

# *Pragmatic Language Games*

Daniel Fried

with slides from Nick Tomlin

# Natural Language Interfaces

## Science Fiction *Her*, 2013

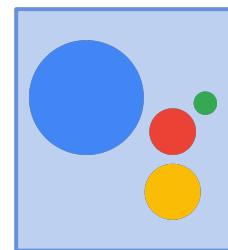


Let's start with your emails. You have several thousand emails regarding LA Weekly, but it looks like you haven't worked there in many years.

Oh yeah, I guess I was saving those because in some of them I thought I might have written some funny stuff.

Yeah, there are some funny ones. I'd say there are about 86 that we should save. We can delete the rest.

## In Reality Google Assistant, 2017



Who are you?

I'm your Google Assistant.

And I can let you know if you'll need a jacket today.

Do I?

Sorry, I don't understand.

# Context in NLP

## Other Language

Language Modeling,  
Structure & Semantics



Write With Transformer distil-gpt2 ⓘ

Understanding searches better  
than ever before

Pandu Nayak  
Google Fellow and Vice President, Search

## This Lecture

### The World Grounding



*“Take me to the airport”*

### Intents and Effects Pragmatics



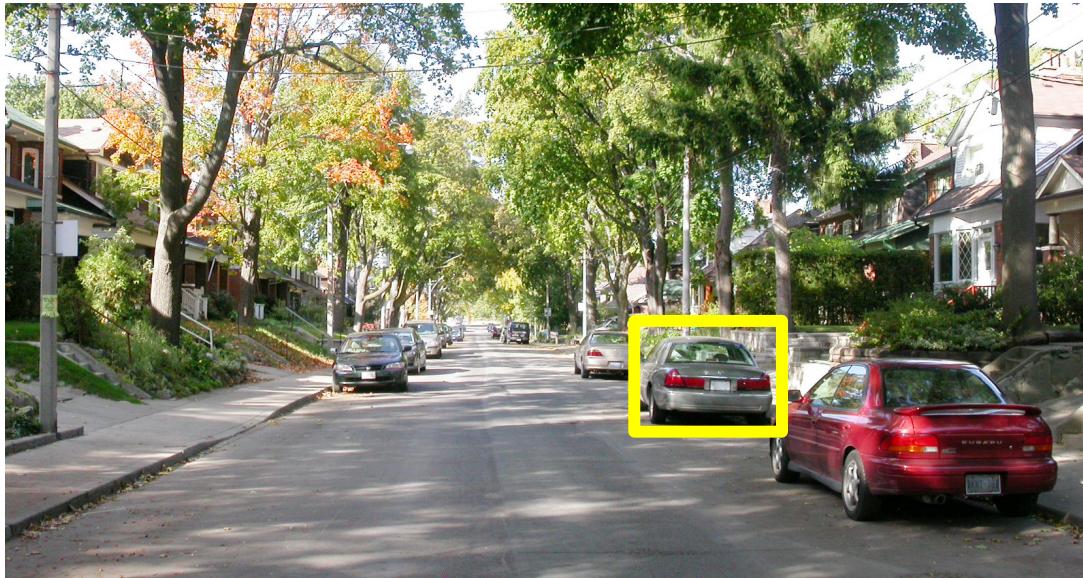
*“My neck hurts”*

# Grounding and Pragmatics

---

Grounding

*“Stop at the second car”*



Pragmatics

*“Stop at the car”*



# Pragmatics and Reasoning

---

Saying something will often... produce certain consequential effects upon the feelings, thoughts, or actions of the audience.

[*How to Do Things with Words*. Austin, 1962]

Our talk exchanges ... are cooperative efforts... One of my avowed aims is to see talking as purposive, indeed rational, behavior.

[*Logic and Conversation*. Grice, 1975]

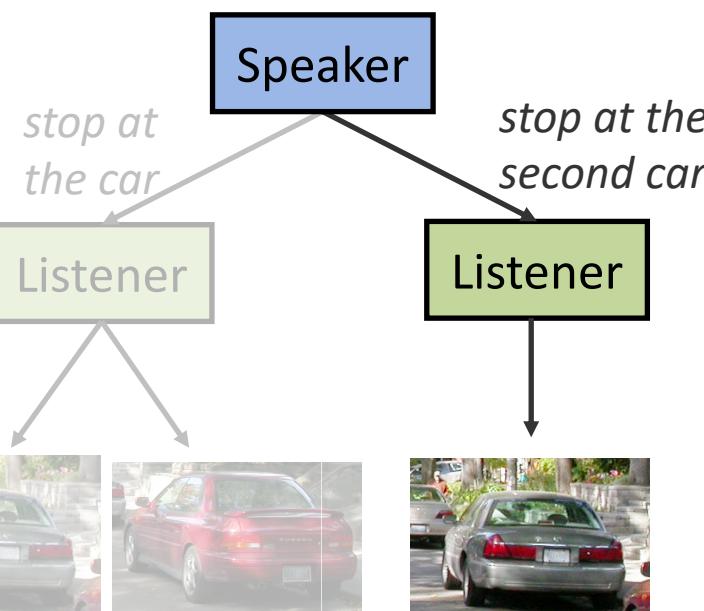
**Language is an act people take to produce effects on others and the world!**

# Pragmatics and Reasoning

## Generation



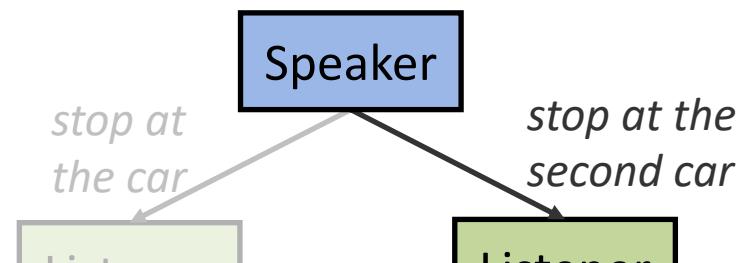
## Interpretation



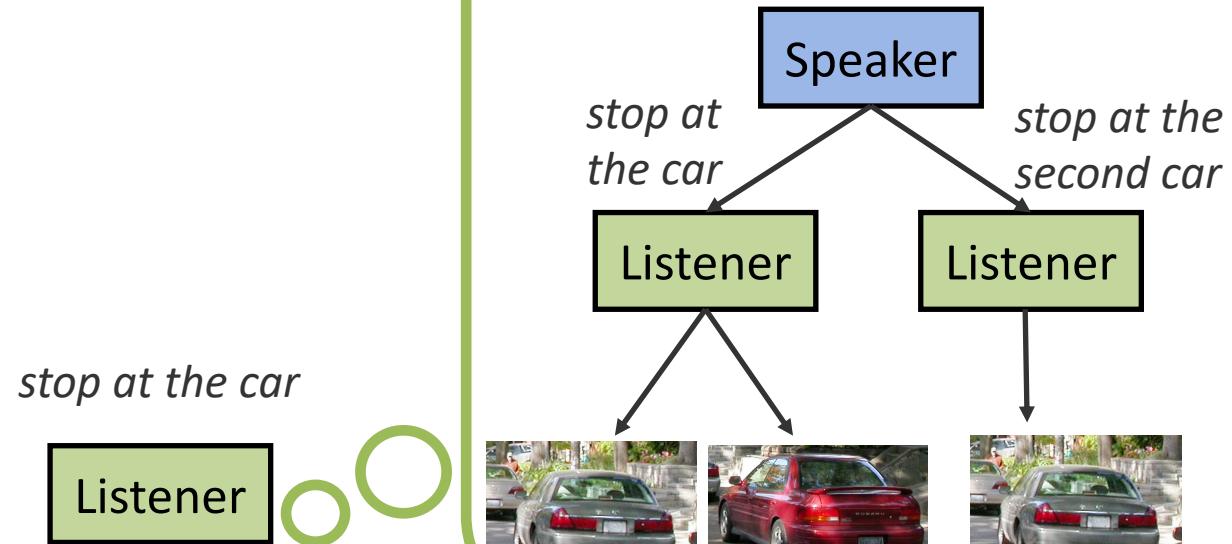
[e.g. Lewis 1969; Golland et al. 2010;  
Frank and Goodman 2012; Degen et al. 2013]

# Pragmatics and Reasoning

## Generation



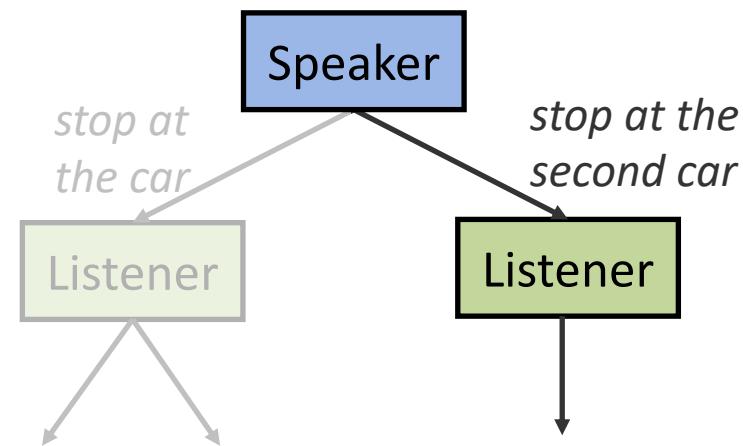
## Interpretation



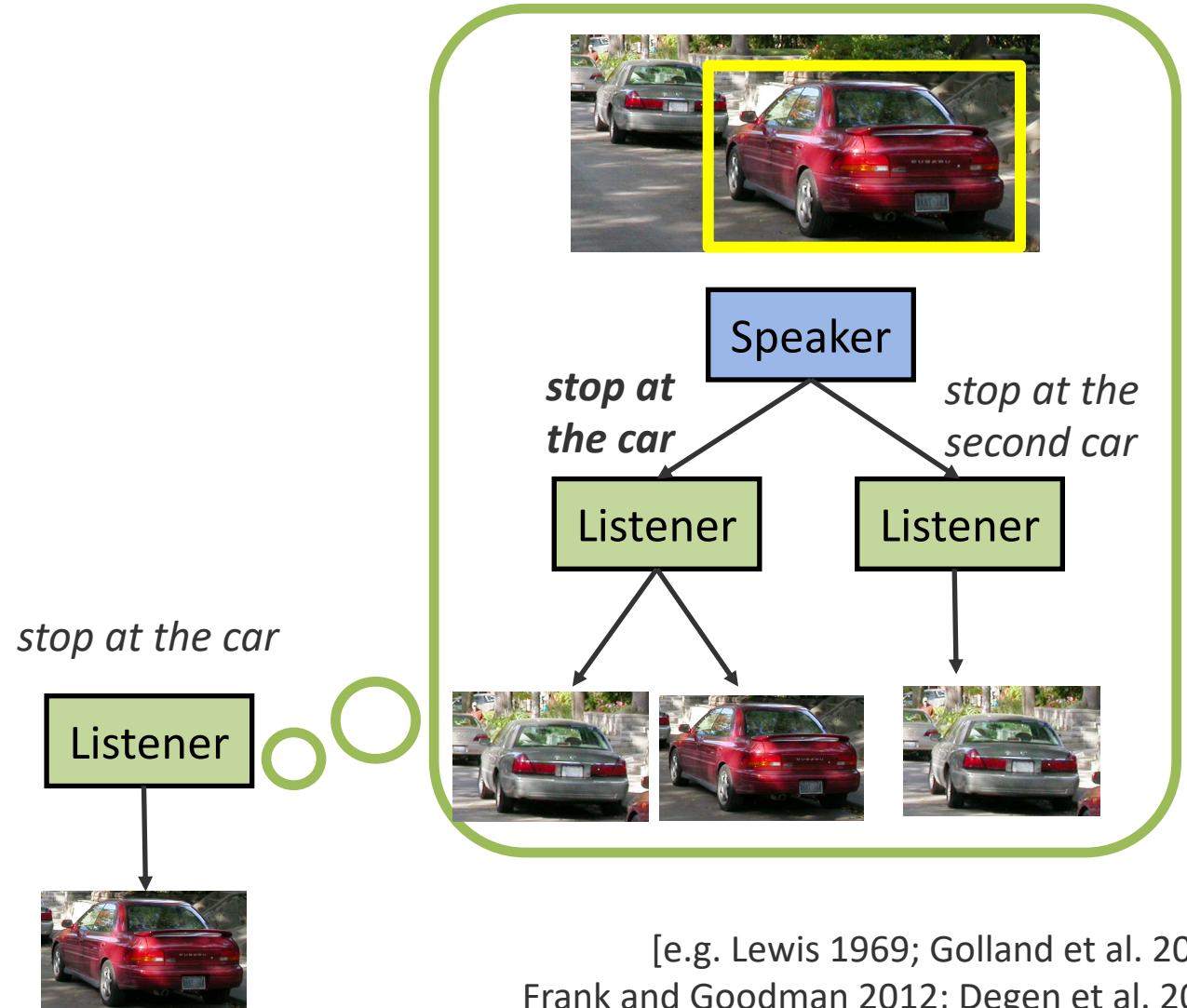
[e.g. Lewis 1969; Golland et al. 2010;  
Frank and Goodman 2012; Degen et al. 2013]

# Pragmatics and Reasoning

Generation



Interpretation



[e.g. Lewis 1969; Golland et al. 2010;  
Frank and Goodman 2012; Degen et al. 2013]

# Reasoning About Alternatives

---

Core Idea:

*Large chunks of linguistic understanding can be attributed to reasoning about alternatives. E.g., if a speaker says X but not Y, then perhaps Y isn't true, or the speaker doesn't want to talk about Y.*

# Reasoning About Alternatives

---

Core Idea:

*Large chunks of linguistic understanding can be attributed to reasoning about alternatives. E.g., if a speaker says X but not Y, then perhaps Y isn't true, or the speaker doesn't want to talk about Y.*

Example:

“I didn’t steal your car.”

# Reasoning About Alternatives

---

Core Idea:

*Large chunks of linguistic understanding can be attributed to reasoning about alternatives. E.g., if a speaker says X but not Y, then perhaps Y isn't true, or the speaker doesn't want to talk about Y.*

Example:

**“I didn’t steal your car.”**

Conveyed meaning:

*Someone stole your car, but it wasn't me.*

# Reasoning About Alternatives

---

Core Idea:

*Large chunks of linguistic understanding can be attributed to reasoning about alternatives. E.g., if a speaker says X but not Y, then perhaps Y isn't true, or the speaker doesn't want to talk about Y.*

Example:

“I didn’t steal your car.”

Conveyed meaning:

*Contrary to what you think, I did not steal your car.*

# Reasoning About Alternatives

---

Core Idea:

*Large chunks of linguistic understanding can be attributed to reasoning about alternatives. E.g., if a speaker says X but not Y, then perhaps Y isn't true, or the speaker doesn't want to talk about Y.*

Example:

“I didn’t steal your car.”

Conveyed meaning:

*I did something to your car, but not stealing it. E.g., I just borrowed it.*

# Reasoning About Alternatives

---

Core Idea:

*Large chunks of linguistic understanding can be attributed to reasoning about alternatives. E.g., if a speaker says X but not Y, then perhaps Y isn't true, or the speaker doesn't want to talk about Y.*

Example:

“I didn’t steal your car.”

Conveyed meaning:

*I stole somebody else's car.*

# Reasoning About Alternatives

---

Core Idea:

*Large chunks of linguistic understanding can be attributed to reasoning about alternatives. E.g., if a speaker says X but not Y, then perhaps Y isn't true, or the speaker doesn't want to talk about Y.*

Example:

“I didn’t steal your car.”

Conveyed meaning:

*I stole something you own, but not your car.*

# Implicatures

---



The New York Times @nytimes



We've deleted an earlier tweet and updated a sentence in our article that implied that only "some experts" view the ingestion of household disinfectants as dangerous. To be clear, there is no debate on the danger.

9:17 AM · Apr 24, 2020 · [Twitter Web App](#)

---

**4.7K** Retweets    **22K** Likes

# Implicatures

---

Q: Does *some* mean *not all*?

A: Not always:

- “Some of the students were late for class; in fact, they all were.”
- “I’d be much happier if some grocery stores had eggs in stock.”

We call this *implicature*. The implicature occurs because a rational listener might assume that the speaker would have said *all* if they meant to, since *all* is the more informative choice.

# Implicatures

---

“The car was stolen.”

- *The speaker doesn't know, or doesn't want to tell, who stole it.*

“Did you invite Alice and Bob?” // “I invited Alice.”

- *The speaker didn't invite Bob.*

“I'm out of gas.” // “There's a station round the corner.”

- *You can get gas there (e.g. it's open).*

“He overslept and failed the test.”

- *Those events happened in that order.*

# What Are People's Goals in Conversation?

---

Grice (1975) claims that many of these phenomena are explained by the tensions between the following **maxims**:

1. **Quantity** – be as informative as possible, give as much information as needed, but no more. (“*The car was stolen.*”)
2. **Quality** - be truthful, and don't give information that is false or unsupported by evidence. (“*Did you invite A and B?*” // “*I invited B.*”)
3. **Relation** – be relevant, and say things that are pertinent to the discussion. (“*I'm out of gas*” // “*There's a station round the corner.*”)
4. **Manner** – be clear, brief, and orderly as possible; avoid unnecessary prolixity. (“*He overslept and failed the test.*”)

# The Cooperative Principle

---

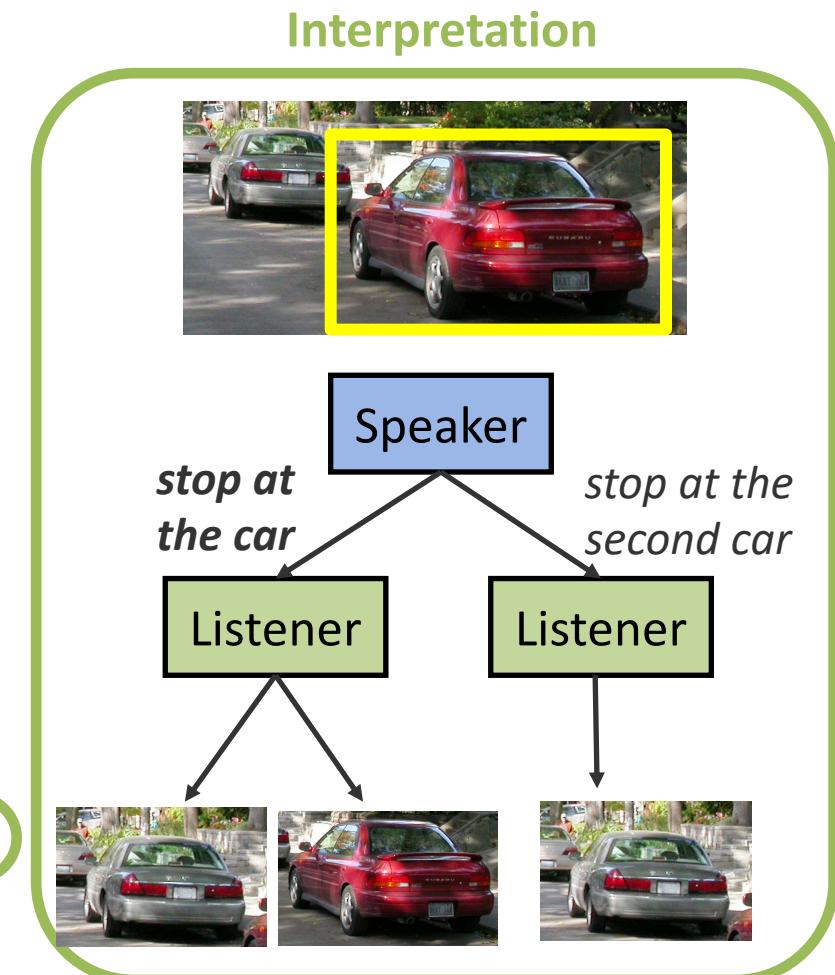
The Cooperative Principle (Grice 1975):

*“Make your 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.”*

**Language is a rational action in a cooperative game.**

# Cooperative Principle via Game Theory

- **Best-response** [Franke 2009; Golland 2010; Jäger 2014]
- Recursive Bayesian agents
  - RSA [Frank and Goodman 2012, 2016]
  - Reward-rational implicit choice [Jeon et al. 2020]
- Other formalisms (info-theoretic):
  - Optimal transport of beliefs [Wang et al. 2020]
  - Rate distortion [Zaslavsky et al. 2020]



# Reasoning with Speakers and Listeners

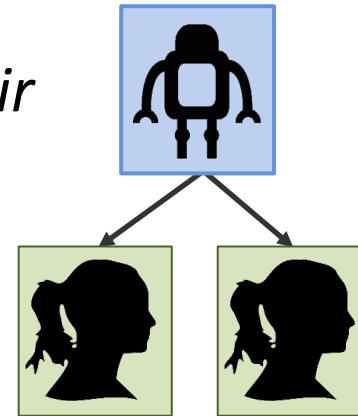
---



# Pragmatics and Generation

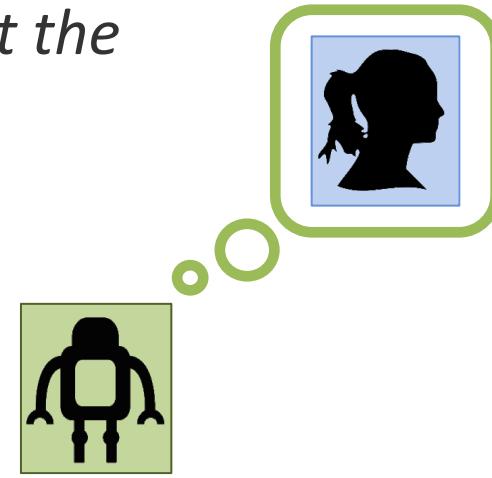


*walk along the wood path to the chair*

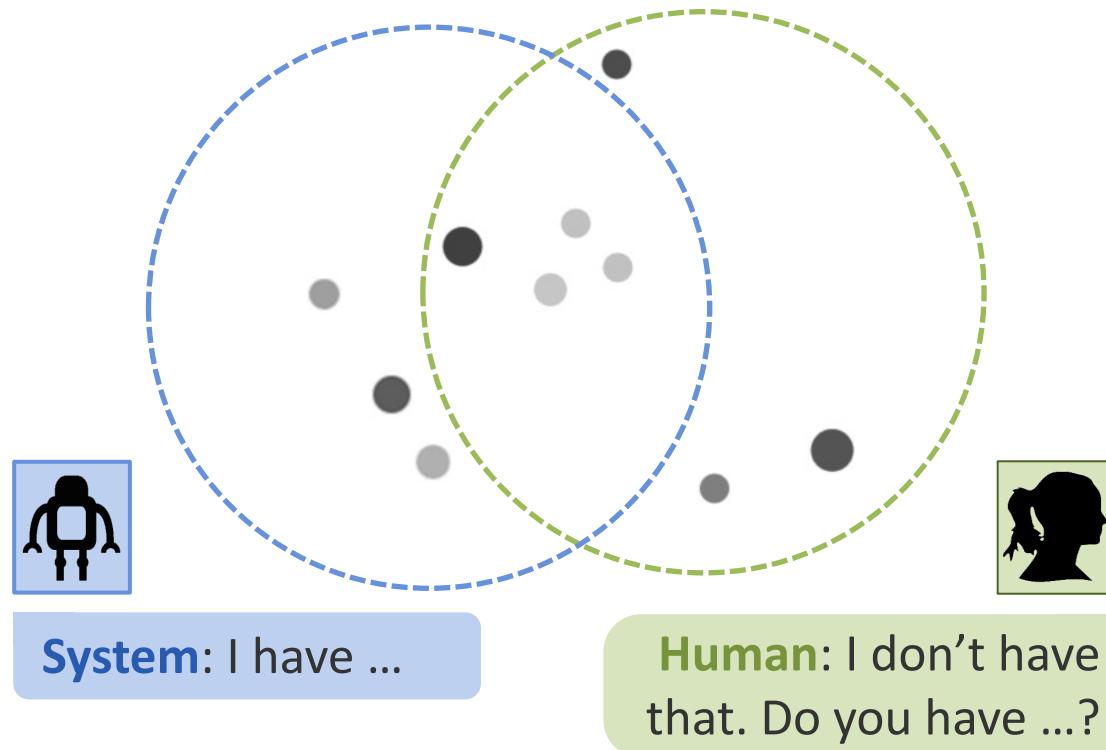


# Pragmatics and Interpretation

*Turn left and take a right at the table. Take a left at the painting and then take your first right.*



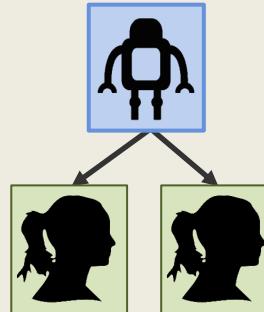
# Pragmatics and Dialogue



# *Pragmatics and...*

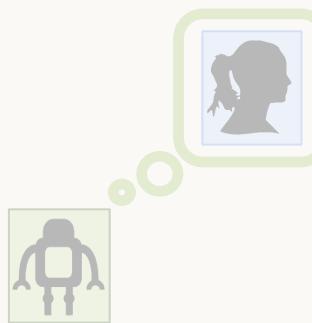
## Generation

[Fried, Andreas, & Klein. 2018]



