A Bayesian Approach to Modeling Grammatical Perspective-Taking

Carolyn Jane Anderson
Wellesley College
Perspective:

- the point-of-view from which an event or object is seen
- a person’s beliefs and attitudes
- a person’s self-identified spatial and temporal location
The cat is coming inside!

The cat is going inside!
What a magnificent creature!

I’d say it’s average at best.
Perspectival Expression:
an expression whose meaning depends on
the beliefs, perception, or location of a
prominent individual chosen by the speaker.
Last weekend Thelma ______ with me to New York. We saw the Statue of Liberty, took a ferry to Staten Island, and ate pizza. We even asked a New Yorker for directions just to see him get mad!

Fill in the blank:
(a) came
(b) went
Last weekend Thelma _____ with me to New York. We saw the Statue of Liberty, took a ferry to Staten Island, and ate pizza. We even asked a New Yorker for directions just to see him get mad!
Showing my exchange student Thelma the US has been fun!

How many states has she been to so far?

She's been to Connecticut, Maine, and New Hampshire. And last weekend Thelma ____ with me to New York. So, counting Vermont, five.

Fill in the blank:

(a) came
(b) went
Showing my exchange student Thelma the US has been fun!

She's been to Connecticut, Maine, and New Hampshire. And last weekend Thelma _____ with me to New York. So, counting Vermont, five.
Last weekend Thelma _____ with me to New York.

How does context affect our preference for *come* over *go*?

How do speakers decide which verb to use?

How do listeners reason about what the speaker is trying to say?
Talk Outline

1. Develop a computational model of perspectival reasoning
2. Run simulations to predict speaker and listener behavior
3. Compare model predictions against evidence from crowdsourced behavioral experiments
   - How do listeners understand perspectival expressions?
   - How do speakers choose perspectival expressions?
4. Open questions
Perspectival motion verbs

Thelma is coming to the bank.

*Come* requires the *perspective holder* to be located at the destination of motion.
Thelma is coming to the bank.
Listener perspective

Thelma is coming to the bank.
Thelma is coming to the bank.
Perspectival motion verbs

*Come* requires the **perspective holder** to be located at the destination of motion.

**Perspective holders:**
- The **listener**
- The **speaker**
- Attitude holders
- Protagonists

*Thelma is coming to the bank.*
The Semantics of Perspectival Expressions

**Indexical Family:**

\[
[[\text{come}]]_{C,g} = \\
\lambda x.\lambda e.\text{move}(e) \land \text{AGENT}(e,x) \land \text{DEST}(e,\text{LOC}(C_{\text{speaker}}))
\]

**Logophoric Binding Family:**

\[
[[\text{come}]]_{C,g} = \\
\lambda a.\lambda x.\lambda e.\text{move}(e) \land \text{AGENT}(e,x) \land \text{DEST}(e,\text{LOC}(a))
\]

**Anaphoric Family:**
(Barlew 2017, Roberts 2015)

\[
[[\text{come}]]_{C,g} = \\
\lambda x.\lambda e.\text{move}(e) \land \text{AGENT}(e,x) \land \text{DEST}(e,\text{LOC}(a))
\]
How do *speakers* pick a perspective?

\[ p(\text{perspective} | \text{meaning}) \]

How do *listeners* reason about whose perspective the speaker is using?

\[ p(\text{perspective} | \text{utterance}) \]

How do *contextual factors* influence perspective identification and selection?

\[ p(\text{perspective} | \text{context}) \]
**Talk Outline**

1. **Develop a computational model of perspectival reasoning as a joint inference problem**
2. Run simulations to predict speaker and listener behavior
3. Compare model predictions against evidence from crowdsourced behavioral experiments
   - How do listeners understand perspectival expressions?
   - How do speakers choose perspectival expressions?
4. Open questions
Cognitive Biases in Perspectival Processing

* Bias towards the speaker’s perspective
  - Speakers are egocentric: they prefer to use their own perspectives (Keysar & Barr 2003, Epley et al. 2004, Heller et al. 2012)
  - Accessing other perspectives is cognitively costly (Lin et al. 2010)
  - Speaker perspective appears to be most common cross-linguistically (Gathercole 1987, Nakazawa 2007, Korotkova 2016, Barlew 2017)

* Bias against perspective shift
  - Perspective shift is costly and language users avoid it (Millis 1995, Harris 2012, Köder et al. 2015, Ferguson et al. 2017)
  - Perspective shift is a risky conversational move (Harris & Potts 2009)
JOINT PERSPECTIVAL REASONING

