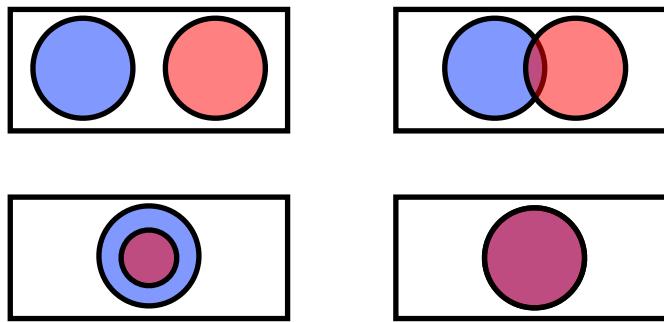


Introductory Bayesian pragmatics

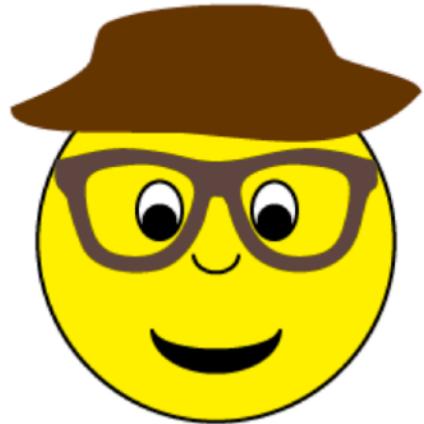


9.19: Computational Psycholinguistics

1 December 2021

Roger Levy

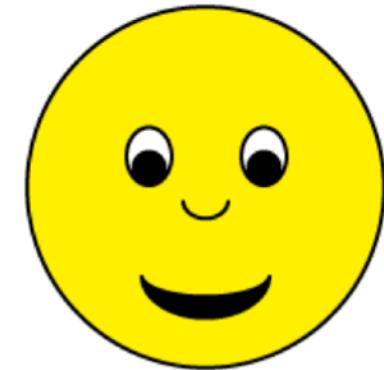
Ad-hoc scalar inference



A



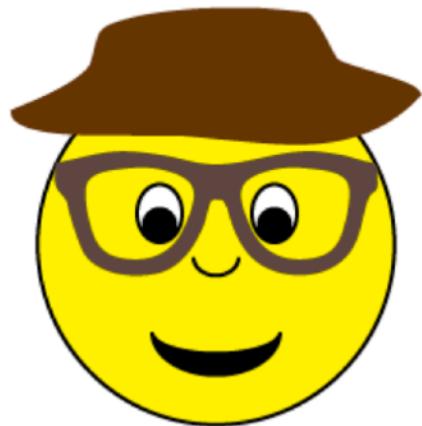
B



C

Bob can only say one word to communicate with you and he says: "glasses"

Ad-hoc scalar inference



A



B

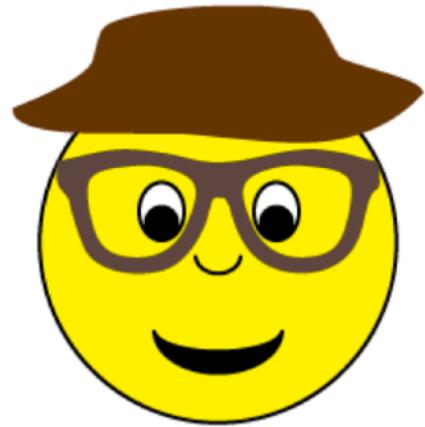


C

Bob can only say one word to communicate with you and he says: "glasses"

Empirical finding: >75% of experimental participants choose character B!

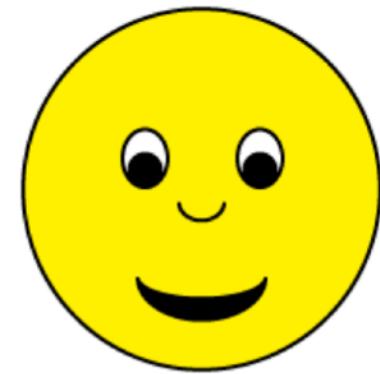
What is *said* and what is *meant*



A



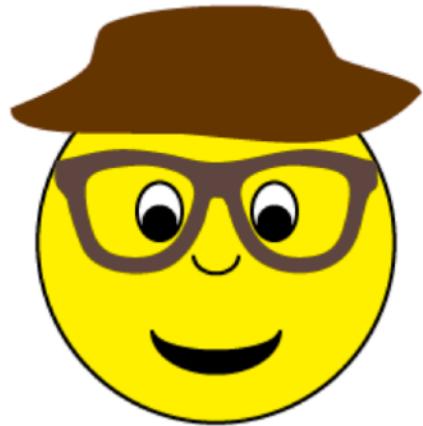
B



C

What is *said* and what is *meant*

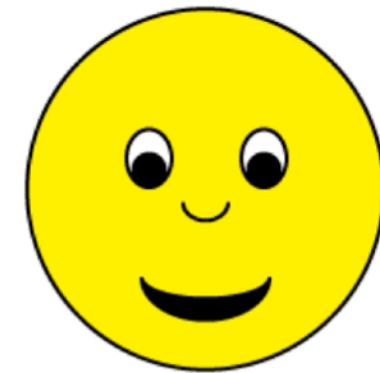
"glasses"



A



B

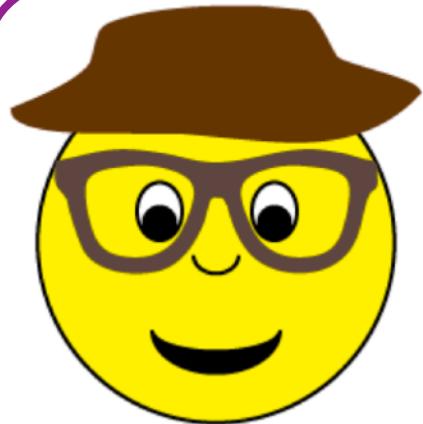


C

What is said and what is meant

"glasses"

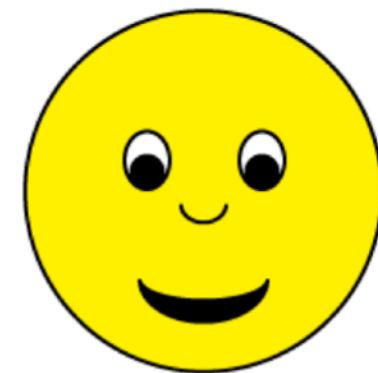
Literally compatible



A



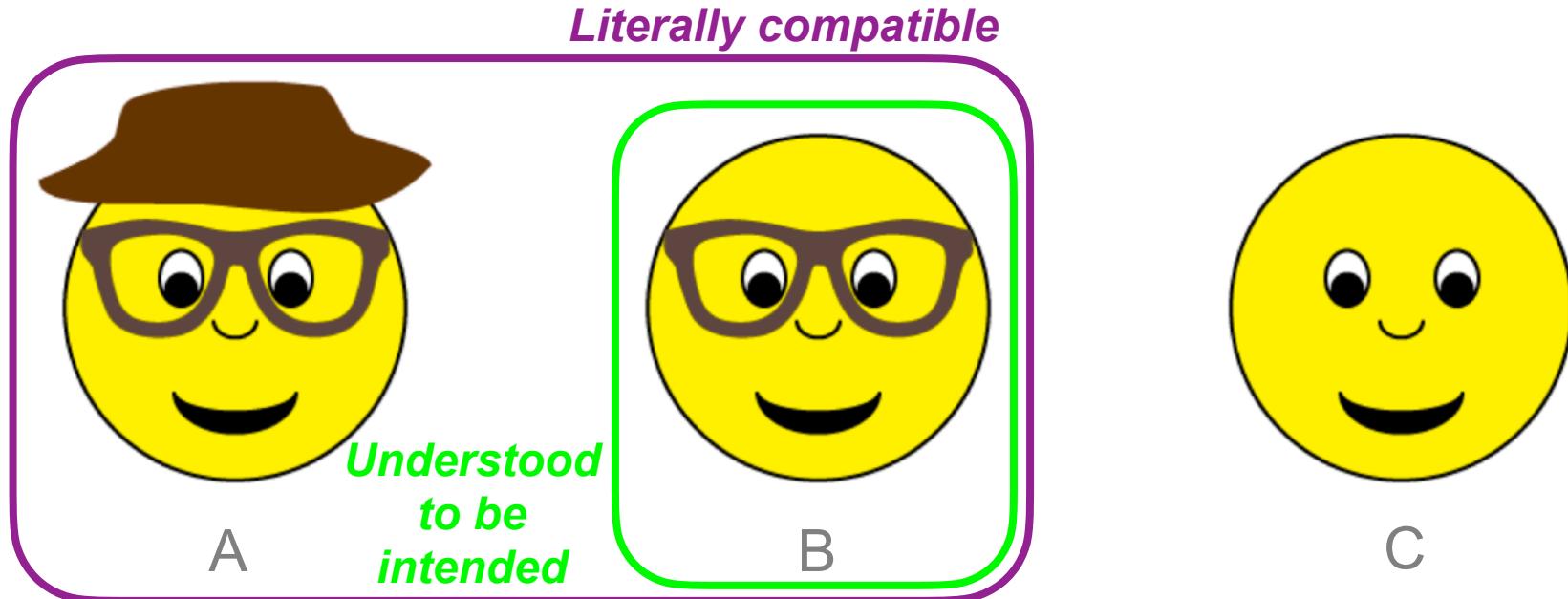
B



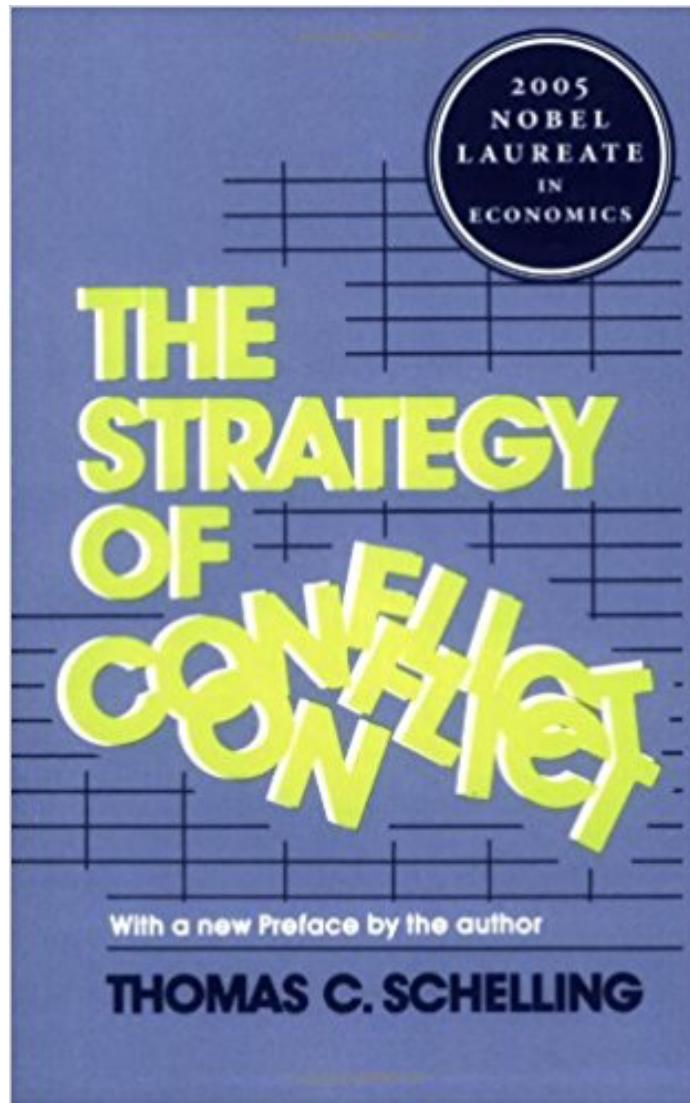
C

What is said and what is meant

"glasses"



Coordination games



Formalizing theories of semantics & pragmatics

Formalizing theories of semantics & pragmatics

- How does human language achieve its unbounded and highly context-dependent expressive capacity?

Formalizing theories of semantics & pragmatics

- How does human language achieve its unbounded and highly context-dependent expressive capacity?
- **Semantics:** the “literal” meanings of words and the rules of composition by which words are combined

Formalizing theories of semantics & pragmatics

- How does human language achieve its unbounded and highly context-dependent expressive capacity?
- **Semantics:** the “literal” meanings of words and the rules of composition by which words are combined
- **Pragmatics:** how a speaker’s communicative intent is inferred from literal meaning in context

Formalizing theories of semantics & pragmatics

- How does human language achieve its unbounded and highly context-dependent expressive capacity?
- **Semantics:** the “literal” meanings of words and the rules of composition by which words are combined
- **Pragmatics:** how a speaker’s communicative intent is inferred from literal meaning in context
 - A. *I could really use a cup of coffee.*

Formalizing theories of semantics & pragmatics

- How does human language achieve its unbounded and highly context-dependent expressive capacity?
- **Semantics:** the “literal” meanings of words and the rules of composition by which words are combined
- **Pragmatics:** how a speaker’s communicative intent is inferred from literal meaning in context
 - A. *I could really use a cup of coffee.*
 - B. *There’s a good place called Area Four nearby.*

Formalizing theories of semantics & pragmatics

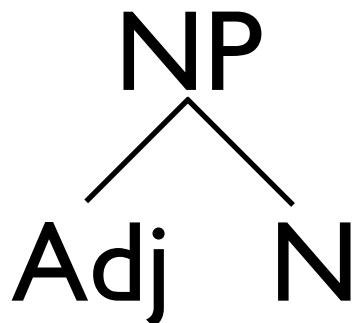
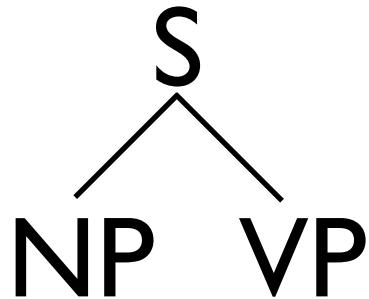
- How does human language achieve its unbounded and highly context-dependent expressive capacity?
- **Semantics:** the “literal” meanings of words and the rules of composition by which words are combined
- **Pragmatics:** how a speaker’s communicative intent is inferred from literal meaning in context
 - A. *I could really use a cup of coffee.*
 - B. *There’s a good place called Area Four nearby.*
- Probabilistic models over rich logical structures finally allow us to formalize joint semantic/pragmatic models

Formalizing theories of semantics & pragmatics

- How does human language achieve its unbounded and highly context-dependent expressive capacity?
- **Semantics:** the “literal” meanings of words and the rules of composition by which words are combined
- **Pragmatics:** how a speaker’s communicative intent is inferred from literal meaning in context
 - A. *I could really use a cup of coffee.*
 - B. *There’s a good place called Area Four nearby.*
- Probabilistic models over rich logical structures finally allow us to formalize joint semantic/pragmatic models
- Allows us to connect insights about linguistic meaning from across cognitive science—linguistics, AI, cognitive psychology, social cognition, philosophy

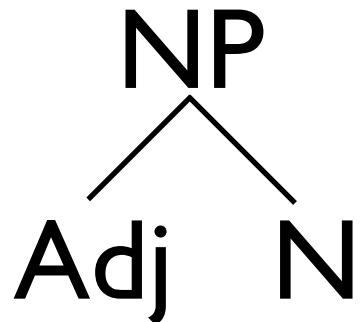
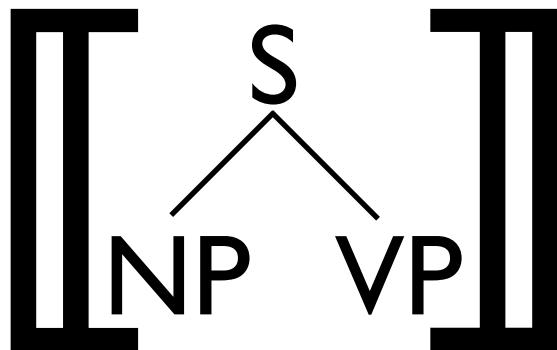
Semantics: principle of **compositionality**

The meaning of a complex expression is determined by the rules by which the expression is formed as applied to the meaning of the expression's subparts.



Semantics: principle of compositionality

The meaning of a complex expression is determined by the rules by which the expression is formed as applied to the meaning of the expression's subparts.

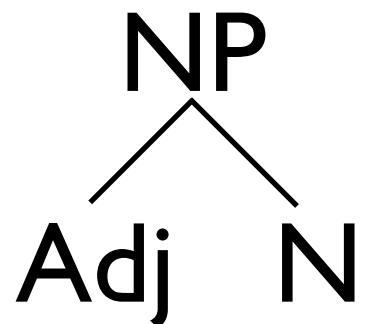


Semantics: principle of compositionality

The meaning of a complex expression is determined by the rules by which the expression is formed as applied to the meaning of the expression's subparts.

$$\boxed{\boxed{S}} = \llbracket VP \rrbracket (\llbracket NP \rrbracket)$$

Diagram: A tree diagram where the root node is labeled 'S'. It branches into two nodes labeled 'NP' and 'VP'. To the left of the tree is a large pair of vertical brackets, and to the right is an equals sign followed by another pair of brackets containing the words 'VP' and 'NP' separated by a parenthesis.



Semantics: principle of compositionality

The meaning of a complex expression is determined by the rules by which the expression is formed as applied to the meaning of the expression's subparts.

$$\boxed{\begin{array}{c} S \\ \diagdown \quad \diagup \\ NP \quad VP \end{array}} = [[VP]([NP])]$$

$$\boxed{\begin{array}{c} NP \\ \diagdown \quad \diagup \\ Adj \quad N \end{array}}$$

Semantics: principle of compositionality

The meaning of a complex expression is determined by the rules by which the expression is formed as applied to the meaning of the expression's subparts.

$$\boxed{\begin{array}{c} S \\ \diagdown \quad \diagup \\ NP \quad VP \end{array}} = [[VP]]([[NP]])$$

$$\boxed{\begin{array}{c} NP \\ \diagdown \quad \diagup \\ Adj \quad N \end{array}} = [[Adj]]([[N]])$$

...

Pragmatics: Grice, 1975

Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged. One might label this the COOPERATIVE PRINCIPLE.

A simple communication game

A simple communication game

- The *speaker* knows which of two states {A,B} holds of the world

A simple communication game

- The ***speaker*** knows which of two states $\{A, B\}$ holds of the world
- She can transmit one of two messages $\{x, y\}$ to the ***listener*** to signal which world state holds

A simple communication game

- The *speaker* knows which of two states $\{A, B\}$ holds of the world
- She can transmit one of two messages $\{x, y\}$ to the *listener* to signal which world state holds
- Speaker and listener have as common ground:

A simple communication game

- The *speaker* knows which of two states $\{A, B\}$ holds of the world
- She can transmit one of two messages $\{x, y\}$ to the *listener* to signal which world state holds
- Speaker and listener have as common ground:
 - A *prior distribution* on world state $P(A), P(B)$

A simple communication game

- The ***speaker*** knows which of two states $\{A, B\}$ holds of the world
- She can transmit one of two messages $\{x, y\}$ to the ***listener*** to signal which world state holds
- Speaker and listener have as common ground:
 - A ***prior distribution*** on world state $P(A), P(B)$
 - Knowledge that messages x and y are equal in cost

A simple communication game

- The ***speaker*** knows which of two states $\{A, B\}$ holds of the world
- She can transmit one of two messages $\{x, y\}$ to the ***listener*** to signal which world state holds
- Speaker and listener have as common ground:
 - A ***prior distribution*** on world state $P(A), P(B)$
 - Knowledge that messages x and y are equal in cost
 - That the game is purely cooperative

A simple communication game

A simple communication game

- A *Pareto optimal strategy* is one that delivers the highest possible reward to all players

A simple communication game

- A *Pareto optimal strategy* is one that delivers the highest possible reward to all players
 - A speaker who knows **A**

A simple communication game

- A *Pareto optimal strategy* is one that delivers the highest possible reward to all players
 - A speaker who knows **A**
 - A speaker who knows **B**

A simple communication game

- A *Pareto optimal strategy* is one that delivers the highest possible reward to all players
 - A speaker who knows **A**
 - A speaker who knows **B**
 - A listener who hears **x**

A simple communication game

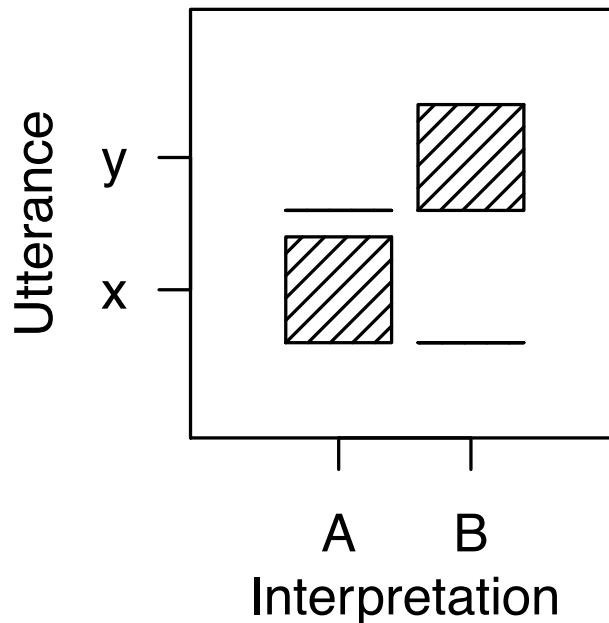
- A *Pareto optimal strategy* is one that delivers the highest possible reward to all players
 - A speaker who knows **A**
 - A speaker who knows **B**
 - A listener who hears **x**
 - A listener who hears **y**

A simple communication game

- A *Pareto optimal strategy* is one that delivers the highest possible reward to all players
 - A speaker who knows **A**
 - A speaker who knows **B**
 - A listener who hears **x**
 - A listener who hears **y**
- There are two Pareto optimal strategies in this game:

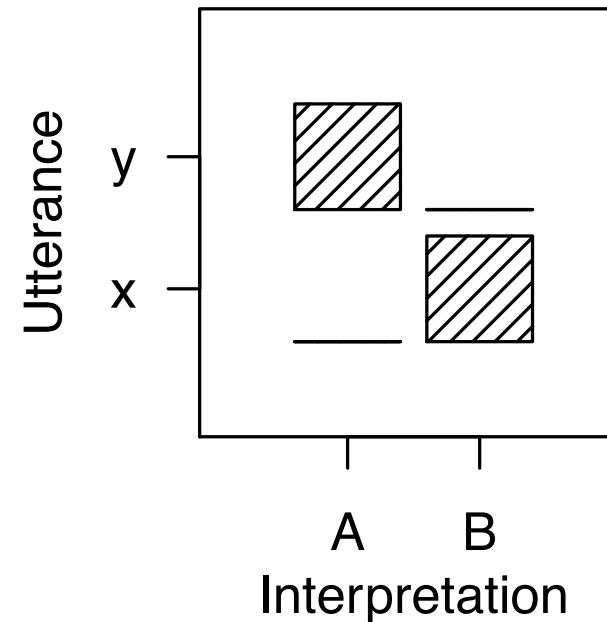
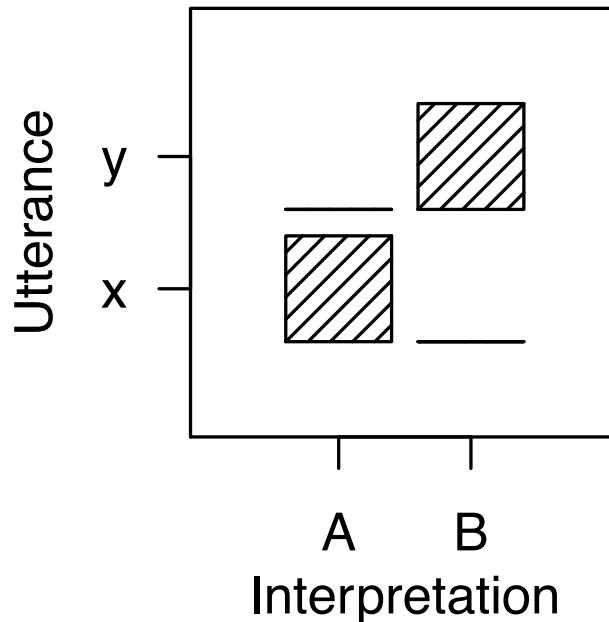
A simple communication game

- A *Pareto optimal strategy* is one that delivers the highest possible reward to all players
 - A speaker who knows **A**
 - A speaker who knows **B**
 - A listener who hears **x**
 - A listener who hears **y**
- There are two Pareto optimal strategies in this game:



A simple communication game

- A *Pareto optimal strategy* is one that delivers the highest possible reward to all players
 - A speaker who knows **A**
 - A speaker who knows **B**
 - A listener who hears **x**
 - A listener who hears **y**
- There are two Pareto optimal strategies in this game:



Efficiency limits of literal meaning

Efficiency limits of literal meaning

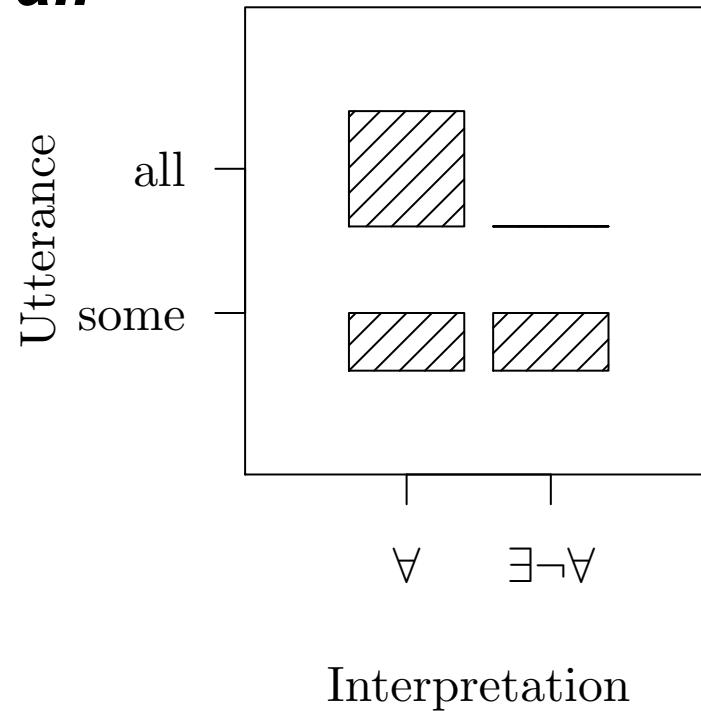
- But literal meanings don't hand us Pareto optimality

Efficiency limits of literal meaning

- But literal meanings don't hand us Pareto optimality
- A simple example: **some** and **all**

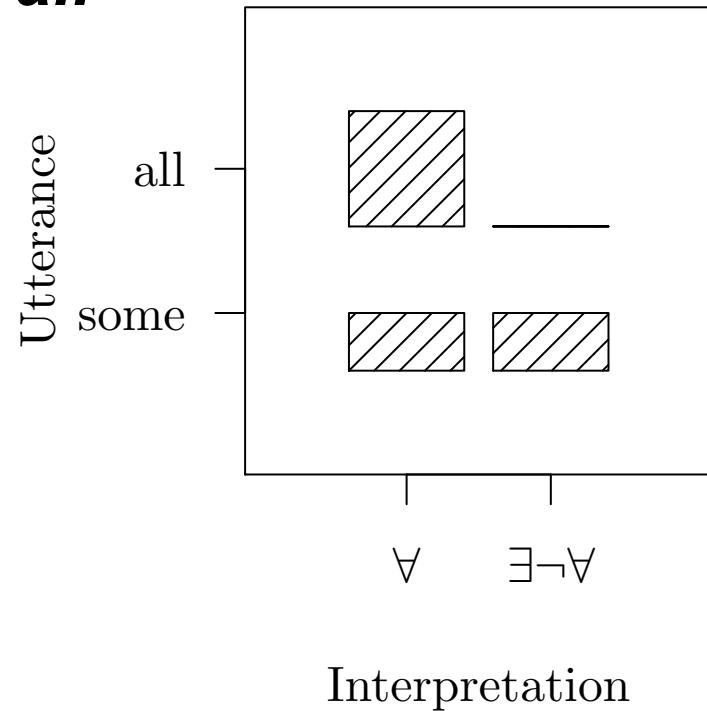
Efficiency limits of literal meaning

- But literal meanings don't hand us Pareto optimality
- A simple example: **some** and **all**



Efficiency limits of literal meaning

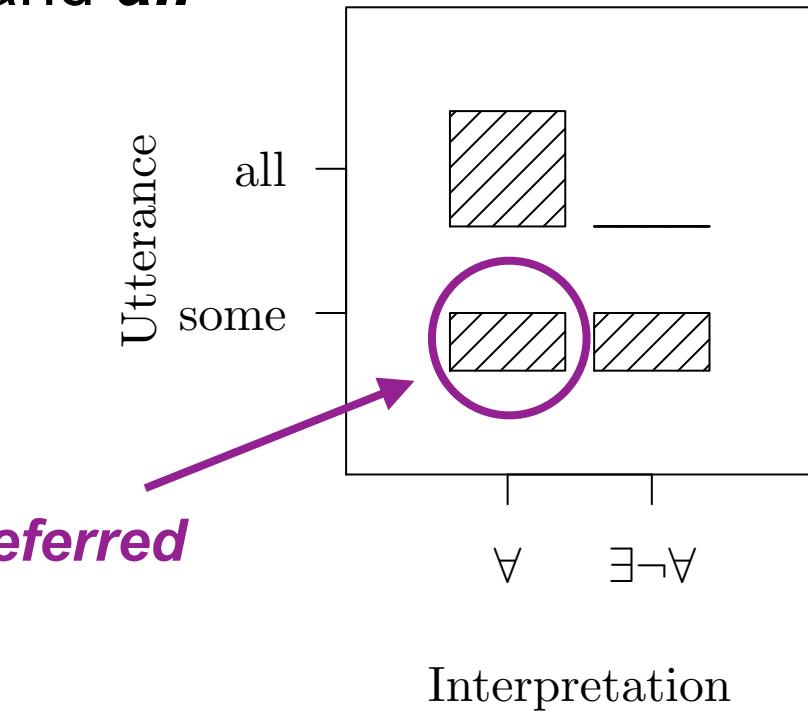
- But literal meanings don't hand us Pareto optimality
- A simple example: **some** and **all**



- “Remarkable” fact: the interpretation of **some** that is responsible for suboptimality is dispreferred!

Efficiency limits of literal meaning

- But literal meanings don't hand us Pareto optimality
- A simple example: **some** and **all**



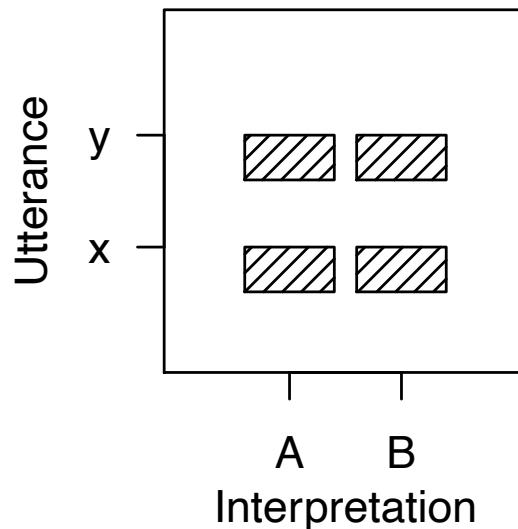
- “Remarkable” fact: the interpretation of **some** that is responsible for suboptimality is dispreferred!

A simple communication game

- There are also many *non-Pareto-optimal* strategies:

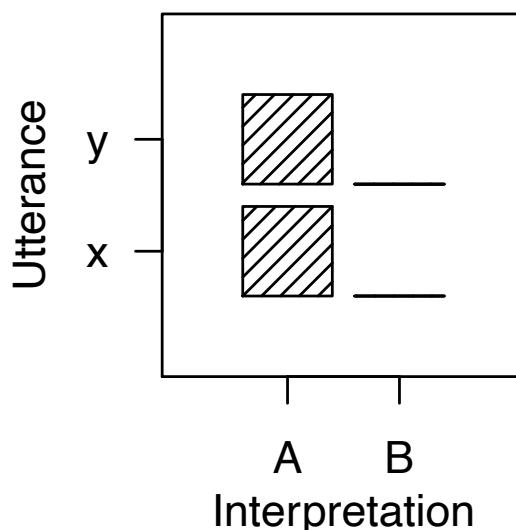
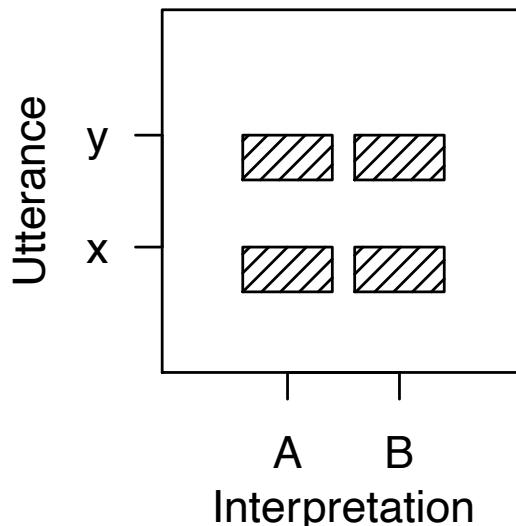
A simple communication game

- There are also many *non-Pareto-optimal* strategies:



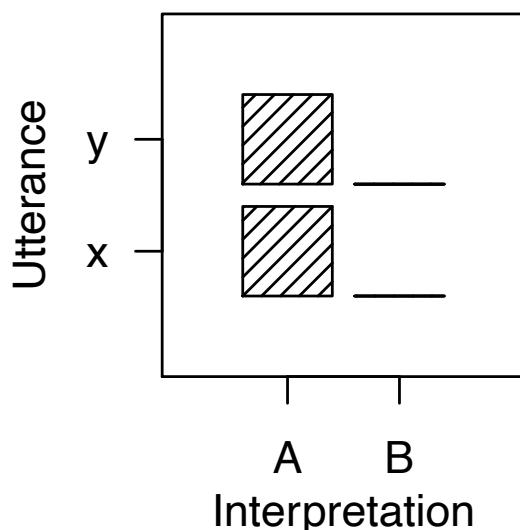
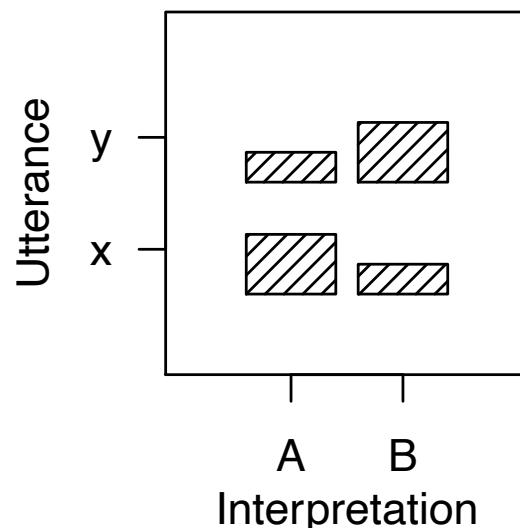
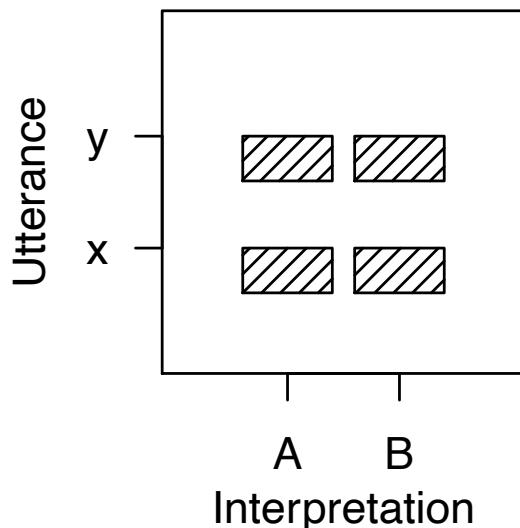
A simple communication game

- There are also many *non-Pareto-optimal* strategies:



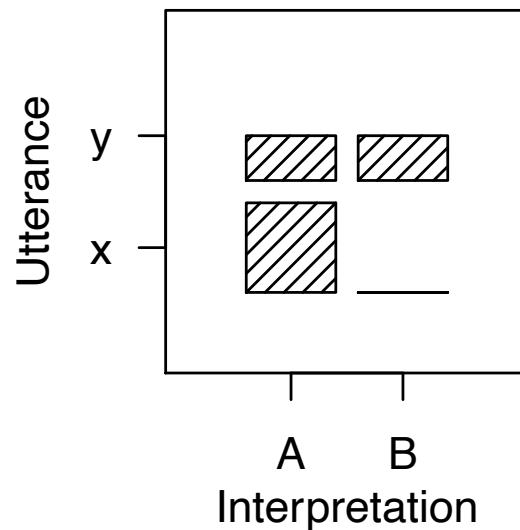
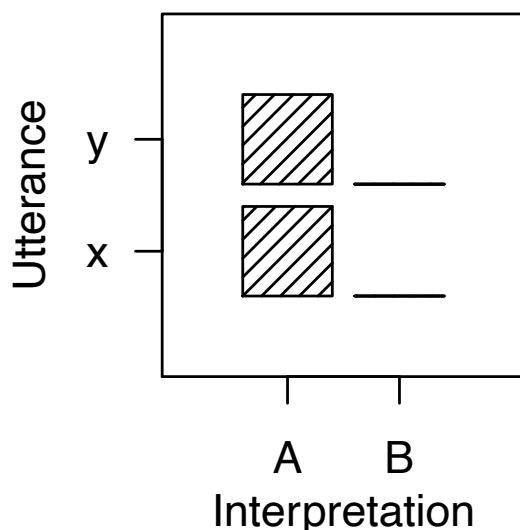
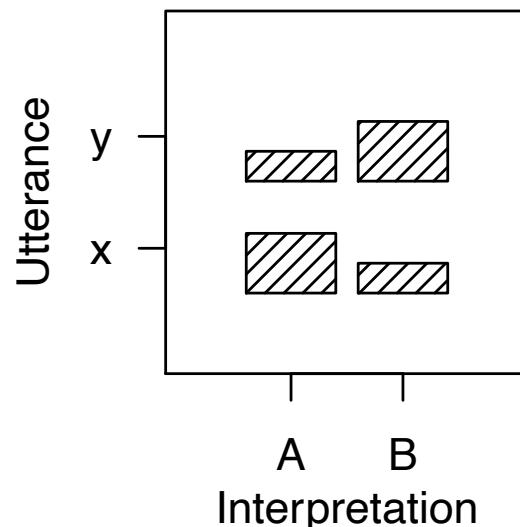
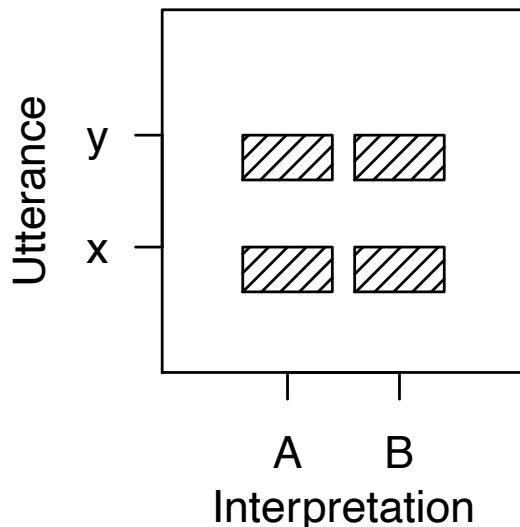
A simple communication game

- There are also many *non-Pareto-optimal* strategies:



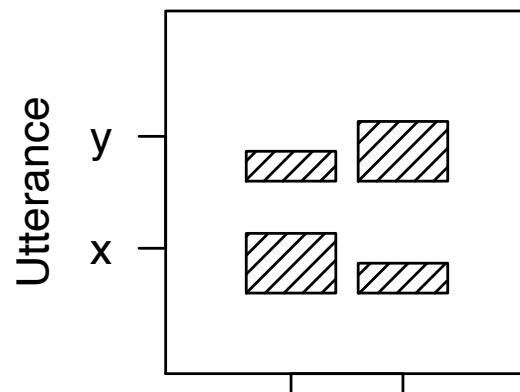
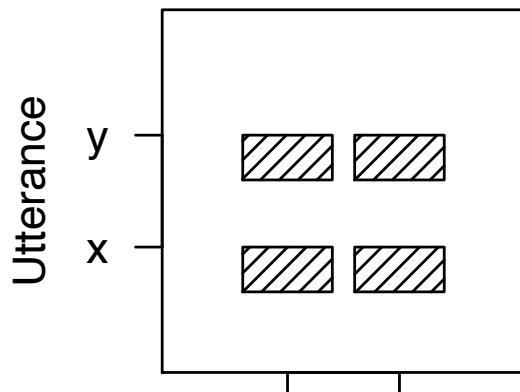
A simple communication game

- There are also many *non-Pareto-optimal* strategies:

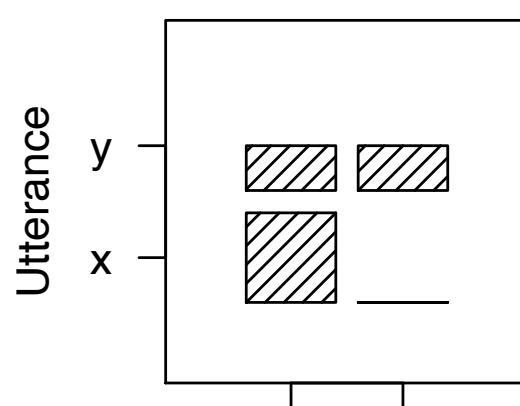
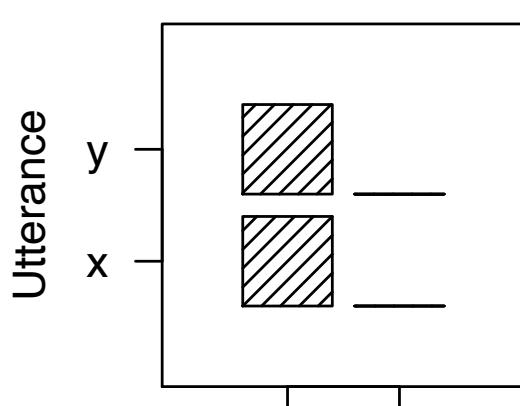


A simple communication game

- There are also many *non-Pareto-optimal* strategies:



...



