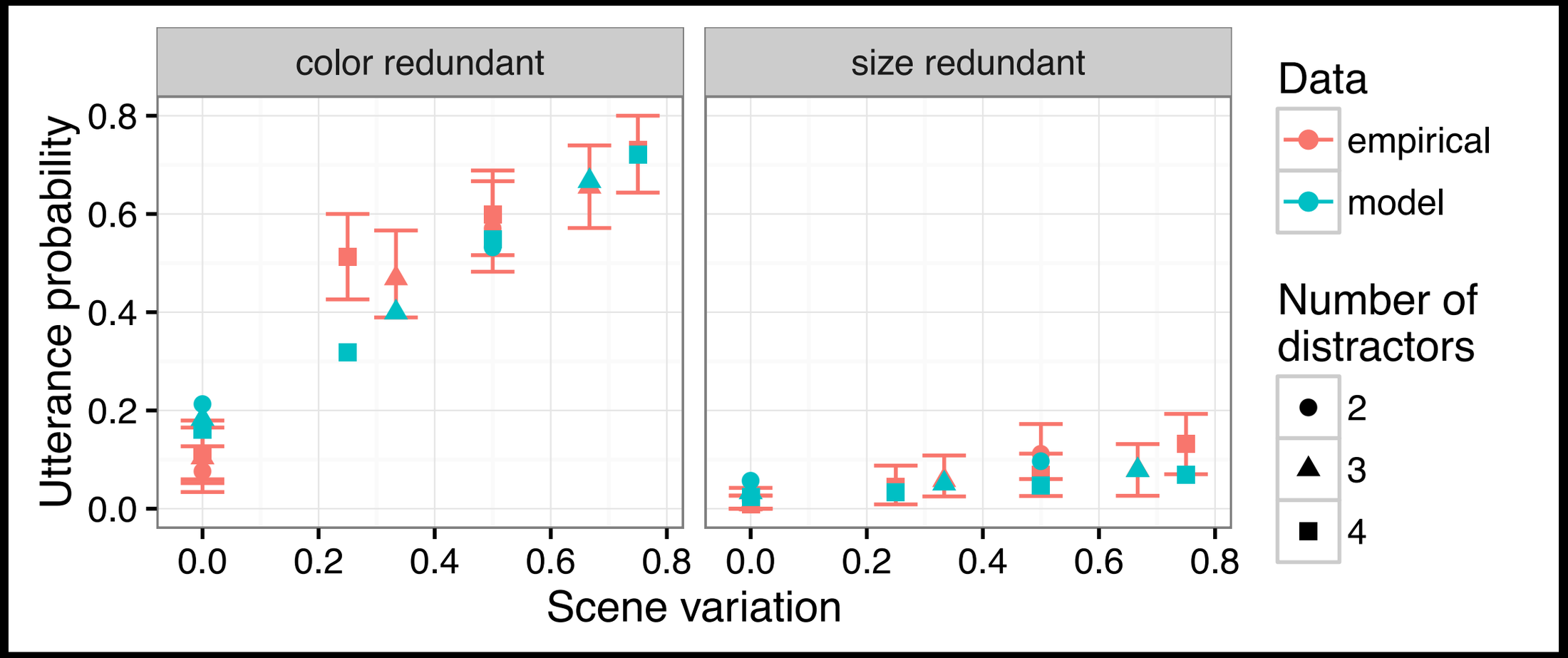
Bayesian data analysis



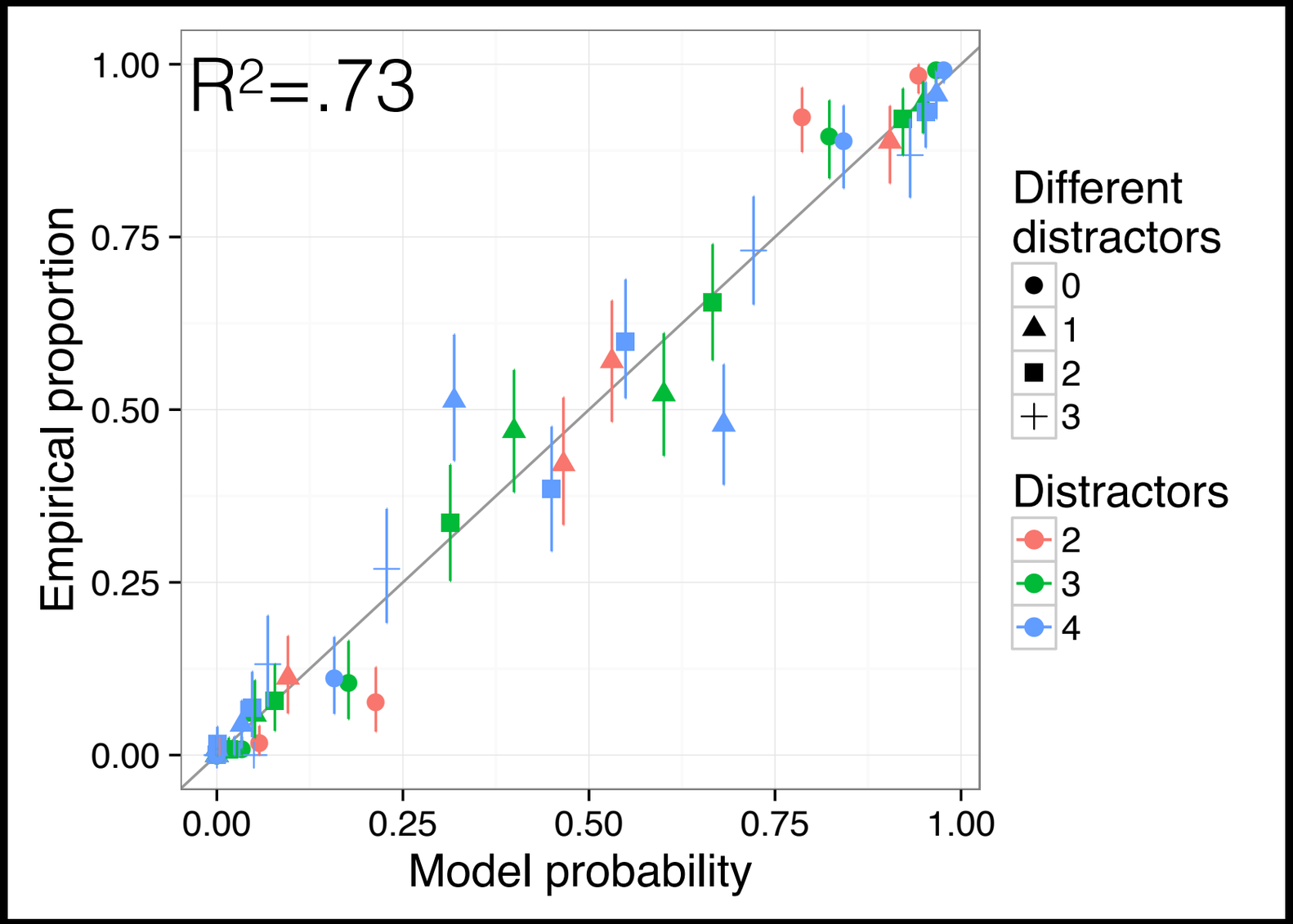
Results


color redundant
size redundant
Utterance probability
0.0
0.2
0.4
0.6
0.8
Scene variation
Data
empirical
Number of distractors
2
3
4

Results

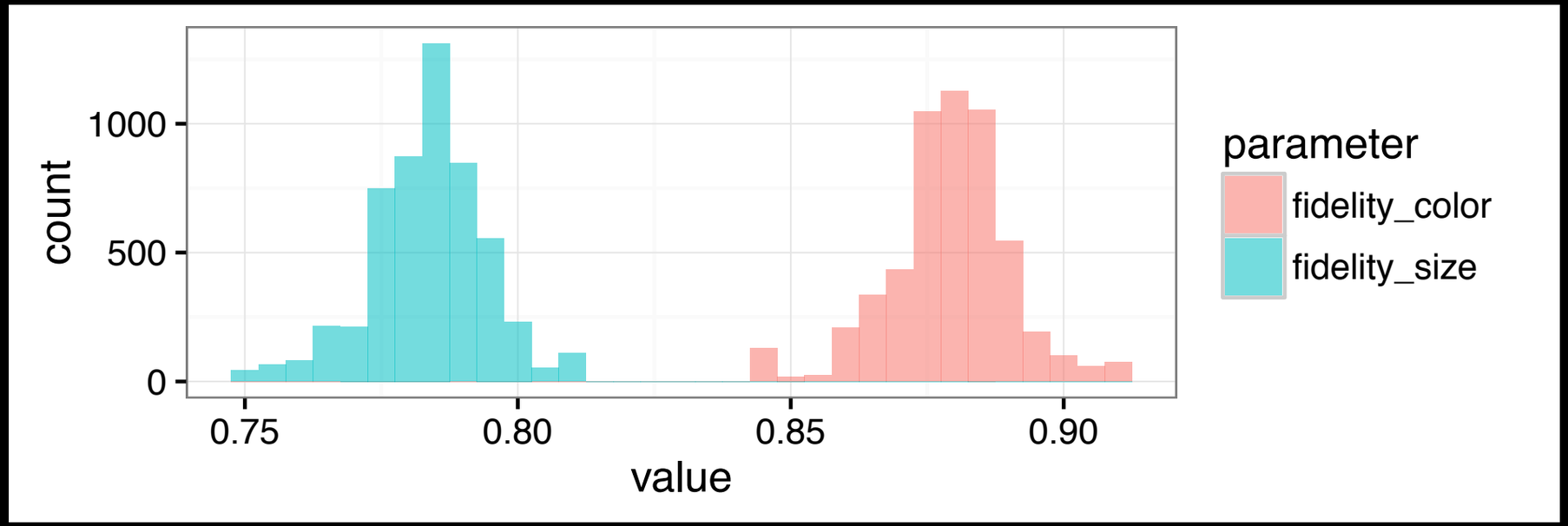

color redundant
size redundant
Utterance probability
0.0
0.2
0.4
0.6
0.8
Scene variation
Data
empirical
model
Number of distractors
2
3
4

Model fit

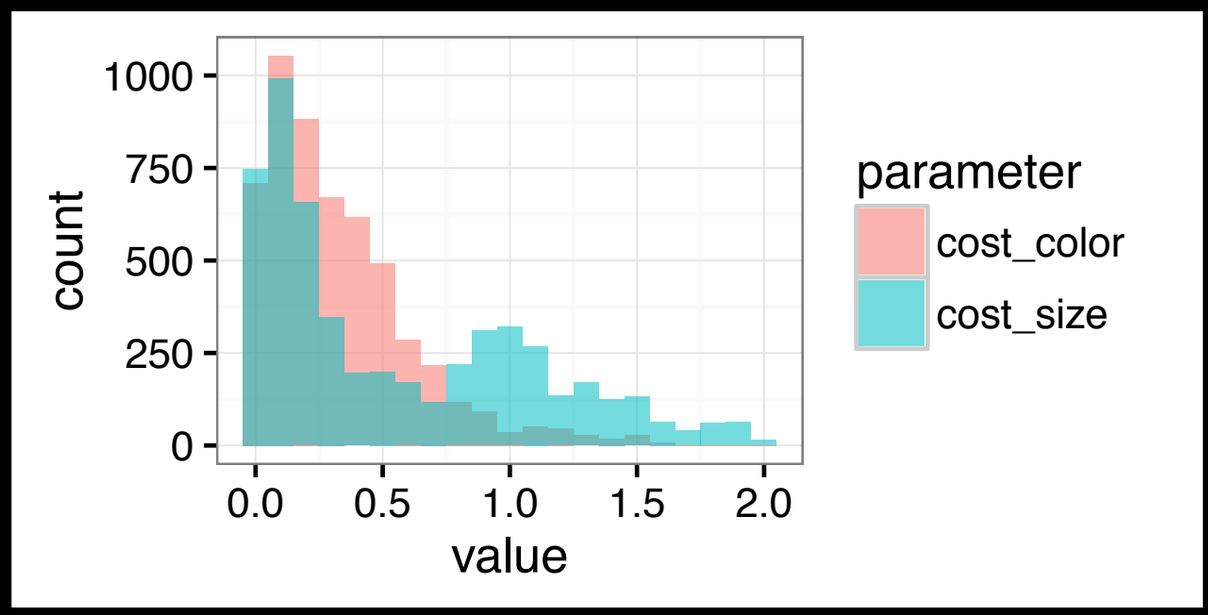

R²=.73
1.00
0.75
0.50
0.25
0.00
Empirical proportion
0.00
0.25
0.50
0.75
1.00
Model probability
Different distractors
0
1
2
3
Distractors
2
3
4

Posteriors over parameters

Fidelity:
inferred size fidelity lower than inferred color fidelity



Cost:
inferred size and color costs similar (with tendency towards costlier size)



Interim summary

if modifiers are noisy, adding modifiers adds utility

RSA with noisy truth functions captures this:

overinformative referring expressions

Interim summary

if modifiers are noisy, adding modifiers adds utility

RSA with noisy truth functions captures this:

~~overinformative referring expressions~~

Interim summary

if modifiers are noisy, adding modifiers adds utility

RSA with noisy truth functions captures this:

~~overinformative referring expressions~~

rationally redundant referring expressions

Interim summary

if modifiers are noisy, adding modifiers adds utility

RSA with noisy truth functions captures this:

~~overinformative referring expressions~~

rationally redundant referring expressions

level of reference

Graf et al 2016; Degen et al under review

typicality effects

Degen et al under review

Part II

Comprehension of scalar expressions

Scalar implicature

(1) John: Why is Ann happy?
Mary: She found some of her marbles.
Inference: Ann found some, but not all her marbles.

Generalization Grice 1975; Horn 1972, 2004
By uttering the weaker alternative from a scale of a weaker and a stronger alternative, the speaker implicates the negation of the stronger alternative.

Scalar implicature

(1) John: Why is Ann happy?
Mary: She found some of her marbles.
Inference: Ann found some, but not all her marbles.

Generalization Grice 1975; Horn 1972, 2004
By uttering the weaker alternative from a scale of a weaker and a stronger alternative, the speaker implicates the negation of the stronger alternative.

Pragmatic interpretation
…some, but not all…

Literal interpretation
…some, and possibly all…

Scalar implicature

(1) John: Why is Ann happy?
Mary: She found some of her marbles.
Inference: Ann found some, but not all her marbles.

Generalization Grice 1975; Horn 1972, 2004
By uttering the weaker alternative from a scale of a weaker and a stronger alternative, the speaker implicates the negation of the stronger alternative.

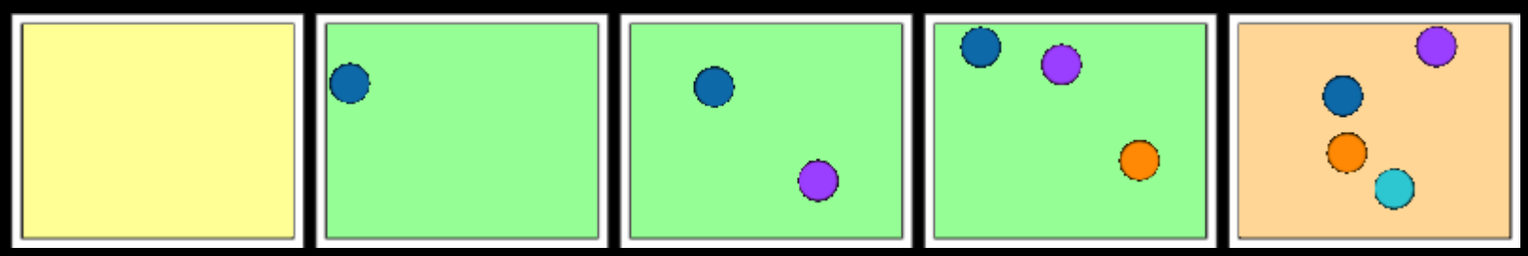
Pragmatic interpretation
…some, but not all…

Literal interpretation
…some, and possibly all…

"UNDERINFORMATIVENESS"

RSA for scalar implicature

$$M = \{m_{\neg\exists}, m_{\exists\neg\forall}, m_{\forall}\}$$
$$U = \{u_{\text{none}}, u_{\text{some}}, u_{\text{all}}\}$$
$$[[u_{\text{none}}]] = \{m_{\neg\exists}\}$$
$$[[u_{\text{some}}]] = \{m_{\exists\neg\forall}, m_{\forall}\}$$
$$[[u_{\text{all}}]] = \{m_{\forall}\}$$



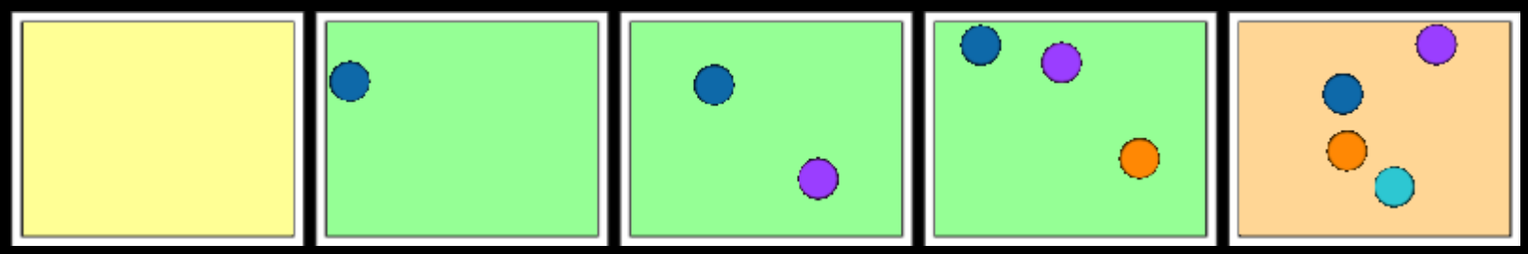
Ann found some of her marbles

RSA for scalar implicature

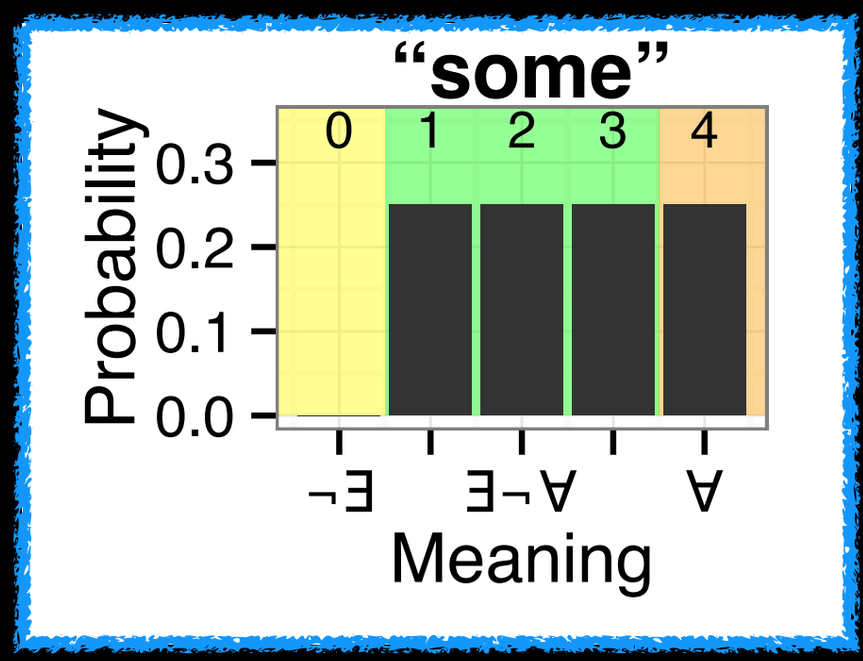
Literal listener

$$P_{L_0}(m|u) = \mathcal{U}(m|\{u \text{ is true of } m\})$$

$$M = \{m_{\neg\exists}, m_{\exists\neg\forall}, m_{\forall}\}$$
$$U = \{u_{\text{none}}, u_{\text{some}}, u_{\text{all}}\}$$
$$[[u_{\text{none}}]] = \{m_{\neg\exists}\}$$
$$[[u_{\text{some}}]] = \{m_{\exists\neg\forall}, m_{\forall}\}$$
$$[[u_{\text{all}}]] = \{m_{\forall}\}$$



