(1) Collect 100 summaries (iteratively over the first year work up to 500 summaries)

20 dialogues, get 10 summaries for each. Replicate what we did.

Short dialogues (at least 4 turns each), two person, select so that they have some of the implicit marker features. Stick with what we said in the proposal.

“at least two differing perspectives that contain at least 6 dialogic features and at least 4 expressions of evaluative stance. Dialogic features may include uses of particular discourse cues, using punctuation such as capitalization, asterisks, exclamation points, or ellipses, name calling, sarcasm, and strong agreement vs.disagreement.”

Death Penalty 10 dialogues

Gay Marriage 10 dialogues

Then expand to other topics next phase possibly.

Use undergrads to produce summaries AND pyramid annotations.

Could collect Turker summaries in a second phase. Need more control to

get started.

Re-use the HIT that Reid and I used with James and collect via Turk, but we can use our own annotators.

Is there a way to align the summaries automatically or semi-automatically with the

original dialogues. NO, or only for some of the stuff. Because of these ‘editorial’ comments like “s2 gets defensive”.

Instructions: summarization

* Training (n = 1)
* underconstrain:
* range =
* summarise the conversation for someone who wasn’t present, but try to avoid a simple play by play: ‘eg X said Y, A said B’
* provide some good examples (but don’t leave these constantly in view)

(2) GET THEM PYRAMID annotated.

What would the Pyramid annotation tool from Columbia let us do.

Discourse connectives using sense disambugation

Tense biagram

Discourse entity features

Sentence complexity