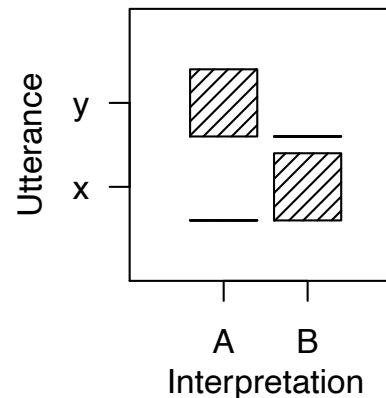
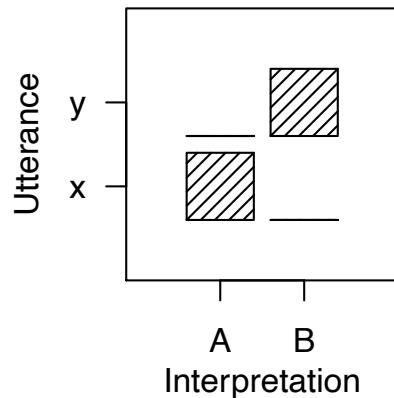
A simple communication game

A simple communication game

- *Efficient* communication would involve getting as close as possible to Pareto-optimal strategies...

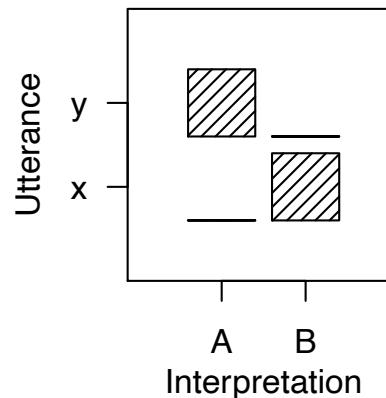
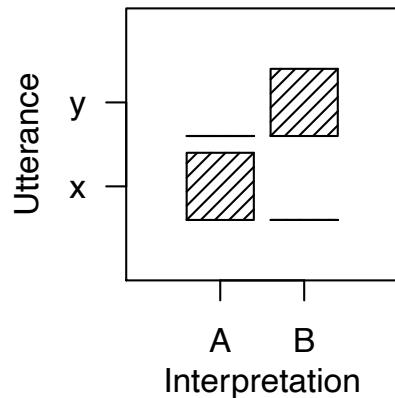
A simple communication game

- *Efficient* communication would involve getting as close as possible to Pareto-optimal strategies...



A simple communication game

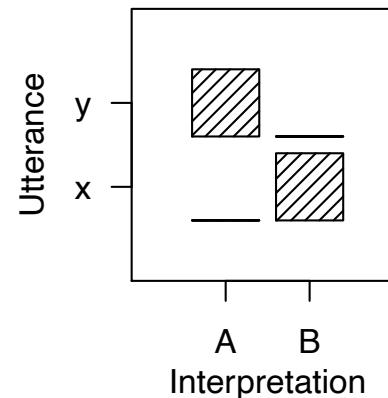
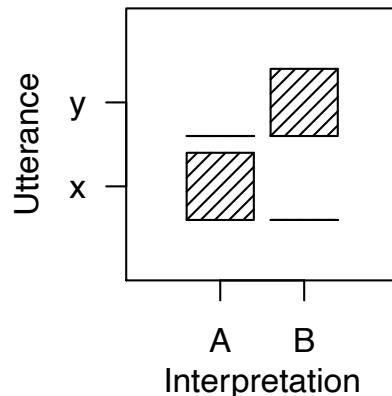
- *Efficient* communication would involve getting as close as possible to Pareto-optimal strategies...



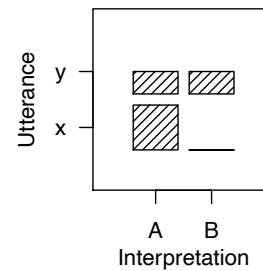
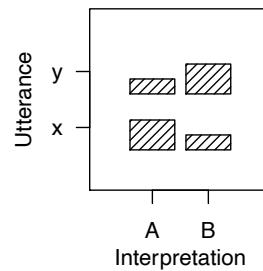
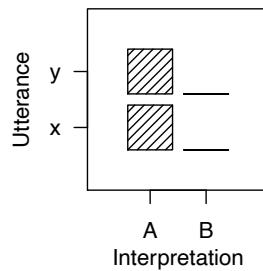
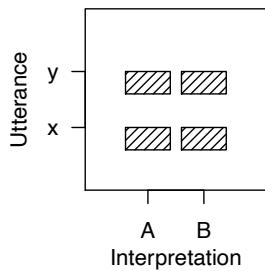
- ...and away from the suboptimal strategies

A simple communication game

- *Efficient* communication would involve getting as close as possible to Pareto-optimal strategies...



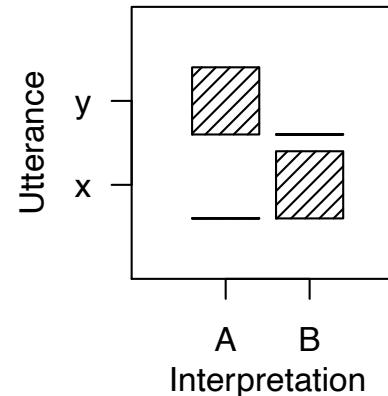
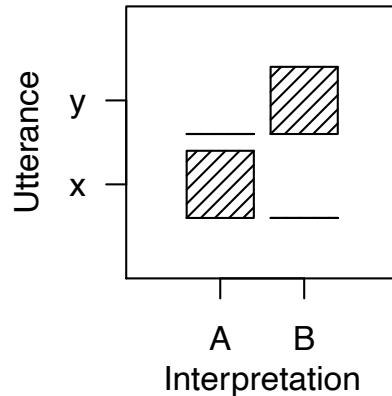
- ...and away from the suboptimal strategies



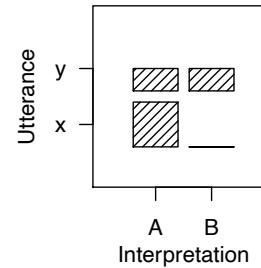
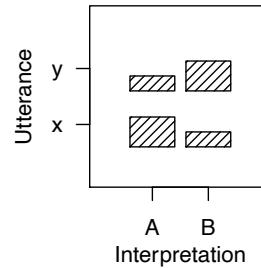
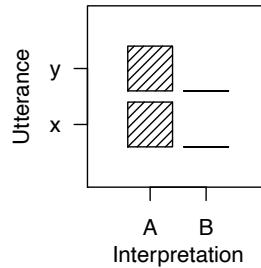
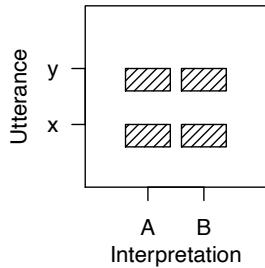
...

A simple communication game

- *Efficient* communication would involve getting as close as possible to Pareto-optimal strategies...



- ...and away from the suboptimal strategies



...

- ...but without conventions, there's no way to do this reliably!

Scalar implicature

Scalar implicature

- Consider the conventions offered us by *some* and *all*

Scalar implicature

- Consider the conventions offered us by ***some*** and ***all***
- Two meanings: \forall , $\exists \neg$

Scalar implicature

- Consider the conventions offered us by **some** and **all**
- Two meanings: \forall , $\exists \neg$
- Two signals:

Scalar implicature

- Consider the conventions offered us by **some** and **all**
- Two meanings: \forall , $\exists \neg$
- Two signals:
 - **all** is compatible only with meaning \forall

Scalar implicature

- Consider the conventions offered us by **some** and **all**
- Two meanings: \forall , $\exists \neg A$
- Two signals:
 - **all** is compatible only with meaning \forall
 - **some** is compatible with both meaning \forall and meaning $\exists \neg A$

Scalar implicature

- Consider the conventions offered us by **some** and **all**
- Two meanings: \forall , $\exists \neg \forall$
- Two signals:
 - **all** is compatible only with meaning \forall
 - **some** is compatible with both meaning \forall and meaning $\exists \neg \forall$
- For simplicity, assume prior $P(\exists \neg \forall) = P(\forall) = 1/2$

Bayesian theories of pragmatics

(Frank & Goodman, 2012; Bergen et al., 2016; Goodman & Frank, 2016; Franke, 2009; Jäger & Ebert, 2009; Jäger, 2011)

Bayesian theories of pragmatics

Assumptions:

Bayesian theories of pragmatics

Assumptions:

- Speaker and listener beliefs represented as probability distributions over world states

Bayesian theories of pragmatics

Assumptions:

- Speaker and listener beliefs represented as probability distributions over world states
- Joint communicative goal:

Bayesian theories of pragmatics

Assumptions:

- Speaker and listener beliefs represented as probability distributions over world states
- Joint communicative goal:
 - align the listener's beliefs with those of the speaker

Bayesian theories of pragmatics

Assumptions:

- Speaker and listener beliefs represented as probability distributions over world states
- Joint communicative goal:
 - align the listener's beliefs with those of the speaker
 - but maintain brevity while doing so!

Bayesian theories of pragmatics

Assumptions:

- Speaker and listener beliefs represented as probability distributions over world states
- Joint communicative goal:
 - align the listener's beliefs with those of the speaker
 - but maintain brevity while doing so!
- Grammar and the literal meanings of words are common knowledge between speaker and listener

Bayesian theories of pragmatics

Assumptions:

- Speaker and listener beliefs represented as probability distributions over world states
- Joint communicative goal:
 - align the listener's beliefs with those of the speaker
 - but maintain brevity while doing so!
- Grammar and the literal meanings of words are common knowledge between speaker and listener
- Speaker and listener can recursively reason (probabilistically) about each other

Scalar implicature

Scalar implicature

- Simple model of *literal interpretation*:

Scalar implicature

- Simple model of *literal interpretation*:
 - Listener rules out meanings incompatible with message

Scalar implicature

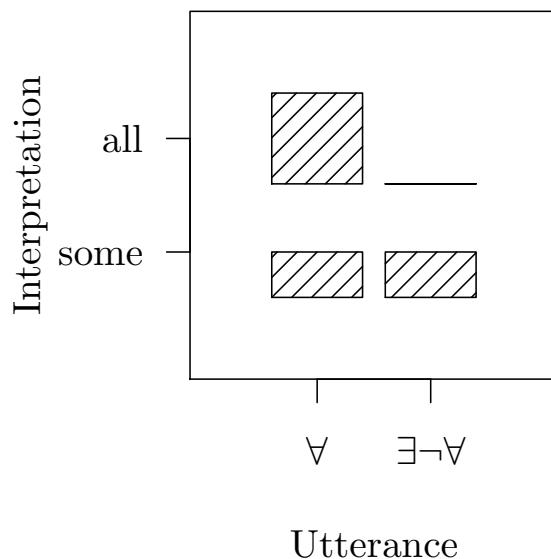
- Simple model of *literal interpretation*:
 - Listener rules out meanings incompatible with message
 - Among meanings compatible with message, prefers those with higher prior probability

Scalar implicature

- Simple model of *literal interpretation*:
 - Listener rules out meanings incompatible with message
 - Among meanings compatible with message, prefers those with higher prior probability
- Literal interpretation matrix for *some/all*:

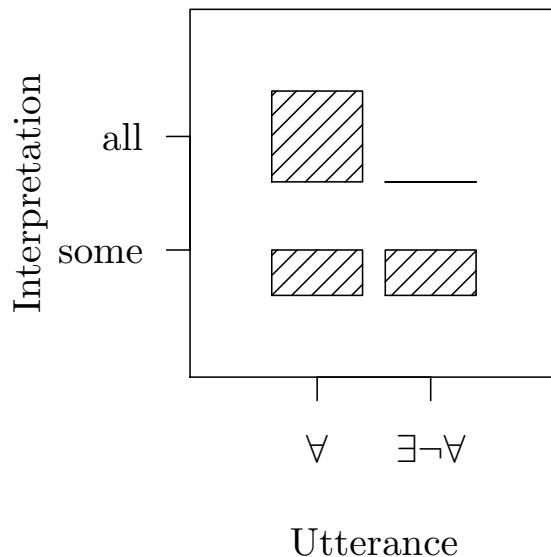
Scalar implicature

- Simple model of *literal interpretation*:
 - Listener rules out meanings incompatible with message
 - Among meanings compatible with message, prefers those with higher prior probability
- Literal interpretation matrix for *some/all*:



Scalar implicature

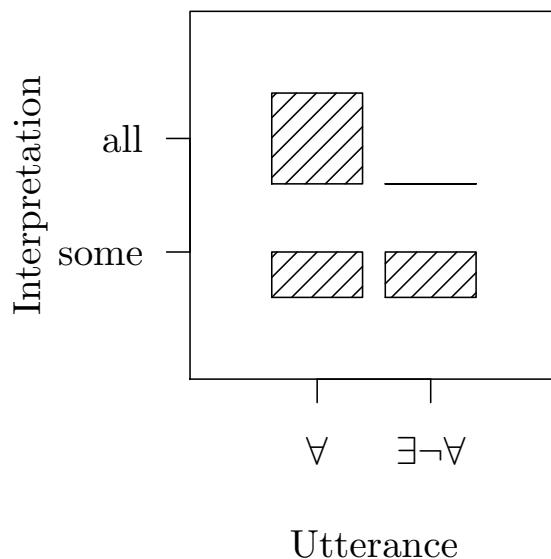
- Simple model of *literal interpretation*:
 - Listener rules out meanings incompatible with message
 - Among meanings compatible with message, prefers those with higher prior probability
- Literal interpretation matrix for *some/all*:



- This is non-Pareto—!

Scalar implicature

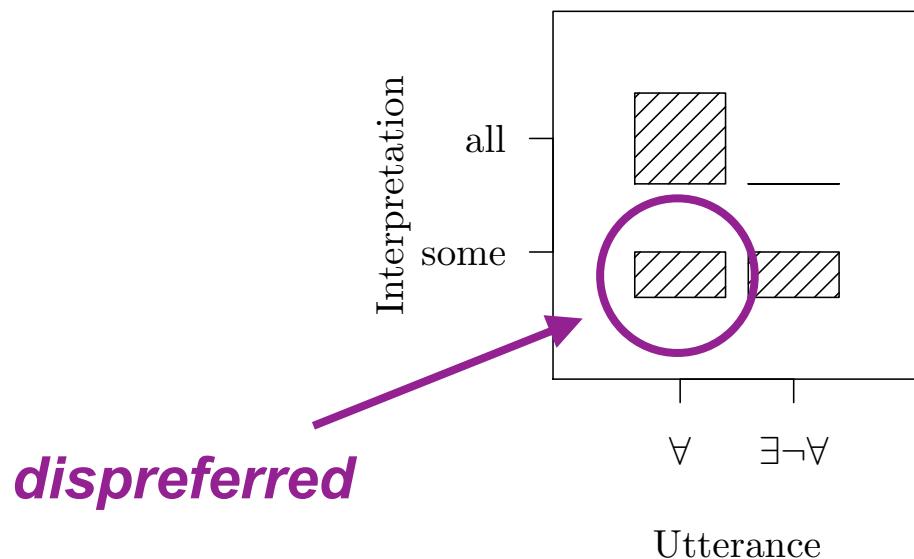
- Simple model of *literal interpretation*:
 - Listener rules out meanings incompatible with message
 - Among meanings compatible with message, prefers those with higher prior probability
- Literal interpretation matrix for *some/all*:



- This is non-Pareto—!
- —and it fails to capture human preferences

Scalar implicature

- Simple model of *literal interpretation*:
 - Listener rules out meanings incompatible with message
 - Among meanings compatible with message, prefers those with higher prior probability
- Literal interpretation matrix for *some/all*:



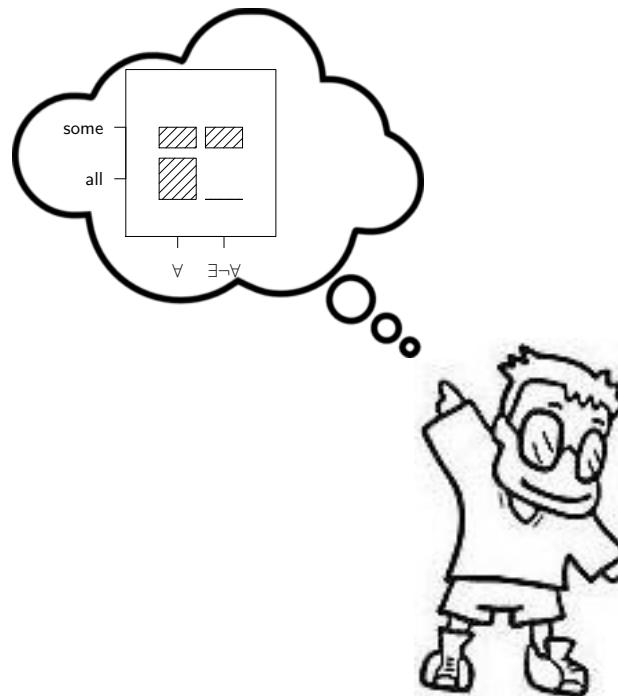
- This is non-Pareto—!
- —and it fails to capture human preferences

The Rational Speech-Act (RSA) model



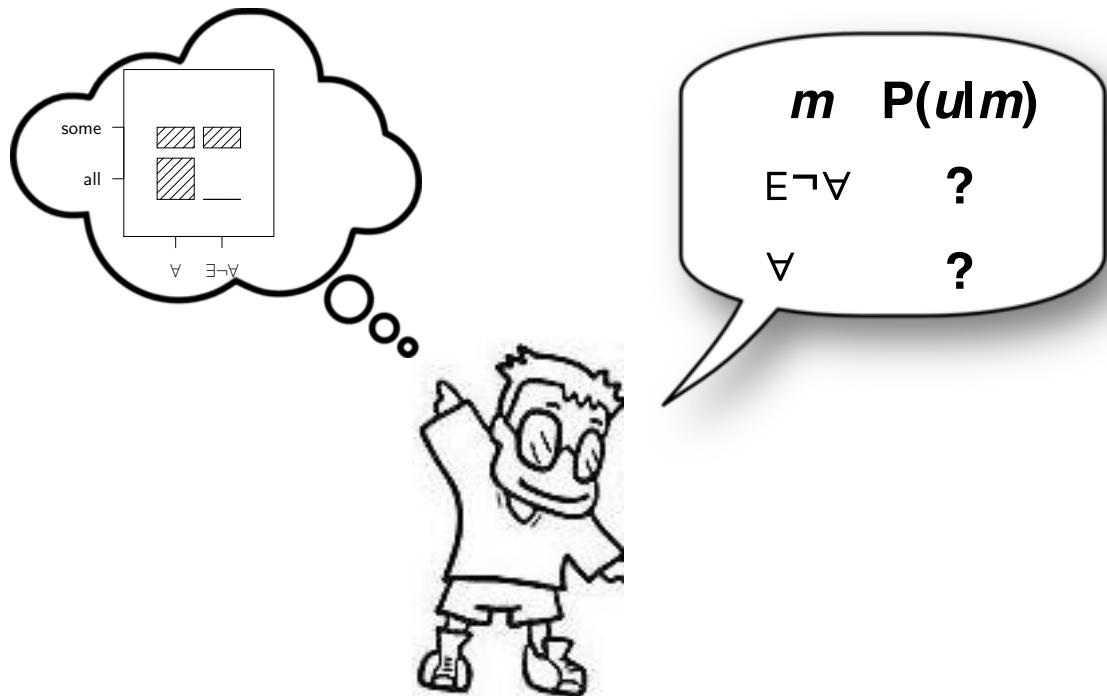
(Frank & Goodman, 2012;
Goodman & Frank, 2016)

The Rational Speech-Act (RSA) model

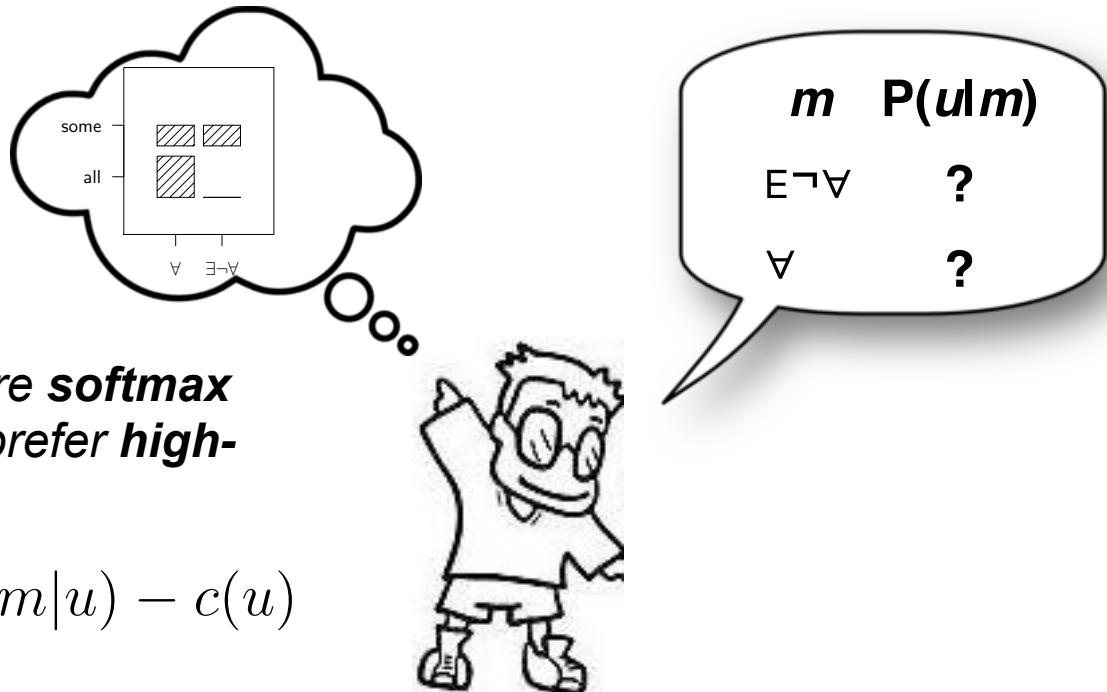


(Frank & Goodman, 2012;
Goodman & Frank, 2016)

The Rational Speech-Act (RSA) model



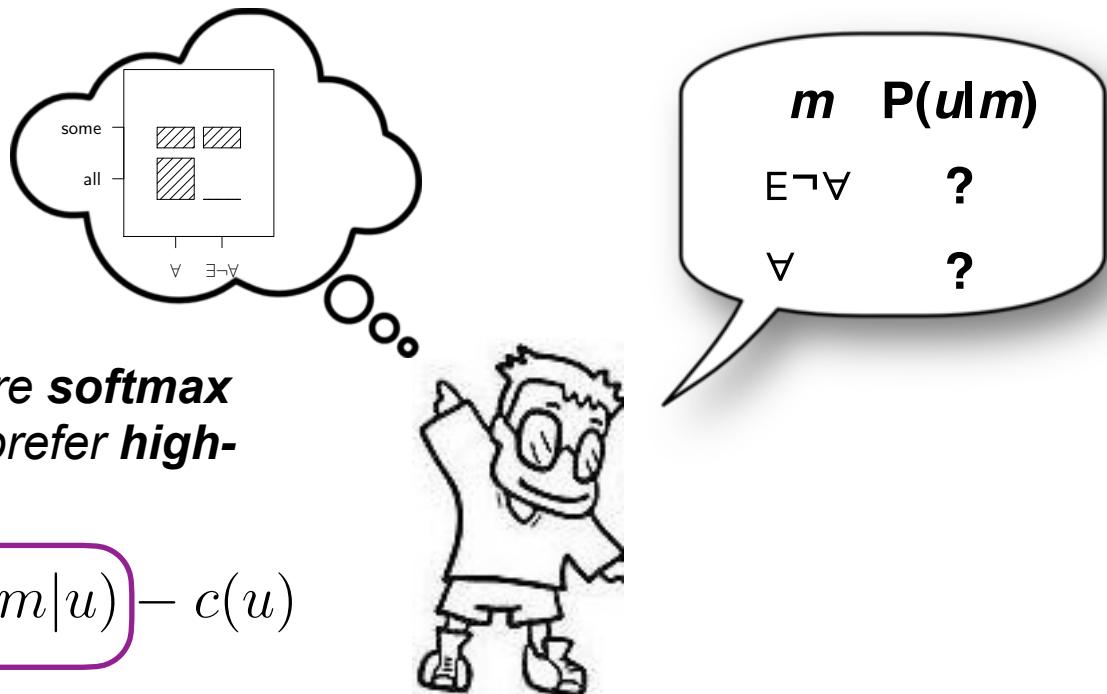
The Rational Speech-Act (RSA) model



*Modeling hypothesis: speakers are **softmax decision-theoretic agents** that prefer **high-utility utterances**:*

$$\text{Utility}^{(1)}(u; m) = \log P_{\text{Listener}}^{(0)}(m|u) - c(u)$$

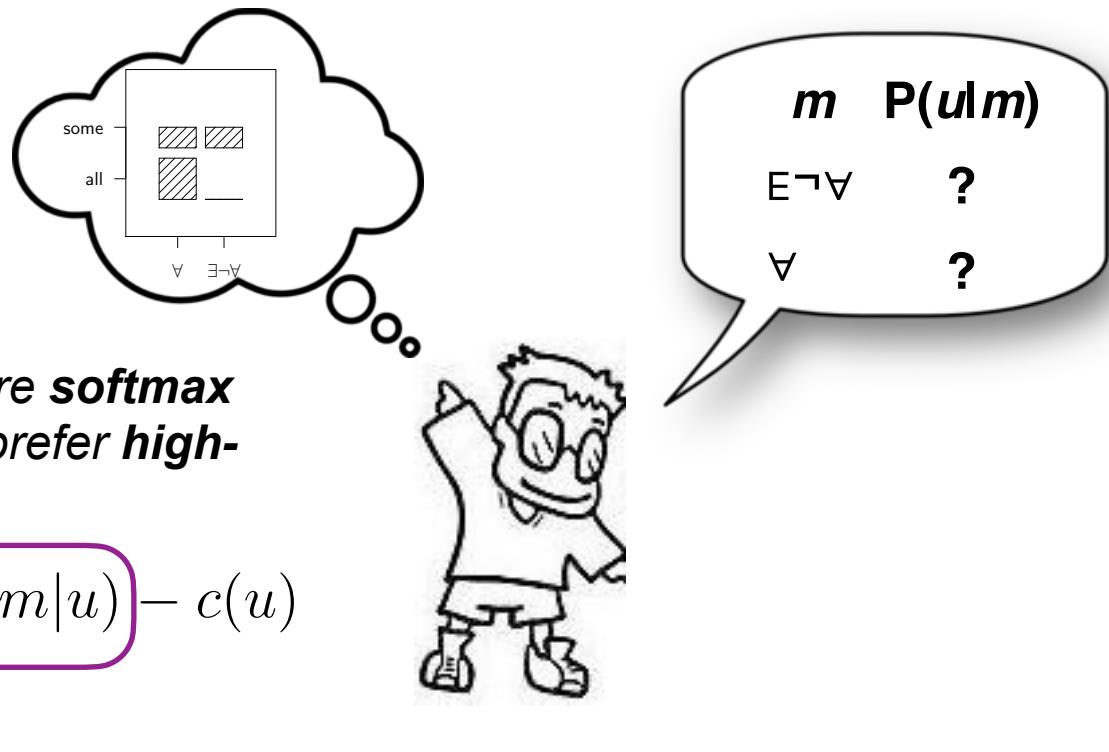
The Rational Speech-Act (RSA) model



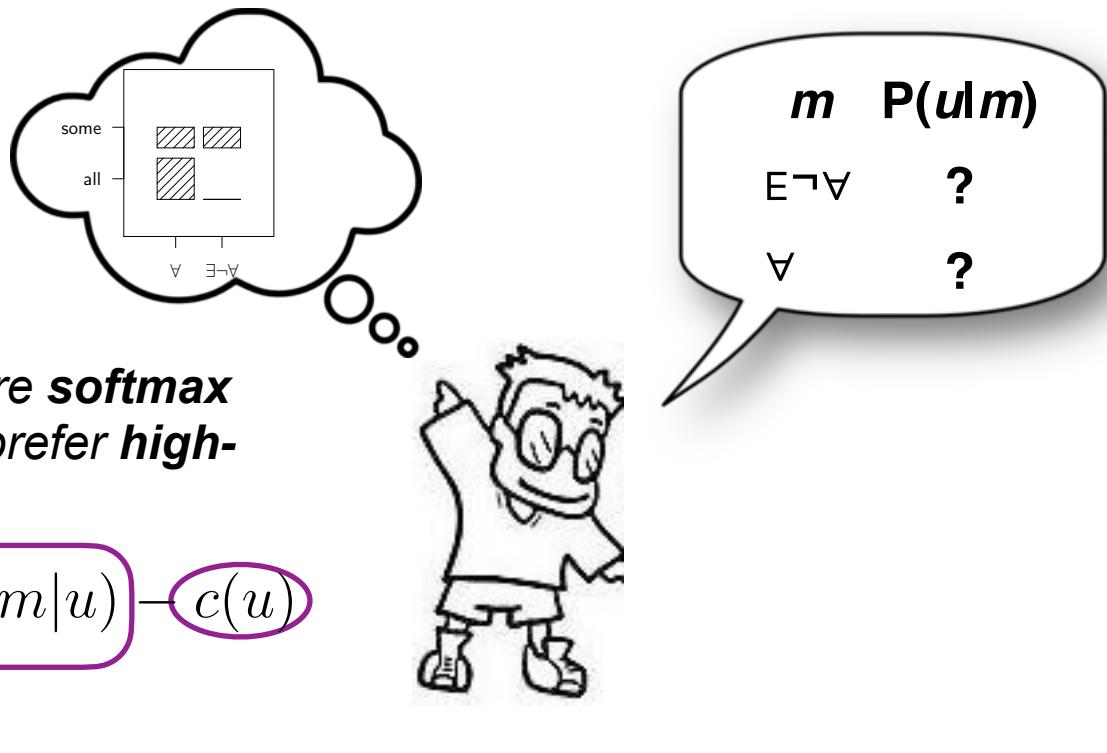
*Modeling hypothesis: speakers are **softmax decision-theoretic agents** that prefer **high-utility utterances**:*

$$\text{Utility}^{(1)}(u; m) = \log P_{\text{Listener}}^{(0)}(m|u) - c(u)$$

The Rational Speech-Act (RSA) model



The Rational Speech-Act (RSA) model



(Frank & Goodman, 2012;
Goodman & Frank, 2016)

The Rational Speech-Act (RSA) model



Modeling hypothesis: speakers are **softmax decision-theoretic agents** that prefer **high-utility utterances**:

$$\text{Utility}^{(1)}(u; m) = \log P_{\text{Listener}}^{(0)}(m|u) - c(u)$$

How surprised would u leave the listener about my intended meaning m ?

How intrinsically **costly** (time-consuming, effortful, taboo, ...) is the utterance u ?

The Rational Speech-Act (RSA) model



Modeling hypothesis: speakers are **softmax decision-theoretic agents** that prefer **high-utility utterances**:

$$\text{Utility}^{(1)}(u; m) = \log P_{\text{Listener}}^{(0)}(m|u) - c(u)$$

How surprised would u leave the listener about my intended meaning m ?

How intrinsically costly (time-consuming, effortful, taboo, ...) is the utterance u ?

$$P_{\text{Speaker}}^{(1)} \propto [\text{Utility}^{(1)}(u; m)]^\lambda$$

(Frank & Goodman, 2012;
Goodman & Frank, 2016)

The Rational Speech-Act (RSA) model



Modeling hypothesis: speakers are **softmax decision-theoretic agents** that prefer **high-utility utterances**:

$$\text{Utility}^{(1)}(u; m) = \log P_{\text{Listener}}^{(0)}(m|u) - c(u)$$

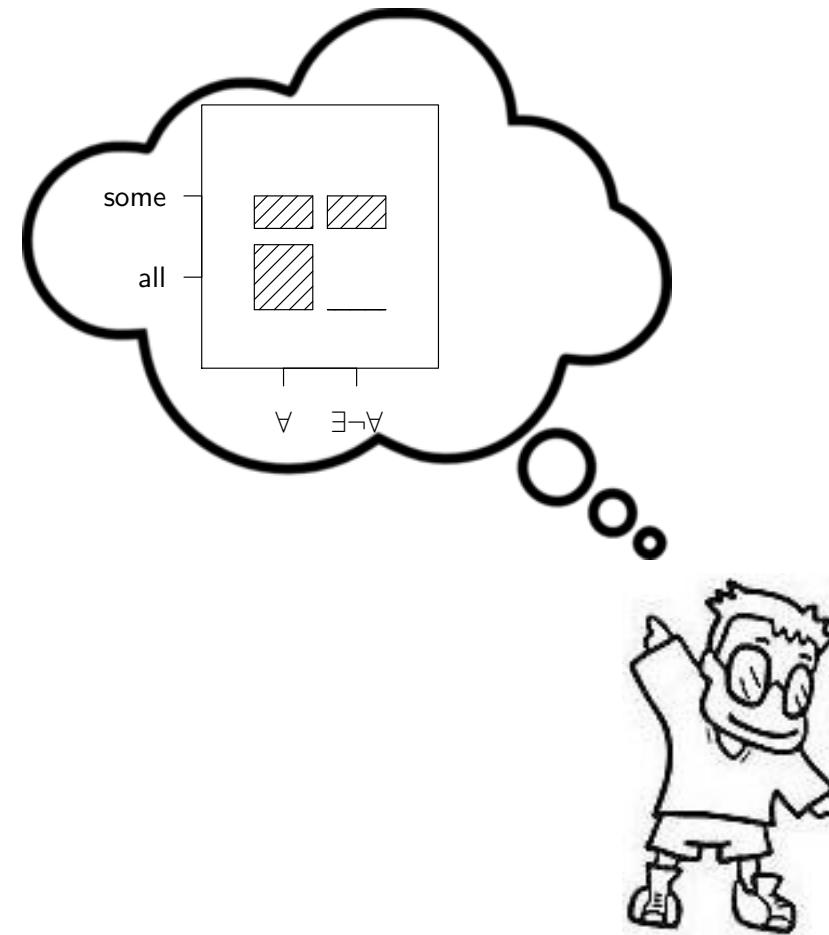
How surprised would u leave the listener about my intended meaning m ?