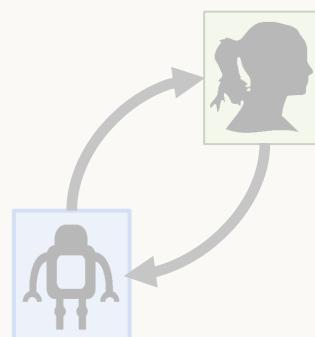
## Interpretation

[Fried\*, Hu\*, Cirik\* et al. NeurIPS 2018]



## Dialogue

[Fried, Chiu, & Klein. EMNLP 2021]



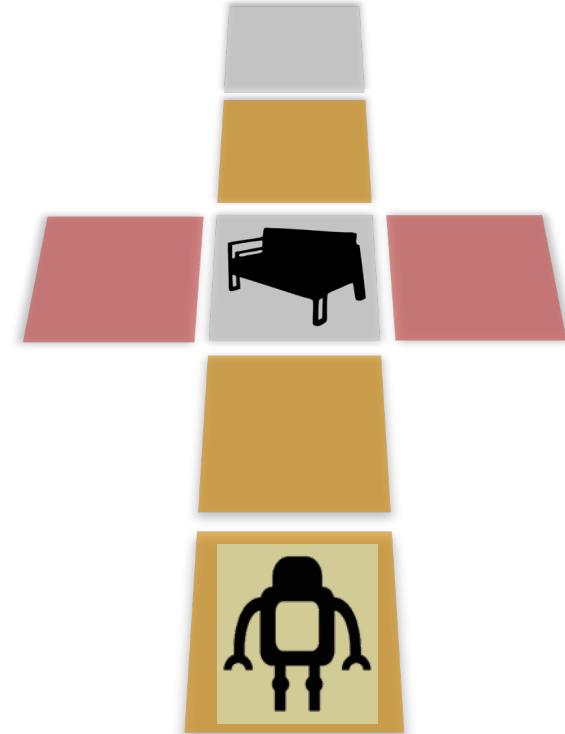
# To Start: Virtual Environments

---

Human View:



Agent View:



SAIL [MacMahon et al., 2006; Chen and Mooney, 2011]

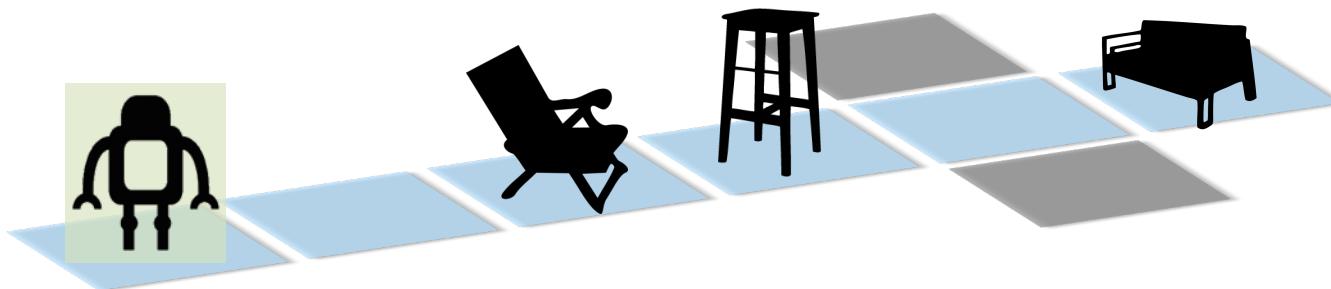
# Interpretation Task

---

Input  
instruction:

*go forward to the grey hallway*

Output  
actions:



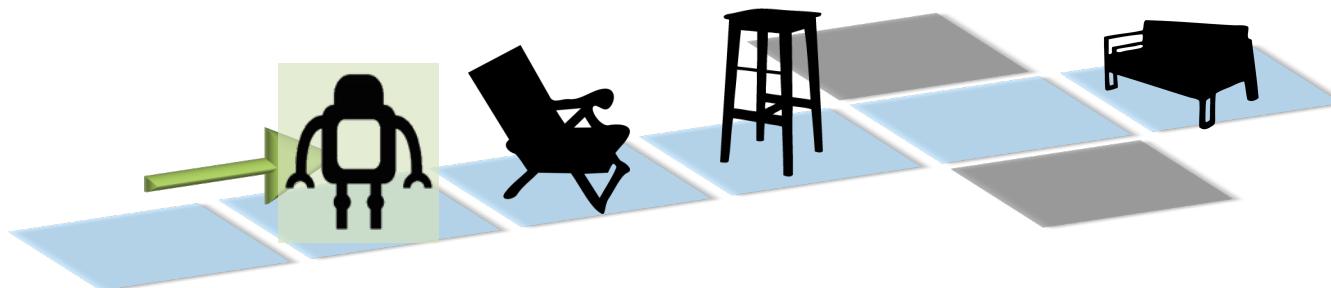
# Interpretation Task

---

Input  
instruction:

*go forward to the grey hallway*

Output  
actions:



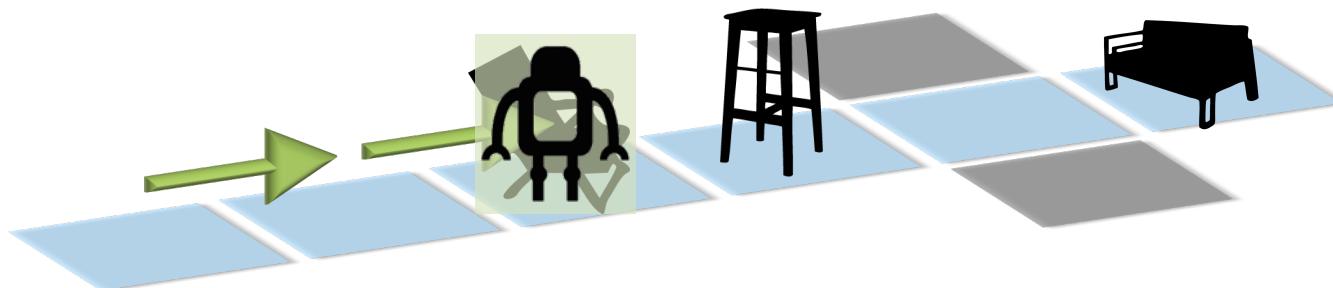
# Interpretation Task

---

Input  
instruction:

*go forward to the grey hallway*

Output  
actions:



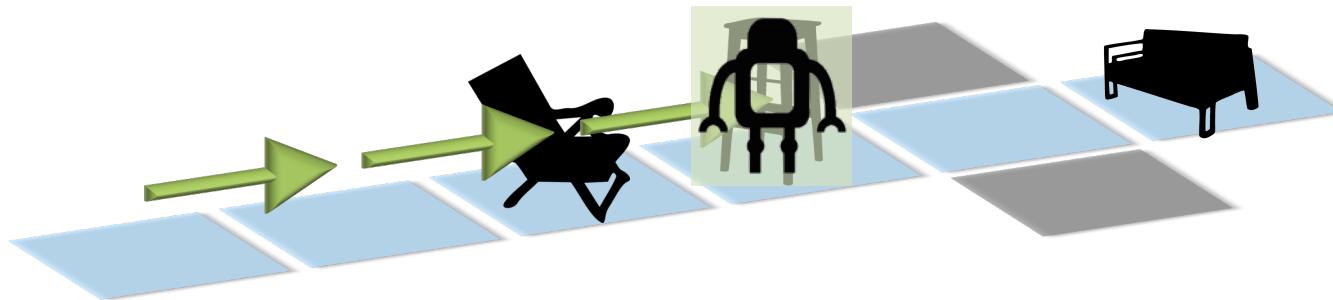
# Interpretation Task

---

Input  
instruction:

*go forward to the grey hallway*

Output  
actions:



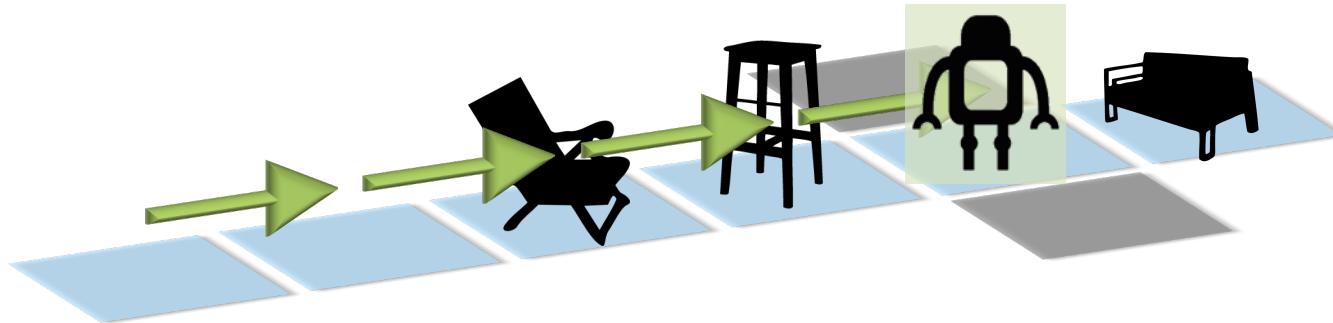
# Interpretation Task

---

Input  
instruction:

*go forward to the grey hallway*

Output  
actions:

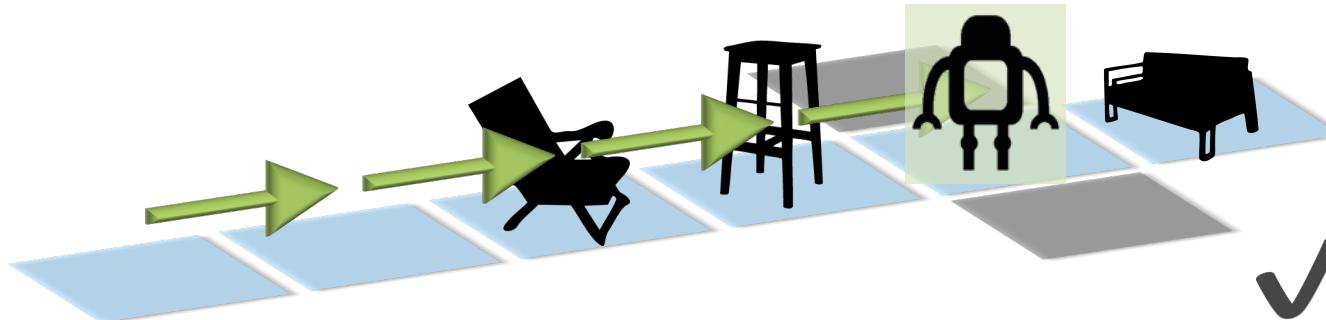


# Interpretation Task

---

Input  
instruction: *go forward to the grey hallway*

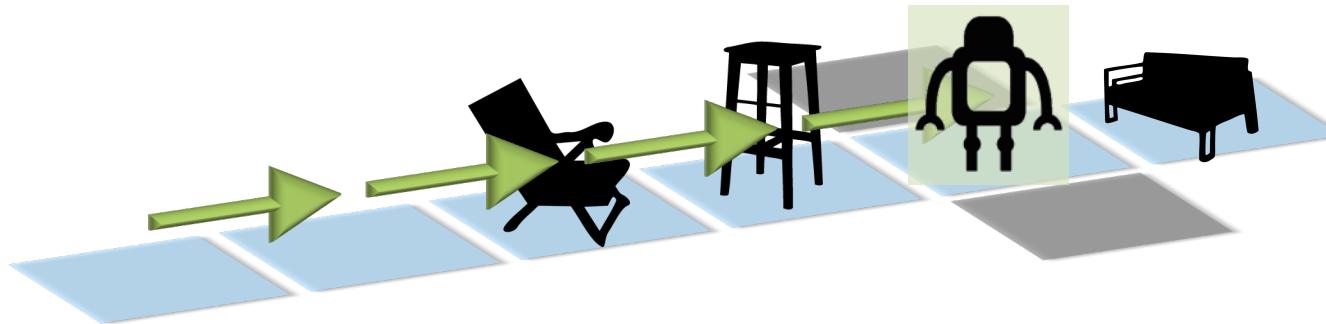
Output  
actions:



# Generation Task

---

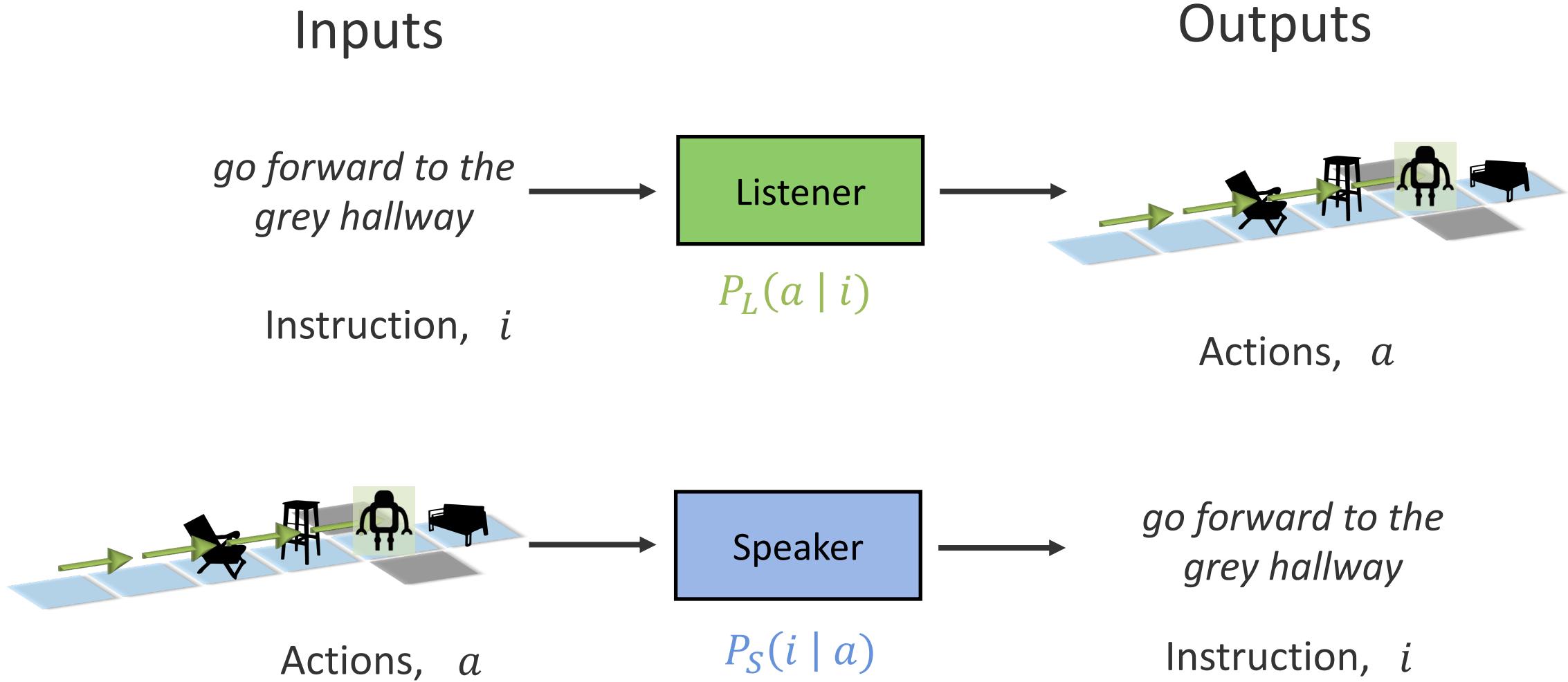
Input  
actions:



Output  
Instruction:

*go forward to the grey hallway*

# Models of Listeners and Speakers

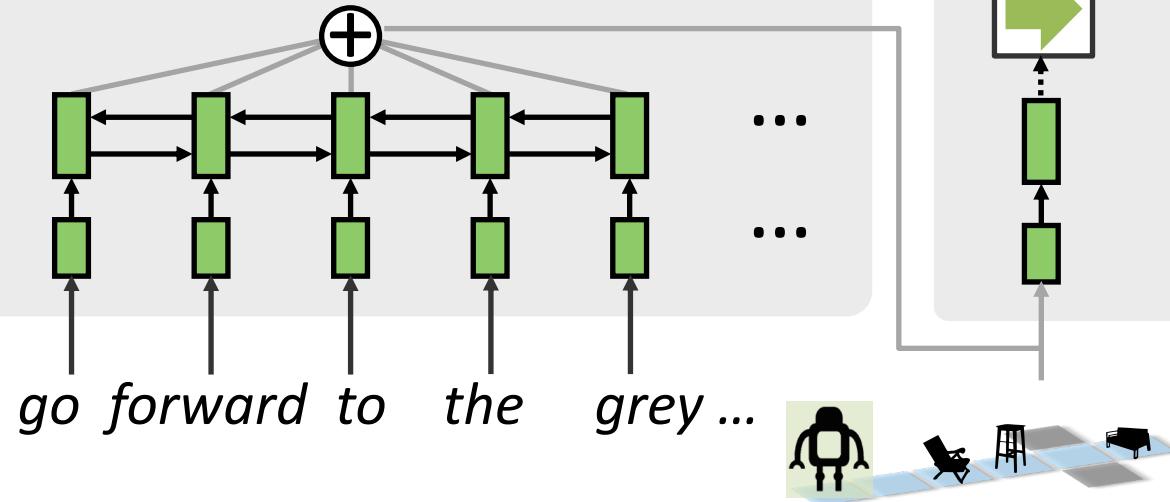


# Base Models

Base  
Listener

$$P_L(a | i) = \prod_t P_L(a_t | a_{1:t-1}, i)$$

LSTM Encoder



LSTM Decoder

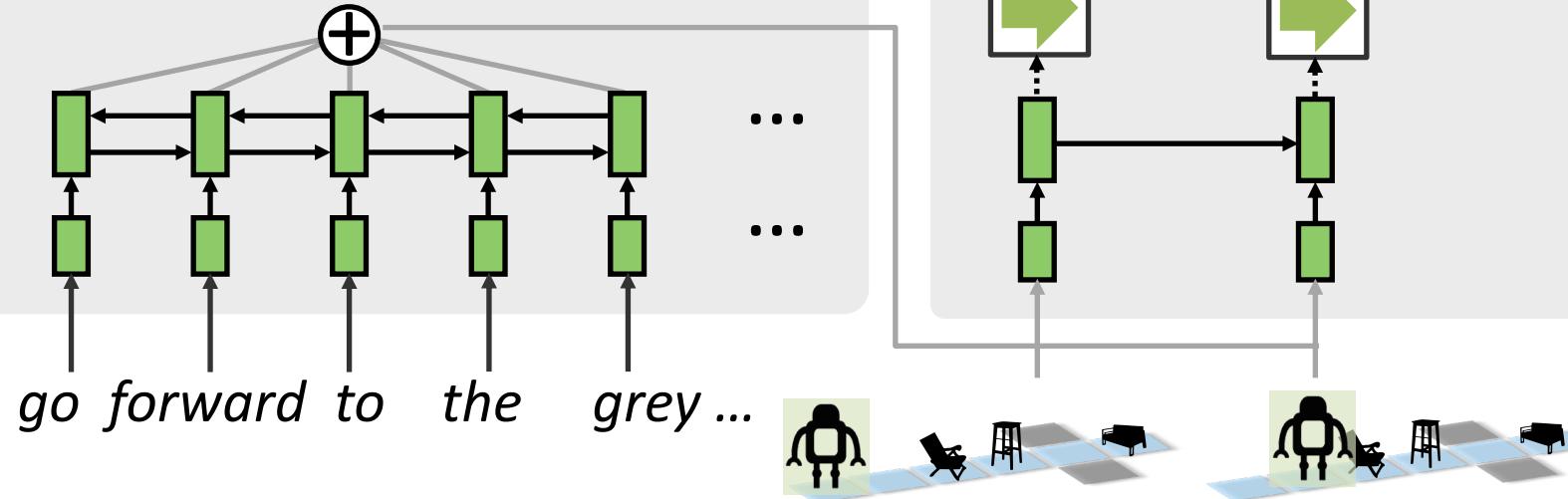
$$P_L(a_1|i)$$

# Base Models

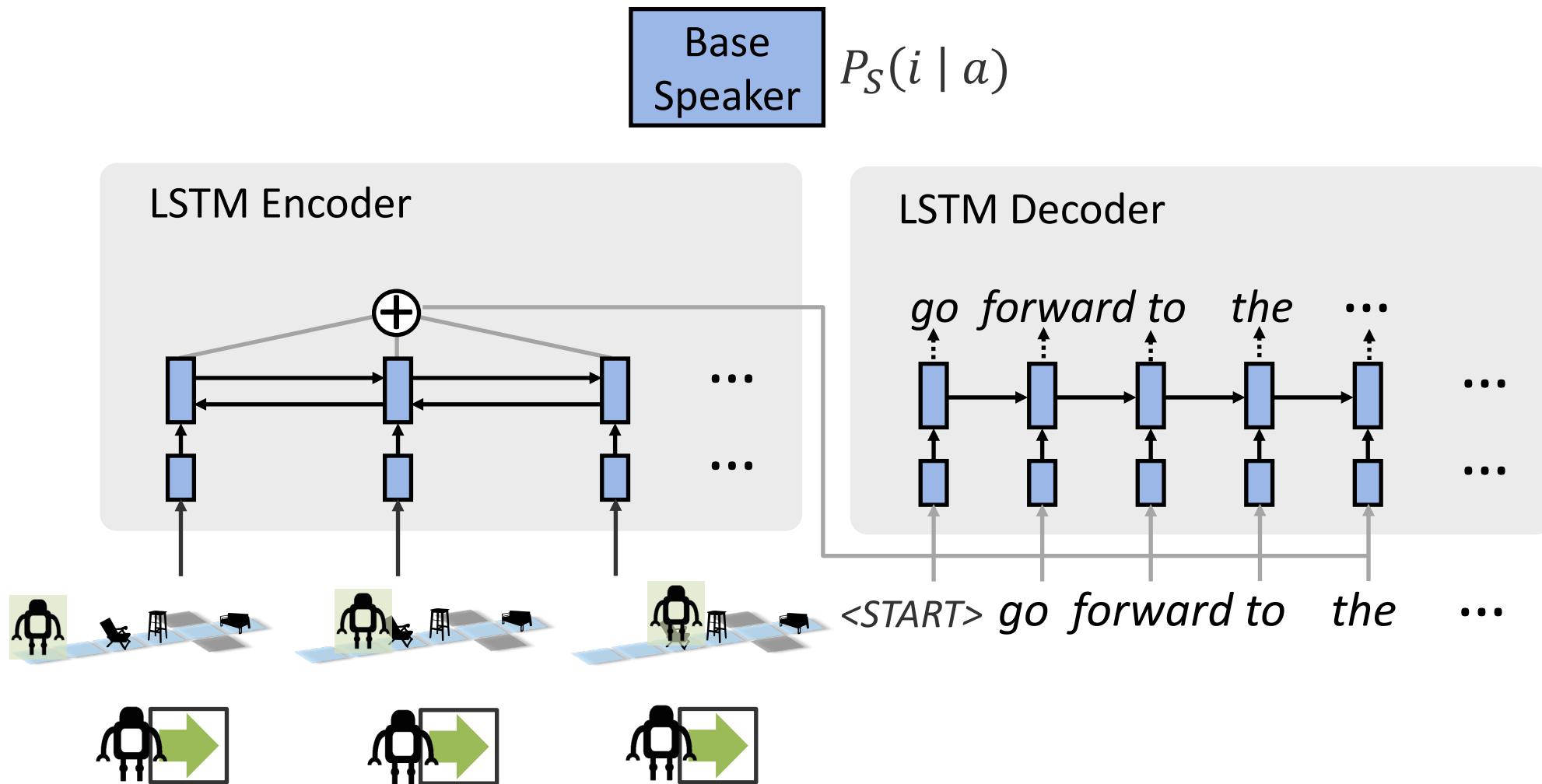
Base  
Listener

$$P_L(a | i) = \prod_t P_L(a_t | a_{1:t-1}, i)$$

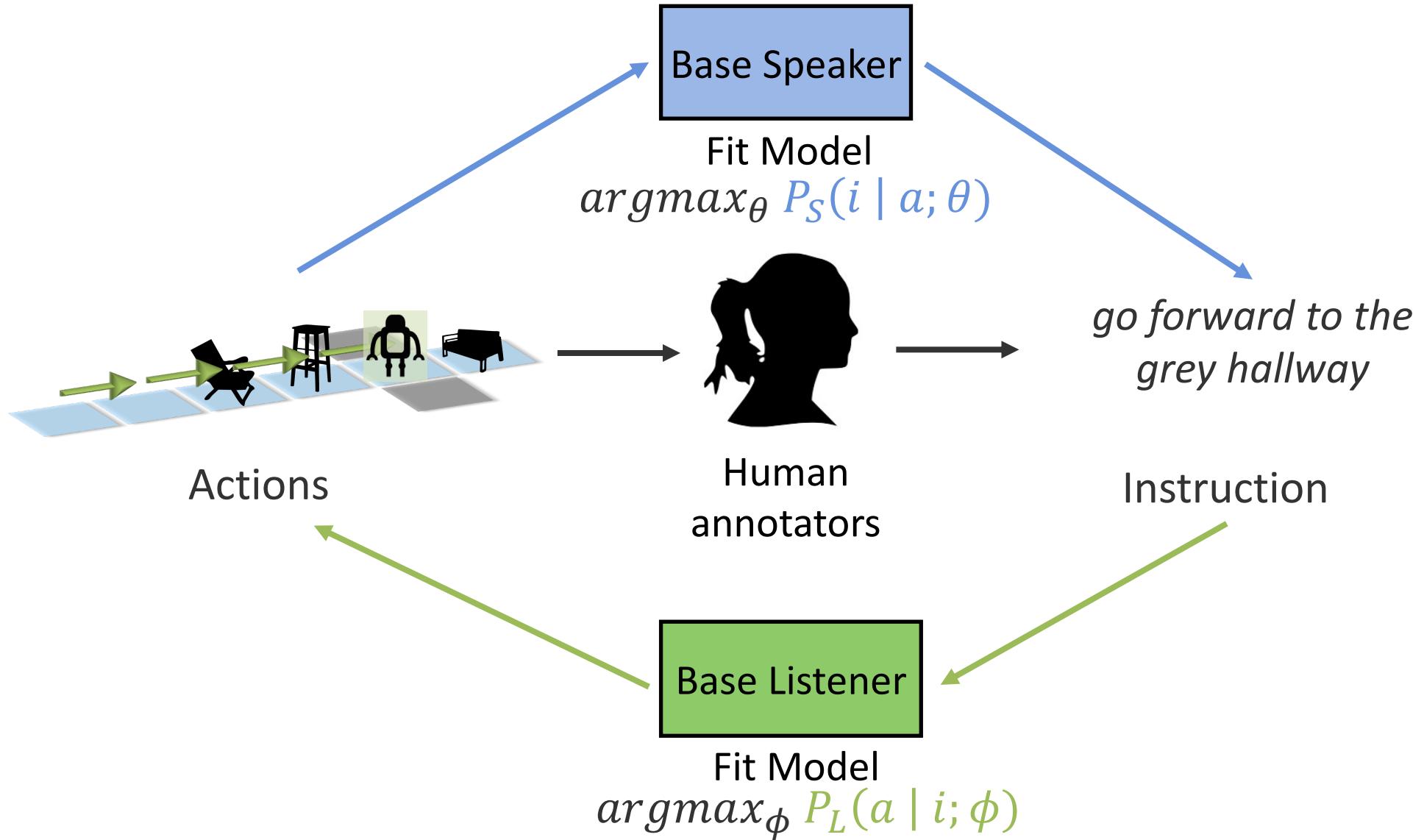
LSTM Encoder



# Base Models

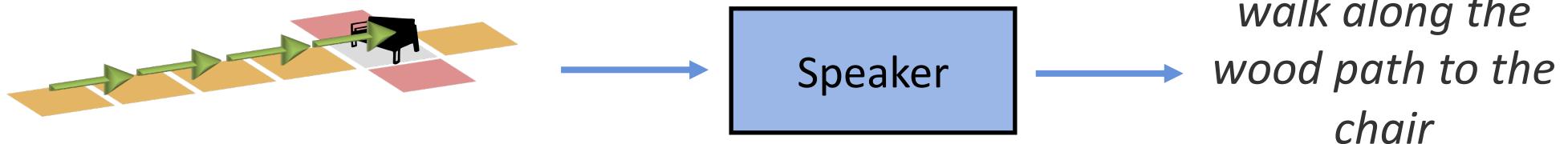


# Training Models on Human Instructions

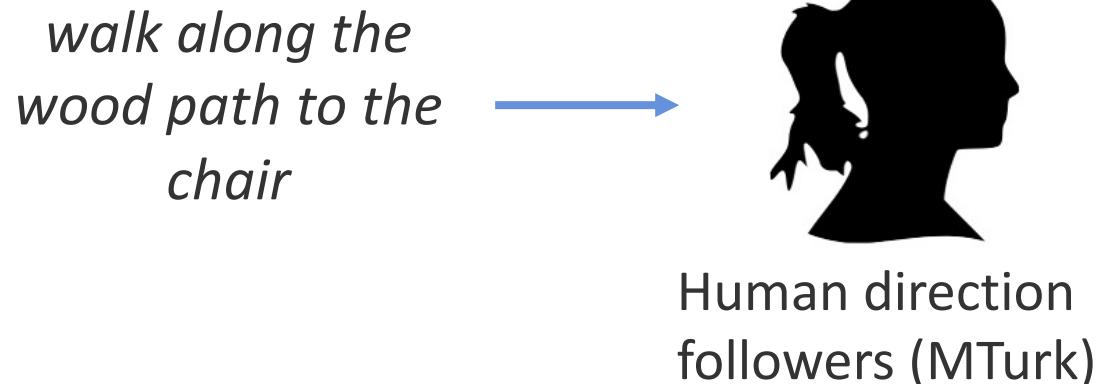


# Speaker Tasks and Evaluation

Speaker produces an instruction



Humans try to interpret it

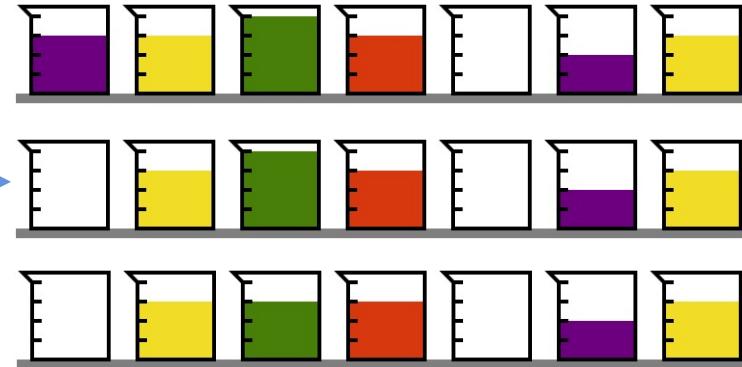


SAIL navigation [MacMahon et al., 2006; Chen and Mooney, 2011]

# Speaker Tasks and Evaluation

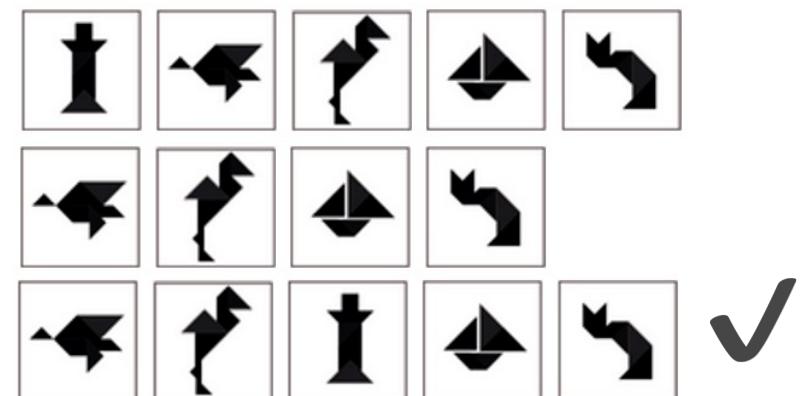
## Alchemy

1. remove all the purple chemical from the beaker on the far left
2. do the same with one unit of green chemical
3. ...



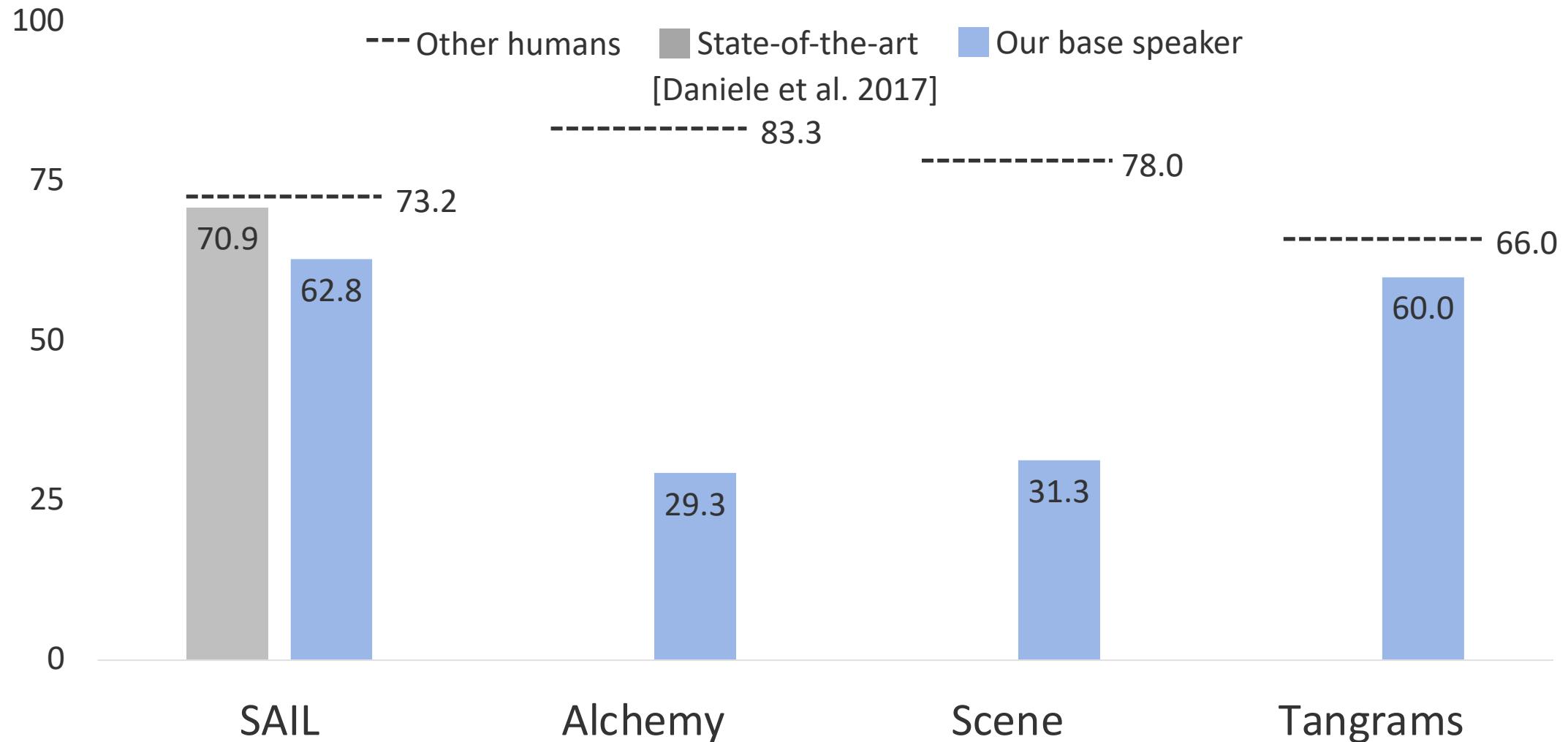
## Tangrams

1. remove first figure
2. add it back into middle spot
3. ...



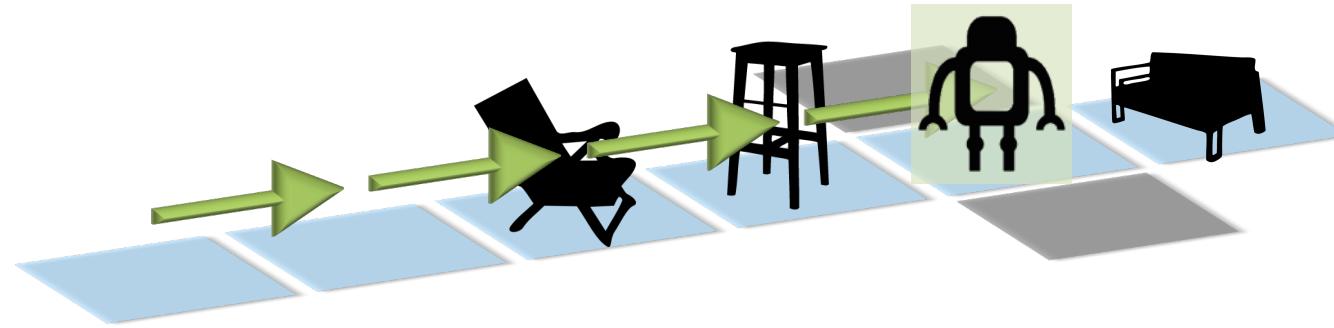
# Generation is Hard to Imitate!

Human accuracy at following instructions from:



# A Failure Mode: Underspecification

---

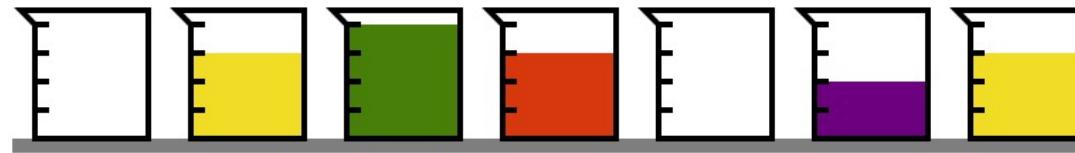
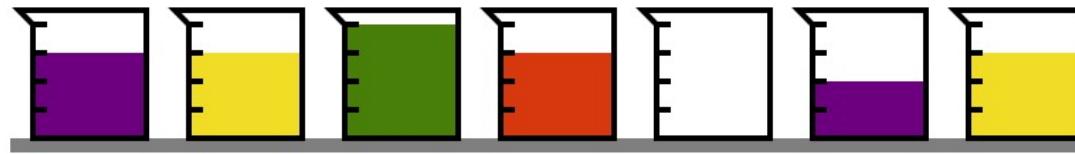


Base  
Speaker

*go forward past the stool ?*

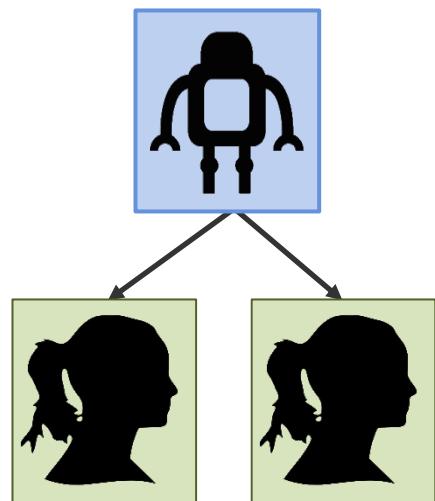
# A Failure Mode: Contextual Ambiguity

---



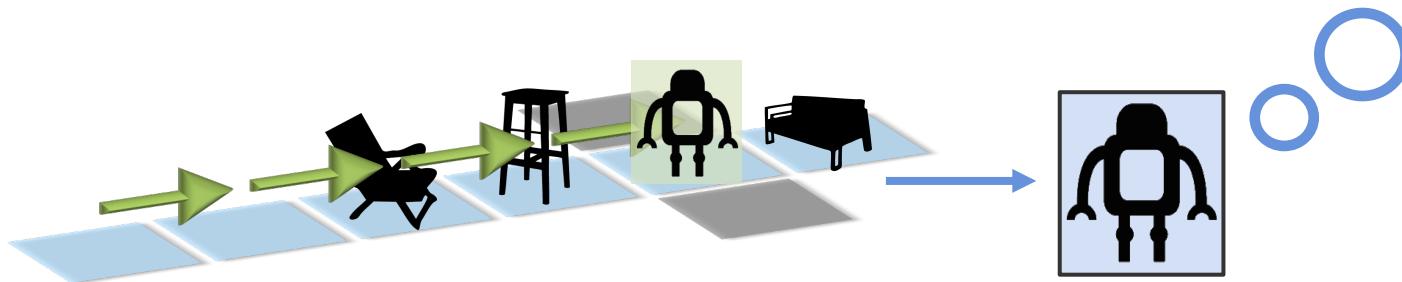
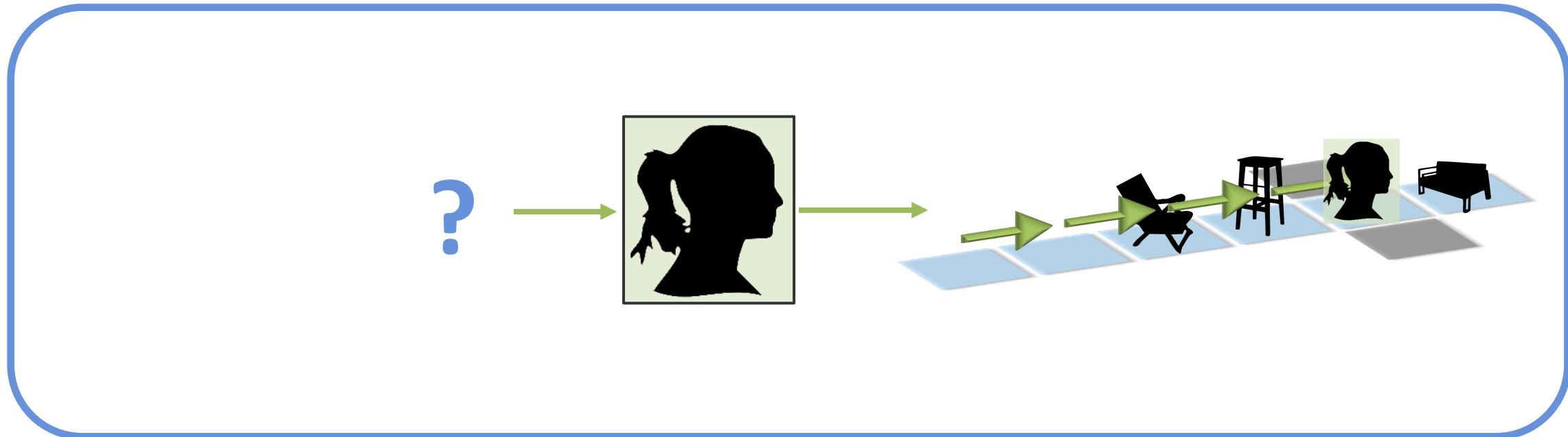
Base  
Speaker

*throw out the purple chemical* X



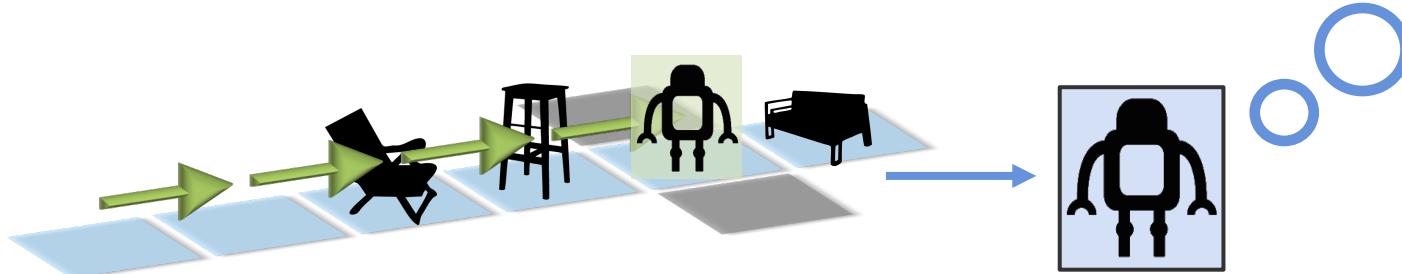
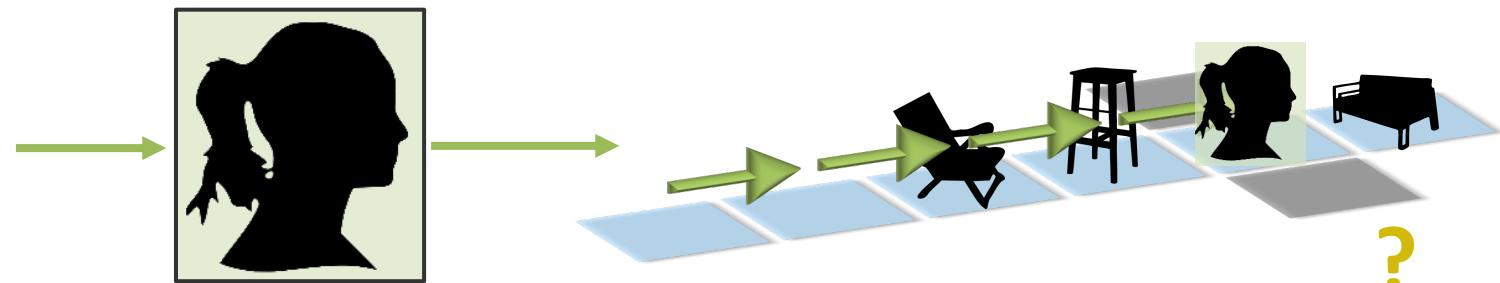
*Making Text Informative  
with Pragmatic Speakers*

# Pragmatic Speakers Simulate Interpretation



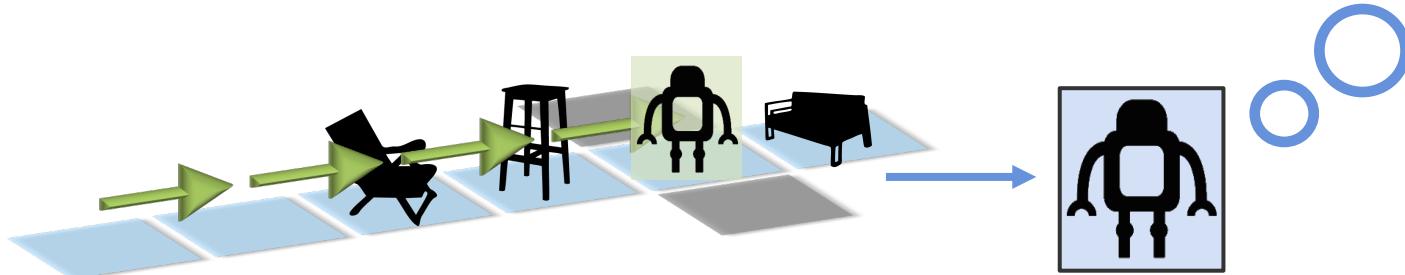
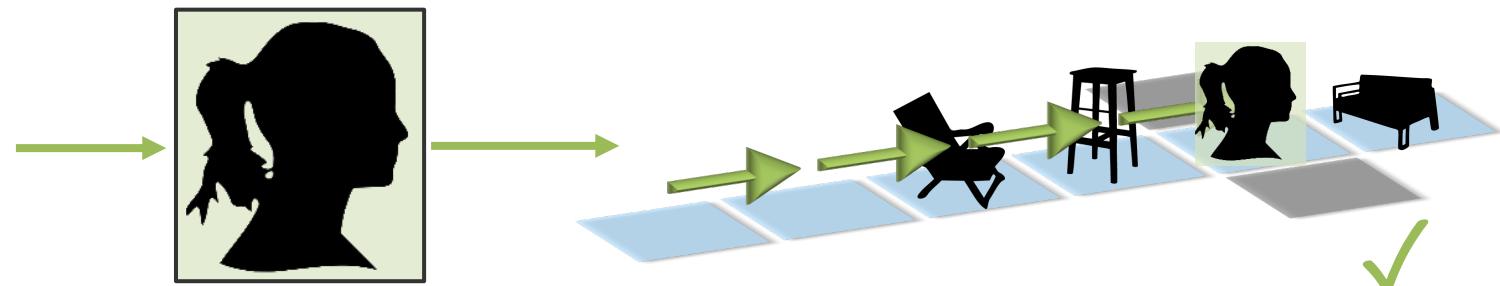
# Pragmatic Speakers Simulate Interpretation