Ann found some of her marbles

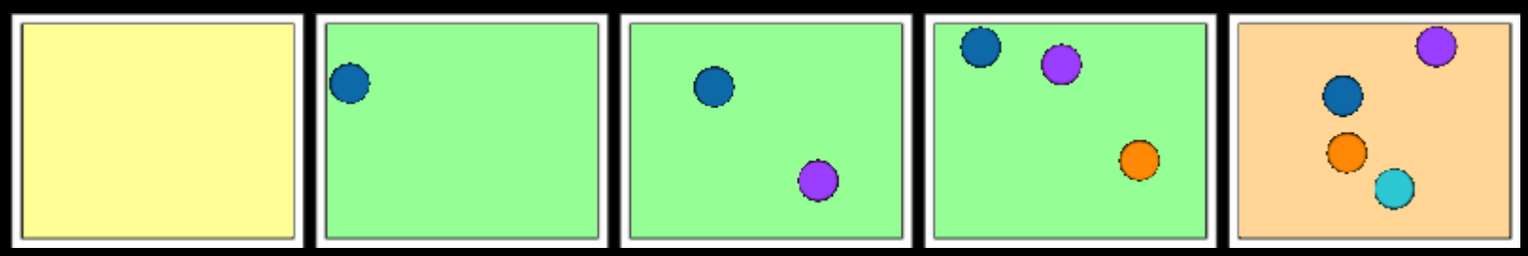


RSA for scalar implicature

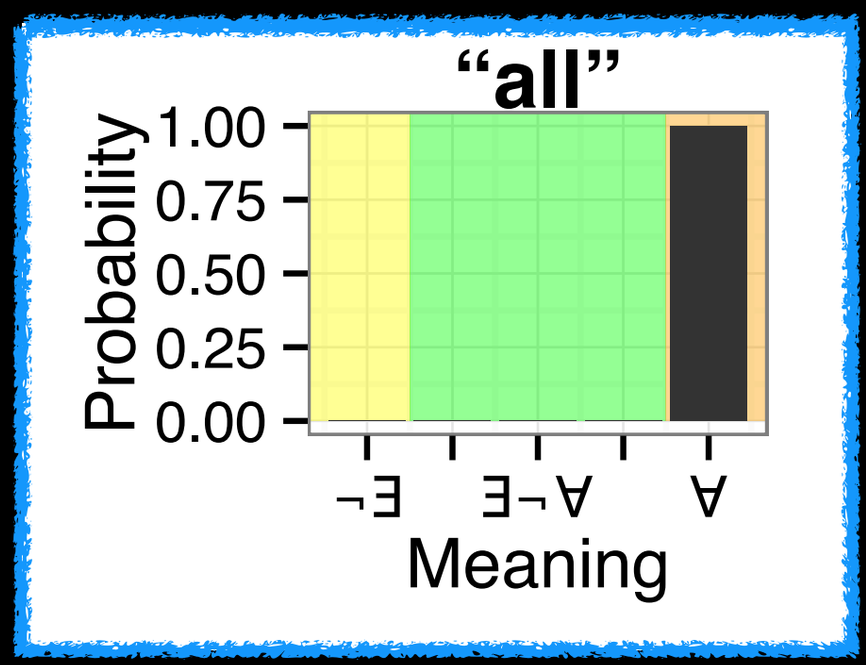
Literal listener

$$P_{L_0}(m|u) = \mathcal{U}(m|\{u \text{ is true of } m\})$$

$$M = \{m_{\neg\exists}, m_{\exists\neg\forall}, m_{\forall}\}$$
$$U = \{u_{\text{none}}, u_{\text{some}}, u_{\text{all}}\}$$
$$[[u_{\text{none}}]] = \{m_{\neg\exists}\}$$
$$[[u_{\text{some}}]] = \{m_{\exists\neg\forall}, m_{\forall}\}$$
$$[[u_{\text{all}}]] = \{m_{\forall}\}$$



Ann found some of her marbles



RSA for scalar implicature

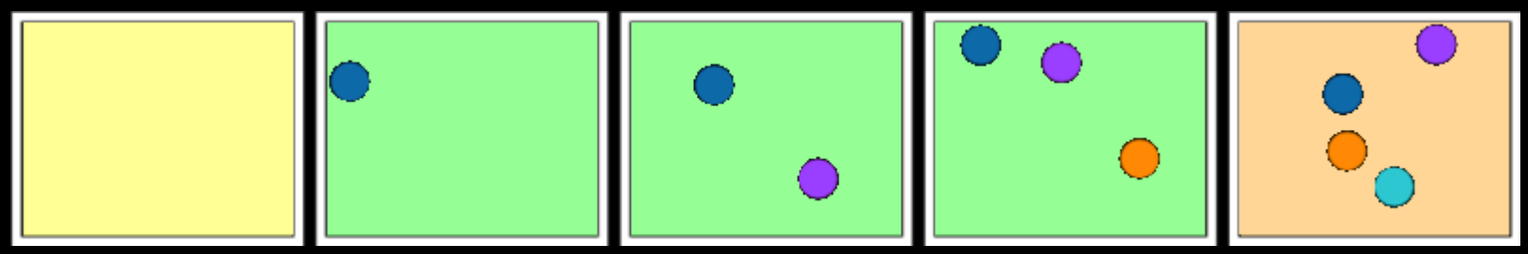
Literal listener

$$P_{L_0}(m|u) = \mathcal{U}(m|\{u \text{ is true of } m\})$$

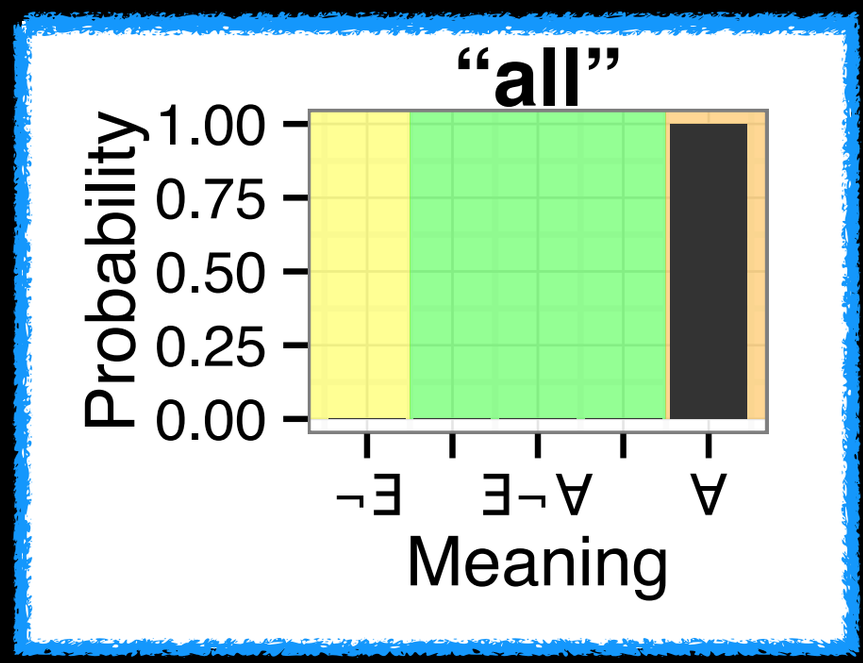
Pragmatic speaker

$$P_{S_1}(u|m) \propto e^{\lambda \cdot (\ln P_{L_0}(m|u))}$$

$$M = \{m_{\neg\exists}, m_{\exists\neg\forall}, m_{\forall}\}$$
$$U = \{u_{\text{none}}, u_{\text{some}}, u_{\text{all}}\}$$
$$[[u_{\text{none}}]] = \{m_{\neg\exists}\}$$
$$[[u_{\text{some}}]] = \{m_{\exists\neg\forall}, m_{\forall}\}$$
$$[[u_{\text{all}}]] = \{m_{\forall}\}$$



Ann found some of her marbles



RSA for scalar implicature

Literal listener

$$P_{L_0}(m|u) = \mathcal{U}(m|\{u \text{ is true of } m\})$$

Pragmatic speaker

$$P_{S_1}(u|m) \propto e^{\lambda \cdot (\ln P_{L_0}(m|u))}$$

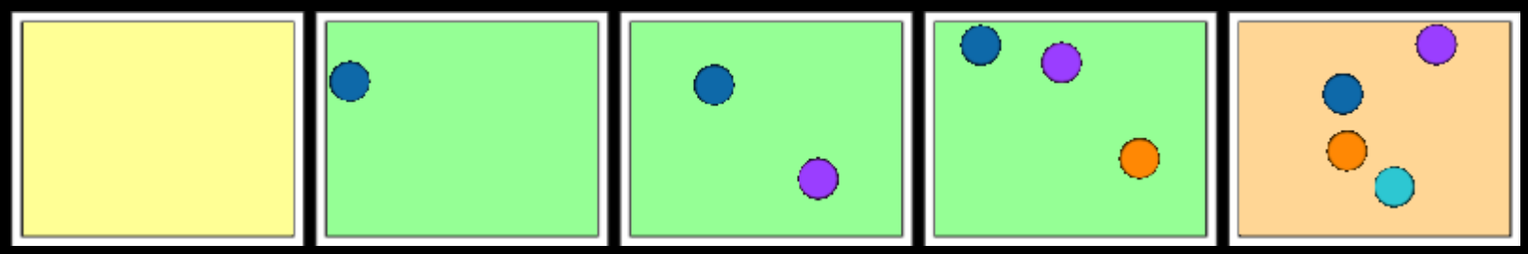
$$M = \{m_{\neg\exists}, m_{\exists\neg\forall}, m_{\forall}\}$$

$$U = \{u_{\text{none}}, u_{\text{some}}, u_{\text{all}}\}$$

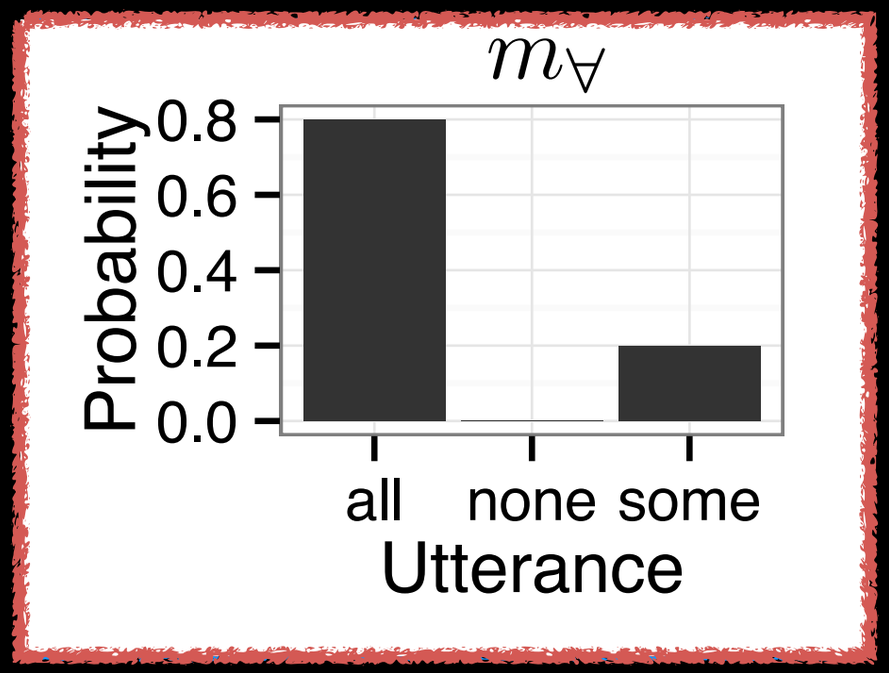
$$[[u_{\text{none}}]] = \{m_{\neg\exists}\}$$

$$[[u_{\text{some}}]] = \{m_{\exists\neg\forall}, m_{\forall}\}$$

$$[[u_{\text{all}}]] = \{m_{\forall}\}$$



Ann found some of her marbles



RSA for scalar implicature

Literal listener

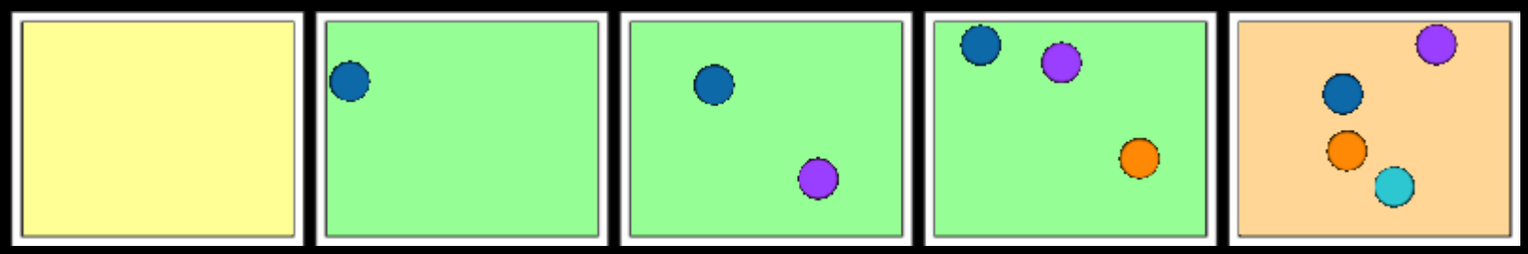
$$P_{L_0}(m|u) = \mathcal{U}(m|\{u \text{ is true of } m\})$$

Pragmatic speaker

$$P_{S_1}(u|m) \propto e^{\lambda \cdot (\ln P_{L_0}(m|u))}$$

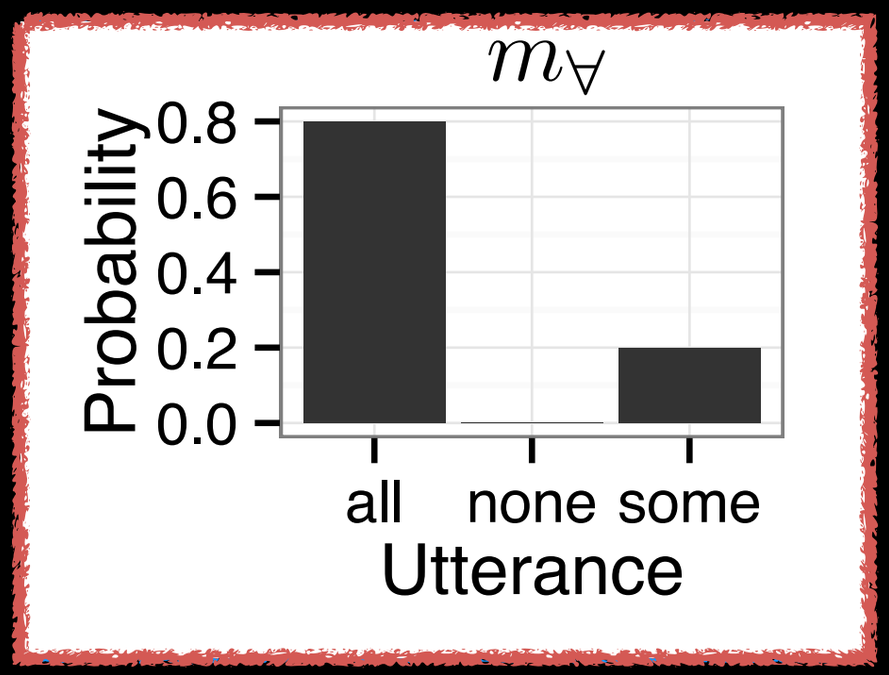
Pragmatic listener

$$P_{L_1}(m|u) \propto P_{S_1}(u|m) \cdot P(m)$$



Ann found some of her marbles

$$M = \{m_{\neg\exists}, m_{\exists\neg\forall}, m_{\forall}\}$$
$$U = \{u_{\text{none}}, u_{\text{some}}, u_{\text{all}}\}$$
$$[[u_{\text{none}}]] = \{m_{\neg\exists}\}$$
$$[[u_{\text{some}}]] = \{m_{\exists\neg\forall}, m_{\forall}\}$$
$$[[u_{\text{all}}]] = \{m_{\forall}\}$$



RSA for scalar implicature

Literal listener

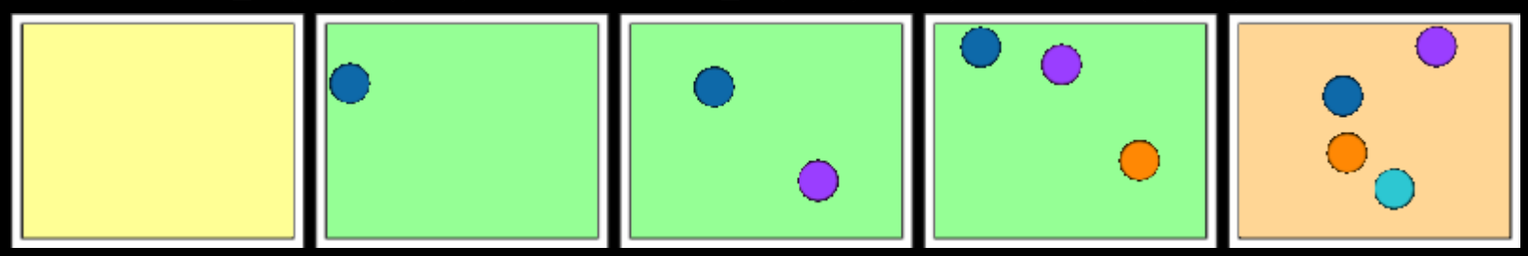
$$P_{L_0}(m|u) = \mathcal{U}(m|\{u \text{ is true of } m\})$$

