

# Annotating anaphoric reference in dialogue: the CODI/CRAC 2022 Shared Task corpus

**Massimo Poesio and Lori Levin**

(Queen Mary University) - (Carnegie-Mellon University)

(Joint work with **Maris Camilleri, Paloma Carretero Garcia, Taiqi He, Mark-Christoph Müller, Carolyn Rose, Michael Strube, Juntao Yu, and Katherine Zhang**)

<http://universalanaphora.org>

<https://arciduca.org/>

CODI/CRAC Shared Task Workshop, October 2022

# Motivations for the CODI/CRAC Shared Tasks and Corpus

# Broadening the scope of anaphoric interpretation

- ▶ Up until fairly recently anaphoric interpretation has focused primarily on one genre (NEWS) and on one type of anaphoric reference (IDENTITY REFERENCE, or COREFERENCE), but this is rapidly changing
- ▶ NEW GENRES:
  - ▶ BIOMEDICAL TEXTS (e.g., CRAFT (Cohen et al., 2017))
  - ▶ ENCYCLOPEDIA articles (e.g., LIVEMEMORIES (Rodriguez et al., 2010) / WIKICOREF (Ghaddar and Langlais, 2016) / PHRASE DETECTIVES (Poesio et al., 2019))
  - ▶ FICTION (e.g., LITBANK (Bamman et al., 2020) / PHRASE DETECTIVES)
- ▶ BENCHMARKS to explore some issues in depth, as well as full-corpus datasets
  - ▶ E.g., Winograd Schema Challenge (Rahman&Ng 2012), GAP (Webster et al, 2018)

## Still under-researched

- ▶ Anaphoric reference in other languages (see the CRAC 2022 Shared Task on Multilingual Coreference Resolution, this conference)
- ▶ **OTHER TYPES OF ANAPHORIC REFERENCE**
  - ▶ Bridging (e.g., ISNOTES, ARRAU, PDT)
  - ▶ Discourse deixis (e.g., ARRAU, PDT)
- ▶ **DIALOGUE**

# Anaphoric Reference in Dialogue

# Anaphora and Reference in Dialogue

A number of issues in interpreting anaphoric reference are specific to reference in dialogue:

- ▶ Disfluencies and repairs
- ▶ Abandoned and discontinuous markables
- ▶ First and second person pronouns (not always deictic!)
- ▶ Questions
- ▶ More ambiguous / vague references
- ▶ Expressions used to establish grounding instead of referring

## Disfluencies and Repairs

U : okay

um

how long will it take engine one to get to **Danville**  
**Dansville**

▶ (trains93 subcorpus, dia18-1)

U : okay so that's four hours

S : um

U : or **three**

S : **[three hours]**

▶ (trains93 subcorpus, dia18-4)

(From the ARRAU corpus, Trains-93 subcorpus (Poesio and Artstein, 2008; Uryupina et al., 2020))

## Discontinuous markables: Completions

(1)	1.1	Inst	So, jetzt nimmst Du [pause] <i>Well, now you take</i>
	1.2	Cnst	eine Schraube <i>a screw</i>
	1.3	Inst	eine < – > orangene mit einem Schlitz <i>an &lt; – &gt; orange one with a slit</i>
	1.4	Cnst	Ja <i>Yes</i>

(From the Bielefeld Toy Plane Corpus (Skuplik, 1999; Poesio and Rieser, 2010))



## Second person pronoun references

S: hello how can I help [you]

▶ (trains93 subcorpus, D93\_26\_3)

U: can [you] do can [you] do things simultaneously here or  
do they have to be done like can I have the same time  
having it the engine

▶ (trains93 subcorpus, D93\_26\_3)

U: when [you] get to Dansville

▶ (trains93 subcorpus, D93\_22\_3)

## First person pronoun references?

U: okay

[we]'re still in Corning

- ▶ (trains93 subcorpus, dia18-4)

