Learning through Discussion: Foundations, Findings, and Future

Carolyn Penstein Rosé
Language Technologies Institute
Human-Computer Interaction Institute
School of Computer Science
Carnegie Mellon

With funding from the National Science Foundation and the Office of Naval Research
Slides with full citations for all referenced work will be available from my website!

Feel free to contact me to discuss further or collaborate!!

Carolyn Penstein Rosé

http://www.cs.cmu.edu/~cprose

cprose@cs.cmu.edu
Developing **technology** capable of shaping conversation and **supporting effective participation in conversation** to achieve positive impact on...
Developing **technology** capable of shaping conversation and **supporting effective participation in conversation** to achieve positive impact on...

**Human learning**
Developing **technology** capable of shaping conversation and **supporting effective participation in conversation** to achieve positive impact on...

Human learning

Health
Developing **technology** capable of shaping conversation and **supporting effective participation in conversation** to achieve positive impact on...

**Human learning**

**Health**

**Wellbeing**
Human learning
Human learning

- Understand what properties of discussions have a positive effect on learning
- Computationally model those selected properties of conversation
  - Makes research more scalable/efficient
  - Enables real time monitoring and dynamic triggering of supportive interventions
- Deployment and evaluation of supportive interventions at scale
Little known fact about me...
Overview

- Foundations
  - Applied machine learning
  - Learning Sciences
- Findings
  - Pitfalls to Avoid
  - Patterns to Leverage
- Future
  - Working together!!


*The International Journal of Computer Supported Collaborative Learning*, Springer.

Let's think about the motivation for Reheat. What process does the steam undergo in the Turbines?

Student: condensation

Steam undergoes Isentropic Expansion in a turbines that converts the heat energy of the steam into work by rotating the blades. What happens to the Quality of steam during Expansion?

Student: I’m not sure

As steam expands in a turbine, its moisture level increases. Does this affect the extent to which the Quality is acceptable for use?

Student: wtf!!!

We should NOT use low quality steam in turbines as it condenses on the turbine blades and damages the turbine running at high speeds.

Does this constrain the amount of work steam can do in turbines?

Student: yes

Good. Since quality of steam falls gradually in the turbines, we design turbines such that the quality remains above acceptable levels [0.85]. ...
** Students learn up to 1.25 standard deviations more when interactive support is provided in the environment. (more than a full letter grade!)
Foundations
Machine Learning

Download LightSIDE Now
Latest Update 5.10.2013

We're happy that you're interested in using LightSIDE as part of your research! Here's what the freely available and completely GPLv3 open source version of LightSIDE does for you:

Easy, Fast Feature Extraction
What do we learn from this?

- It’s easy to find low level, simple features that give us leverage for making a prediction.
- Frequently, they correlate with the real factors of interest.
- But their effect is very specific to the data that we have trained the model on.
- The models will not generalize well!!
What do we learn from this?

- It’s easy to find low level, simple features that give us leverage for making a prediction.
- Frequently, they correlate with the real factors of interest.
- But their effect is very specific to the data that we have trained the model on.
- The models will not generalize well!!

- **We find the same thing in a learning context!!!!**
Low level features we used to believe in

- General indicators of interactivity
- Turn length
- Conversation length
- Number of student questions
- Student to tutor word ratio
- Student initiative


So, what do we know about how people learn?
Learning about Learning by Listening to Learners

- Think aloud protocols, verbal protocol analysis (Ericsson & Simon, 1993; van Someren et al., 1994; Chi, 1997)
  - Understanding problem solving processes
  - Assessing mental models

- Self-explanation studies (Chi et al., 1994)
  - Understanding sense making and learning processes of individuals

- Tutorial Dialogue (Graesser et al., 1998; Evens & Michael, 2003; Rosé et al., 2001b)
  - Understanding learning from human tutoring

- Collaborative Process Analysis (Berkowitz & Gibbs, 1983; Stahl, 2006; de Wever et al., 2006; van der Pol et al., 2006; Weinberger & Fischer, 2006; Rosé et al., 2008)
  - Understanding collaborative learning processes

http://www.slideshare.net/ALHalverson/the-think-aloud-ta-method
What do we mean by learning?

