Dialogue Structure Annotation for Multi-Floor Interaction

David Traum, Cassidy Henry, Stephanie Lukin, Ron Artstein, Felix Gervitz, Kimberly A. Pollard, Claire Bonial, Su Lei, Clare R. Voss, Matthew Marge, Cory J. Hayes, Susan G. Hill

The work depicted here was sponsored by the U.S. Army. Statements and opinions expressed do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

USC Institute for Creative Technologies

ARL

University of Southern California

Outline

1. Conceptual Framework
   - Meso-level dialogue structure
   - Multi-floor Dialogue & multi-communicators
   - Multi-floor dialogue structure
2. Multi-floor Dialogue Structure Annotation scheme
3. Data
   - Doman: Human-robot collaboration
   - 2 Wizards
   - Example Annotations
   - Corpus Statistics
4. Structure Patterns
5. Uses of data and Future work

Types of Dialogue Structure (Traum & Nakatani 1999)

Structure Content

- Intentional
- Linguistic
- Relational/Rhetorical
- Attentional State
- Turn-taking/floor management
- Grounding
- Participant structure

Structure Granularity

- Micro – within a single turn
- Meso – short subdialogue
- Macro – full conversation

Meso-level Dialogue Structure Annotations

Structure Types

- Intentional:

  Transaction Units – smallest unit of specified and performed action, including all dialogue needed to accomplish this

- Relational/Rhetorical :

