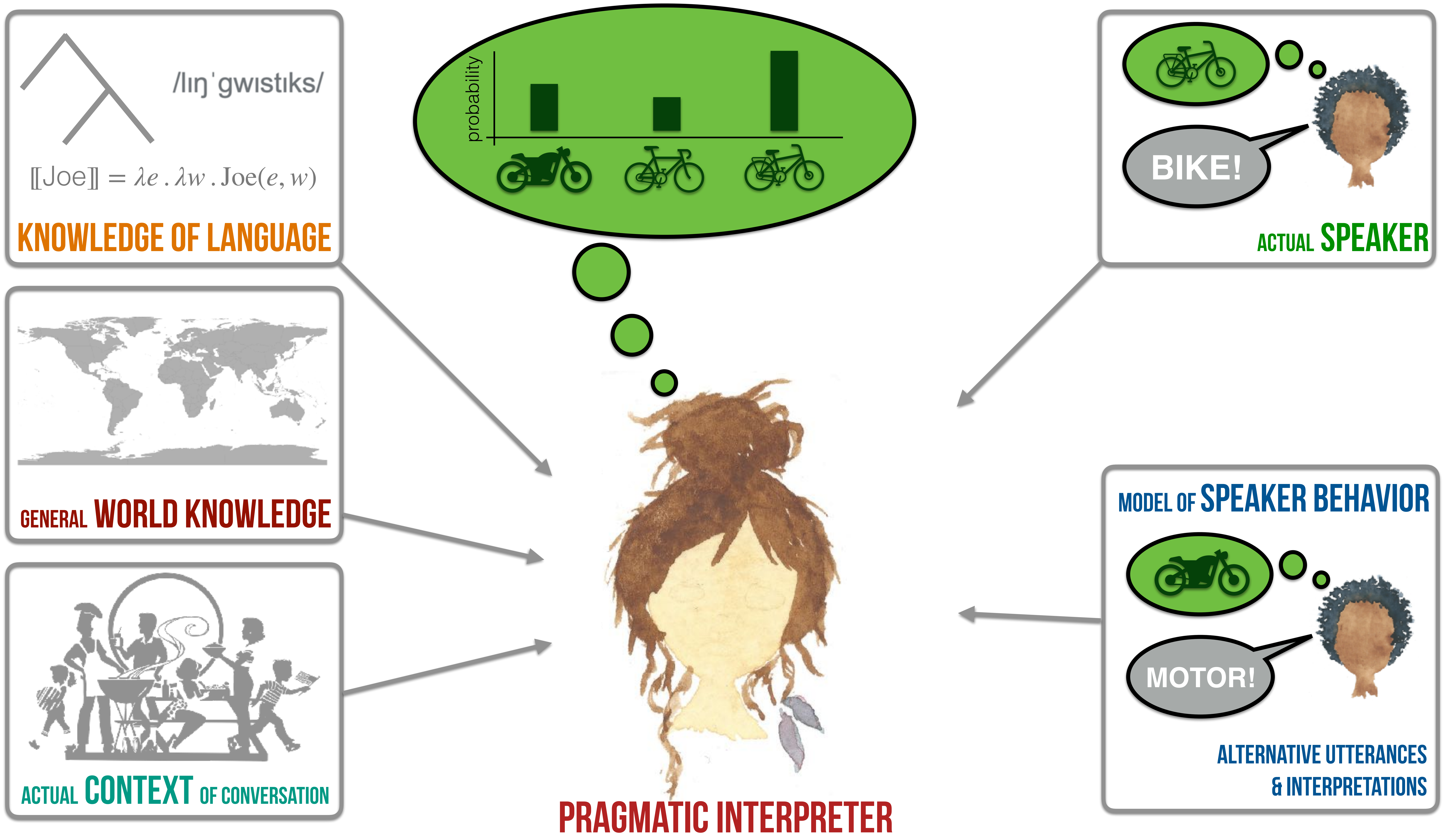