- **Cognitive**
  - Individual is the unit of analysis
  - The goal is for an individual to acquire concepts and skills
  - Discussion provides stimulus for the individual’s learning
  - The focus of the work is to understand the psychological processes located within the individual’s head

- **Cognitive Apprentice**
  - Individual is still the unit of analysis
  - The goal is still for an individual to acquire concepts and skills
  - Learning occurs through interaction
    - That could be phrased as “learning is inherently social”
  - Discussion is one medium on interaction, especially when the interaction is between humans
  - The focus of the work is to understand how scaffolding influences cognitive processes, especially acquisition of concepts and skills

- **Sociocultural**
  - The individual within a community is the focus
  - The goal is for a student to learn to participate in the community, relationships are an important outcome
  - Learning still occurs within interaction
  - The focus of the work is to understand the paths of participation that lead from peripheral participation to core participation within the community

- **Communities of Practice**
  - The community itself is the unit of analysis (or an organization, or a small group)
  - The goal is for the group to learn to function well together
  - Learning still occurs within interaction, but the interaction now is an end, not just a means
  - The focus is on understanding the processes that exemplify optimal group interactions
A Spectrum of Complexity

Individuals  Pairs  Small Groups  Organizations  Communities  Nations
Causality

*Hypothesis:* If I manipulate variable A, it will have effect B on variable C.

How do we test this?
A Spectrum of Research Methodologies

Simplification/Abstraction

Preserving Complexity

Quantitative

Individuals  Pairs  Small Groups  Organizations  Communities  Nations

Qualitative
How can we use measures derived from process analysis?

- Assuming we are approaching our analysis quantitatively
  - Test for mediation vs moderation
    - Process variable correlates with dependent variable
    - Experimental manipulation has an effect on a process variable
    - Experimental manipulation has an effect on a dependent variable
  - If we add the process variable in as a covariate, and the effect of condition becomes non-significant, then the process variable mediates the effect of condition
  - Otherwise it is just a moderator
What if we don’t agree on what learning is?

- Our definition of learning affects:
  - How we measure learning
  - What role discussion plays in learning
  - What about discussion is relevant for learning
  - What kinds of manipulations we are interested in
  - Whether we can even study learning experimentally
  - What we’ll decide is conducive to learning
The think aloud method

FIGURE 7.1: The analysis process

From van Someren et al., 1994
Caution!!!

Cognitive and Social Processes are Intertwined
Findings
Purpose of Interaction in Learning Contexts

- Reward structures encourage students to focus on performance over learning.
- Well crafted instruction provides opportunities for learning.
- Opportunities only help if students take them.

Take Home Message:
Introducing reflection points provides opportunities for students to take advantage of learning resources.
Important Ingredients for Learning

- Carefully Structured Conceptual Knowledge
- Reflection through Rich Interaction
Souflé Framework
(Howley et al., 2013)

3 Dimensions:
- Transactivity
- Engagement
- Authoritativeness
Souflé Framework
(Howley et al., 2013)
• Definition of Transactivity
  • building on an idea expressed earlier in a conversation
  • using a reasoning statement

I think the tube will get heavier because water is going in.

That’s true, but the important point is that water can flow in, but starch can’t flow out.
Transactivity (Berkowitz & Gibbs, 1983)

- **Findings**
  - Moderating effect on learning (Joshi & Rosé, 2007; Russell, 2005; Kruger & Tomasello, 1986; Teasley, 1995)
  - Moderating effect on knowledge sharing in working groups (Gweon et al., 2011)

- **Computational Work**
  - Can be automatically detected in:
    - Threaded group discussions (Kappa .69) (Rosé et al., 2008)
    - Transcribed classroom discussions (Kappa .69) (Ai et al., 2010)
    - Speech from dyadic discussions (R = .37) (Gweon et al., 2012)
      - Predictable from a measure of speech style accommodation computed by an unsupervised Dynamic Bayesian Network (Jain et al., 2012)
### Accountable Talk  
(Resnick, Michaels & O’Connors, 2010)