## Ambiguous references

- 1.1 M : all right system  
1.2 : we've got a more complicated problem  
1.4 : first thing \_I'd\_ like you to do  
1.5 : is send engine E2 off with a boxcar to Corning  
to pick up oranges  
1.6 : uh as soon as possible  
2.1 S : okay  
3.1 M : and while it's there it should pick up the tanker  
(2) 4.1 S : okay  
4.2 : and that can get  
4.3 : we can get that done by three  
5.1 M : good  
5.3 : can we please send engine E1 over to Dansville  
to pick up a boxcar  
5.4 : and then send [it]<sub>i</sub> right back to Avon  
6.1 S : okay  
6.2 : it'll get back to Avon at 6

(From the TRAINS corpus (Heeman and Allen, 1995; Poesio and Artstein, 2005))

## Vague / unclear references

U : um

can you do can you do things simultaneously here  
or do they have to be done like one like  
can I have the same time having [it] the engine

▶ (trains93 subcorpus, D93\_26\_3)

U : and fill up

and engine two to Dansville

S : okay

U : pick up those three boxcars and bring them to Corning  
and engine three with those two boxcars to Corning

S : okay

U : okay and then fill them up

S : uh-huh

U : uh and bring them to Elmira

S : okay let me do [this] real quick to make sure you  
have enough time six a.m. seven okay you have done [it]

▶ (trains93 subcorpus, D93\_25\_5)

# Questions

U : so **how long** will it take if I take the two boxcars  
ship from Elmira to Corning load up the oranges take  
the two tankers to Elmira and then make orange juice  
there

S : **[eight hours]**

U : eight hours

▶ (trains93 subcorpus, D93\_26\_5)

## Referring expressions used for grounding

U : so how long will it take if I take the two boxcars  
ship from Elmira to Corning load up the oranges take  
the two tankers to Elmira and then make orange juice  
there

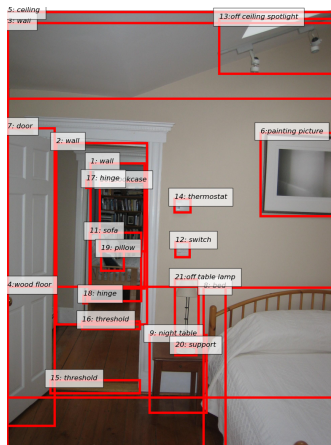
S : [eight hours]

U : **eight hours**

▶ (trains93 subcorpus, D93\_26\_5)

# Situated reference

(From the *Tell-me-more* corpus, Studied in joint work by Loaiciga, Dobnik and Schlangen - see other invited talk at CRAC 2022)



1. it's a bedroom scene with the bed partially visible
2. the bed has a curved wooden headboard with slots like a fence
3. there is framed art hanging above the bed
4. to the left of the bed is a door, which is open
5. there is a small square nightstand next to the bed which has a lamp on top of it

## Previous efforts at annotating anaphoric reference in dialogue

- ▶ Donna Byron's annotation of the TRAINS-93 corpus (Byron, 2002)
- ▶ Christoph Mueller's annotation of the Switchboard corpus (Müller and Strube, 2003a)
- ▶ French ANCOR corpus (Muzerelle et al., 2014)
- ▶ **ARRAU** (Poesio and Artstein, 2008; Uryupina et al., 2020)



# Annotation of anaphora and reference in dialogue in ARRAU

- ▶ The ARRAU corpus (Poesio and Artstein, 2008; Uryupina et al., 2020) is a corpus covering several genres including two dialogue genres
  - ▶ The TRAINS corpus
  - ▶ The Pear Stories corpus
- ▶ Its annotation guidelines cover most of the information included in the current Universal Anaphora proposal
- ▶ The corpus is annotated with the MMAX2 tool (Müller and Strube, 2003b) which makes it possible to annotate
  - ▶ Discontinuous markables
  - ▶ Ambiguity

# ARRAU: Discontinuous markables in MMAX2

The screenshot displays the MMAX2 software interface, which is used for linguistic annotation. The window title is "[three][loaded boxcars] [dirty]". The interface is divided into several panels:

- Settings Panel:** Contains various annotation options. The "Comment" field is set to "erring (grounding)". The "min\_words" field is set to "boxcars". The "min\_ids" field is set to "word\_226". Under "Gender", "neuter" is selected. Under "Number", "plur" is selected. Under "Person", "per3" is selected. Under "Gram\_fnc", "unsure-gf" is selected. Under "< > Reference", "non\_referring" is selected. Under "non\_ref\_type", "unknown" is selected. There are also checkboxes for "Suppress check" (checked) and "Warn on extra attributes" (unchecked). Buttons for "Apply" and "Undo changes" are present.
- Transcript Panel:** Shows a transcript of a conversation. The text is as follows:  
S: hello can I help you  
U: um okay I need to transport four boxcars of oranges to Bath by two PM  
S: okay four boxcars of oranges to Bath  
U: mm  
S: what- at what time  
U: by two PM  
S: okay  
U: how long will it take me to get engine two along with two boxcars from Elmira to Corning  
S: two hours  
U: okay  
um  
how long will it take engine one to get to Danville Dansville  
S: engine one from Avon  
U: mm-hm  
S: uh it will take three hours  
U: okay and how long will that take will it take for um engine one at Dansville  
S: mm-hm  
U: to bring two boxcars along and go to Corning  
S: another hour  
U: okay  
S: so four hours in all| from Avon- starting from Avon  
U: okay now I have four boxcars at Corning um  
S: okay  
U: how long will it take me to load um the oranges at Corning onto the two boxcars from Elmira and the two boxcars from Dansville  
S: one hour you can do all of it in one hour  
U: okay and then to transport the four boxcars carried by engine two with the oranges to Bath from Corning  
S: okay if um you can only pull **three loaded boxcars**  
U: **three**  
S: yeah **loaded boxcars**  
U: by one engine  
S: with one engine right  
U: with one engine  
okay um so  
how long will it take me to uh carry at- use engine two to carry three box car- cars of loaded oranges from Corning to Bath  
S: b- two hours you've got another boxcar there right  
U: right  
S: y- I mean another uh engine  
U: an engine  
S: yeah  
U: but two engines can n't go at the same time

Below the transcript, a list of levels is shown: [Corning], [you], [three loaded boxcars], [three][loaded boxcars], [one engine].

At the bottom of the settings panel, the text "Auto-apply is ON" is displayed in red.

# ARRAU: Ambiguity in MMAX2

The screenshot shows the MMAX2 software interface with the following settings for an utterance:

- Reference:  unmarked,  new,  old,  non\_referring,  undef\_reference
- Category:  unmarked,  person,  animate,  substance,  medicine,  concrete
- Ref\_type:  phrase,  segment
- On\_map:  no,  yes
- Object: engine\_e1
- Phrase\_Antecedent:  single\_phrase,  multiple\_phrases
- Single\_phrase\_antecedent: 14
- Related\_object:  no,  yes
- Generic: generic-no
- Ambiguity:  unambiguous,  ambiguous,  ambiguous\_antecedent
- Ref\_type\_2:  phrase,  segment,  new,  non\_referring
- Category\_2:  unmarked,  person,  animate,  substance,  medicine,  concrete
- On\_map\_2:  no,  yes
- Object\_2: boxcar\_from\_dansville
- Phrase\_Antecedent\_2:  single\_phrase,  multiple\_phrases

Buttons: Apply, Undo changes

Auto-apply is ON

MMAX2MarkableBrowser [1]

Levels: Order  KWIC view Font  Auto-select

Avon]

Avon]

Poesio and Artstein, 2005)

The transcript window shows a dialogue with the following text:

M : all right system  
we've got a more complicated problem  
uh  
first thing I'd like you to do  
is send engine E2 off with a boxcar to Corning to pick up oranges  
uh as soon as possible  
S : okay  
M : and while it's there it should pick up the tanker  
S : okay  
and that can get  
we can get that done by three  
M : good  
uhh  
can we please send engine E1 over to Dansville to pick up a boxcar  
and then send it right back to Avon  
S : okay  
it'll get back to Avon at 6  
M : m kay  
uh  
how long will it take it to uh /  
pick up bananas at Avon  
and  
deliver them to Corning  
S : okay  
to pick up bananas at Avon

Annotations in the transcript include a green highlight under "send engine E1 over to Dansville to pick up a boxcar" and a yellow highlight under "it right back to Avon".