Pragmatic speaker

$$P_{S_1}(u|m) \propto e^{\lambda \cdot (\ln P_{L_0}(m|u))}$$

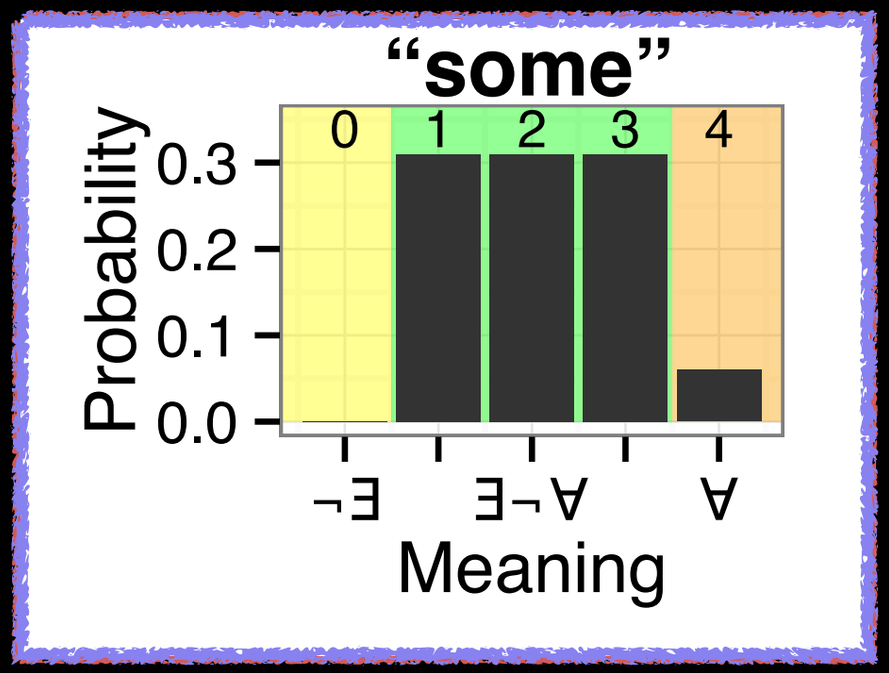
Pragmatic listener

$$P_{L_1}(m|u) \propto P_{S_1}(u|m) \cdot P(m)$$



Ann found some of her marbles

$$M = \{m_{\neg\exists}, m_{\exists\neg\forall}, m_{\forall}\}$$
$$U = \{u_{\text{none}}, u_{\text{some}}, u_{\text{all}}\}$$
$$[[u_{\text{none}}]] = \{m_{\neg\exists}\}$$
$$[[u_{\text{some}}]] = \{m_{\exists\neg\forall}, m_{\forall}\}$$
$$[[u_{\text{all}}]] = \{m_{\forall}\}$$



RSA for scalar implicature

Literal listener

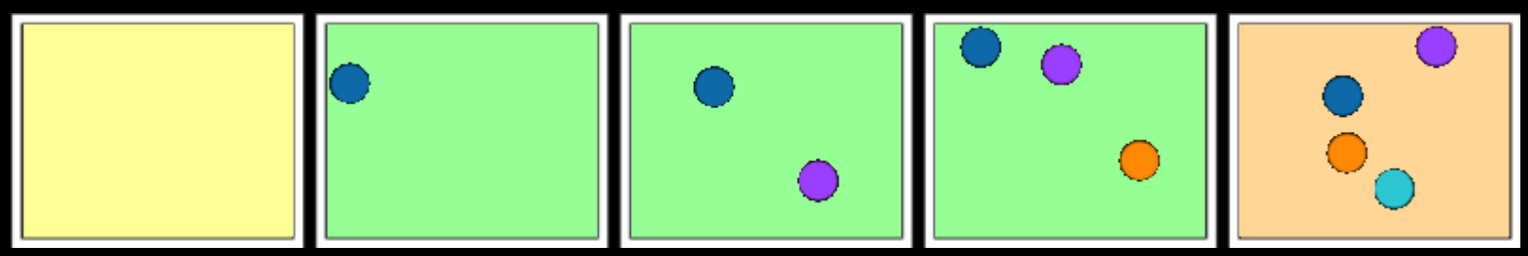
$$P_{L_0}(m|u) = \mathcal{U}(m|\{u \text{ is true of } m\})$$

Pragmatic speaker

$$P_{S_1}(u|m) \propto e^{\lambda \cdot (\ln P_{L_0}(m|u))}$$

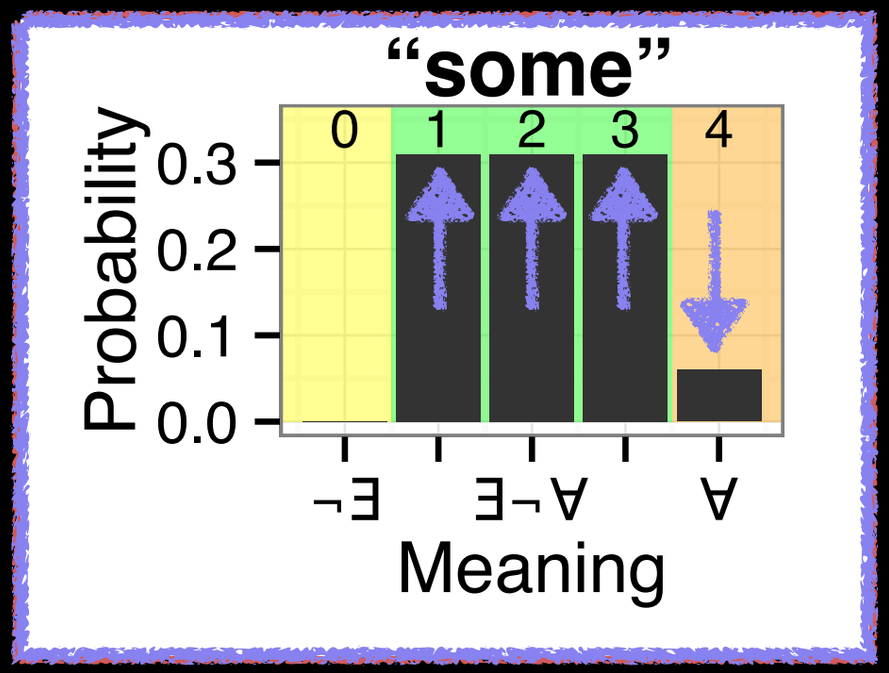
Pragmatic listener

$$P_{L_1}(m|u) \propto P_{S_1}(u|m) \cdot P(m)$$



Ann found some of her marbles

$$M = \{m_{\neg\exists}, m_{\exists\neg\forall}, m_{\forall}\}$$
$$U = \{u_{\text{none}}, u_{\text{some}}, u_{\text{all}}\}$$
$$[[u_{\text{none}}]] = \{m_{\neg\exists}\}$$
$$[[u_{\text{some}}]] = \{m_{\exists\neg\forall}, m_{\forall}\}$$
$$[[u_{\text{all}}]] = \{m_{\forall}\}$$



Context in scalar implicature

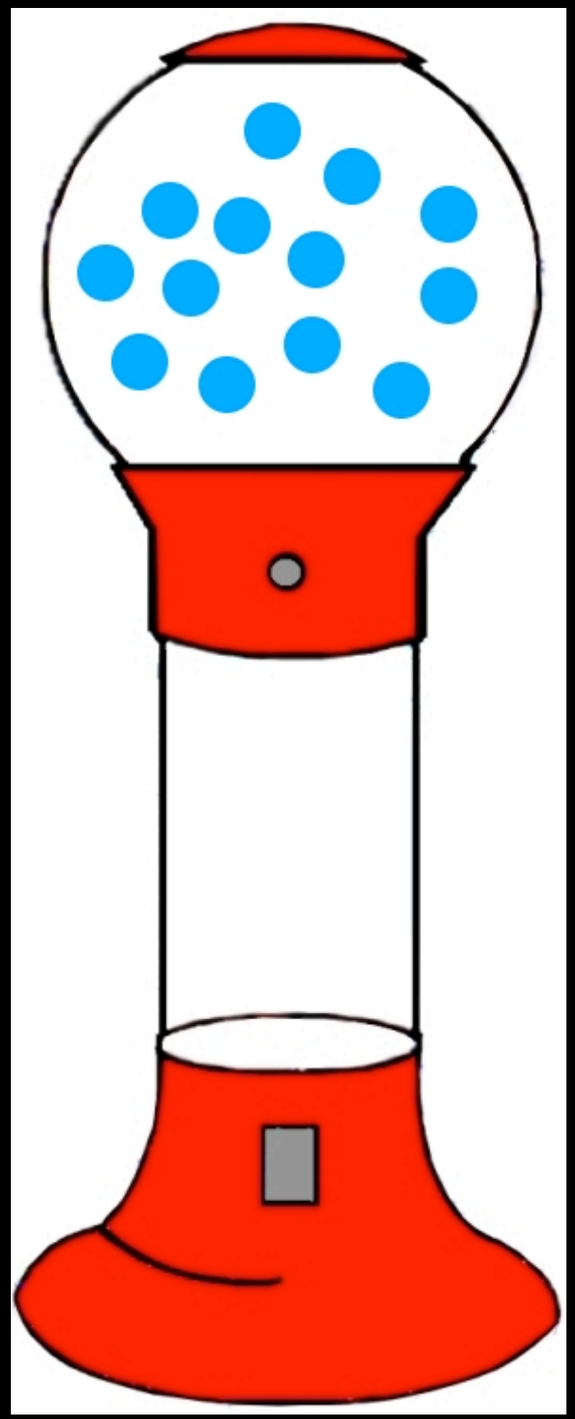
- pragmatic factors:
  - world knowledge/prior beliefs Degen et al 2015; Javangula & Degen in prep
  - speaker knowledge Bergen & Grodner 2012; Goodman & Stuhlmüller 2013; Breheny et al 2013
    - utterance alternatives Degen & Tanenhaus 2015; 2016; Rees & Bott 2018; Bott & Chemla 2016
    - conversational goal / Question Under Discussion (QUD) Zondervan 2010; Degen 2013; Degen & Goodman 2014
- syntactic/semantic factors:
  - partitive, linguistic mention, determiner strength, subjecthood, information structure, prosody, monotonicity Degen 2015; Breheny et al 2006; Katsos et al 2005; de Marneffe & Tonhauser 2016;

Context in scalar implicature

- pragmatic factors:
  - world knowledge/prior beliefs Degen et al 2015; Javangula & Degen in prep
  - speaker knowledge Bergen & Grodner 2012; Goodman & Stuhlmüller 2013; Breheny et al 2013
    - **utterance alternatives** Degen & Tanenhaus 2015; 2016; Rees & Bott 2018; Bott & Chemla 2016
    - **conversational goal / Question Under Discussion (QUD)** Zondervan 2010; Degen 2013; Degen & Goodman 2014
- syntactic/semantic factors:
  - partitive, linguistic mention, determiner strength, subjecthood, information structure, prosody, monotonicity Degen 2015; Breheny et al 2006; Katsos et al 2005; de Marneffe & Tonhauser 2016;

Number alternatives in processing “some”

Degen & Tanenhaus 2015

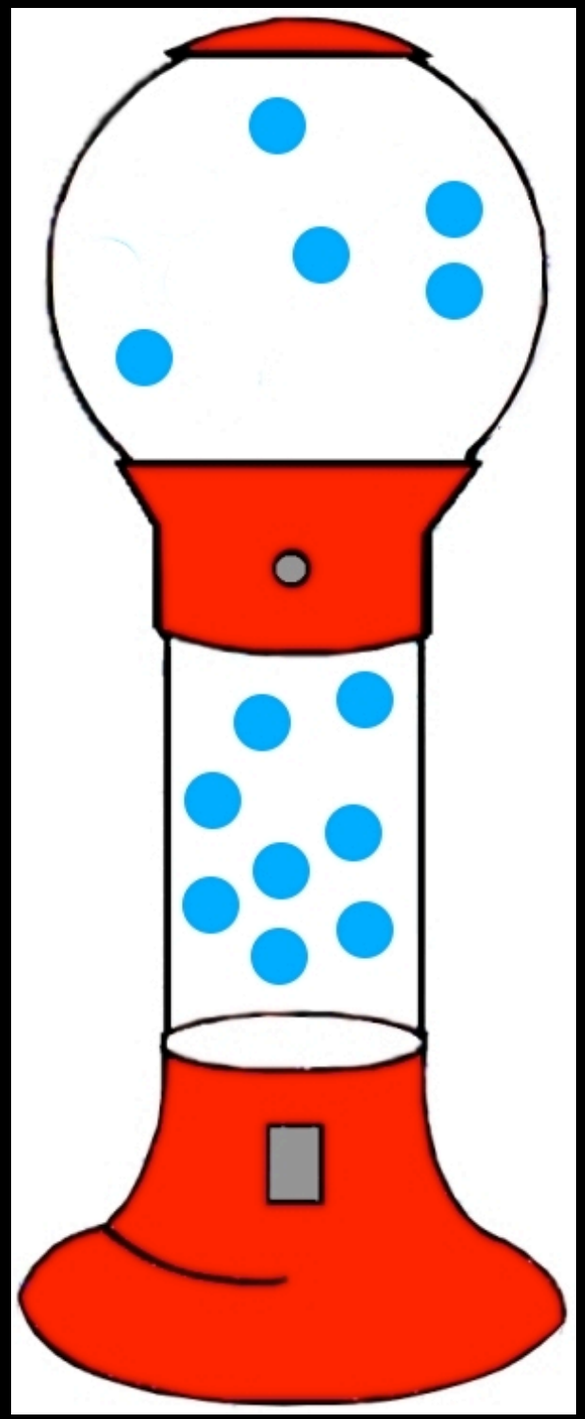


Naturalness ratings

- Do listeners have **expectations** about use of *some*? P(*some* | M)
- Do expectations depend on the contextual availability of **number alternatives**? P(*some* | M, C)

Number alternatives in processing “some”

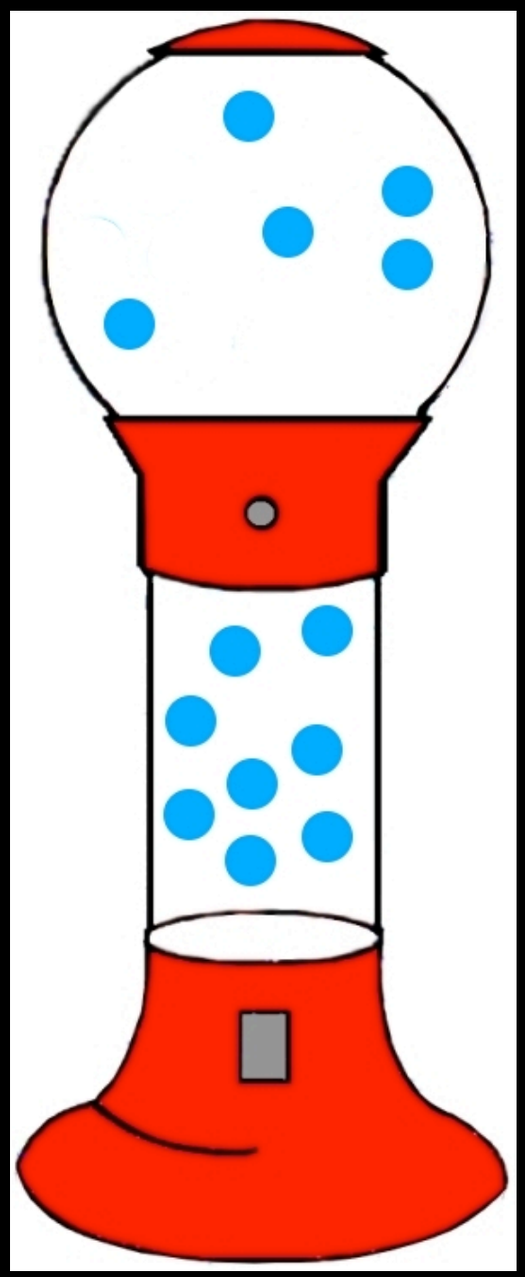
Degen & Tanenhaus 2015



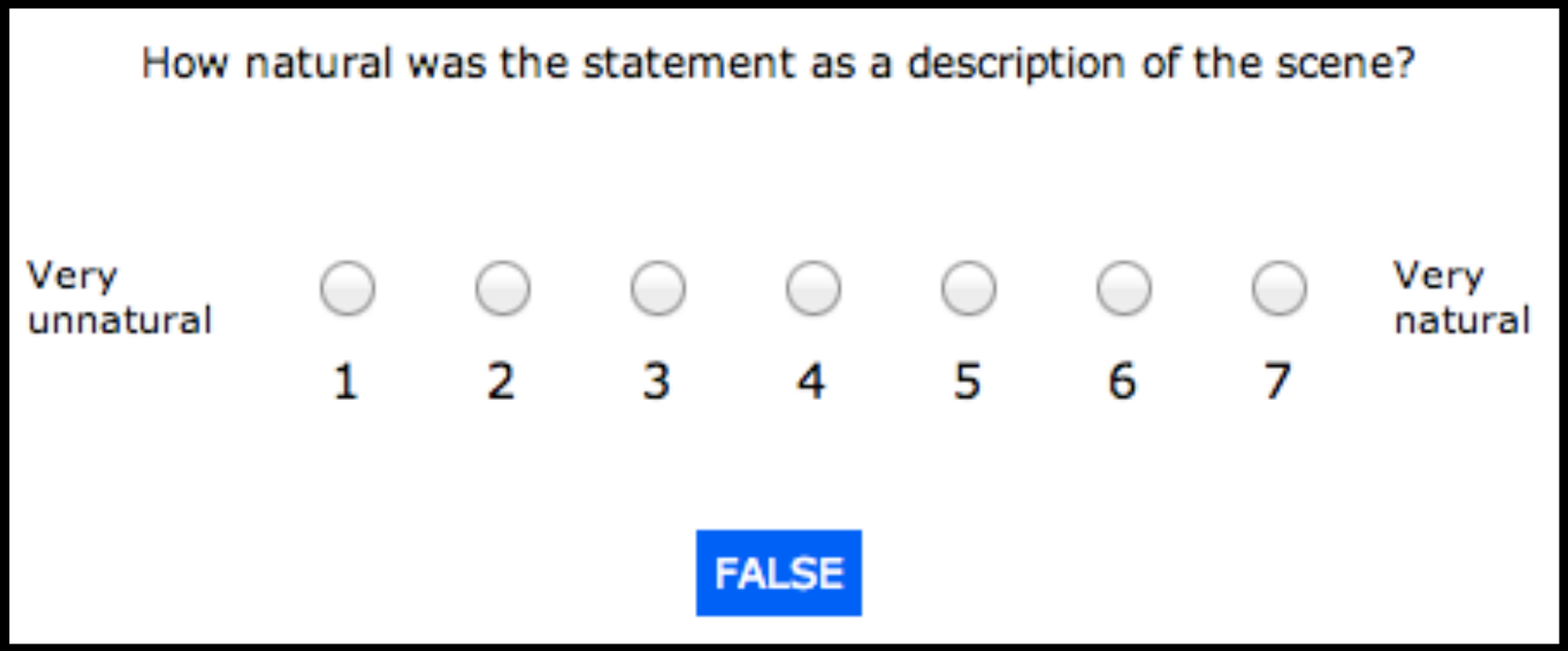
Naturalness ratings

- Do listeners have **expectations** about use of *some*? P(*some* | M)
- Do expectations depend on the contextual availability of **number alternatives**? P(*some* | M, C)

