Grounding and Clarification

September 15, 2015

Computational Pragmatics
Institute for Logic, Language, and Computation
University of Amsterdam

Contents

Transcription Exercise

Choice of Conversation

- Many of you chose scripted conversations. These are typically less chaotic, more grammatical, more structured than naturally occurring ones. (unless they are scripted to be chaotic!)

- Interviews are a good source for available natural conversations. But the setting is also quite special.

- In particular, many of you chose antagonistic situations, where some aspects of cooperative discourse do not show up.

Collaboration in Cooperative Dialogue

(1) B: really like a::hm (0.2) ⟩how do you call it⟨ (0.2)
     A: up[(right)
     B: [straight
     A: [up up
     B: [up up straight [yeah exactly

(from last year's course)

Intonation

- Remember that the stress underline is a binary feature.
- So if you annotate both strong and weak emphasis with them, it looks odd.
- Most longer words have a stress; unless it is particularly pronounced it does not need to be annotated.
- But it can also be used to great effect!

(2)   =Chi<u>ca</u>go! (0.3) It's called Chi<u>ca</u>go.

(3)   and making <u>something</u> of this <u>life</u>?

(4)   <u>eve</u>rybody come in <u>kil</u>lers, <u>cri</u>minals (0.3) <u>drug</u> dealers

Turn Grabbing

Interlocutors can forcefully take the turn.

> (5)   A: She's got MS and she's been 20 years with MS=
>   B: =Right, OK, I need you to listen carefully.

> (6)   A:partly because he was caught in a sexual [scandal]
>   B: [I'm gonna]
>   interrupt you right there

Turn Yielding

- ▶ If the next speaker is selected, they can take their time to speak.
- ▶ Or, if someone wants to yield the turn, they pause.

(7)   A: I think careers are twentieth century invention and
       I don't want one (1.5)
       You don't need to worry about me

(8)   A: How much? (0.8) How much money are they paying you?
      B: (1.0) I don't know what you're talking about.

(9)   A: Uhm one two three four (.) but spelled out?
      B: (1.5) That's not gonna help you.

Turn Holding

- Speakers can use fillers to signal that their turn isn't over.
- Or just keep talking and correcting on the way.

(10) A: Uuuh. Like The Whip? (0.5) Uuuh. The Nae Nae?

(11) A: <u>illegal</u> immigrants– I'm talking about illegal immigrants.
         I'm not talking about immi– I'm talking about <u>illegal</u>
         (0.5) immigr– I'm an immigrant. You're an immi–
         we're all immigrants. [I'm talking about <u>illegal</u>.
                                            I'm talking–]
      B: [Illegal immigrants commit less]

Transition Relevant Points

An important part is that TRPs are anticipated by the
interlocutor vying for the turn.

(12)	A: You're a very naive person.	[You're]
	B:	[The PEW research says]
	B: (0.2) that there arr uh-	

| (13) | A: No, it's not. | [No] |
| | B: | [Why would they] |

(14)	A: that doesn't make you feel much better now,	[does it?]
	B:	[Where's]
	Diondra?	

Transition Relevant Points

An important part is that TRPs are anticipated by the interlocutor vying for the turn.

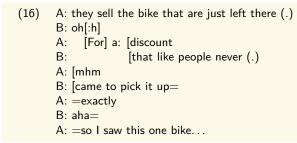(12)   A: You're a very naive person. | [You're]
        B: [The PEW research says]
        B: (0.2) that there arr uh-

(13)   A: No, it's not. | [No]
        B: [Why would they]

(14)   A: that doesn't make you feel much better now,| [does it?]
        B: [Where's]
           Diondra?

Clarification

Clarification Requests facilitate the grounding process. As they indicate a communicative problem, they need to be dealt with (almost) immediately.

```
(15)   A: That's science=
       B: =that's not [s-]
       A:              [sp]ort science
       B: What? S-
       A: I saw this on sport science=
       B: =You saw this on sport [science
       A:                        [I saw] this on sport science.
```

CR: Another Example

```
(16)   A: they sell the bike that are just left there (.)
       B: oh[:h]
       A:   [For] a: [discount
       B:              [that like people never (.)
       A: [mhm
       B: [came to pick it up=
       A: =exactly
       B: aha=
       A: =so I saw this one bike...
```

(from last year's course)

Clarification: Grounding

Sometimes it is important that grounding is secured.