<table>
<thead>
<tr>
<th>Utterance</th>
</tr>
</thead>
</table>
| **1**  
Teacher  
OK, does anyone want to respond to that? Who wants to respond and can prove that, kind of respond with their own ideas or can add another idea to it? Frank, go ahead. |
| **2**  
Frank  
I um I agree with what you said because this for example like if you put-- if you had big um, can- like if you got a big cup of water and you put a- an eraser in there, like a- a ah, like the eraser over there, if you put something like that in a big cup of water, the water level would rise a lot, and, if you put in a copper cube, and it’s not even gonna- it’s not going to rise that much even though that copper cube will weigh more than a eraser. |
| **3**  
Teacher  
Ahh ok, anyone agree with Frank's idea? I like that he kind of, talked about another object and he chose an object that we know has a different volume than the copper cube. He also chose an example of an object that has a much greater volume. And can someone explain or repeat for us what Frank thought would happen if I put an object this big, in water? James, what he think would happen if I put an object this big, into a cup of water? |
| **4**  
James  
He said that if you-- if you like a big- bur--like a big bottle of like water |
| **5**  
Teacher  
Uh huh |
| **6**  
James  
and you put the eraser in it, then it would probably like, rise a lot. Then-- |
| **7**  
Teacher  
Woah, woah, rephrase that – it has a what?  
Press for Reasoning |
| **8**  
James  
a different vol-- a different volume... |
| **9**  
Teacher  
And which has a greater volume?  
Say More |
| **10**  
James  
The eraser. |
| **11**  
Teacher  
The eraser. Which means it takes up more space. So if I were to put this into, a big container of water for example, if I were to put it into my little pitcher here, if I had this filled up and I dropped it in, well, which direction would my water have to go?  
Revoice |
Empirical Support for Accountable Talk

- When teachers of math, science, and reading use structured teacher-lead discussion methods...
  - steep changes in student achievement (Bill, Leer, Reams & Resnick, 1992; Chapin & O’Connor, 2004)
  - Retention for up to 3 years (Adey & Shayer, 1993, 2001; Shayer, 1999; Topping & Trickey, 2007a, 2007b)
  - Transfer across domains for up to 3 years (Bill, Leer, Reams & Resnick, 1992; Adey & Shayer, 1993, 2001; Shayer, 1999; Chapin & O’Connor, 2004)
  - Students perform better on non-verbal reasoning tests e.g. Ravens (Mercer, Wegerif & Dawes, 1999)
  - Reasoning itself improves (Kuhn & Zillmer, in press; Lin et al 2012)
• Definition of Transactivity
  • building on an idea expressed earlier in a conversation
  • using a reasoning statement

- Homozygous for both. One parent is orange and the other is not. Orange is dominant.
- I agree because all the kids are orange also.

- Self-oriented vs. other-oriented
- Representational vs. transformational
- Consensus-oriented vs. Conflict-oriented
<table>
<thead>
<tr>
<th>Example Teacher Utterance</th>
<th>Accountable Talk Move</th>
<th>Transact Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explain your thinking.</td>
<td>SAY MORE</td>
<td>SELF ORIENTED, REPRESENTATIONAL,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CONSENSUS ORIENTED</td>
</tr>
<tr>
<td>What's it prove? Put it into words.</td>
<td>PRESS FOR REASONING</td>
<td>SELF ORIENTED, REPRESENTATIONAL,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CONSENSUS ORIENTED</td>
</tr>
<tr>
<td>Let me see if I understand correctly. Are you saying they were all adopted?</td>
<td>REVOICE</td>
<td>SELF ORIENTED, TRANFORMATIONAL,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CONSENSUS ORIENTED</td>
</tr>
<tr>
<td>If capital ‘G’’s dominant, wouldn’t all babies be orange?</td>
<td>CHALLENGE</td>
<td>SELF ORIENTED, TRANFORMATIONAL,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CONFLICT ORIENTED</td>
</tr>
<tr>
<td>Can you repeat what she said?</td>
<td>RESTATE</td>
<td>OTHER ORIENTED, REPRESENTATIONAL,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CONSENSUS ORIENTED</td>
</tr>
<tr>
<td>Help him out Stephen. Can you add to what he said?</td>
<td>ADD MORE</td>
<td>OTHER ORIENTED, REPRESENTATIONAL,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CONSENSUS ORIENTED</td>
</tr>
<tr>
<td>Kelly, are they right? Do you agree or disagree with what they said?</td>
<td>AGREE/DISAGREE</td>
<td>OTHER ORIENTED, REPRESENTATIONAL,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CONFLICT ORIENTED</td>
</tr>
<tr>
<td>In your own words, explain why she’s right or wrong.</td>
<td>EXPLAIN OTHER</td>
<td>OTHER ORIENTED, TRANFORMATIONAL,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CONFLICT ORIENTED</td>
</tr>
</tbody>
</table>
8 Years of Positive Results