The gumball paradigm



You got some of the gumballs

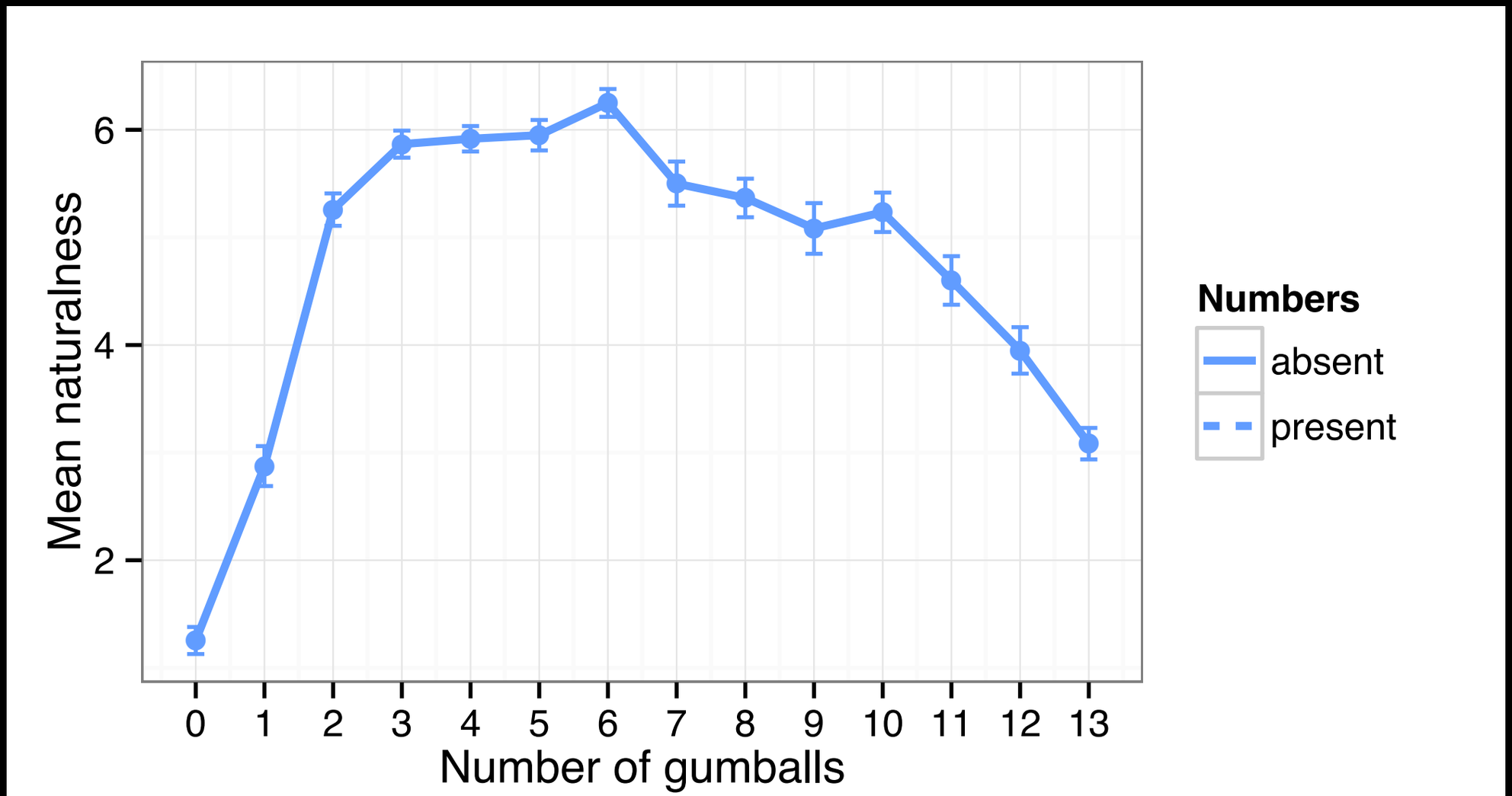


360 participants on MTurk
Independent variables:

- set size in lower chamber: 0 - 13
- quantifier: *some, all, none,* (*one, two*,…)
- presence of number terms

Expectations of use for *some*

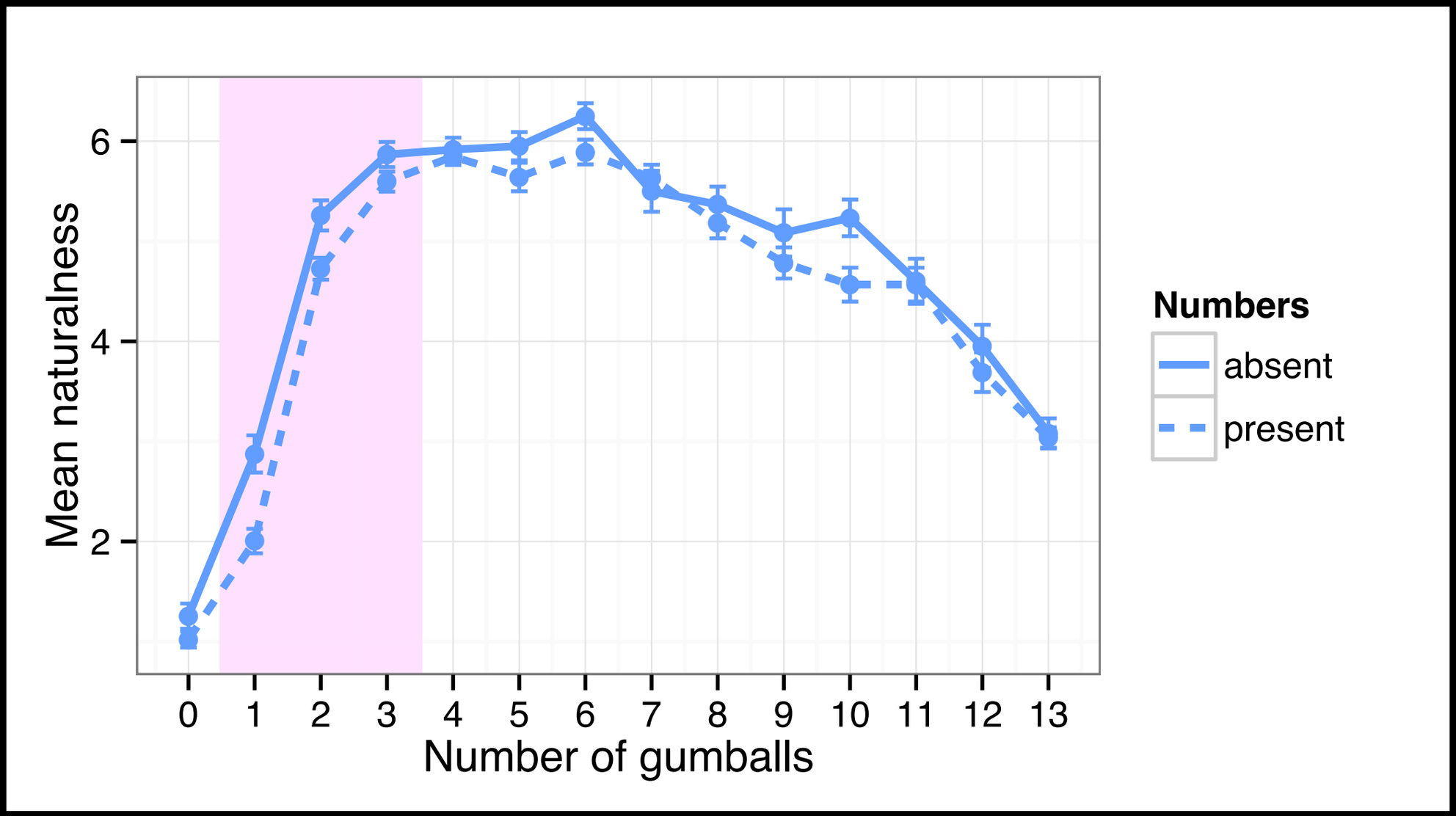
cf van Tiel 2014



some is a dispreferred alternative for small sets ($p < .0001$)

Expectations of use for *some*

cf van Tiel 2014

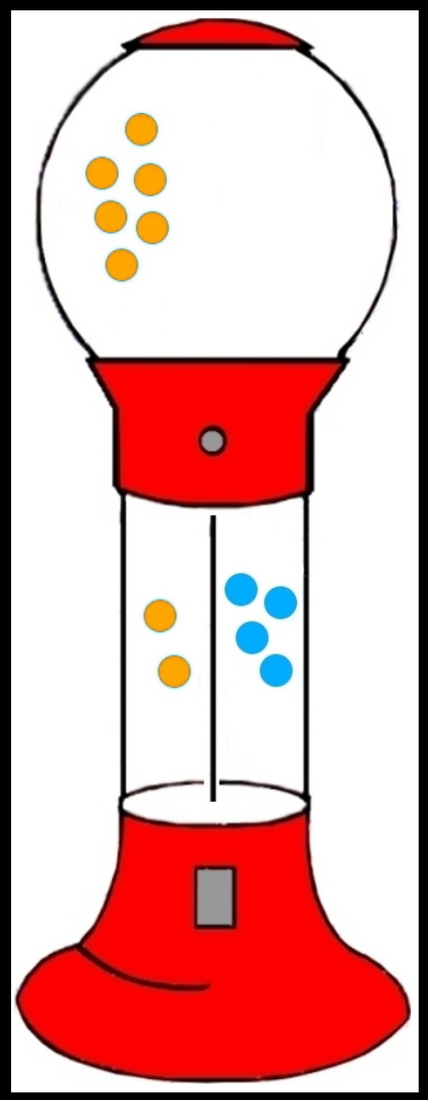


some is a dispreferred alternative for small sets ($p < .0001$)

especially when numbers are available alternatives ($p < .01$)

Processing alternatives online

Degen & Tanenhaus 2016

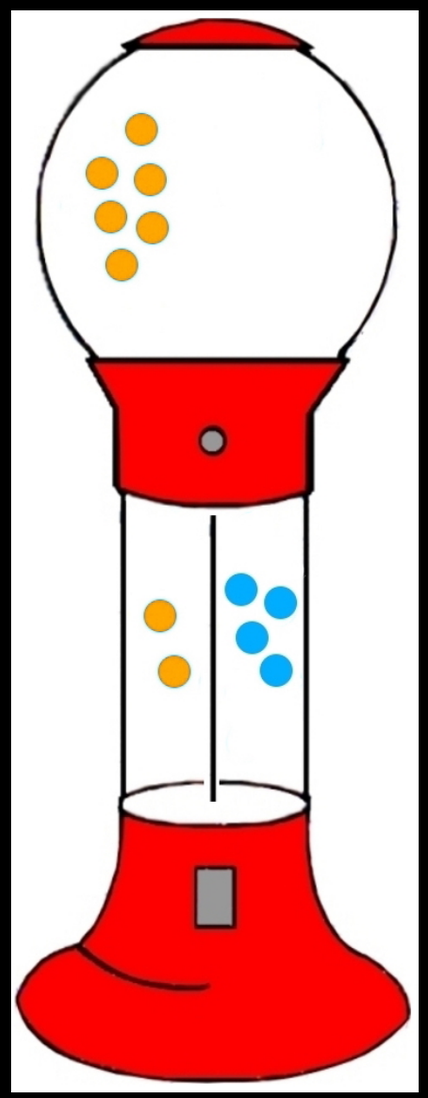


You got some of the orange gumballs

Processing alternatives online

Degen & Tanenhaus 2016

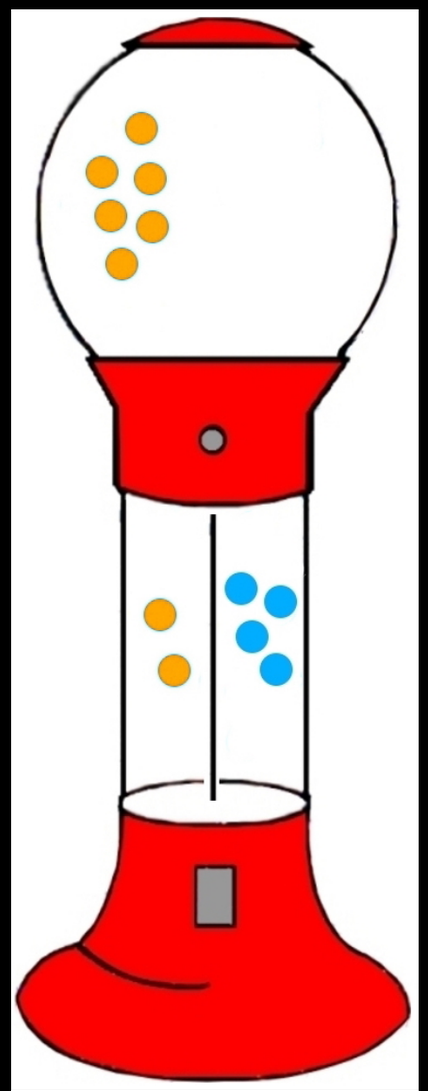
absent: *some/all* **present**: *some/all/two/three/four/five*