> (17) A: see if she's breathing
>   B: hhh .hh (0.2) rea::lly, rea:lly struggling
>   A: ⟩Right⟨ is she breathing at all.
>   B: (0.4) Hardly (.) she's gone from sort of: from white
>   to all purply blue now
>   A: hh okay
>   B: She's slightly gasping.
>   A: She's slightly gasping?

Backchanneling

- The listening interlocutor gives feedback cues that signal continued attention, comphrehension, understanding and (possibly) agreement.
- They may be nonlinguistic (*e.g.*, nodding, gaze, eye-contact), but also linguistic. In the latter case these cues are called backchannels.
- Backchannels are not considered turns, but are subordinate to the (continued) turn of the speaking interlocutor.

> (18) A: to see if we can help [her] over the phone okay?
> B: [Right]
> A: I need you to get

Corpora / Resources
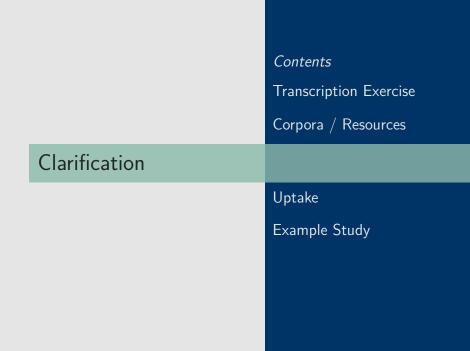
Corpora

- Many other people have made that effort on a much larger scale. . .

- The British National Corpus (BNC).
  - → Unconstrained conversation (people carrying around recorders in their daily lives).
  - → Mostly unannotated.
  - → Some subsets have been annotated (Clarification Requests by Matt Purver; Non-sentential Utterances by Raquel Fernández)

- The AMI Meeting Corpus.
  - → Four people roleplaying a design-team for a TV remote.
  - → Very rich annotation.

- The ICSI Meeting Corpus.
  - → Recordings of actual meetings.
  - → Also richly annotated.

More Corpora

- ▶ The Switchboard corpus.
  - → Unconstrained conversation between two strangers on telephone.
  - → Dialouge-act annotation.

- ▶ The Bielefeld corpus of Clarification Requests (German).
  - → An instructor telling someone how to assemble a paper airplane.
  - → Annotation of Clarification Requests.

- ▶ COCA (Corpus of Contemporary American English)
  - → Mostly transcripts of talkshows.

- ▶ RST Treebank.
  - → Newspaper articles annotated for Discourse Relations.

- ▶ The Settlers Corpus.
  - → Chatlogs of 4 people playing Settlers of Catan (online interface).
  - → Annotated for Discourse Relations.

Even More Corpora

- MICASE (Michigan Corpus of Academic Spoken English)

- Santa Barbara Corpus of Spoken American English.

- Transcripts of the US Supreme Court cases.

- Dumps of Wikipedia Discussion Pages.

- Proceedings of the UN or the EU (multiple languages!)

- Some smaller corpora in the NLTK distribution.

Contents

Clarification

Common Ground and Grounding

Common Ground

The common ground of (a group of at least 2) interlocutors is what they **jointly take for granted**. (Stalnaker 1978)

The process through which common ground is established is called **grounding**. (Clark 1996)

Common Ground and Grounding

Common Ground

The common ground of (a group of at least 2) interlocutors is what they **jointly take for granted**. (Stalnaker 1978)

The process through which common ground is established is called **grounding**. (Clark 1996)

Grounding (Clark 1996)

	Level	Joint Action
1	contact	A executes a behavior and B attends it
2	perception	A produces a signal and B perceives it
3	understanding	A conveys a meaning and B understands it
4	uptake	A proposes a project and B accepts/considers it

Shared Bases

The Shared Basis Model (Clark 1996)

A proposition p is common ground for members of community C iff there is a *shared basis* b for p, that is:

1. every member of C beliefs (individually) that b,

2. b indicates to every member of C that every member of C (individually) beliefs b,

3. b indicates to every member of C that p.