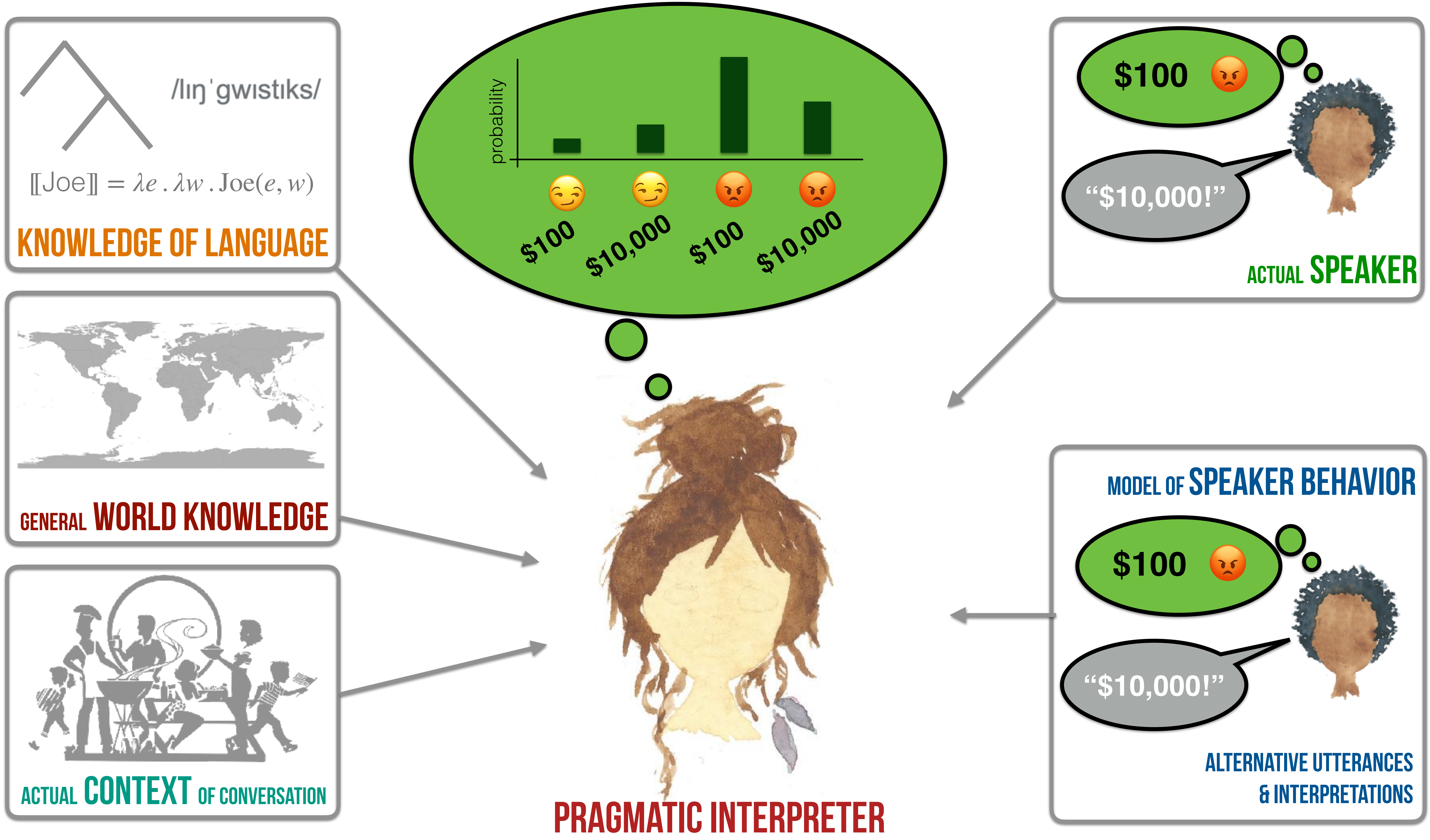
Computational Pragmatics