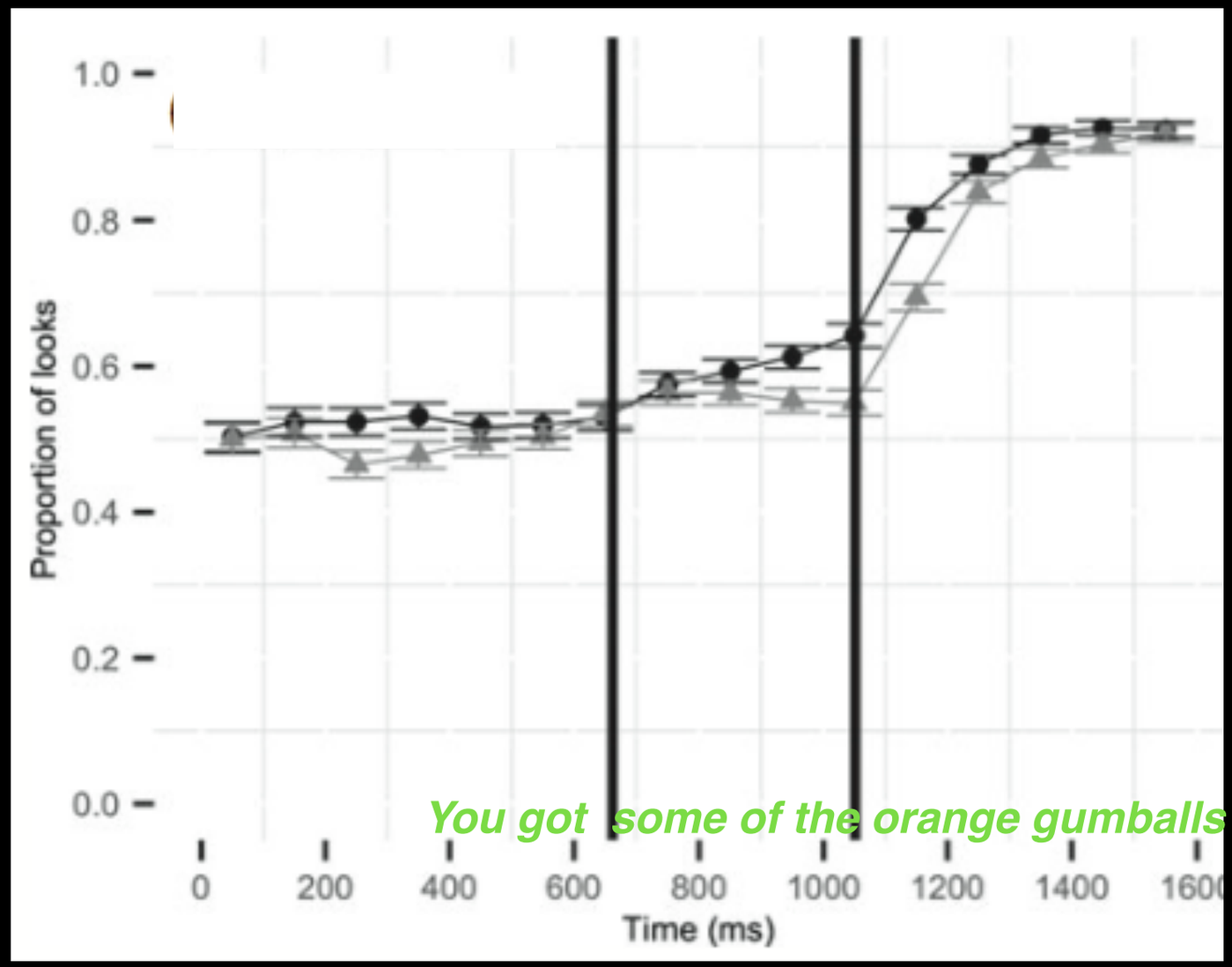


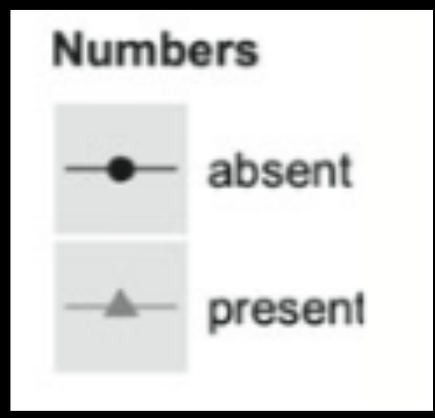
Processing alternatives online

Degen & Tanenhaus 2016

absent: *some/all* **present**: *some/all/two/three/four/five*



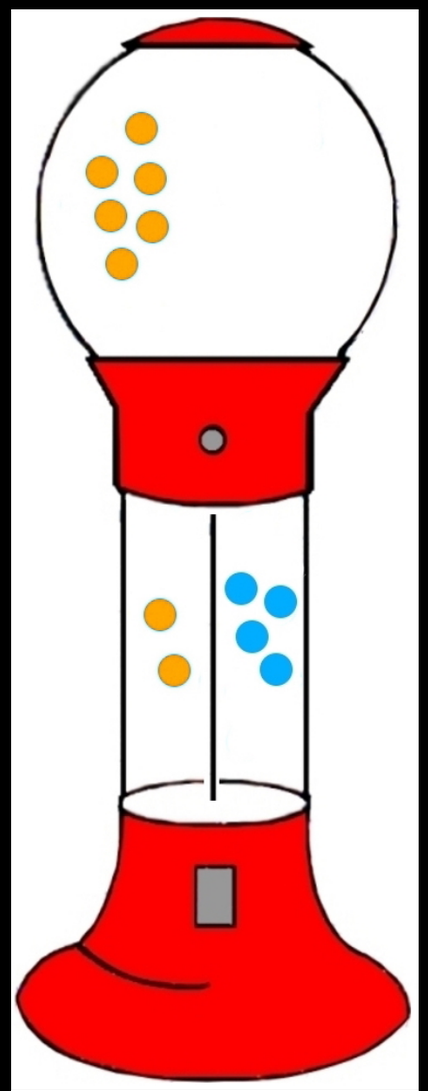


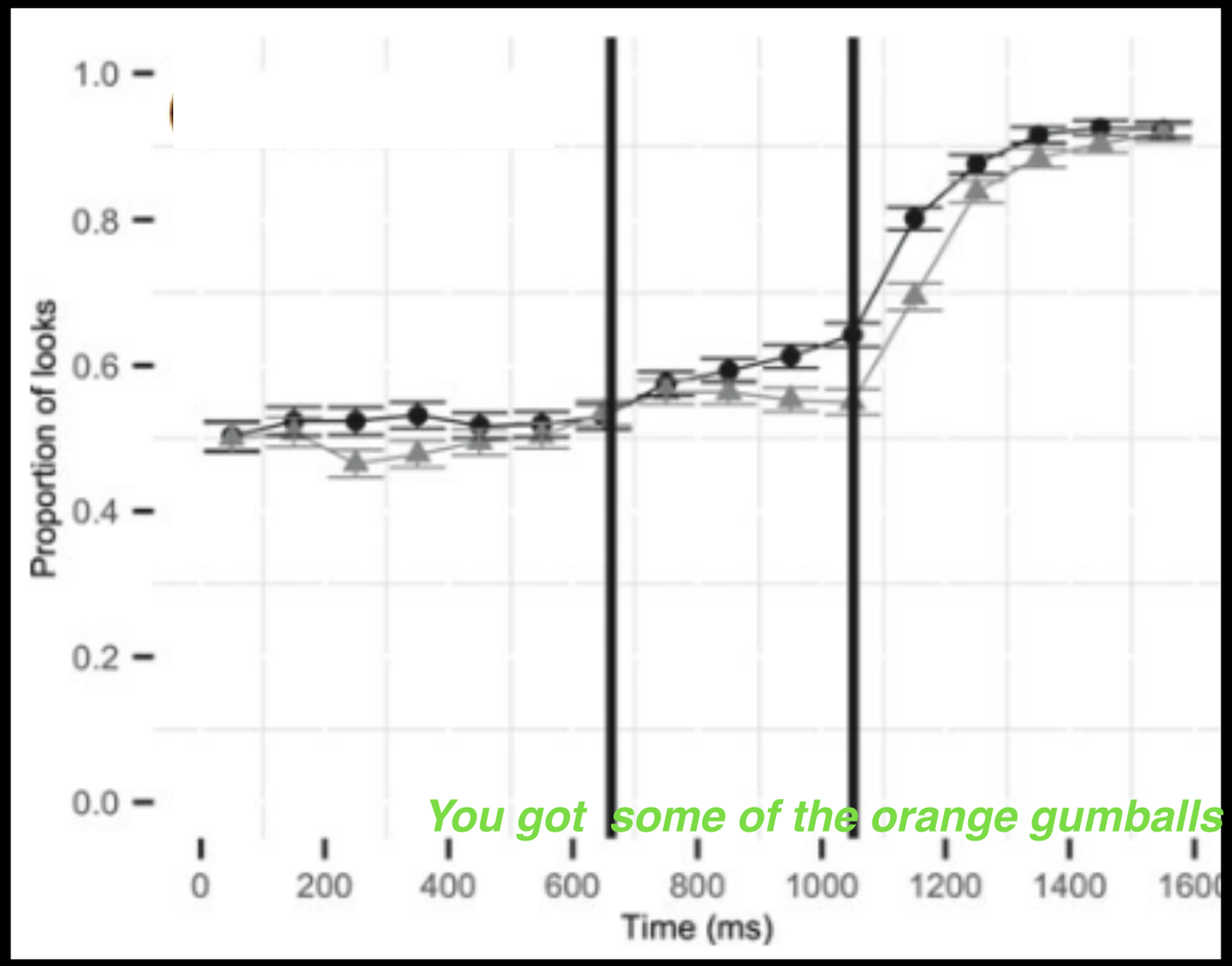


Processing alternatives online

Degen & Tanenhaus 2016

absent: *some/all* **present**: *some/all/two/three/four/five*



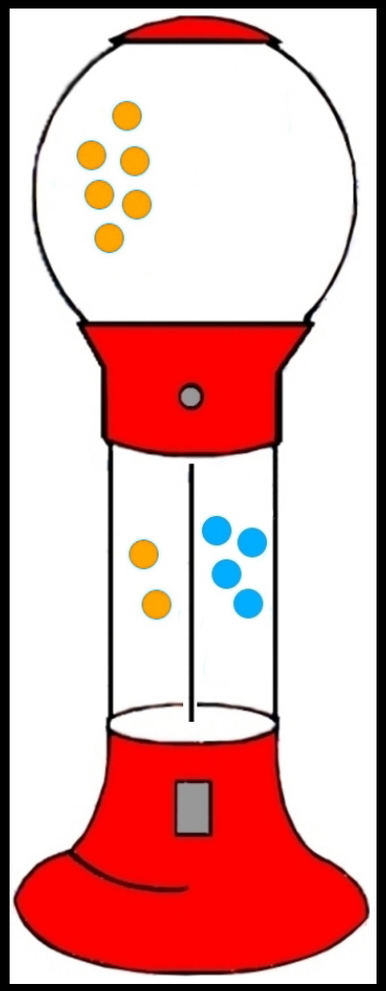


implicatures are slower to process when number alternatives are contextually available

Context in RSA: alternatives

Absent: U = {none, some, all}

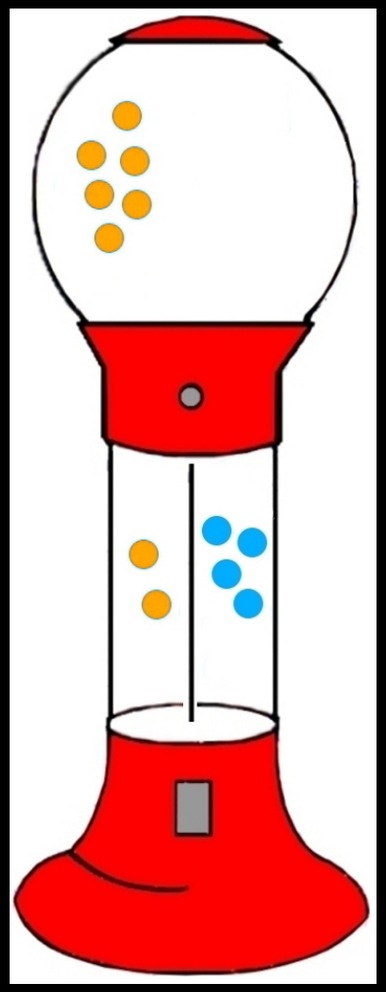
Present: U = {none, some, all, one, two, three, four}



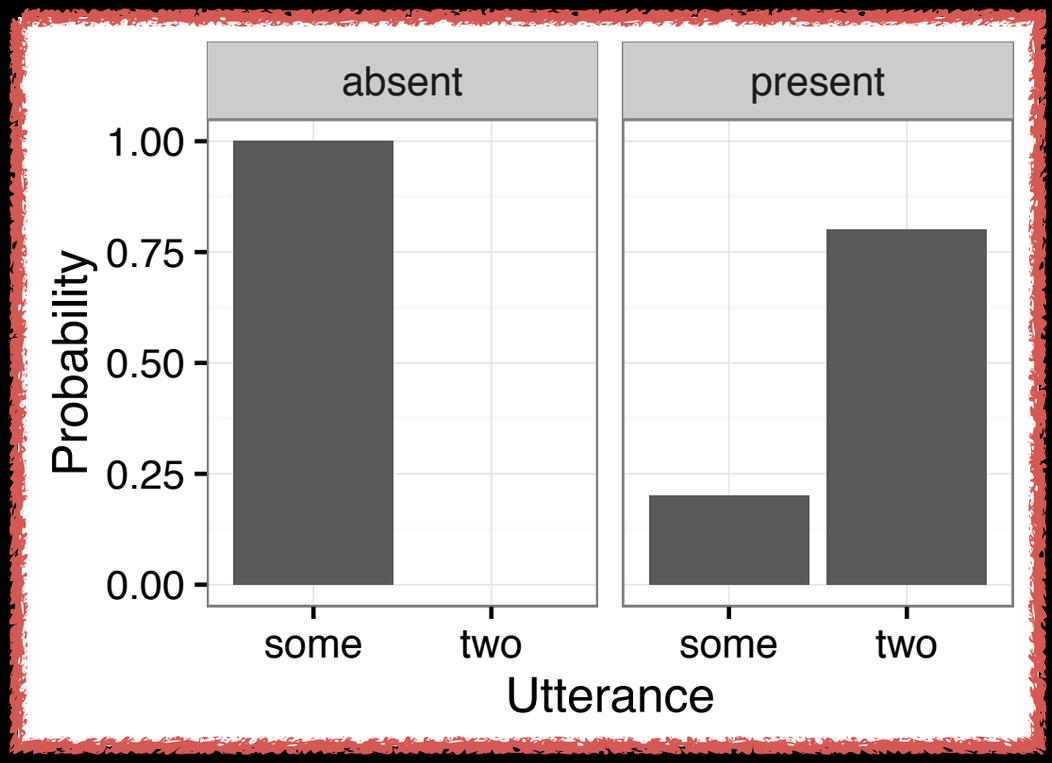
Context in RSA: alternatives

Absent: U = {none, some, all}

Present: U = {none, some, all, one, two, three, four}



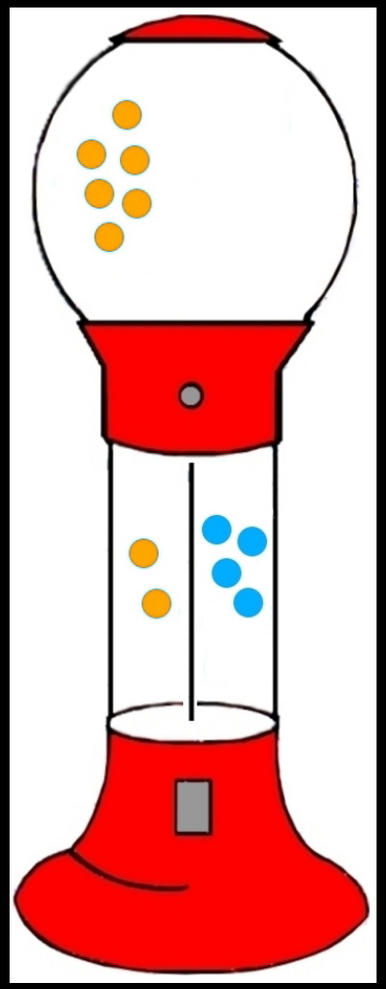
Pragmatic speaker
state = 2



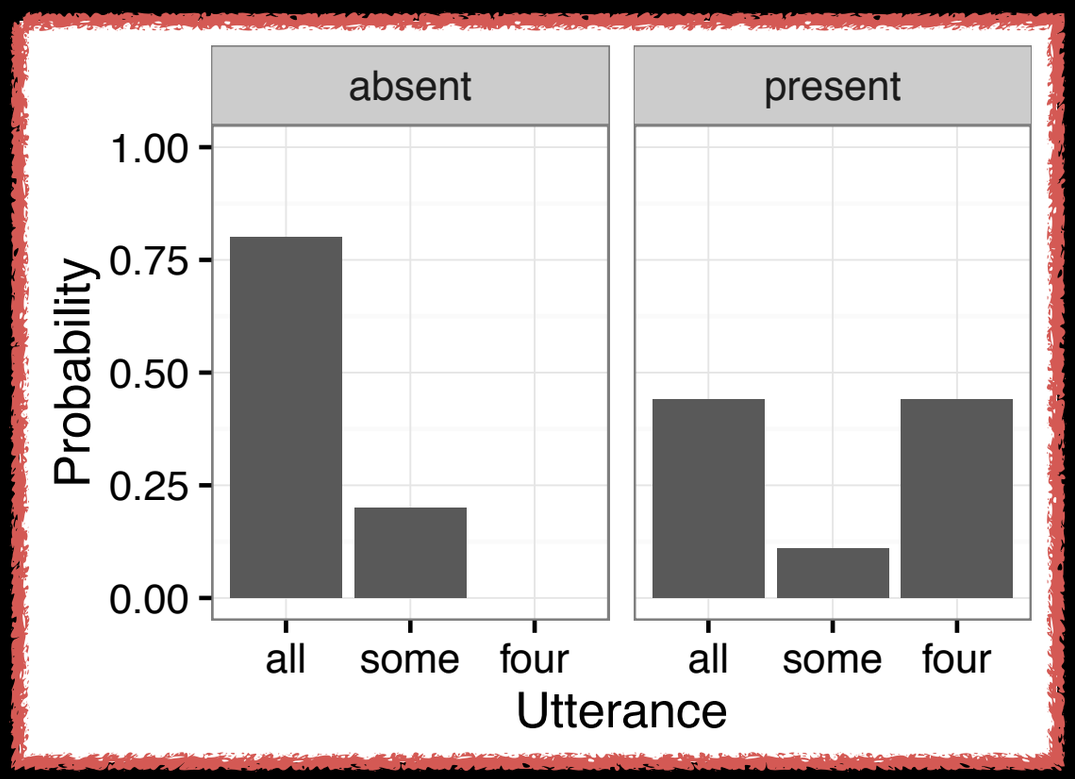
Context in RSA: alternatives

Absent: U = {none, some, all}

Present: U = {none, some, all, one, two, three, four}



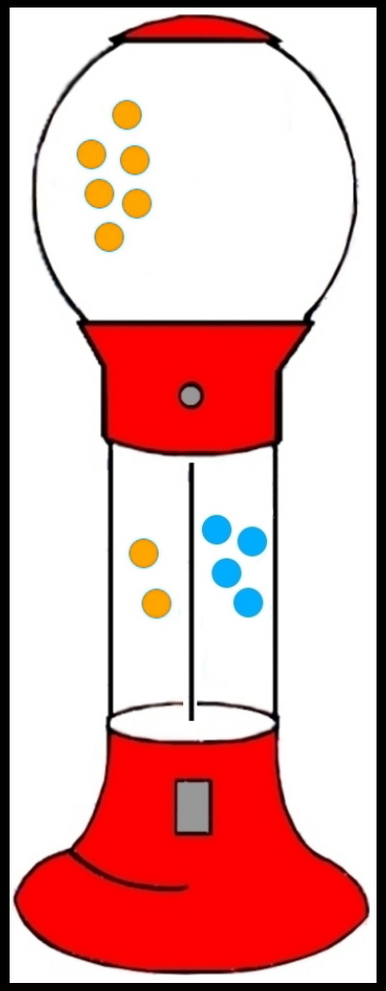
Pragmatic speaker
state = 4



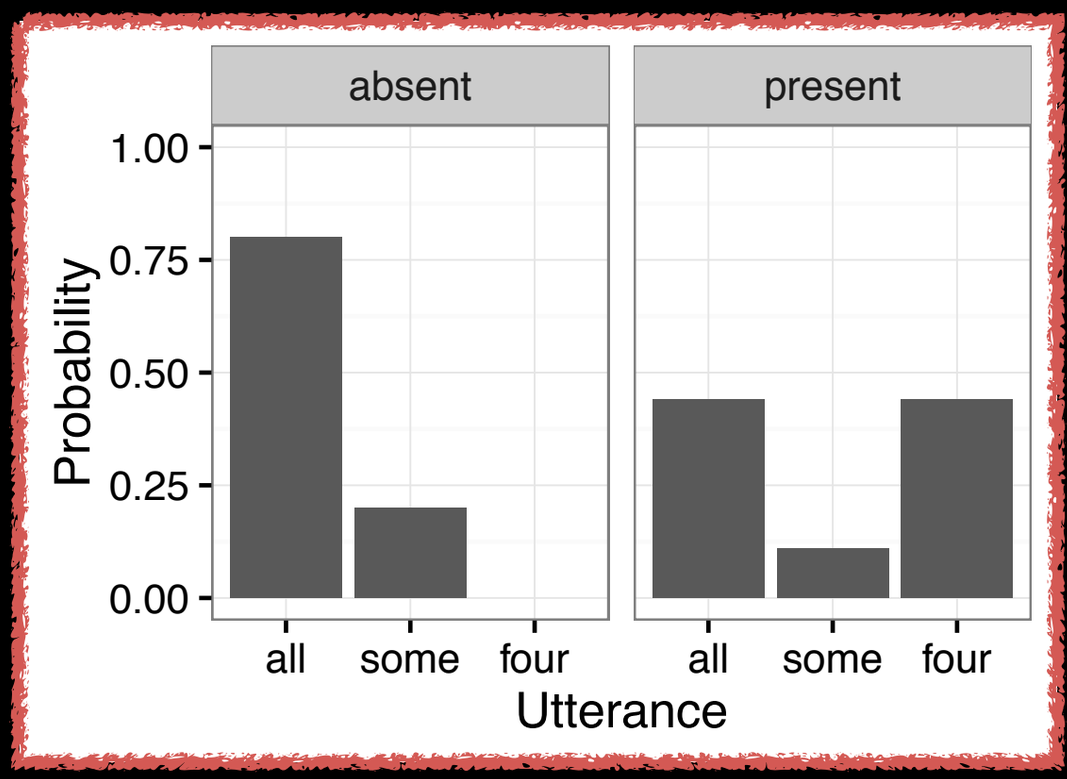
Context in RSA: alternatives

Absent: U = {none, some, all}

Present: U = {none, some, all, one, two, three, four}



Pragmatic speaker
state = 4

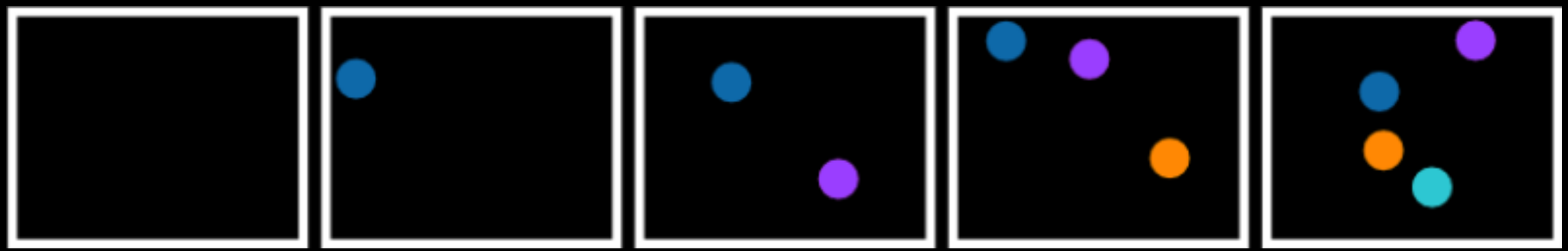


both “some” and “all” less expected when numbers are present

QUD effects on scalar implicature

Degen & Goodman 2014

Does the QUD modulate scalar implicature strength?



Implicit QUD

QUD effects on scalar implicature

Degen & Goodman 2014

Does the QUD modulate scalar implicature strength?

Implicit QUD

all? Did the speaker find all of the marbles?

QUD effects on scalar implicature

Degen & Goodman 2014

Does the QUD modulate scalar implicature strength?

Implicit QUD

all? Did the speaker find all of the marbles?

QUD effects on scalar implicature

Degen & Goodman 2014

Does the QUD modulate scalar implicature strength?

Implicit QUD

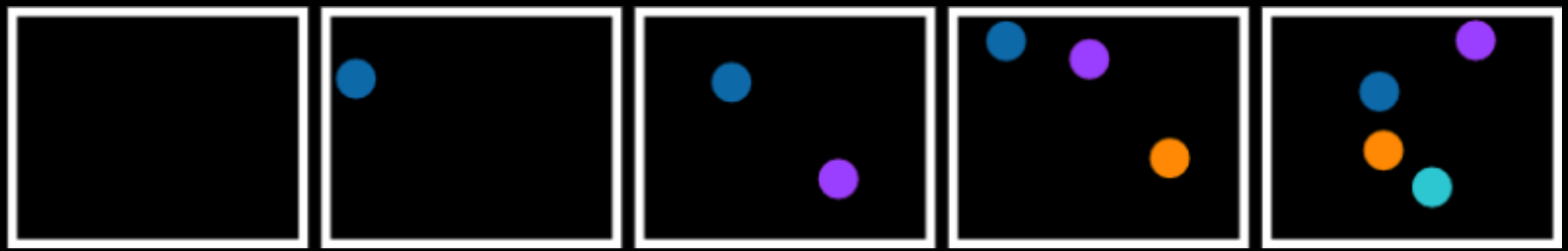
all? Did the speaker find all of the marbles?