Shared Bases

(visual)
1. I see the table.
2. You see the table.
3. I see that 2.
4. You see that 1.

CG: *There is a table here.*

Shared Bases

(visual)
1. I see the table.
2. You see the table.
3. I see that 2.
4. You see that 1.

CG: *There is a table here.*

(logic)
A: p.
B: $\texttt{accept}(p)$.
Basis: $p \wedge \texttt{accept}(p)$.
CG: p.

Failure to Ground

- A and B might pay insufficient attention to each other;
- A might mumble or B might not hear A properly;
- A might speak in a complicated manner or B might not know all words in A's utterance;
- A might propose an infeasible project or B might fail to see the relevance of A's proposal.

Failure to Ground

- ▶ A and B might pay insufficient attention to each other;
- ▶ A might mumble or B might not hear A properly;
- ▶ A might speak in a complicated manner or B might not know all words in A's utterance;
- ▶ A might propose an infeasible project or B might fail to see the relevance of A's proposal.

↝ Evidence for failure on some such level are clarification requests, utterances where the speaker requests that the other party repeats or elaborates on some action.

Grounding and Clarification

Evidence for Failure

Level	Joint Action	Ex. Clarification
1 contact	A and B pay attention to another.	*Are you talking to me?*
2 percept.	A produces a signal; B perceives it.	*What did you say?*
3 underst.	A conveys a meaning; B recognizes it.	*What did you mean?*
4 uptake	A intends a project; B considers it.	*What do you want?*

We have seen that grounding can fail at each level and that clarification requests can evince that.

So if we are investigating grounding it seems appropriate to look for clarifications.

Classifying Clarification Requests: Levels

Classification by Level

Level	Type of Problem	Example
1 contact	channel	*huh?*
2 percept.	acoustic	*pardon?*
3 underst.	lexical	*What's a double torx?*
	parsing	*Did <u>you</u> have a telescope, or the man?*
	reference resolution:	
	– NP-reference	*Which square?*
	– Deictic-reference	*Where is 'there'?*
	– Action-reference	*What's to kowtow?*
4 uptake	recognizing and	*You want me to give you this?*
	evaluating intention	*Why?*

Rodríguez & Schlangen. *Form, Intonation and Function of Clarification Requests in German Task-Oriented Spoken Dialogues.* Proceedings of SemDial04 (Catalog).

Classifying Clarification Requests: Form

Classification by Form

I want to go to Paris.

Class	Description	Example
non	Non-Reprise	*What did you say?*
wot	Conventional	*Pardon?*
frg	Reprise Fragment	*Paris?*
slu	Reprise Sluice	*Where?*
lit	Literal Reprise	*You want to go to Paris?*
sub	Wh-Substituted Reprise	*You want to go where?*
gap	Gap	*You want to go to . . . ?*
fil	Gap Filler	*I want to go to .. – Paris?*
oth	Other	other.

Purver, Ginzburg & Healy. *On the Means for Clarifiation in Dialogue.* Proceedings of SIGdial01.

A Full Annotation Scheme

Form

distance	{1, 2, 3, 4, 5, more}
mood	{none, decl, polar-q, wh-q, alt-q, imp, other}
completeness	{none, particle, partial, complete}
rel-antec	{none, addition, repet, reformul, indep}
boundary-tone	{none, rising, falling, no-appl}

Function

source	{none, acous, lex, parsing, np-ref, deictic-ref, act-ref, int+eval, src-3, src-2+3, src-2+4, src-3+4, src-all}
extent	{none, yes, no}
severity	{none, cont-conf, cont-rep, no-react}
answer	{none, ans-repet, ans-y/n, ans-elab, ans-reformul, ans-w-defin, no-react}
happiness	{none, happy-yes, happy-no, happy-ambig}

Uptake

Austin

Austin on Uptake

"I cannot be said <u>to have warned</u> an audience unless it hears what I say and <u>takes</u> what I say in a certain sense. An <u>effect must be achieved</u> on the audience if the illocutionary act is to be carried out [. . .] So the performance of an illocutionary act involves the securing of *uptake*." (Austin 1962, p. 116f, underlining mine)