How intrinsically costly (time-consuming, effortful, taboo, ...) is the utterance u ?

$$P_{\text{Speaker}}^{(1)} \propto [\text{Utility}^{(1)}(u; m)]^\lambda$$

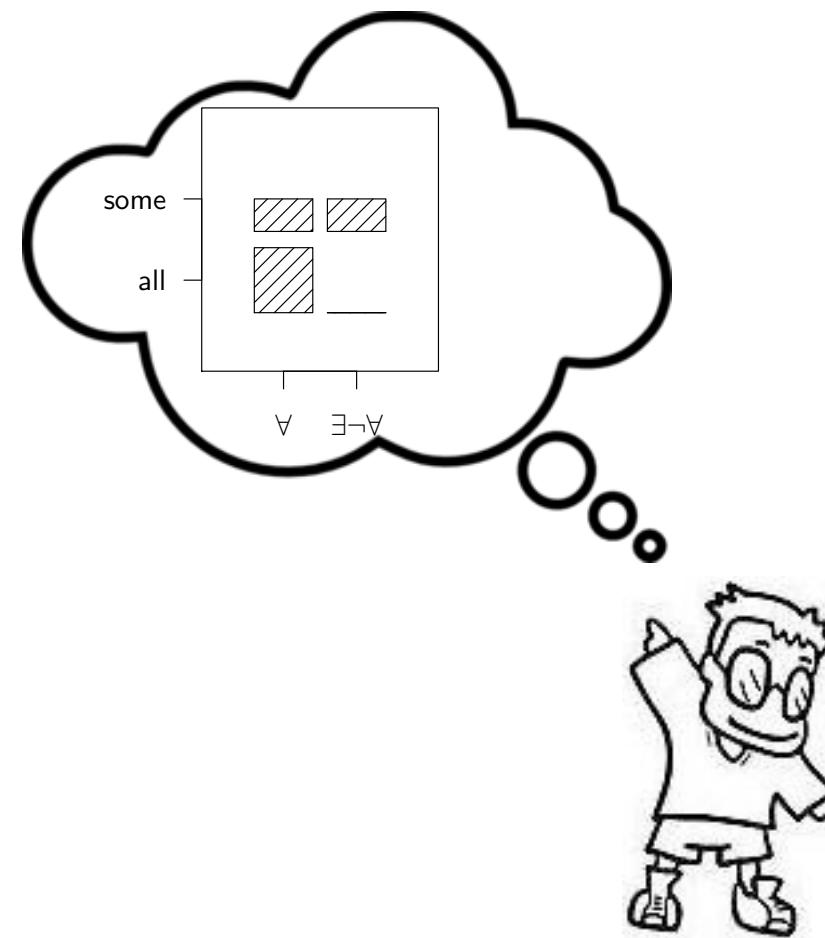
Softmax optimality parameter

The Rational Speech-Act (RSA) model



(Frank & Goodman, 2012;
Goodman & Stühlmüller, 2013)

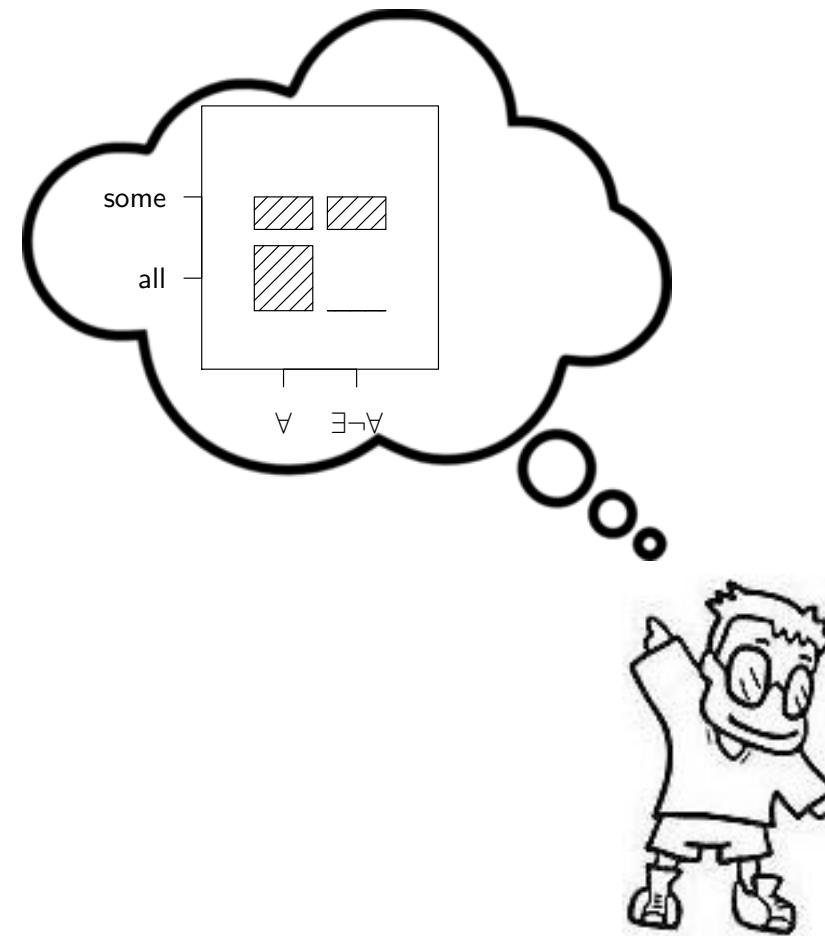
The Rational Speech-Act (RSA) model



$$P_{Speaker}^{(1)}(u|m) \propto \left[P_{Listener}^{(0)}(m|u) e^{-c(u)} \right]^\lambda$$

(Frank & Goodman, 2012;
Goodman & Stühlmüller, 2013)

The Rational Speech-Act (RSA) model

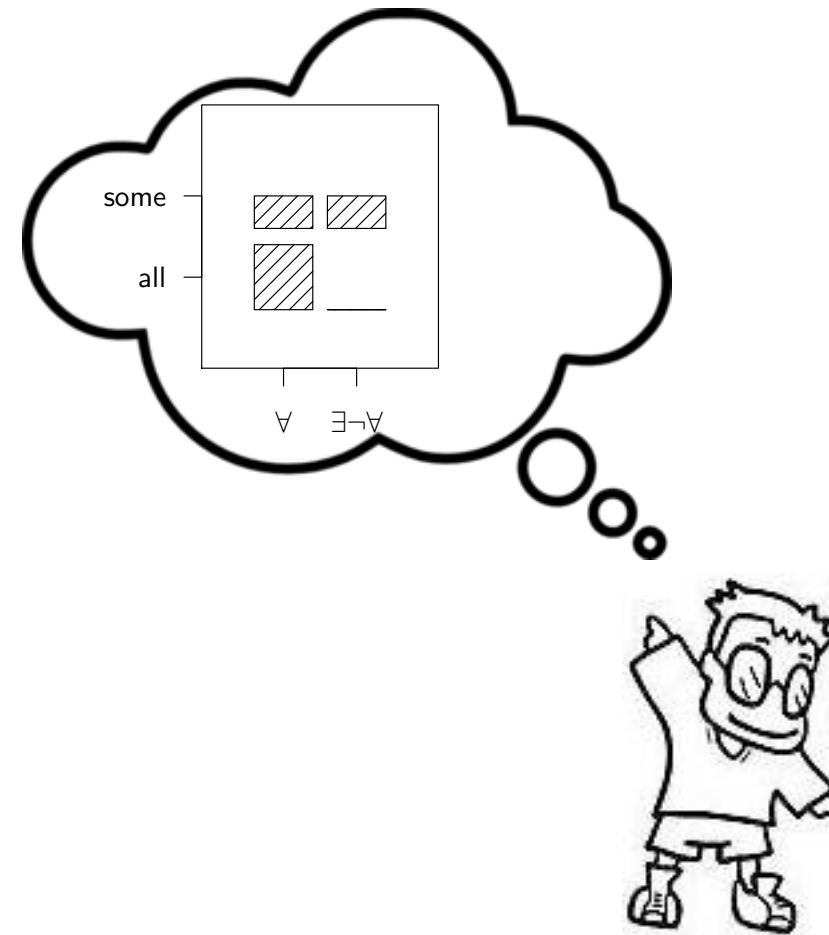


Utterance cost

$$P_{Speaker}^{(1)}(u|m) \propto \left[P_{Listener}^{(0)}(m|u) e^{-c(u)} \right]^\lambda$$

(Frank & Goodman, 2012;
Goodman & Stühlmuller, 2013)

The Rational Speech-Act (RSA) model



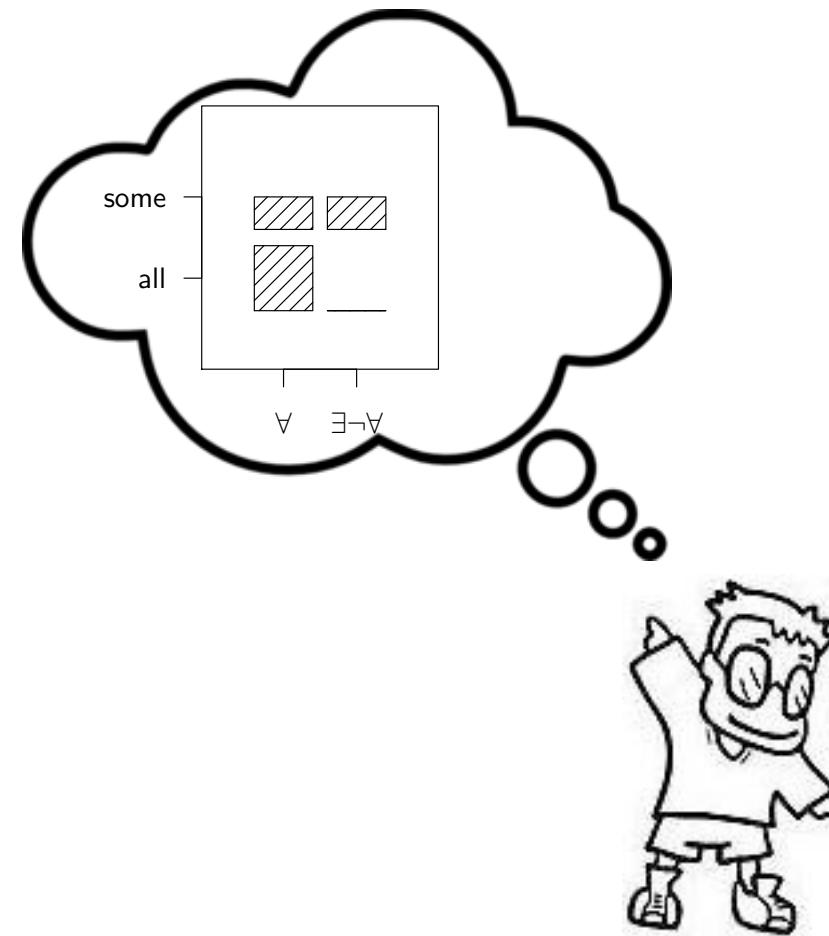
$$c(\text{some}) = c(\text{all}) = 0$$

Utterance cost

$$P_{\text{Speaker}}^{(1)}(u|m) \propto \left[P_{\text{Listener}}^{(0)}(m|u) e^{-c(u)} \right]^\lambda$$

(Frank & Goodman, 2012;
Goodman & Stühlmuller, 2013)

The Rational Speech-Act (RSA) model



$$c(\text{some}) = c(\text{all}) = 0$$

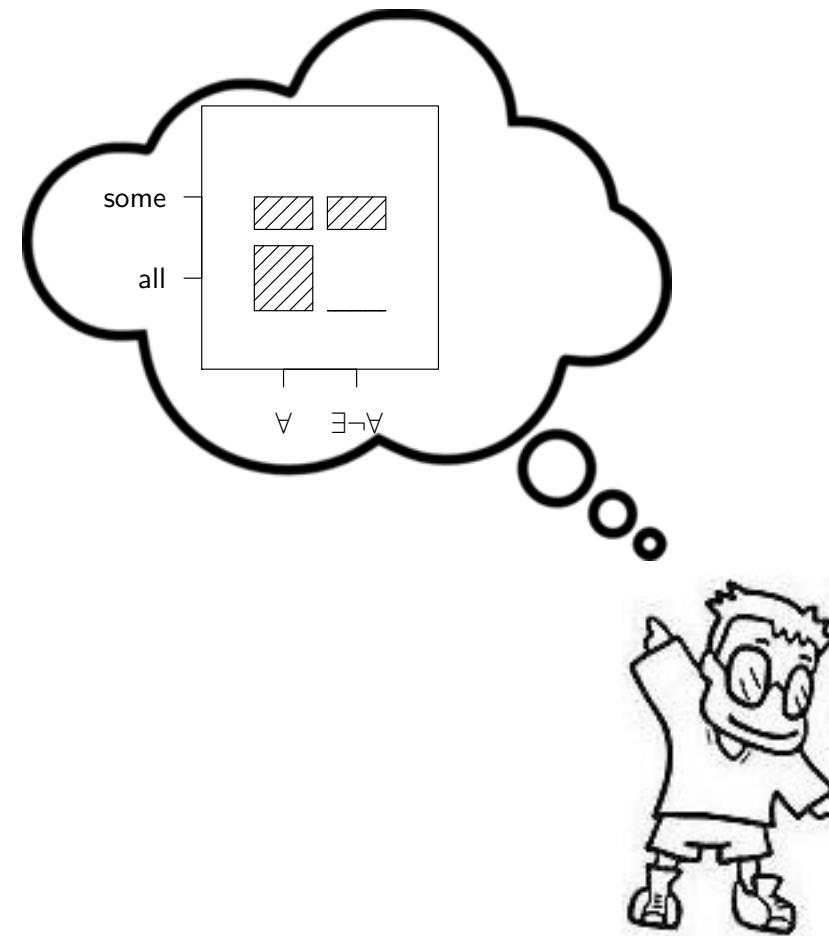
Utterance cost

$$P_{\text{Speaker}}^{(1)}(u|m) \propto \left[P_{\text{Listener}}^{(0)}(m|u) e^{-c(u)} \right]^\lambda$$

“Greedy optimality” parameter

(Frank & Goodman, 2012;
Goodman & Stühlmüller, 2013)

The Rational Speech-Act (RSA) model



$$c(\text{some}) = c(\text{all}) = 0$$

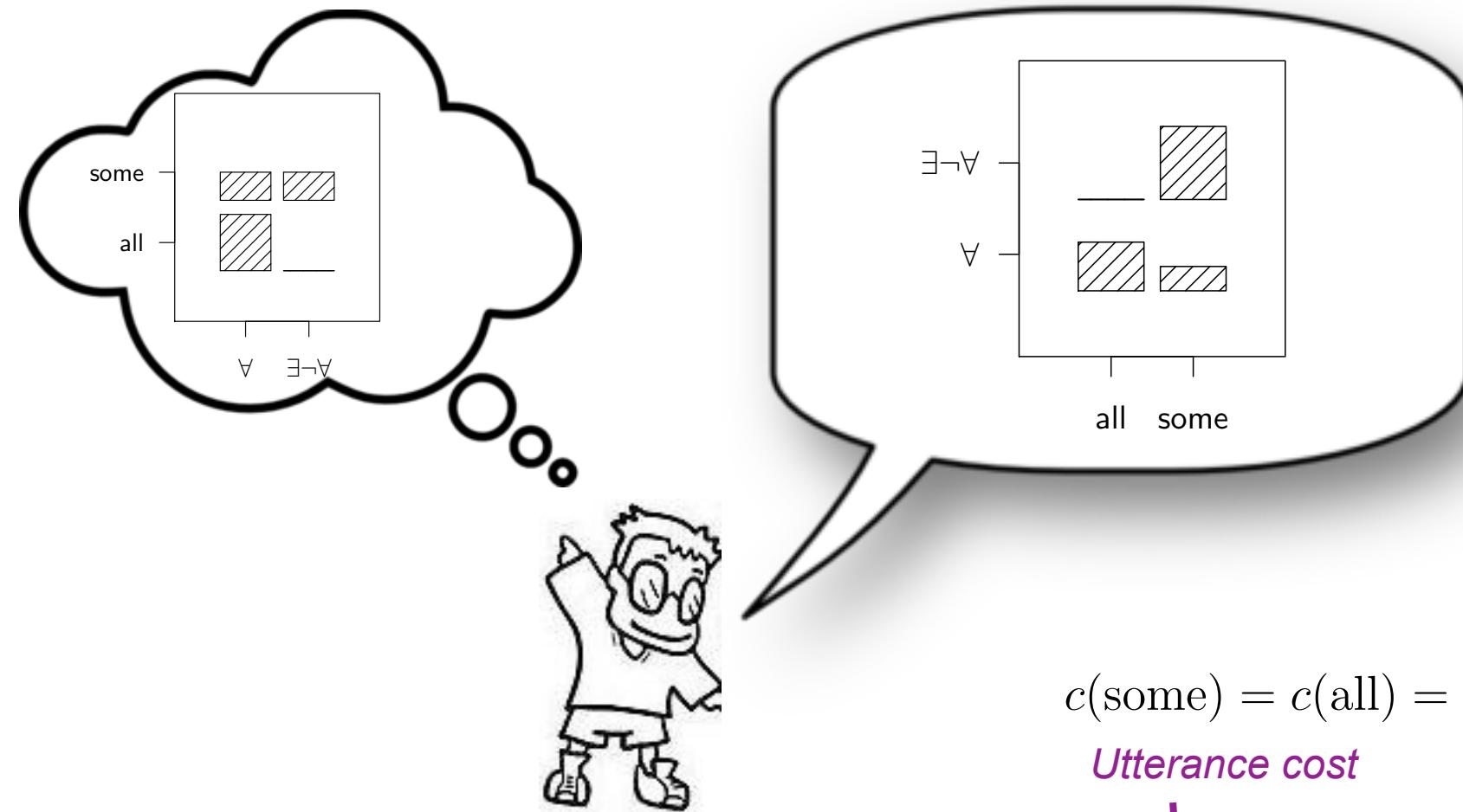
Utterance cost

$$P_{\text{Speaker}}^{(1)}(u|m) \propto \left[P_{\text{Listener}}^{(0)}(m|u) e^{-c(u)} \right]^\lambda \quad \lambda = 1$$

“Greedy optimality” parameter

(Frank & Goodman, 2012;
Goodman & Stühlmüller, 2013)

The Rational Speech-Act (RSA) model



$$c(\text{some}) = c(\text{all}) = 0$$

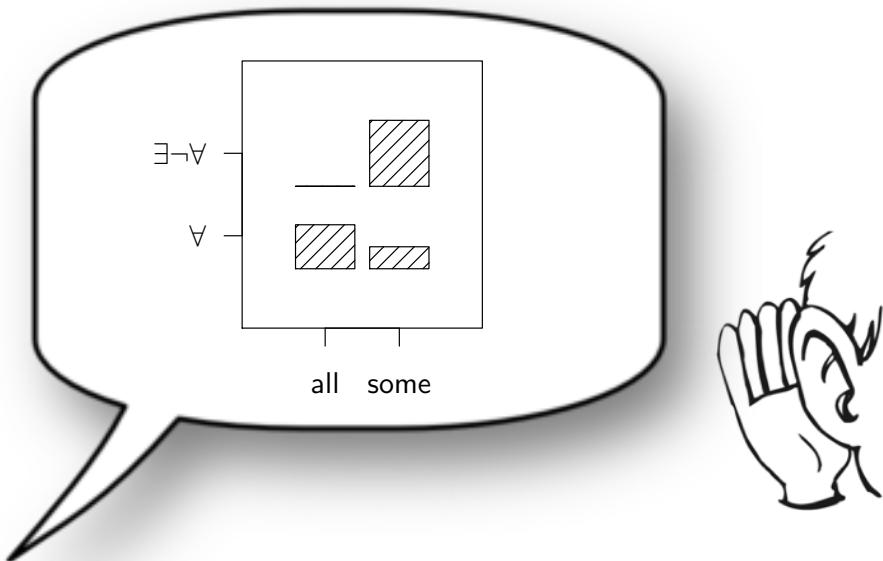
Utterance cost

$$P_{\text{Speaker}}^{(1)}(u|m) \propto \left[P_{\text{Listener}}^{(0)}(m|u) e^{-c(u)} \right]^\lambda \quad \lambda = 1$$

“Greedy optimality” parameter

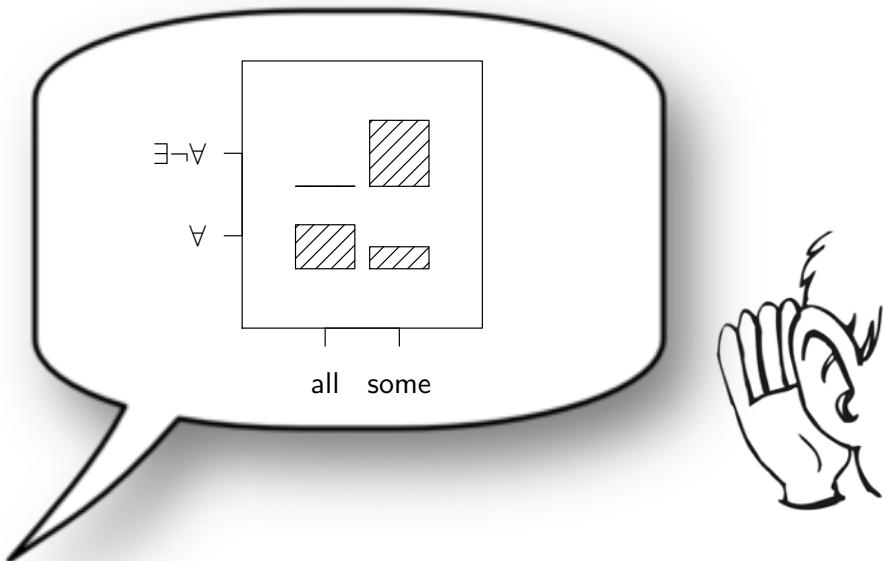
(Frank & Goodman, 2012;
Goodman & Stühlmüller, 2013)

Scalar implicature in RSA: listening



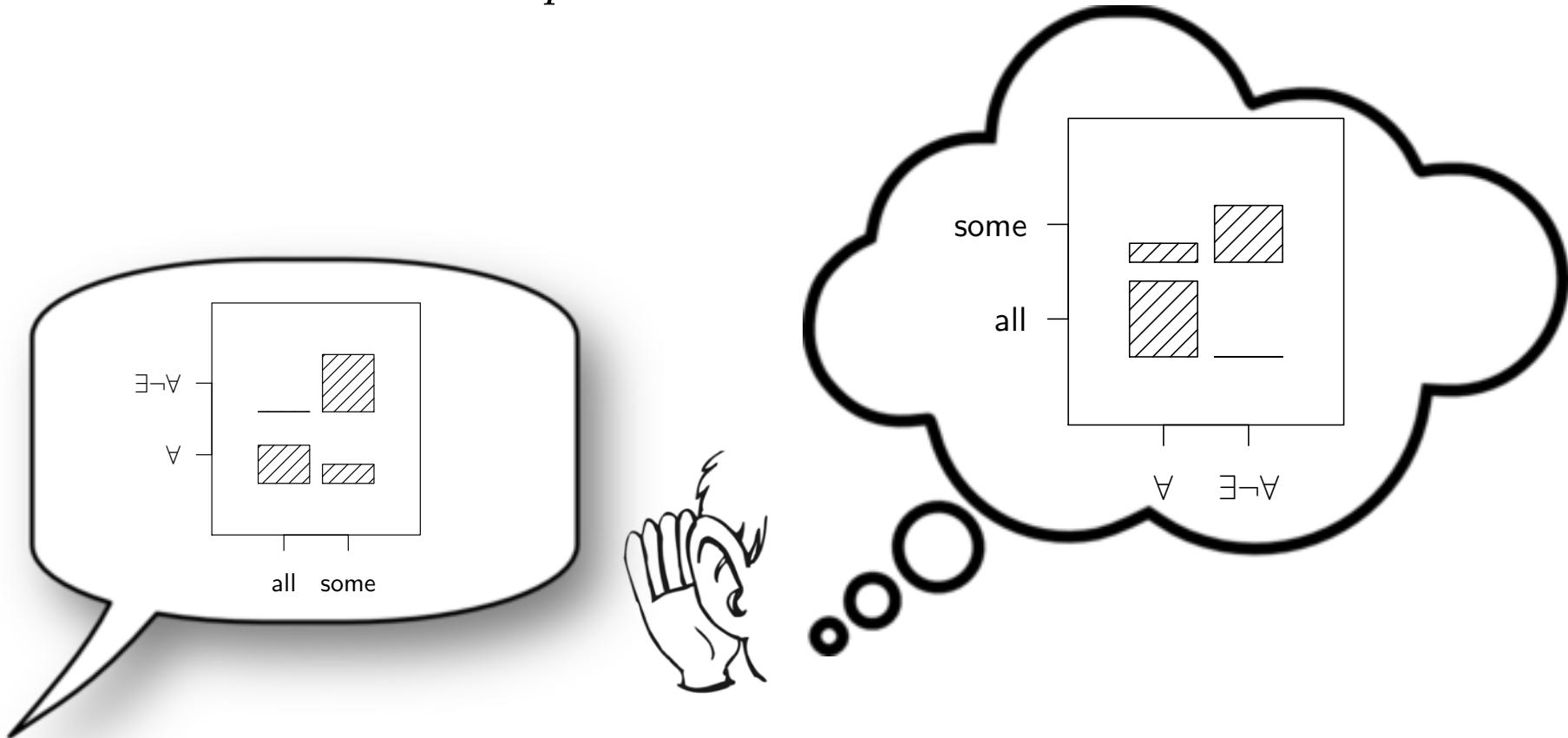
Scalar implicature in RSA: listening

$$P_{Listener}^{(1)}(m|u) \propto P_{Speaker}^{(1)}(u|m)P(m)$$



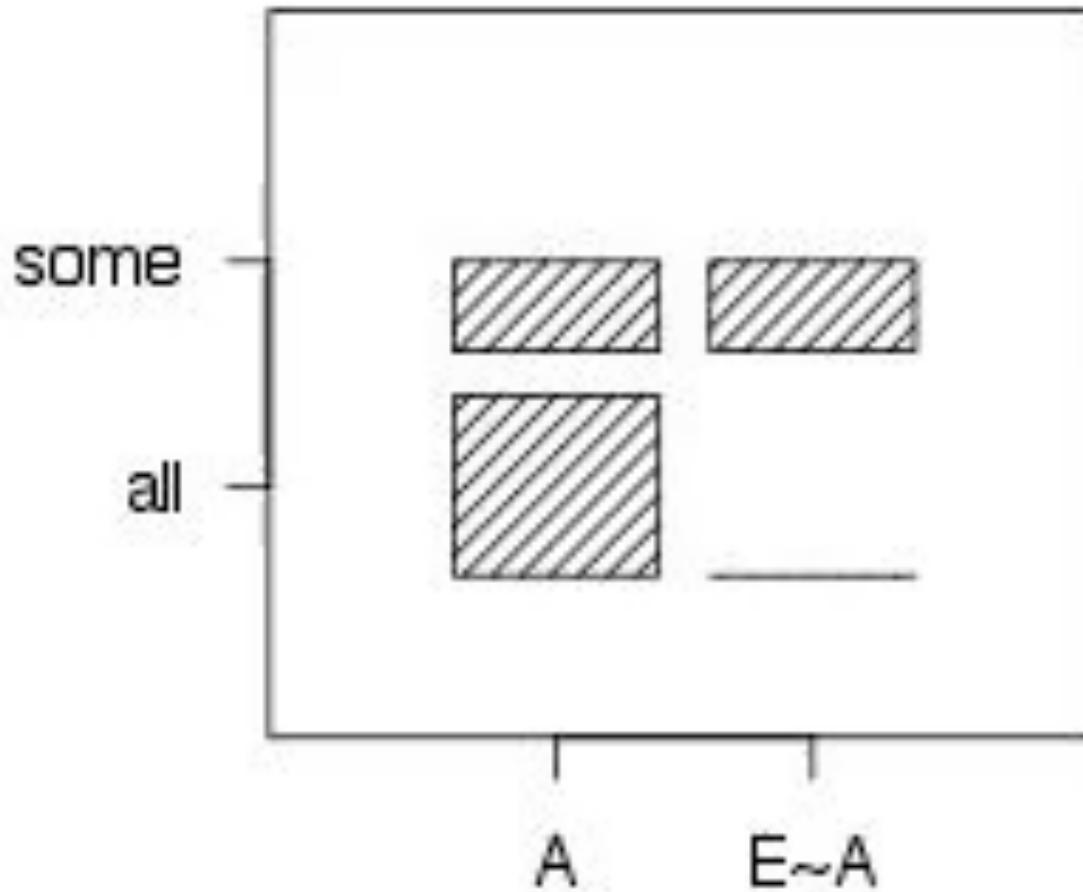
Scalar implicature in RSA: listening

$$P_{Listener}^{(1)}(m|u) \propto P_{Speaker}^{(1)}(u|m)P(m)$$



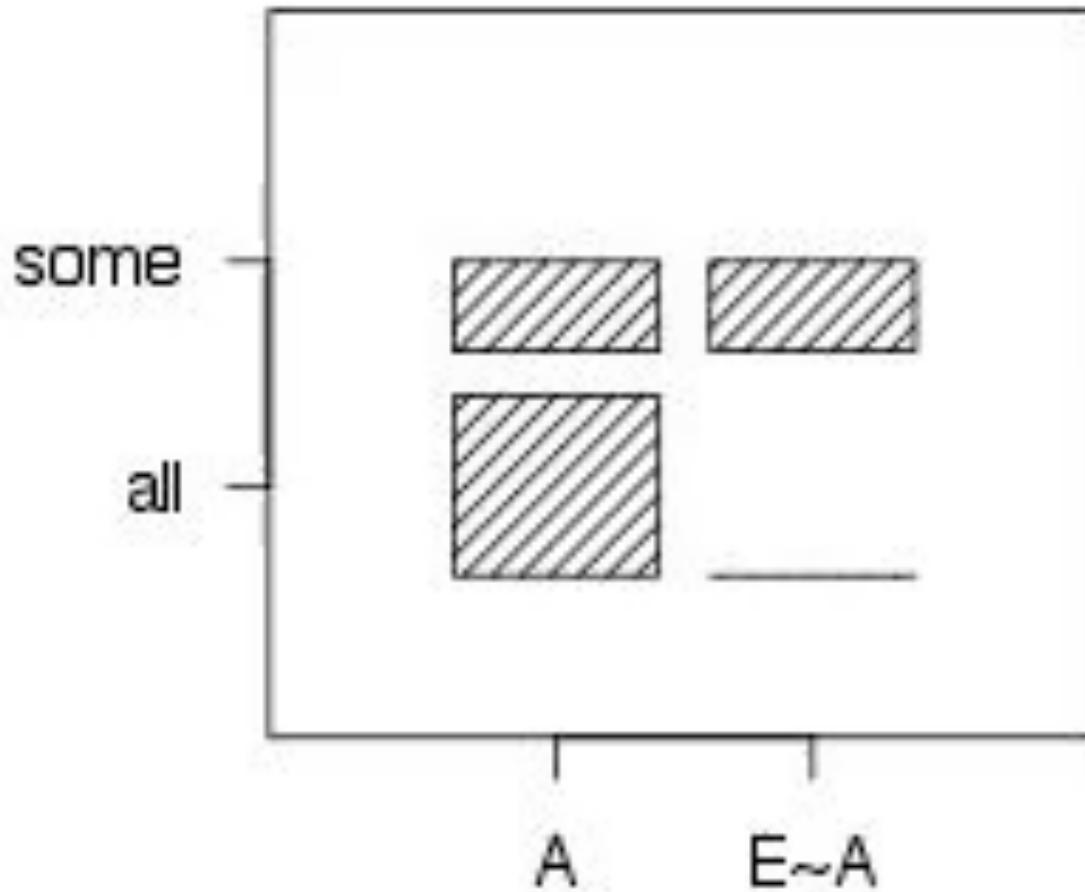
Speaker—listener recursion in RSA

- The process of recursion strengthens the implicature



Speaker—listener recursion in RSA

- The process of recursion strengthens the implicature



Conceptual framing

- Speaker and listener got (close) to a Pareto-optimal strategy by combining two ingredients:
 - Language knowledge (lexicon/grammar) as the *raw materials* for initial solutions to the communication game
 - General principles of socio-cognitive reasoning to craft these raw materials into more efficient solutions
- These two ingredients together allow discourse participants to do so *much more* than either one alone

Levinson's (2000) typology of implicature

Levinson's (2000) typology of implicature

Q-implicature

(Horn's Q)

What isn't said isn't meant

Levinson's (2000) typology of implicature

Q-implicature

(Horn's Q)

What isn't said isn't meant

Pat has three children



Pat has **exactly three** children

Levinson's (2000) typology of implicature

Q-implicature

(Horn's Q)

What isn't said isn't meant

Pat has three children



Pat has **exactly three** children

I injured a child



I didn't injure **my** child

Levinson's (2000) typology of implicature

Q-implicature

(Horn's Q)

What isn't said isn't meant

Pat has three children



Pat has **exactly three** children

M-implicature

(Horn's "division of pragmatic labor")

*Align utterance simplicity
with situation stereotypicality*

I injured a child



I didn't injure **my** child

Levinson's (2000) typology of implicature

Q-implicature

(Horn's Q)

What isn't said isn't meant

Pat has three children



Pat has **exactly three** children

M-implicature

(Horn's "division of pragmatic labor")

*Align utterance simplicity
with situation stereotypicality*

I started the car



...by **just** turning the key

I injured a child



I didn't injure **my** child

Levinson's (2000) typology of implicature

Q-implicature

(Horn's Q)

What isn't said isn't meant

Pat has three children



Pat has **exactly three** children

M-implicature

(Horn's "division of pragmatic labor")

*Align utterance simplicity
with situation stereotypicality*

I started the car



...by **just** turning the key

I injured a child



I didn't injure **my** child

I got the car to start



...I needed to do **more than**
just turn the key

Levinson's (2000) typology of implicature

Q-implicature

(Horn's Q)

What isn't said isn't meant

Pat has three children



Pat has **exactly three** children

M-implicature

(Horn's "division of pragmatic labor")

*Align utterance simplicity
with situation stereotypicality*

I started the car



...by **just** turning the key

I injured a child



I didn't injure **my** child

I got the car to start



...I needed to do **more than**
just turn the key

I-implicature

(Horn's R, sort of)

*Interpret utterances as
the prototypical case*

Levinson's (2000) typology of implicature

Q-implicature

(Horn's Q)

What isn't said isn't meant

Pat has three children



Pat has **exactly three** children

I injured a child



I didn't injure **my** child

M-implicature

(Horn's "division of pragmatic labor")

*Align utterance simplicity
with situation stereotypicality*

I started the car



...by **just** turning the key

I got the car to start



...I needed to do **more than**
just turn the key

I-implicature

(Horn's R, sort of)

*Interpret utterances as
the prototypical case*

The cup is on the table



It's **in contact with** the table

Levinson's (2000) typology of implicature

Q-implicature

(Horn's Q)

What isn't said isn't meant

Pat has three children



Pat has **exactly three** children

I injured a child



I didn't injure **my** child

M-implicature

(Horn's "division of pragmatic labor")

*Align utterance simplicity
with situation stereotypicality*

I started the car



...by **just** turning the key

I got the car to start



...I needed to do **more than**
just turn the key

I-implicature

(Horn's R, sort of)

*Interpret utterances as
the prototypical case*

The cup is on the table



It's **in contact with** the table

I injured a finger



I injured **my own** finger

Levinson's (2000) typology of implicature

Q-implicature

(Horn's Q)

What isn't said isn't meant

Pat has three children



Pat has **exactly three** children

I injured a child



I didn't injure **my** child

M-implicature

(Horn's "division of pragmatic labor")

*Align utterance simplicity
with situation stereotypicality*

I started the car



...by **just** turning the key

Q/I Tradeoff

I got the car to start



...I needed to do **more than**
just turn the key

I-implicature

(Horn's R, sort of)

*Interpret utterances as
the prototypical case*

The cup is on the table



It's **in contact with** the table

I injured a finger



I injured **my own** finger

Levinson's (2000) typology of implicature

Q-implicature

(Horn's Q)

What isn't said isn't meant

Pat has three children



Pat has **exactly three** children

I injured a child



I didn't injure **my** child

M-implicature

(Horn's "division of pragmatic labor")

*Align utterance simplicity
with situation stereotypicality*

I started the car



...by **just** turning the key

I got the car to start



...I needed to do **more than**
just turn the key

I-implicature

(Horn's R, sort of)

*Interpret utterances as
the prototypical case*

The cup is on the table



It's **in contact with** the table

I injured a finger



I injured **my own** finger

Levinson's (2000) typology of implicature

Q-implicature

(Horn's Q)

What isn't said isn't meant

Pat has three children



Pat has **exactly three** children

I injured a child



I didn't injure **my** child

M-implicature

(Horn's "division of pragmatic labor")

*Align utterance simplicity
with situation stereotypicality*

I started the car



...by **just** turning the key

I got the car to start



...I needed to do **more than**
just turn the key

I-implicature

(Horn's R, sort of)

*Interpret utterances as
the prototypical case*

The cup is on the table



It's **in contact with** the table

I injured a finger



I injured **my own** finger

Can we explain this typology from basic principles in a probabilistic pragmatic framework, respecting linguistic form, semantic composition, and world knowledge?