Speaker:
\[ p(\text{perspective, utterance} \mid \text{meaning}) \]

Listener:
\[ p(\text{perspective, meaning} \mid \text{utterance}) \]
The Rational Speech Acts Model (Frank & Goodman 2012)

- Captures reasoning over alternatives in a rational approach: agents maximize communicative success
- Speakers and listeners use Bayesian reasoning: they reason using recursive mental models of each other’s behavior
- Easy to implement computationally to generate experimentally testable quantitative predictions
Applied to a variety of phenomena:

**scalar implicatures** (Bergen et al. 2012, Degen et al. 2015, Potts et al. 2016, Brochhagen et al. 2016 … )

**politeness** (Yoon et al. 2016, Yoon et al. 2017)

**irony** (Cohn-Gordon & Bergen 2019)

**hyperbole** (Kao et al. 2014)

**social meaning** (Qing & Cohn-Gordon 2018)

**visual perspective** (Hawkins et al. 2021)

**grammatical perspective** (Anderson & Dillon 2018; Anderson & Dillon forthcoming)

---

**The Rational Speech Acts model** (Frank & Goodman 2012)
Basic Rational Speech Acts Model

**Speaker’s goal:** $p(\text{utterance} \mid \text{meaning})$

Guess which sentence is most likely to communicate the intended meaning to the listener.

**Listener’s goal:** $p(\text{meaning} \mid \text{utterance})$

Guess which meaning the speaker’s sentence is supposed to communicate.

A meaning is represented as a possible world sampled from the world set.
Basic Rational Speech Acts Model

**Speaker’s goal: \( p(\text{utterance} \mid \text{world}) \)**

\[
p(\text{utterance} \mid \text{world}) = \frac{p(\text{world} \mid \text{utterance}) \cdot p(\text{utterance})}{p(\text{world})}
\]

**Listener’s goal: \( p(\text{world} \mid \text{utterance}) \)**

\[
p(\text{world} \mid \text{utterance}) = \frac{p(\text{utterance} \mid \text{world}) \cdot p(\text{world})}{p(\text{utterance})}
\]

*By Bayes’ Rule: \( p(A \mid B) = \frac{p(B \mid A) \cdot p(A)}{p(B)} \)*
A Recursive Model

Speaker: \( p(u \mid w) \propto p(w \mid u) p(u) \)

Listener: \( p(w \mid u) \propto p(u \mid w) p(w) \)
**Basic Rational Speech Acts model**

Pragmatic Listener: \( p(\text{world} \mid \text{utterance}) \)

\[
p(w \mid u) \propto p(u \mid w)p(w)
\]

\[
L_1(w \mid u) \propto \text{Speaker}(u \mid w) \ p(w)
\]

where \( w = \text{world} \) and \( u = \text{utterance} \)
Pragmatic Speaker: \( p(\text{utterance} \mid \text{world}) \)
\[ p(u \mid w) \propto p(w \mid u) \ p(u) \]

\( S_1(u \mid w) \propto \text{Max}(\text{Listener}(w \mid u) \ p(u)) \)

where \( w = \text{world} \) and \( u = \text{utterance} \)
Basic Rational Speech Acts model

Literal Listener: \( p(\text{world} \mid \text{utterance}) \)

\[
p(w \mid u) \propto p(u \mid w)p(w)
\]

\[
L_0(w \mid u) \propto [\![u]\!]^wp(w)
\]

where \( w = \text{world} \) and \( u = \text{utterance} \)
**Basic Rational Speech Acts model**

**Literal Listener:** $p(\text{world} \mid \text{utterance})$

$L_0(w \mid u) \propto [[u]]^wp(w)$

**Pragmatic Speaker:** $p(\text{utterance} \mid \text{world})$

$S_1(u \mid w) \propto \text{Max}(\text{Listener}(w \mid u) \ p(u))$

**Pragmatic Listener:** $p(\text{world} \mid \text{utterance})$

$L_1(w \mid u) \propto \text{Speaker}(u \mid w) \ p(w)$
Thelma is coming.

What is Sam trying to communicate?

Whose perspective is Sam using?

$p(\text{world} \mid \text{utterance})$

$p(\text{perspective} \mid \text{utterance})$
Thelma is coming to the park.

Sam must be at the park.
Thelma is coming.

Thelma’s destination must be the zoo.
Whose perspective is Sam using?

What is Sam trying to communicate?

Whose perspective is Sam using?