*go forward  
past the stool*

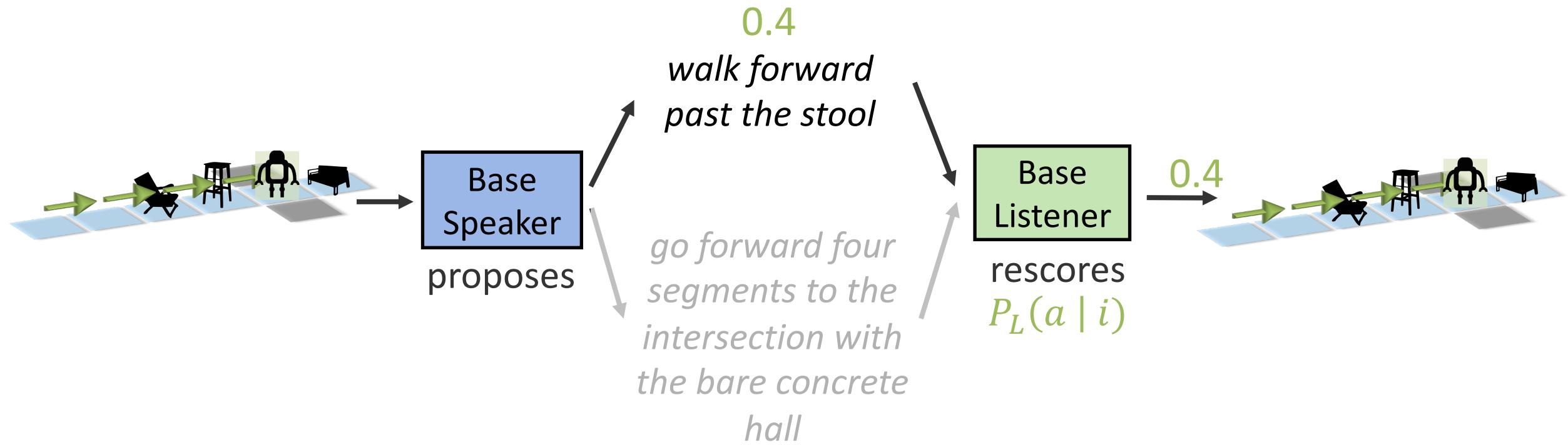


# Pragmatic Speakers Simulate Interpretation

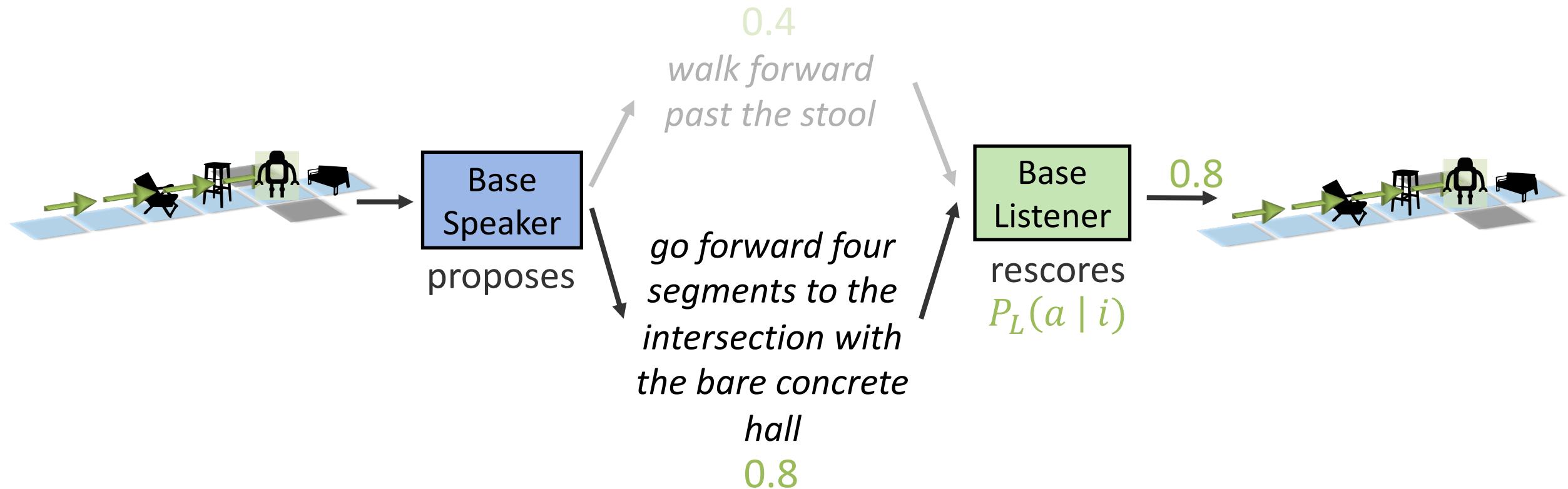
*go forward four segments to the intersection with the bare concrete hall*



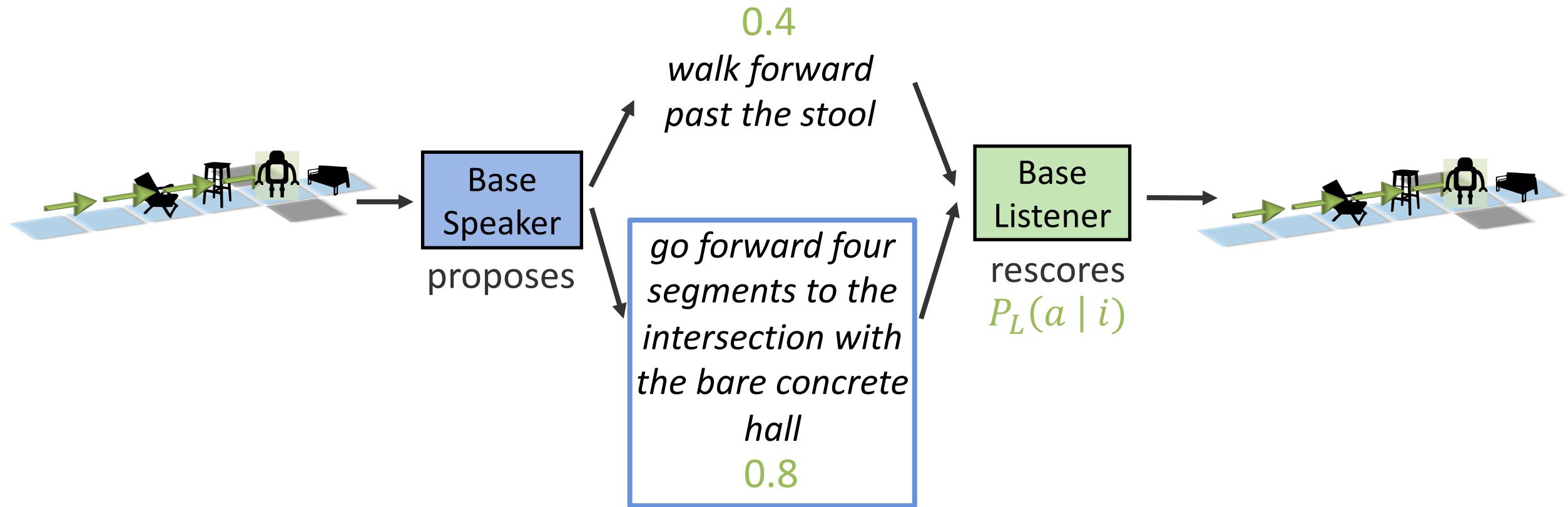
# Building a Pragmatic Speaker



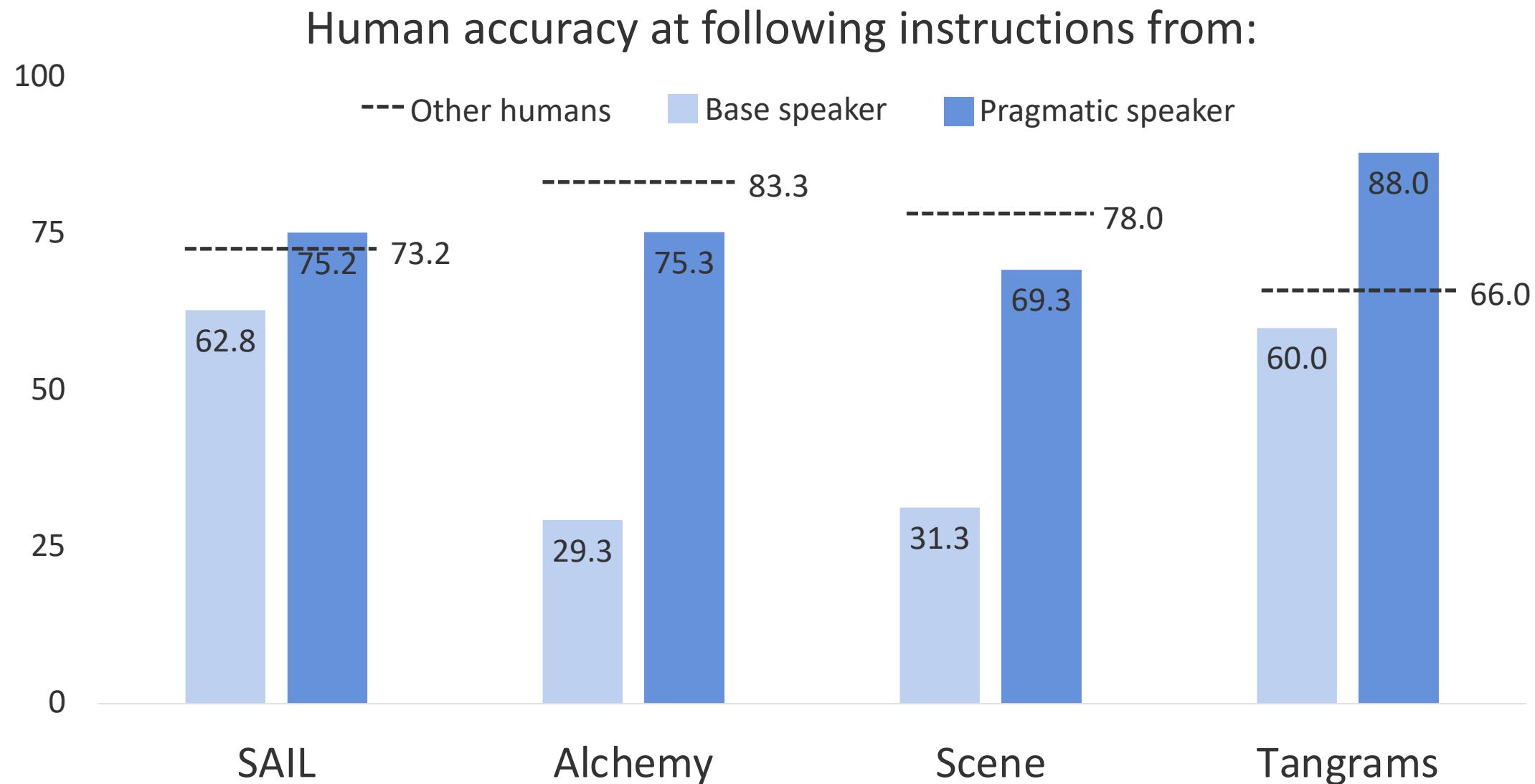
# Building a Pragmatic Speaker



# Building a Pragmatic Speaker



# Speaker Results



# Pragmatics and Communicative Success

---



Base  
Speaker

*throw out the purple chemical*

X

Pragmatic  
Speaker

*throw out the first purple chemical*

✓

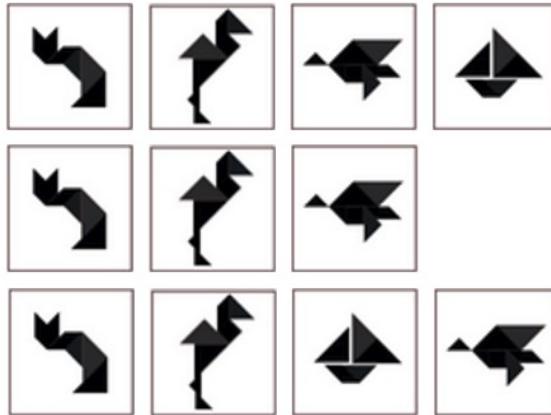
Human

*remove all the purple chemical  
from the beaker on the far left*

✓

# Pragmatics and Communicative Success

---



Base  
Speaker

*remove the last figure  
add it back*

X

Pragmatic  
Speaker

*remove the last figure  
add it back in the 3rd position*

✓

Human

*take away the last item  
undo the last step*

X

# Pragmatic Speakers in Other Domains

## Document Summarization

Input:

*... The 1-0 scoreline that took Barcelona through to the Champions League quarterfinals made their clash with Manchester City all seem rather academic. ....*

Pragmatic Output:

*Barcelona beat Manchester City 1-0 in the Champions League.*

## Image Captioning

Input:



Pragmatic Output:

*two giraffes standing in a large enclosure with a building in the background*

## Visual Navigation

Input:

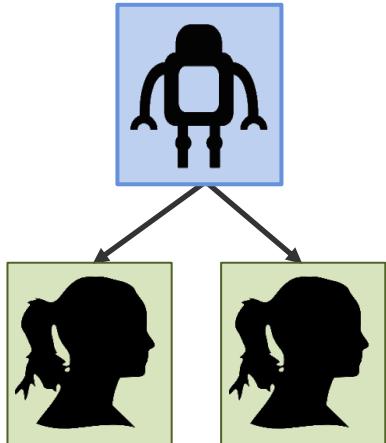


Pragmatic Output:

*walk past the dining room table and chairs and take a right into the living room. stop once you are on the rug.*

# Takeaways

---



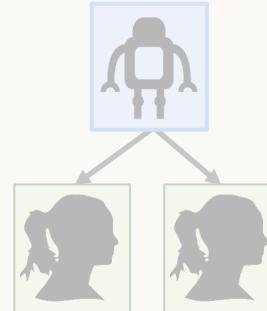
*Simulating people's interpretations makes language more informative.*

*Pragmatics allows models to sometimes outperform their training data.*

# *Pragmatics and...*

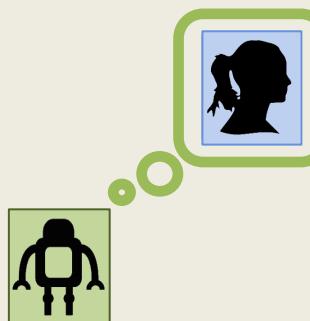
## Generation

[Fried, Andreas, & Klein. NAACL 2018]



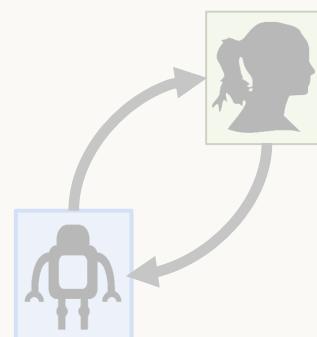
## Interpretation

[Fried\*, Hu\*, Cirik\* et al. NeurIPS 2018]



## Dialogue

[Fried, Chiu, & Klein. EMNLP 2021]

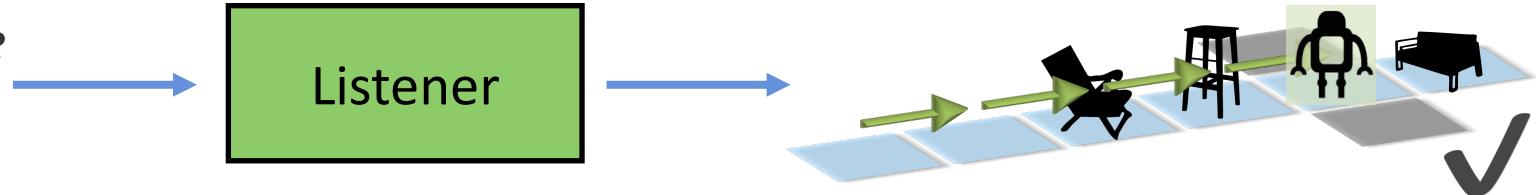


# Listener Tasks

---

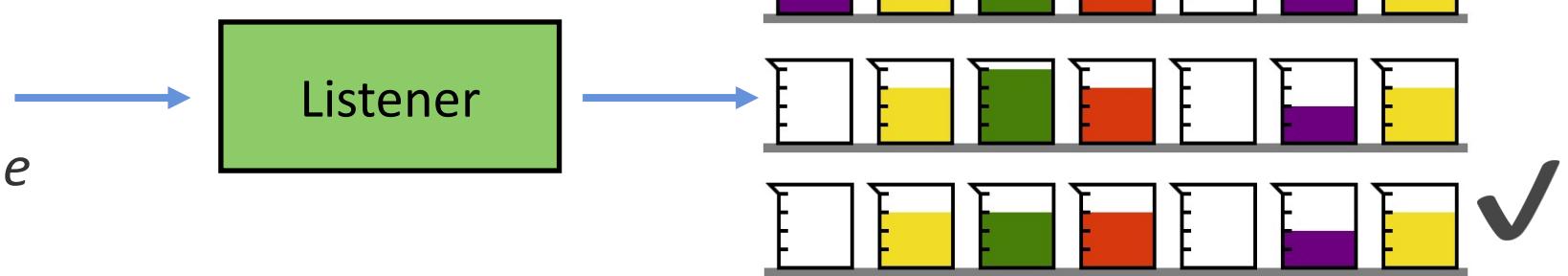
## Navigation

*go forward to the  
grey hallway*

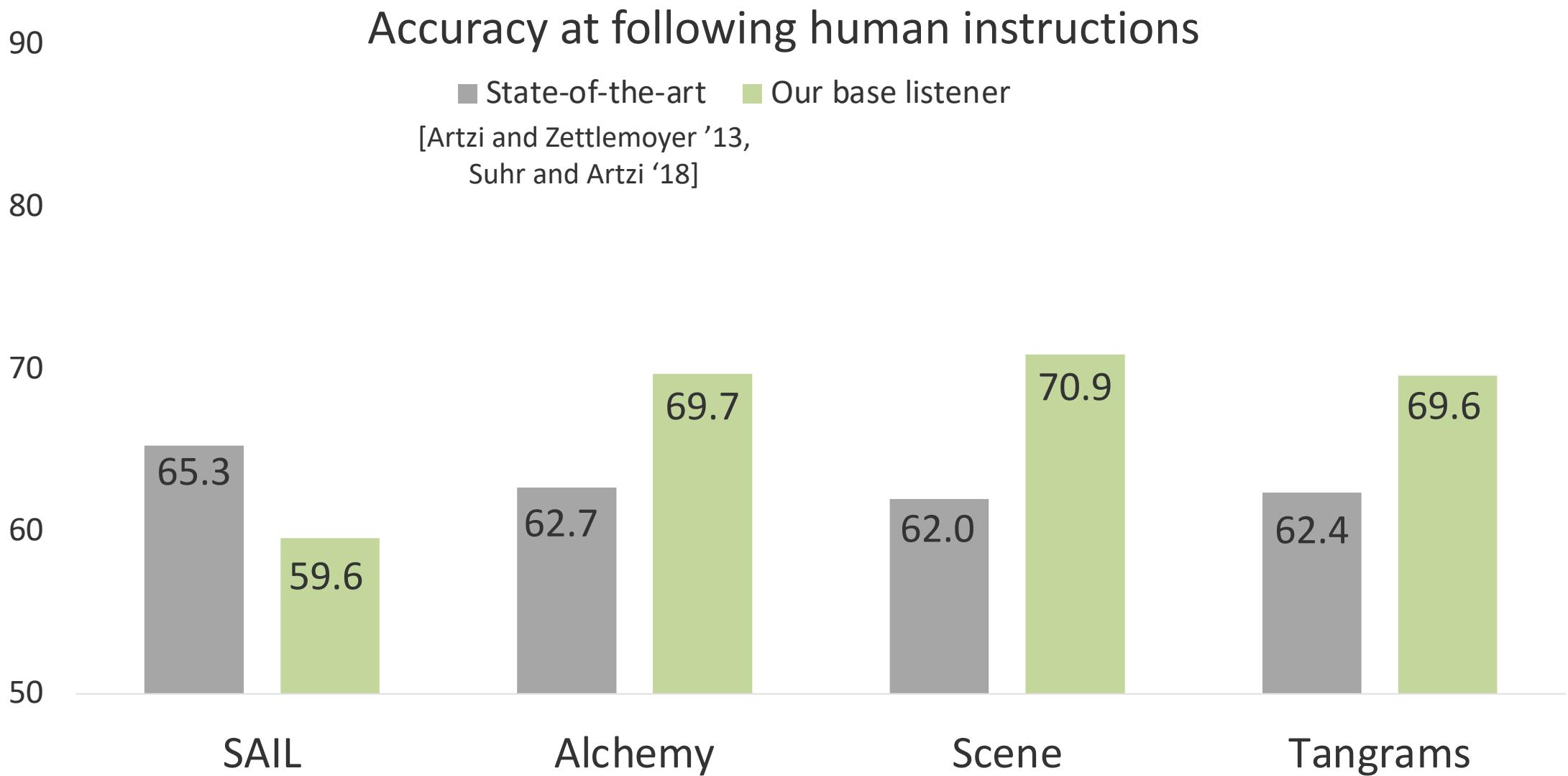


## Contextual Execution: Alchemy

- remove all the purple  
chemical from the  
beaker on the far left*
- do the same with one  
unit of green chemical*