Austin

Austin on Uptake

"I cannot be said <u>to have warned</u> an audience unless it hears what I say and <u>takes</u> what I say in a certain sense. An <u>effect must be achieved</u> on the audience if the illocutionary act is to be carried out [. . .] So the performance of an illocutionary act involves the securing of *uptake*." (Austin 1962, p. 116f, underlining mine)

However, sometimes Uptake is refused:

Refusing Uptake

"my attempt to make a bet by saying 'I bet you sixpence' is abortive unless you say 'I take you on' or words to that effect; my attempt to marry by saying 'I will' is abortive if the woman says 'I will not'"' (Austin 1962, p. 37)

Austin. *How To Do Things With Words*. Harvard University Press, 1962.

Austin: Ratification and Cancellation

(constructed)		
	✓ I warned you	and you were more careful.
	✓ I warned you	and you didn't care.
	✓ I asked you	and you answered.
	✓ I asked you	and you didn't answer.
	✓ I bet you	and you accepted.
	? I bet you	and you refused.
	✓ I married you	and you married me.
	✗ I married you	and you refused.
	? I declared war to you	and you refused.

Austin: Ratification and Cancellation

(constructed)		
	✓ I warned you	and you were more careful.
	✓ I warned you	and you didn't care.
	✓ I asked you	and you answered.
	✓ I asked you	and you didn't answer.
	✓ I bet you	and you accepted.
	? I bet you	and you refused.
	✓ I married you	and you married me.
	✗ I married you	and you refused.
	? I declared war to you	and you refused.

It seems that there are two dimensions to uptake:

▶ Recognition: The audience taking the illocutionary act as such.
→ intention recognition ("weak" uptake).

▶ Transfer: The audience adopting/ratifying the illocution.
→ intention transfer ("strong" uptake).

Clark

Clark on Uptake

"3: B is <u>recognizing</u> A's <u>request</u>
4: B is <u>considering</u> A's <u>proposal</u>" (Clark 1996, p. 152, ul mine)

"When Jane produces 'Who is it?' she <u>means</u> (at level 3) that Kate is to say who she is and, thereby, <u>proposes</u> (at level 4) that Kate tell her who she is" (ibid., p. 199, ul mine)

Clark

Clark on Uptake

"3: B is <u>recognizing</u> A's <u>request</u>
4: B is <u>considering</u> A's <u>proposal</u>" (Clark 1996, p. 152, ul mine)

"When Jane produces 'Who is it?' she <u>means</u> (at level 3) that Kate is to say who she is and, thereby, <u>proposes</u> (at level 4) that Kate tell her who she is" (ibid., p. 199, ul mine)

But, again, Uptake can be refused:

Refusing Uptake

"when respondents are unwilling or unable to comply with the project as proposed, they can <u>decline to take it up</u>" (ibid., p. 204, ul mine)

"such joint projects [questions] become <u>complete only through up-take</u>, so completion requires [. . .] [an] answer." (ibid., p. 198, ul mine)

Clark. *Using Language*. Cambridge University Press, 1996.

Clark: Joint Construals

Principle of Joint Construal (Clark 1996)

For each signal, the speaker and addressees try to create a joint construal of what the speaker *is to be taken to mean* by it.
(Clark 1996, p. 212, emphasis mine)

NOT: What the speaker means!

Construing

"By this principle, [...] she is trying to create a construction that the two of them are willing to accept as what he meant by it." (Clark 1996, p. 212, emphasis mine)

Clark: Construing in Uptake

Construals

A: Sit down.
B: Yes, sir.
  ⤳ Order.
B: Thanks!
  ⤳ Offer.
B: Good idea!
  ⤳ Advice.

A: Sit down.
B: I'm not doing what you tell me!
  ⤳ Order.
B: No, thanks!
  ⤳ Offer.
B: I think I'd rather stand.
  ⤳ Advice.

