

Introduction

To annotate the MaGE corpus, we used the presentation-acceptance model that is defined in Clark's *Using Language*, Chapter 8¹. In this model, there is a presenter who offers information, and an acceptor, who accepts or understands it. Presenters and acceptors each have a set of grounding markers that is appropriate for their roles. Thus, presenter turns are annotated with presenter grounding markers while acceptor roles are annotated with acceptor grounding markers. These markers are defined in the Grounding Markers section below.

A conversation between two speakers consists of several projects. In our model, a project is a series of turns that begins when a speaker introduces a topic, and ends when both speakers have achieved common ground, marked by the 'closure' grounding marker, on the topic.

Grounding Markers

We use the following grounding markers for annotation:

1. **Presentation:** A signal or piece of information. Presentations are always offered by a presenter. Once someone is in a presenter role, anything they say is a presentation, unless it is clearly only a question (probe) or a backchannel. In Clark's model, the definition of a presentation is similarly broad. In that model, each acceptance by the acceptor is also itself a presentation, resulting in a nested structure of presentations and acceptances. In contrast, our model follows a flat structure, and does not treat acceptances as presentations themselves.

In our study, we are interested in studying turns that shifted the floor towards a certain speaker (i.e., turns where they took the presenter role from the other speaker). While annotating turns, it is sometimes tricky to tell when the floor has shifted, and if a certain turn should be labelled with *[acceptor grounding marker]* or *[acceptor grounding marker, presentation]*, where the second label is a floor shift. In general, if the turn serves more than the purpose of the first *[acceptor grounding marker]* and offers additional information that is carried on past this turn, in a way that the acceptor has now taken on initiative to present, then this turn can be considered a floor shift and should be labelled a presentation.

2. **Backchannel:** Short acknowledgements. Usually < 1 second, such as "Mh-mmm", "Okay", etc. In our annotation, both presenters and acceptors can backchannel. Although the corpus is annotated with presenter backchannels, we did not count presenter backchannels in our final study.
3. **Probe:** A question, or a signal offered without certainty of acknowledgement (understanding) from the other speaker. The second type of signal could be a reference to an object or a concept, such as, "You know that assignment...?", where the speaker is

referring to the assignment without certainty that the other speaker knows what they are referring to.

Probes are labelled with an additional grounding marker if one part of the turn is a question and the other part is some additional information. For example,

Learner: "So, I didn't really use that but it's just there. Should I just delete it? I don't know."
(Probe, Presentation)

In Clark's model, probes do not fall into any of the four main classes of positive evidence (assertions, presuppositions, displays or exemplifications of understanding), but are treated as an entirely different class of actions (*Communication Probes*, Clark pg. 234). For this reason, we have used probes as a signal for presenters and acceptors in our model. Clark's communicative probes are further broken down into four categories. We do not distinguish between types of probes, and instead treat all questions and those signals made without certainty of succeeding as probes.

- 4. Uptake:** The acceptor's next relevant turn. These are usually used when acceptors are offering information without making it a presentation (i.e., they do not take enough initiative to become the presenter). For example,

Teacher: "So once you drag off of it, then this never executes this, uh, as- assignment to false never happens..."
(Presentation)
Learner: "I don't get it. Sorry."
(Uptake)

If the acceptor uses an uptake to take initiative and continue offering information, then the 'uptake' turn could also be labelled as a presentation. For example,

Learner: "Yeah, it was due yesterday."
(Answer)
Teacher: "Then you've already done it"
(Presentation)
Learner: "Yeah. Yeah, it was nice that we could adapt most of our code."
(Uptake, Presentation)
Teacher: "Mm-hmm."
(Backchannel)
Learner: "Like it wasn't completely new, or the binary search part was new, but the other part."
(Presentation)

In Clark's model, uptakes or presuppositions of understanding, one of the four main classes of positive evidence. We use uptakes as they are defined in this model. (Clark pg. 228)

- 5. Answer:** The response to a presenter's probe. Sometimes answers will span many turns, with presenter turns in between. For example,