# The CODI/CRAC 2022 Shared Task and Corpus

# The CODI/CRAC Shared Tasks on Anaphoric Reference in Dialogue (2021, 2022)

- ▶ OBJECTIVE: to promote research on a wide range of anaphoric reference phenomena in dialogue
- ▶ TASKS
  - ▶ Identity anaphora
  - ▶ Bridging reference
  - ▶ Discourse deixis
- ▶ Created a NEW DATASET: the CODI/CRAC Corpus, including data from four well-known dialogue datasets
  - ▶ The focus of this talk
- ▶ Developed a new scorer (Yu et al, 2022)
  - ▶ See separate presentation
- ▶ <https://codalab.lisn.upsaclay.fr/competitions/614>
- ▶ See talks in the CODI/CRAC workshop for more details

# The CODI/CRAC 2022 Corpus

- ▶ Data from four well-known dialogue datasets:
  - ▶ SWITCHBOARD (Stolcke et al, 2000; Nissim et al, 2004)
  - ▶ AMI (Carletta, 2006)
  - ▶ LIGHT (Urbanek et al, 2019)
  - ▶ PERSUASION (Wang et al, 2019)
- ▶ Covering the range of anaphoric references included in the current Universal Anaphora specification
  - ▶ Identity anaphora
  - ▶ Bridging reference
  - ▶ Discourse deixis

Annotated according to the ARRAU3 Guidelines

- ▶ Creation:
  - ▶ First release created for 2021 Shared Task
  - ▶ Annotation guidelines thoroughly discussed and revised
  - ▶ Completely checked this year; new data annotated
- ▶ Available from <https://github.com/UniversalAnaphora>

# Dialogue datasets included in the CODI/CRAC 2022 corpus

# Switchboard

- ▶ <https://catalog ldc.upenn.edu/LDC97S62>
- ▶ One of the best known dialogue corpora
- ▶ 1,155 five-minute spontaneous telephone conversations between two participants not previously acquainted with each other. In these conversations, callers question receivers on provided topics, such as child care, recycling, and news media. 440 speakers participate in these 1,155 conversations, producing 221,616 utterances (Godfrey, Holliman and McDaniel 1992)
- ▶ Annotated for dialogue acts by Stolcke et al (1997)
  - ▶ <https://convokit.cornell.edu/documentation/switchboard.html>
- ▶ Annotated for information status by Nissim et al (2004)



# Switchboard in MMAX

Comment

min\_words way

min\_ids word\_168

Gender  unmarked  male  female  neuter  undersp-gen

Number  unmarked  plur  sing  mass  undersp-num  unsure-num

Person  per3  per1  per2

<-> Reference  unmarked  new  old  non\_referring  undef\_reference

Category  unmarked  person  animate  substance  medicine  concrete

<-> Ref\_type  phrase  segment

Segment\_antecedent utterances:7 utterances:6

<-> Related\_object  no  yes

Generic generic-no

<-> Ambiguity  unambiguous  ambiguous  ambiguous\_antecedent

gov\_verb phrased

adj\_arg na

an\_arg na

Apply Undo changes

Auto-apply is OFF

Levels Order KWIC view Font Auto-select

that way

he difference)

he difference)

.:ate]

here]

File Settings Display Tools Plugins Info Show ML Panel

B: Yeah,

actually, [I], [I] agree that [that]'s an interesting thing.

[I], [I] believe [that]'s [a, for, for [civil suits] [it]'s not unanimous, right,

but for, for [criminal suits] [it]'s unanimous by [federal law]].