- **Foundational study**: students work with a partner and dialogue agent for support Learn **1.24 s.d.** more than individuals without support (Kumar et al., 2007a)

- Results inform **iterative design** of agent behavior
  - *Personalized agents* increase supportiveness and help exchange between students (Kumar et al., 2007b)
  - Agents are more effective when **students have control** over timing of the interaction (Chaudhuri et al., 2008; Chaudhuri et al., 2009)
  - Agents that employ *Balesian social strategies* are more effective than those that do not (Kumar et al., 2010; Ai et al., 2010)
  - Students are sensitive to agent *rhetorical strategies* such as displayed bias (Ai et al., 2010), *displayed openness* to alternative perspectives (Kumar et al., 2011), and targeted elicitation (Howley et al., 2012)
  - **Accountable talk agents** (Dyke et al., 2013; Adamson et al., in press)

- **Bazaar architecture** enables efficient, principle based agent development (Kumar & Rosé, 2011; Adamson & Rosé, 2012)
Exploring the Connection between Cognitive and Social Processes
Theoretical Framework

• Basic concepts of power and social distance explain social processes operating in interactions
Theoretical Framework

- Basic concepts of power and social distance explain social processes operating in interactions
Theoretical Framework

- Basic concepts of power and social distance explain social processes operating in interactions
- Social processes are reflected through patterns of language variation
Theoretical Framework

- Basic concepts of power and social distance explain social processes operating in interactions
- Social processes are reflected through patterns of language variation
- If we understand this connection, we can model language more effectively
Basic concepts of power and social distance explain social processes operating in interactions.

Social processes are reflected through patterns of language variation.

If we understand this connection, we can model language more effectively.

Models that embody these structures will be able to predict social processes from interaction data.
Theoretical Framework

- We gain influence in interaction through manipulation of horizontal and vertical social distance

- We manipulate distance through signaling
# Power, Relationships, and Transactivity

<table>
<thead>
<tr>
<th>Power, Cognitive Conflict, And Learning</th>
<th>Socio-Cognitive Conflict and Transactivity</th>
<th>Power Balance And Transactivity</th>
<th>Friendship, Transactivity, And Learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Piaget</td>
<td>Berkowitz &amp; Gibbs</td>
<td>Kruger &amp; Tomasello</td>
<td>Azmitia &amp; Montgomery</td>
</tr>
</tbody>
</table>
Modeling the Cognitive-Social Connection

Findings from Sociolinguistics → DBM Model → Model connecting speech style accommodation and Transactivity

(Jain et al., 2012) (Gweon et al., 2013)

Reflecting Perspectives and Relationships

Findings from Developmental Psychology

Reflecting Evidence of Consensus Building
Key insight: Accommodation occurs over time as a reflection of social processes that extend over time

Speech Observation (O): Obtained from speech feature such as pitch, loudness, voice probability, harmonic to noise ratio, voice quality. We used openSMILE (Eyben et al., 2010) to extract these features.

We use BNT (Murphy, 2001) to learn parameters for our DBN models.
Models

Explore a space of model defined by two independent factors:

1. **Direct Links**: Direct influence of one speaker’s style on another speaker’s style (Prior work mostly focused on this aspects by finding similarity between speech of partners).