Teacher: "How was spring break?" *(Probe)*

Learner: "Spring break was good. I went to New York with a friend from California. She flew and and yeah, but it was so cold. There was a blizzard." *(Answer)*

Teacher: "Oh yeah, I heard." *(Backchannel)*

Learner: "So bad, but whatever, company was good, so-" *(Answer)*

Longer answers are sometimes indications that the acceptor (in this example, the learner) is trying to take the floor. If the acceptor continues answering for more turns, or if there is a shift in initiative because of the acceptor offering information for multiple turns, this could be a floor shift where the acceptor becomes the new presenter. In the example below, the learner is initially the presenter. The teacher takes the presenter role through their second answer.

Learner: "... that has border layout?" *(Probe)*

Teacher: "You can use...like you can try all different, but border layout would definitely work." *(Answer)*

Learner: "Okay. Do you think that'd be best or do you think a different layout like..." *(Presentation, Probe)*

Teacher: "I think maybe grid layouts work the best because your focus is on the Tetris board - " *(Answer, Presentation)*

Learner: "Mh-mmm." *(Backchannel)*

Learner: "... but you would be placed in the center." *(Presentation)*

In Clark's model, answers are displays of understanding. We use answers in the same way as in Clark's model.

- 6. Repetition:** Acceptors may repeat all or part of the presenter's last turn to signal understanding. For example,

Learner: "...I write some comments on the same line, not before the, before the coding or after the coding." *(Presentation)*

Teacher: "Mh-mmm, like this" *(Backchannel)*

Learner: "Yeah, like this" *(Repetition)*

- 7. Paraphrase:** Acceptors may paraphrase the presenter's last turn to signal understanding. For example,

Teacher: "Um, and then *lastpoint* is just a global variable." *(Presentation)*

Learner: "Mh-mmm." *(Backchannel)*

Teacher: "So it's accessible out here or in there." *(Presentation)*

Learner: "Okay. So it's a global variable that's been in the computer forever. Or no, like -" *(Paraphrase)*

Repetitions are paraphrases are both exemplifications of understanding in Clark’s model and are used in the same way in our model. (Clark pg. 228)

8. **Closure:** Closure is achieved when the presenter and acceptor acknowledge that they have both come to a mutual understanding about the topic they were understanding. This may span over a few turns. For example,

Teacher: “ I can’t get into it, but I think they do have solutions for you” *(Presentation)*
 Learner: “Okay, great.” *(Closure)*
 Teacher: “Yeah.” *(Backchannel, Closure)*

In our annotation, both speakers must achieve closure for the project to be consider closed. Sometimes, one speaker will try to close the project, but find that the other speaker continues speaking about the topic. For example,

Learner: “Thank you.” *(Closure)*
 Teacher: “Yeah, of course. Um, well if you don’t have anymore questions...” *(Presentation)*
 Learner: “Not really.” *(Uptake)*
 Teacher: “I don’t have anything for you.” *(Closure)*
 Learner: “Yeah, I’ll see you next week.” *(Closure)*
 In this case, the project continues until both speakers achieve closure on this part of the project.

Grounding Markers for Specific Roles

Some grounding markers are specific to presenters or acceptors, while others can be used by both. The table below shows which speakers can use which grounding markers.

Presenter	Acceptor
Presentation Backchannel Probe Closure	Uptake Backchannel Probe Repetition Paraphrase Closure

Labelling Conventions

We used the following shorthand notation to label grounding markers.

Grounding Marker	Notation
Presentation	pres
Uptake	uptk
Backchannel	bc
Probe	prob
Repetition	rpt
Paraphrase	prph
Closure	clos

In addition, the label indicates the speaker of the turn in its prefix. We used the prefix ‘*t/’ if the speaker was a teacher, and ‘*l/’ if the speaker was a learner.

Below are two labelled turns.

Teacher: “Um, and then *lastpoint* is just a global variable.” *t/pres
Learner: “Mh-mmm.” *l/bc

If a turn has multiple labels, the labels are written next to each other. For example:

Learner: “Okay. Do you think that'd be best or do you think a different layout like...” *l/pres *l/prob

Other Notes

1. At the beginning of conversations, and in between presentations, there is sometimes no clear presenter. For example, many conversations begin this way:

Teacher: “How are you doing?”
Learner: “Good, how are you?”
Teacher: “I’m good.”