Clark: Construing in Uptake

Construals

A: Sit down.	A: Sit down.
B: Yes, sir.	B: I'm not doing what you tell me!
$\leadsto$ Order.	$\leadsto$ Order.
B: Thanks!	B: No, thanks!
$\leadsto$ Offer.	$\leadsto$ Offer.
B: Good idea!	B: I think I'd rather stand.
$\leadsto$ Advice.	$\leadsto$ Advice.

Apparently, construals

. . . require (semantic) understanding.

. . . are prior to acceptance / refusal.

Illocutionary Force and Effect

For an illocutionary act, we separate the recognition of its force from the achieving of its effect.

Weak and Strong Uptake

- A speech act is **weakly taken up** if the hearer has recognized the illocutionary force.
- A speech act is **strongly taken up** if the illocutionary effect on the hearer has been achieved.

Illocutionary Force and Effect

For an illocutionary act, we separate the recognition of its force from the achieving of its effect.

Weak and Strong Uptake

- A speech act is **weakly taken up** if the hearer has recognized the illocutionary force.
- A speech act is **strongly taken up** if the illocutionary effect on the hearer has been achieved.

- Even if a bet has not been established, to even make the refusal the hearer needed to recognize the betting force.
- And then it is grounded between the interlocutors that the attempt to bet has taken place.

Disentangling Uptake: Grounding

Refined Uptake Level

	Level	Joint Action
1	contact	A and B pay attention to another.
2	perception	A produces a signal; B perceives it.
3	understanding	A conveys a meaning; B recognizes it.
4.1	int. recognition	A intends a project; B understands it.
4.2	int. transfer	A proposes a project; B accepts it.

Disentangling Uptake: Grounding

Refined Uptake Level

	Level	Joint Action
1	contact	A and B pay attention to another.
2	perception	A produces a signal; B perceives it.
3	understanding	A conveys a meaning; B recognizes it.
4.1	int. recognition	A intends a project; B understands it.
4.2	int. transfer	A proposes a project; B accepts it.

- ▶ Level 3 is semantics only.
- ▶ Intention Recognition / Construal is a separate step.
  - → Misconstruals are a new source of failure.
  - → Rhetorical questions have assertive force.
- ▶ Full Grounding is achieved upon acceptance of the project.
  - → Ratifying a bet.
  - → Adopting a belief.
  - → Answering a Question.

New Sources of Errors

The question at hand is: What might preclude strong uptake?
How and why might it fail?

New Sources of Errors

The question at hand is: What might preclude strong uptake?
How and why might it fail?

The intuitive answer is that speakers do not accept something if
they can't or don't want to.

New Sources of Errors

The question at hand is: What might preclude strong uptake? How and why might it fail?

The intuitive answer is that speakers do not accept something if they can't or don't want to.

> **Conditions for Joint Purposes (Clark 1996)**
>
> A and B can adopt a **joint purpose** p if:
>
> **Identification** A and B know about p;
>
> **Ability** A and B are able to do the participatory actions in p;
>
> **Willingness** A and B must be willing to engage in p;
>
> **Mutual Belief** The previous three are common ground for A and B.

Evidence

To make this theory more plausible, we are looking for evidence.
We have introduced new sources for errors, so we are searching
for clarification requests.

* We need to separate level 3 from level 4.1.
▶ Show that level 4.1 can fail.
▶ Investigate where level 4.2 fails
* Separate level 4.2 CRs from content questions.

Evidence

To make this theory more plausible, we are looking for evidence. We have introduced new sources for errors, so we are searching for clarification requests.

- * We need to separate level 3 from level 4.1.
- ► Show that level 4.1 can fail.
- ► Investigate where level 4.2 fails
- * Separate level 4.2 CRs from content questions.

Evidence might look like this:

(con.)   A: Can you get the butter?
   a. B: Salted or unsalted? ⤳ failure on level 3.
   b. B: Should I bring it to you? ⤳ failure on level 4.1.
   c. B: Why would I? ⤳ failure on level 4.2.
   d. B: Sure. *[fetches butter]* ⤳ full grounding.