  Relations between utterances within a transaction

Annotations

- TUs: cluster of utterances
  - Not necessarily sequential

- Relations: Label 2nd part utterance with
  - Antecedent
  - Relation type

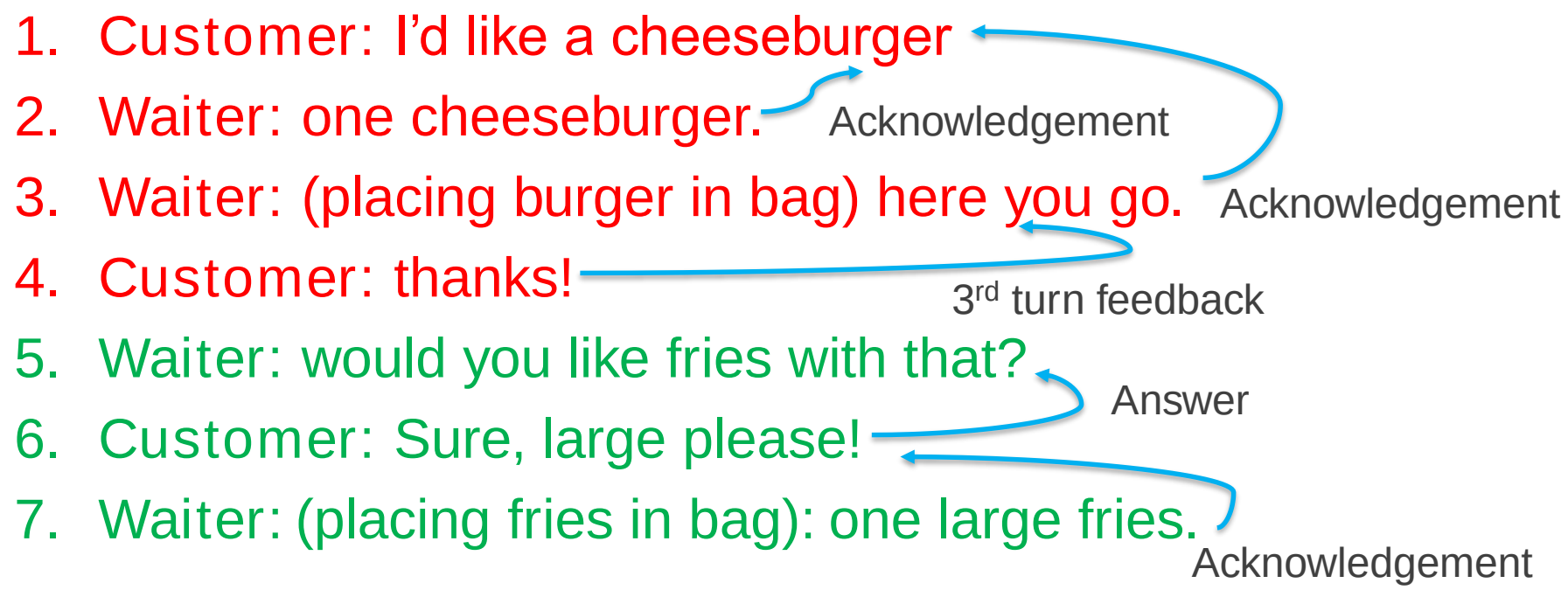
Example: At a lunch counter

- Customer: I'd like a cheeseburger
- Waiter: one cheeseburger.
- Waiter: (placing burger in bag) here you go.
- Customer: thanks!
- Waiter: would you like fries with that?
- Customer: Sure, a large one please!
- Waiter: (placing fries box in bag): one large fries.

Example: Transaction Units (TUs)

- Customer: I'd like a cheeseburger
- Waiter: one cheeseburger.
- Waiter: (placing burger in bag) here you go.
- Customer: thanks!
- Waiter: would you like fries with that?
- Customer: Sure, a large one please!
- Waiter: (placing fries box in bag): one large fries.

Example: Relations



Example: TU Structures

1. Customer: I'd like a cheeseburger
2. Waiter: one cheeseburger.
3. Waiter: (placing burger in bag) here you go.
4. Customer: thanks!
5. Waiter: would you like fries with that?
6. Customer: Sure, large please!
7. Waiter: (placing fries in bag): one large fries.

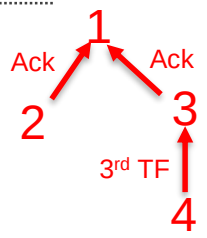
Acknowledgement

Acknowledgement

3rd turn feedback

Answer

Acknowledgement



Floor and Participant Structure

Participants and Floors

- Single floor Dyadic (A,B)
- Single floor Multiparty: (A,B,C,…)
- Multiple floors (with different sets of participants): {(A B), (C D E)}

Kinds of Interactions between Floors

- Same purpose, distinct participants
- Co-located, observable
  - Participants play different roles for different floors (e.g. active participant vs overhearer)
- Some Shared participant(s)
  - multi-communicating (Rentch et al)

- Multi-floor dialogue:
  - Same purpose
  - Some Multi-communicating participant(s)
  - Content flows across floors

Examples of (observable) Multi-floor dialogue

Indirect Action

Live Interpretation

Multi-floor Relation types

- Expansions - relate utterances that are produced by the same participant within the same floor.
- Responses - relate utterances by different participants in the same floor.
- Translations - relate utterances in different floors

Examples:

1. (A,B) A->B: I'll have a cheeseburger
2. (A,B) A->B: and a small coke

1. (A,B) A->B: a small coke
2. (A,B) B->A: no coke, pepsi

1. (A,B) A->B: I'll have a cheeseburger
2. (B,C) B->C: Cheeseburger!!

Relations by type (1)

Expansions

a) Continue
b) (self-) Correction
c) Link-next
d) Summarization

Translation

a) Translation <from,to>
b) Partial
c) Quotation
d) Comment

Relations by type (2) Responses

a. Processing: positive feedback at perception level
b. acknowledgement: positive feedback of understanding
c. clarification: negative feedback of understanding
d. question-response
e. reciprocal response: e.g. “hello” -> “hello”
f. 3rd turn feedback: response to feedback
g. other