Q/I tradeoff in rational speech-act theory

I injured a child→it WASN'T my child

Q/I tradeoff in rational speech-act theory

I injured a child→*it WASN'T my child*

Assumed alternative set:

*I injured **my** child*

*I injured **a** child*

*I injured **someone else's** child*

Q/I tradeoff in rational speech-act theory

I injured a child → *it WASN'T my child*

$$c(\text{my}) = c(a) = 0$$

$$c(\text{someone else's}) = 1$$

$$P(\text{MINE}) = \frac{1}{2}$$

Assumed alternative set:

I injured my child

I injured a child

I injured someone else's child

Q/I tradeoff in rational speech-act theory

I injured a child → *it WASN'T my child*

$$c(\text{my}) = c(a) = 0$$

$$c(\text{someone else's}) = 1$$

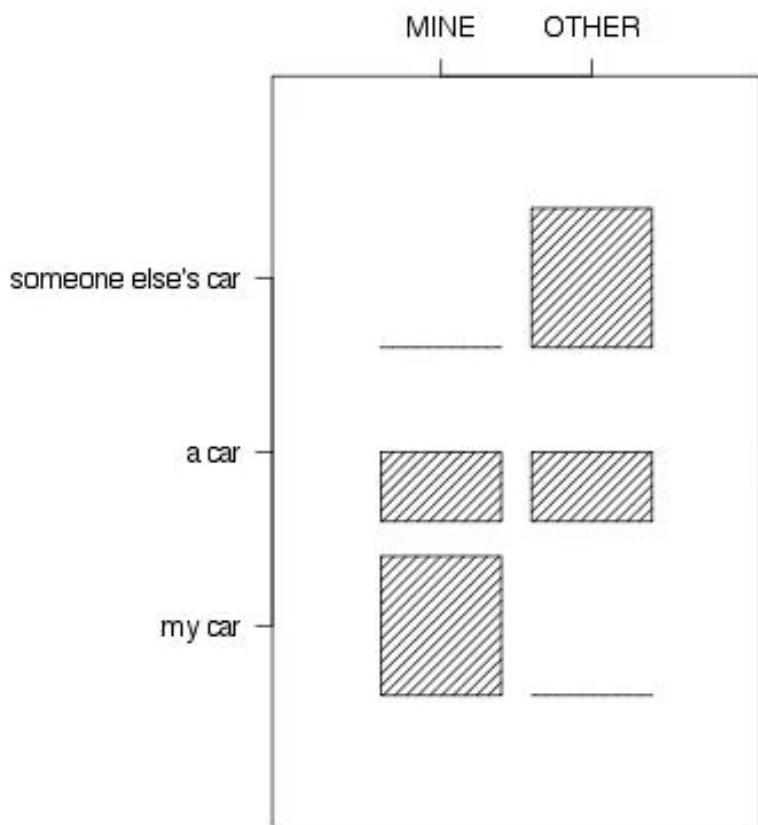
$$P(\text{MINE}) = \frac{1}{2}$$

Assumed alternative set:

I injured my child

I injured a child

I injured someone else's child



Q/I tradeoff in rational speech-act theory

I injured a child → *it WASN'T my child*

$$c(\text{my}) = c(a) = 0$$

$$c(\text{someone else's}) = 1$$

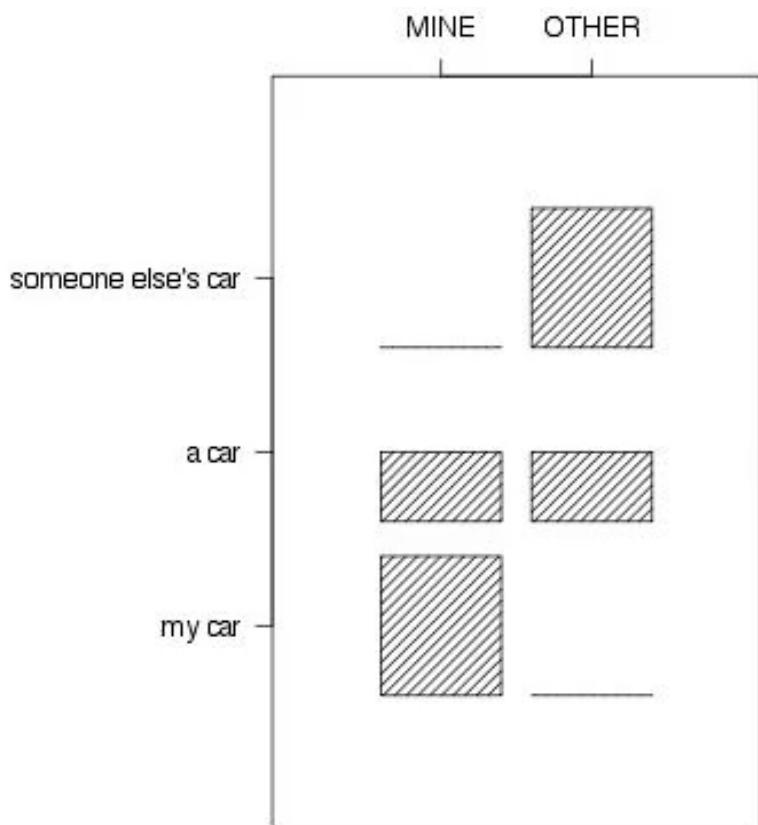
$$P(\text{MINE}) = \frac{1}{2}$$

Assumed alternative set:

I injured my child

I injured a child

I injured someone else's child



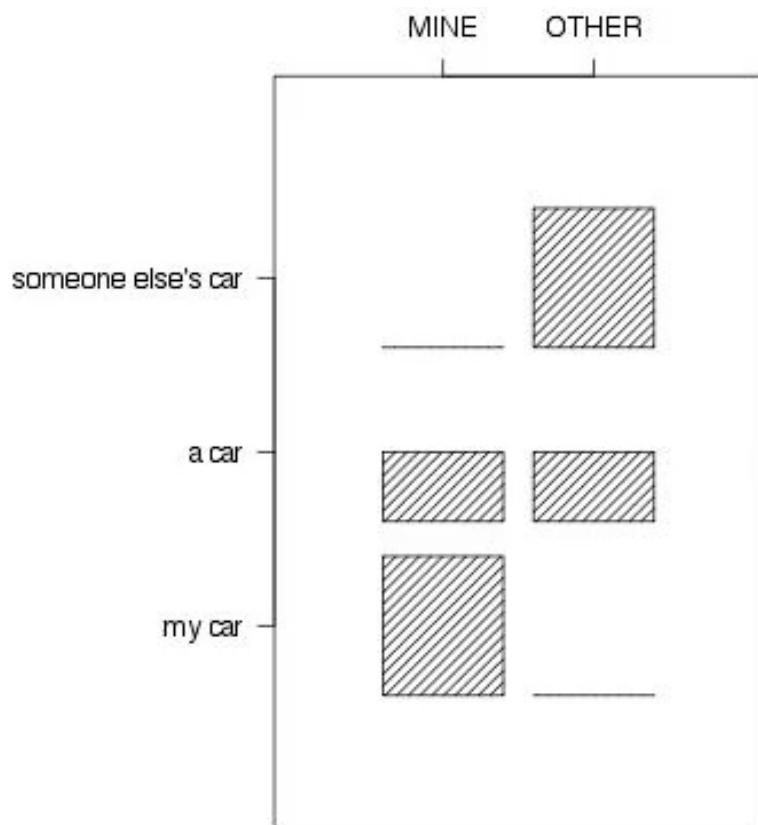
Q/I tradeoff in rational speech-act theory

I injured a child→it WASN'T my child

$$c(\text{my}) = c(a) = 0$$

$$c(\text{someone else's}) = 1$$

$$P(\text{MINE}) = \frac{1}{2}$$



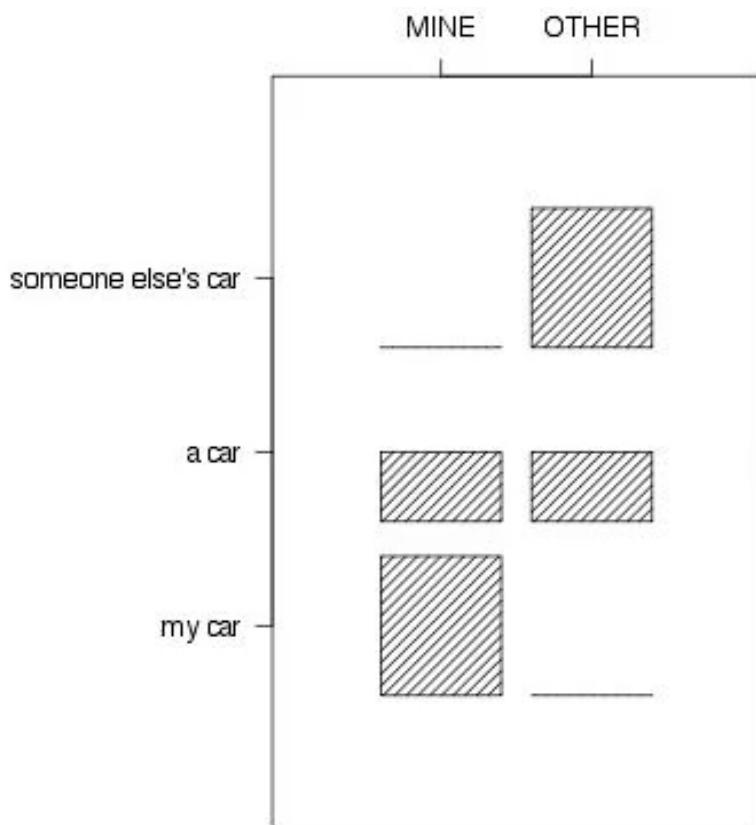
Q/I tradeoff in rational speech-act theory

I injured a child→it WASN'T my child I broke a finger→it WAS my finger

$$c(\text{my}) = c(a) = 0$$

$$c(\text{someone else's}) = 1$$

$$P(\text{MINE}) = \frac{1}{2}$$



Q/I tradeoff in rational speech-act theory

I injured a child→it WASN'T my child I broke a finger→it WAS my finger

$$c(\text{my}) = c(a) = 0$$

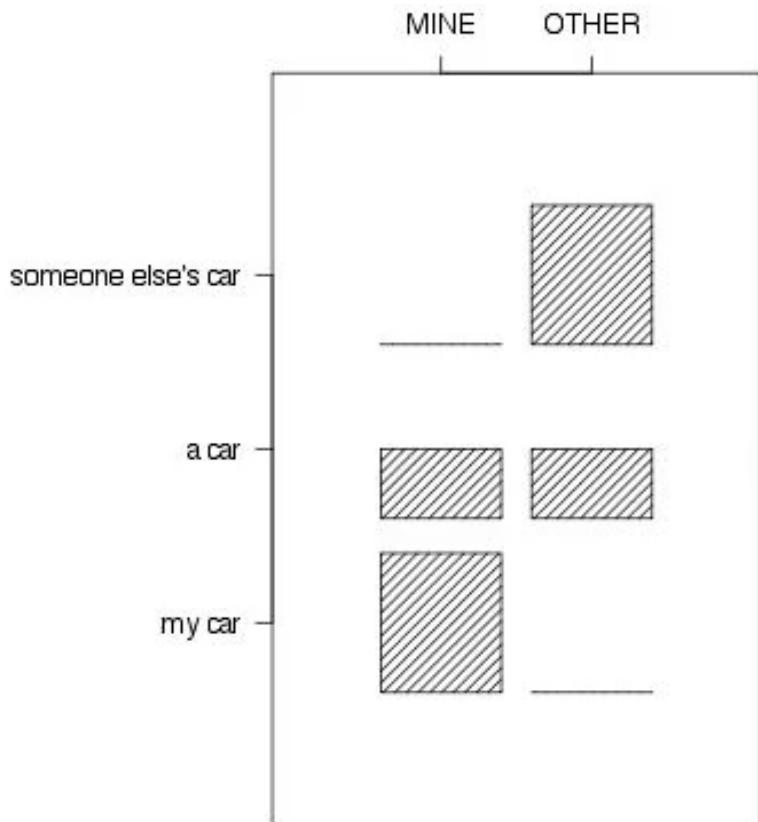
$$c(\text{someone else's}) = 1$$

$$P(\text{MINE}) = \frac{1}{2}$$

$$c(\text{my}) = c(a) = 0$$

$$c(\text{someone else's}) = 1$$

$$P(\text{MINE}) = \frac{5}{6}$$



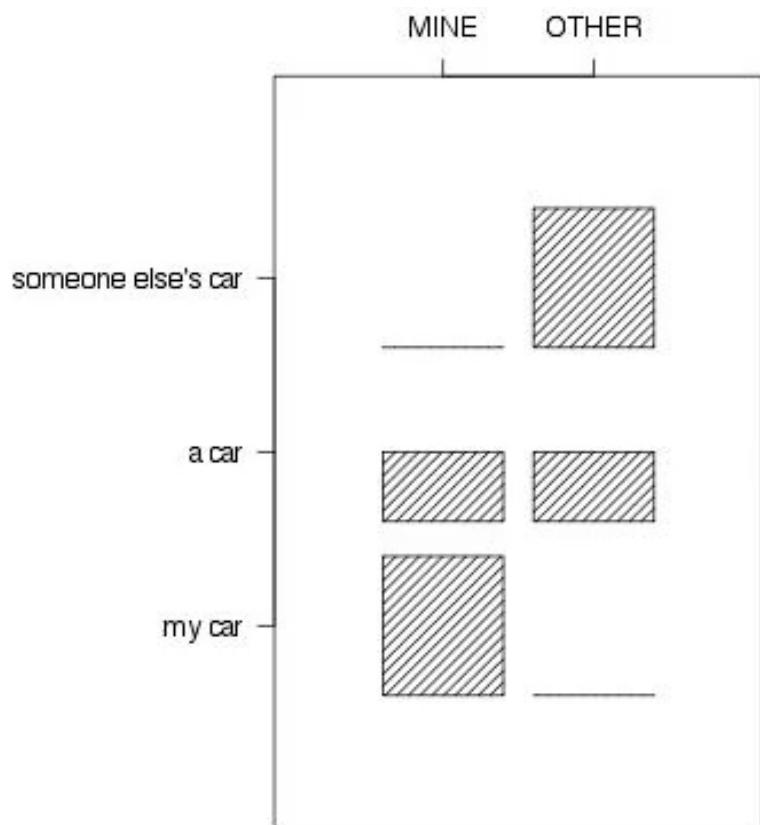
Q/I tradeoff in rational speech-act theory

I injured a child→it WASN'T my child I broke a finger→it WAS my finger

$$c(\text{my}) = c(a) = 0$$

$$c(\text{someone else's}) = 1$$

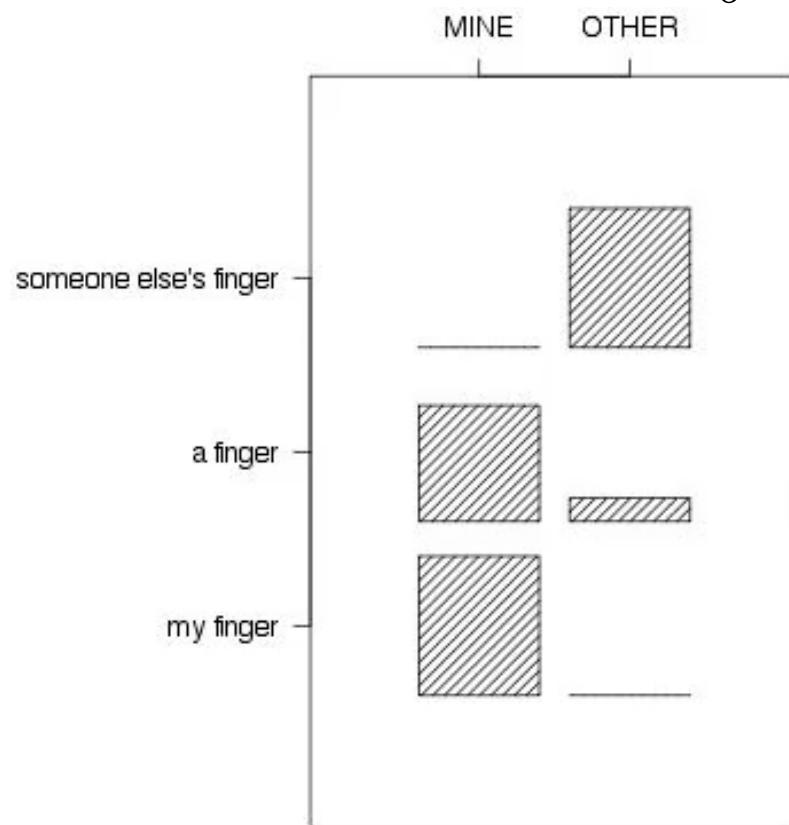
$$P(\text{MINE}) = \frac{1}{2}$$



$$c(\text{my}) = c(a) = 0$$

$$c(\text{someone else's}) = 1$$

$$P(\text{MINE}) = \frac{5}{6}$$



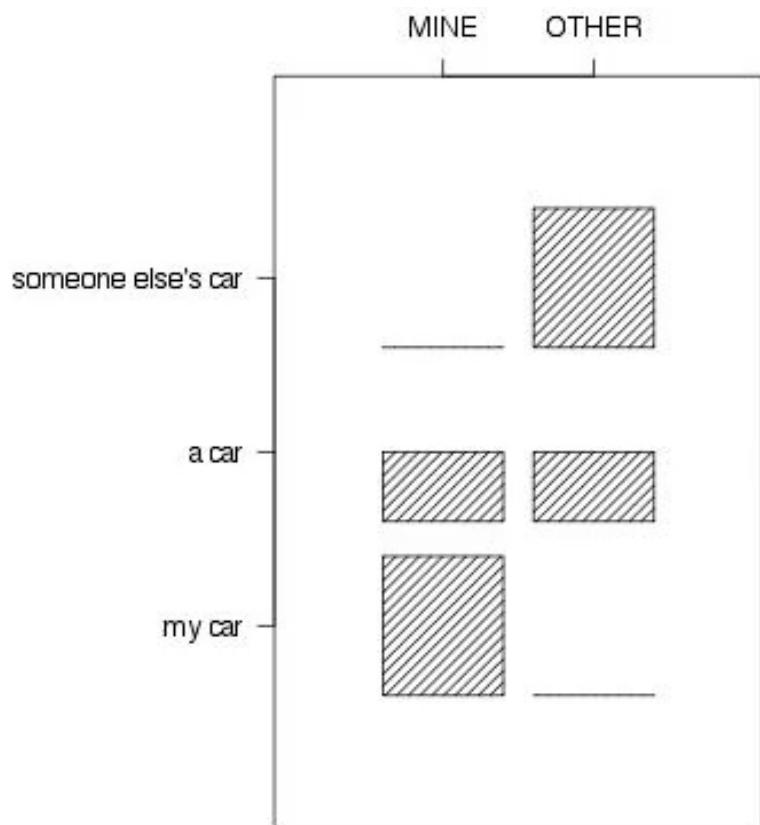
Q/I tradeoff in rational speech-act theory

I injured a child→it WASN'T my child I broke a finger→it WAS my finger

$$c(\text{my}) = c(a) = 0$$

$$c(\text{someone else's}) = 1$$

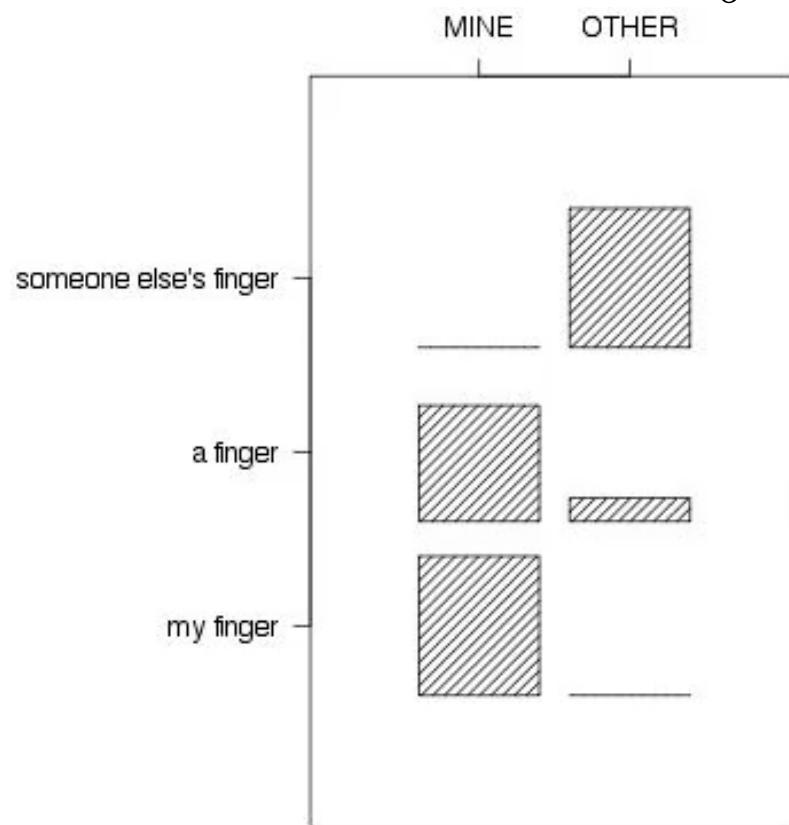
$$P(\text{MINE}) = \frac{1}{2}$$



$$c(\text{my}) = c(a) = 0$$

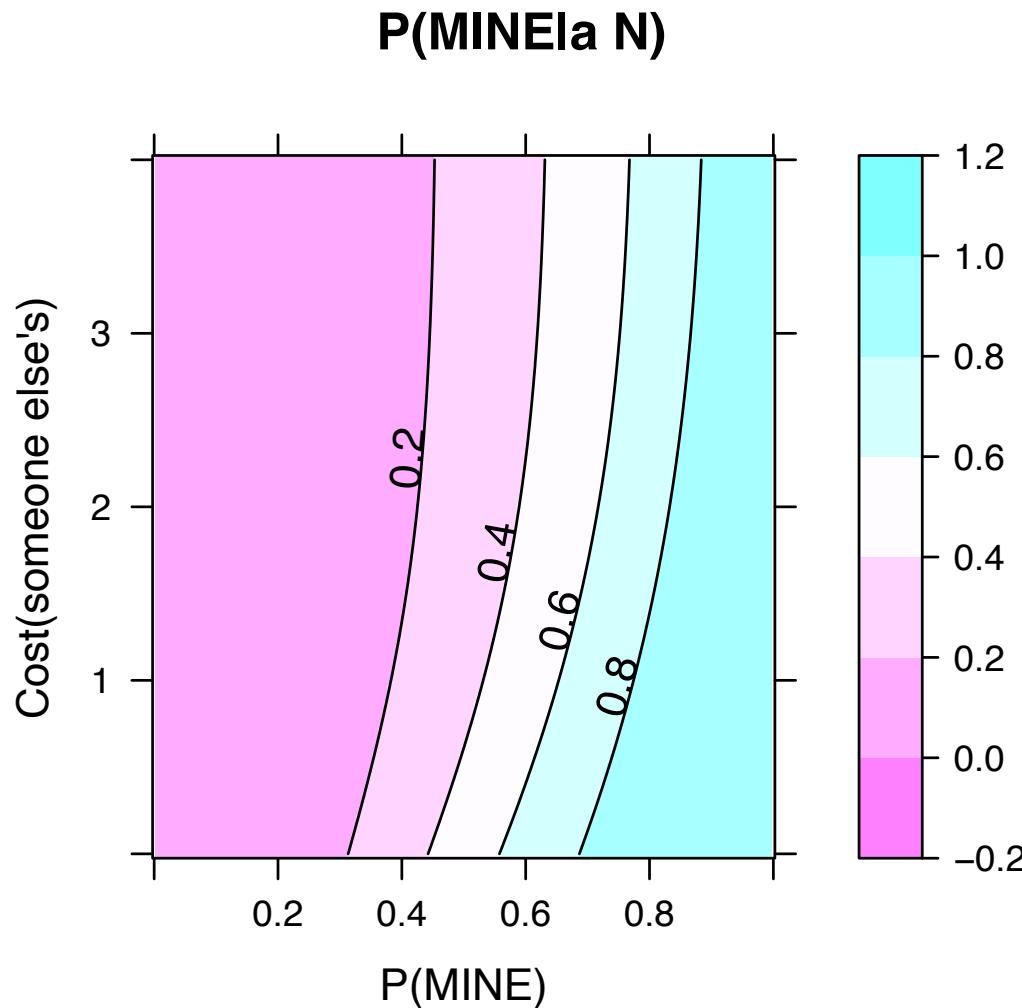
$$c(\text{someone else's}) = 1$$

$$P(\text{MINE}) = \frac{5}{6}$$



Q/I tradeoff in rational speech-act theory

- Prior probability and simplicity trade off against one another
- But they aren't symmetric!



A rich testbed for exploring Q/I tradeoff

The man injured a finger The man injured a child

A rich testbed for exploring Q/I tradeoff

The man injured a finger The man injured a child

His?

Someone else's?

A rich testbed for exploring Q/I tradeoff

The man injured a finger

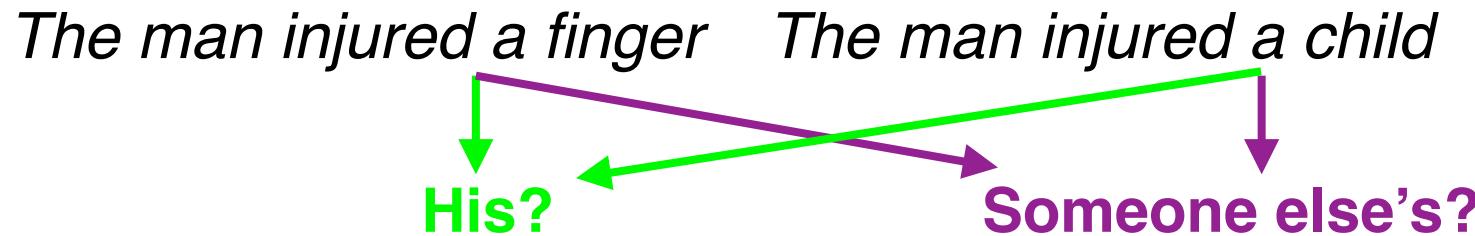
His?

The man injured a child

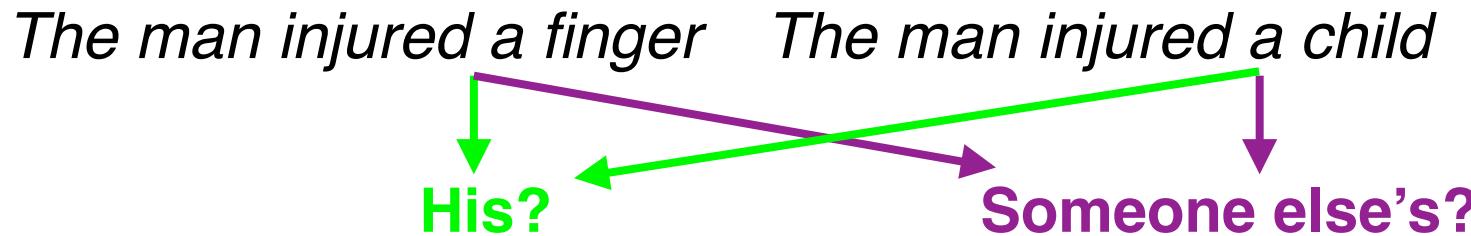
Someone else's?



A rich testbed for exploring Q/I tradeoff

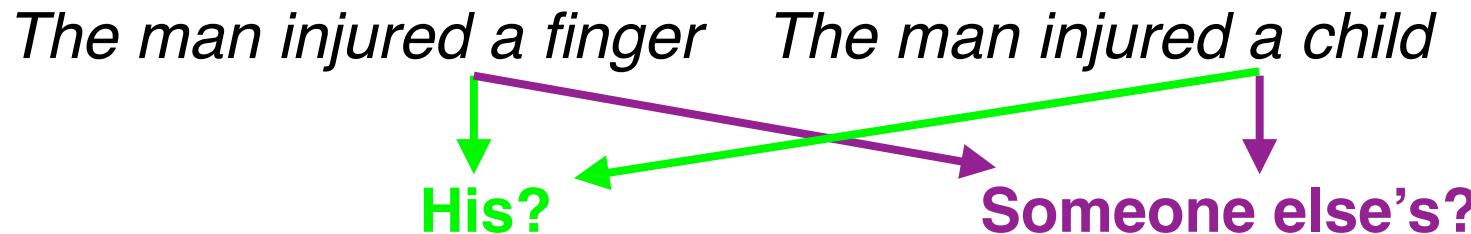


A rich testbed for exploring Q/I tradeoff



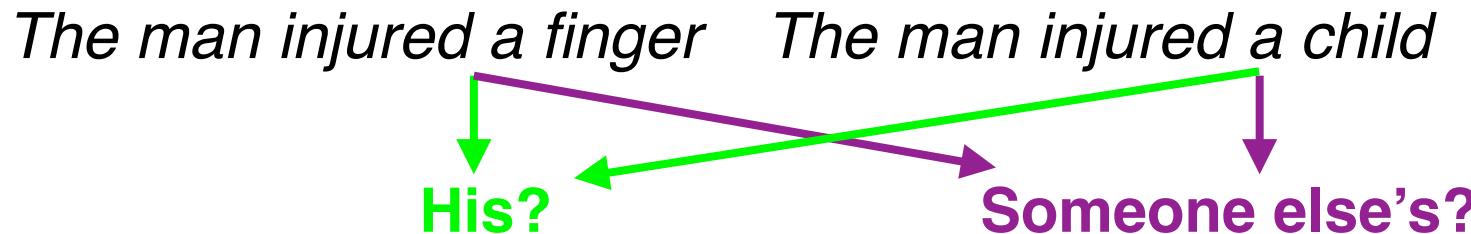
- Five predictions from the rational speech-act model:

A rich testbed for exploring Q/I tradeoff



- Five predictions from the rational speech-act model:
 1. Judgments should track prior event probabilities

A rich testbed for exploring Q/I tradeoff

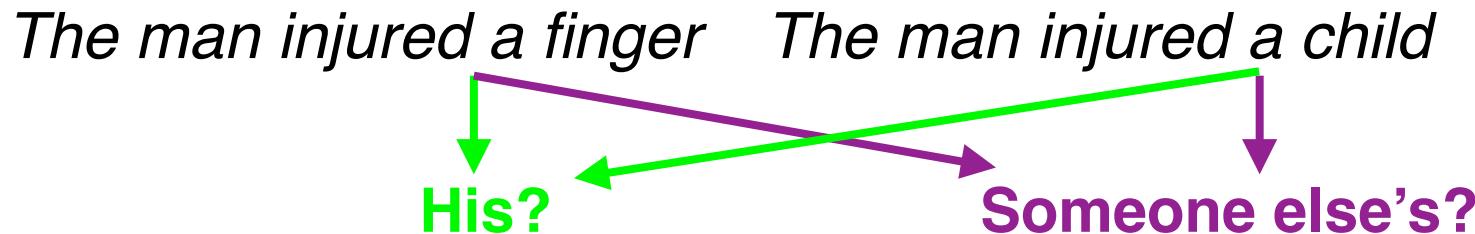


- Five predictions from the rational speech-act model:
 1. Judgments should track prior event probabilities

The man broke a nose

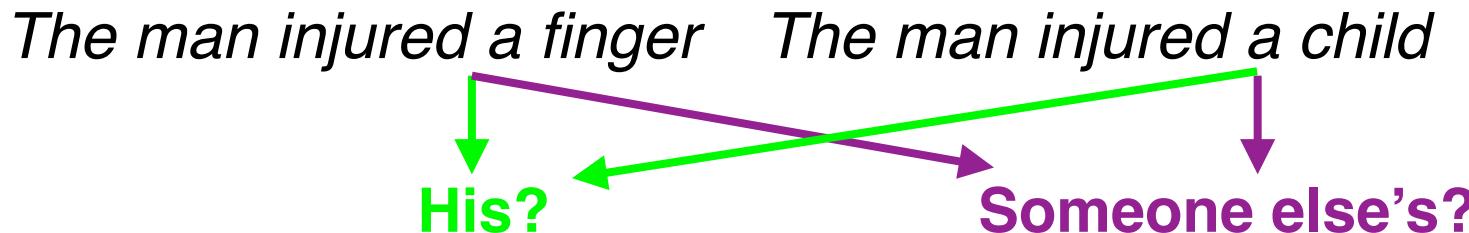
The python broke a nose

A rich testbed for exploring Q/I tradeoff



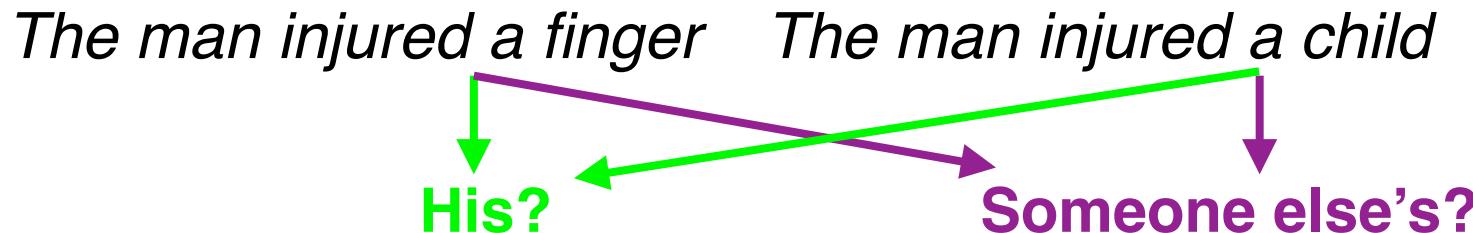
- Five predictions from the rational speech-act model:
 1. Judgments should track prior event probabilities
The man broke a nose *The python broke a nose*
 2. Judgments should be other skewed relative to prior

A rich testbed for exploring Q/I tradeoff



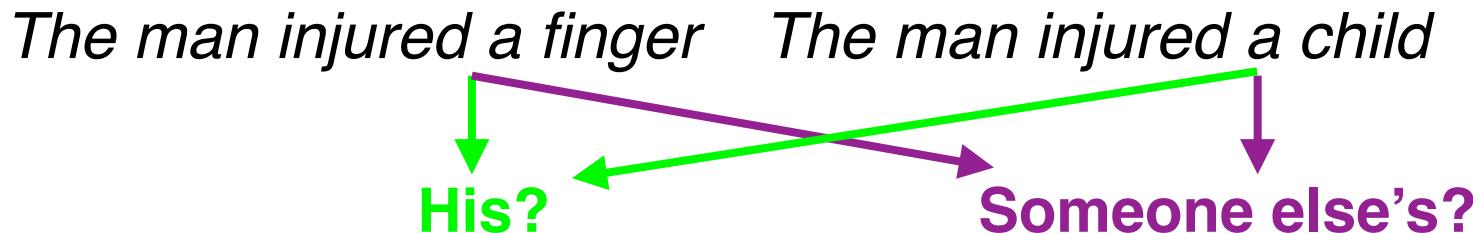
- Five predictions from the rational speech-act model:
 1. Judgments should track prior event probabilities
The man broke a nose *The python broke a nose*
 2. Judgments should be *other* skewed relative to prior
 3. Relational nouns should favor *own* judgments

A rich testbed for exploring Q/I tradeoff



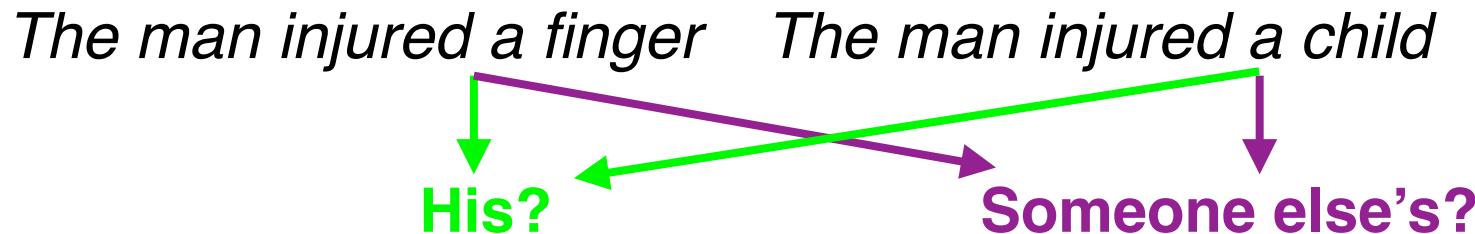
- Five predictions from the rational speech-act model:
 1. Judgments should track prior event probabilities
The man broke a nose *The python broke a nose*
 2. Judgments should be other skewed relative to prior
 3. Relational nouns should favor own judgments
The man injured a child *The father injured a child*

A rich testbed for exploring Q/I tradeoff



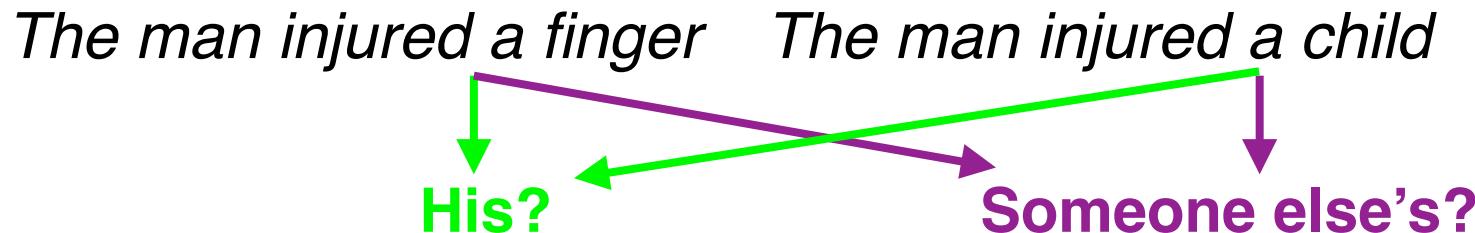
- Five predictions from the rational speech-act model:
 1. Judgments should track prior event probabilities
The man broke a nose *The python broke a nose*
 2. Judgments should be *other* skewed relative to prior
 3. Relational nouns should favor *own* judgments
The man injured a child *The father injured a child*
 4. “Only-one-of” nouns should favor *other* judgments