# Strong Listener Models

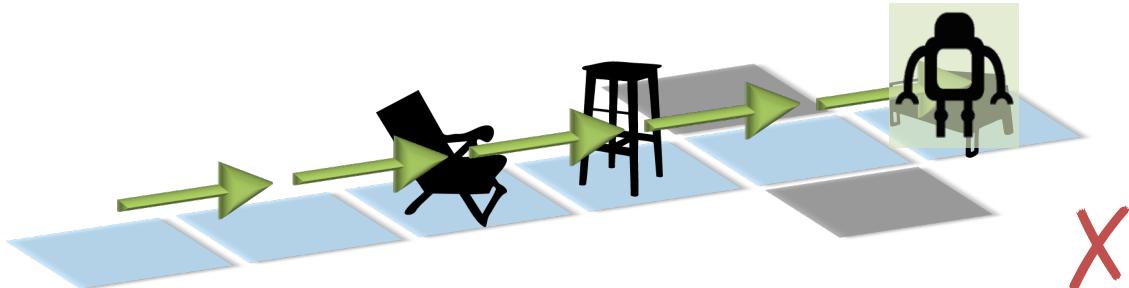


# A Failure Mode for Listeners: Ambiguity

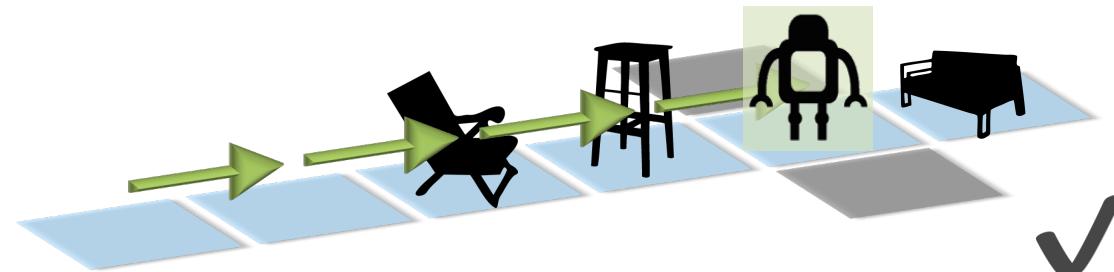
Instruction

*walk along the blue carpet and you pass two objects*

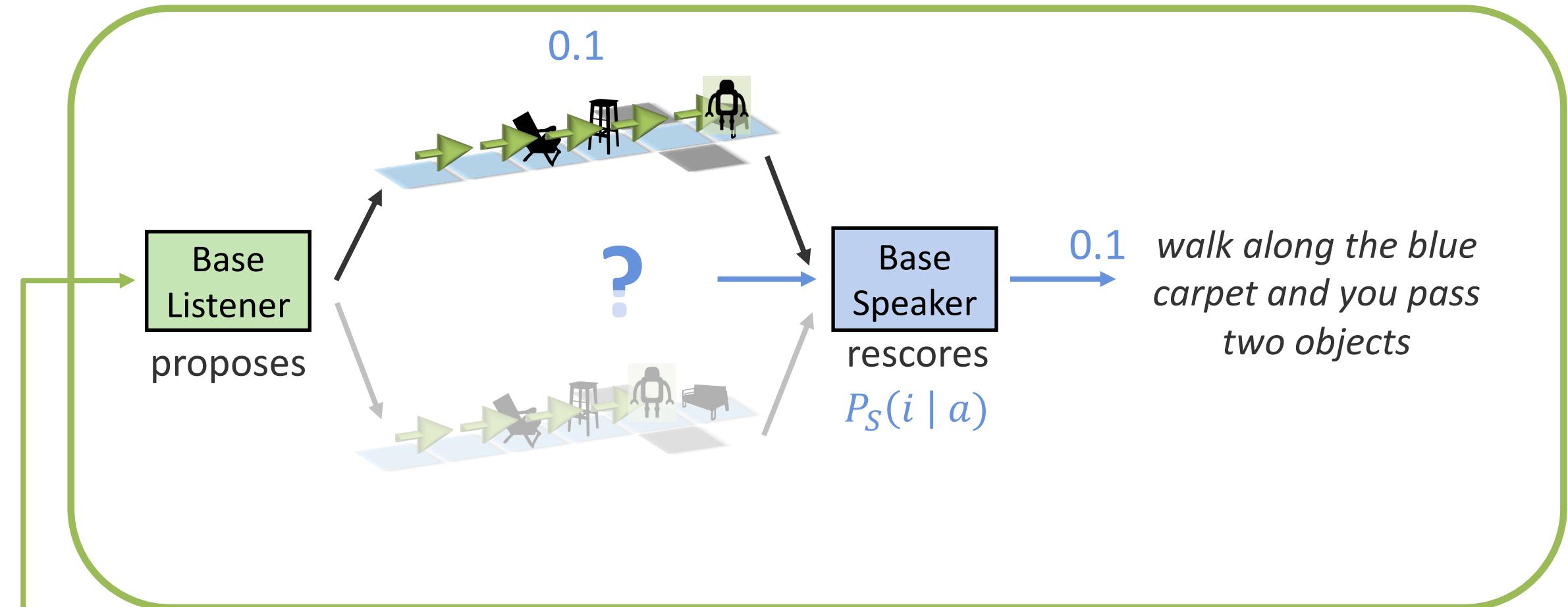
Base  
Listener



Correct



# Building a Pragmatic Listener

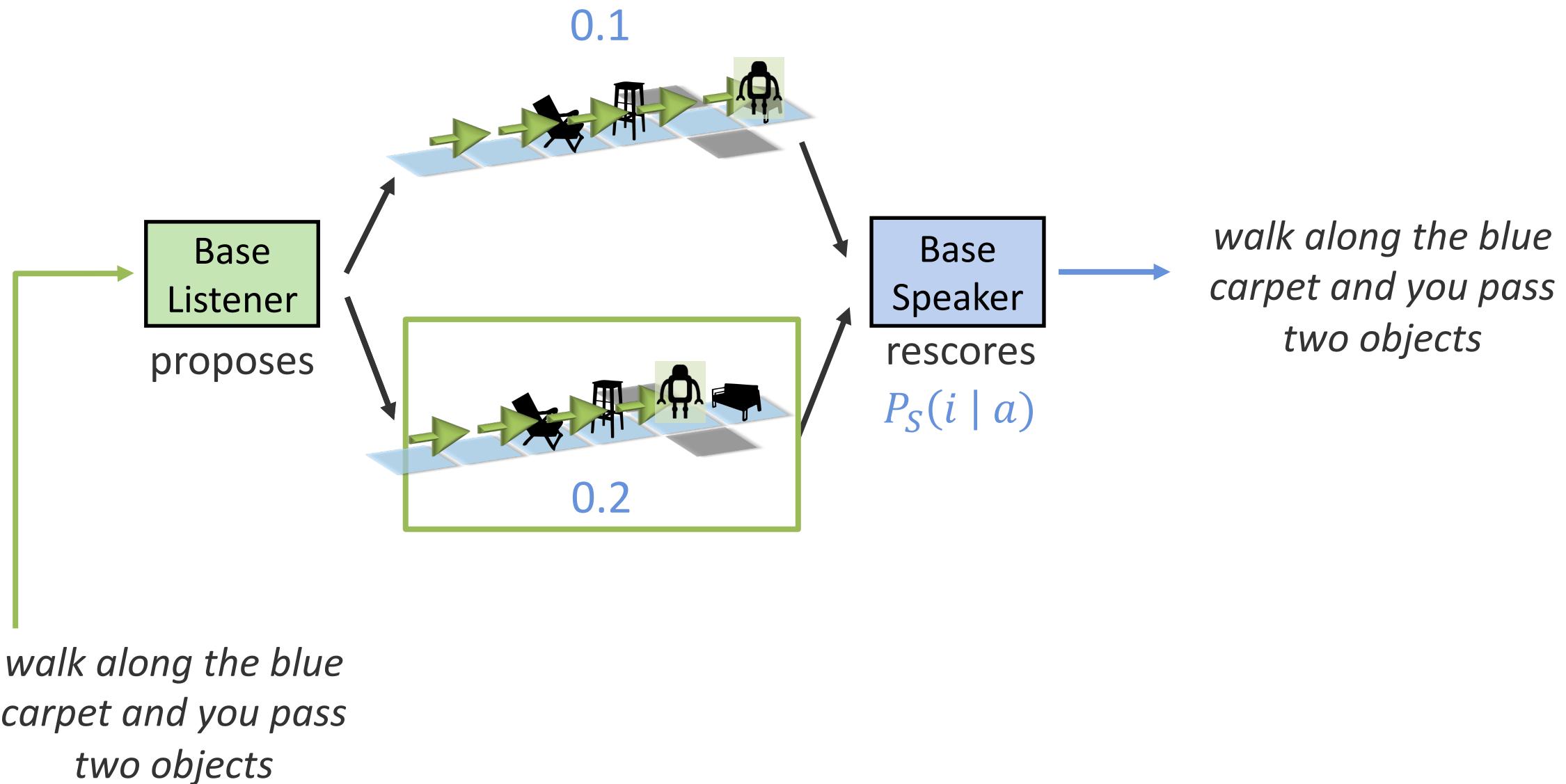


*walk along the blue carpet and you pass two objects*

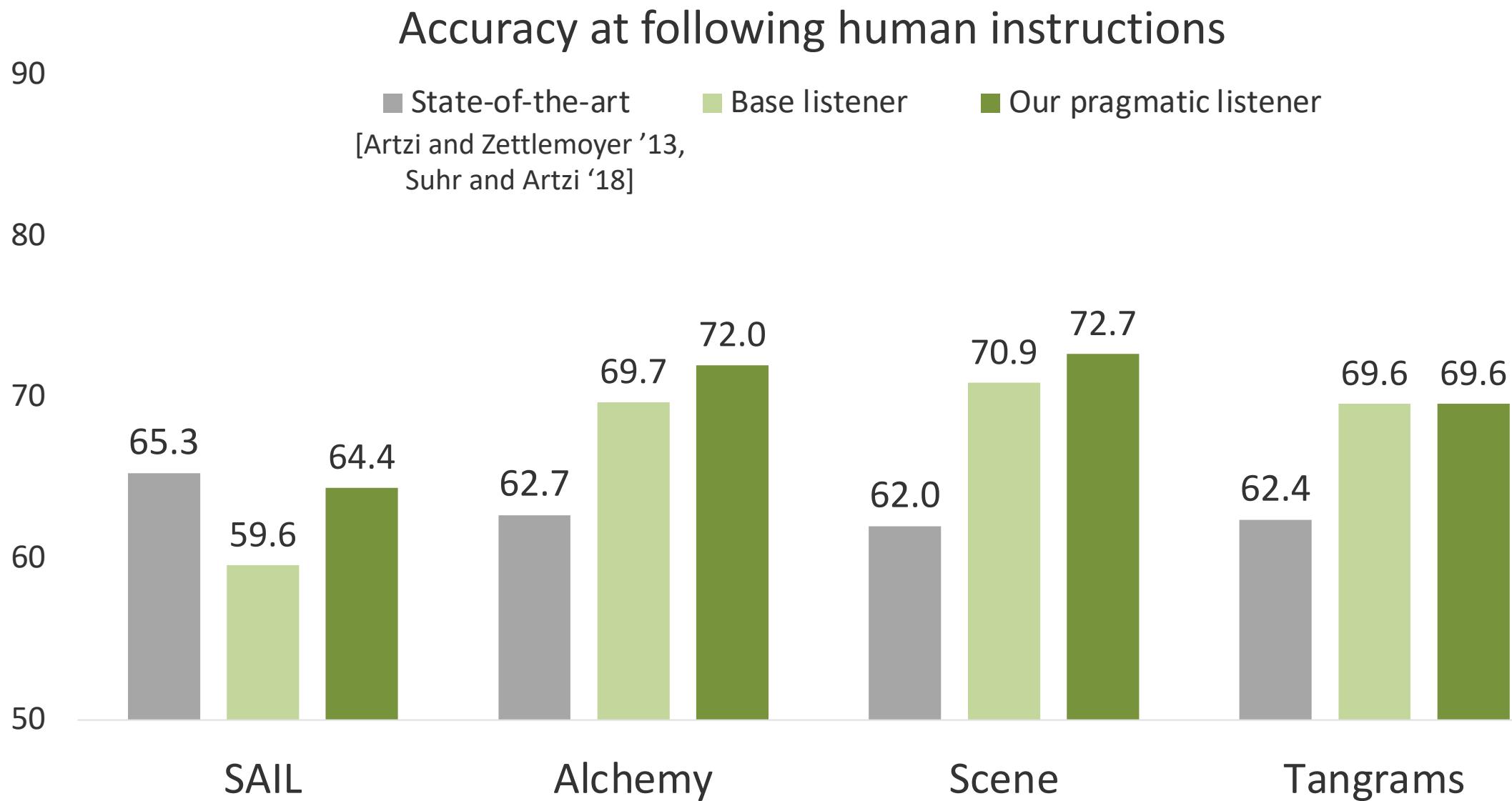
Pragmatic Listener

*walk along the blue carpet and you pass two objects*

# Building a Pragmatic Listener



# Listener Results



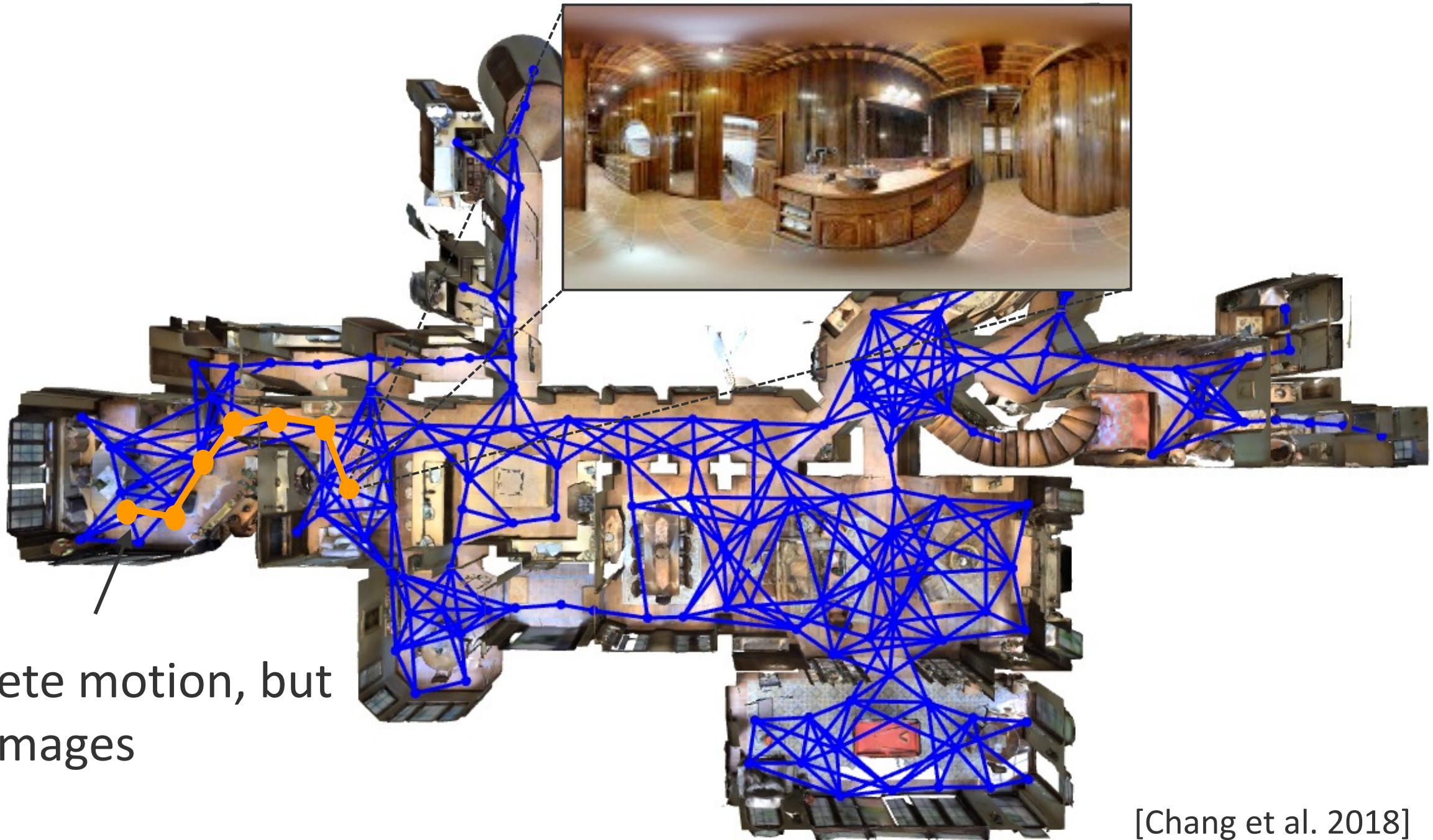
# Visually-Grounded Listeners

---



*Turn left and take a right at the table. Take a left at the painting and then take your first right. Wait next to the exercise equipment.*

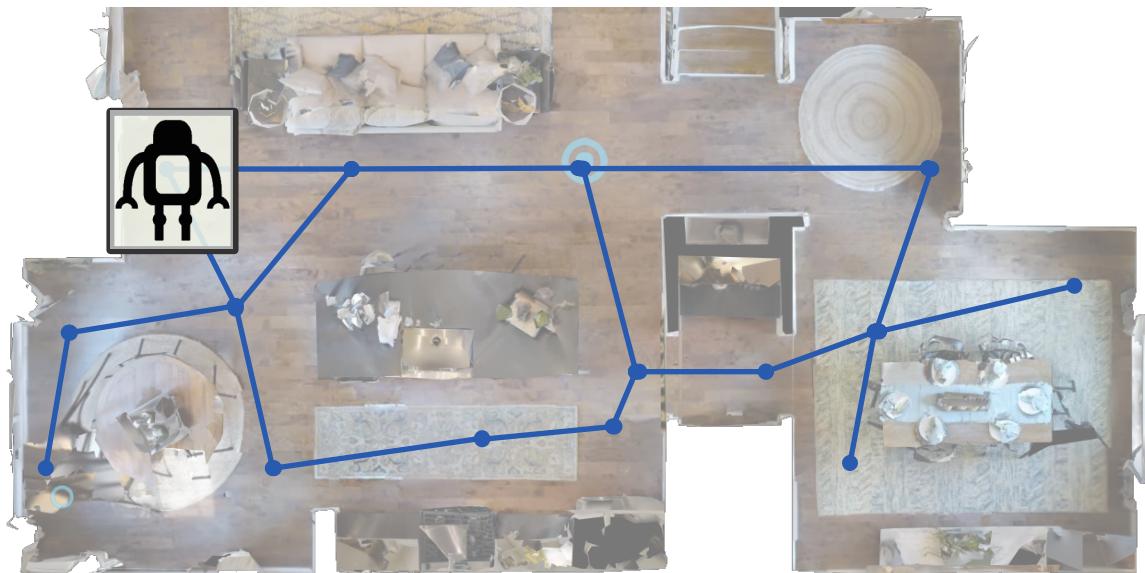
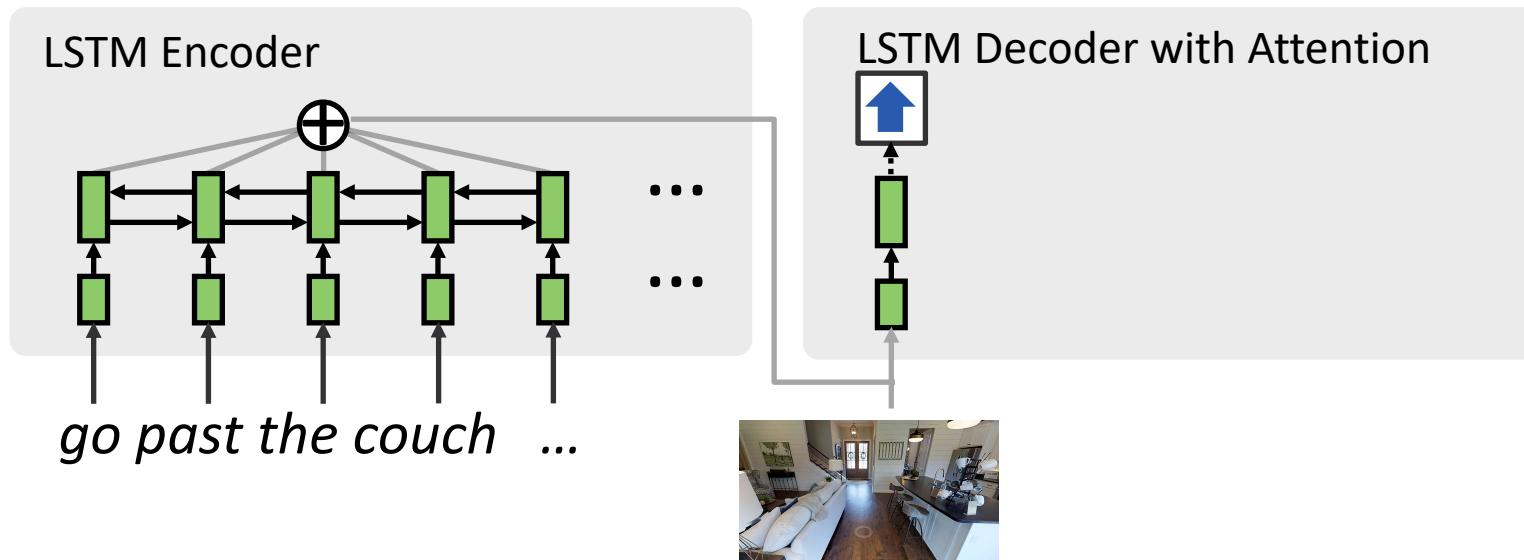
[Vision-and-Language Navigation Task. Anderson et al., 2018]



Discrete motion, but  
real images

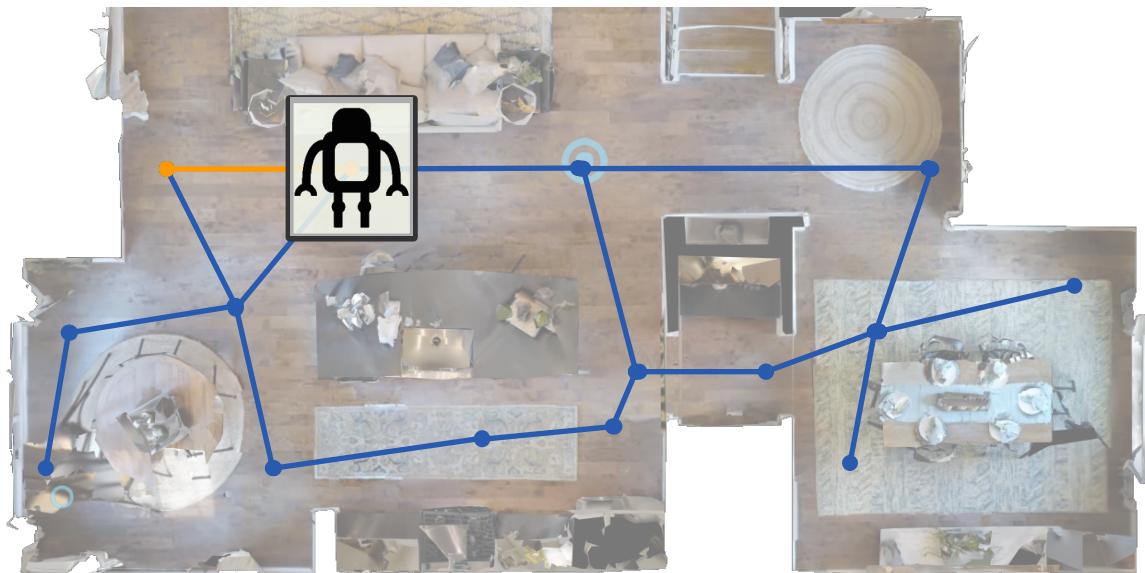
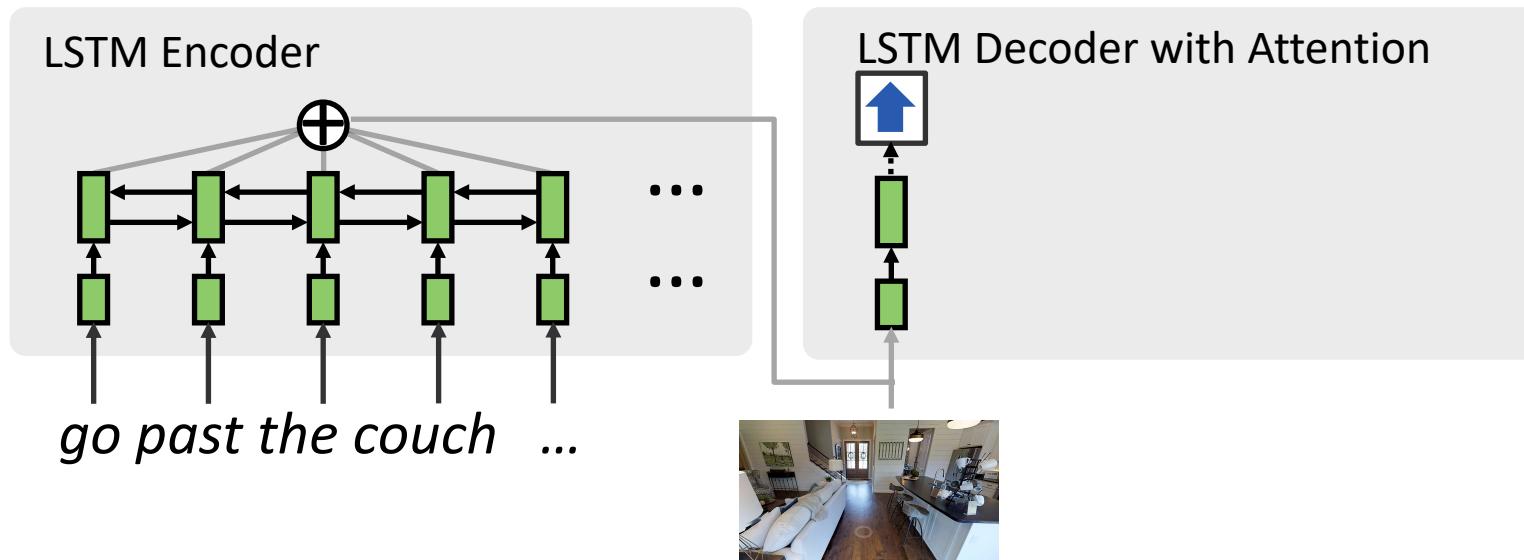
[Chang et al. 2018]

# Base Listener Model



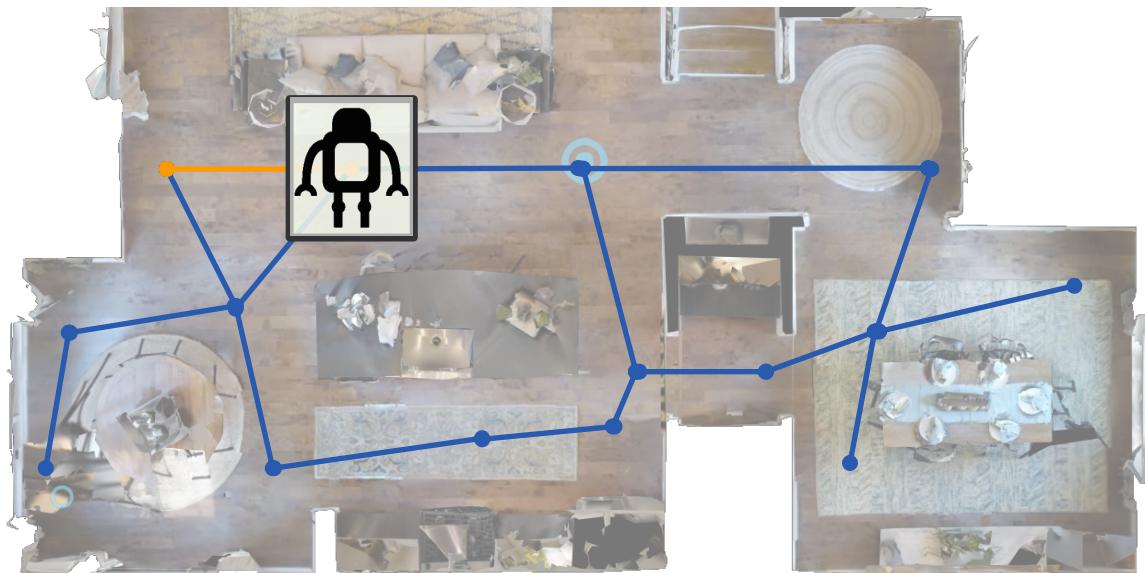
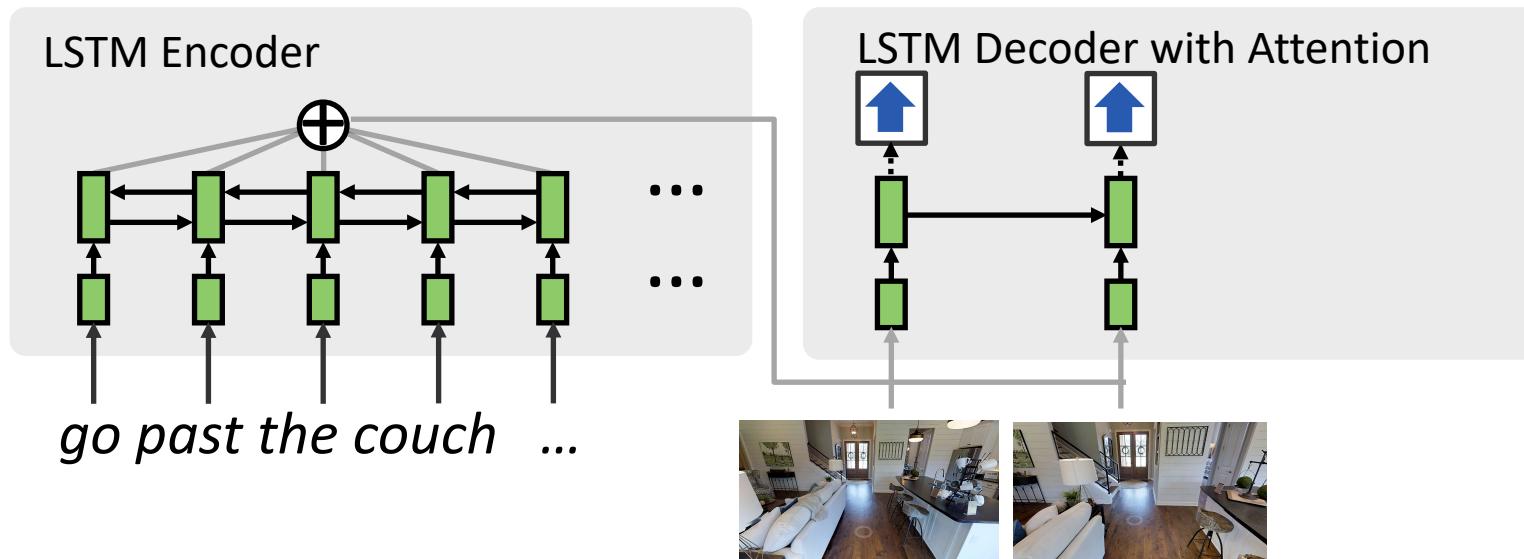
[Anderson et al., 2018]