Response sub-relations

acknowledgment

- ack-done
- ack-doing
- ack-wilco
- ack-understand
- ack-try
- ack-unsure
- ack-cant

clarification

- req-clar
- clar-repair
- missing info
- nack
- req-repeat
- clar-repeat

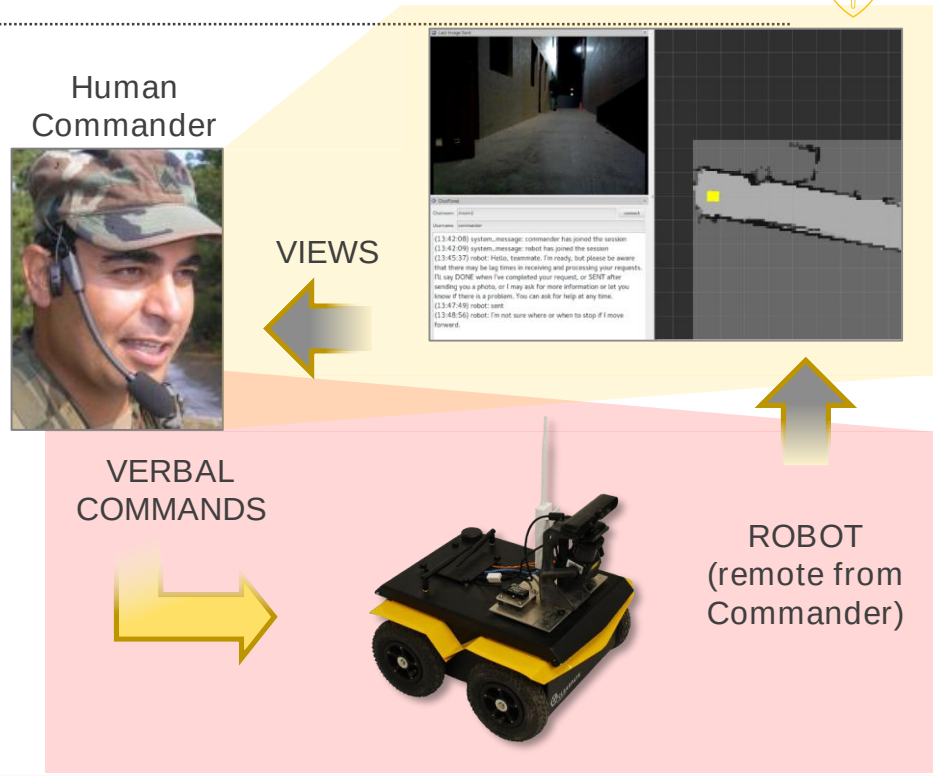
question-response

- answer
- Non-Answer-Response (NAR)

Domain: Human-Robot Collaboration

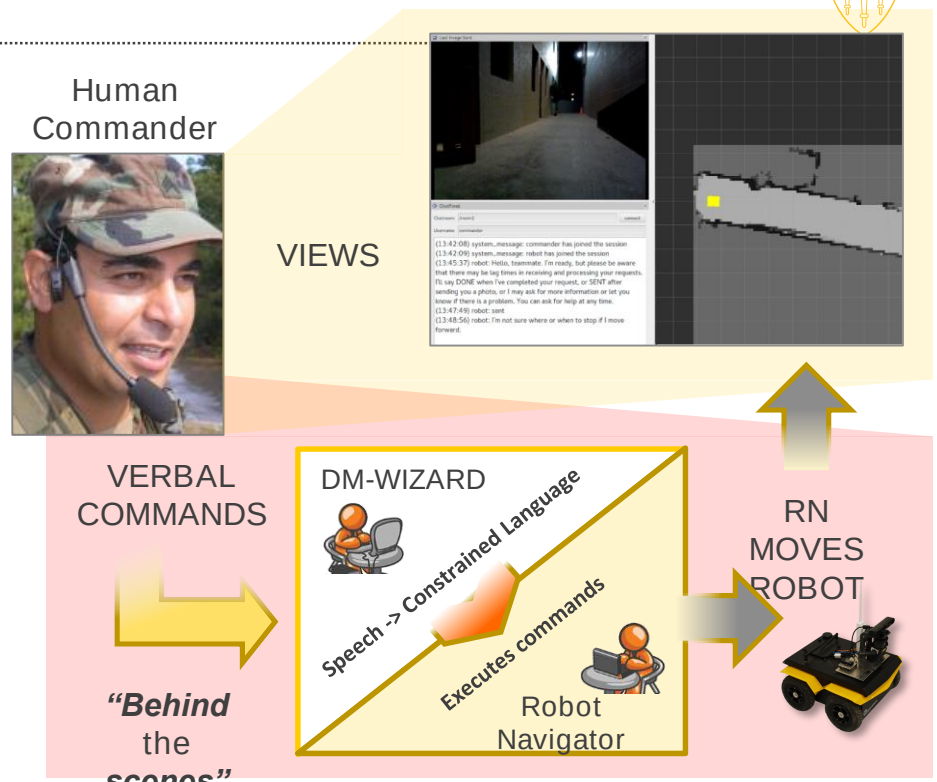
Remote reconnaissance task
- Unfamiliar environment
- Bandwidth limitations
- User and robot not co-present