A: [I] don't know if [it]'s even true that [it]'s always unanimous because [I] thought [there] were [[cases]where, uh, [I] don't know if [it]'s [the difference between [[felonies] and [misdemeanors]]], but [where[ [it] was okay for [a state] to have [it] like [[eleven] out of [[twelve]]]].

B: Yeah,

[I] actually thought [I] 'd read [that], too

and [that]'s why [I] phrased [it] [that way].

A: Yeah.

B: [I] wonder,

[the difference], yeah, [the difference] could be that for, uh, [state], if [there] is such [a thing], [[state felonies] versus] [federal felonies] that [[federal, [things that are under [federal jurisdiction]]]

# The AMI Meeting Corpus

- ▶ <https://groups.inf.ed.ac.uk/ami/corpus/>
- ▶ Another well-known dialogue corpus
- ▶ 100 hours of meeting recordings between several participants. The recordings include signals from close-talking and far-field microphones, individual and room-view video cameras, and output from a slide projector and an electronic whiteboard.
- ▶ Several types of annotation, including dialogue acts, topics, summaries, named entities, focus of attention

# The Light Corpus

- ▶ Amazon, Facebook, Google, are all creating dialogue corpora
- ▶ Light is a recently created corpus on Parl.ai:  
<https://parl.ai/projects/light/>
- ▶ The LIGHT project is a large-scale fantasy text adventure game research platform for training agents that can both talk and act, interacting either with other models or with humans.
- ▶ The Light corpus was entirely created through crowdsourcing at different levels

# The Light corpus

<b>Persona 1: Servant</b>	<b>Persona 2: King</b>
I come from the lower class. I do what I am told without question. I can not read. I have not seen my family in a long time.	I am a king of the whole empire. I give rules and pursuit them. I am brave and fearless.
Carrying: a duster, a small bucket, a rag Wearing: a shirt	Carrying: a crown, a scepter

**Setting:** The main foyer is massive. A grand staircase sits to the back of the foyer leading to the upstairs. At the front of the foyer stand two servants ready to help anyone who comes to visit. To the left of the room there is a doorway leading into a corridor. To the right there is a door leading to another corridor for the King's servants. At the foot of the stairs there is a bearskin rug that is staring at you almost as if still hungry. The walls are lined with portraits of the king and his family.

**Servant:** my humble king. What am I to do to serve you?

**King:** Ahhh. My loyal servant. Polish my scepter.

*Action:* give scepter to servant

**Servant:** Yes my lord. I will polish it immediately. Am I to return it to you personally?

*Action:* put scepter in small bucket

**King:** Yes. Yes. Of course. Also check the jewels in my crown. They seem loose.

*Action:* give crown to servant

**Servant:** But sire I am not qualified to do that. Would you prefer I take it to someone?

**King:** Oh fine then.

*Emote:* gesture sigh

**Servant:** I am sorry sir the rug startled me

*Action:* drop crown

**King:** Haha! That's bear I slain on my latest hunting trip. He's a mighty beast!

# The Light Corpus in MMAX

One-click annotation Panel Settings

turn utterance scene

Comment

min\_words our

min\_ids word\_8

Gender  unmarked  fem  masc  neut  undersp-gen

Number  unmarked  plur  sing  mass  undersp-num  unsure-num

Person  per3  unmarked  per1  per2

Gram\_fnc

< > Reference  unmarked  new  old  non\_referring  undef\_reference

Category  unmarked  person  animate  substance  medicine  concrete  on

Object

< > Related\_object  no  yes

Apply Undo changes

Auto-apply is OFF

MMAX2MarkableBrowser [1]

Levels Order KWIC view Font Auto-select

- acolyte
- god
- great Lord of Lightning
- Lightning
- our

File Settings Display Tools Plugins Info Show ML Panel

SETTING: temple hallway/Inside Temple

*The inner sanctum of the Acolytes of the temple. A table where offerings of food and gems for who they worship. Surrounded in paintings of mythical beasts on the walls. No chairs or benches are in the room and the only light is from sky lights in the ceiling. A place of worship for an ancient sect. They acolytes make offerings and do their ceremonies. No one outside the sect is even suppose to know about them or their temple.*

CHARACTER: [acolyte]

*I am an acolyte. I help with religious festivities. I help the Bishop.*

CHARACTER: [god]

*I am a god in the eyes of the villagers. I am regarded very highly as a god. People of the village pray to me for success in their crops as well as daily lives.*

----- End header -----

acolyte: Oh [great Lord of [Lightning]] !

