

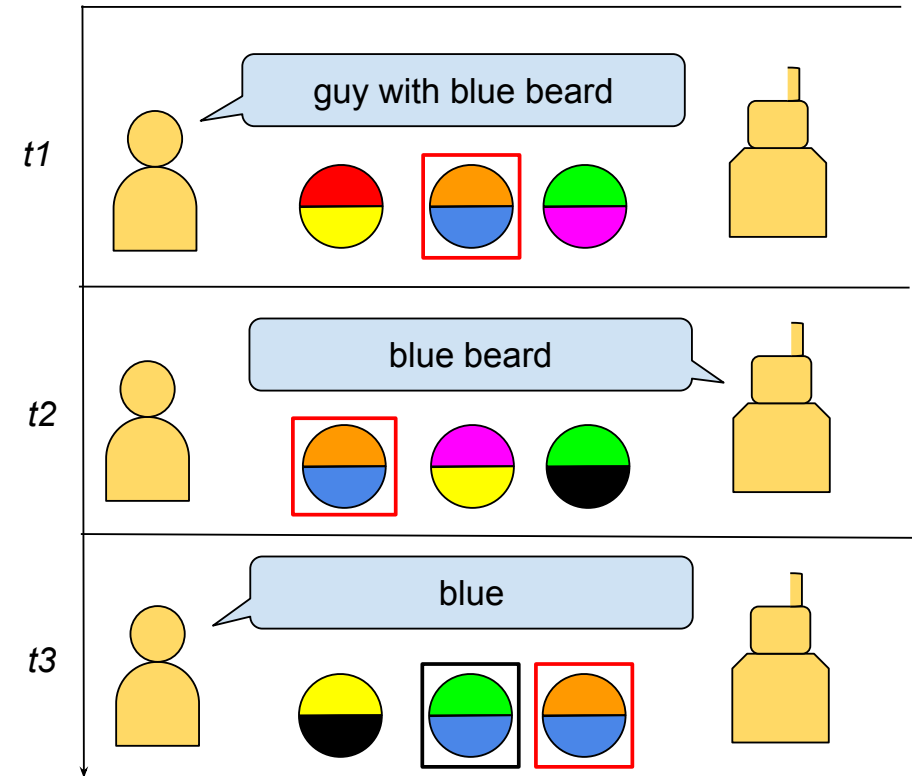
SPOTTER: A Framework for Investigating Convention Formation in a Visually Grounded Human-Robot Reference Task

Jaap Kruijt, Peggy van Minkelen, Lucia Donatelli, Piek Vossen, Elly Konijn,
Thomas Baier

Vrije Universiteit Amsterdam

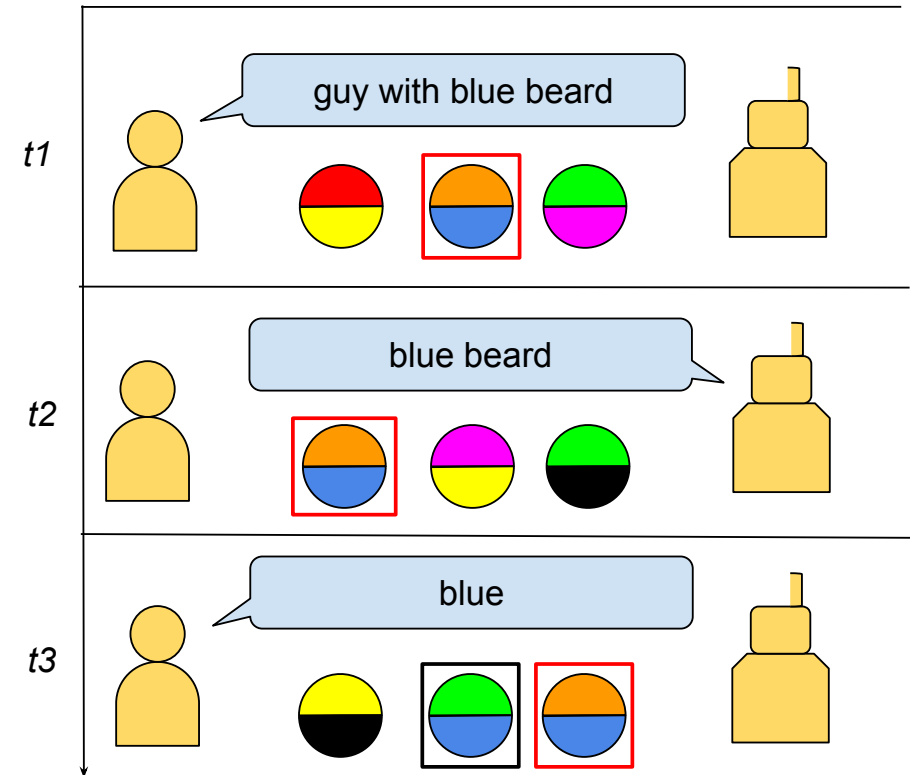
Introduction

- **Conventions** reflect **common ground** between conversation partners
- Linguistic conventions increase communicative efficiency
- Conventions may also help resolve **ambiguity**
- Important features in long-term communication