This will last for a few turns before one of the speakers takes initiative to begin a presentation. In this case, we label both speakers’ turns with acceptor grounding markers since neither speaker is making a presentation. We labelled these chunks as “buffer” chunks in some CSV documents. However, we did not use buffers for any analysis in the study.

2. Some of the CSV documents have additional information that was automatically generated but not used in the study.

Speaker	Start Time	End Time	Turn	Grounding Marker	Presenter	Project No
Learner	3	6	I almost finished the text version?	*l/ans *l/pres	learner	1
Teacher	6	10	Nice that's awesome.	*t/uptk	learner	1

In the above example, the learner is the presenter, as indicated by the “*l/pres’ tag in the first turn. The ‘Presenter’ column, which is automatically generated, indicates the presenter by checking for ‘pres’ tags in the grounding markers and looking for the speaker of the turn. If a ‘pres’ marker belonging to another speaker is found in a subsequent turn (i.e, the floor has shifted to this new speaker), the ‘Presenter’ column is updated. This update is shown in the example below. When both speakers reach closure, and there is no clear presenter immediately following the closure, the ‘Presenter’ column will say ‘buffer’ to indicate that this is a buffer chunk.

Learner	176	181	Yeah.	*l/bc	learner	1
Teacher	181	185	Um, like you don't have to stick to their design...	*t/ans*t/pres	teacher	1

The ‘Project No.’ column is also automatically generated and increments in value every time a project comes to a close (i.e, ‘clos’ tags are found for both speakers). However, the generating the ‘Project No.’ was a little confusing when also considering buffer chunks. Buffer chunks are currently indicated by ‘-1’ in the ‘Project No.’ column, but if you choose to use this ‘Project No.’ column or buffer chunks in subsequent analysis, it may be useful to revise this code.

Related Work

Clark and Schaefer’s Contribution model² consists of two phases: a presentation phase and an acceptance phase. In the presentation phase, **A** presents an utterance **u** for to **B** and looks for evidence **e** from **B** that **B** has understood **u**. In the acceptance phase, **B** accepts **u** by offering a signal **e’**. Once this pair of signals is complete, **A** and **B** have achieved common ground. The model allows signals to be offered in a recursive way, as embedded contributions. For example, **B** can reject or repair **A**’s utterance **u** in the acceptance phase. Then, **B**’s turn itself becomes a presentation that needs to be accepted, and this structure of embedded contributions can continue infinitely.

The Grounding Acts model^{3,4} is a reformulation of the contribution model which treats sets of utterances, rather than presentation-acceptance pairs, as grounding units. Each utterance corresponds to one or more grounding act, similar to the grounding markers used in our model. The different types of acceptance are collapsed into one grounding act. Unlike the Contribution model, the Grounding Acts model does not trigger recursive or embedded grounding units to address previous grounding acts.

Our model uses elements from both these models. It follows the presentation-acceptance structure and types of evidence from the Contribution model, but does not allow recursive pairs.

In other words, it follows the 'flat' structure of the Grounding Acts model. In our model, the discourse unit is a project, which may span between a few turns to the entire conversation.

The Contribution model also orders types of evidence in order of their strength. The Grounding Acts model does not, since there is only one type of evidence. Although our model uses the same classes of evidence as the Contribution model, we do not explicitly use the ordering of evidence in any analysis.

References

¹ Herbert H. Clark. 1996. *Using Language*. Cambridge University Press.

² Herbert H Clark and Edward F Schaefer. 1989. Contributing to discourse. *Cognitive Science*, 13(2):259–294.

³ David R Traum. 1999. Computational models of grounding in collaborative systems. In *AAAI Fall Symposium on Psychological Models of Communication in Collaborative Systems*, pages 124–131.

⁴ David Traum. 1994. *A Computational Theory of Grounding in Natural Language Conversation*. Ph.D. thesis, Computer Science Dept., University of Rochester.