# Base Listener Model



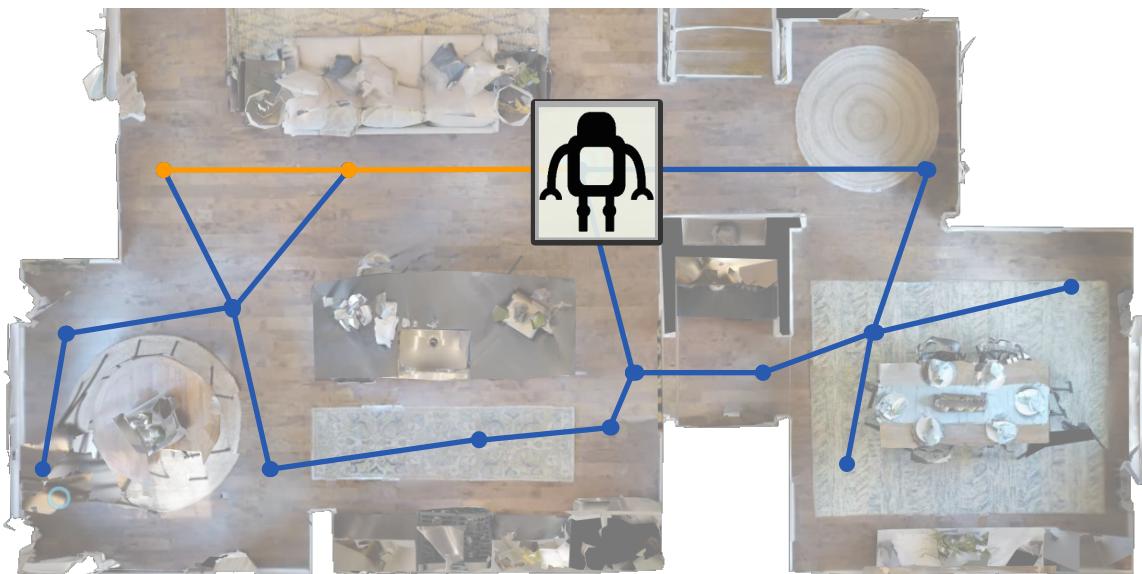
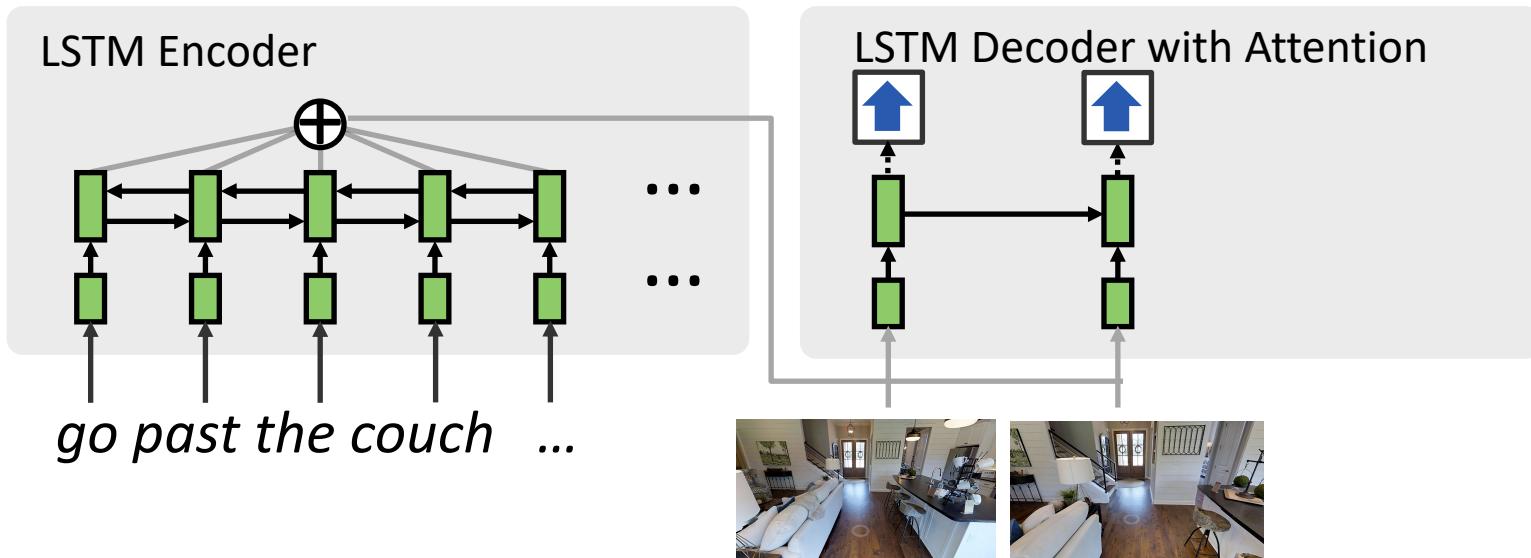
[Anderson et al., 2018]

# Base Listener Model



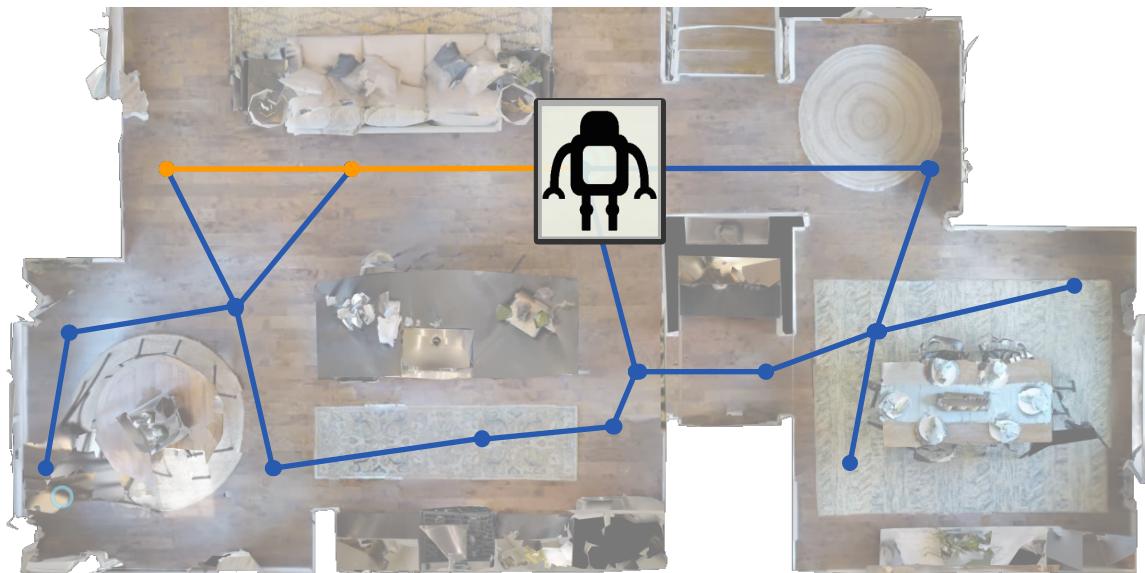
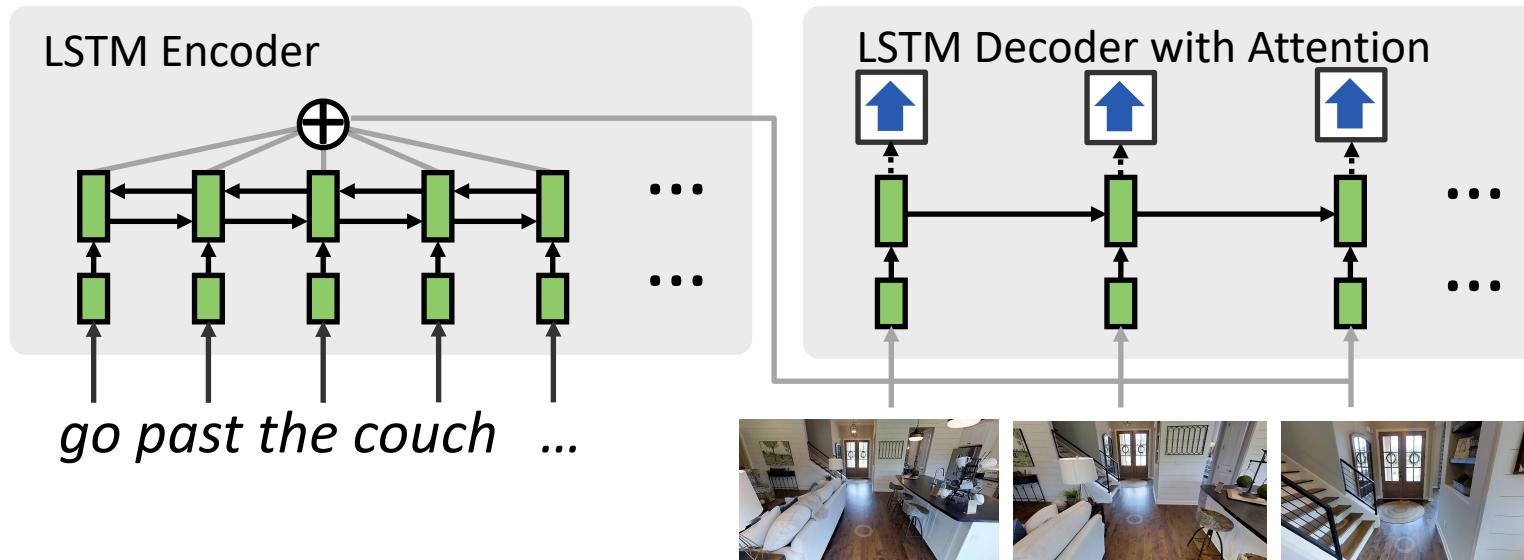
[Anderson et al., 2018]

# Base Listener Model



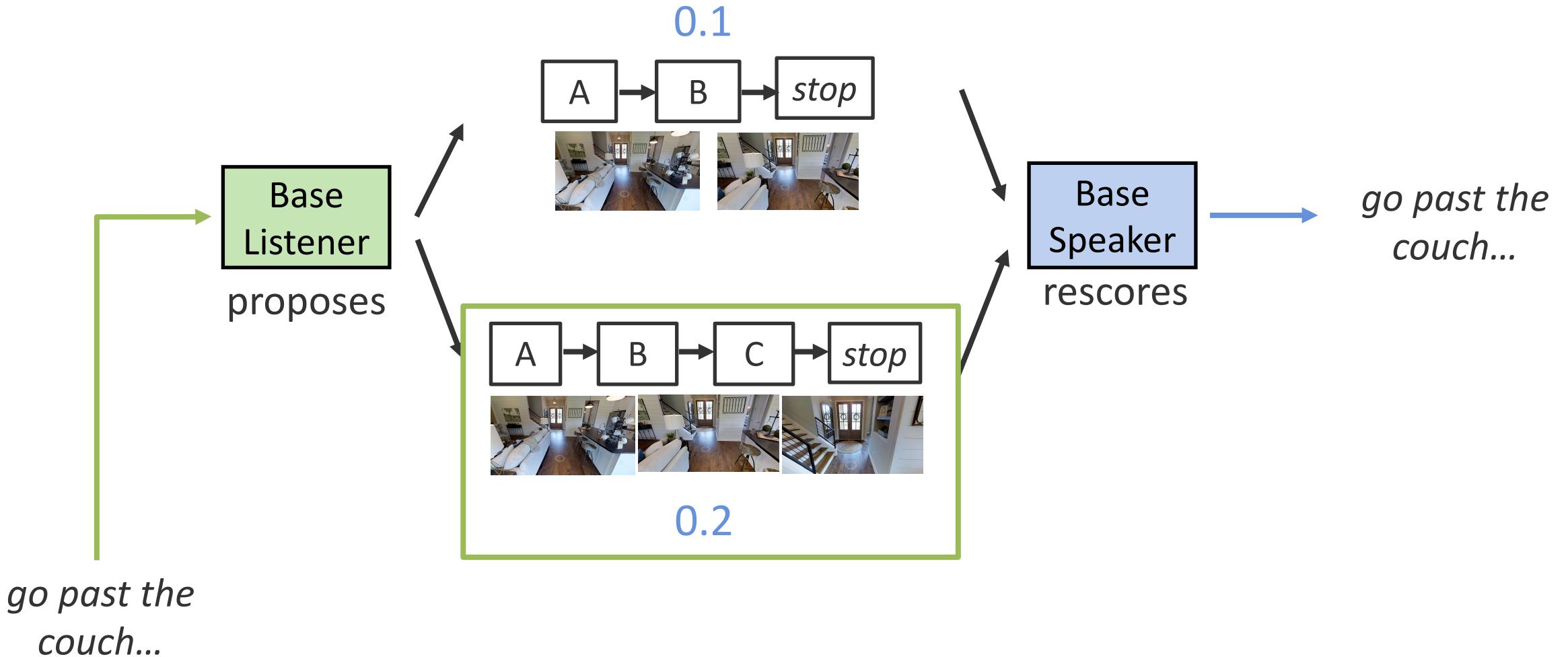
[Anderson et al., 2018]

# Base Listener Model



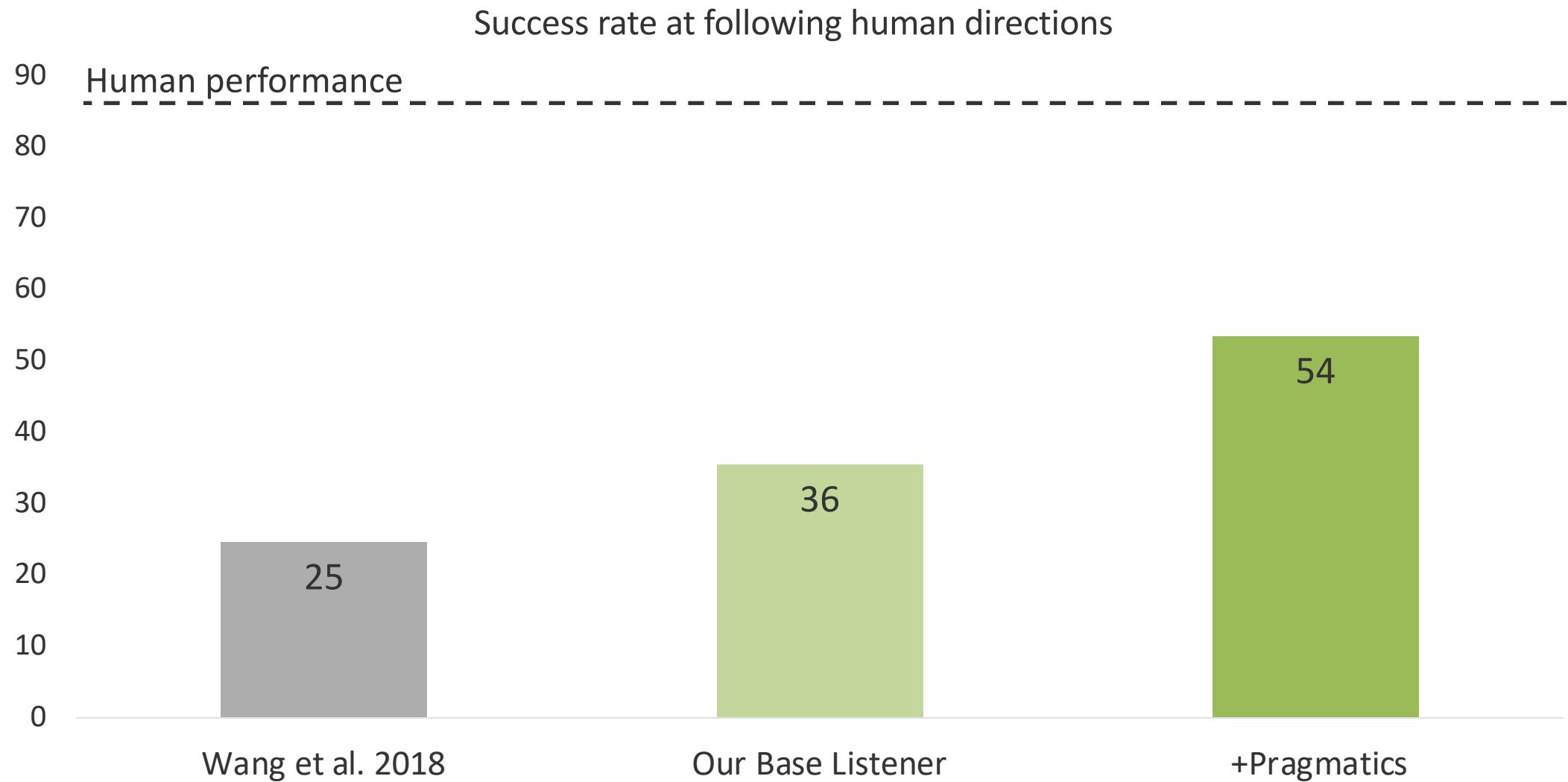
[Anderson et al., 2018]

# Pragmatics for Visual Navigation



# Comparison to Prior Work

---



*Walk past hall table. Walk into bedroom. Make left at table clock.  
Wait at bathroom door threshold.*



Base listener

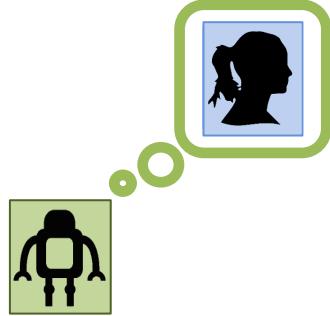
*Walk past hall table. Walk into bedroom. Make left at table clock.  
Wait at bathroom door threshold.*



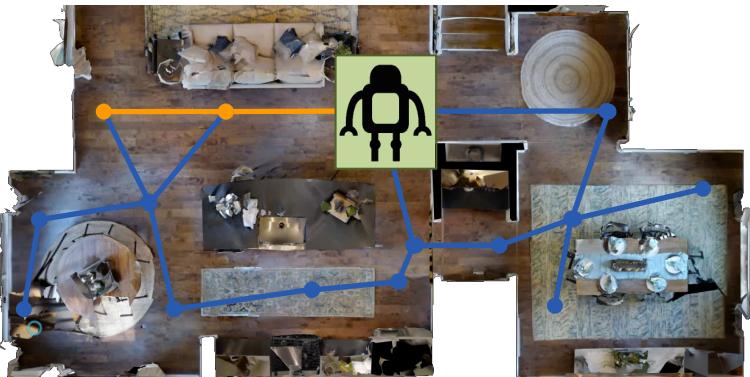
Pragmatic listener

# Takeaways

---



*Simulating why a speaker said what they did helps resolve ambiguity.*

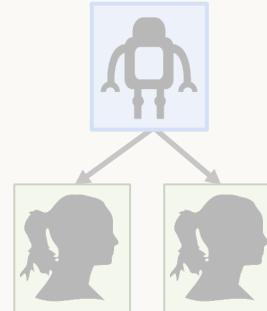


*Pragmatics improves most in complex environments where grounding is harder.*

# *Pragmatics and...*

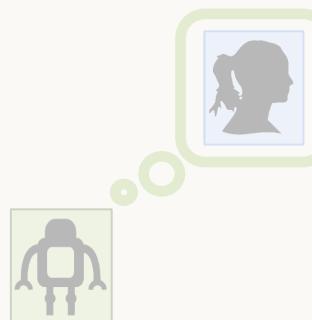
## Generation

[Fried, Andreas, & Klein. NAACL 2018]



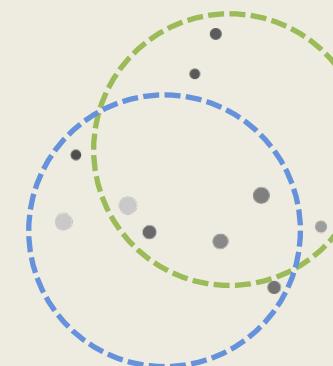
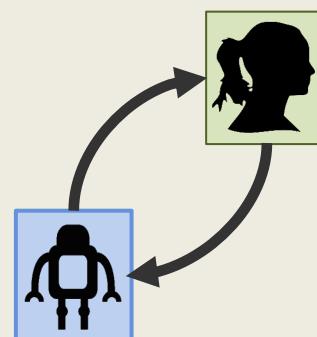
## Interpretation

[Fried\*, Hu\*, Cirik\* et al. NeurIPS 2018]

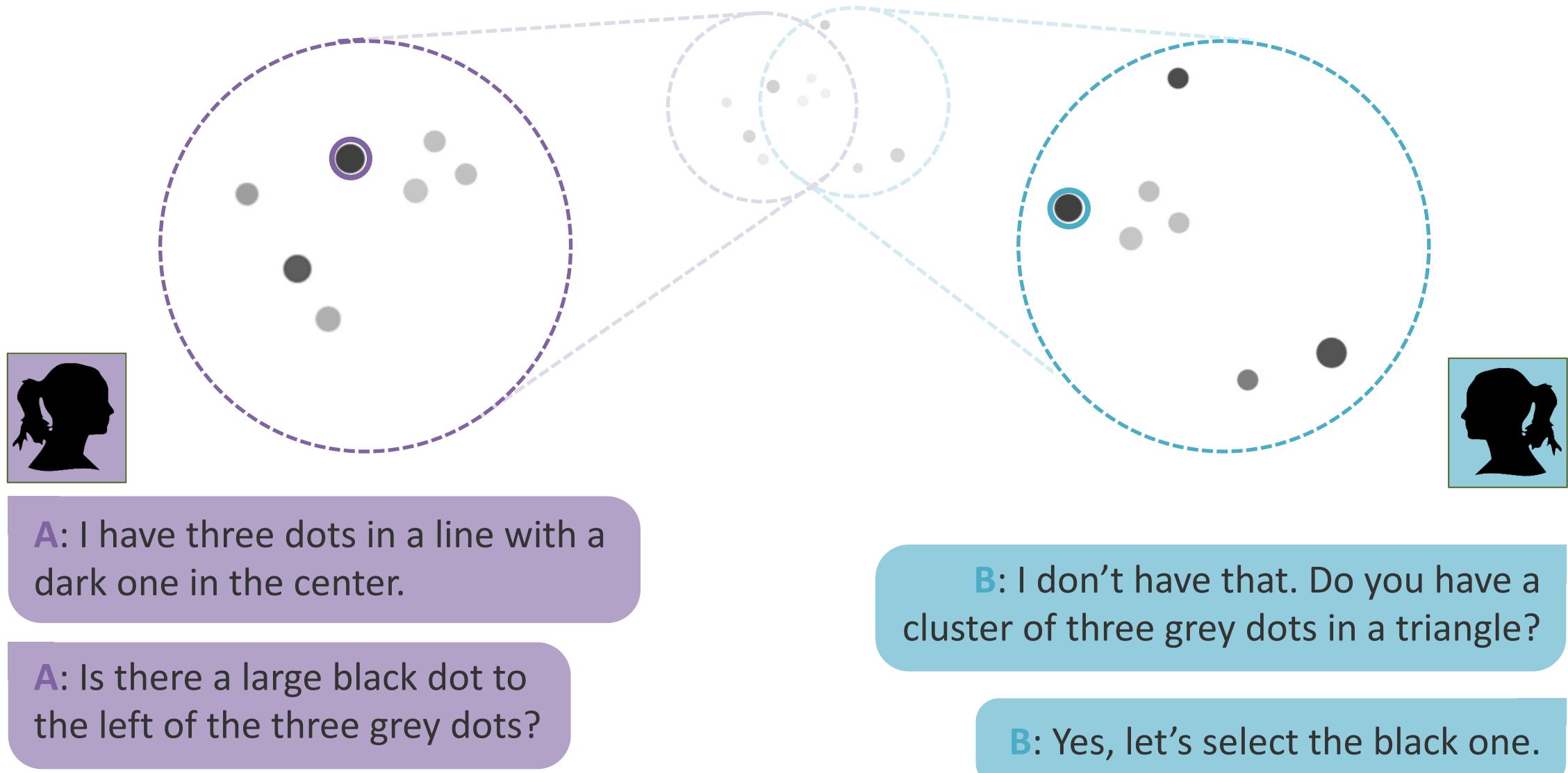


## Dialogue

[Fried, Chiu, & Klein. EMNLP 2021]

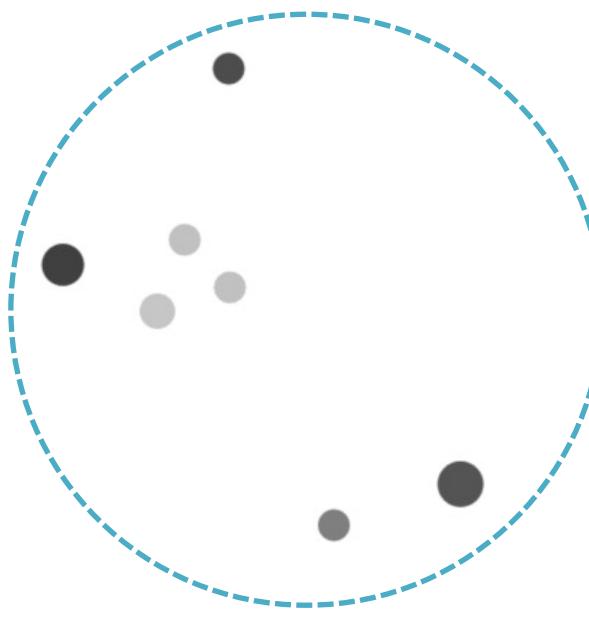


# Grounded Collaborative Dialogue



# Decomposing Into Subtasks

---



A: I have three dots  
in a line with a dark →  
one in the center.