Hear [our] prayers !

god: [I] hear [your] prayers] and will answer [them] .

[What] have [you] brought [me] ?

acolyte: [I]ll] be honest - [this] is [the first time [you] have appeared for real in like - forever] ?

So [there] s [some debate] in [the texts] so [we] decided to cover [all [our] bases] .

Navigation icons: back, forward, search, etc.

# The Persuasion for Good Corpus

- ▶ <https://convokit.cornell.edu/documentation/persuasionforgood.html>
- ▶ A collection of online conversations generated by Amazon Mechanical Turk workers, where one participant (the persuader) tries to convince the other (the persuadee) to donate to a charity.
- ▶ 1017 conversations, along with demographic data and responses to psychological surveys from users.
- ▶ Annotations: various speaker-level annotations (demographics, big five personality traits, etc)

# The annotation scheme used in the CODI/CRAC corpus

# The annotation scheme

- ▶ We used the ARRAU3 scheme, which covers the phenomena in the current version of the Universal Anaphora initiative
- ▶ Followed the ARRAU3 guidelines (see below)



# Universal Anaphora: Aims

- ▶ There are by now a number of substantial anaphorically annotated corpora, covering several languages, according to broadly comparable guidelines, but created without real coordination
- ▶ The objective of the Universal Anaphora initiative is to attempt to replicate the success of Universal Dependencies for anaphora

# Universal Anaphora: Progress so far

- ▶ Website:  
<http://universalanaphora.org>
- ▶ Proposal regarding coverage:  
[https://github.com/UniversalAnaphora/UniversalAnaphora/blob/main/Universal\\_Anaphora\\_1\\_0\\_\\_\\_Proposal\\_for\\_Discussion.pdf](https://github.com/UniversalAnaphora/UniversalAnaphora/blob/main/Universal_Anaphora_1_0___Proposal_for_Discussion.pdf)
- ▶ See also the very extensive CorefUD document,  
[https://github.com/UniversalAnaphora/UniversalAnaphora/blob/main/nedoluzhko\\_et\\_al\\_CorefUD20201\\_tr66.pdf](https://github.com/UniversalAnaphora/UniversalAnaphora/blob/main/nedoluzhko_et_al_CorefUD20201_tr66.pdf)
- ▶ Markup proposals
- ▶ UA Scorer:  
<https://github.com/juntaoy/universal-anaphora-scorer>

# Universal Anaphora: Anaphoric phenomena covered

## Core: IDENTITY ANAPHORIC REFERENCE VS PREDICATION

- (3) [Mary]<sub>i</sub> bought [a new dress]<sub>j</sub> but [it]<sub>j</sub> didn't fit [her]<sub>i</sub>.
- (4) [Mary]<sub>i</sub> is [a teacher].

## ALSO COVERING ZEROs, CLITICS

- (5) [IT] Giovanni è in ritardo, così mi ha chiesto se posso incontrarlo al cinema.  
*[EN] John is late so he asked me if I can meet him at the movies. (Poesio et al., 2016, ex. 9, p. 29)*

# Universal Anaphora: Anaphoric phenomena covered

## Beyond the core: SPLIT ANTECEDENT PLURALS

- (6) 'In THAT direction,' the Cat said, waving its right paw round, 'lives [a Hatter]<sub>i</sub>; and in THAT direction,' waving the other paw, 'lives [a March Hare]<sub>j</sub>. Visit either you like: [they]<sub>i,j</sub> 're both mad.' (From *Alice in Wonderland*, *Phrase Detectives* corpus.)