I found all / some of the marbles.

QUD effects on scalar implicature

Degen & Goodman 2014

Does the QUD modulate scalar implicature strength?



Implicit QUD

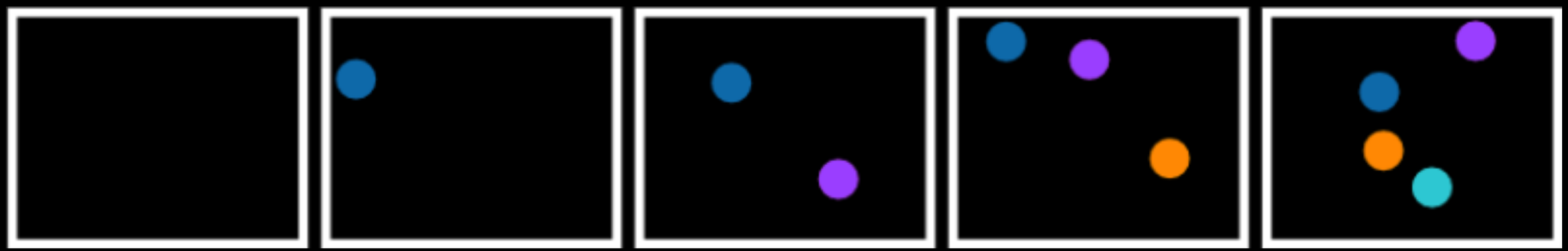
all? Did the speaker find all of the marbles?

I found all / some of the marbles.

QUD effects on scalar implicature

Degen & Goodman 2014

Does the QUD modulate scalar implicature strength?



Implicit QUD

all? Did the speaker find all of the marbles?

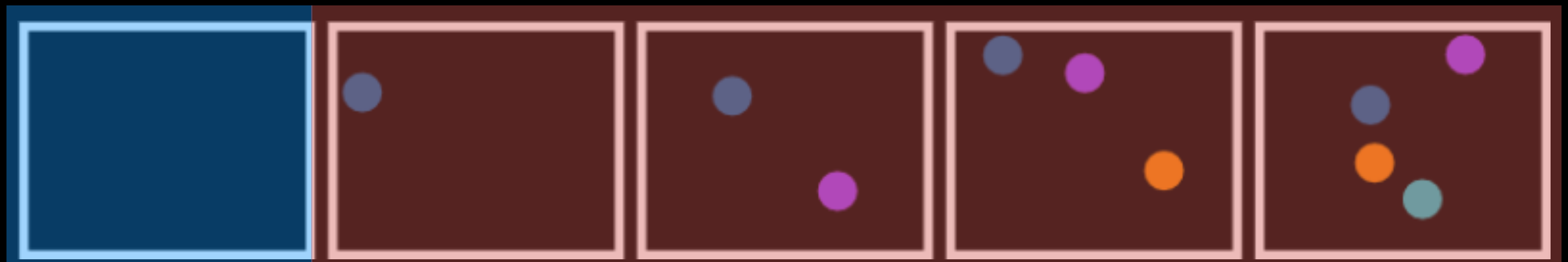
I found all / some of the marbles.

any? Did the speaker find any of the marbles?

QUD effects on scalar implicature

Degen & Goodman 2014

Does the QUD modulate scalar implicature strength?



Implicit QUD

all? Did the speaker find all of the marbles?

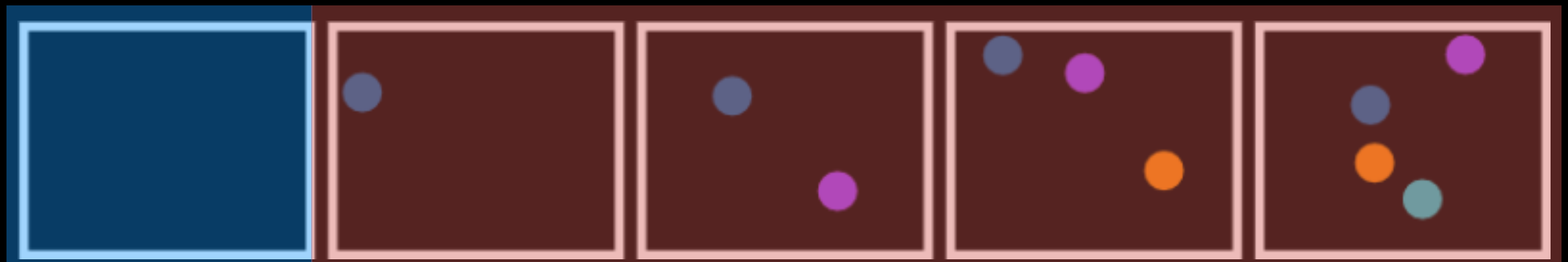
I found all / some of the marbles.

any? Did the speaker find any of the marbles?

QUD effects on scalar implicature

Degen & Goodman 2014

Does the QUD modulate scalar implicature strength?



Implicit QUD

all? Did the speaker find all of the marbles?

I found all / some of the marbles.

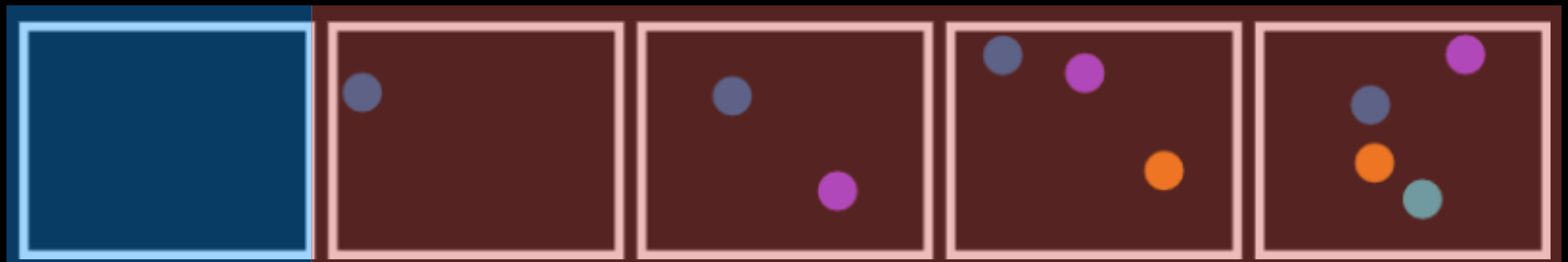
any? Did the speaker find any of the marbles?

I found all / some of the marbles.

QUD effects on scalar implicature

Degen & Goodman 2014

Does the QUD modulate scalar implicature strength?



Implicit QUD —> manipulated via cover stories

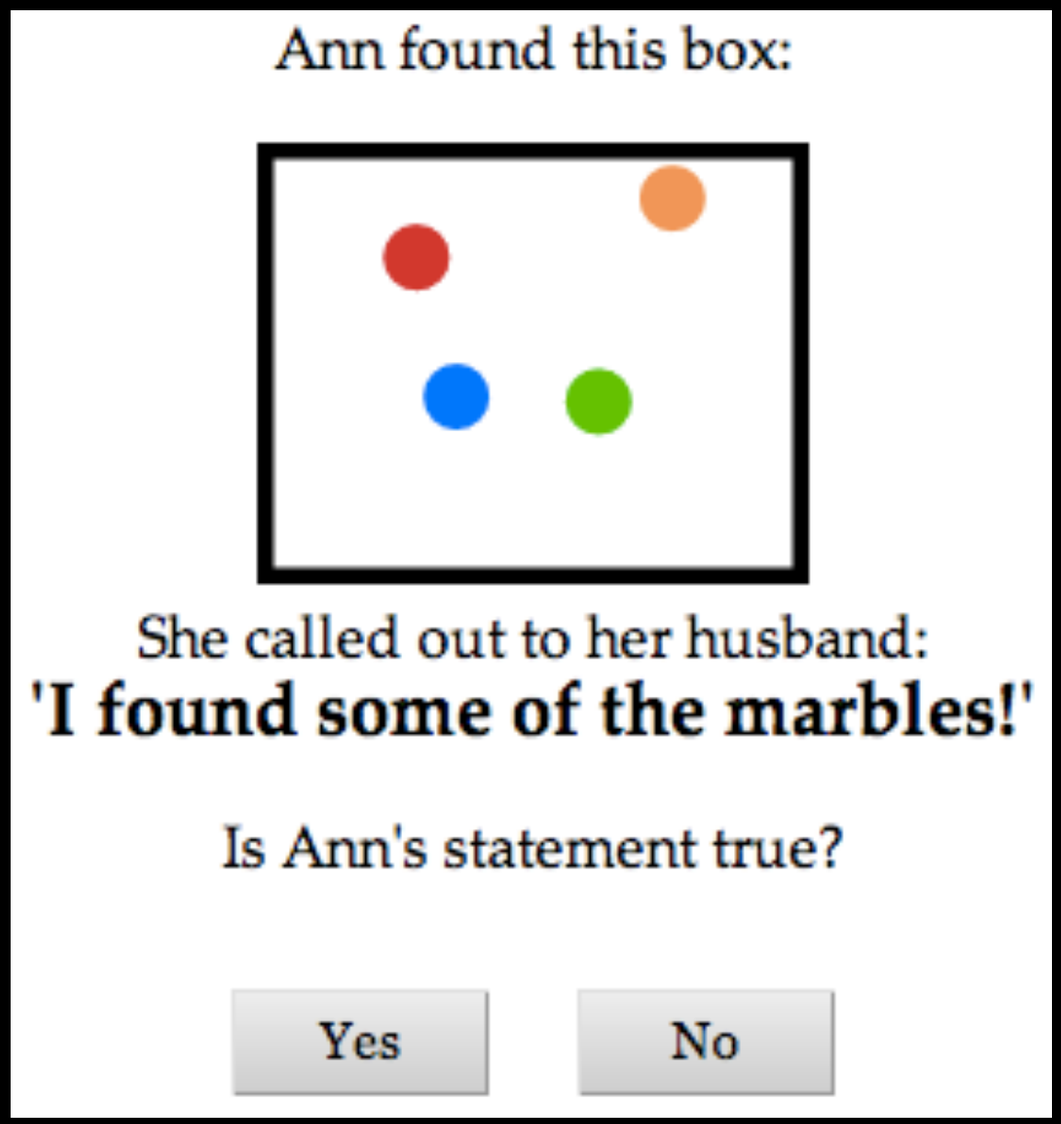
all? Did the speaker find all of the marbles?

I found all / some of the marbles.

any? Did the speaker find any of the marbles?

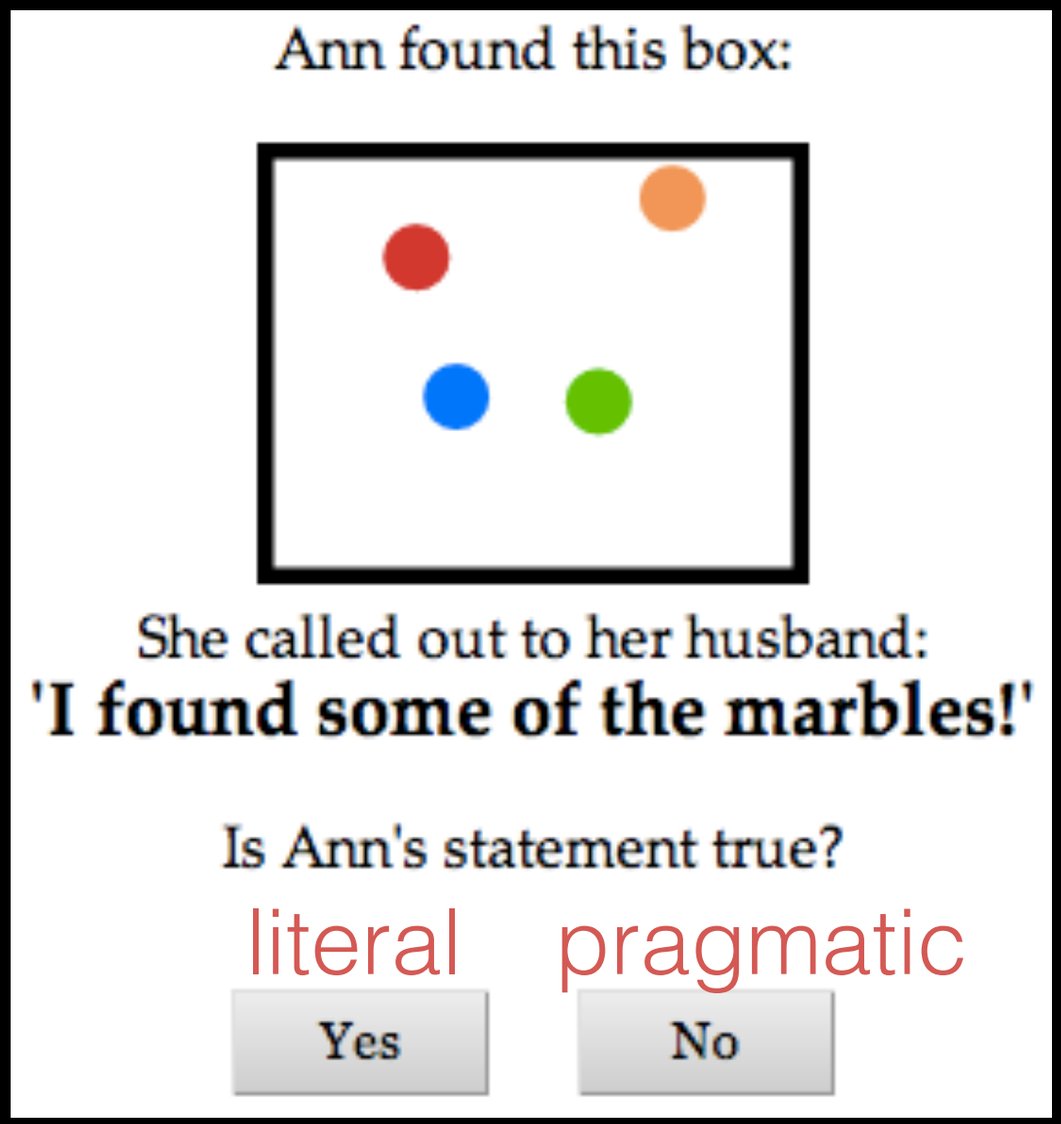
I found all / some of the marbles.