Introduction

- Existing work has examined convention formation in Human-Human settings:
 - Reduction in utterance length
- What role do linguistic conventions play in **Human-Robot Communication**?
 - In creating efficient communication?
 - In resolving ambiguity?
- To answer these questions, we created a novel experimental framework: SPOTTER



SPOTTER Framework

- Shared Picture Observation Task for Testing Entity References
- Communicative referencing task with **repeated references** (cf. Clark & Wilkes-Gibbs 1986)
- Contrasting **inner circle (InC)** references with **outer circle (OutC)** references which occur only once
 - Common ground within a changing context

1. *...a bald man...*
2. *...the bald man...*
3. *...bald man...*
4. *..baldie...*

InC

OutC

— 1.
— 2.
— 3.
... a bald man with large ears... 4.

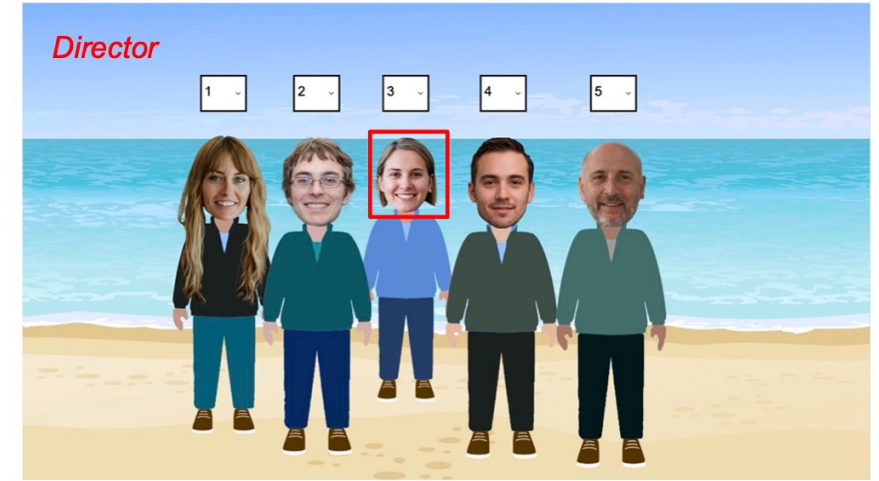
Above: A bald man is referenced in four rounds, with a convention being formed over time. Once a second bald man appears in the fourth round, the convention can help resolve potential ambiguity.

Research Goals

- Main questions:
 - Contrasting the development of references for the InC and OutC
 - Contrasting convention formation in HRI with existing work in H-H interaction
- Contributions:
 - **Novel framework design** for investigating common ground within changing contexts (InC vs OutC). The framework is adaptable and flexible in its use
 - **Annotated dataset** of first results on HRI pilot study in Dutch
 - **In-depth analysis** of pilot study, showing research gaps and providing pointers for future work

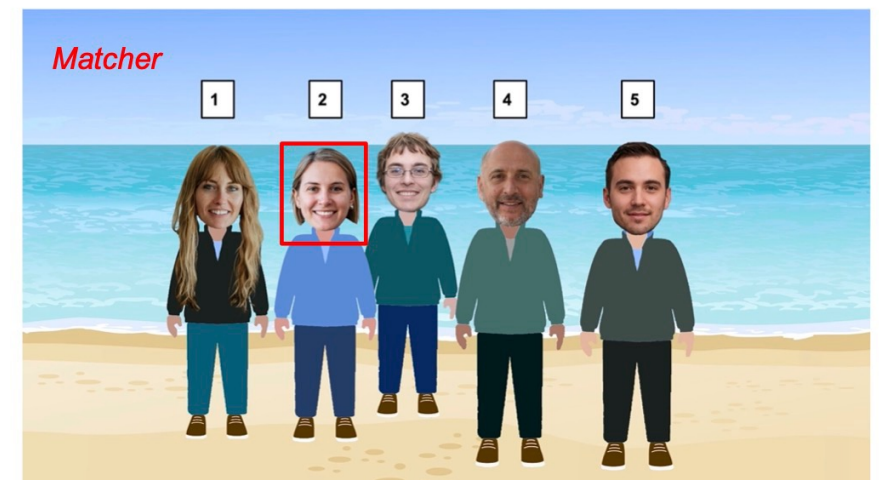
SPOTTER design

- Two-player game
- Six **rounds** with 2 different **scenes** each
- 5 **characters** per round in 5 **positions**
- Goal: locate positions in other player's scene through communication using **mentions**