B: I don't have that. Do  
you have a group of →  
three grey dots?

A: Is there a large  
black dot to the left of → B:???  
the three grey dots?

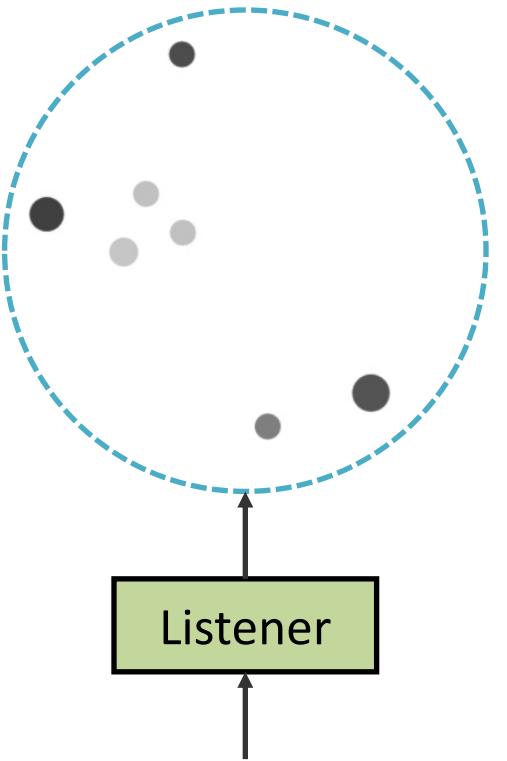
# Decomposing Into Subtasks

---

don't have A that have three dots there B all don't have that. Do you have a group with black dot to the left of a group? If → three grey dots in the center three grey dots?

A: Is there a large black dot to the left of → B: ???  
the three grey dots?

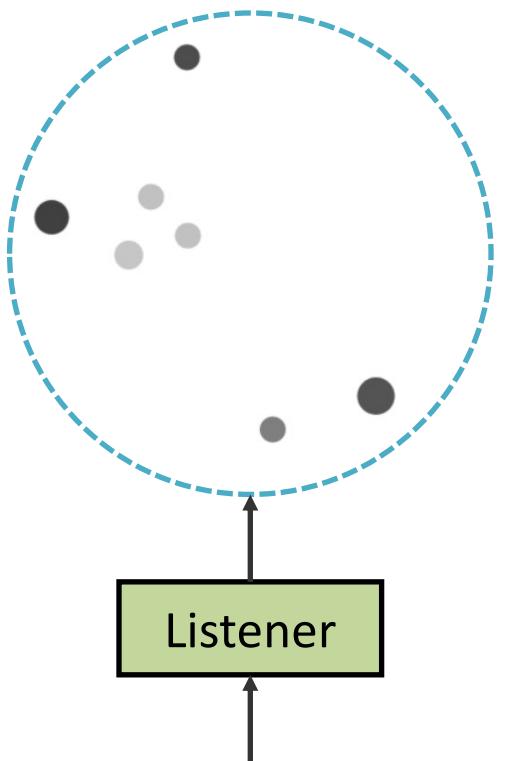
# Decomposing Into Subtasks



don't have that. Do you have a group of three grey dots?

A: Is there a large  
black dot to the left of → B:????  
the three grey dots?

# Decomposing Into Subtasks

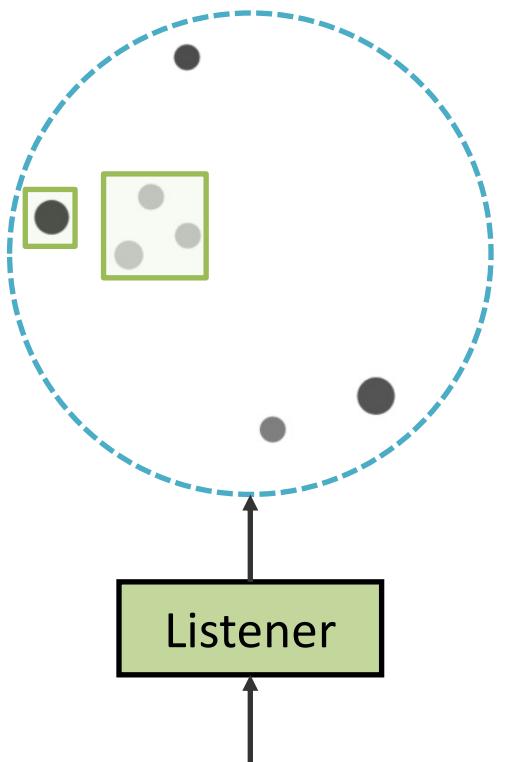


don't have that. Do you have a group of three grey dots? → black dot to the left of → B:???

A: Is there a large  
the three grey dots?

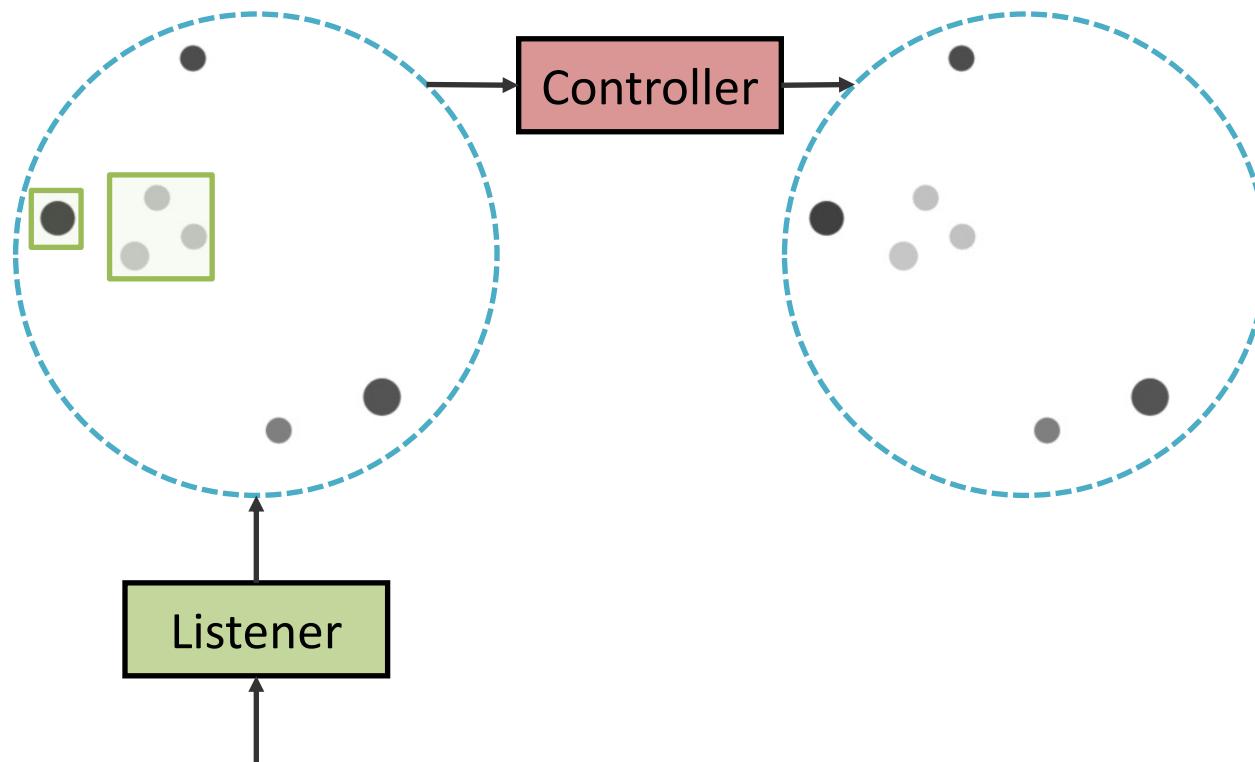
# Decomposing Into Subtasks

---



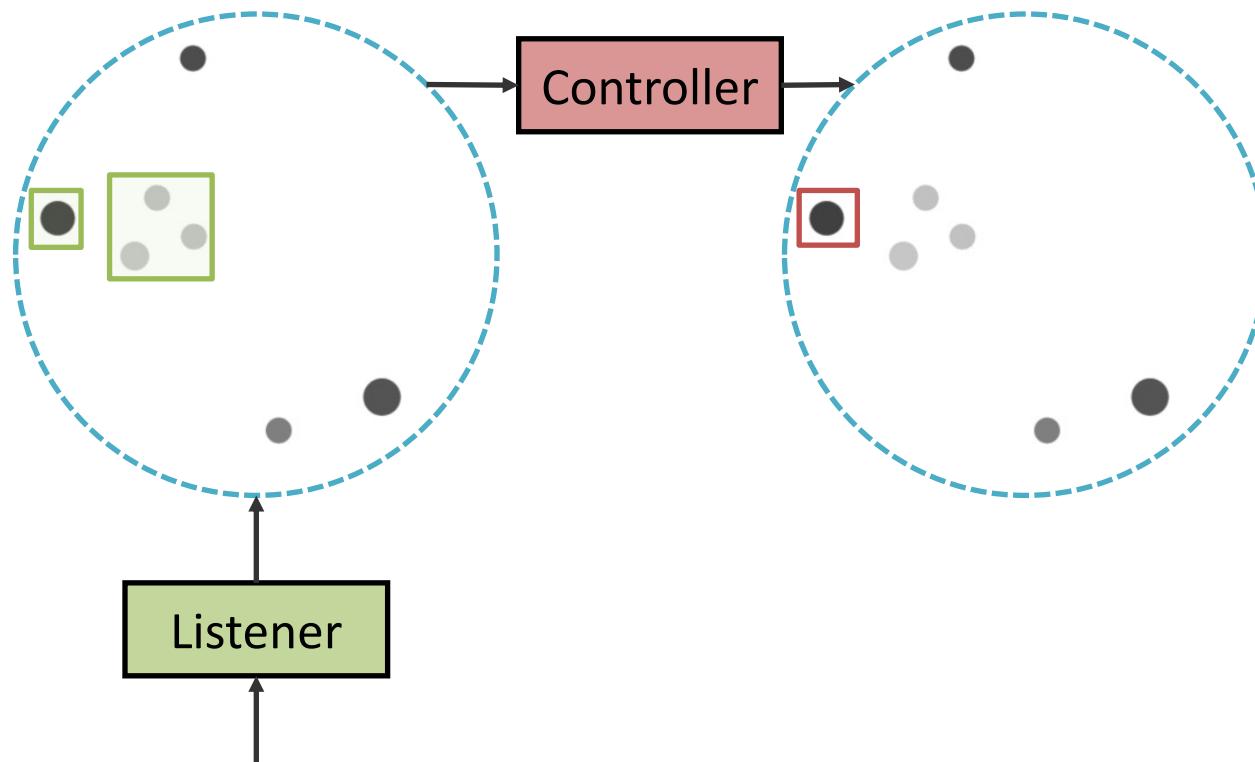
don't have that. Do  
you have a group of → A: Is there a large  
three grey dots? → black dot to the left of → B:???  
the three grey dots?

# Decomposing Into Subtasks



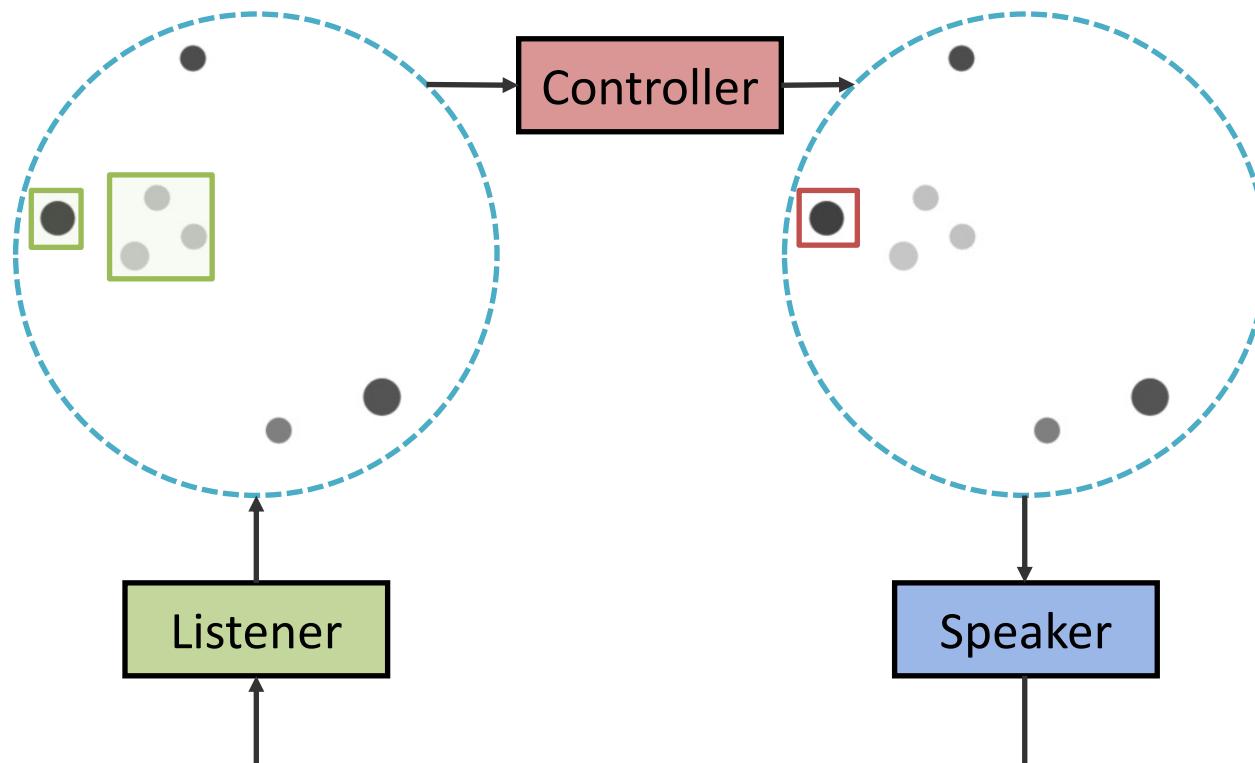
don't have that. Do  
you have a group of → A: Is there a large  
three grey dots? → black dot to the left of → B:???  
the three grey dots?

# Decomposing Into Subtasks



don't have that. Do  
you have a group of → A: Is there a large  
three grey dots? → black dot to the left of → B:???  
the three grey dots?

# Decomposing Into Subtasks

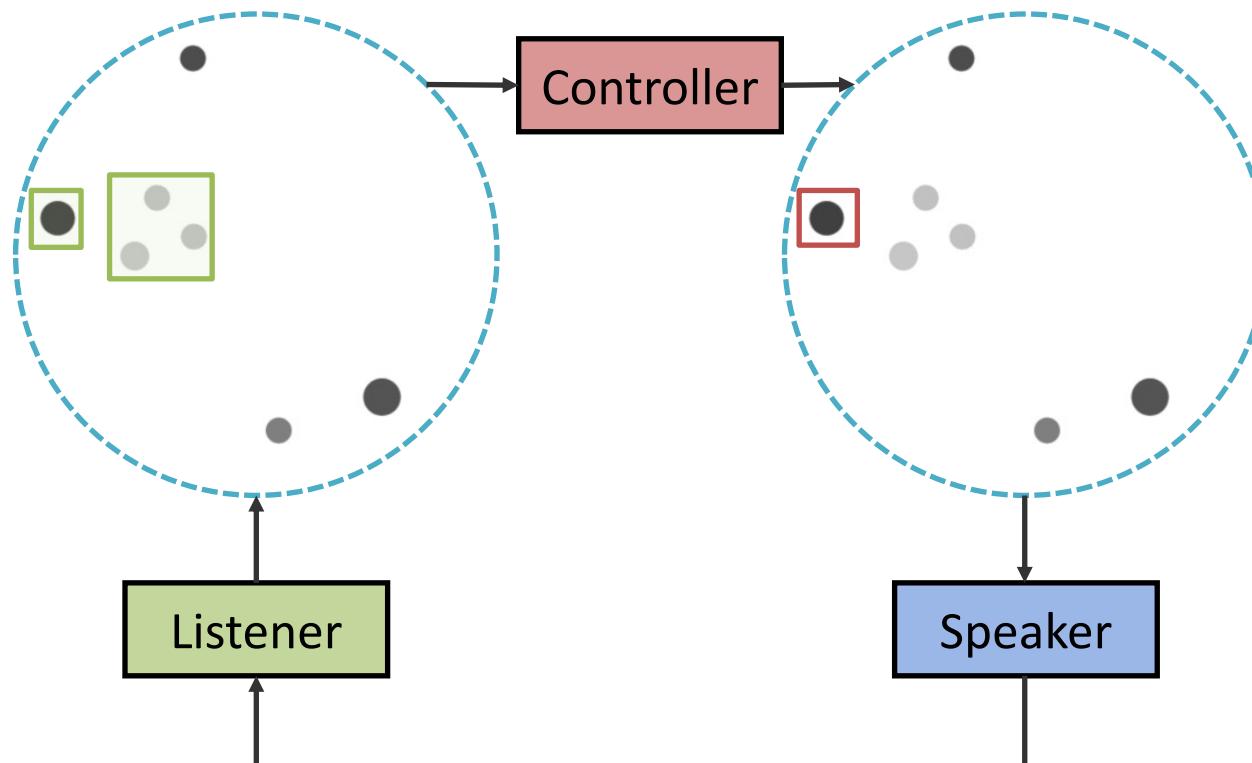


don't have that. Do  
you have a group of → black dot to the left of →  
three grey dots?

A: Is there a large  
black dot to the left of →  
the three grey dots?

B: Yes, let's select  
the black one.

# Decomposing Into Subtasks

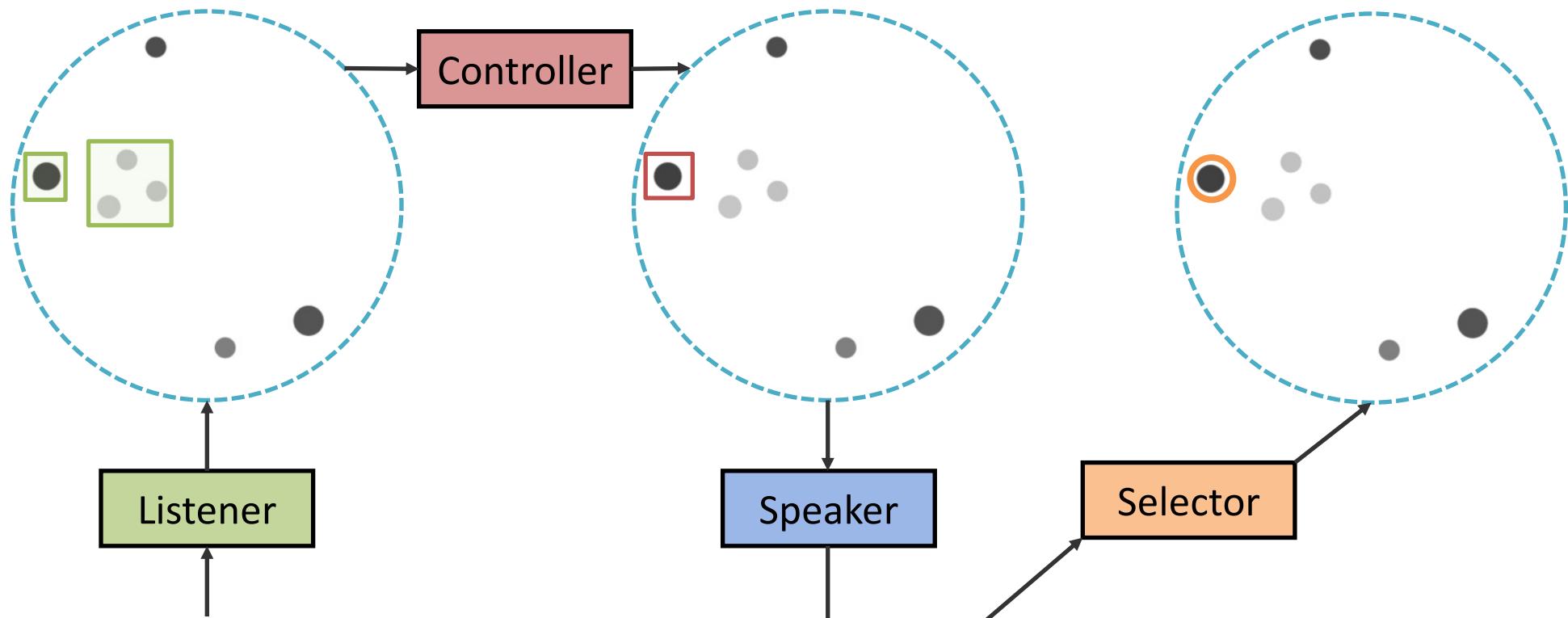


don't have that. Do  
you have a group of → black dot to the left of →  
three grey dots?

A: Is there a large  
black dot to the left of →  
the three grey dots?

B: Yes, let's select  
the black one.

# Decomposing Into Subtasks

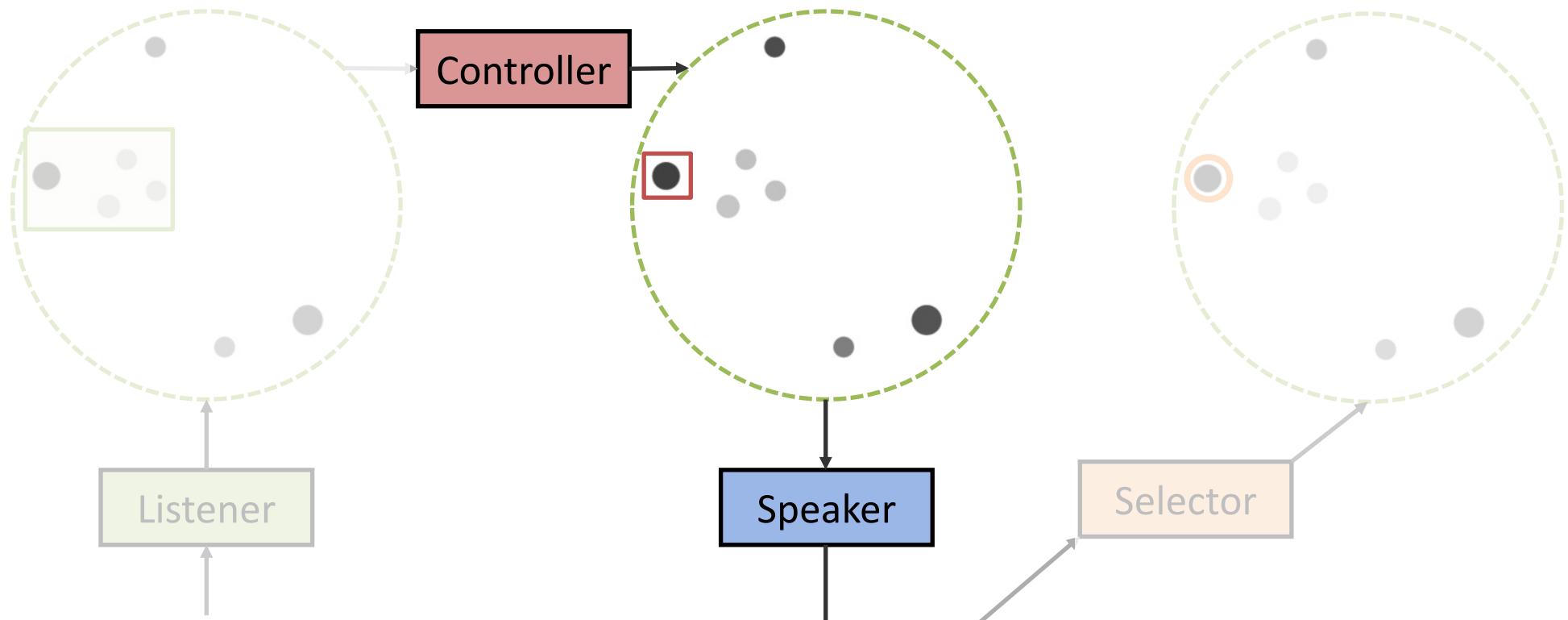


don't have that. Do  
you have a group of → black dot to the left of →  
three grey dots?

A: Is there a large  
black dot to the left of →  
the three grey dots?

B: Yes, let's select  
the black one.