- What would the human users want to say?
  - Need to collect a corpus in order to train and evaluate the system.
- How would users naturally collaborate with this robot teammate?

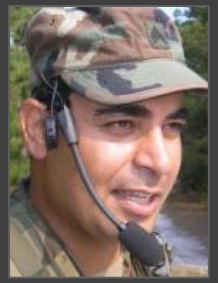


Multi-floor data collection setting

- Robot assisted by two human "wizards"
  - Dialogue Manager (DM) is the language "brain" of the robot
  - Robot Navigator (RN) moves robot based on instructions



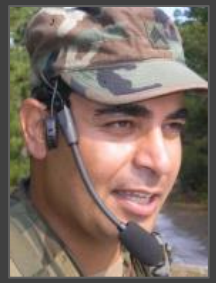
Example Interaction



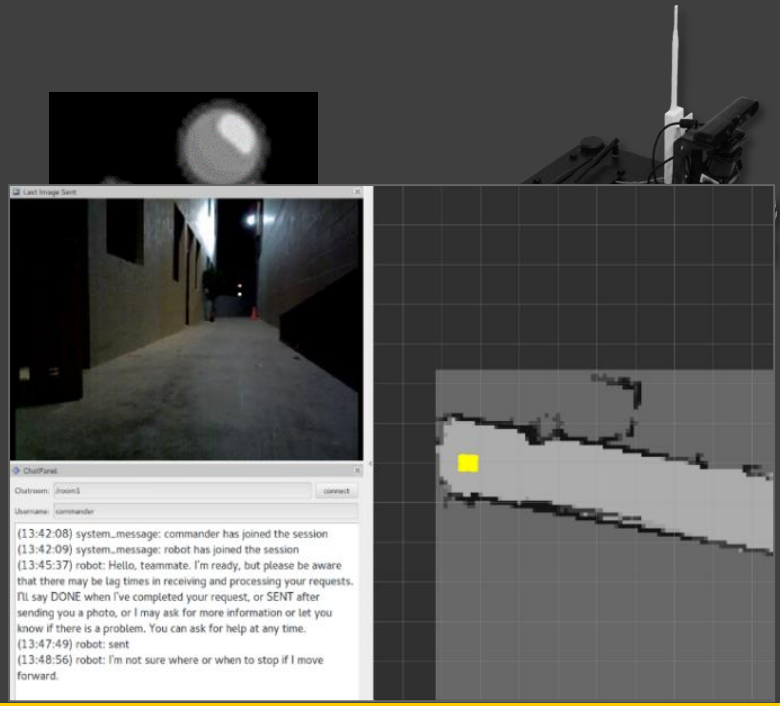
DM

RN

Commander



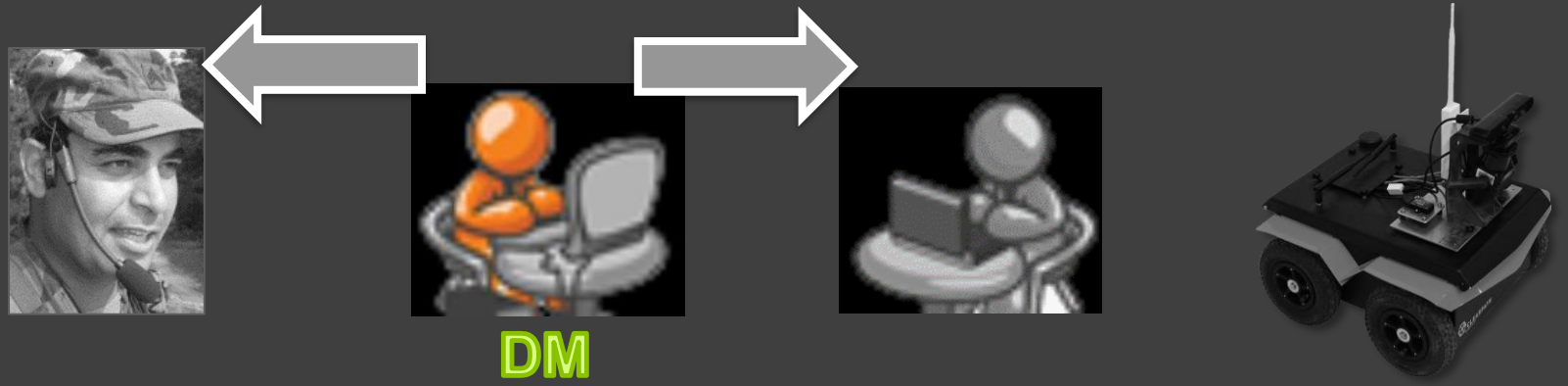
CMD



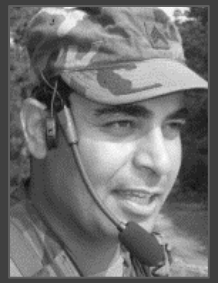
Commander – Human Participant

- Verbally Instructs a Robot
- Sees text message responses, LIDAR map, and images sent from onboard robot