A rich testbed for exploring Q/I tradeoff



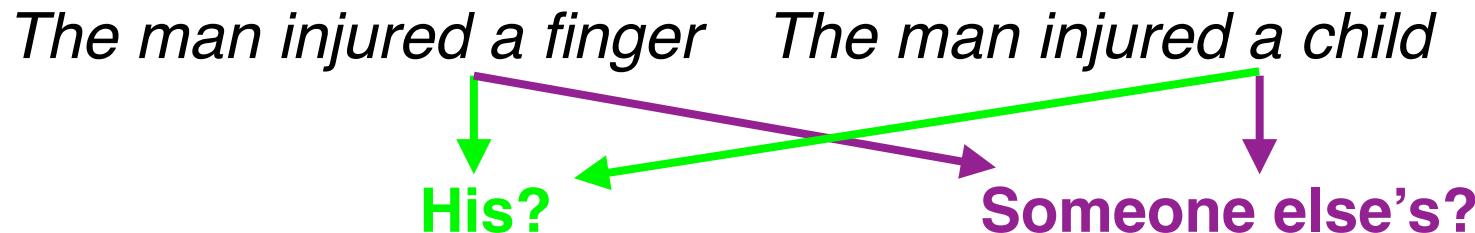
- Five predictions from the rational speech-act model:
 1. Judgments should track prior event probabilities
The man broke a nose *The python broke a nose*
 2. Judgments should be *other* skewed relative to prior
 3. Relational nouns should favor *own* judgments
The man injured a child *The father injured a child*
 4. “Only-one-of” nouns should favor *other* judgments
The man broke a finger *The man broke a nose*

A rich testbed for exploring Q/I tradeoff



- Five predictions from the rational speech-act model:
 1. Judgments should track prior event probabilities
The man broke a nose *The python broke a nose*
 2. Judgments should be *other* skewed relative to prior
 3. Relational nouns should favor *own* judgments
The man injured a child *The father injured a child*
 4. “Only-one-of” nouns should favor *other* judgments
The man broke a finger *The man broke a nose*
 5. Allowing null determiners should favor *own* judgments

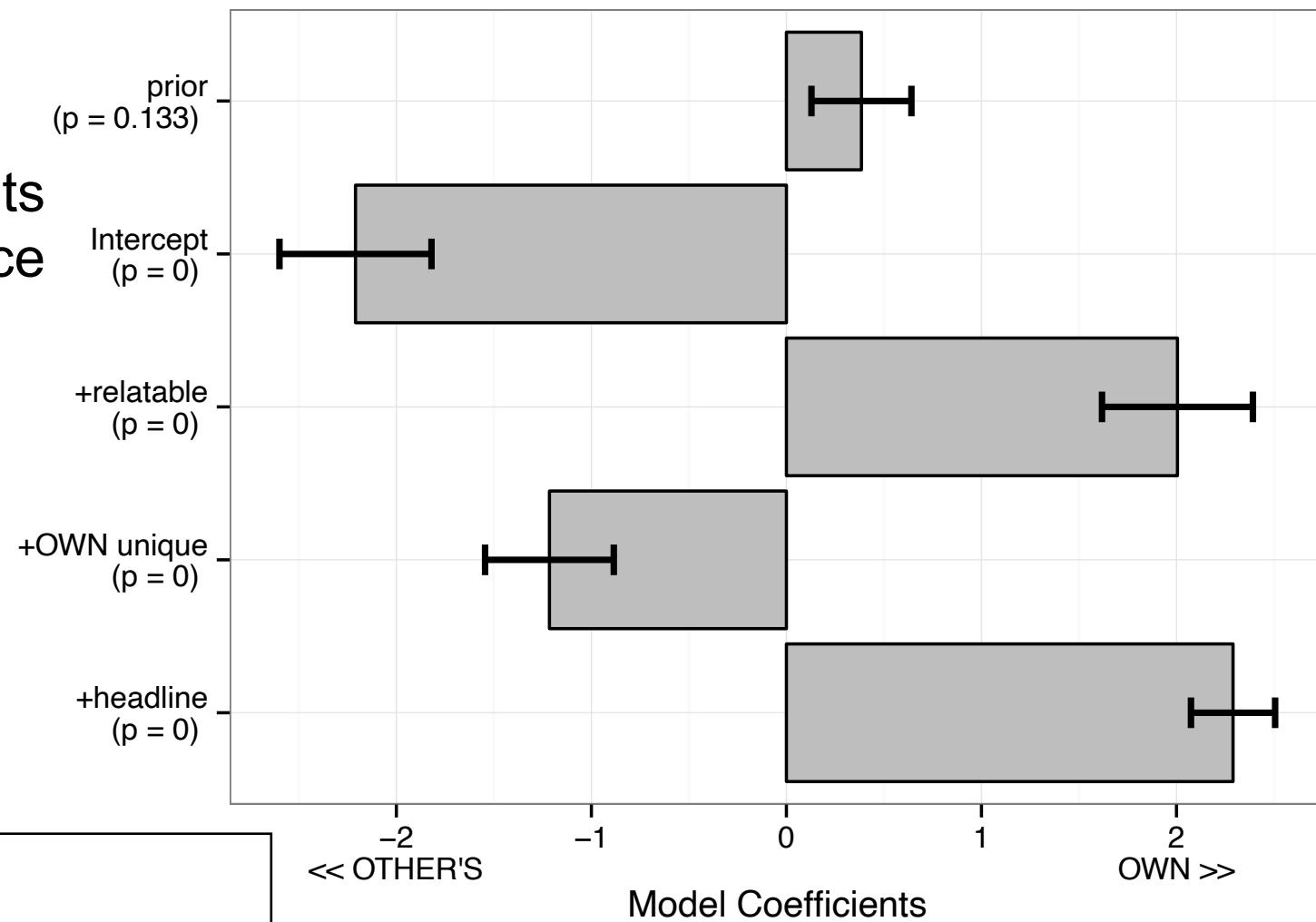
A rich testbed for exploring Q/I tradeoff



- Five predictions from the rational speech-act model:
 1. Judgments should track prior event probabilities
The man broke a nose *The python broke a nose*
 2. Judgments should be *other* skewed relative to prior
 3. Relational nouns should favor *own* judgments
The man injured a child *The father injured a child*
 4. “Only-one-of” nouns should favor *other* judgments
The man broke a finger *The man broke a nose*
 5. Allowing null determiners should favor *own* judgments
The man injured a child *Man injured child*

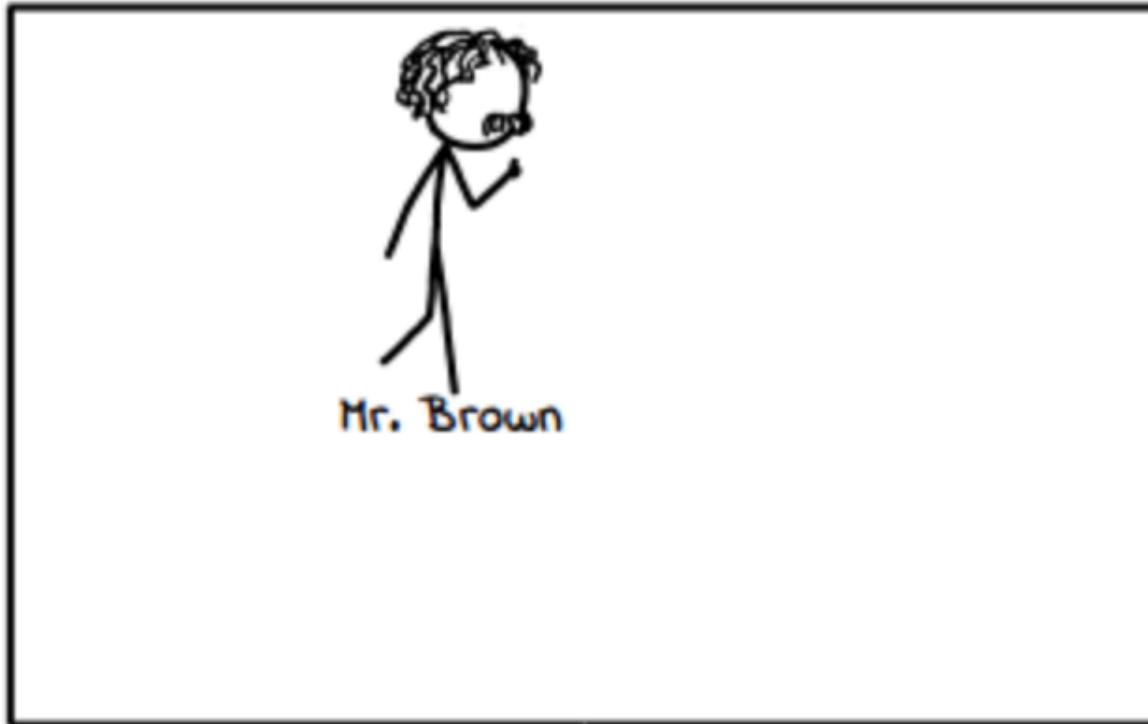
A rich testbed for exploring Q/I tradeoff

- 1348 judgments of 108 sentence prompts
- Multivariate mixed-effects logistic regression analysis



1. Effect of prior
2. Overall *other* skew
3. Relational nouns favor *own*
4. “Only-one-of” nouns favor *other*
5. Null determiners favor *own*

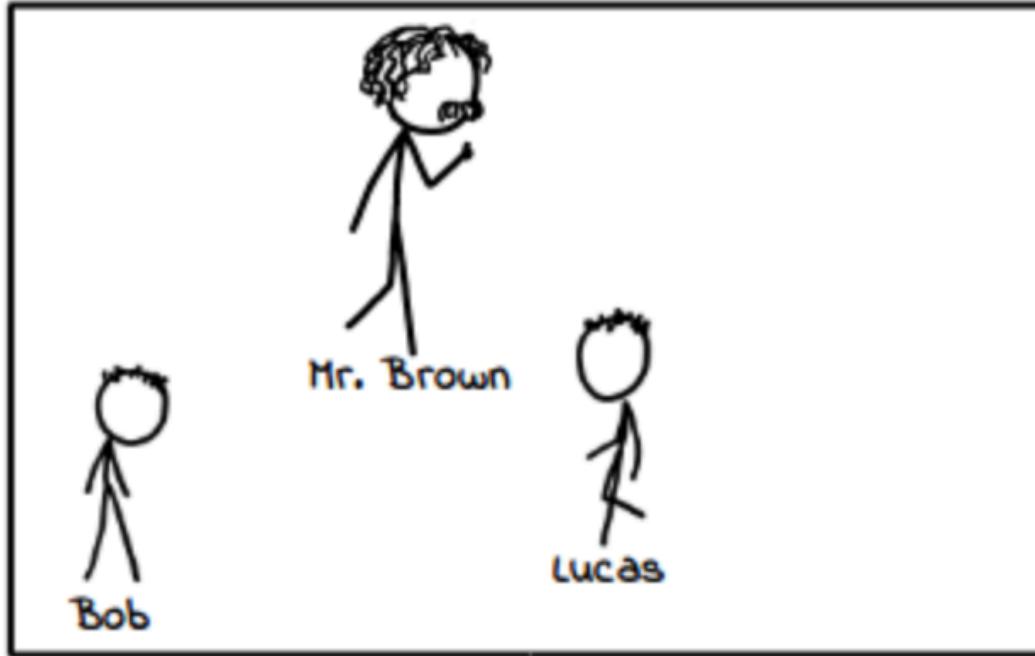
A controlled Q/I experiment



"Mr. Brown is the principal at an element##### school."

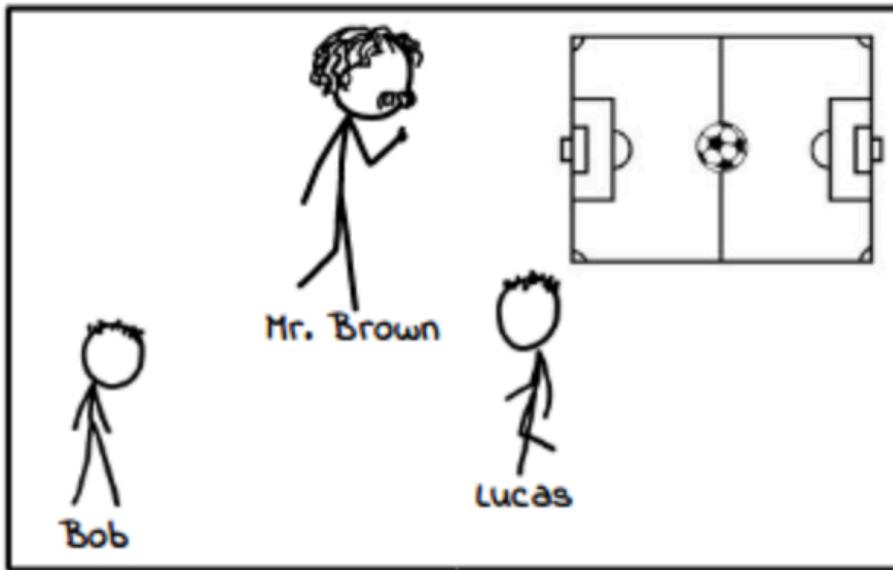
('#####' indicates phone static.)

A controlled Q/I experiment



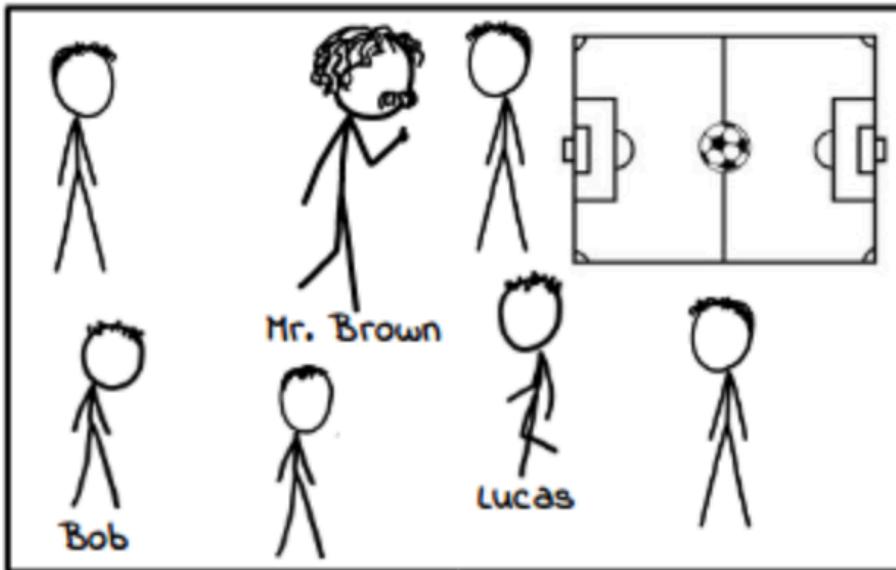
"He has two sons called Bob and Lucas who are 8 and 9 ye#####old."

A controlled Q/I experiment



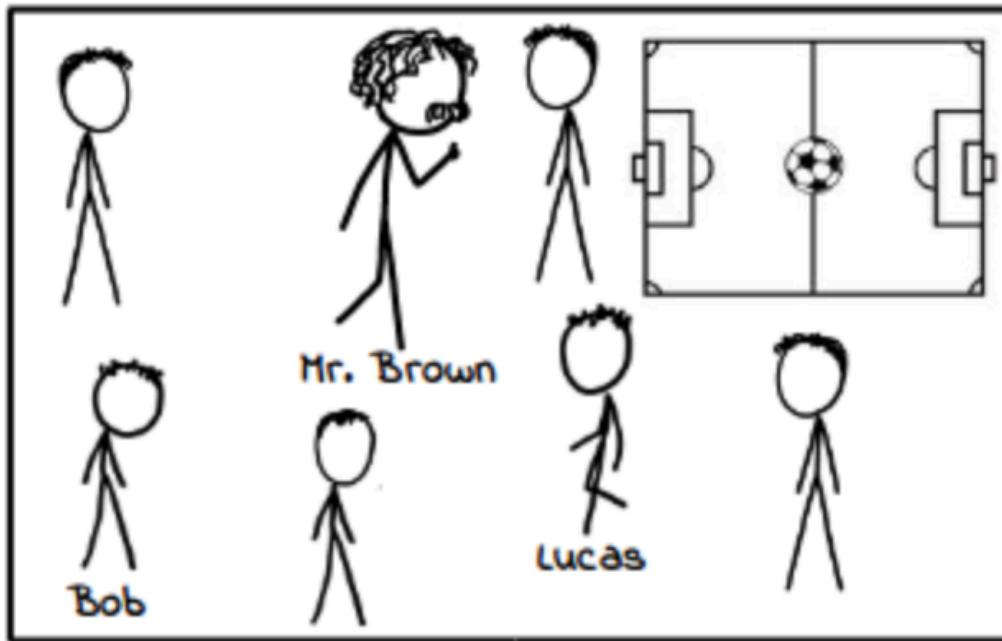
"Bob and Lucas are Mr. Brown's only children and they participated in the school's soccer tourna##### last weekend."

A controlled Q/I experiment



"After the #####nament, Mr. Brown played tag with Bob, Lucas, and 4 other boys when the accident happened:"

A controlled Q/I experiment



"**Brown knocks over#####son!**"

What do you think the speaker is trying to convey with this sentence?

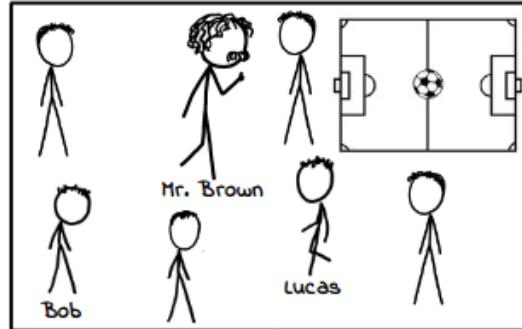
- That Mr. Brown knocked over **someone else's child**.
- That Mr. Brown knocked over **his own child**.

What do you think the speaker is trying to convey with this sentence?

- That Mr. Brown knocked over **someone else's child**.
- That Mr. Brown knocked over **his own child**.

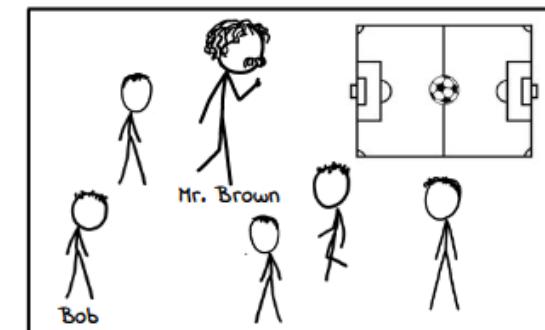
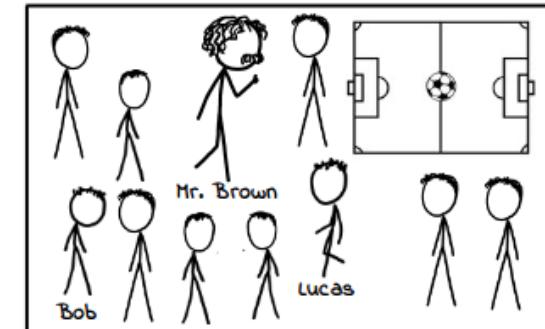
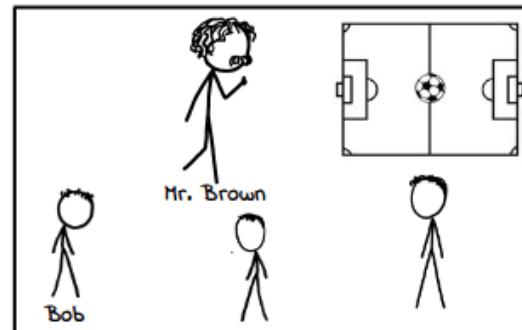
Prior P(own) HIGH

|sons of Brown|=2



Prior P(own) LOW

|sons of Brown|=1

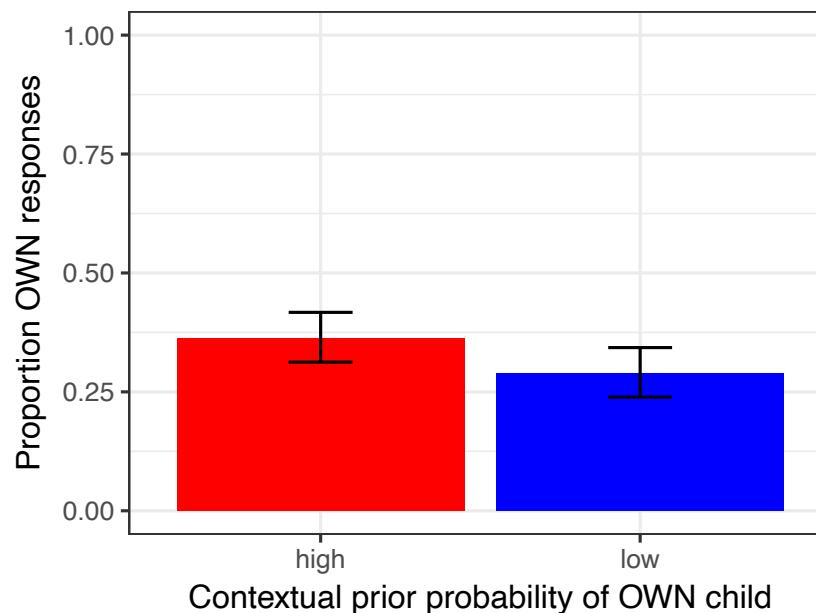
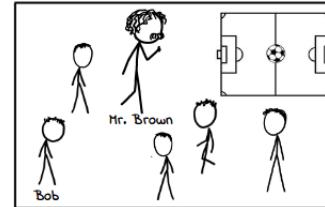
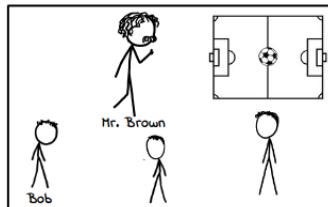
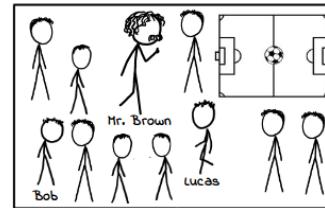
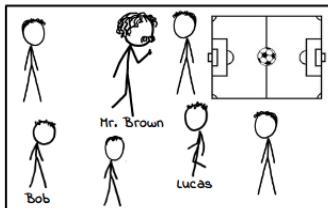


Utterance condition	unrelatable noun (boy)	relatable noun (son)
(prior elicitation)	<i>Brown knocks over#####boy!</i>	<i>Brown knocks over#####son!</i>
Posterior elicitation	<i>Brown knocks over a boy!</i>	<i>Brown knocks over a son!</i>
"Headline" interpretation	<i>Brown knocks over boy!</i>	<i>Brown knocks over son!</i>

Checking predicted effects

- Effect of contextual prior probability of OWN child

Prior P(own) HIGH Prior P(own) LOW

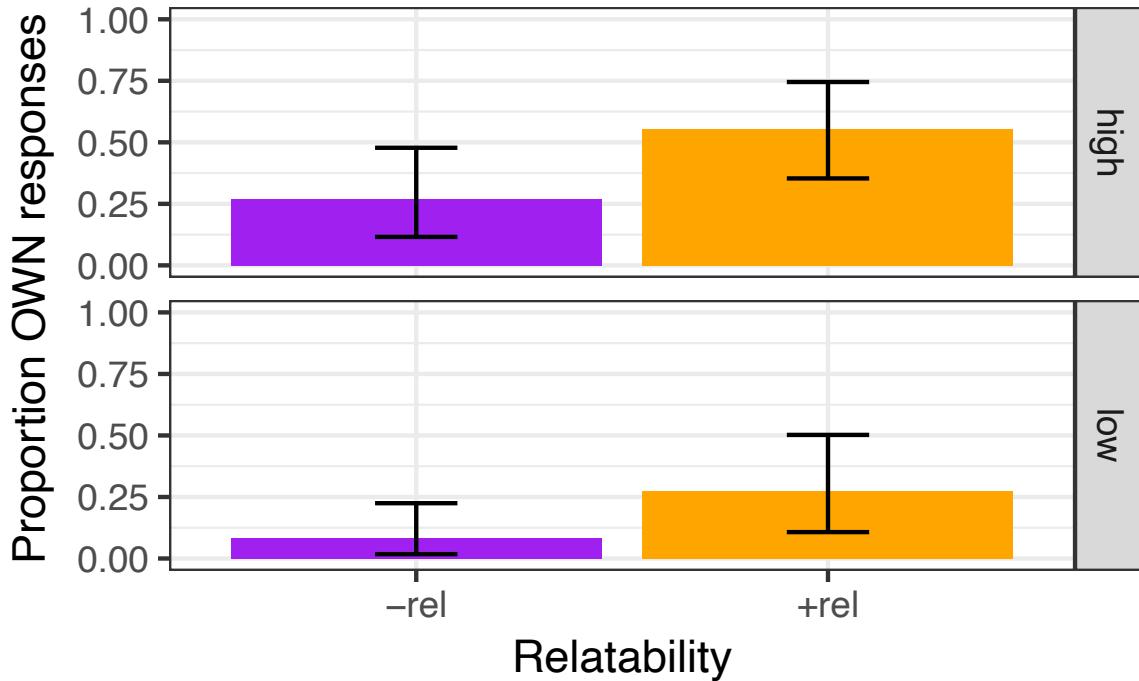


Checking predicted effects

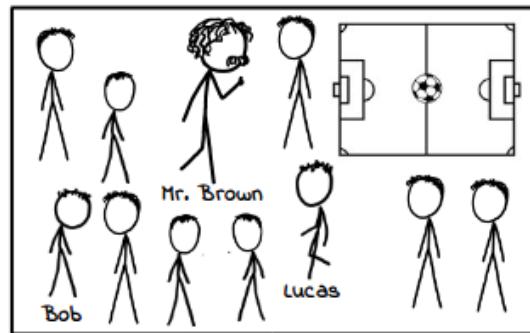
- The relatability effect

unrelatable noun (<i>boy</i>)	relatable noun (<i>son</i>)
<i>Brown knocks over a boy!</i>	<i>Brown knocks over a son!</i>

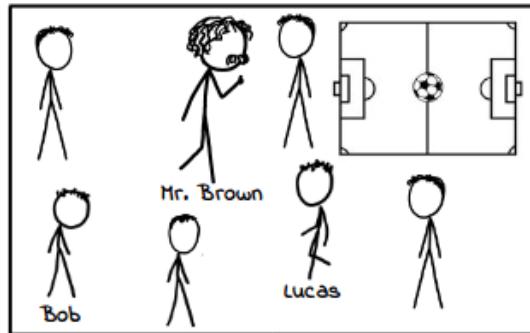
Relatability:  -rel  +rel



Prior P(own) HIGH



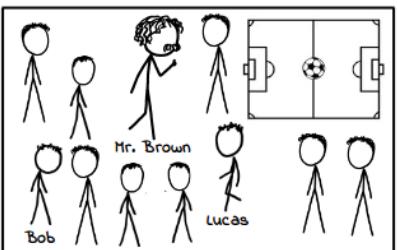
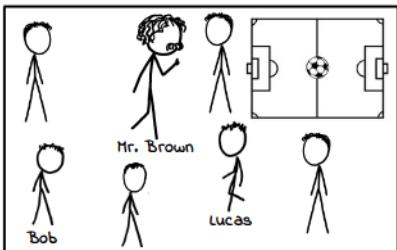
Prior P(own) LOW



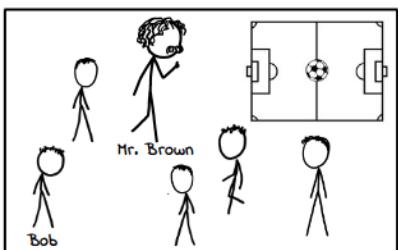
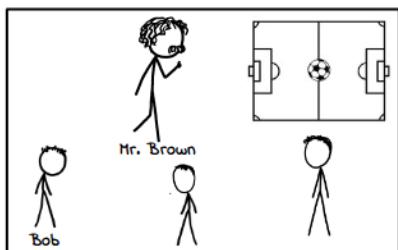
Checking predicted effects

- The effect of contextual uniqueness

$|\text{sons of Brown}|=2$

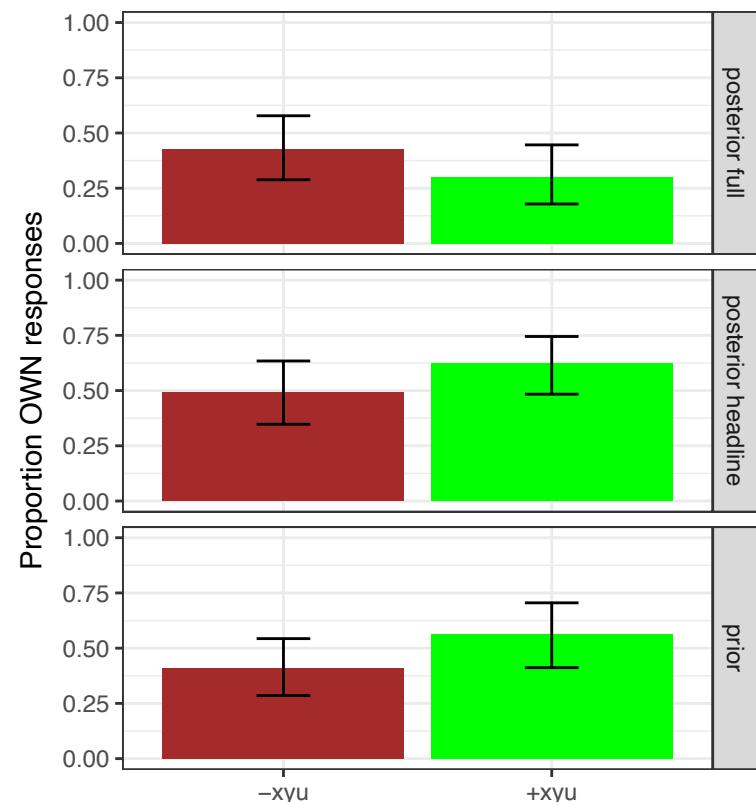


$|\text{sons of Brown}|=1$



Posterior	Brown knocks over a son!
"Headline"	Brown knocks over son!
Prior	Brown knocks over#####son!

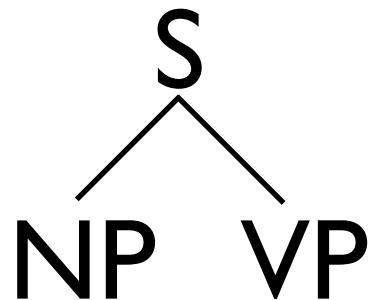
X-Y relationship contextually unique? -xyu +xyu



A little more about logical semantics

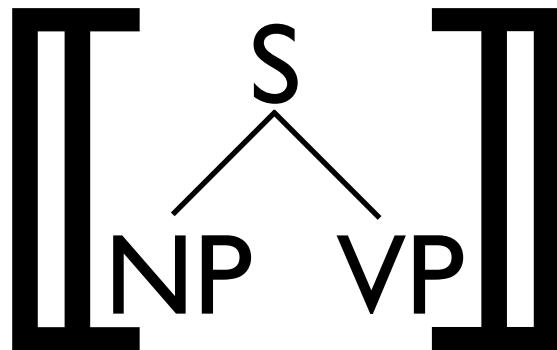
- **Logical** approaches to semantics influential in linguistics, psychology & AI
 - Linguistic expressions have **semantic types**
 - **Nouns**: sets (equiv: functions from entities to truth values, or $e \rightarrow t$)
 - **NPs** (oversimplified): individuals e
 - **Intransitive verbs**: sets, or $e \rightarrow t$ functions
 - Syntactic rules have corresponding semantic rules too
 - e.g., set membership check for $S \rightarrow NP\ VP$

Super-brief introduction to logical semantics



Super-brief introduction to logical semantics

*denotation
function*



Super-brief introduction to logical semantics

*denotation
function*

$$\boxed{\begin{array}{c} S \\ \diagdown \quad \diagup \\ NP \quad VP \end{array}} = \llbracket VP \rrbracket(\llbracket NP \rrbracket)$$

Super-brief introduction to logical semantics

*denotation
function*

$$\boxed{\begin{array}{c} S \\ \diagdown \quad \diagup \\ NP \quad VP \end{array}} = \llbracket VP \rrbracket(\llbracket NP \rrbracket)$$

$$\llbracket \text{Mary sings} \rrbracket = \llbracket \text{sings} \rrbracket(\llbracket \text{Mary} \rrbracket)$$

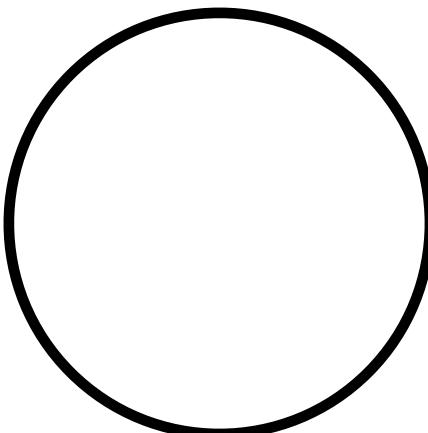
Super-brief introduction to logical semantics

*denotation
function*

$$\boxed{\begin{array}{c} S \\ \diagdown \quad \diagup \\ NP \quad VP \end{array}} = \llbracket VP \rrbracket(\llbracket NP \rrbracket)$$

$$\llbracket \text{Mary sings} \rrbracket = \llbracket \text{sings} \rrbracket(\llbracket \text{Mary} \rrbracket)$$

set of entities
that sing



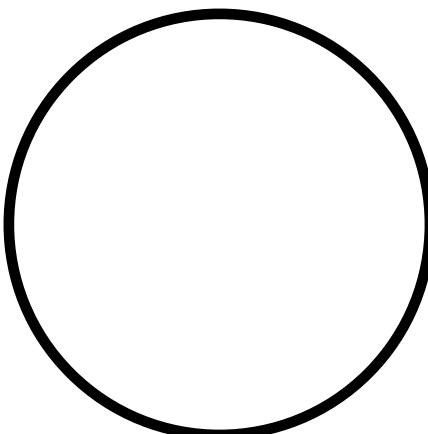
Super-brief introduction to logical semantics

*denotation
function*

$$\boxed{\boxed{\text{NP} \quad \text{VP}}} = \llbracket \text{VP} \rrbracket(\llbracket \text{NP} \rrbracket)$$

$$\llbracket \text{Mary sings} \rrbracket = \llbracket \text{sings} \rrbracket(\llbracket \text{Mary} \rrbracket)$$

set of entities
that sing



Mary

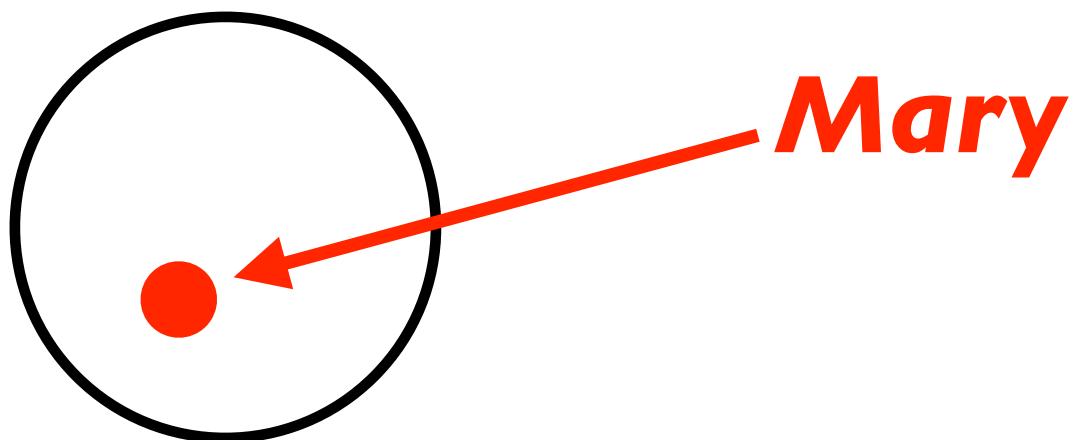
Super-brief introduction to logical semantics

*denotation
function*

$$\boxed{\boxed{\text{NP} \quad \text{VP}}} = \llbracket \text{VP} \rrbracket(\llbracket \text{NP} \rrbracket)$$

$$\llbracket \text{Mary sings} \rrbracket = \llbracket \text{sings} \rrbracket(\llbracket \text{Mary} \rrbracket)$$

set of entities
that sing



Background: logical semantics

Warning!