\( p(\text{perspective, world} \mid \text{utterance}) \)

Thelma is coming.
JOINT PERSPECTIVAL REASONING

**Listener:**
\[ p(\text{perspective, world | utterance}) \]

**Speaker:**
\[ p(\text{perspective, utterance | world}) \]
Pragmatic listener:
\[ p(\text{world, perspective} | \text{utterance}) \propto p(\text{utterance, perspective} | \text{world}) \, p(\text{world}) \]

\[ L_1(w, a | u) \propto S_1(u, a | w) \, p(w) \]

where \( w = \text{world} \), \( u = \text{utterance} \), \( a = \text{perspective} \)
Perspectival RSA model

Pragmatic speaker:

\[ p(\text{utterance, perspective} \mid \text{world}) \propto p(\text{world} \mid \text{utterance, perspective}) p(\text{utterance, perspective}) \]

\[ S_1(u, a \mid w) \propto \text{Max}(L_0(w \mid u, a) p(u \mid a) p(a)) \]

where \( w = \text{world}, u = \text{utterance}, a = \text{perspective} \)
Perspectival RSA model

Literal listener:
\[ p(\text{world} | \text{perspective}, \text{utterance}) \propto p(\text{perspective}, \text{utterance} | \text{world}) p(\text{world}) \]

\[ L_0(w | u, a) \propto [[u]]^{w,a} p(w) \]

where \( w = \text{world}, u = \text{utterance}, a = \text{perspective} \)
**Perspectival RSA model**

**Literal listener:**
\[ L_0(w \mid u, a) \propto [[u]]^{w,a} p(w) \]

**Pragmatic speaker:**
\[ S_1(u, a \mid w) \propto \text{Max}(L_0(w \mid u, a) p(u \mid a) p(a)) \]

**Pragmatic listener:**
\[ L_1(w, a \mid u) \propto S_1(u, a \mid w) p(w) \]

where \( w = \text{world}, u = \text{utterance}, a = \text{perspective} \)
Perspectival RSA model

**Perspective Cost Function:** Penalizes non-speaker perspectives. Motivated by egocentricity biases (Lin & Epley 2010, Köder et al. 2015) and Harris (2012)’s Speaker Default heuristic.

**Literal Listener:**
\[ L_0(w \mid u, a) \propto [u]^{w,a} p(w) \]

**Pragmatic Speaker:**
\[ S_1(u, a \mid w) \propto \text{Max}(L_0(w \mid u, a) p(u \mid a) p(a) - \text{Cost}(a)) \]

**Pragmatic Listener:**
\[ L_1(w, a \mid u) \propto S_1(u, a \mid w) p(w) \]
1. Develop a computational model of perspectival reasoning as a joint inference problem

2. Run simulations to predict speaker and listener behavior

3. Compare model predictions against evidence from crowdsourced behavioral experiments
   • How do listeners understand perspectival expressions?
   • How do speakers choose perspectival expressions?

4. Open questions
Thelma is coming.

What is Sam trying to communicate?

Whose perspective is Sam using?
Perspective Set

Lucy Listener

Sam Speaker
Utterance Set

\[
\begin{align*}
\{ \text{Thelma is} \} & \quad \{ \text{coming} \} \quad \{ \text{to Northampton} \} \\
\{ \text{I am} \} & \quad \{ \text{going} \} \\
\{ \text{You are} \} & \quad \{ \text{walking} \}
\end{align*}
\]
Utterance semantics

The literal listener looks up the literal meaning of the utterance according to the perspective and world.

*come*: 1 if the perspective holder is at the destination of motion and the subject is in motion.

*walk*: 1 if the subject is in motion.

*go*: 1 if the subject is in motion and the perspective holder is not at the destination of motion.
World Set

**BOTH**

- Thelma is coming to the market.

**LISTENER**

- Thelma is coming to the market.

**SPEAKER**

- Thelma is coming to the market.

**NONE**

- ??Thelma is coming to the market.
Generating predictions

❖ Model is implemented in the WebPPL probabilistic programming language
❖ Simulations run using Markov Chain Monte Carlo sampling
❖ Tested perspective cost function settings from 0 - 1
Thelma is coming to Northampton
Thelma is going to Northampton
Thelma is walking to Northampton

Simulated with perspectival semantics for *go* and perspective cost settings of 0, 0.5, and 1
Given the sentence *Thelma is coming to Northampton*, the model predicts highest marginal probability for the world where both speaker and listener are located at the destination.
Talk Outline

- Part I: Encoding Perspectives
  - The semantic landscape of perspectival expressions
- Part II: Taking Perspectives
  - Perspective identification
    - A rational approach to perspectival reasoning
    - Model predictions
    - Comprehension studies
  - Perspective selection
    - Production study
- Concluding thoughts
Who else is at the zoo?