2. **Indirect Links**: The influence of the relational gestalt between the two speakers that motivates the stylistic accommodation.
Evaluation

- SASDM Model that includes both symmetric Indirect Influence and Direct influence links performs best
Modeling the Cognitive-Social Connection

Findings from Sociolinguistics -> DBM Model -> (Jain et al., 2012)  

Findings from Developmental Psychology -> Model connecting speech style accommodation and Transactivity -> (Gweon et al., 2013)

Reflecting Perspectives And Relationships

Reflecting Evidence of Consensus Building
Correlation between Accommodation Score and Discussion Process Variables

<table>
<thead>
<tr>
<th>Model #</th>
<th>Acc &amp; Reas</th>
<th>Acc &amp; Transactivity</th>
<th>Acc &amp; Other Oriented Transactivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.15</td>
<td>0.11</td>
<td>0.36*</td>
</tr>
<tr>
<td>2</td>
<td>0.10</td>
<td>0.07</td>
<td>0.35*</td>
</tr>
<tr>
<td>3</td>
<td>0.12</td>
<td>0.07</td>
<td>0.30*</td>
</tr>
<tr>
<td>4</td>
<td>0.18</td>
<td>0.12</td>
<td>0.37*</td>
</tr>
<tr>
<td>5</td>
<td>0.18</td>
<td>0.13</td>
<td>0.36*</td>
</tr>
<tr>
<td>6</td>
<td>0.15</td>
<td>0.10</td>
<td>0.33*</td>
</tr>
</tbody>
</table>
$R^2 = 0.13$
Souflé Framework
(Howley et al., 2013)
Souflé Framework
(Howley et al., 2013)
- System of Engagement
  - Showing openness to the existence of other perspectives
  - Less final / Invites more discussion

- Example:
  - [M] Nuclear is a good choice
  - [HE] I consider nuclear to be a good choice
  - [HC] There’s no denying that nuclear is a superior choice
  - [NA] Is nuclear a good choice?
Engagement (Martin & White, 2005)

- **Findings**
  - *Correlational analysis:* Strong correlation between displayed openness of group members and articulation of reasoning (R = .72) (Dyke et al., in press)
  - *Intervention study:* Causal effect on propensity to articulate ideas in group chats (effect size .6 standard deviations) (Kumar et al., 2011)
  - Mediating effect of idea contribution on learning in scientific inquiry (Wang et al., 2011)
Souflé Framework
(Howley et al., in press)

Authority

Transactive Knowledge Integration

Person

Engagement

Authority

Engagement
Analysis of Authoritativessess

Water pipe analogy:
Water = Knowledge or Action
Source = Authoritative speaker
Sink = Non-authoritative Speaker

Authoritativeness Ratio = \frac{Source\ Actions}{Actions}
The Negotiation Framework
(Martin & Rose, 2003)

Additionally...
ch (direct challenge to previous utterance)
o (all other moves, backchannels, etc.)
The Negotiation Framework
(Martin & Rose, 2003)

- **K1**: giving knowledge, information, opinions, or facts
- **K2**: requesting knowledge, information, opinions, or facts
- **A1**: Narrating or performing your own non-verbal action
- **A2**: Instructing, suggesting, or requesting non-verbal action
The Negotiation Framework
(Martin & Rose, 2003)

- **K1**: Giving knowledge, information, opinions, or facts
- **K2**: Requesting knowledge, information, opinions, or facts
- **A1**: Narrating or performing your own non-verbal action
- **A2**: Instructing, suggesting, or requesting non-verbal action

Source or Sink?  Secondary
Primary  Knowledge  Type of Content?  Action
The Negotiation Framework

(Martin & Rose, 2003)

- **K1**: giving knowledge, information, opinions, or facts
- **K2**: requesting knowledge, information, opinions, or facts
- **A1**: Narrating or performing your own non-verbal action
- **A2**: Instructing, suggesting, or requesting non-verbal action

Source or Sink?
- Primary
- Secondary

Type of Content?
- Knowledge
- Action
The Negotiation Framework
(Martin & Rose, 2003)

Additionally...
ch (direct challenge to previous utterance)
o (all other moves, backchannels, etc.)
The Negotiation Framework
(Martin & Rose, 2003)

Additionally...

ch (direct challenge to previous utterance)
O (all other moves, backchannels, etc.)