# Universal Anaphora: Anaphoric phenomena covered

## Beyond Identity: BRIDGING (ASSOCIATIVE) REFERENCE

- (7) There was not a moment to be lost: away went Alice like the wind, and was just in time to hear it say, as it turned a corner, 'Oh my ears and whiskers, how late it's getting!' She was close behind it when she turned the corner, but the Rabbit was no longer to be seen: she found herself in [a long, low hall, which was lit up by a row of lamps hanging from [the roof]]. There were doors all round the hall, but they were all locked; and when Alice had been all the way down one side and up the other, trying every door, she walked sadly down [the middle], wondering how she was ever to get out again.

# Universal Anaphora: Anaphoric phenomena covered

## Beyond Identity: EVENT ANAPHORA AND DISCOURSE DEIXIS

- (8) So she was considering in her own mind (as well as she could, for the hot day made her feel very sleepy and stupid), whether the pleasure of making a daisy-chain would be worth the trouble of getting up and picking the daisies, when suddenly a White Rabbit with pink eyes ran close by her. There was nothing so VERY remarkable in [that]; ...
- (9) There was nothing so VERY remarkable in that; nor did Alice think it so VERY much out of the way to hear the Rabbit say to itself, 'Oh dear! Oh dear! I shall be late!' (when she thought it over afterwards, it occurred to her that she ought to have wondered at [this], but at the time [it] all seemed quite natural); ...

# The Annotation

# The annotation process

- ▶ 2021 release annotated by the QMUL group according to the ARRAU3 guidelines
- ▶ A new version of the manual produced  
<https://sites.google.com/view/arrau/corpus>
- ▶ 2022 release:
  - ▶ Joint work between QMUL and CMU
  - ▶ extensive discussion of guidelines
  - ▶ reliability tests run
  - ▶ 2021 corpus completely checked
  - ▶ each document in the new test set doubly annotated



## Reliability results

Small reliability study:

- ▶ 3 documents from Persuasion already annotated last year 377 markables, 1241 words
- ▶ Compare two annotators from CMU with each others and with last year's annotation
- ▶ Using Krippendorff's  $\alpha$  for identity as suggested by Passonneau

	morphosyntax	identity	bridging	genericity
All	0.74	0.80	0.49	0.43
Taiqi-Katherine	0.78	0.81	0.44	0.59

## Issues: non-dialogue specific

- ▶ Structure of measure NPs / quasi-partitives
  - ▶ [4mg of [oestradiol]]: two markables
  - ▶ [a lot of good] one markable (*a lot of* determiner)
  - ▶ how about [tons of good]?
- ▶ Genericity

# Issues: dialogue specific

- ▶ Ambiguity
  - ▶ Ambiguity of first and second person pronouns
  - ▶ Ambiguity name / reference in AMI
- ▶ Role of referring expressions in grounding
- ▶ Associative DDs

# Ambiguity name/reference

The screenshot shows a chat window with a title bar containing the path: MMAX2 1.14.005 /Corpora/CODL\_CRAC21/Release\_1/AM/dev/ES2003a.mmax [modified]. The window has a menu bar with 'File', 'Settings', 'Display', 'Tools', 'Plugins', 'Info', and 'Show ML Panel' (checked). The chat content is as follows:

**PM:** Okay , well [I] think [were] ready to begin .

Right . [[my] names [Adam Duguid]] , [were] [here] because of [real reaction] ; um , [we] have in [the group]

**ME:** Oh , [Ebenezer Ademesoye] .

Would [you] like [me] to spell [that] ?

S

**PM:** Um , yeah , go for [it] [mate] .

**ME:** Um , N \_ E \_ Z \_

**PM:** N \_ E \_ Z \_ .

**ME:** E \_ R \_ .

**PM:** [Ebenezer] .

A yellow line is drawn from the word 'we' in the second PM message to the name 'Ebenezer Ademesoye' in the ME response, illustrating the ambiguity of the reference.