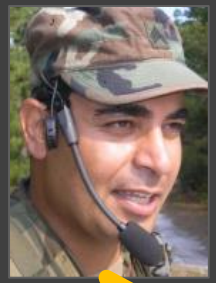
Wizard #1 – Dialogue Manager



Wizard #2 – Robot Navigator



Example Interaction

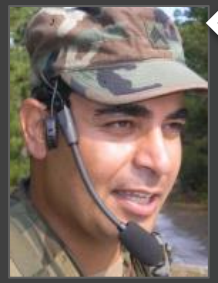


DM

RN

Proceed forward

Example Interaction

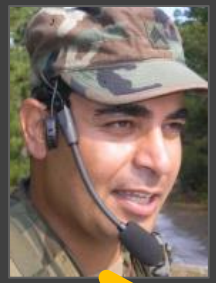


DM

RN

How far? You can tell me to move to an object that you see, or a distance

Example Interaction

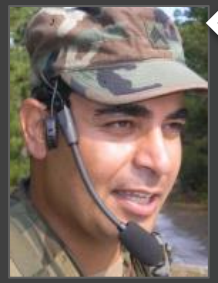


DM

RN

Proceed forward three feet

Example Interaction

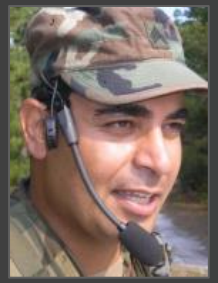


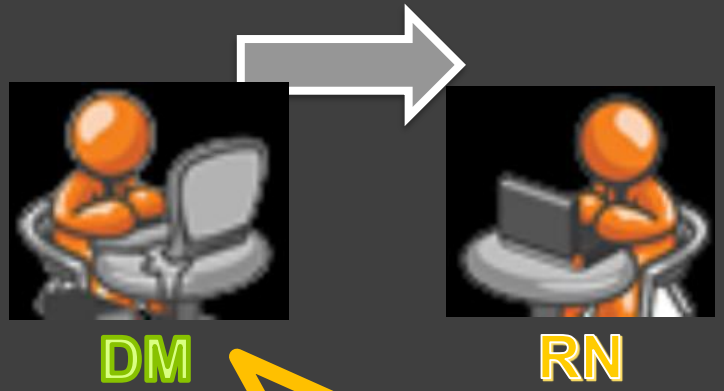
DM

RN