Authoritativeness:
K1 + A2
K1 + K2 + A1 + A2
<table>
<thead>
<tr>
<th>Doctor Bob</th>
<th>BTW: Is it safe to keep increasing Tmax?</th>
<th>05.12.17</th>
</tr>
</thead>
<tbody>
<tr>
<td>sa08</td>
<td>no</td>
<td>05.12.33</td>
</tr>
<tr>
<td>Sa04</td>
<td>nope. you'll blow out your turbine</td>
<td>05.12.33</td>
</tr>
<tr>
<td>Doctor Bob</td>
<td>Good. Material properties constrain the maximum temperature we can use in a Cycle.</td>
<td>05.12.39</td>
</tr>
<tr>
<td></td>
<td>For our cycle, Tmax cannot be more than 570C</td>
<td></td>
</tr>
</tbody>
</table>
Set up!

<table>
<thead>
<tr>
<th>Doctor Bob</th>
<th>BTW: Is it safe to keep increasing Tmax?</th>
</tr>
</thead>
<tbody>
<tr>
<td>sa08</td>
<td>no</td>
</tr>
<tr>
<td>Sa04</td>
<td>nope. you'll blow out your turbine</td>
</tr>
<tr>
<td>Doctor Bob</td>
<td>Good. Material properties constrain the maximum temperature we can use in a Cycle. For our cycle, Tmax cannot be more than 570°C</td>
</tr>
</tbody>
</table>
Authoritativeness (Martin & Rose, 2003)

Findings

- Authoritativeness measures display how students respond to aggressive behavior in groups (Howley et al., in press)
- Authoritativeness predicts learning ($R = .64$) and self-efficacy ($R = .35$) (Howley et al., 2011)
- Authoritativeness predicts trust in doctor-patient interactions ($R$ values between .25 and .35) (Mayfield et al., 2014)

Computational Work

- Detectable in collaborative learning chat logs ($R = .86$)
- Detectable in transcribed dyadic discussions in a knowledge sharing task ($R = .95$) (Mayfield & Rosé, 2011)
- Detectable in transcribed doctor-patient interactions ($R = .96$) (Mayfield et al., under review)
Future
Researching Attrition in Massively Open Online Courses

Losing many acquaintances last week

Losing many acquaintances this week

Having many significant acquaintances

Accountable Talk Course Survival Curves

Fantasy Course Survival Curves

Python Course Survival Curves

SURE, WE LOSE SOME--BUT WHO'S COUNTING?

DROPOUT FACTORIES

What would MOOC training wheels look like?
What will support look like in MOOCs?

- **Challenges**
  - Asynchronous
  - Split attention
  - Differing norms of participation

- **Idea: Personal Mentoring Paradigm**
  - Inspired by recommender systems
Thank You!

Questions?
Style based models:  
*The Independent Speaker Model*

The dynamics of each speaker are independent of the other speaker.

**Model:** ISM  
**Direct Link:** NoDirect  
**Indirect Link:** NoIndirect
Style based models:  
*The Cross-speaker Dependence Model*

A speaker’s style depends on their partner’s style at the previous turn.

**Model:** CSDM  
**Direct Link:** Direct  
**Indirect Link:** NoIndirect

---

88
Accommodation state models:

The Symmetric Accommodation State Model

Both partners styles depend on mutual accommodation to one another.

**Mode:** SASM  
**Direct Link:** NoDirect  
**Indirect Link:** Symetric
Accommodation state models: The Asymmetric Accommodation State Model

Accommodation state associated with every speaker turn

Model: AASM
Direct Link: NoDirect
Indirect Link: Asymmetric
Accommodated style dependence models:

The Symmetric Accommodated Style Dependence Model

A speaker’s style depends both on mutual accommodation and the partner’s style in the previous turn.

**Model:** SASDM

**Direct Link:** Direct

**Indirect Link:** Symmetric
Accommodated style dependence models:

**Asymmetric Accommodated Style Dependence Model**

The accommodation state associated with every speaker and a speaker’s style depends on the partner’s style.