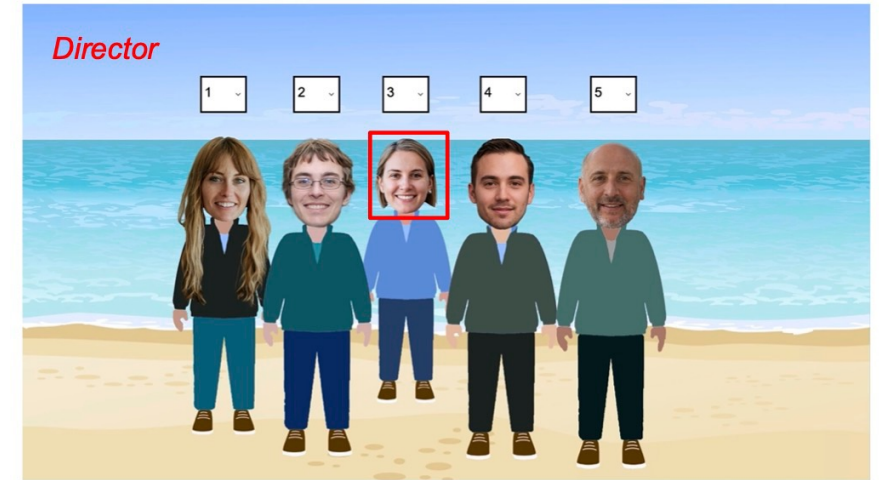
"The woman with the short hair is in spot 3."

"She is in spot 2 for me."



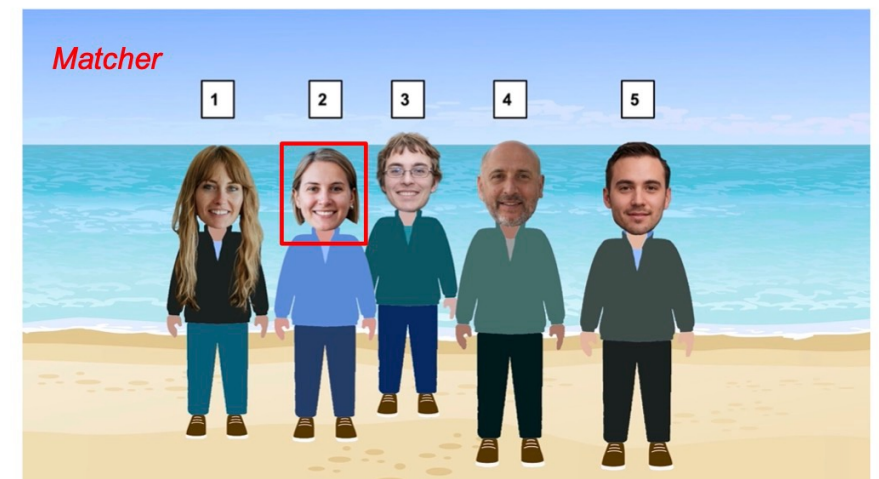
SPOTTER design

- Characters:
 - 3 **main characters** recur each round (InC)
 - 2 **side characters** appear only once (OutC)
- Characters have set of **attributes**
 - E.g. beard, glasses
 - Overlap between attributes creates ambiguity
- (How) do conventions form for the InC?
 - Do conventions decrease in length
- What happens to mentions in case ambiguity is introduced?



“The woman with the short hair is in spot 3.”

“She is in spot 2 for me.”