Thelma is coming to the zoo.
MODELS OF GRAMMATICAL PERSPECTIVE PROCESSING

- Heuristic approach: listeners use simple rules
  - Speaker Default: assume speakers use their own perspectives whenever possible
  - Maintain Perspective: assume speakers avoid perspective shift

- Perspectival reasoning approach: listeners reason over multiple perspectives
  - Listeners use a range of contextual evidence to infer the speaker’s adopted perspective

- Two-stage system (Harris 2012): listeners rely on heuristics when possible, but switch to a more costly reasoning system when necessary
Perspectival Rational Speech Acts Model: If listeners consider multiple perspectives simultaneously, they should assign highest marginal posterior probability to the world where both listener and speaker are at the destination. Predicts a Convergent Perspective Boost.
HEURISTIC APPROACH

Speaker default heuristic (Harris 2012): If listeners assume that the speaker is the perspective holder, they should assign equal marginal posterior probability to all worlds where the speaker is at the destination.

Predicts a Simple Speaker Advantage.
**Thelma is coming hypotheses**

**Perspectival Rational Speech Acts Model:** If listeners consider multiple perspectives simultaneously, they should assign highest marginal posterior probability to the world where both listener and speaker are at the destination. *Predicts a Convergent Perspective Boost.*

**Speaker default heuristic** (Harris 2012): If listeners assume that the speaker is the perspective holder, they should assign equal marginal posterior probability to all worlds where the speaker is at the destination. *Predicts a Simple Speaker Advantage.*
Thelma is coming

Speaker default heuristic (Harris 2012): If listeners assume that the speaker is the perspective holder, they should assign equal marginal posterior probability to all worlds where the speaker is at the destination. Predicts a convergent perspective boost.

Perspectival rational speech acts model: If listeners consider multiple perspectives simultaneously, they should assign highest marginal posterior probability to the world where both listener and speaker are at the destination. Predicts a convergent perspective boost.

PRSA Speaker Default

Simple speaker advantage

Perspective reselection

No compatible perspective

Convergent perspective boost

Perspective cost determines difference

No compatible perspective
Experimental design

- Eight conditions: 4 worlds x 2 sentences
- **Perspectival condition**: Thelma is coming to the zoo.
- **Plain condition**: Thelma is driving to the zoo.
Thelma is coming to the bank.
Thelma is coming to the bank.
Experimental design

- Eight conditions: 4 worlds x 2 sentences
- 6 items in each condition
- **Perspectival condition:** Thelma is coming to the zoo.
- **Plain condition:** Thelma is driving to the zoo.
- Dependent variables: reaction times and acceptability.
- Monolingual American English-speaking participants recruited on Prolific
- 3 experiments:
  - Expt. 1A: n= 80, Expt. 1B: n=64, Expt. 1C: n=120
The cat on the chair closest to you is orange.
Reaction Time Results

Experiment 1a: By-participant mean RT differences

Experiment 1b: By-participant mean RT differences

Experiment 1c: By-participant mean RT differences

Perspectival – non-perspectival RT
Acceptability Results

1A

1B

1C
Comprehension Results: $p(w|u)$

Summary:

- Participants were fastest to accept scenes following *come* in the both condition.
- The acceptance rate in the listener *come* condition was low, high speaker bias.
1. Develop a computational model of perspectival reasoning as a joint inference problem

2. Run simulations to predict speaker and listener behavior

3. Compare model predictions against evidence from crowdsourced behavioral experiments
   - How do listeners understand perspectival expressions?
   - How do speakers choose perspectival expressions?

4. Open questions
Perspective Selection: $p(a \mid w)$

Thelma is coming to the zoo.

??

Thelma is going to the zoo.
Experimental design

❖ Experiment 2A: 4 conditions (Both, Speaker, Listener, None)
❖ Experiment 2B: 6 conditions (same 4 + Speaker Moving Listener, Speaker Moving None)
❖ 12 items in each condition
❖ Participants completed the speaker's speech bubble in a text box.
❖ Prompt: Thelma is...
❖ Monolingual American English-speaking participants recruited on Prolific
❖ 2 experiments:
  ❖ Expt. 2A: n= 40, Expt. 2B: n=56
Thelma is coming

hypotheses

Speaker default heuristic (Harris 2012): If listeners assume that the speaker is the perspective holder, they should assign equal marginal posterior probability to all worlds where the speaker is at the destination.

Predicts a Convergent Perspective Boost.