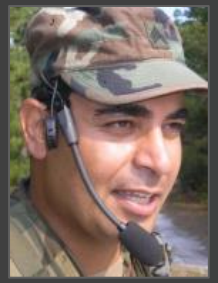
Executing…

Example Interaction





Example Interaction



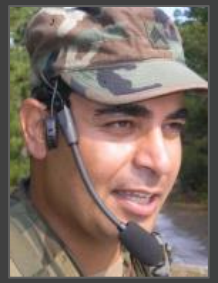
DM

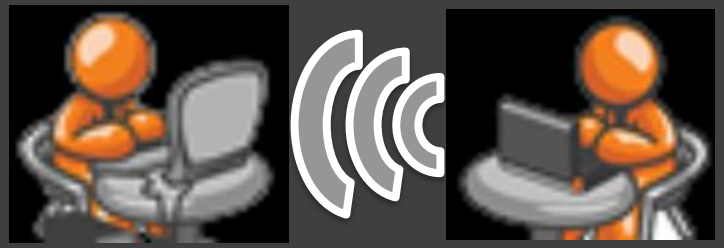
RN

move

moves robot forward 3 feet

Example Interaction



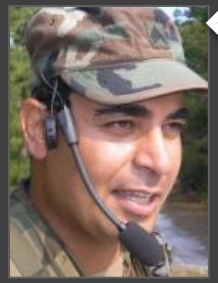


DM

RN

done

Example Interaction



DM

RN

done

Data - Transcripts

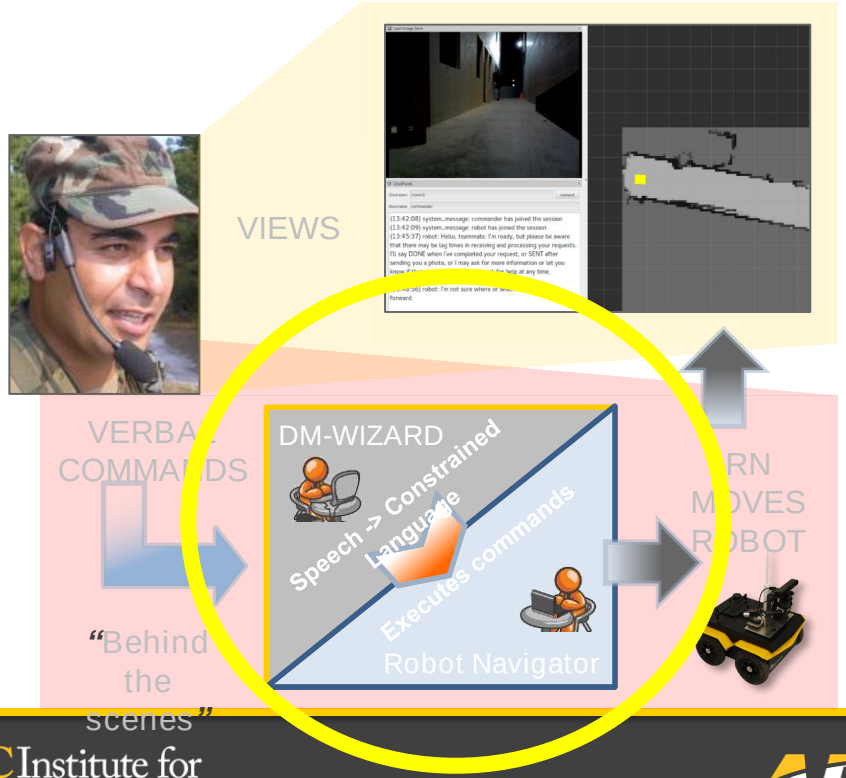
- Time aligned transcripts of 4 data streams
  - 2 audio streams
    - CMD and RN
  - 2 text streams
    - DM->CMD, DM->RN
- Two conversational floors present