# Decomposing Into Subtasks

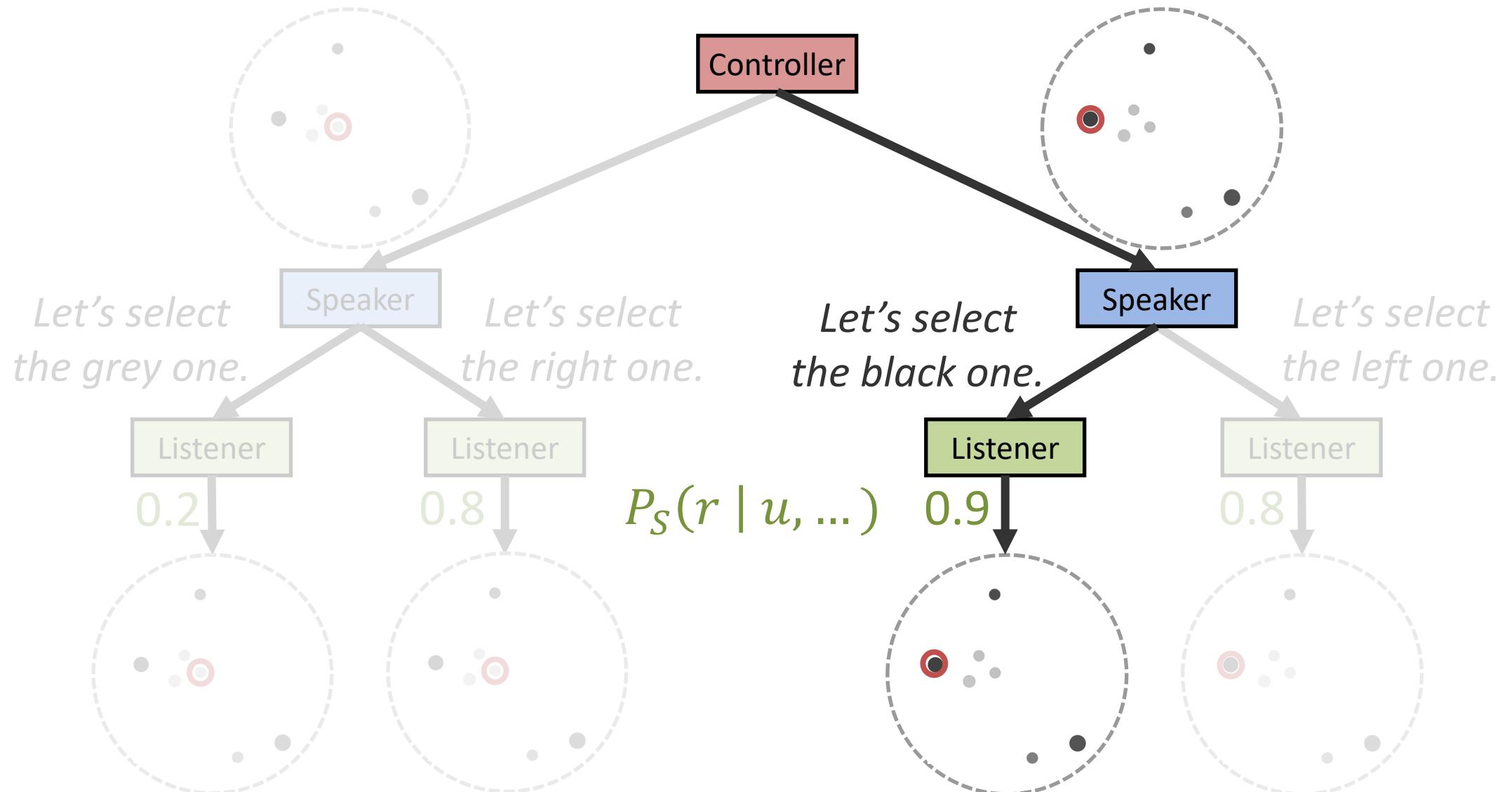


don't have that. Do  
you have a group of  
three grey dots?

A: Is there a large  
black dot to the left of  
the three grey dots? →

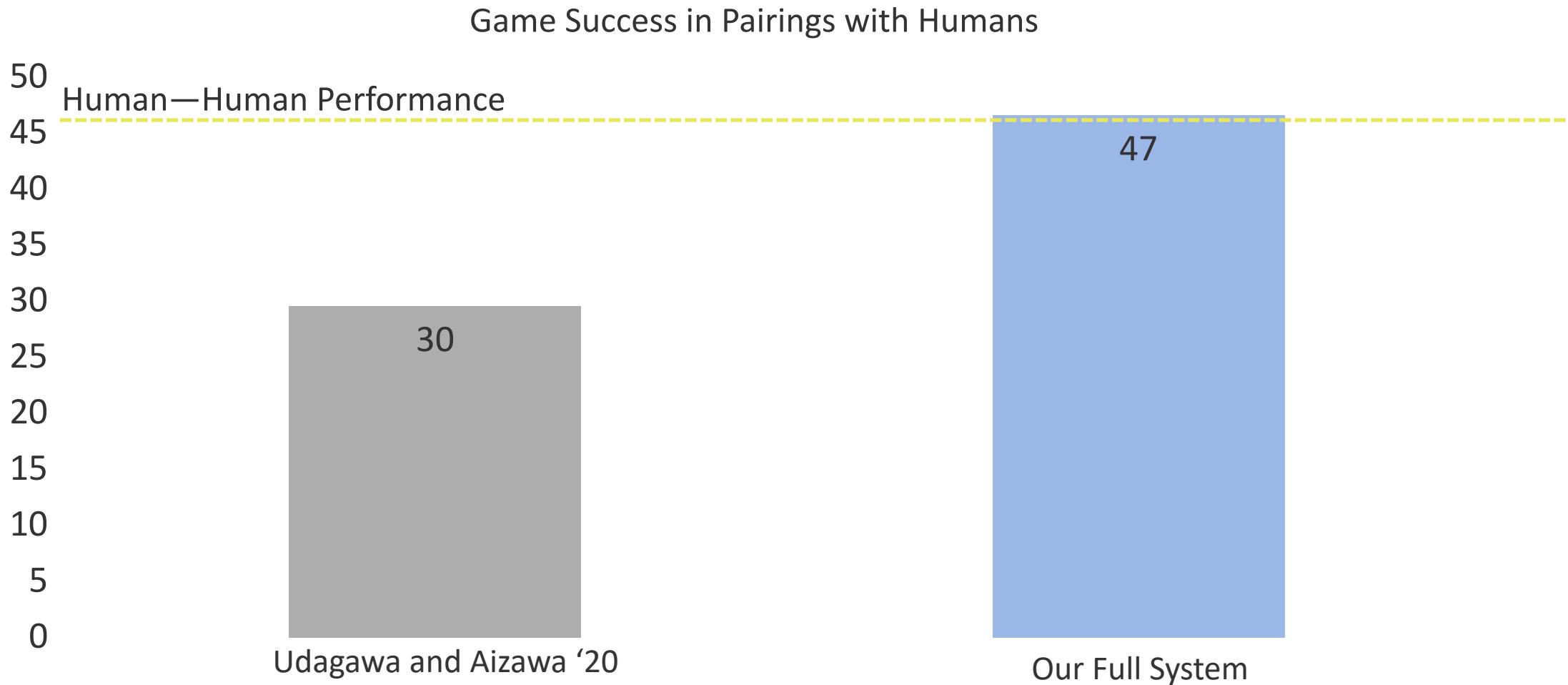
B: Yes, let's select  
the black one.

# Pragmatic Generation



# Full System Evaluation

---



# Demo

S Find One In Common! × +

localhost:5000/dialogue/?uid=U\_97547191ca4347b884ec3fd41df531a3

Time Remaining: 6:00

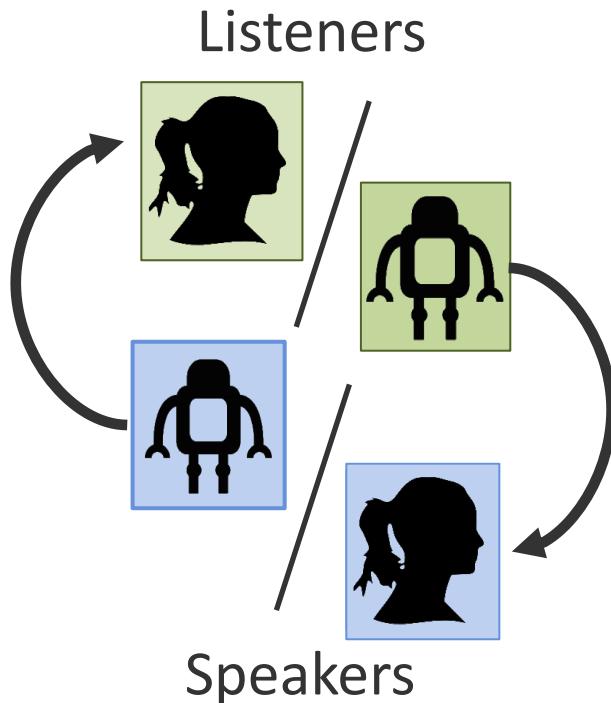
[02/12/21 08:57:44] <You entered the room.>  
[02/12/21 08:57:46] <Your partner has joined the room.>

Your view

Waiting on your partner to take a turn...

# Final Takeaways

---



*Language is a cooperative,  
multiagent process.*

*Language systems improve when they  
plan against simulated humans.*

# Collaborators

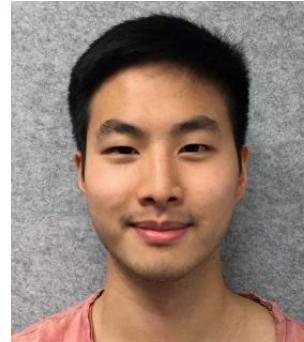
---



Jacob Andreas



Taylor Berg-Kirkpatrick



Justin Chiu



Volkan Cirik



Trevor Darrell



Ronghang Hu



Dan Klein



Louis-Philippe  
Morency



Anna Rohrbach



Kate Saenko



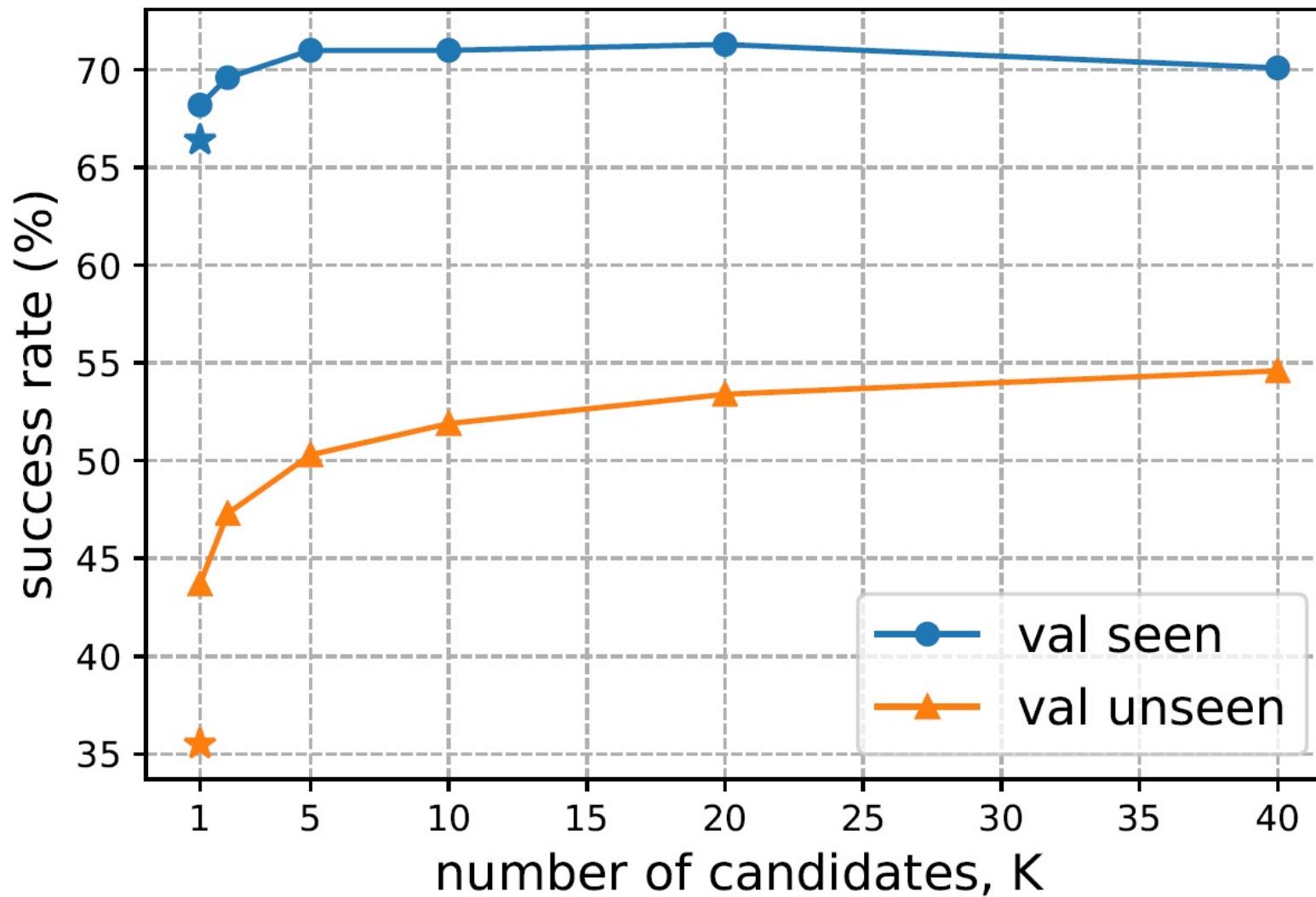
Sheng Shen

# Thanks! Questions?

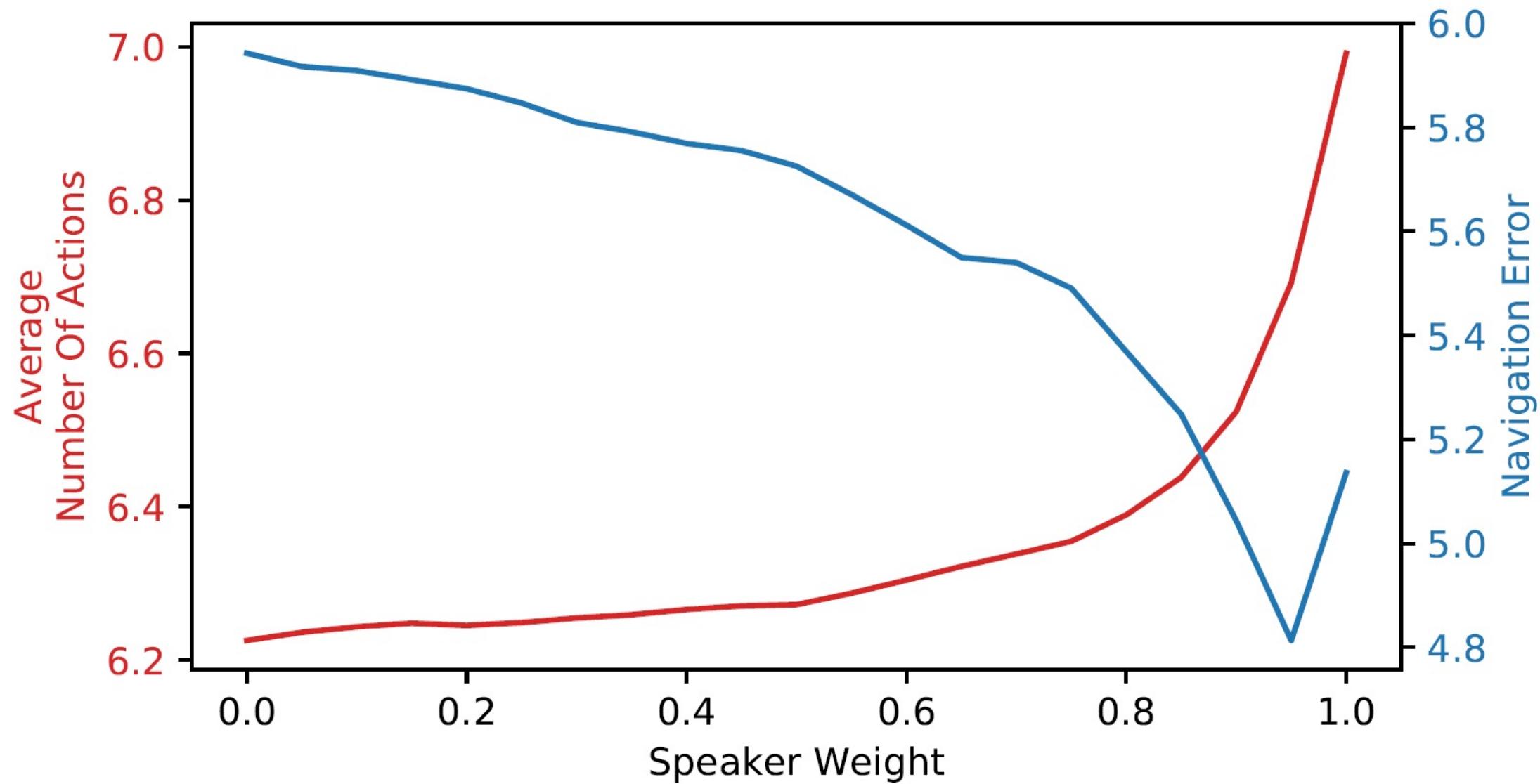
`dfried@andrew.cmu.edu`  
`dpfried.github.io`



# Candidates in Pragmatic Inference

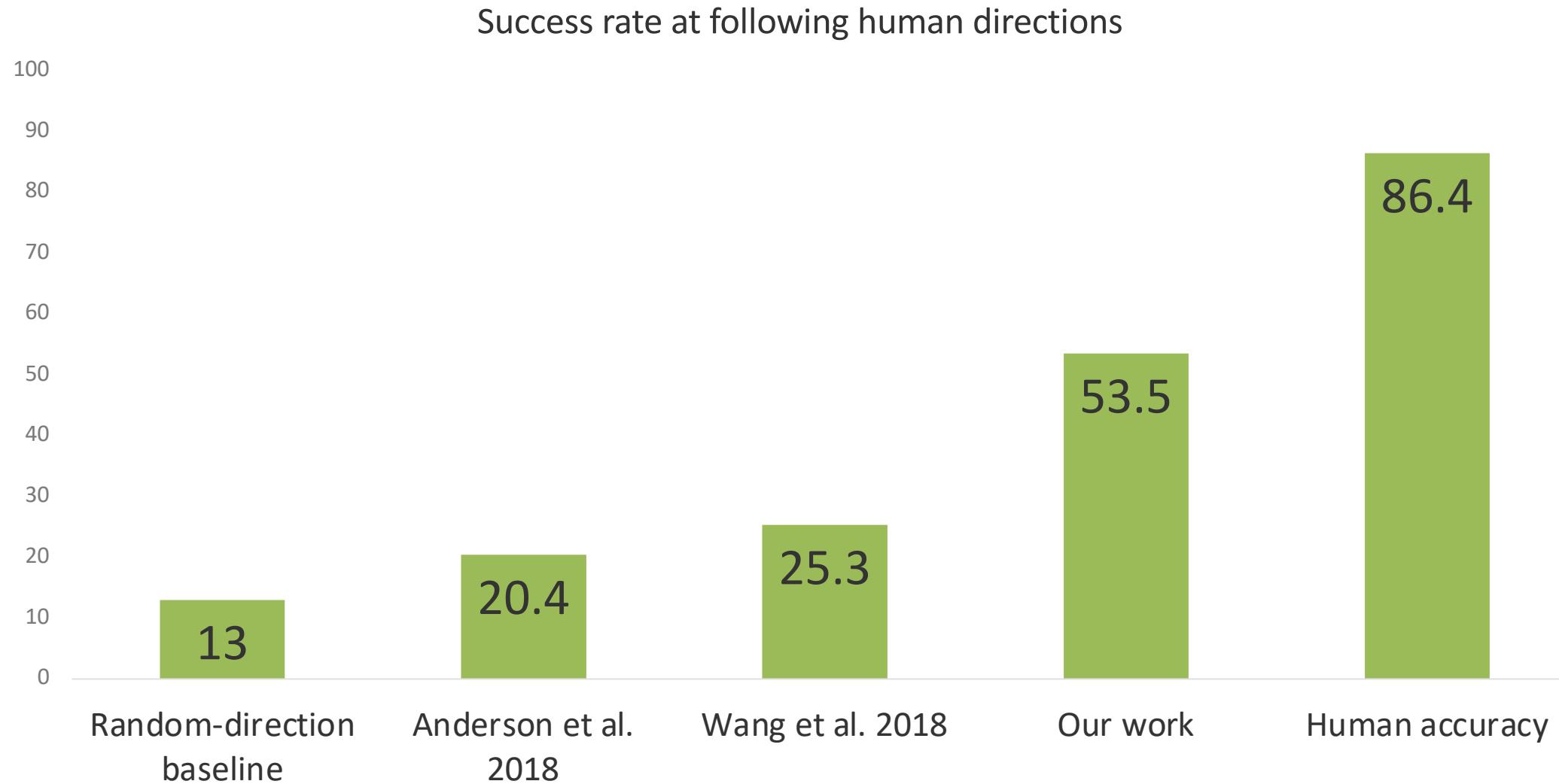


# Pragmatic Path Length and Error Reduction



# Comparison to Prior Work: Test

---



# Pragmatic Speakers in Other Domains

---

## Generation from Meaning Representations

### Input:

Name[Fitzbillies],  
EatType[Coffee Shop],  
PriceRange[Cheap],  
CustomerRating[5 out of 5],  
Area[Riverside],  
FamilyFriendly[Yes]

### Base Speaker:

*Fitzbillies is a family friendly coffee shop located near the river.*

### Pragmatic Speaker:

*Fitzbillies is a family friendly coffee shop that serves English food. It is located in riverside area. It has a customer rating of 5 out of 5 and is cheap.*

## Abstractive Summarization

### Input:

*The 1-0 scoreline that took Barcelona through to the Champions League quarterfinals made their clash with Manchester City all seem rather academic.*

### Base Speaker:

*Barcelona beat Barcelona 1-0 in the Champions League quarterfinals.*

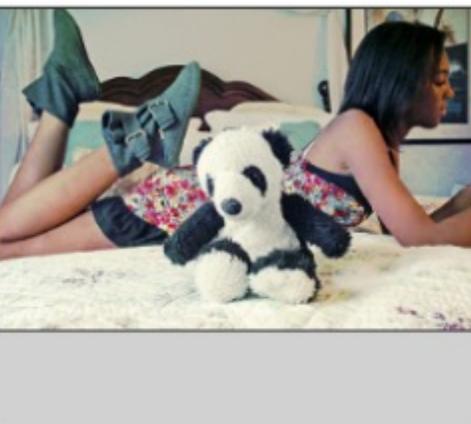
### Pragmatic Speaker:

*Barcelona beat Manchester City 1-0 in the Champions League.*

# Other Communication Tasks

Collaborative visual dialogue (common ground):

A's view:



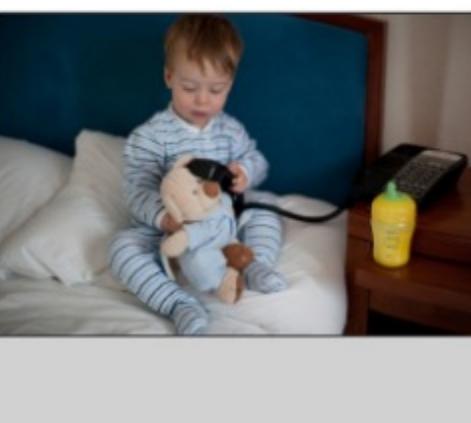
A: Do you have a man with two dogs  
on a bed?

B: With a purple wall in the  
background?

A: Yes

...

*Later, in another round:*



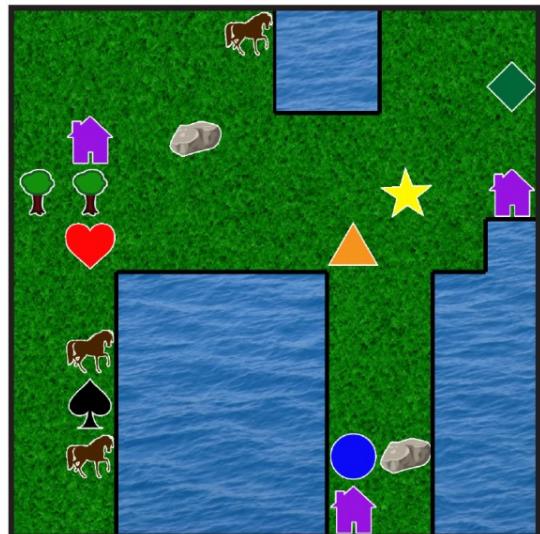
B: I have the man with dogs now

# Other Communication Tasks

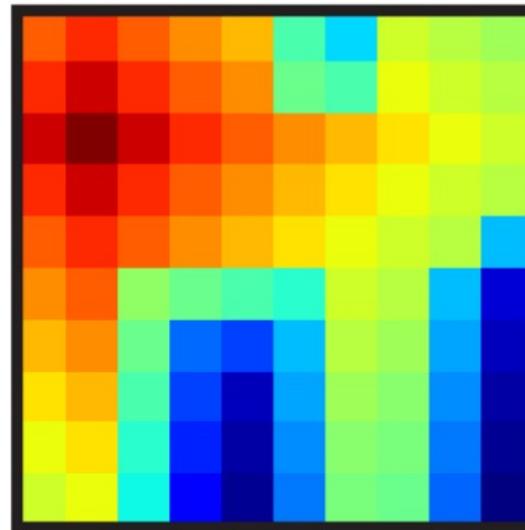
Describing learned behavior:

Reward Functions

World State



Learned Reward Function



*Reach the northernmost house and avoid the water.*

[PuddleWorld, Janner et al. 2017]

Policies and Rollouts



*Build a tower next to the base.*

[MiniRTS, Hu et al. 2019]

# A Modular Neural Architecture