- The previous slide is a serious oversimplification of NP semantics & composition with VP into S, e.g.:
 - *Dogs growl* (dogs is a generic)
 - *Every student studied hard* (quantifiers every need different treatment)
- But, this example hopefully gives a sense of the context in which set-based semantics is put to use

Adjectives: a range of semantic types

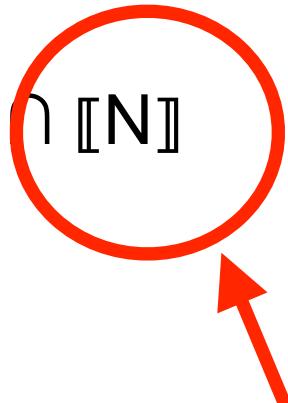
- **Intersective:** *living, blue*
- **Scalar:**
 - **Relative:** *short, expensive*
 - **Absolute:** *dangerous, full*
- **Non-intersective:** *possible, alleged*
- **Anti-intersective:** *former, counterfeit*

Intersective adjectives

- Examples: living, blue
- The criterion imposed by the adjective is independent of the noun
- Formally, $[\text{Adj N}] = [\text{Adj}] \cap [\text{N}]$

Intersective adjectives

- Examples: living, blue
- The criterion imposed by the adjective is independent of the noun
- Formally, $\llbracket \text{Adj N} \rrbracket = \llbracket \text{Adj} \rrbracket \cap \llbracket \text{N} \rrbracket$



$\llbracket X \rrbracket$: “***meaning
of X***”

Scalar adjectives

Scalar adjectives

- Examples: *short, expensive, dangerous, full*

Scalar adjectives

- Examples: *short, expensive, dangerous, full*
- Constrain referent's value on a **scale**

Scalar adjectives

- Examples: *short, expensive, dangerous, full*
- Constrain referent's value on a **scale**
- **Relative adjs** (*short, expensive*): scale constraint based on comparison class

Scalar adjectives

- Examples: *short, expensive, dangerous, full*
- Constrain referent's value on a **scale**
- **Relative adjs** (*short, expensive*): scale constraint based on comparison class
 - Highly context-sensitive; compare:

Scalar adjectives

- Examples: *short, expensive, dangerous, full*
- Constrain referent's value on a **scale**
- **Relative adjs** (*short, expensive*): scale constraint based on comparison class
 - Highly context-sensitive; compare:
Bill is a big man!

Scalar adjectives

- Examples: *short, expensive, dangerous, full*
- Constrain referent's value on a **scale**
- **Relative adjs** (*short, expensive*): scale constraint based on comparison class
 - Highly context-sensitive; compare:
Bill is a big man!
Bill is a big mouse!

Scalar adjectives

- Examples: *short, expensive, dangerous, full*
- Constrain referent's value on a **scale**
- **Relative adjs** (*short, expensive*): scale constraint based on comparison class
 - Highly context-sensitive; compare:
Bill is a big man!
Bill is a big mouse!
Bill is a big elephant!

Scalar adjectives

- Examples: *short, expensive, dangerous, full*
- Constrain referent's value on a **scale**
- **Relative adjs** (*short, expensive*): scale constraint based on comparison class
 - Highly context-sensitive; compare:

Bill is a big man!

Bill is a big mouse!

Bill is a big elephant!

Bill is a big basketball player!

Scalar adjectives

- Examples: *short, expensive, dangerous, full*
- Constrain referent's value on a **scale**
- **Relative adjs** (*short, expensive*): scale constraint based on comparison class
 - Highly context-sensitive; compare:
Bill is a big man!
Bill is a big mouse!
Bill is a big elephant!
Bill is a big basketball player!
- **Absolute adjs** (*dangerous, full, empty*): constraint is tied to scale boundary

Scalar adjectives

- Examples: *short, expensive, dangerous, full*
- Constrain referent's value on a **scale**
- **Relative adjs** (*short, expensive*): scale constraint based on comparison class
 - Highly context-sensitive; compare:
Bill is a big man!
Bill is a big mouse!
Bill is a big elephant!
Bill is a big basketball player!
- **Absolute adjs** (*dangerous, full, empty*): constraint is tied to scale boundary
 - Less context sensitive?

Scalar adjectives

- Examples: *short, expensive, dangerous, full*
- Constrain referent's value on a **scale**
- **Relative adjs** (*short, expensive*): scale constraint based on comparison class
 - Highly context-sensitive; compare:
Bill is a big man!
Bill is a big mouse!
Bill is a big elephant!
Bill is a big basketball player!
- **Absolute adjs** (*dangerous, full, empty*): constraint is tied to scale boundary
 - Less context sensitive?
The glass is empty.

Scalar adjectives

- Examples: *short, expensive, dangerous, full*
- Constrain referent's value on a **scale**
- **Relative adjs** (*short, expensive*): scale constraint based on comparison class
 - Highly context-sensitive; compare:
Bill is a big man!
Bill is a big mouse!
Bill is a big elephant!
Bill is a big basketball player!
- **Absolute adjs** (*dangerous, full, empty*): constraint is tied to scale boundary
 - Less context sensitive?
The glass is empty.
The gas tank is empty.

Scalar adjectives

- Examples: *short, expensive, dangerous, full*
- Constrain referent's value on a **scale**
- **Relative adjs** (*short, expensive*): scale constraint based on comparison class
 - Highly context-sensitive; compare:
Bill is a big man!
Bill is a big mouse!
Bill is a big elephant!
Bill is a big basketball player!
- **Absolute adjs** (*dangerous, full, empty*): constraint is tied to scale boundary
 - Less context sensitive?
The glass is empty.
The gas tank is empty.
The auditorium is empty.

Non-intersective adjectives

- Examples: *alleged, possible*
- Adj's meaning contribution depends fundamentally on modified noun
- Adj “releases” referent from constraints on the noun

Anti-intersective adjectives

- Examples: *counterfeit, former*
- Adjective's meaning contribution depends fundamentally on modified noun
- Adjective adds commitment that the referent is NOT in noun's denotation

Adjectives: a range of semantic types

- **Intersective:** *living, blue*
- **Scalar:**
 - **Relative:** *short, expensive*
 - **Absolute:** *dangerous, full*
- **Non-intersective:** *possible, alleged*
- **Anti-intersective:** *former, counterfeit*

Adjectives: a range of semantic types

- **Intersective:** *living, blue*
- **Scalar:**
 - **Relative:** *short, expensive*
 - **Absolute:** *dangerous, full*
- **Non-intersective:** *possible, alleged*
- **Anti-intersective:** *former, counterfeit*

Today

Degree semantics for scalar adjectives

Mary is tall

Degree semantics for scalar adjectives

- The meaning of a scalar adjective like *tall* does two things

Mary is tall

Degree semantics for scalar adjectives

- The meaning of a scalar adjective like *tall* does two things
 1. Projects a referent onto some **value** on a **scale**

Mary is tall

Degree semantics for scalar adjectives

- The meaning of a scalar adjective like *tall* does two things
 1. Projects a referent onto some **value** on a **scale**

Mary is tall



Degree semantics for scalar adjectives

- The meaning of a scalar adjective like *tall* does two things
 1. Projects a referent onto some **value** on a **scale**

Mary is tall



Mary

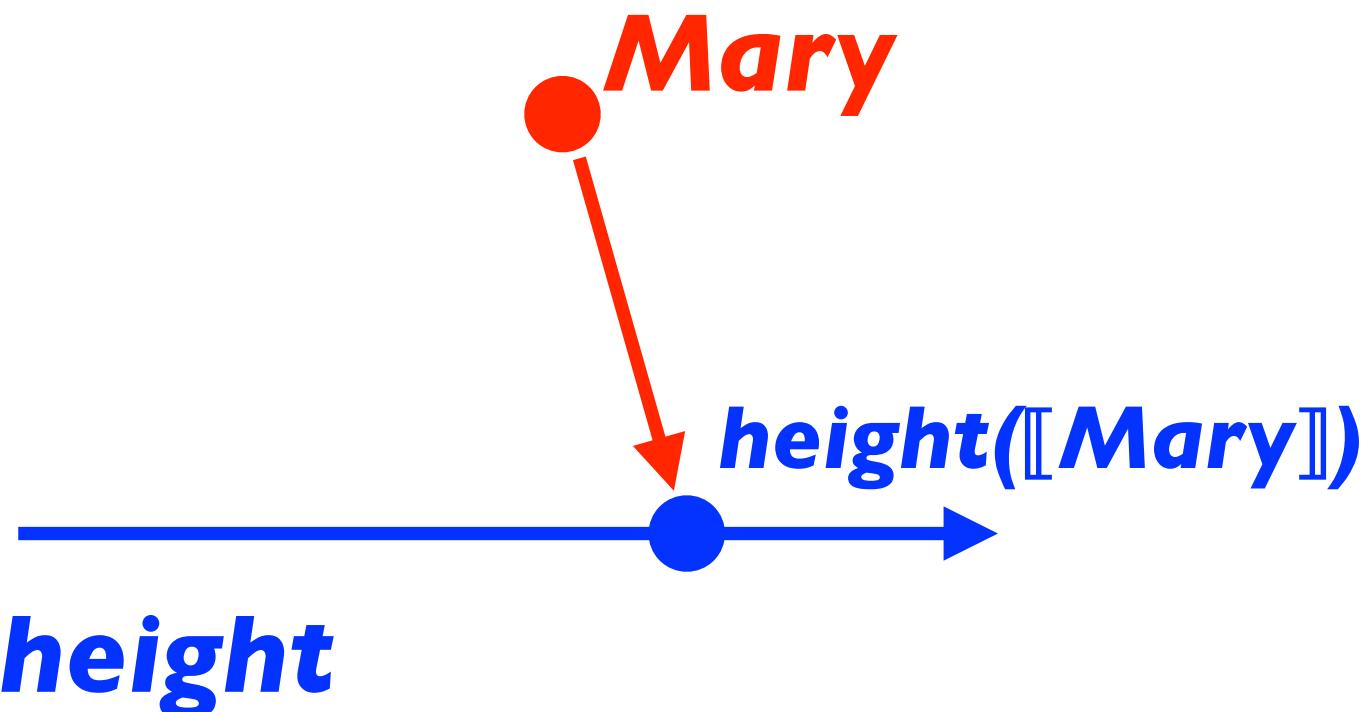


height

Degree semantics for scalar adjectives

- The meaning of a scalar adjective like *tall* does two things
 1. Projects a referent onto some **value** on a **scale**

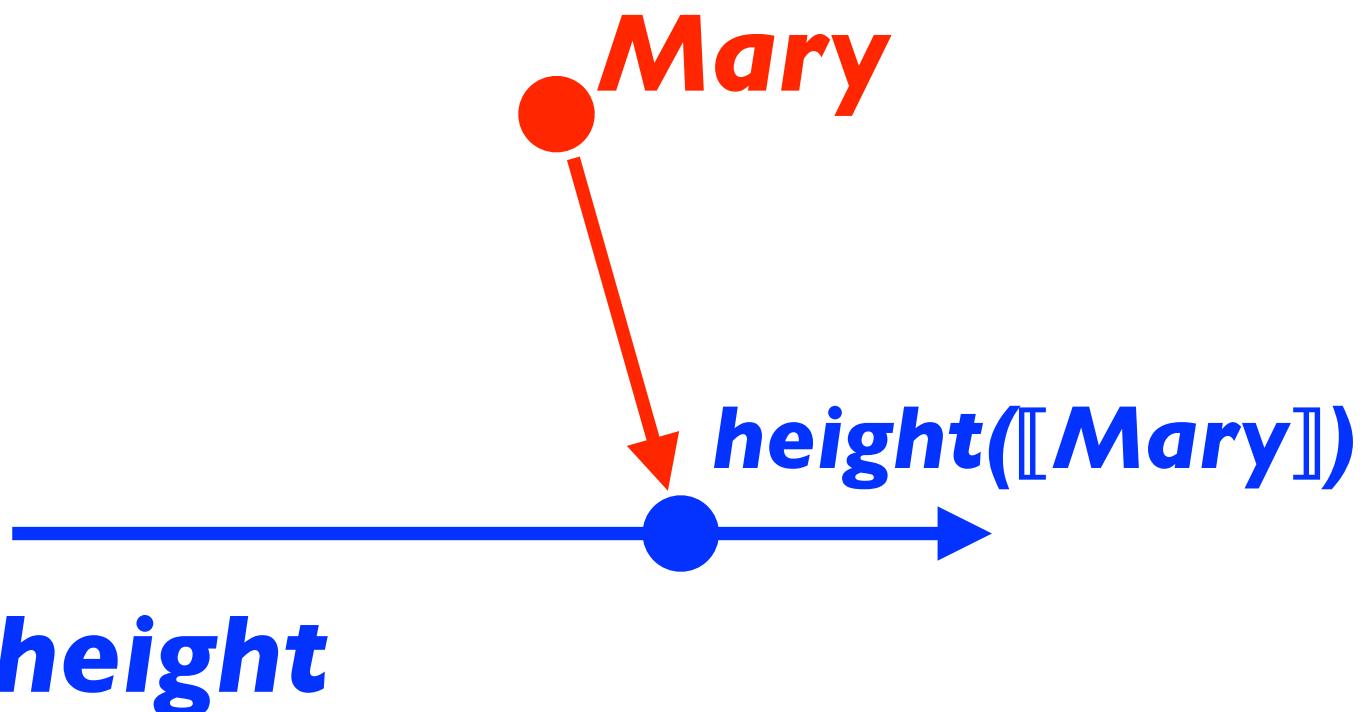
Mary is tall



Degree semantics for scalar adjectives

- The meaning of a scalar adjective like *tall* does two things
 1. Projects a referent onto some **value** on a **scale**
 2. Predicates that that **value** is greater than some **threshold θ**

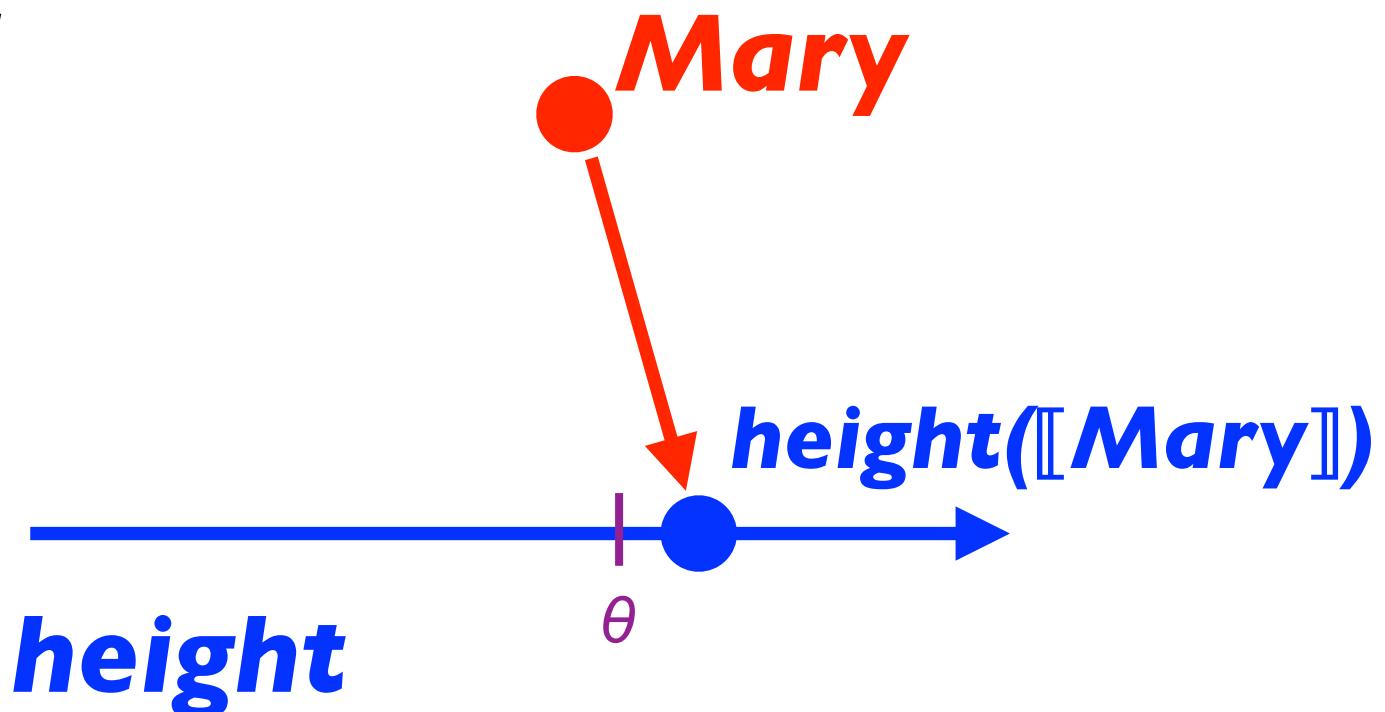
Mary is tall



Degree semantics for scalar adjectives

- The meaning of a scalar adjective like *tall* does two things
 - Projects a referent onto some **value** on a **scale**
 - Predicates that that **value** is greater than some **threshold θ**

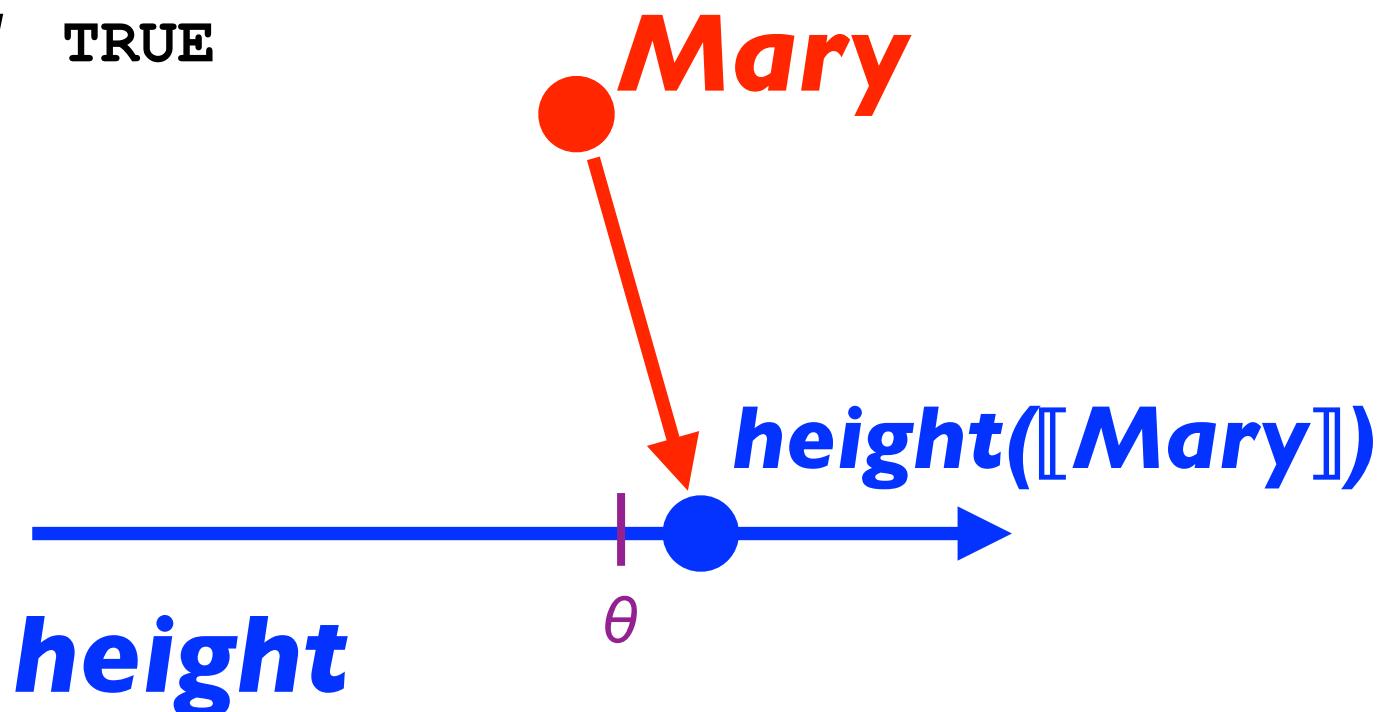
Mary is tall



Degree semantics for scalar adjectives

- The meaning of a scalar adjective like *tall* does two things
 - Projects a referent onto some **value** on a **scale**
 - Predicates that that **value** is greater than some **threshold θ**

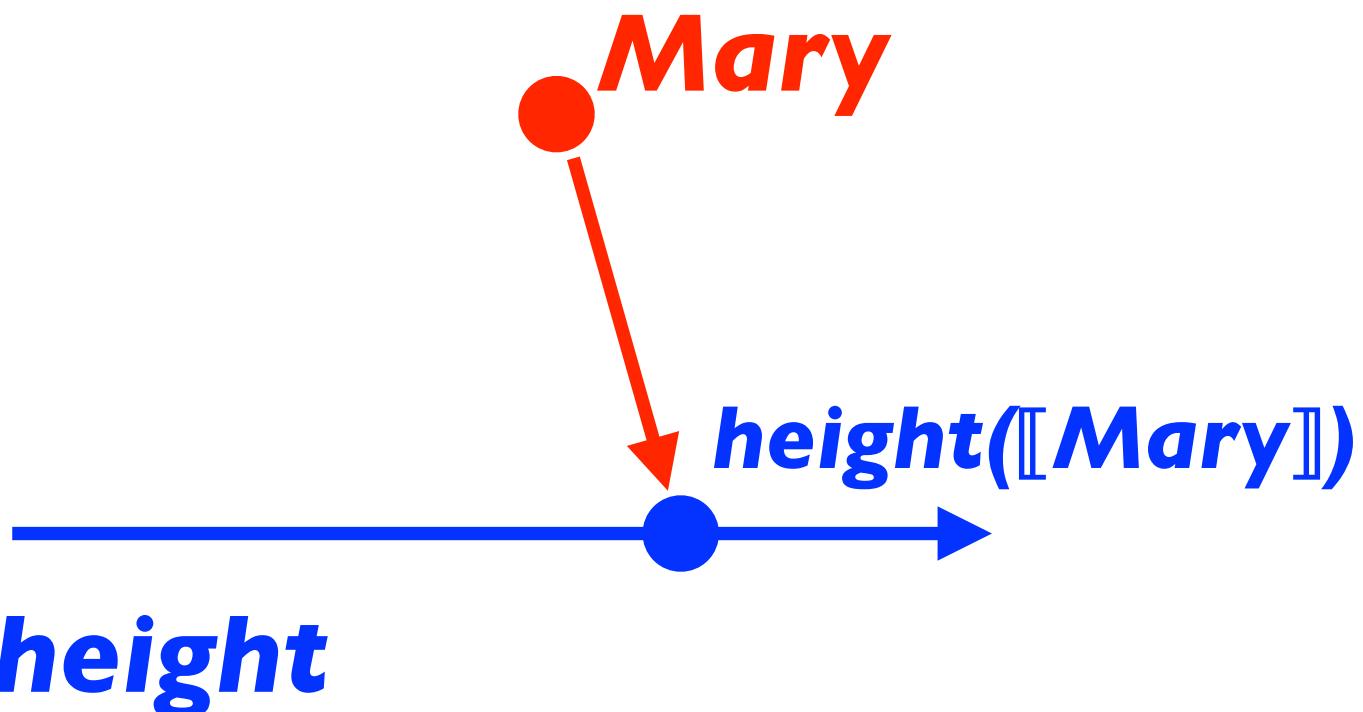
Mary is tall TRUE



Degree semantics for scalar adjectives

- The meaning of a scalar adjective like *tall* does two things
 1. Projects a referent onto some **value** on a **scale**
 2. Predicates that that **value** is greater than some **threshold θ**

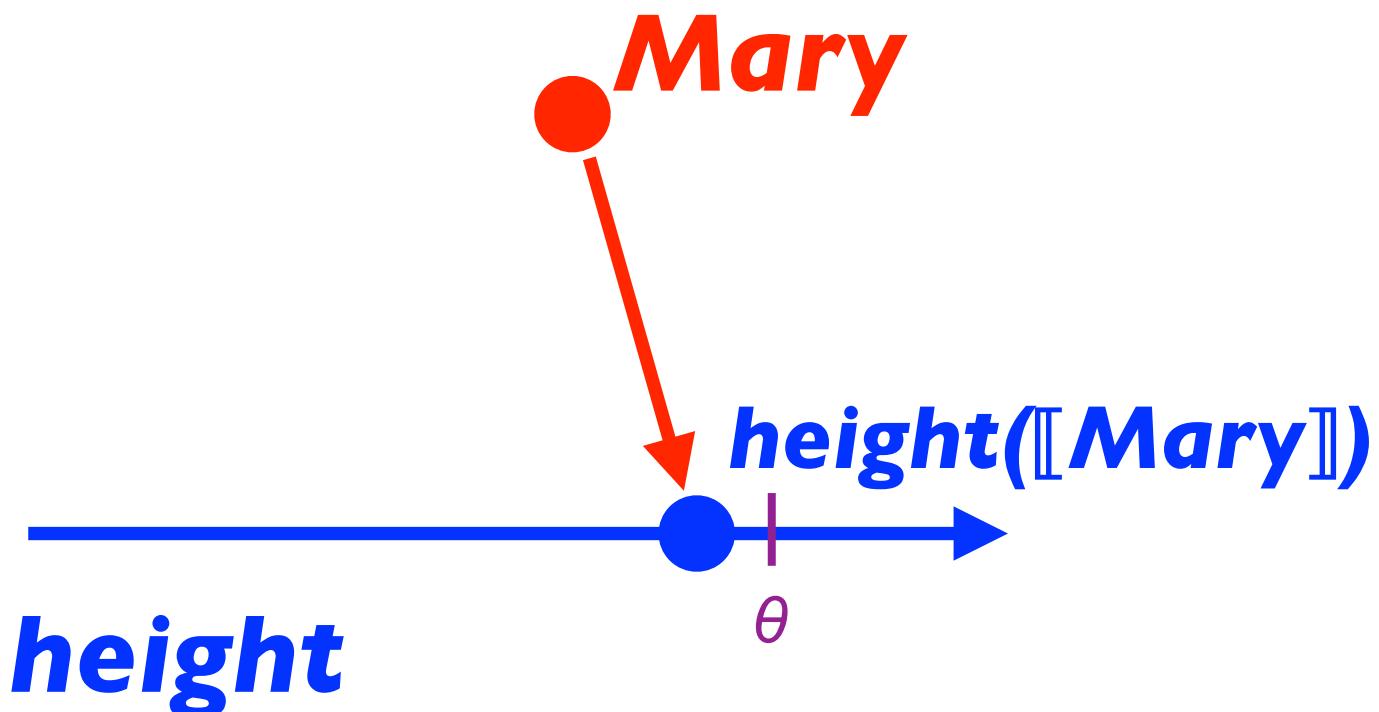
Mary is tall



Degree semantics for scalar adjectives

- The meaning of a scalar adjective like *tall* does two things
 1. Projects a referent onto some **value** on a **scale**
 2. Predicates that that **value** is greater than some **threshold θ**

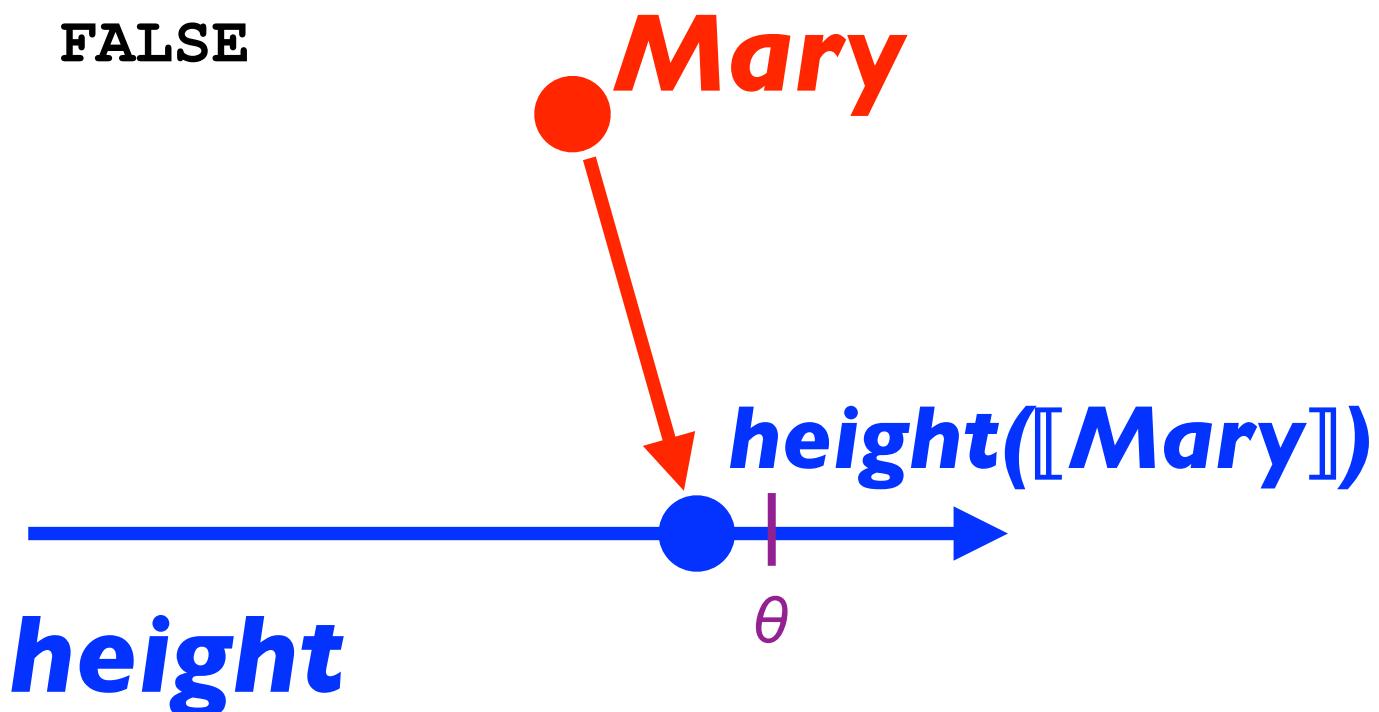
Mary is tall



Degree semantics for scalar adjectives

- The meaning of a scalar adjective like *tall* does two things
 - Projects a referent onto some **value** on a **scale**
 - Predicates that that **value** is greater than some **threshold θ**

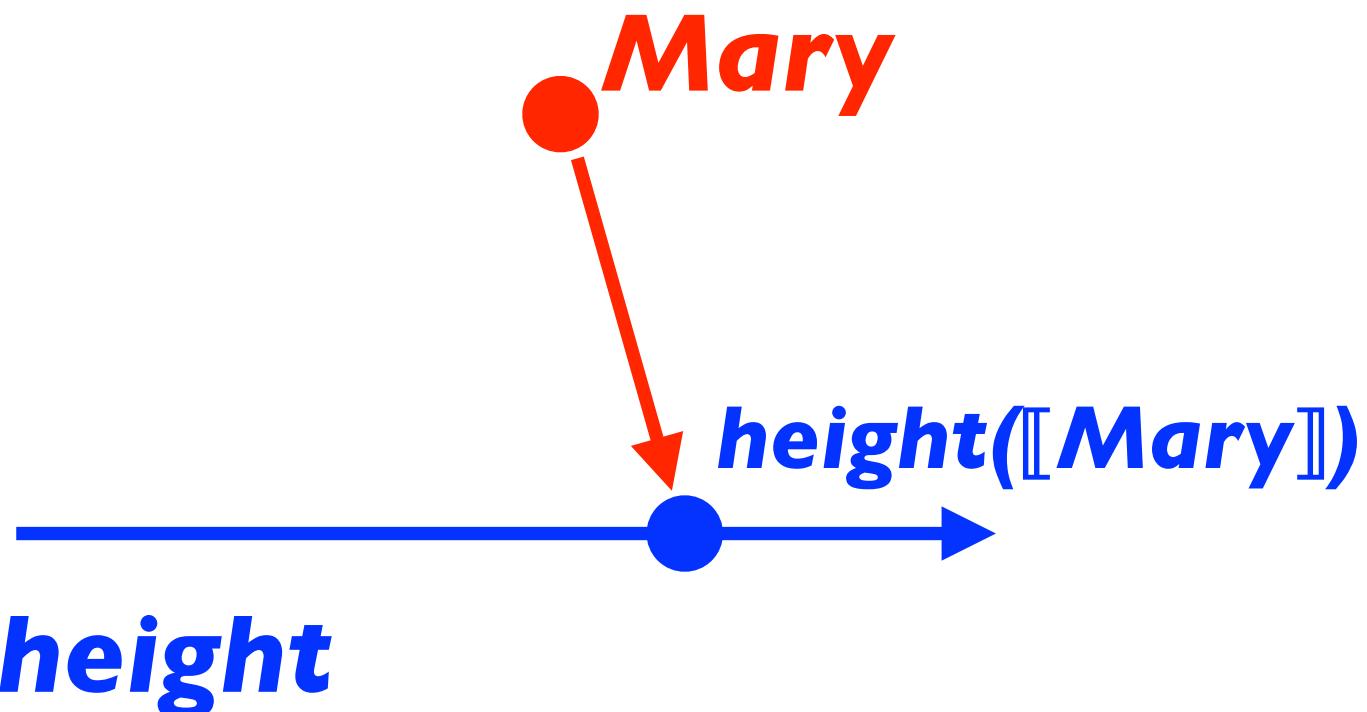
Mary is tall FALSE



Degree semantics for scalar adjectives

- The meaning of a scalar adjective like *tall* does two things
 1. Projects a referent onto some **value** on a **scale**
 2. Predicates that that **value** is greater than some **threshold θ**

Mary is tall



Degree semantics for scalar adjectives

- The meaning of a scalar adjective like *tall* does two things
 1. Projects a referent onto some **value** on a **scale**
 2. Predicates that that **value** is greater than some **threshold θ**

Mary is tall

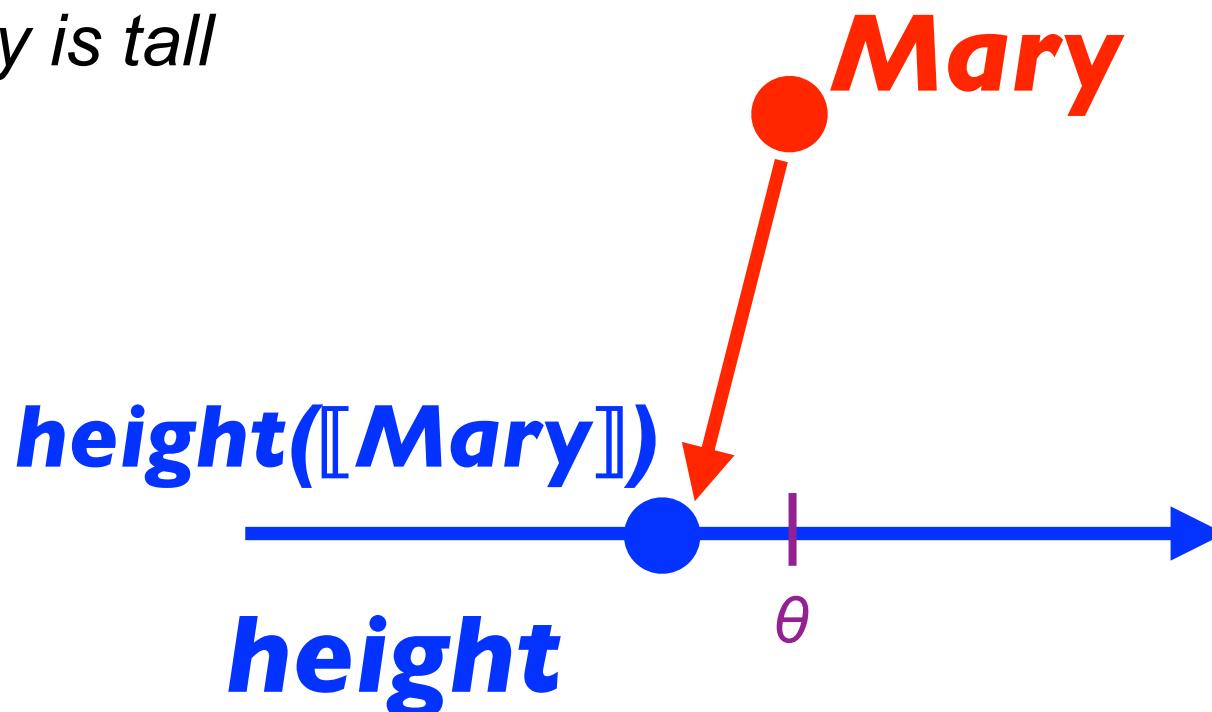


height

Degree semantics for scalar adjectives

- The meaning of a scalar adjective like *tall* does two things
 - Projects a referent onto some **value** on a **scale**
 - Predicates that that **value** is greater than some **threshold θ**

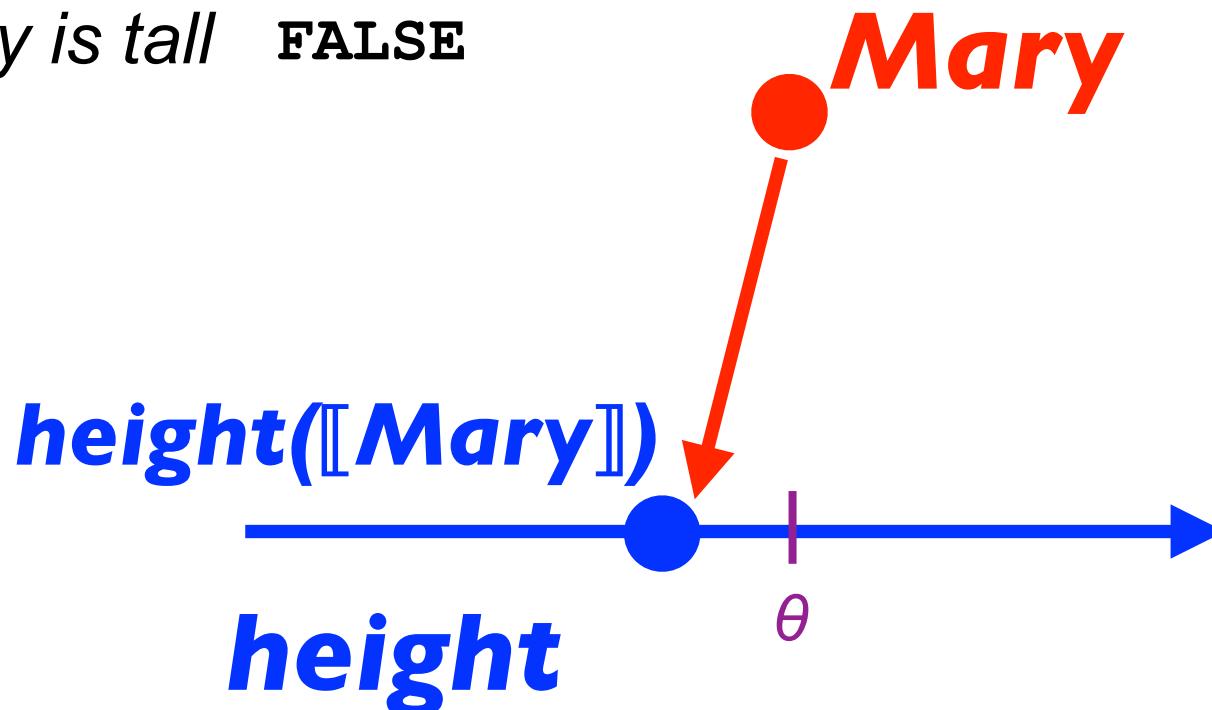
Mary is tall



Degree semantics for scalar adjectives

- The meaning of a scalar adjective like *tall* does two things
 - Projects a referent onto some **value** on a **scale**
 - Predicates that that **value** is greater than some **threshold θ**

Mary is tall FALSE



Observations regarding degree semantics

- Differences in scale structure can predict validity of compositions

Observations regarding degree semantics

- Differences in scale structure can predict validity of compositions



fullness

Observations regarding degree semantics

- Differences in scale structure can predict validity of compositions



fullness

- ✓ *The glass is perfectly full.*
- ✓ *The glass is perfectly empty.*

Observations regarding degree semantics

- Differences in scale structure can predict validity of compositions



fullness

- ✓ *The glass is perfectly full.*
- ✓ *The glass is perfectly empty.*



danger

Observations regarding degree semantics

- Differences in scale structure can predict validity of compositions



fullness

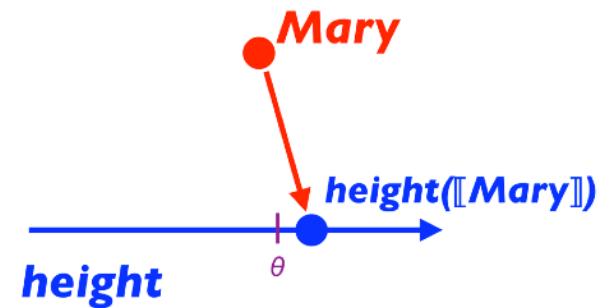
- ✓ *The glass is perfectly full.*
- ✓ *The glass is perfectly empty.*



danger

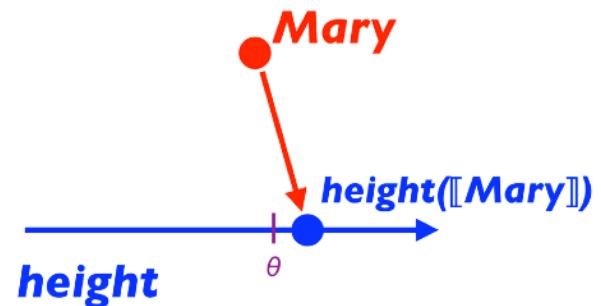
- ✓ *The neighbourhood is perfectly safe.*
- * *The neighbourhood is perfectly dangerous.*

What the degree semantics doesn't say

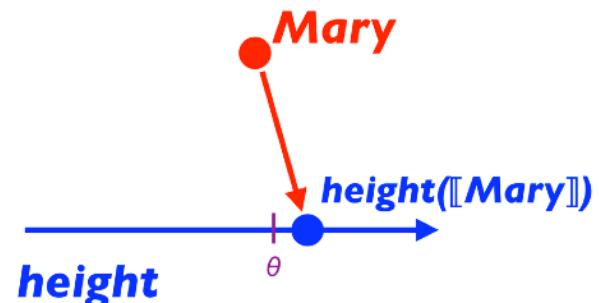


What the degree semantics doesn't say

- This is a very elegant model



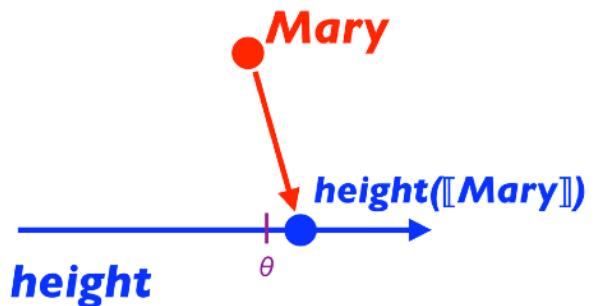
What the degree semantics doesn't say



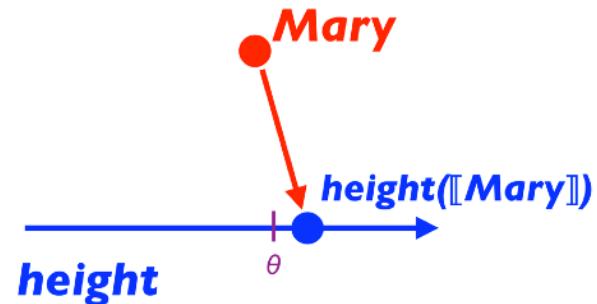
- This is a very elegant model
 - The abstractness of the model allows for context-sensitivity

What the degree semantics doesn't say

- This is a very elegant model
 - The abstractness of the model allows for context-sensitivity
- But it doesn't say *how* this context-sensitivity is achieved!