Examples

(19)   A: And we're going to discuss [. . .] who's gonna do
          what and just clarify
       B: **Are you asking me whether I wanna be in there?**
       A: I was just mentioning it to you in case you wanted to
       B: Don't wanna.
    ⤳ *failure of weak uptake*

(20)   A: I know Vic has cream in his [food] and
       B: **How do you know?**
       A: Well it said so on the menu, that's why.
    ⤳ *failure of strong uptake (proposer: ability / knowledge)*

(21)   A: Daddy can we swop places now?
       B: **Why?**
       A: Cos I wanna sit next to you and Lee.
    ⤳ *failure of strong uptake (proposer: willingness / reason)*

Examples from the British National Corpus.

Example Study

Annotation Study

- ▶ We extract candidates for uptake-level CRs from the AMI corpus.

- ▶ The AMI has dialogue act annotation.
- ▶ The AMI has adjacency pair annotation.

- ▶ Intuitions:
  - → CRs are proper questions (*i.e.*, demand an answer).
  - → CRs are backward-looking (*i.e.*, have a source utterance).
  - → The person asked usually responds.

JJ Schlöder and R Fernández. *Clarifying Intentions in Dialogue: A Corpus Study.* International Conference on Computational Semantics (IWCS) 2015, London, 2015.

Annotation Study

Heuristic for CR retrieval

Take all utterances Q where:

a. Q is turn-initial and annotated as an 'Elicit-' type of dialogue act, spoken by a speaker B.

b. Q is the second part of an adjacency pair; the first part (the *source*) is spoken by another speaker A.

c. Q is the first part of another adjacency pair; the second part (the *answer*) is spoken by A as well.

▶ This heuristic retrieves 338 utterances from the AMI Meeting Corpus, 248 of which we annotated as actually being CRs (73%).

▶ In addition, we find 195 additional CRs if we change criterion (a.) to also retrieve 'Comments about understanding (und).'

Some Examples

(22) A: I think that's all.
      B: **Meeting's over?**
         int-rec

(23) A: Just uh do that quickly.
      B: **How do you do it?**
         int-ad

(24) A: I'd say two.
      B: **Why?**
         int-ad

Annotation Scheme

- **not CR.** The utterance is definitely not a CR.
- **low CR.** The utterance indicates a problem with *semantic / propositional content*.
  - Word meaning.
  - Acoustic channel.
  - Reference resolution.
- **intention recognition CR.** The CR speaker *has not fully understood (or is trying to guess)* the previous speaker's goal/intention.
  - Prototypical case: Speech act determination.
- **intention adoption CR.** The CR speaker *does not yet accept the previous speaker's intention*.
  - The speker wants/needs more information.
  - The speaker has incompatible beliefs.
- **ambiguous.** If none of the above apply with some certainty.

Annotation Results

Inter-Annotator Agreement

- 5-way task: $\kappa = 0.76$.
- CR *vs* not-CR (boolean): $\kappa = 0.85$.

Results and Comparison

Category	Count	incl. 'und'	RS 2004
not CR	90 (27%)	-	-
ambiguous	40 (12%)	40 (9%)	14.3%
low-level	78 (23%)	273 (62%)	63%
intent. recognition	53 (16%)	53 (12%)	22.2%
intent. adoption	77 (23%)	77 (17%)	
Total	338 (100%)	443 (100%)	

Rodríguez & Schlangen. *Form, Intonation and Function of Clarification Requests in German Task-Oriented Spoken Dialogues.* Proceedings of SemDial04 (Catalog).

Coverage

- It is also interesting how well a heuristic retrieves its targets.

Coverage

Our heuristic retrieves a CR for approximately 1.1% of all turns in the AMI Corpus. We estimate that it misses about 50% of actual CRs.

- Previous studies have generally indicated higher numbers: 4% (Purver) – 5.8% (Rodríguez & Schlangen).

Conclusion

- With such a sub-corpus and a small-scale annotation study, Raquel and I could show that:
  - → The difference between intention recognition and intention adoption can be told apart by appropriately instructed humans.
  - → The difference is evinced by clarification questions.

- One can derive specialised sub-corpora from general purpose corpora.

- Rich annotation can be very helpful for this!

- For your own projects: Be creative and ask me for resources.