**Model:** AASDM  
**Direct Link:** Direct  
**InDirect Link:** Asymmetric
Task: 2 students debating about the cause of the Ottoman Empire fall

Participants: Male undergraduates between 18 and 25 years old

Each student equipped with mic

8 min sessions, recorded and transcribed

Evaluation

• Goal: evaluate style accommodation
  – Factor out lexical similarity that results from speaking on the same topic
  – Measure success by separating Real pairs and Constructed pairs
  – No gold standard labels needed

• For each Constructed pair
  – Start with a real pair
  – Select one speaker S
  – Replace each of S’s turns with the most lexically similar turn from another conversation

• 3 fold cross validation
  – For each test conversation, compute an Accommodation score for the Real pair and each Constructed pair
Fake Partner Construction

Real Partners (Speaker1 and Speaker2)

Speakers from other debates

Fake partner for speaker1

Fake partner for speaker2

Turn1

Turn2

Turn1

Turn2
Evaluation

• Goal: evaluate style accommodation
  – Factor out lexical similarity that results from speaking on the same topic
  – Measure success by separating Real pairs and Constructed pairs
  – No gold standard labels needed

• For each Constructed pair
  – Start with a real pair
  – Select one speaker S
  – Replace each of S’s turns with the most lexically similar turn from another conversation

• 3 fold cross validation
  – For each test conversation, compute an Accommodation score for the Real pair and each Constructed pair
Evaluation

• Results evaluated with an ANOVA model
  – Dependent Variable: Accommodation Score
  – Independent Factors: RealVsConstructed, DirectLink, IndirectLink

• Significant main effect of the RealVsConstructed factor
  – $F(1,780) = 18.22, p < .0001$
  – Real pairs demonstrate higher Accommodation scores than Constructed pairs
Evaluation: Effect of Indirect Links

- Posthoc comparison of interaction between IndirectLink and RealVsConstructed factors
  - Significant difference only when IndirectLinks are present (both Symetric and Asymetric)

<table>
<thead>
<tr>
<th>Model</th>
<th>Direct Link</th>
<th>Indirect Link</th>
<th>Real $\mu(\sigma)$</th>
<th>Constructed $\mu(\sigma)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SASDM</td>
<td>Direct</td>
<td>Symmetric</td>
<td>.54(.23)</td>
<td>.44(.29)</td>
</tr>
<tr>
<td>SASM</td>
<td>NoDirect</td>
<td>Symmetric</td>
<td>.54(.23)</td>
<td>.44(.29)</td>
</tr>
<tr>
<td>AASM</td>
<td>NoDirect</td>
<td>Asymmetric</td>
<td>.6(.24)</td>
<td>.51(.3)</td>
</tr>
<tr>
<td>AASDM</td>
<td>Direct</td>
<td>Asymmetric</td>
<td>.61(.24)</td>
<td>.48(.3)</td>
</tr>
<tr>
<td>CSDM</td>
<td>Direct</td>
<td>NoInDirect</td>
<td>.6(.26)</td>
<td>.52(.3)</td>
</tr>
<tr>
<td>ISM</td>
<td>NoDirect</td>
<td>NoInDirect</td>
<td>.56(.25)</td>
<td>.51(.32)</td>
</tr>
</tbody>
</table>
Evaluation: Effect of Direct Links

<table>
<thead>
<tr>
<th>Model</th>
<th>Direct Link</th>
<th>Indirect Link</th>
<th>Real $\mu(\sigma)$</th>
<th>Constructed $\mu(\sigma)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SASDM</td>
<td>Direct</td>
<td>Symmetric</td>
<td>.54(.23)</td>
<td>.44(.29)</td>
</tr>
<tr>
<td>SASM</td>
<td>NoDirect</td>
<td>Symmetric</td>
<td>.54(.23)</td>
<td>.44(.29)</td>
</tr>
<tr>
<td>AASM</td>
<td>NoDirect</td>
<td>Asymmetric</td>
<td>.6(.24)</td>
<td>.51(.3)</td>
</tr>
<tr>
<td>AASDM</td>
<td>Direct</td>
<td>Asymmetric</td>
<td>.61(.24)</td>
<td>.48(.3)</td>
</tr>
<tr>
<td>CSDM</td>
<td>Direct</td>
<td>NoDirect</td>
<td>.6(.26)</td>
<td>.52(.3)</td>
</tr>
<tr>
<td>ISM</td>
<td>NoDirect</td>
<td>NoDirect</td>
<td>.56(.25)</td>
<td>.51(.32)</td>
</tr>
</tbody>
</table>

- Posthoc comparison of interaction between DirectLink and RealVsConstructed
  - Significant separation irrespective of whether DirectLinks are present or not.