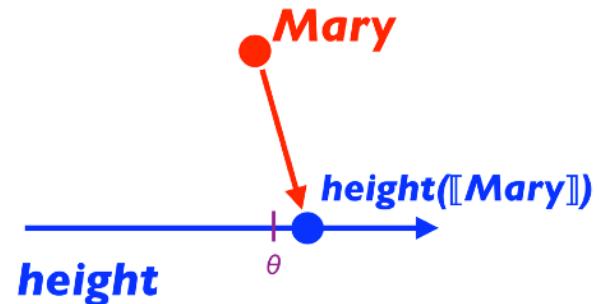


What the degree semantics doesn't say



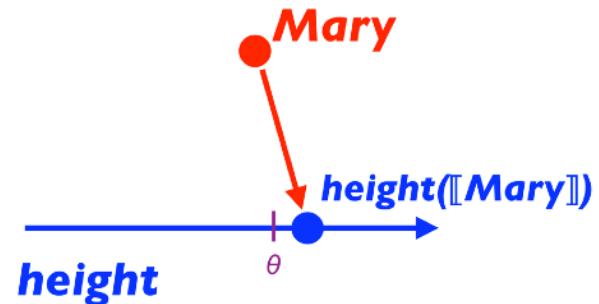
- This is a very elegant model
 - The abstractness of the model allows for context-sensitivity
- But it doesn't say *how* this context-sensitivity is achieved!
 - How does *tall elephant* turn out to mean something different from *tall mouse*?

What the degree semantics doesn't say



- This is a very elegant model
 - The abstractness of the model allows for context-sensitivity
- But it doesn't say *how* this context-sensitivity is achieved!
 - How does *tall elephant* turn out to mean something different from *tall mouse*?
 - How can the same *individual* be evaluated as either tall or not tall in different contexts?

What the degree semantics doesn't say



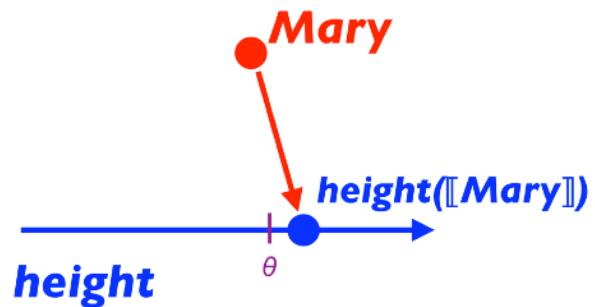
- This is a very elegant model
 - The abstractness of the model allows for context-sensitivity
- But it doesn't say *how* this context-sensitivity is achieved!
 - How does *tall elephant* turn out to mean something different from *tall mouse*?
 - How can the same *individual* be evaluated as either tall or not tall in different contexts?



What the degree semantics doesn't say

- This is a very elegant model
 - The abstractness of the model allows for context-sensitivity
- But it doesn't say *how* this context-sensitivity is achieved!
 - How does *tall elephant* turn out to mean something different from *tall mouse*?
 - How can the same *individual* be evaluated as either tall or not tall in different contexts?

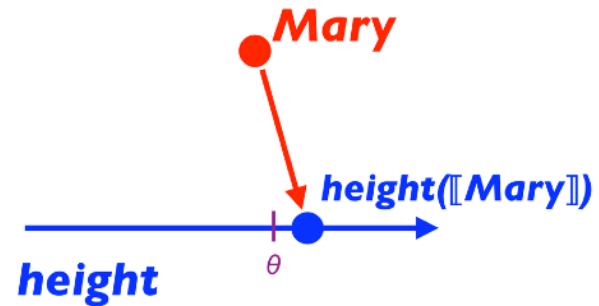
Stephen Curry is tall.



What the degree semantics doesn't say

- This is a very elegant model
 - The abstractness of the model allows for context-sensitivity
- But it doesn't say *how* this context-sensitivity is achieved!
 - How does *tall elephant* turn out to mean something different from *tall mouse*?
 - How can the same *individual* be evaluated as either tall or not tall in different contexts?

Stephen Curry is tall.



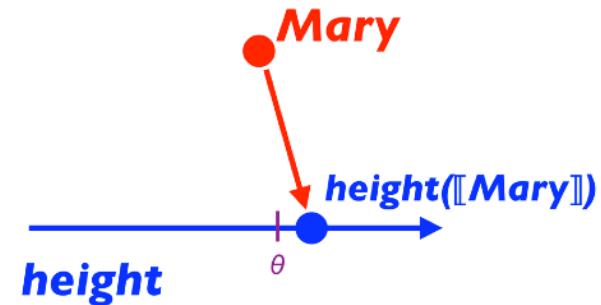
What the degree semantics doesn't say

- This is a very elegant model
 - The abstractness of the model allows for context-sensitivity
 - But it doesn't say *how* this context-sensitivity is achieved!
 - How does *tall elephant* turn out to mean something different from *tall mouse*?
 - How can the same *individual* be evaluated as either tall or not tall in different contexts?

Stephen Curry is tall.

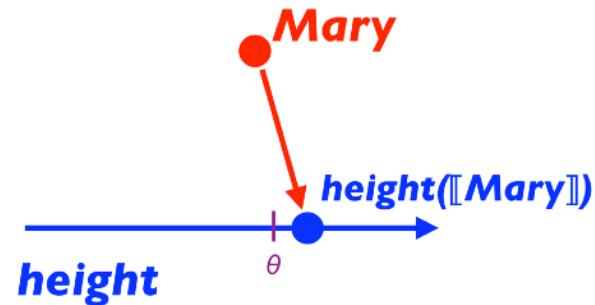


(*Stephen Curry is 6'2"; this is the 12th percentile of NBA player heights*)



What the degree semantics doesn't say

- This is a very elegant model
 - The abstractness of the model allows for context-sensitivity
 - But it doesn't say *how* this context-sensitivity is achieved!
 - How does *tall elephant* turn out to mean something different from *tall mouse*?
 - How can the same *individual* be evaluated as either tall or not tall in different contexts?



Stephen Curry is tall.

Stephen Curry is a tall basketball player.

(*Stephen Curry is 6'2"; this is the 12th percentile of NBA player heights*)



Towards a pragmatic model for scalar adjectives

- Desiderata
 - Inference on a continuum of possible scalar values
 - A threshold representation

The Lassiter & Goodman model

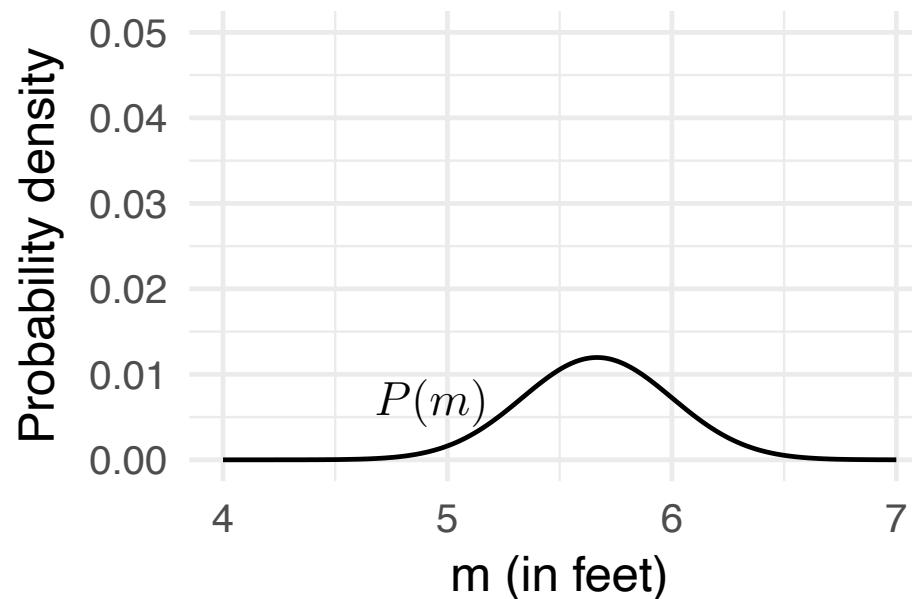
- The literal-listener model of interpretation:

$$L_0(m|u, \theta) \propto \begin{cases} P(m) & m \geq \theta \\ 0 & \text{otherwise} \end{cases}$$

The Lassiter & Goodman model

- The literal-listener model of interpretation:

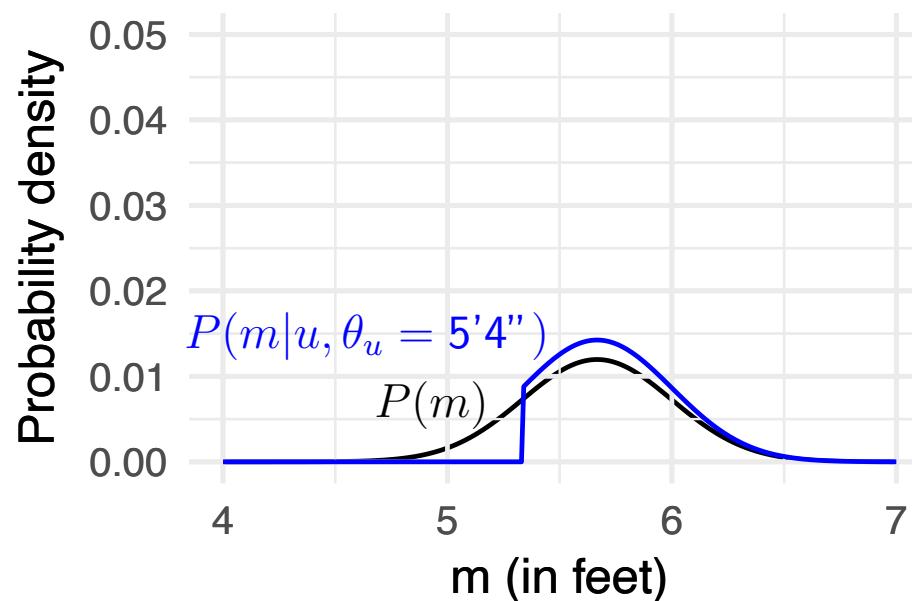
$$L_0(m|u, \theta) \propto \begin{cases} P(m) & m \geq \theta \\ 0 & \text{otherwise} \end{cases}$$



The Lassiter & Goodman model

- The literal-listener model of interpretation:

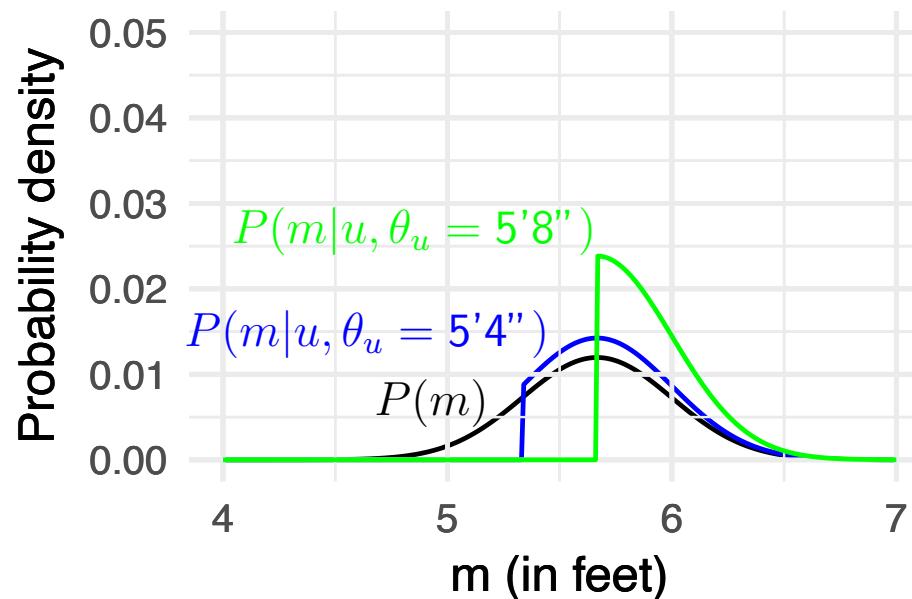
$$L_0(m|u, \theta) \propto \begin{cases} P(m) & m \geq \theta \\ 0 & \text{otherwise} \end{cases}$$



The Lassiter & Goodman model

- The literal-listener model of interpretation:

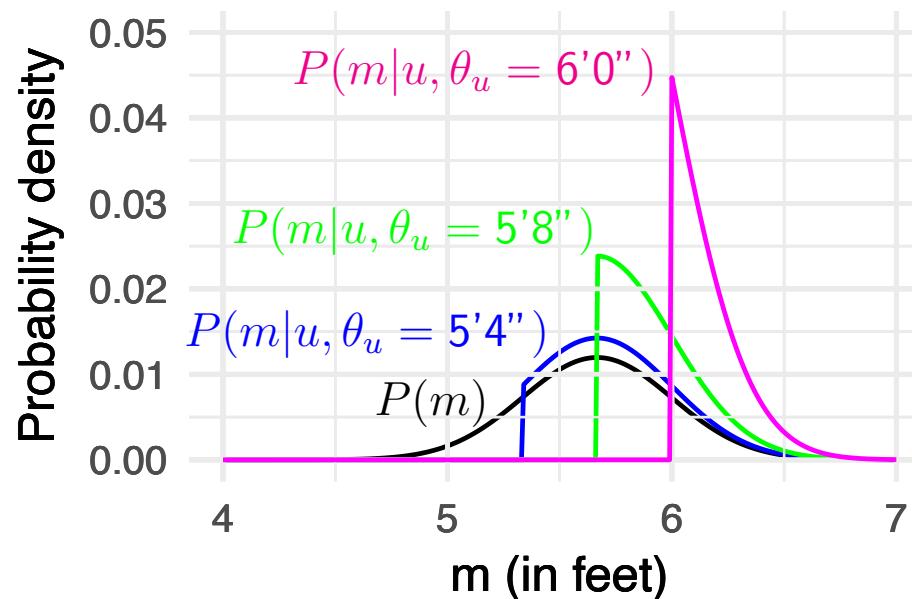
$$L_0(m|u, \theta) \propto \begin{cases} P(m) & m \geq \theta \\ 0 & \text{otherwise} \end{cases}$$



The Lassiter & Goodman model

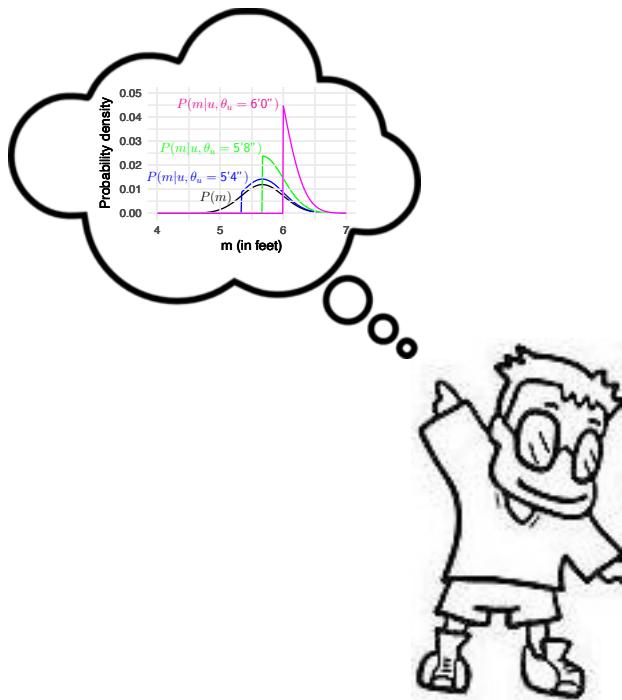
- The literal-listener model of interpretation:

$$L_0(m|u, \theta) \propto \begin{cases} P(m) & m \geq \theta \\ 0 & \text{otherwise} \end{cases}$$



A speaker model

- Assume a set of *alternative utterances* available to speaker
 - For “Pat ate some of the cookies”, alternatives were *some/all*
 - For “I injured a finger”, alternatives were *a/my/someone else’s*
- Here, we assume alternatives (to start) *tall* and **silence** (\emptyset)



tall?
 $\emptyset?$

Lassiter & Goodman's cost assumption:

$$\text{cost}(\text{tall}) = \text{cost}(\emptyset) + 2$$

$$\text{Utility}(u|m, \theta_u) = \log L_0(m|u, \theta_u) - \text{cost}(u)$$

$$S_1(u|m, \theta_u) \propto e^{\text{Utility}(u|m, \theta_u)}$$

$$S_1(u|m, \theta_u) \propto \frac{L_0(m|u, \theta_u)}{e^{\text{cost}(u)}}$$

A pragmatic listener

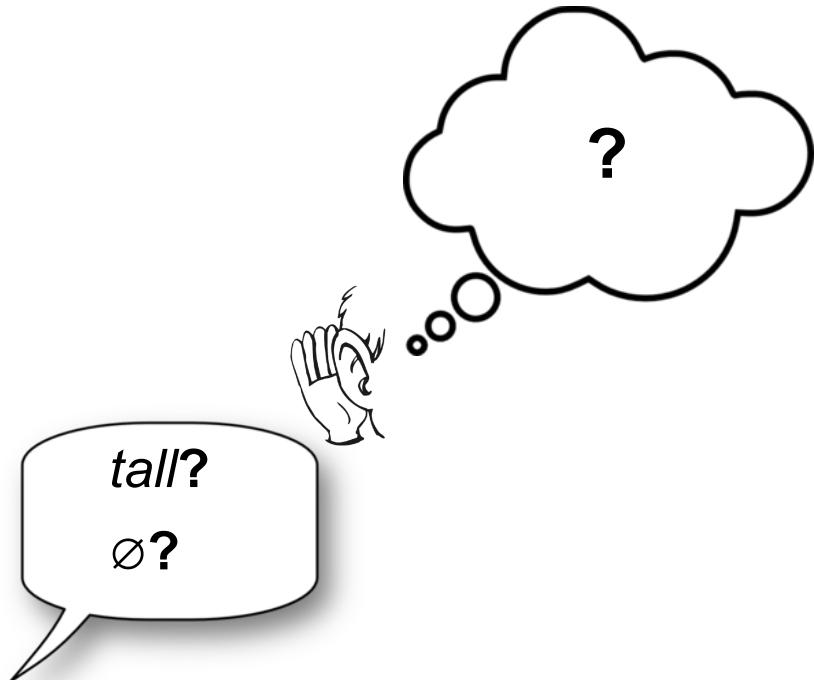
tall?

\emptyset ?

A pragmatic listener

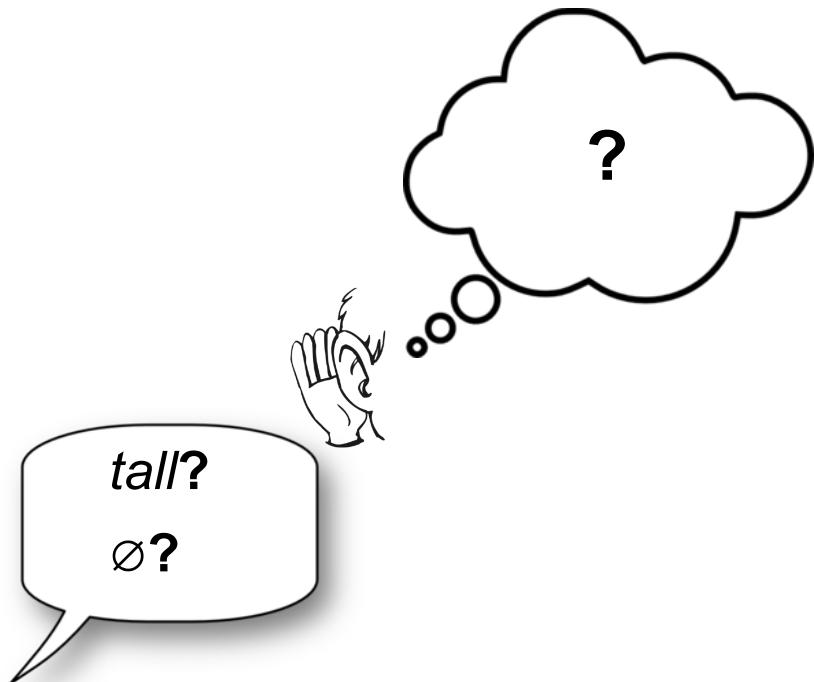


A pragmatic listener



A pragmatic listener

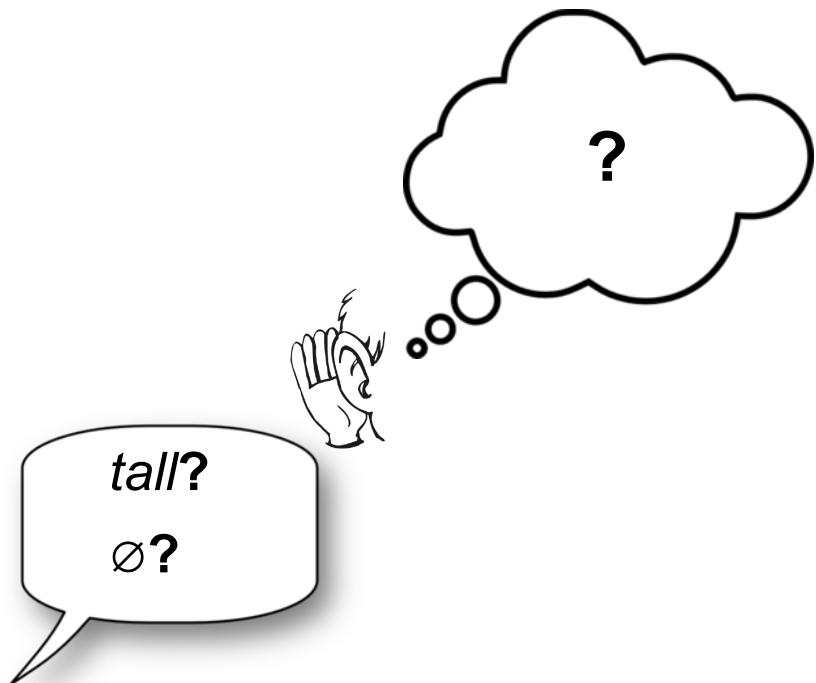
Pragmatic listener is a standard Bayesian comprehender:



A pragmatic listener

Pragmatic listener is a standard Bayesian comprehender:

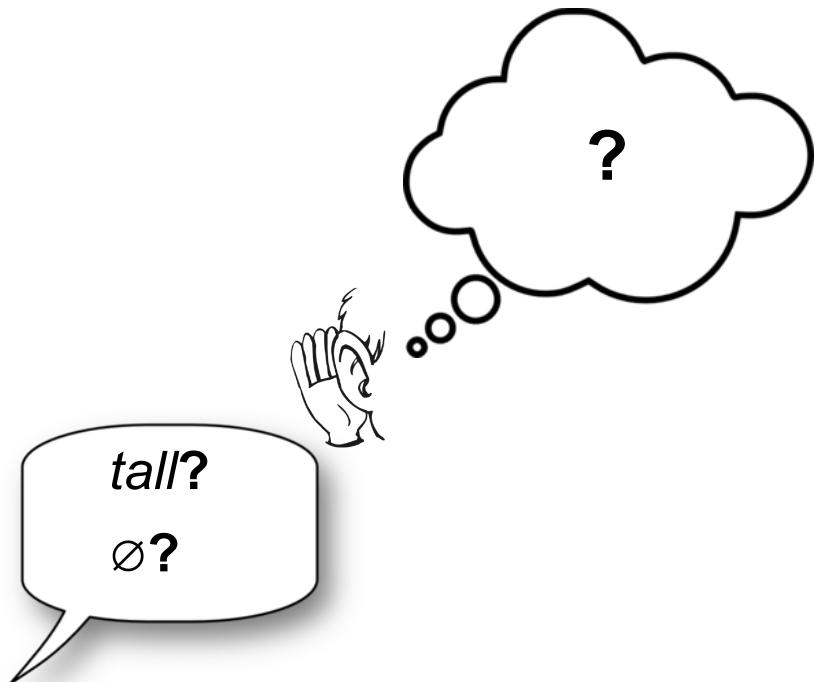
$$L_1(m, \theta_u | u) \propto S_1(u|m, \theta_u)P(m, \theta_u)$$



A pragmatic listener

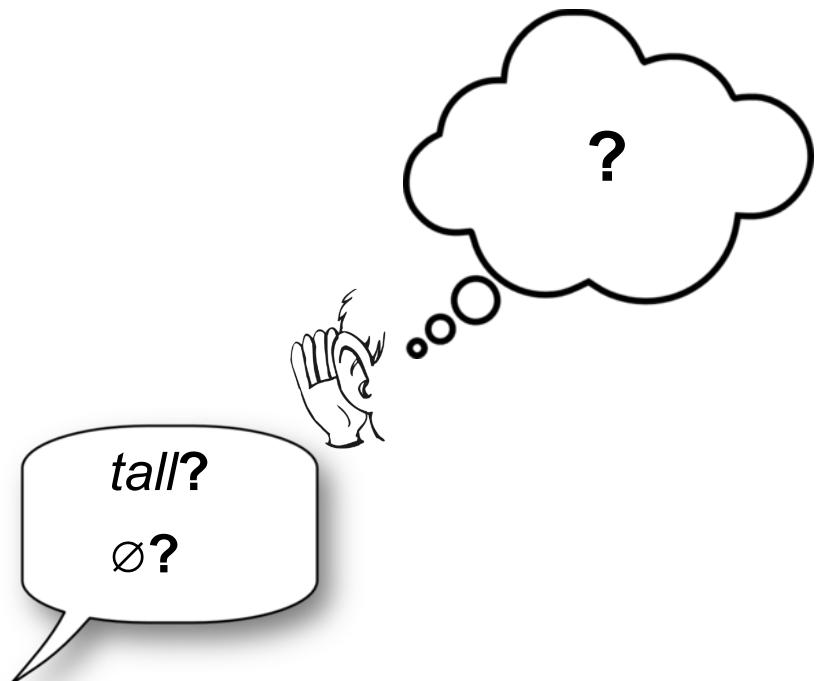
Pragmatic listener is a standard Bayesian comprehender:

$$L_1(m, \theta_u | u) \propto S_1(u|m, \theta_u) P(m, \theta_u)$$



A pragmatic listener

Pragmatic listener is a standard Bayesian comprehender:

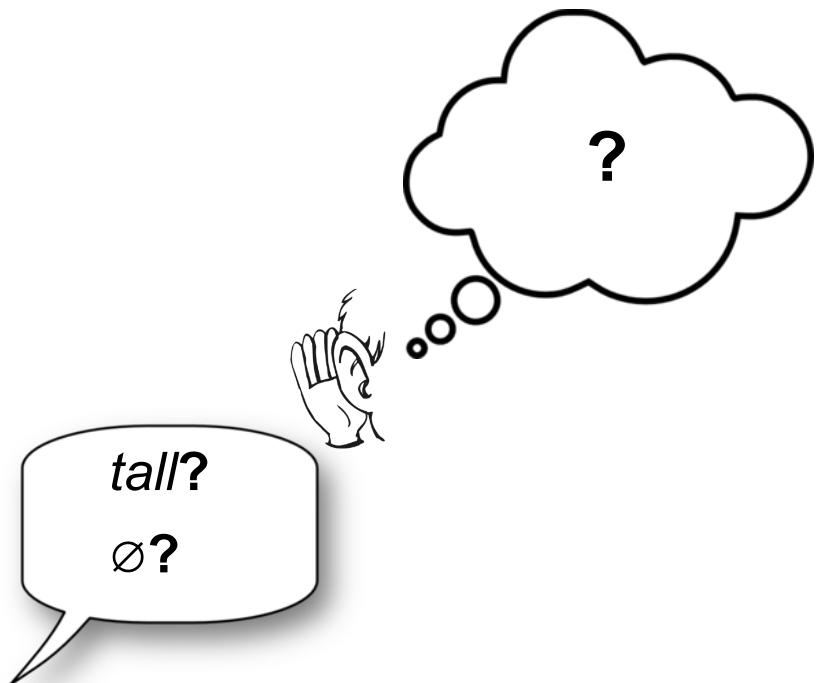


$$L_1(m, \theta_u | u) \propto S_1(u|m, \theta_u) P(m, \theta_u)$$

What do we do with this joint distribution?

A pragmatic listener

Pragmatic listener is a standard Bayesian comprehender:



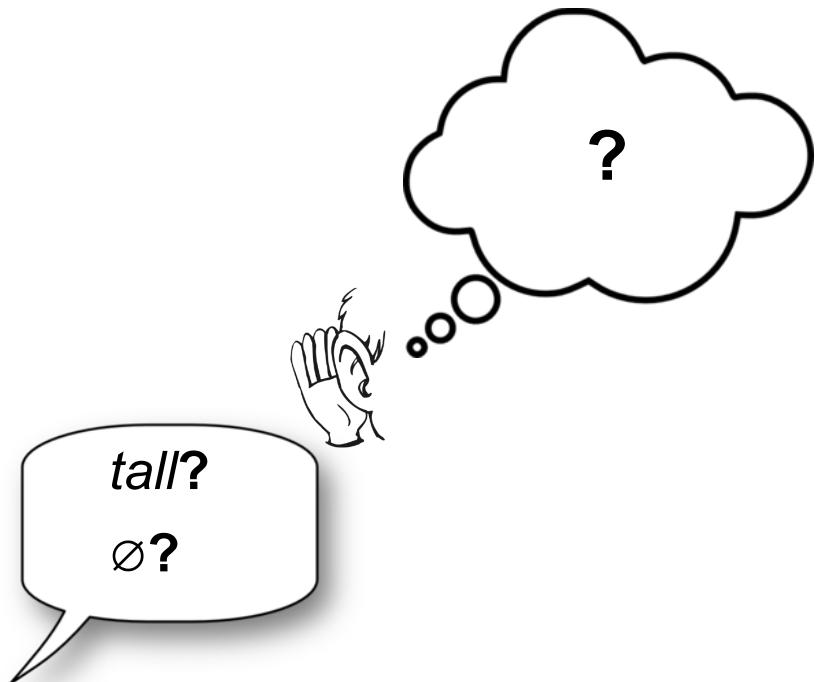
$$L_1(m, \theta_u | u) \propto S_1(u|m, \theta_u) P(m, \theta_u)$$

What do we do with this joint distribution?

Proposal: they are conditionally independent...

A pragmatic listener

Pragmatic listener is a standard Bayesian comprehender:



$$L_1(m, \theta_u | u) \propto S_1(u|m, \theta_u) P(m, \theta_u)$$

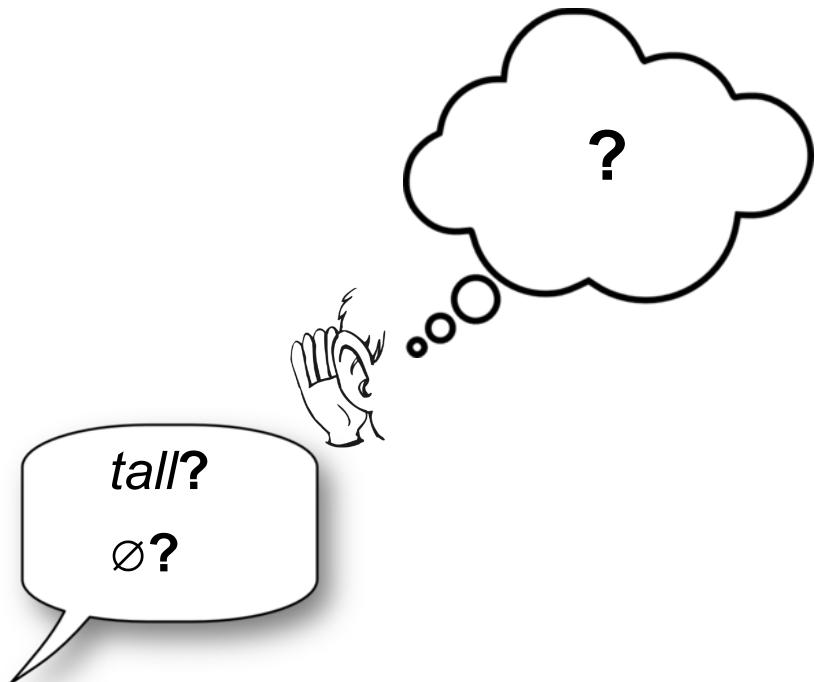
What do we do with this joint distribution?

Proposal: they are conditionally independent...

$$L_1(m, \theta_u | u) \propto S_1(u|m, \theta_u) P(m) P(\theta_u)$$

A pragmatic listener

Pragmatic listener is a standard Bayesian comprehender:



$$L_1(m, \theta_u | u) \propto S_1(u|m, \theta_u) P(m, \theta_u)$$

What do we do with this joint distribution?