Perspectival Rational Speech Acts Model: If listeners consider multiple perspectives simultaneously, they should assign highest marginal posterior probability to the world where both listener and speaker are at the destination.

Predicts a Convergent Perspective Boost.

PRSA Speaker Default

Perspective Cost Determines Difference

No Compatible Perspective

Speaker's Perspective Compatible

Speaker's Perspective Incompatible
Production Results

**PRSA**

**Speaker Default**

**Experiment 2a Results**

**Experiment 2b Results**
Speaker Moving Conditions

**Speaker Moving Listener**

- House
- Person (S) saying, "I am coming to the market."
- Market stand

**Speaker Moving None**

- House
- Person (S) saying, "I am coming to the market."
- Market stand
Speaker Default

Speaker's Perspective Incompatible

PRSA

Ambiguity Elimination Advantage

Compatible But Ambiguous

No Compatible Perspective
## Production Results

### Expt. 2b Results

<table>
<thead>
<tr>
<th>Utterance</th>
<th>Listener</th>
<th>Speaker Moving Listener</th>
<th>Speaker Moving None</th>
</tr>
</thead>
<tbody>
<tr>
<td>Come</td>
<td><img src="image1" alt="Graph" /></td>
<td><img src="image2" alt="Graph" /></td>
<td><img src="image3" alt="Graph" /></td>
</tr>
<tr>
<td>Go</td>
<td><img src="image4" alt="Graph" /></td>
<td><img src="image5" alt="Graph" /></td>
<td><img src="image6" alt="Graph" /></td>
</tr>
<tr>
<td>Walk</td>
<td><img src="image7" alt="Graph" /></td>
<td><img src="image8" alt="Graph" /></td>
<td><img src="image9" alt="Graph" /></td>
</tr>
</tbody>
</table>

### Marginal Probability: $p(u|w)$

<table>
<thead>
<tr>
<th>Utterance</th>
<th>Listener</th>
<th>Speaker Moving Listener</th>
<th>Speaker Moving None</th>
</tr>
</thead>
<tbody>
<tr>
<td>Go</td>
<td><img src="image10" alt="Graph" /></td>
<td><img src="image11" alt="Graph" /></td>
<td><img src="image12" alt="Graph" /></td>
</tr>
<tr>
<td>Walk</td>
<td><img src="image13" alt="Graph" /></td>
<td><img src="image14" alt="Graph" /></td>
<td><img src="image15" alt="Graph" /></td>
</tr>
</tbody>
</table>

### Raw Response Counts

<table>
<thead>
<tr>
<th>Utterance</th>
<th>Listener</th>
<th>Speaker Moving Listener</th>
<th>Speaker Moving None</th>
</tr>
</thead>
<tbody>
<tr>
<td>Come</td>
<td><img src="image16" alt="Graph" /></td>
<td><img src="image17" alt="Graph" /></td>
<td><img src="image18" alt="Graph" /></td>
</tr>
<tr>
<td>Go</td>
<td><img src="image19" alt="Graph" /></td>
<td><img src="image20" alt="Graph" /></td>
<td><img src="image21" alt="Graph" /></td>
</tr>
<tr>
<td>Manner</td>
<td><img src="image22" alt="Graph" /></td>
<td><img src="image23" alt="Graph" /></td>
<td><img src="image24" alt="Graph" /></td>
</tr>
</tbody>
</table>

---

**PRSA**

**Speaker Default**

**Expt. 2b Results**
Converging Evidence: Watson et al. (2021)

- Adapts a probabilistic model of visual perspective-taking from Heller et al. 2016 and Mozuraitis et al. 2018 to grammatical perspective-taking
- 4x2 online production experiment: same world conditions (Both, Speaker, Listener, None) x speaker certainty about the listener's location
- Results support a Convergent Perspective Boost

Summary

Since speakers do shift perspective, listeners must reason about the perspective they’ve adopted.

❖ The comprehension data supports the predicted reasoning behavior of rational listeners.

Yet speakers seem to be guided by heuristics like Speaker Default more strongly than listeners expect.

❖ The production data suggests that speakers don’t shift perspective at the rate predicted by listeners.

This goes against a fundamental assumption of the RSA framework: that speakers and listeners learn from and mirror each other’s behavior.

But it mirrors Kehler & Rohde (2019)'s proposal for pronoun resolution.

It is also compatible with Harris (2012)'s proposed two-stage system: listeners may use the costly reasoning system more than speakers.
Acknowledgements

Data Coders:

Tessa Masis
UMass Amherst

Carina Bolaños Lewen
Wesleyan College

Brian Dillon
UMass Amherst
Co-author