Commander (Audio Stream 1)	DM->Commander (Chat Room 1)	DM->RN (Chat Room 2)	RN (Audio Stream 2)
face the doorway on your right			
and take a picture			
	there's a door ahead of me on the right and one just behind me on the right. which would you like me to face?		
the door ahead of you on the right			
		move to face the door ahead of you on the right, image	
	executing...		
			image sent
	sent		

Left floor: CMD, DM

Commander (Audio Stream 1)	DM->Commander (Chat Room 1)	DM->RN (Chat Room 2)	RN (Audio Stream 2)
face the doorway on your right			
and take a picture			
	there's a door ahead of me on the right and one just behind me on the right. which would you like me to face?		
the door ahead of you on the right			
		move to face the door ahead of you on the right, image	
	executing...		
			image sent
	sent		

Right Floor: DM, RN



Commander (Audio Stream 1)	DM->Commander (Chat Room 1)	DM->RN (Chat Room 2)	RN (Audio Stream 2)
face the doorway on your right			
and take a picture			
	there's a door ahead of me on the right and one just behind me on the right. which would you like me to face?		
the door ahead of you on the right			
		move to face the door ahead of you on the right, image	
	executing...		
			image sent
	sent		

DM translates (to) left and right

Commander (Audio Stream 1)	DM->Commander (Chat Room 1)	DM->RN (Chat Room 2)	RN (Audio Stream 2)
face the doorway on your right			
and take a picture			
	there's a door ahead of me on the right and one just behind me on the right. which would you like me to face?		
the door ahead of you on the right			
		move to face the door ahead of you on the right, image	
	executing...		
			image sent
	sent		

Corpus Statistics

Basics

- 60 dialogues
  - 20 participants
  - 3 dialogues each
  - ~20 hours
- 11454 Total Utterances
  - 3,573 from commanders
  - 5,154 from DM
  - 2,727 from RN

Dialogue Structure Annotations

- 2,230 Transaction Units
- 11,058 Relations
- 644 Unique TU Tree structures
  - Classified into 5 types

Frequent Relations

Type	Subtype	#	%
Translation		4282	39
	Translate-r	2355	21
	Translate-l	1911	17
	comment	21	<1
Expansion		1583	14
	Continue	1175	11
	Link-next	337	3
	correction	50	<1
	summarize	20	<1

Type	Subtype	#	%
Response		5193	47
	acknowledge	3998	36
	clarification	569	5
	processing	315	3
	Question-response	212	2
	other	48	<1
	3rd turn feedback	37	<1
	reciprocal	14	<1

Structural Types of Transactions (TUs)

- Minimal TU: single instruction, acks, no repair
- Extended-Link TU: multiple instructions, with expansions
- Repair TU: contains at least one repair
  - successfully resolved or
  - abandoned
- QA TU: starts with question & response rather than instruction
  - simple question,
  - later instruction
- Other TU: none of the above (e.g. no response or translation)

Example minimal TU

	Left Floor		Right Floor		Annotations		
Utt #	**Commander**	**DM→CMD**	**DM→RN**	**RN**	**TU #**	**Antecedent**	**Relation**
1	move forward three feet				1		
2		ok			1	1	ack-wilco
3			move forward 3 feet		1	1	translation-r
4				done	1	3	ack-done
5		I moved forward 3 feet			1	4	translation-l

Structural Types of Transactions (TUs)

- Minimal TU: single instruction, acks, no repair
- Extended-Link TU: multiple instructions, with expansions
- Repair TU: contains at least one repair
  - successfully resolved or
  - abandoned
- QA TU: starts with question & response rather than instruction
  - simple question,
  - later instruction
- Other TU: none of the above (e.g. no response or translation)