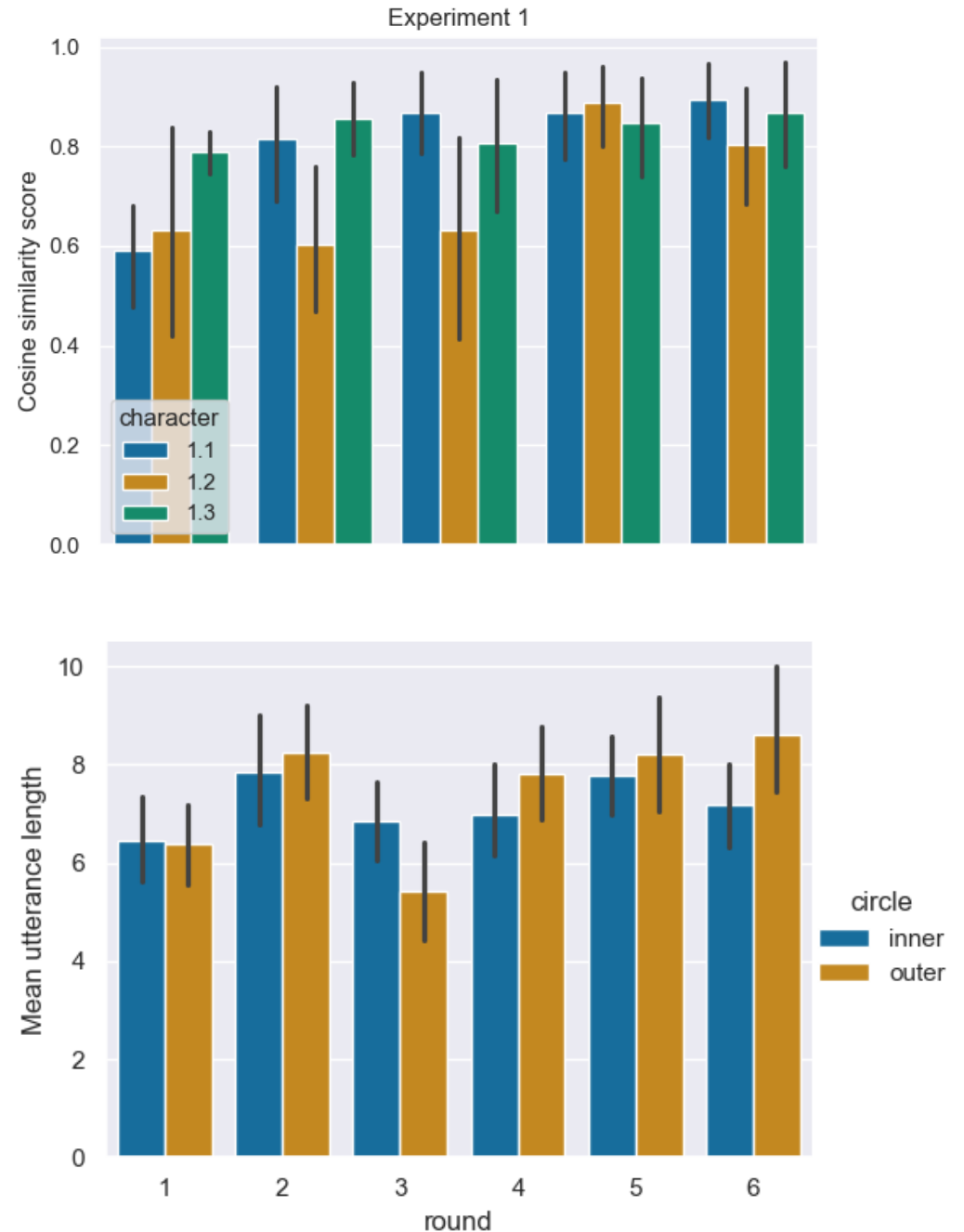
Pilot Study Dataset

- Two HRI in-person pilot studies
 - Wizard-of Oz ('faked' robot)
 - In Dutch
- Data recorded and transcribed using OpenAI's Whisper
- Annotations for i.a.:
 - Mentions
 - Transaction Units and Relations
 - Dialog Acts
- Released for use by research community

	Ex1	Ex2
	# in utterances	
# Utterances	1646	3332
# Mentions	304	567
# Turns	846	1666
Mdn TU length	2	3
# Repair TU's	29	51
	# in TU's	
Mdn # TU's	34	31
Top 2 DA's	statement, pos_answer	statement, back-channeling

Pilot Study Results

- Conventions formed for InC
- InC used as anchor point for OutC
- On average longer mentions for OutC in later rounds
- No reduction in utterance length for InC



Conclusion

- Results on our HRI study differ from results in H-H studies
- Possible effect of perceived robot intelligence?
- Interesting findings regarding the interplay between InC and OutC referring expressions
- First results on pilot study: findings show importance of follow-up studies

Conclusion

- Our SPOTTER framework allows for many different forms of experiments:
 - HRI/Human-Human
 - Online/In person
- Annotated dataset can be used for more analysis and model training
- Future work:
 - Online Human-Human reproduction of pilot studies
 - HRI experiment with autonomous robot
 - Testing influence of multiple interactions of gameplay

Thank you!

j.m.kruijt@vu.nl