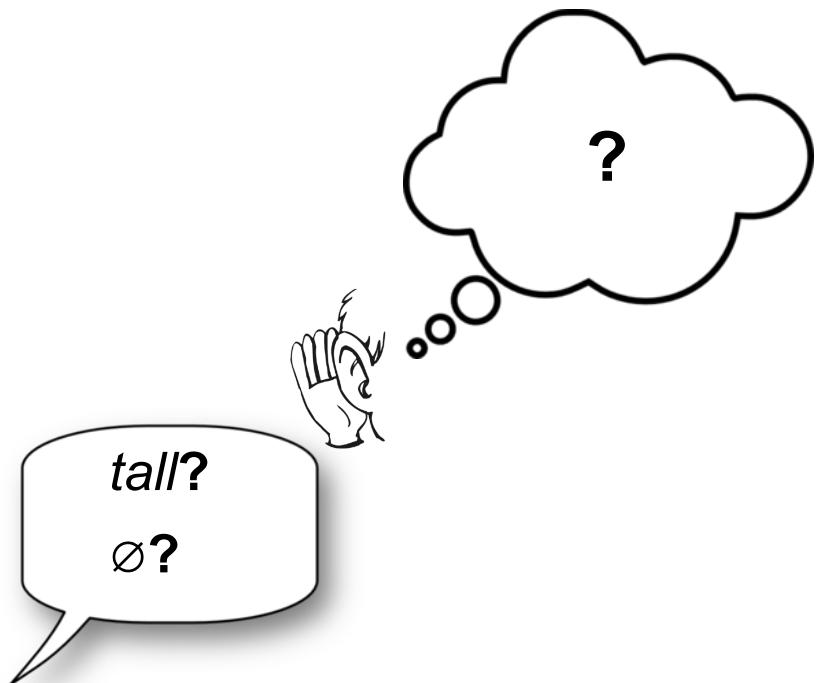
Proposal: they are conditionally independent...

$$L_1(m, \theta_u | u) \propto S_1(u|m, \theta_u) P(m) P(\theta_u)$$

*...and θ_u has a **uniform** prior:*

A pragmatic listener

Pragmatic listener is a standard Bayesian comprehender:



$$L_1(m, \theta_u | u) \propto S_1(u|m, \theta_u) P(m, \theta_u)$$

What do we do with this joint distribution?

Proposal: they are conditionally independent...

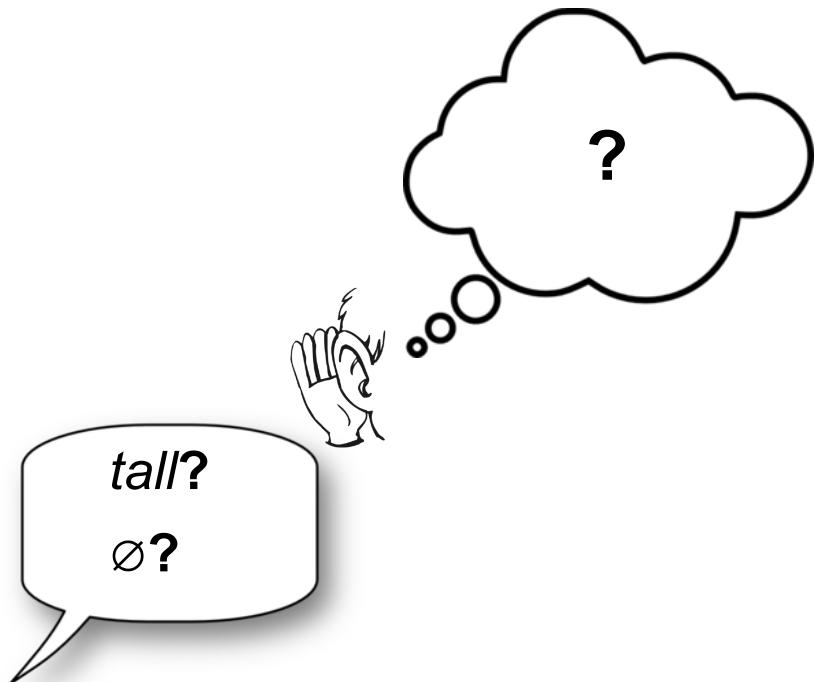
$$L_1(m, \theta_u | u) \propto S_1(u|m, \theta_u) P(m) P(\theta_u)$$

*...and θ_u has a **uniform** prior:*

$$L_1(m, \theta_u | u) \propto S_1(u|m, \theta_u) P(m)$$

A pragmatic listener

Pragmatic listener is a standard Bayesian comprehender:



$$L_1(m, \theta_u | u) \propto S_1(u|m, \theta_u) P(m, \theta_u)$$

What do we do with this joint distribution?

Proposal: they are conditionally independent...

$$L_1(m, \theta_u | u) \propto S_1(u|m, \theta_u) P(m) P(\theta_u)$$

*...and θ_u has a **uniform** prior:*

$$L_1(m, \theta_u | u) \propto S_1(u|m, \theta_u) P(m)$$

This is a proposal of non-trivial theoretical depth and interest; let's discuss!

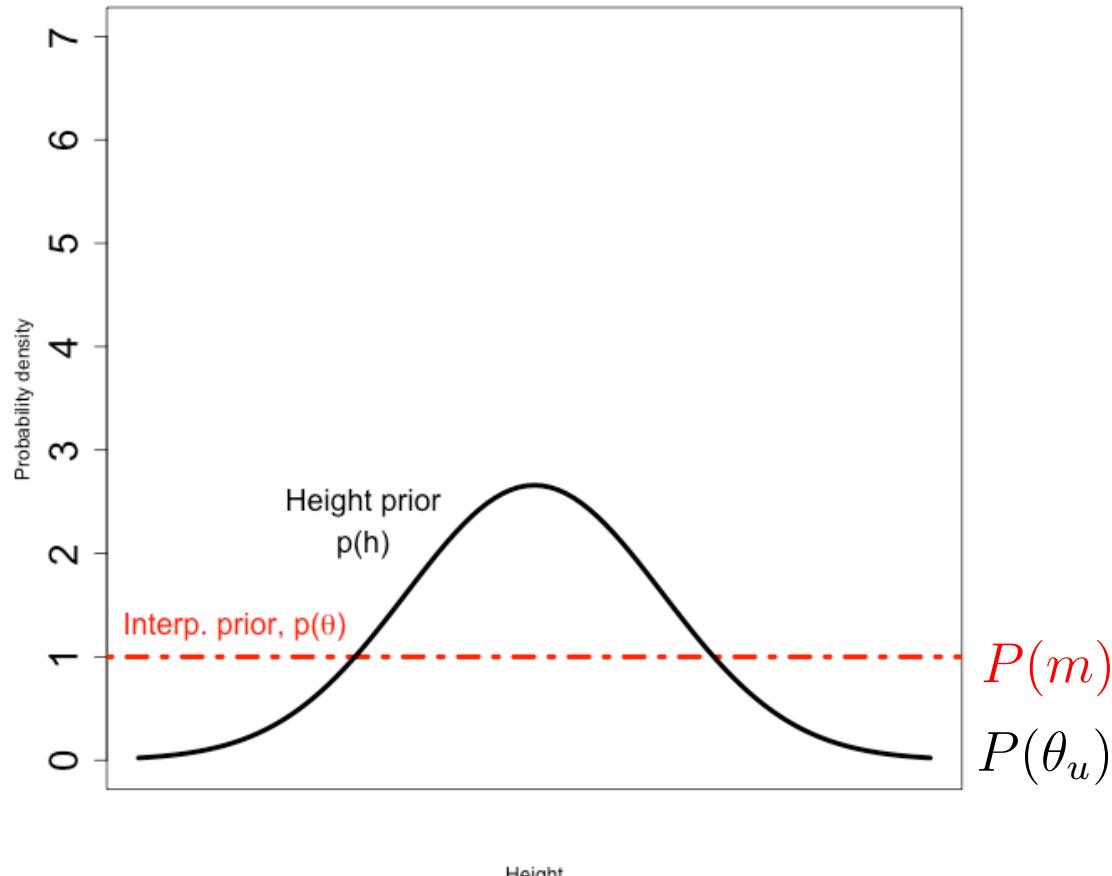
Visualizing pragmatic listener inferences

$$L_1(m, \theta_u | u) \propto S_1(u|m, \theta_u) P(m) \underline{P(\theta_u)}$$



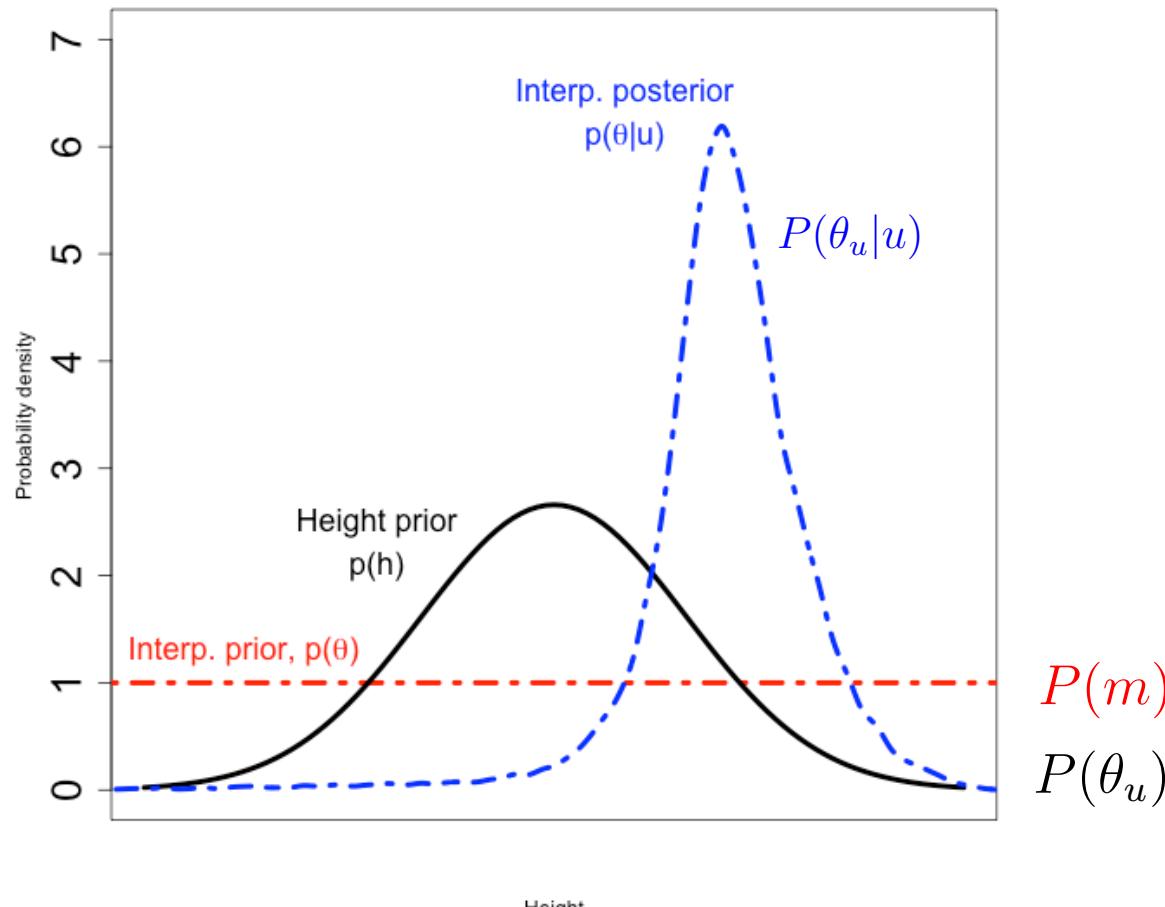
Visualizing pragmatic listener inferences

$$L_1(m, \theta_u | u) \propto S_1(u|m, \theta_u) \underline{P(m)P(\theta_u)}$$



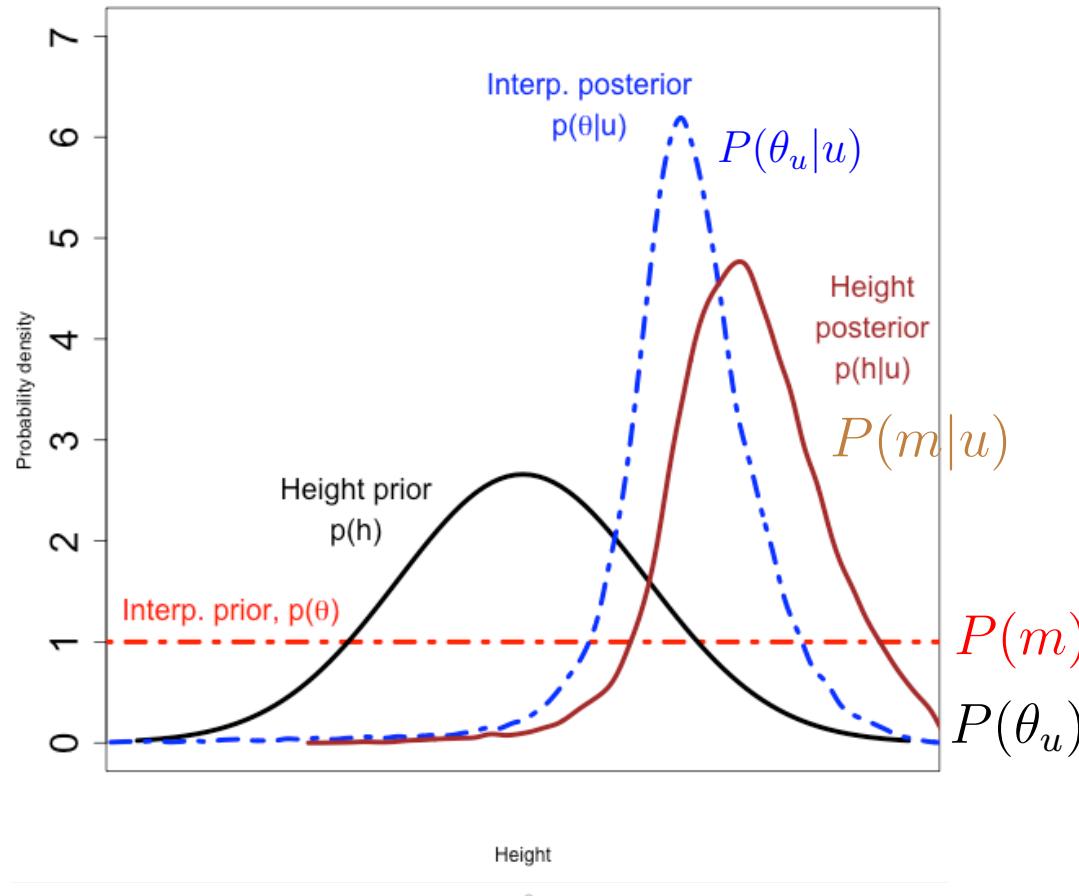
Visualizing pragmatic listener inferences

$$L_1(m, \underline{\theta_u} | u) \propto S_1(u|m, \theta_u) P(m) P(\underline{\theta_u})$$

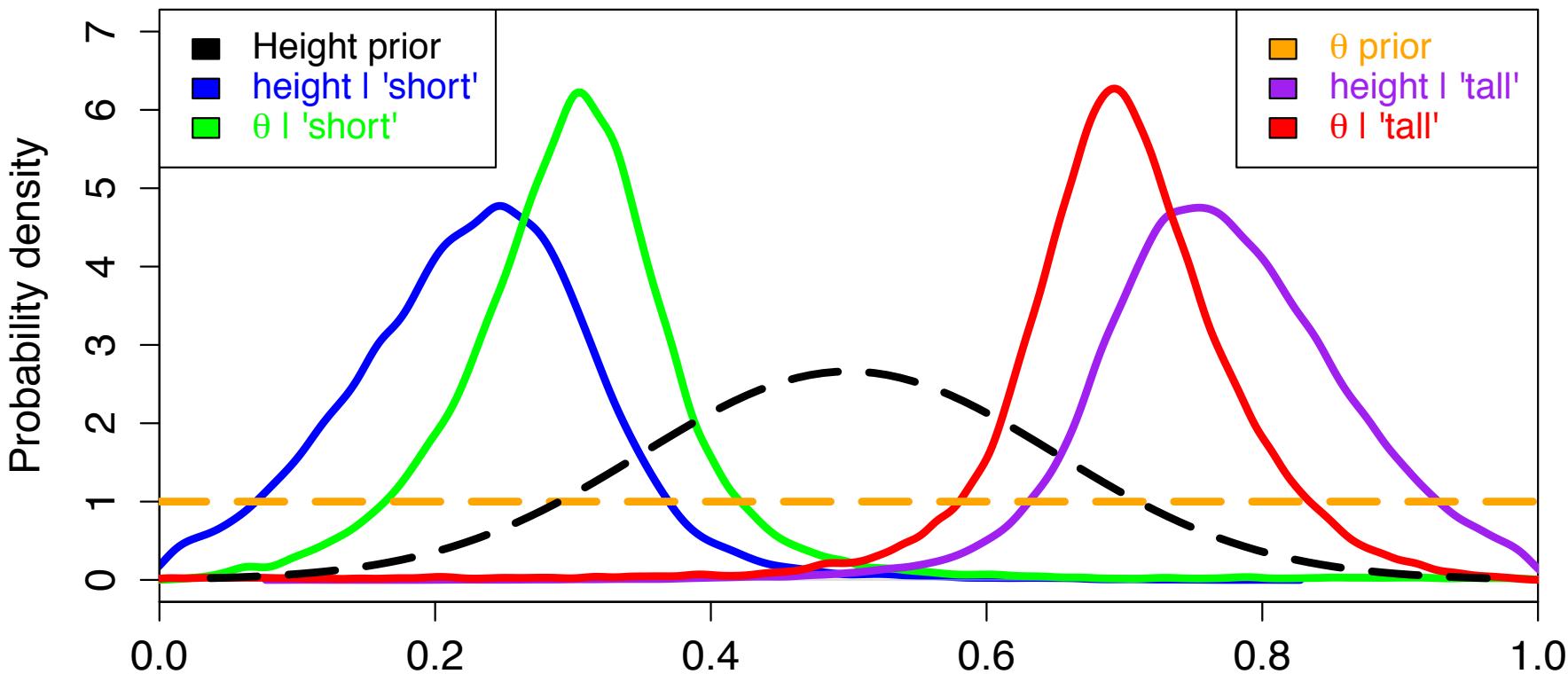


Visualizing pragmatic listener inferences

$$L_1(m, \theta_u | u) \propto S_1(u|m, \theta_u) P(m) P(\theta_u)$$



Antonyms

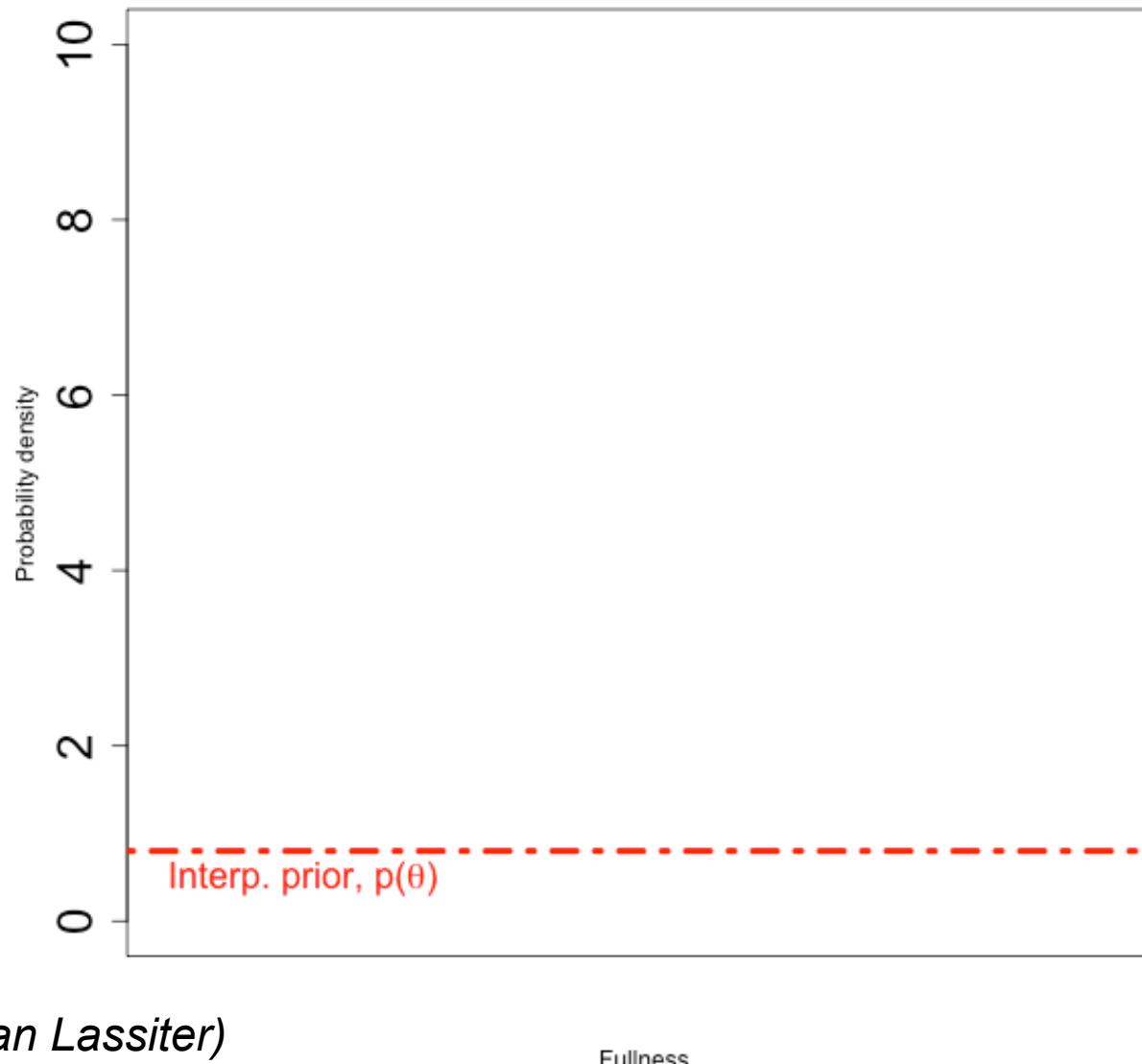


Absolute adjectives

- *full/empty, wet/dry, safe/dangerous, ...*
 - meanings are less (not?) context-dependent
 - meanings are sharp(er)
 - reference classes apparently not relevant to interpretation

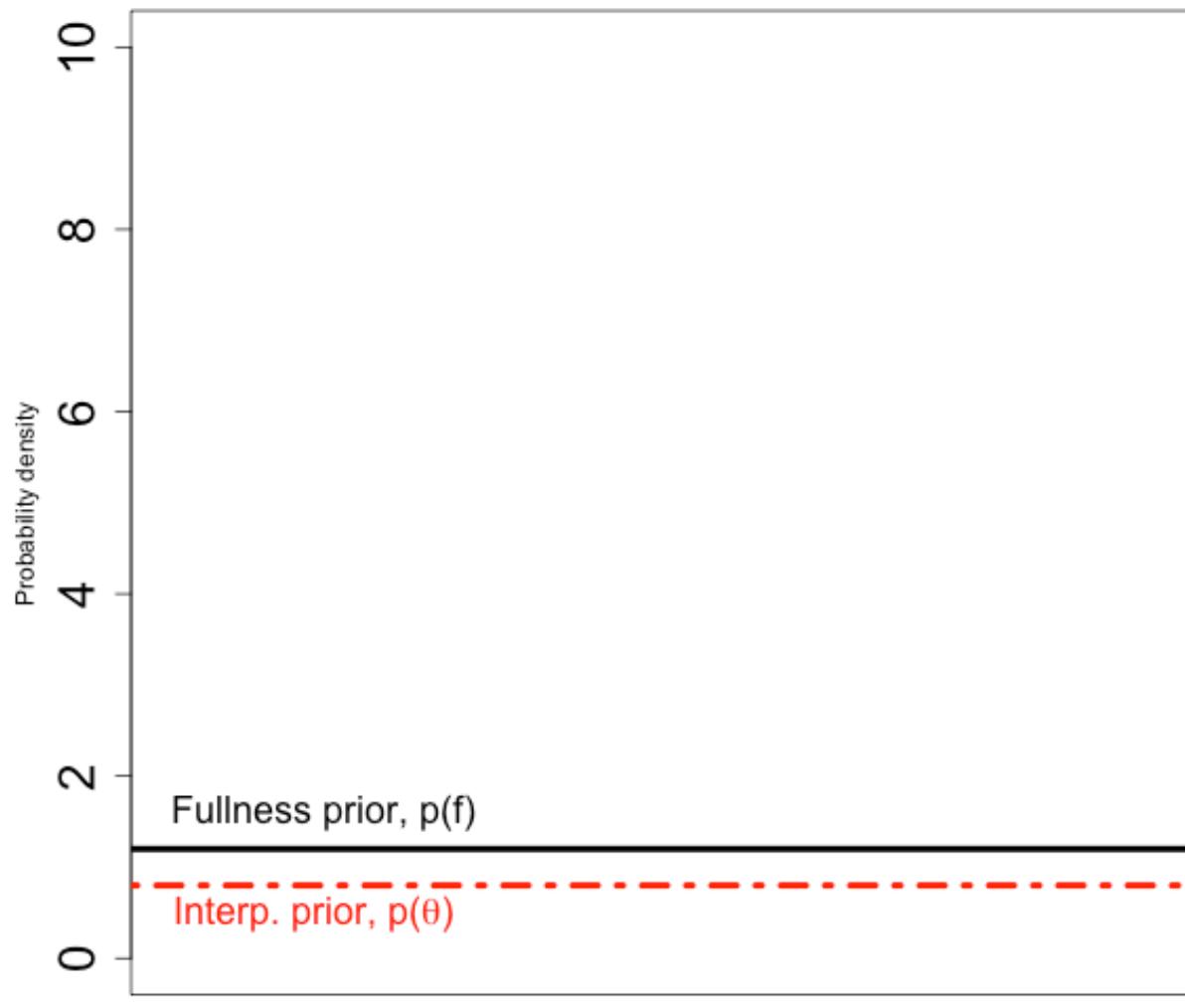
The pragmatic model on *full*

- Crucially, fullness is a *bounded* scale!



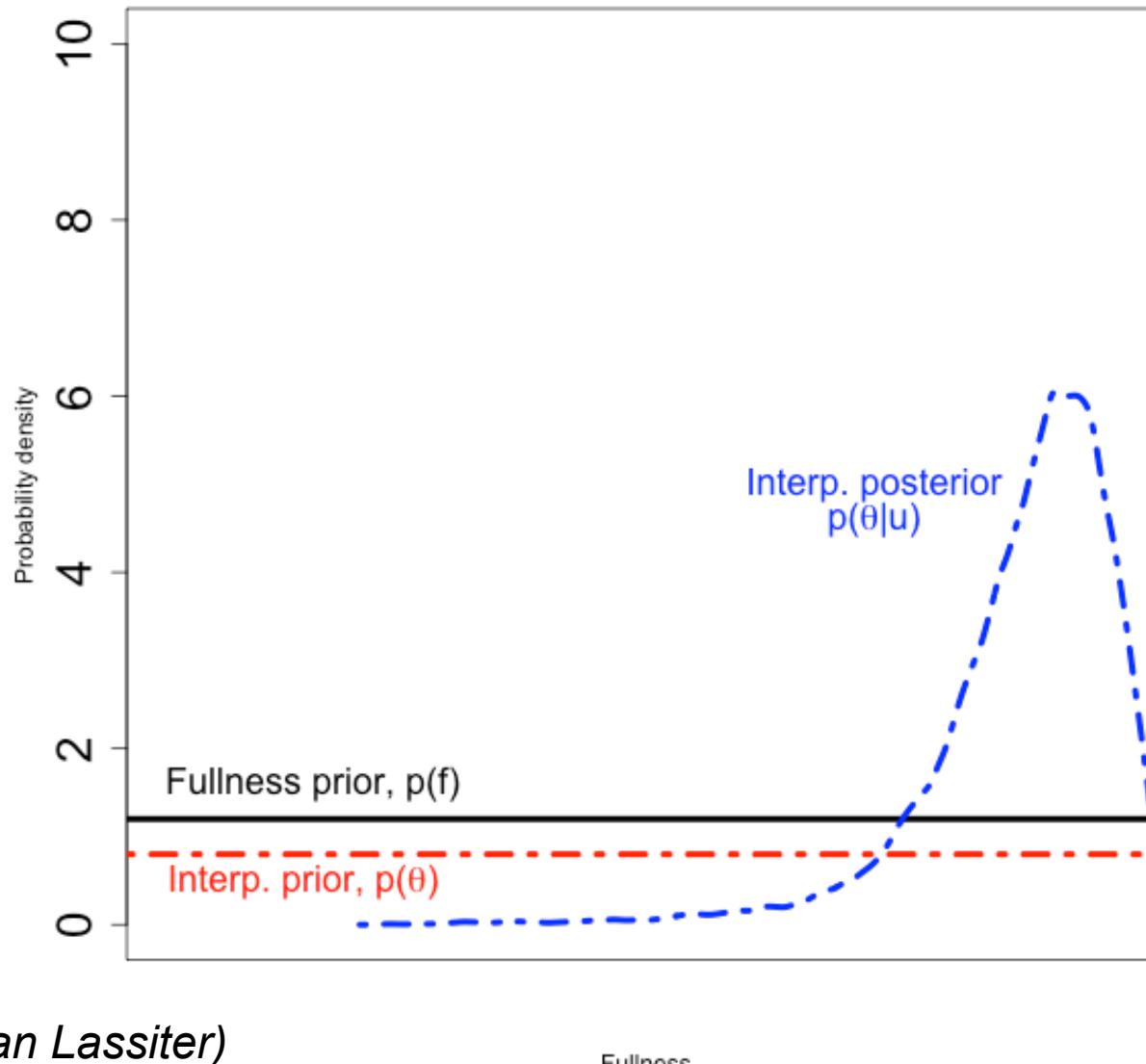
The pragmatic model on *full*

- Crucially, fullness is a *bounded* scale!



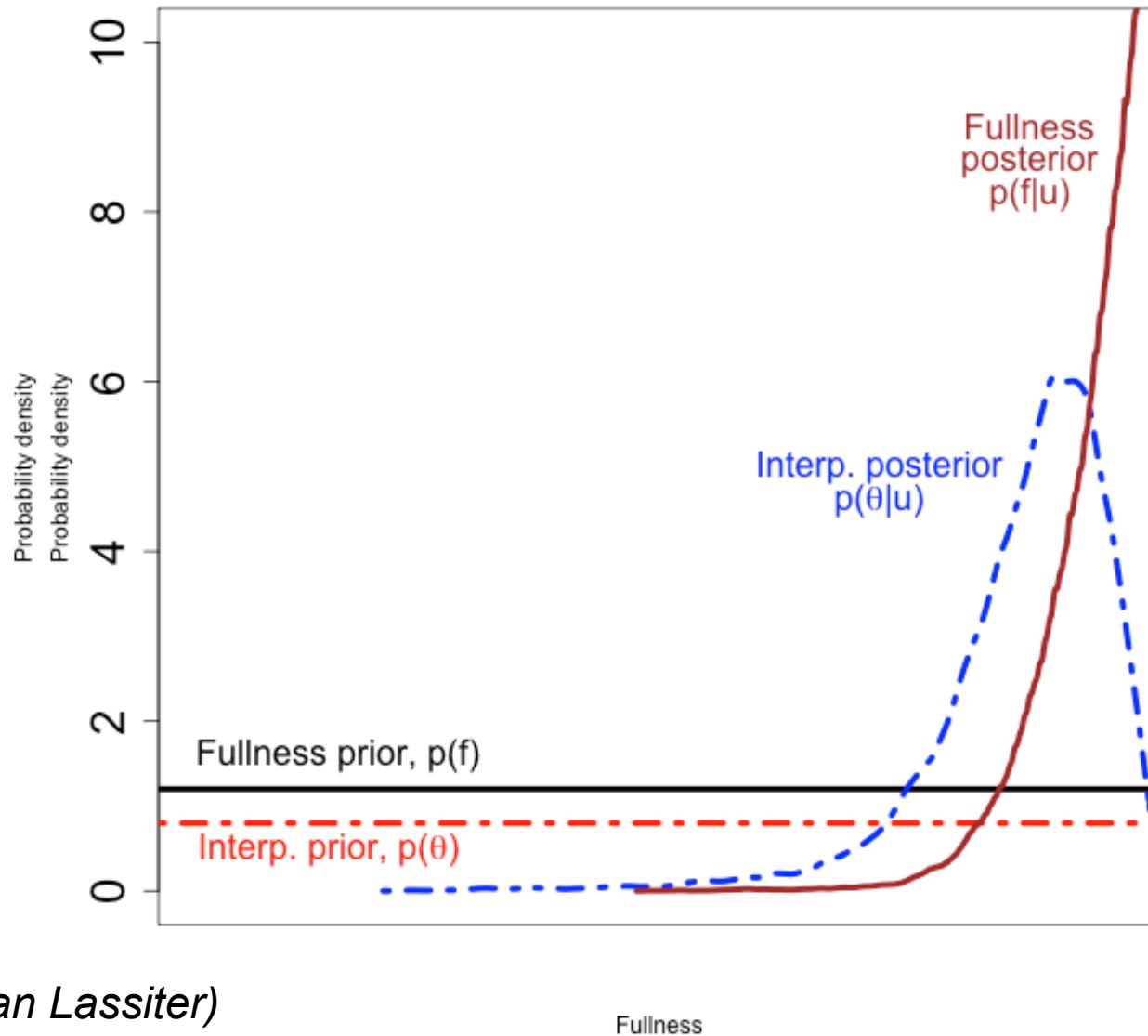
The pragmatic model on *full*

- Crucially, fullness is a *bounded* scale!



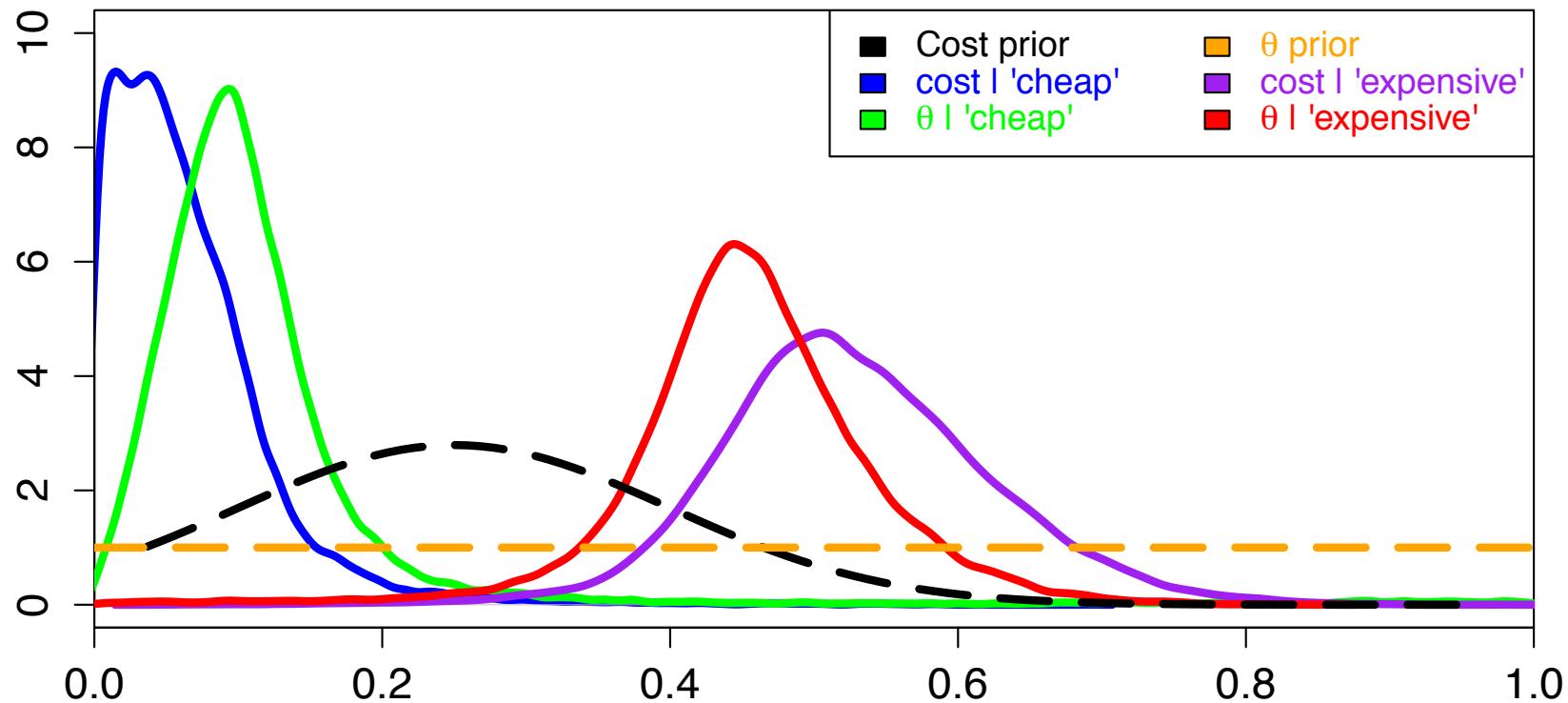
The pragmatic model on *full*

- Crucially, fullness is a *bounded* scale!



Bounds on scales

- On the Lassiter & Goodman model, asymmetries in the interpretations of adjectives arise naturally as a consequence of the prior



Summary

- Scalar adjectives are a simple example, but pose an additional challenge for pragmatics models
- Some part of the *literal meaning of an utterance* must get contextually determined
- This is one of the simplest examples of interleaving of semantic representation and probabilistic pragmatic inference
- Pieces of the puzzle:
 - Logical semantic representations
 - Latent-variable treatment of pieces of these representations
 - Prior probabilities on likely speaker meanings
 - Joint, utility-driven posterior inference on latent semantic variables and speaker meaning