Example Extended-Link TU

	Left Floor		Right Floor		Annotations		
Utt #	Commander	DM→CMD	DM→RN	RN	TU	Ant	Rel
1	face west				1		
2	and take a photo				1	1	continue
3			face west, photo		1	2*	translation-r
4		executing...			1	2*	ack-doing
5				image sent	1	3	ack-done
6		sent			1	5	translation-l

Structural Types of Transactions (TUs)

- Minimal TU: single instruction, acks, no repair
- Extended-Link TU: multiple instructions, with expansions
- Repair TU: contains at least one repair
  - successfully resolved or
  - abandoned
- QA TU: starts with question & response rather than instruction
  - simple question,
  - later instruction
- Other TU: none of the above (e.g. no response or translation)

Example Repair TU

	Left Floor		Right Floor		Annotations		
Utt #	Commander	DM→CMD	DM→RN	RN	TU	Ant	Relation
1	move to where you see the first cone				1		
2		I'm not sure which object you are referring to. Can you describe it in another way, using color or its location?			1	1	request- clarification
3	move to the cone on the right a red cone on the right				1	2	clarification- repair
4			move to face the cone on the right		1	3	translation-r
5		executing...			1	3	ack-doing
6	take another picture				2		
7				done	1	4	ack-done
8		done			1	7	translation-l
9			image		2	6	translation-r
10				image sent	2	9	ack-done

Structural Types of Transactions (TUs)

- Minimal TU: single instruction, acks, no repair
- Extended-Link TU: multiple instructions, with expansions
- Repair TU: contains at least one repair
  - successfully resolved or
  - abandoned
- QA TU: starts with question & response rather than instruction
  - simple question,
  - later instruction
- Other TU: none of the above (e.g. no response or translation)

Example Q&A TUS

	Left Floor		Right Floor		Annotations		
Utt #	Commander	DM→Commander	DM→RN	RN	TU	Ant	Rel
1	how many window openings do you see in front of you				1		
2		three			1	1	answer
3	do you see a yellow flashlight				2		
4		processing...			2	3	processing
5		I'm not sure			2	3	answer
6		If you describe an object, you can help me to learn what it is.			2	3	non-answer response

Structural Types of Transactions (TUs)

- Minimal TU: single instruction, acks, no repair
- Extended-Link TU: multiple instructions, with expansions
- Repair TU: contains at least one repair
  - successfully resolved or
  - abandoned
- QA TU: starts with question & response rather than instruction
  - simple question,
  - later instruction
- Other TU: none of the above (e.g. no response or translation)

Examples of Other TU

	Left Floor		Right Floor		Annotations		
Utt #	Commander	DM→Commander	DM→RN	RN	TU	Ant	Rel
1	i'm ready				1		
2		I'm also ready			1	1	Reciprocal-response
3		Would you like me to send a picture so you can see the room?"			2		
4	Turn 90 degrees left				3		

Frequency of TU Structures (% of corpus)

- Minimal TU (48%)
- Extended-Link TU (26%)
- Repair TU (11%)
  - 9% successfully resolved
  - 2% abandoned
- QA TU (~5)%
  - 4% simple question
  - 1% lead to instruction
- Other TU (11%)

Applications of Annotated Data

- Examination of Dialogue Structure Overlap (Henry et al WiNLP 2018)
- Stylistic differences across individuals and conditions (Lukin et al Sigdial 2018)
- Automating NLU and dialogue management (Gervits et al ACL 2018 Demo)

Future Work

- More data collection – in simulation, further annotation
- Analysis of other levels – dialogue act type, parameter type, etc.
- Analysis of other multi-floor dialogue corpora
  - Simultaneous interpretation
  - Observability of other floors
    - Observable (e.g. restaurant ordering)
    - Semi-observable (e.g. interpretation to another language)
    - Non-observable (Botlanguage)

Thank You!

- Questions?