# Proposed solutions

- ▶ Ambiguity of first and second person pronouns: use generic interpretation
- ▶ Referring expressions used for grounding: mark as non-referring

## Some statistics

		Docs	Tokens	Markables	DO	Bridging	D. Deix
LIGHT	train	20	11495	3907	2132	381	72
	dev	21	11824	3941	2181	424	84
	test	38	22017	7330	3770	812	128
AMI	train	7	33741	8918	4579	853	230
	dev	3	18260	4870	2350	638	118
	test	3	16562	3990	2007	432	118
PERS.	train	21	9185	2743	1242	248	95
	dev	27	12198	3697	1715	316	133
	test	33	14719	4233	2111	304	105
SWITCH.	train	11	14992	4024	1679	589	128
	dev	22	35027	9392	3991	1165	265
	test	12	14605	3888	1606	464	107
<b>Total</b>		218	214625	60933	29363	6626	1583

# Discussion and Conclusions

## Related work

- ▶ A sister project: annotation in Situated Dialog, focusing in particular on reference to the visual situation (Loaiciga, Dobnik, Schlangen)
- ▶ the BOLT corpus of online discussions



# Discussion

- ▶ Work on annotating anaphora and reference in dialogue still limited
- ▶ CODI/CRAC Shared Task corpus an attempt to push forward the state of the art both in terms of annotation guidelines and of developing models
- ▶ Still a number of issues
  - ▶ Annotation guidelines need work
  - ▶ Tools still limited
  - ▶ Not many systems can handle e.g., discontinuous markables / discourse deixis

# Open Issues

- ▶ Associative DDs
- ▶ Sums

# Conclusions

- ▶ Objectives of this talk
  - ▶ Advertise that new dialogue corpora annotated for both anaphora reference are available
  - ▶ Mention some of the issues
- ▶ Current state of the art on anaphoric annotation designed primarily for written text, we need new annotation schemes, new tools, but also new computational models
- ▶ Some issues, particularly related to combining anaphora and reference (deixis)

## References

- David Bamman, Olivia Lewke, and Anya Mansoor. 2020. An annotated dataset of coreference in english literature. In *Proc. of LREC*. European Language Resources Association (ELRA), Association for Computational Linguistics (ACL).
- Donna Byron. 2002. Resolving pronominal references to abstract entities. In *Proc. of the ACL*, pages 80–87.
- Kevin Bretonnel Cohen, Arrick Lanfranchi, Miji Joo-young Choi, Michael Bada, William A. Baumgartner Jr., Natalya Panteleyeva, Karin Verspoor, Martha Palmer, and Lawrence E. Hunter. 2017. Coreference annotation and resolution in the Colorado Richly Annotated Full Text (CRAFT) corpus of biomedical journal articles. *BMC Bioinformatics*, 18(372).
- Abbas Ghaddar and Philippe Langlais. 2016. Wikicoref: An english coreference-annotated corpus of wikipedia articles. In *Proc. of LREC*, pages 136–142. European Language Resources Association (ELRA), Association for Computational Linguistics.
- Christoph Müller and Michael Strube. 2003a. A machine learning approach to pronoun resolution in spoken dialogue. In *Proc. of the 41st ACL*, pages 168–175.
- Christoph Müller and Michael Strube. 2003b. Multi-level annotation in MMAX. In *Proc. of the 4th SIGDIAL*, pages 198–207.
- Judith Muzerelle, Anaïs Lefevre, Emmanuel Schang, Jean-Yves Antoine, Aurore Pelletier, Denis Maurel, Iris Eshkol, and Jeanne Villaneau. 2014. Ancor \_centre, a large free spoken french coreference corpus. In *Proc. of LREC*.
- Massimo Poesio and Ron Artstein. 2008. Anaphoric annotation in the ARRAU corpus. In *Proc. of LREC*, Marrakesh.
- Massimo Poesio, Jon Chamberlain, Udo Kruschwitz, Silviu Paun, Alexandra Uma, and Juntao Yu. 2019. A crowdsourced corpus of multiple judgments and disagreement on anaphoric interpretation. In *Proc. of NAACL*, page 1778–1789, Minneapolis. 