Resolving vague meanings (chapter V of problang.org)

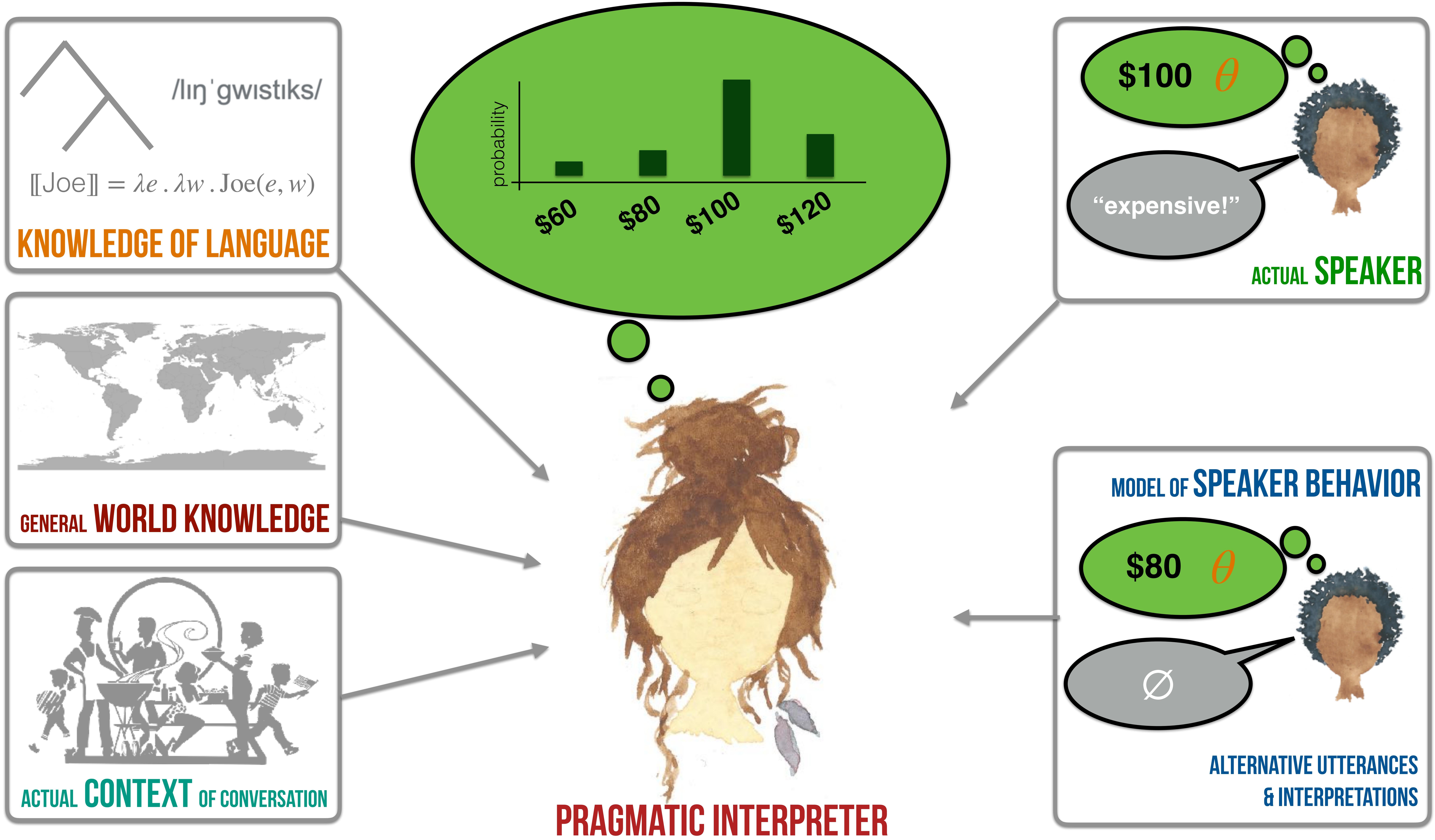
Session 8



/lɪŋˈgwɪstɪks/
[[Joe]] = λe . λw . Joe(e, w)
KNOWLEDGE OF LANGUAGE
GENERAL WORLD KNOWLEDGE
ACTUAL CONTEXT OF CONVERSATION
probability
PRAGMATIC INTERPRETER
BIKE!
ACTUAL SPEAKER
MODEL OF SPEAKER BEHAVIOR
MOTOR!
ALTERNATIVE UTTERANCES
& INTERPRETATIONS