Task and results



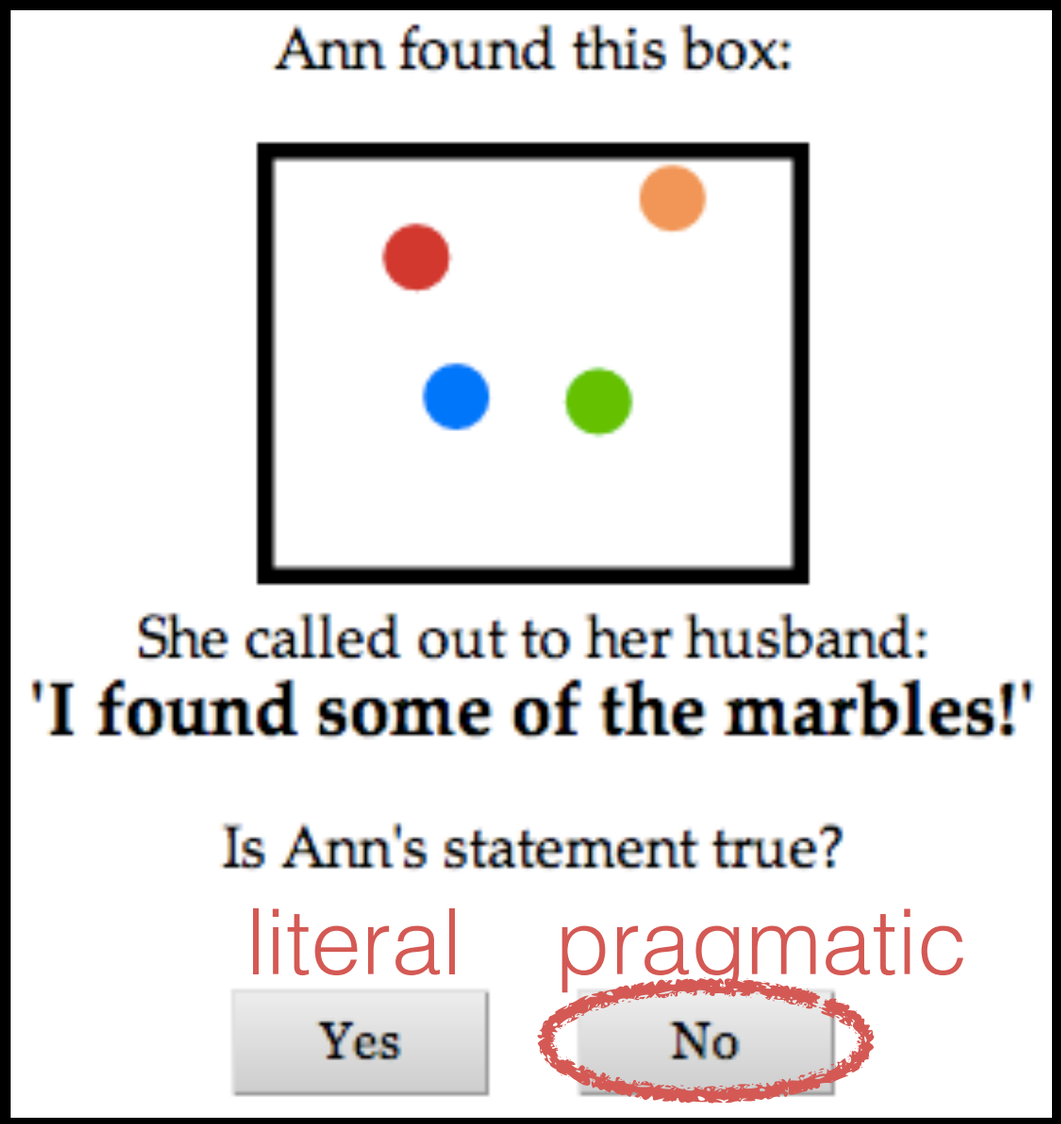
48 participants on Mechanical Turk

Task and results



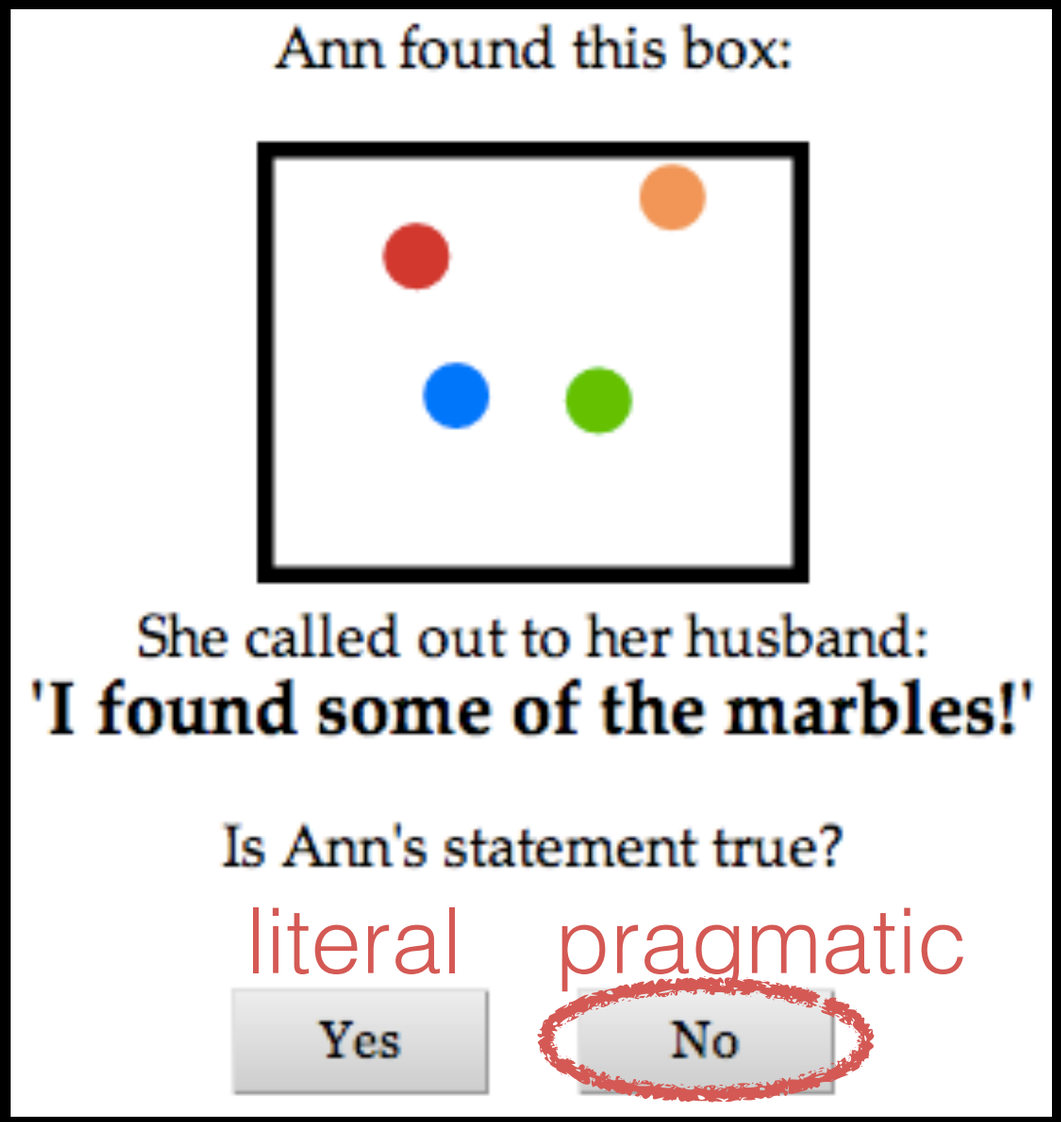
48 participants on Mechanical Turk

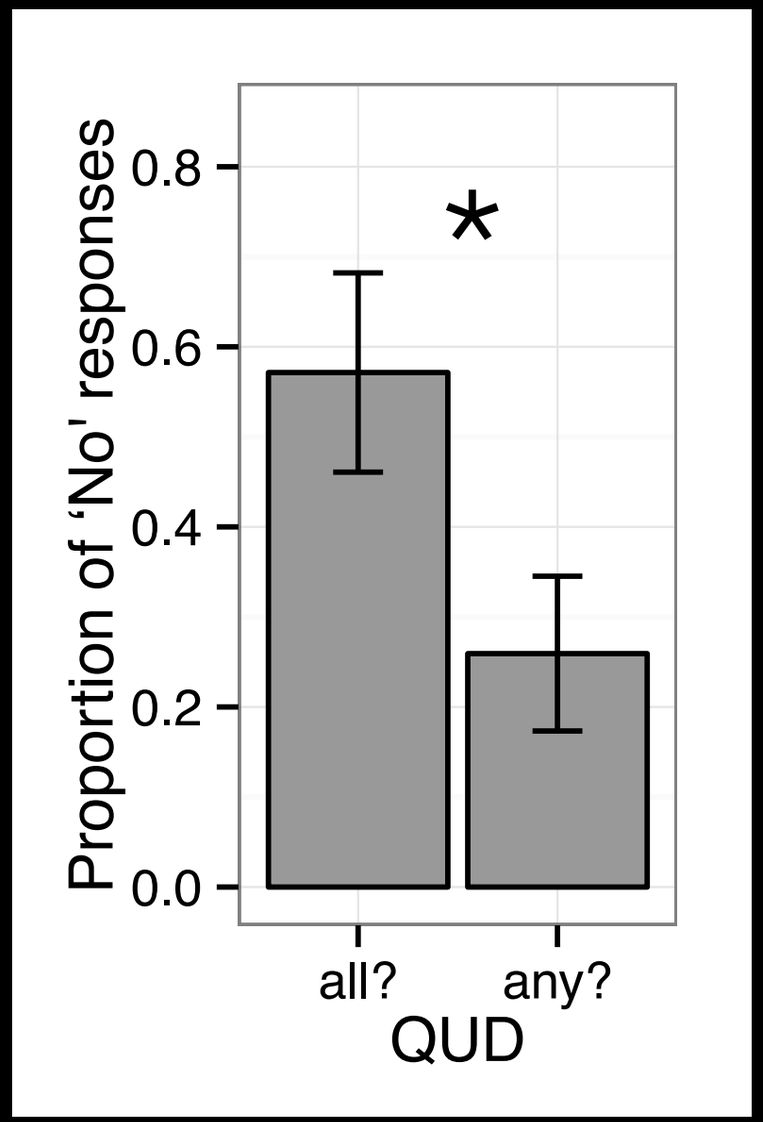
Task and results



48 participants on Mechanical Turk

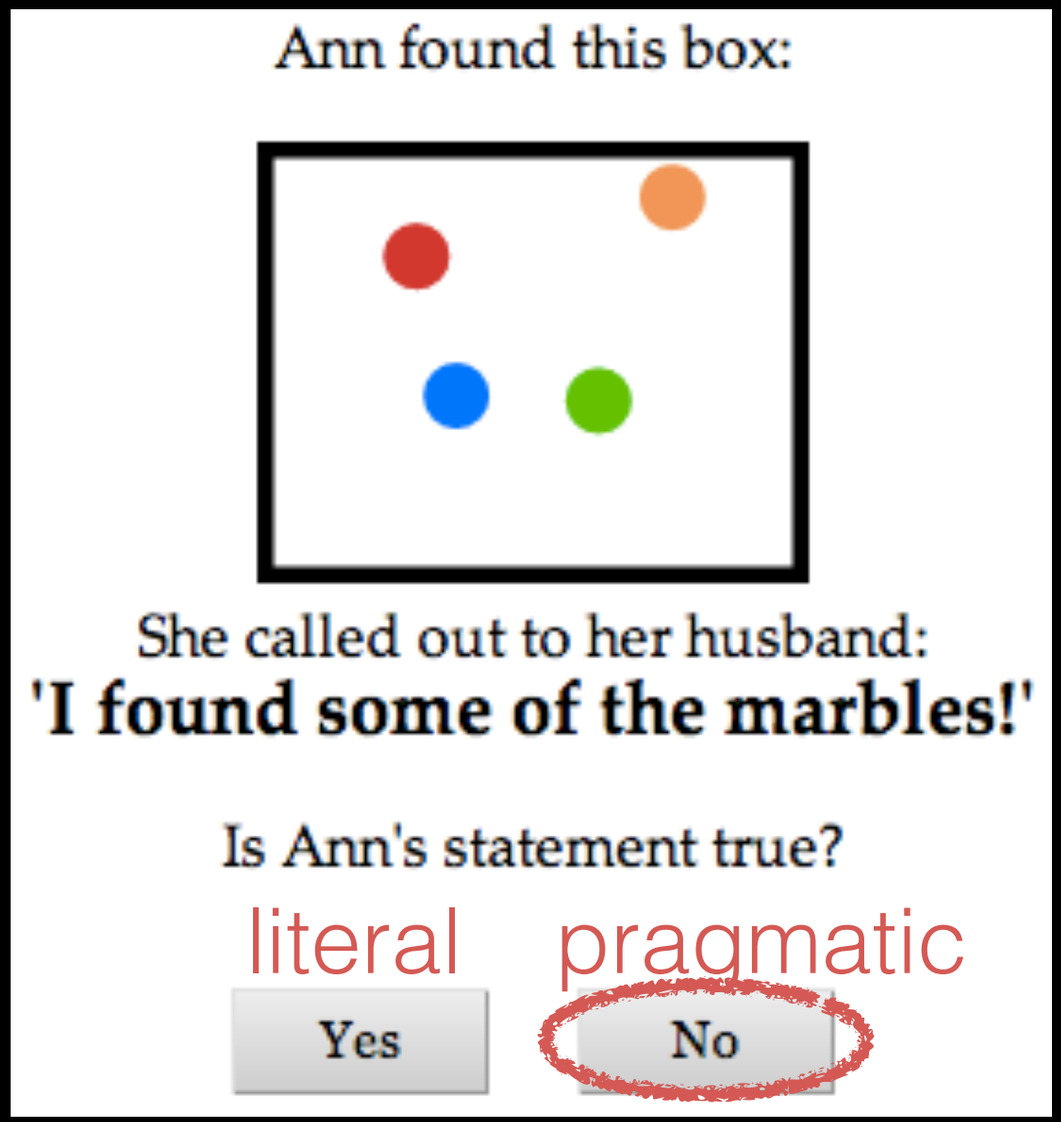
Task and results



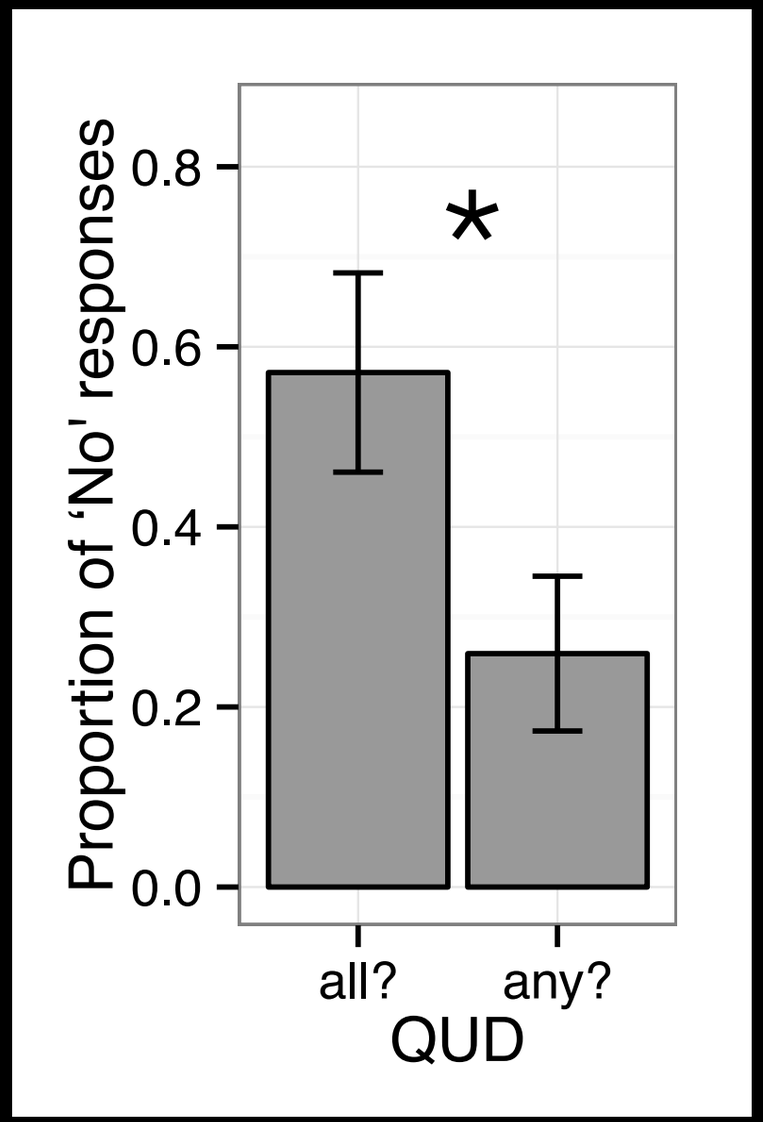


48 participants on Mechanical Turk

Task and results

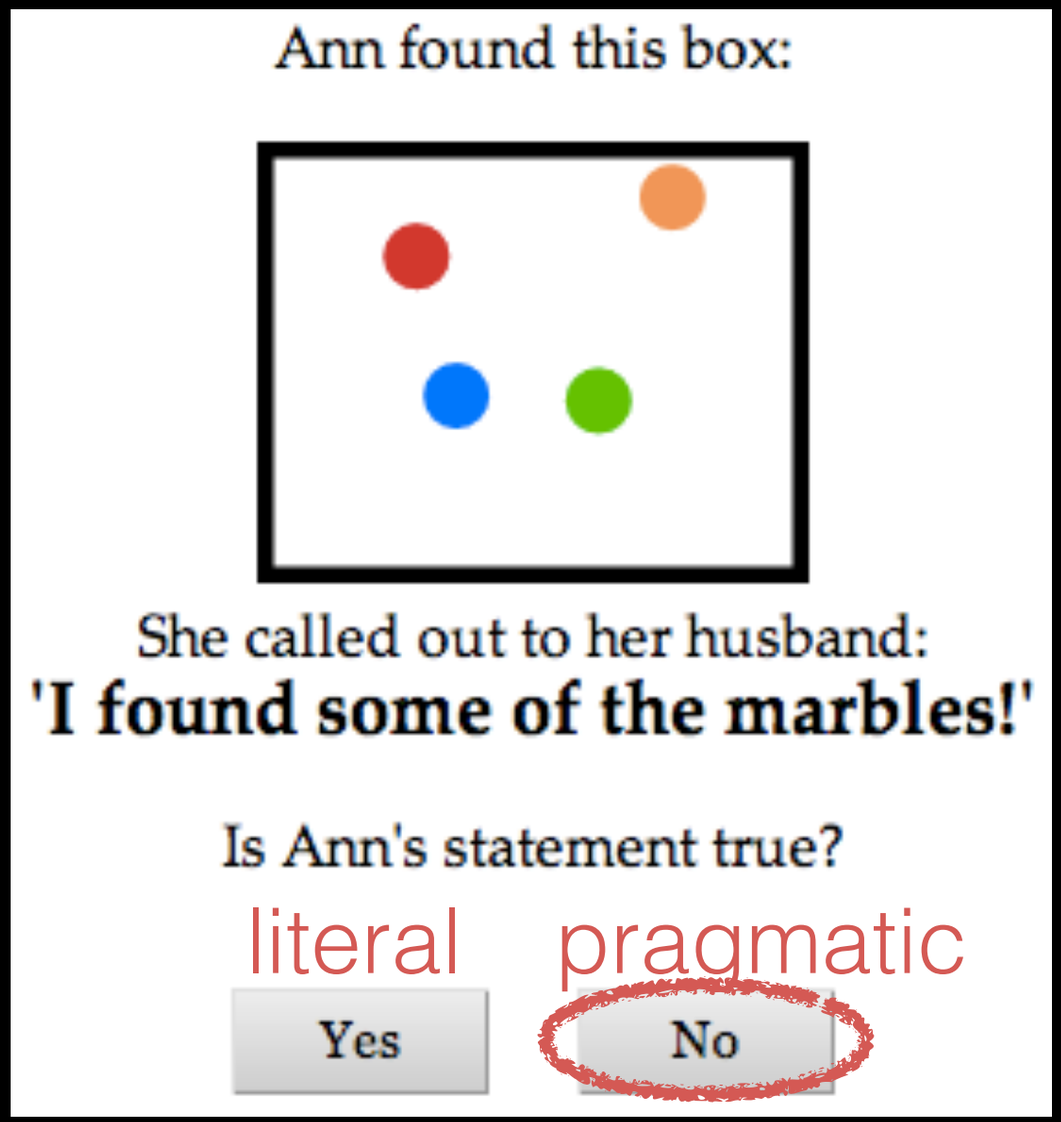


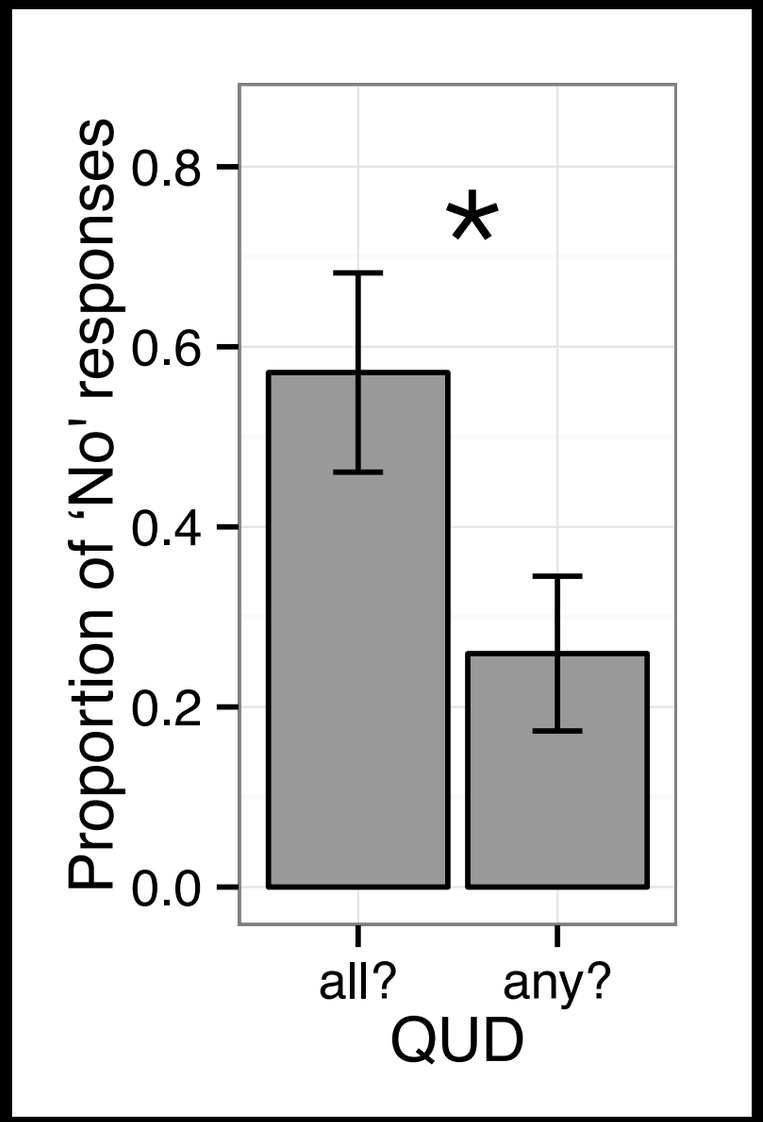
48 participants on Mechanical Turk



The QUD modulates scalar inference strength

Task and results





48 participants on Mechanical Turk

The QUD modulates scalar inference strength

see also Degen 2013

Context in RSA: QUD

Degen & Goodman 2014

$Q = \{q_{\text{all?}}, q_{\text{any?}}\}$

$P_{L_1}(m|u, q) \propto P_{S_1}(u|m, q) \cdot P(m)$

(assuming uniform prior on QUDs and independence of QUD and actual state of the world)

