Lecture 7:
A theory of discourse structure

- copies of handout from last time
- A2o printout: the Kasparov diary for annotation
- Load up the YouTube intention video?
- No class Tues.

? Lessons learned from prior studies?
(no time, most likely)

Last lecture we looked at some suggestive discourse examples that provided evidence of hidden structure. I want to remind you of this one: Ed mengra's resolution lens

#2(a): lines 3-5:

1. Albert...
2. He
3. Actually
4. Anyways ...
5. It
6. They're

Talking about Albert (next sentence could be about Albert or about the book's arrival).

Signals entry into new topic? (quarks, only plural referent)

Book

Segment

@ talking about the contents of the book

Signal of return to prior topic

(Why does Spencer bother?)

(under "it" or "it" in 5 is "he"? )
"Any other observations from class?"

(Ask A. to report from last time)

(See paper on discussing so as discuss works although broader citation better for web page?)

So what we're observing here

- existence of discourse segments
- incoherence when expected segment structure violated

Now, before I kind of argued that you wouldn't expect to see a discourse like Zc in real life.

So you might conceive of incoherence as an acceptability criterion

But incoherent

But incoherent discourses happen all the time in real life!

And actually, they are not always a bad thing at all.

Consider #4 on handout. I'm picking at this particular incoherent discourse due to the pedagogical point I'm driving at.

- incoherent, b/c of interruption

(and it is an interruption, or
(or alternatively the interleaving of two conversations)

Ow. "them" would refer to the "kids"

(although maybe if the kids are badly enough behaved, you would want to put the kids away)

- incoherent, but normal

Speaker wants, in some sense, to be incoherent.

If the kids are going about to microwave the new puppy, the speaker does not want to wait to finish their gripping tale of the Sage of & the Groceries before telling the kids to knock it off.

- listeners (humans or systems) must recognize (or adapt to)

incoherence.

In this case, presumably the speaker gives a lot of cues that the listener can pick up on: turning, change in volume,
So, two points so far: that a theory of global discourse structure should take into account:
- structural units or segments
- coherence; incoherence (maybe amounts to understanding the relations between segments).

Now, one more thought-provoking discourse example, this one involving two people (finally).

>5. One ass &

or actually, once you read it, it seems more like one human being and one somewhat flawed AI system.

Read:
- clearly this is a conversation going off the rails. [bulletproof glass joke from Grishman?]

What's the problem?
B's responses are, from some perspective, totally rational:
- answer to 1st q is truthful
- second utterance by A isn't a q, so B doesn't technically have to do anything,
  but B acknowledges A anyway, which is nice & polite.

Crucially, though, B is failing to recognize A's intentions when making these sorts, and that failure to take attention into account is causing all sorts of hard feelings, etc.

So: third elt: importance of recognizing intentions.
After all, perhaps the most important thing to note about conversation is that people generally have them for a reason (even if those reasons aren't very good).

Geroz; Sidner [1986] theory: has really influenced my thinking when I pay attention to daily conversation.
- primary atoms: discourse segment purpose (DSP)
  - a single intention, the recognition of which motivates the discourse segment.
  
- ok to use phrasing that is culturally specific?

Eff in 5: DSPs find out the time
- check that B knows the time
- test on AI system (een if true)
This explains why A continues the conversation, and does so by repeating the point. They're doing another attempt to get their intent recognized.

DSP relations: relations when DSPS... DSP1 dominates DSP2 if satisfying minimal set of DSP1 (trees) DSP2 forthers DSP1.

This kind of relation naturally suggests a tree.

ex: [Jungreis & Martin] (from back in the day when people called travel agents)

DSP1: caller: agent books flight

DSP2: a: c gives departure

DSP3: a: c gives destination

DSP4: a: c explains which

"Ithaca."

Note: a 'coherent' conversation could be 'incoherent' w.r.t. participants.

Should coherence be a property of the conversation mutually or the participants individually?

Note that the intentions of both parties are in the tree.

Also note that this structure is a property of the joint conversation, not necessarily the viewpoint that either participant themselves has.

So the fact that our agendas are what's interesting makes a lot of sense; you get the same structure whether or not the caller is a valid customer or, say, someone from a rival company that's trying to tie up the phone line,

DSP tree structure = embedded discourse segments.

But none of this theory so far explains why some referents seem to somehow disappear.

How to account for?

Well, parse trees relate to pushdown automata, so it's useful to think of trees being related to stacks. (semi-)natural
Attentional component:

A conversation has a stack of focus spaces, one per discourse segment.
Contains DSP salience entities

- pop f.s. when DSPs are recognized/realized.
- no "they've"

Pop f.s. with possible referents
- explains why there are missing (or inconvenient) referents, like "the thing"
- disappearing.
- ease cognitive load by reducing the # of possible referents you have to look for.

Q: How deep are people's stacks be?
   (let about Walker's quine theory)

Q: Do parts of trees stay around beyond boundary of the conversation
   (i.e., our multiple days)

It's probably true that true intentions last across sessions, and shared
humans keep some always "at the bottom of the stack".

Incoherence: stack doesn't match the tree.

Some implications:

- hearers must act: understanding of DSP for speaker to be satisfied.
  ("closure": Clark '96, Norman '88).

- Ack can consist of attention
  (if I keep staring @ you after I've told you what time
  it is this seems weird)
  (if you don't look @ me while I'm trying to convey a key point)

- speakers can close segments w/ acts (thanks, great)

- hearers can provide next expected contribution (disclaim a compliment,
  to show you knew it was just meant to be polite)

- paraphrasing, repeating

Q: If it's all about getting your DSP recognized, why don't people just 'say what
they mean'?

A: B/c there are other factors (and hence intentions) at play also
controlling phrasing.

ex: status, common knowledge

<Show youtube video - From YouTube directly! Better bandwidth>