/lɪŋˈgwɪstɪks/
⟦Joe⟧ = $\lambda e\,.\,\lambda w\,.\,\mathrm{Joe}(e, w)$
KNOWLEDGE OF LANGUAGE
GENERAL WORLD KNOWLEDGE
ACTUAL CONTEXT OF CONVERSATION
probability
\$100
\$10,000
\$100
\$10,000
\$100
“\$10,000!”
ACTUAL SPEAKER
MODEL OF SPEAKER BEHAVIOR
\$100
“\$10,000!”
ALTERNATIVE UTTERANCES
& INTERPRETATIONS
PRAGMATIC INTERPRETER



/lɪŋˈgwɪstɪks/
⟦Joe⟧ = λe . λw . Joe(e, w)
KNOWLEDGE OF LANGUAGE
GENERAL WORLD KNOWLEDGE
ACTUAL CONTEXT OF CONVERSATION
probability
$60
$80
$100
$120
PRAGMATIC INTERPRETER
$100 θ
"expensive!"
ACTUAL SPEAKER
MODEL OF SPEAKER BEHAVIOR
$80 θ
∅
ALTERNATIVE UTTERANCES
& INTERPRETATIONS

adjectival vagueness

“Geert-Jan is tall.”

🇳🇱 ♂ 🍺 comparison class

$[\![\text{tall}]\!] = \{x \mid \text{height}(x) \geq 185\,\text{cm}\}$

“Hanako is tall.”

🇯🇵 ♀ 子 comparison class

$[\![\text{tall}]\!] = \{x \mid \text{height}(x) \geq 120\,\text{cm}\}$

$$[\![\mathrm{tall}]\!]^{\theta} = \{x \mid \mathrm{height}(x) \geq \theta\}$$

- contextually supplied **threshold** θ
- depends on **comparison class** (possibly implicit)
- depends on (general, statistical) **world knowledge**

joint inference of state & threshold

Vanilla RSA

$$P_{L_0}(s \mid u) = P(s \mid [\![u]\!])$$

$$P_{S_1}(u \mid s) \propto \exp\left(\alpha\left(\log P_{L_0}(s \mid u) - \mathrm{C}(u)\right)\right)$$

$$P_{L_1}(s \mid u) \propto P(s)\ P_{S_1}(u \mid s)$$

θ-inference model

$$P_{L_0}(s \mid u, \theta) = P(s \mid [\![u]\!]^{\theta})$$

$$P_{S_1}(u \mid s, \theta) \propto \exp\left(\alpha\left(\log P_{L_0}(s \mid u, \theta) - \mathrm{C}(u)\right)\right)$$

$$P_{L_1}(s, \theta \mid u) \propto P(s)\ P(\theta)\ P_{S_1}(u \mid s, \theta)$$

θ-inference model

$$P_{L_0}(s \mid u, \theta) = P(s \mid [\![u]\!]^{\theta})$$

$$P_{S_1}(u \mid s, \theta) \propto \exp\left(\alpha\left(\log P_{L_0}(s \mid u, \theta) - \mathrm{C}(u)\right)\right)$$

$$P_{L_1}(s, \theta \mid u) \propto P(s)\ P(\theta)\ P_{S_1}(u \mid s, \theta)$$

- **state** s is the degree to which some object has the relevant property
- **state priors** $P(s)$ capture statistical world knowledge
- **threshold priors** $P(\theta)$ are uniform
- **relevant utterances** u are “adjective” and “silence”
- **utterance costs** $C(u)$ make “silence” cheaper than “adjective”