Context in RSA: QUD

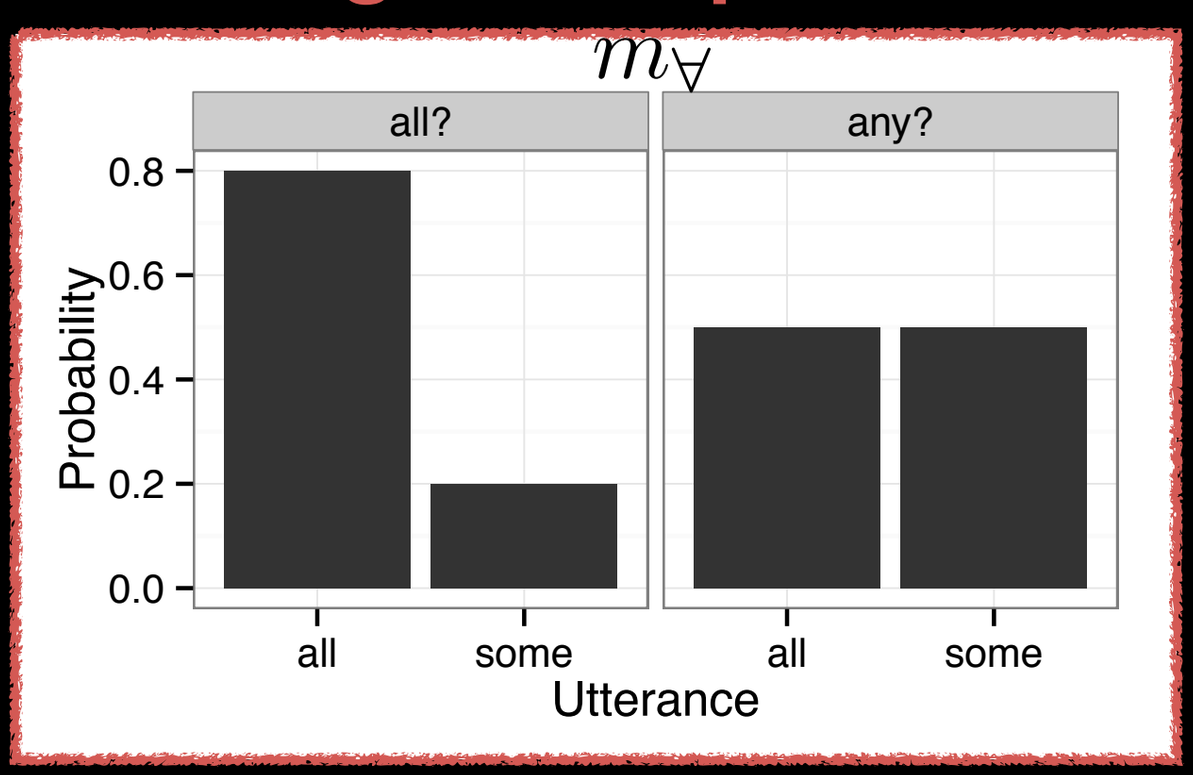
Degen & Goodman 2014

$Q = \{q_{\text{all?}}, q_{\text{any?}}\}$

$P_{L_1}(m|u, q) \propto P_{S_1}(u|m, q) \cdot P(m)$

(assuming uniform prior on QUDs and independence of QUD and actual state of the world)

Pragmatic speaker



Context in RSA: QUD

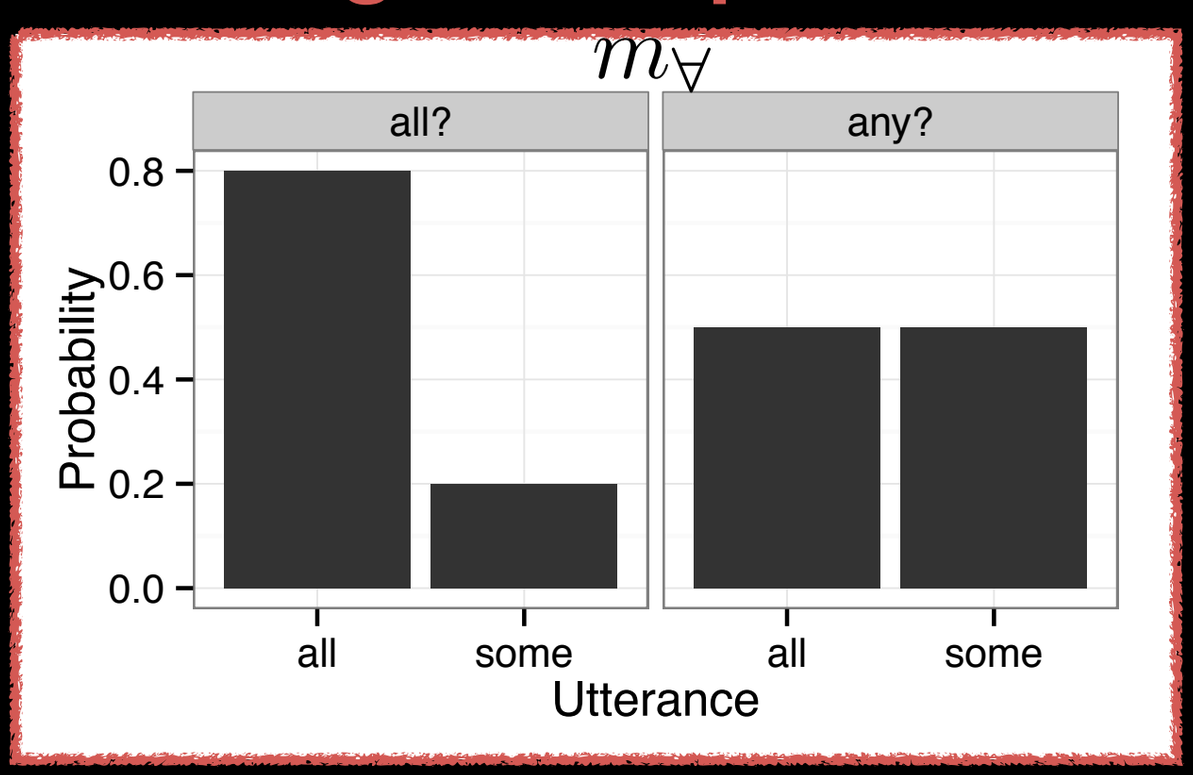
Degen & Goodman 2014

$$Q = \{q_{\text{all?}}, q_{\text{any?}}\}$$

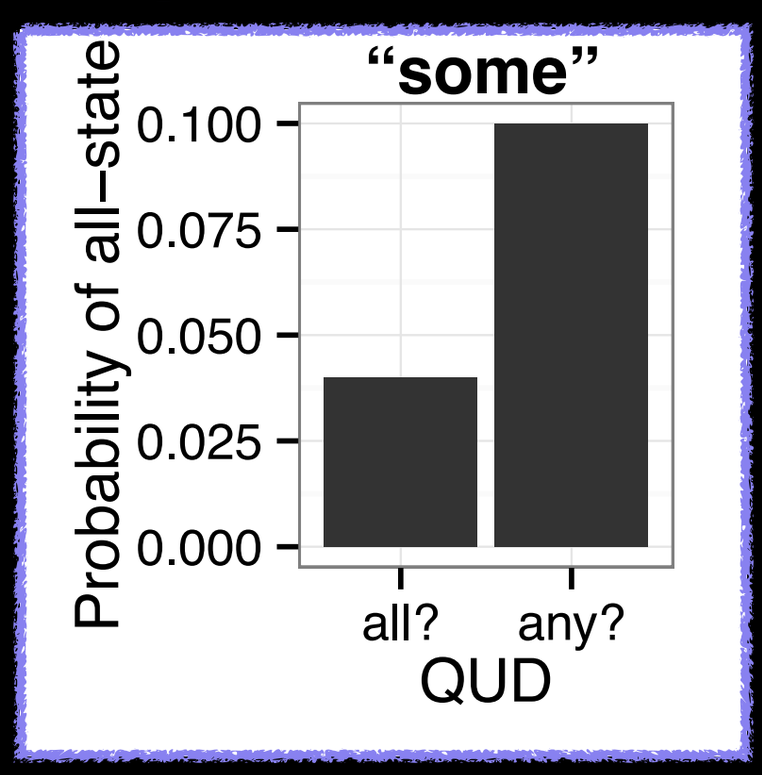
$$P_{L_1}(m|u, q) \propto P_{S_1}(u|m, q) \cdot P(m)$$

(assuming uniform prior on QUDs and independence of QUD and actual state of the world)

Pragmatic speaker



Pragmatic listener



Context in RSA: QUD

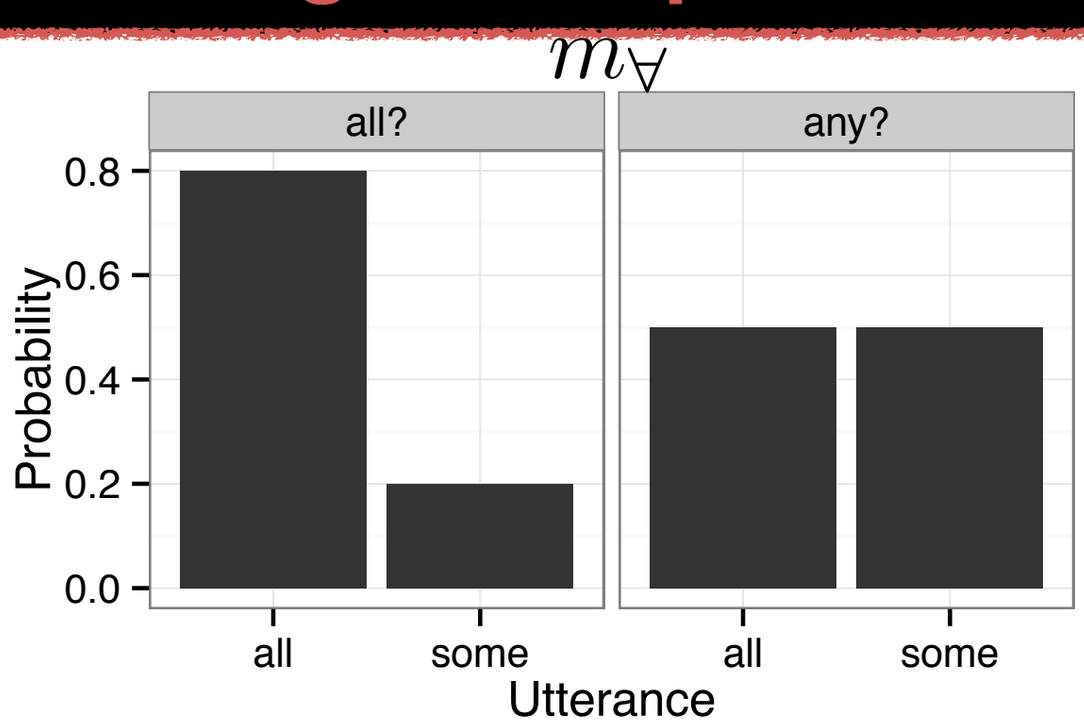
Degen & Goodman 2014

$$Q = \{q_{\text{all?}}, q_{\text{any?}}\}$$

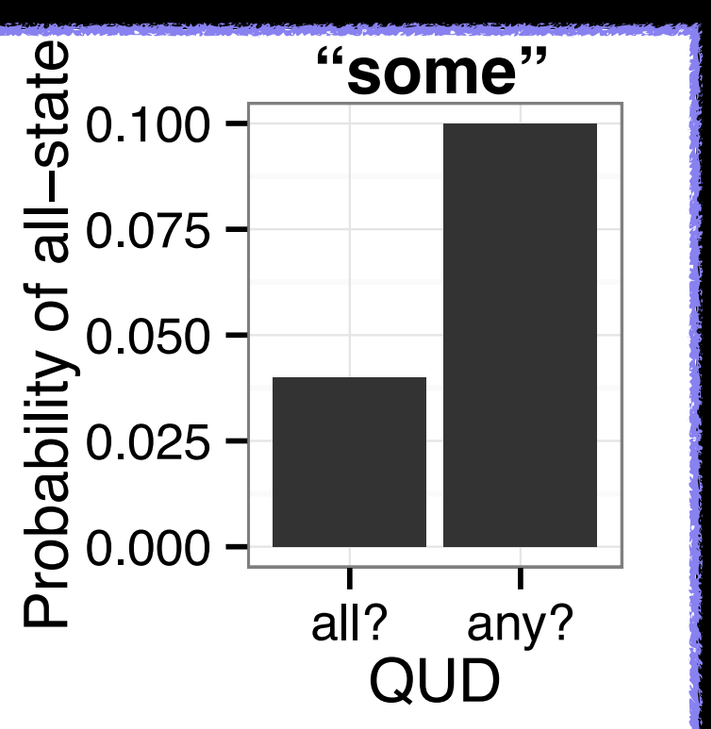
$$P_{L_1}(m|u, q) \propto P_{S_1}(u|m, q) \cdot P(m)$$

(assuming uniform prior on QUDs and independence of QUD and actual state of the world)

Pragmatic speaker $m_\forall$



Pragmatic listener



RSA captures QUD effects in scalar implicature

Summary

I. **Overinformativeness** in production

redundant referring expressions are rational when modifiers are noisy

II. **Underinformativeness** in production

listeners make efficient use of context in drawing scalar inferences

Summary

I. **Overinformativeness** in production

redundant referring expressions are rational when modifiers are noisy

II. **Underinformativeness** in production

listeners make efficient use of context in drawing scalar inferences

RSA as the most promising current framework

RSA?

Default theory
Levinson 2000

Relevance theory
Sperber & Wilson 1995;
Carston 1998

extreme informational privilege

extreme parallelism

Literal-first hypothesis
Huang & Snedeker 2009

Constraint-based account
Degen & Tanenhaus 2015

Levels of analysis

Marr 1982

RSA?

Default theory

Levinson 2000

Relevance theory

Sperber & Wilson 1995;
Carston 1998

extreme informational privilege

extreme parallelism

Literal-first hypothesis

Huang & Snedeker 2009

Constraint-based account

Degen & Tanenhaus 2015

Levels of analysis

Marr 1982

RSA?

Default theory
Levinson 2000

Relevance theory
Sperber & Wilson 1995;
Carston 1998

extreme informational privilege

extreme parallelism

Literal-first hypothesis
Huang & Snedeker 2009

Constraint-based account
Degen & Tanenhaus 2015

mechanistic

Levels of analysis

Marr 1982

RSA?

computational

Default theory

Levinson 2000

Relevance theory

Sperber & Wilson 1995;
Carston 1998

extreme informational privilege

extreme parallelism

Literal-first hypothesis

Huang & Snedeker 2009

Constraint-based account

Degen & Tanenhaus 2015

mechanistic

Challenges / limitations / future directions

- online processing — Tessler & Levy 2019; Cohn-Gordon et al 2019; Augurzky et al 2019
- resource limitations — Franke & Degen 2016
- beyond one-shot utterances — Hawkins et al 2015
- more interesting structure
- adaptation — Schuster & Degen 2019; in prep
- linking functions — Augurzky et al 2019; Qing et al 2018; Waldon & Degen in prep

http://www.problang.org

Probabilistic language understanding
An introduction to the Rational Speech Act framework

By Gregory Scontras, Michael Henry Tessler, and Michael Franke

The present course serves as a practical introduction to the Rational Speech Act modeling framework. Little is presupposed beyond a willingness to explore recent progress in formal, implementable models of language understanding.

Main content

The literal listener rule can be written as follows:

```
// set of states (here: objects of reference)
// we represent objects as JavaScript objects to demarcate them from utterances
// internally we treat objects as strings nonetheless
var objects = [{color: "blue", shape: "square", string: "blue square"},
               {color: "blue", shape: "circle", string: "blue circle"},
               {color: "green", shape: "square", string: "green square"}]

// set of utterances
var utterances = ["blue", "green", "square", "circle"]

// prior over world states
var objectPrior = function() {
  var obj = uniformDraw(objects)
  return obj.string
}

// meaning function to interpret the utterances
var meaning = function(utterance, obj){
  _.includes(obj, utterance)
}

// literal listener
var literalListener = function(utterance){
  Infer({model: function(){
    var obj = objectPrior();
    var uttTruthVal = meaning(utterance, obj);
    condition(uttTruthVal == true)
    return obj
  }})
}

viz.table(literalListener("blue"))
```

run

(state)	probability
blue circle	0.5
blue square	0.5

Exercises:

I. In the model above, `objectPrior()` returns a sample from a `uniformDraw` over the possible objects of reference. What happens when the listener's beliefs are not uniform over the

Scontras, Tessler, & Franke

http://www.problang.org

Probabilistic language understanding
An introduction to the Rational Speech Act framework

By Gregory Scontras, Michael Henry Tessler, and Michael Franke

The present course serves as a practical introduction to the Rational Speech Act modeling framework. Little is presupposed beyond a willingness to explore recent progress in formal, implementable models of language understanding.

Main content

The literal listener rule can be written as follows:

```
// set of states (here: objects of reference)
// we represent objects as JavaScript objects to demarcate them from utterances
// internally we treat objects as strings nonetheless
var objects = [{color: "blue", shape: "square", string: "blue square"},
               {color: "blue", shape: "circle", string: "blue circle"},
               {color: "green", shape: "square", string: "green square"}]

// set of utterances
var utterances = ["blue", "green", "square", "circle"]

// prior over world states
var objectPrior = function() {
  var obj = uniformDraw(objects)
  return obj.string
}

// meaning function to interpret the utterances
var meaning = function(utterance, obj){
  _.includes(obj, utterance)
}

// literal listener
var literalListener = function(utterance){
  Infer({model: function(){
    var obj = objectPrior();
    var uttTruthVal = meaning(utterance, obj);
    condition(uttTruthVal == true)
    return obj
  }})
}

viz.table(literalListener("blue"))
```

run

(state)	probability
blue circle	0.5
blue square	0.5

Exercises:

I. In the model above, `objectPrior()` returns a sample from a `uniformDraw` over the possible objects of reference. What happens when the listener's beliefs are not uniform over the

Scontras, Tessler, & Franke

XPrag.de summer school:

"Probabilistic pragmatics"

Thank you

Collaborators

Elisa Kreiss
Noah Goodman
Caroline Graf
Robert Hawkins
Mike Tanenhaus
Michael Franke
Greg Scontras
Michael Henry Tessler

Funding

- Swiss National Science Foundation Early Postdoc.Mobility grant to JD
- XPrag.de DFG priority program grant to JD & MF
- James S. McDonnell grant to NG
- ONR grant N00014-13-1-0788 to NG
- NIH grant HD27206 to MT