Top-down and bottom-up cues to speech acts

Ye Tian[1] & Chris Cummins[2]
[1]Laboratoire de linguistique formelle, Université Paris Diderot ; [2]Department of Linguistics and English Language, University of Edinburgh

When we are presented with an utterance, we not only interpret its semantic meaning, but also its discourse purpose, and, on the other side of the coin, its relevance to the broader context. These two notions are often construed in the frameworks of speech act and Question Under Discussion (QUD) (Ginzburg, 2012; Roberts, 2012; to appear). This reasoning can take two approaches: a top-down approach, where we reason about the speaker's discourse goals and consider how the utterance contributes to their realisation; or a bottom-up approach, where we use the content of the utterance itself as a basis for inferring what the speaker is attempting to achieve. Computationally, these two approaches can be synthesised within a cue-based approach to speech act recognition/ disambiguation, in which high-level and low-level considerations are both used as probabilistic cues to the successful classification of a speech act. However, little is known about how these two sources of information are integrated by hearers in offline interpretation and in online processing. In this paper, we argue that a better understanding of this process is necessary, not only because of its theoretical implications for our analysis of discourse in general, but also because of its methodological implications for experimental semantics and pragmatics.

We all know that our understanding of the overarching discourse goals influences the speech act interpretation of an utterance. What we don't know is what hearers do when extended context is unavailable. They may use a bottom-up approach and reason about contextual relevance using local cues, or they may use a top-down approach and imagine a (stack of) discourse goals which allow rich inferences to be drawn on the current utterance. For example, if someone hears (1) out-of-the-blue, they might accommodate the QUD (2) and interpret (1) literally; or they might imagine a richer discourse goal, accommodate the QUD (3), and interpret (1) to imply that the date wasn't good.

(1) The coffee was not bad.
(2) Was the coffee bad?
(3) How did your date go?

Experiment: inferences about prior and subsequent context for speech acts

The goal of our research is to shed light on the interplay between high-level and low-level factors that bear upon the recognition of speech acts. A first step is to determine the degree of variability in the inferences participants are able or willing to draw based on decontextualized utterances, about the nature of the current discourse context. We uses a Cloze task in which participants are presented with isolated utterances and asked to suggest the context in which these utterances took place – either providing the preceding turn, the following turn, or both, according to their preference. If the bottom-up approach is primary, we would expect more consistent and literal interpretations; if the top-down approach is primary, we would expect more varied and enriched interpretations.

We constructed 34 potentially ambiguous utterances as experimental items. They were constructed such that each could be interpreted as instantiating at least two distinct speech acts, as shown for examples (5) to (7) below. In addition, 26 unambiguous fillers were

constructed. 63 participants were recruited via Amazon Mechanical Turk. They read each utterance and could fill in the preceding utterance, the following utterance, or both.

(5) I am on my way. (ambiguous between answer to a "where" question and acceptance of a request)
(6) I'm not working on Saturday. (ambiguous between answer to a question and acceptance of a suggestion/invitation)
(7) Are you wearing that shirt? (ambiguous between question and complaint)

Results: for each response, we coded the fine-grained speech act (e.g. "information-seeking question", "offer" etc.), as well as whether the approach is "bottom-up" (literal interpretations) or "top-down" (enriched interpretations). The participants filled the preceding context in 77% of responses and the following context in 82% of responses. On average, each utterance received 4 different interpretations. Although declaratives received more distinct interpretations than interrogatives did, this may be a reflection of the content of the utterances rather than their sentence type. Among the interrogatives, we included some items that have been argued strongly to cue particular "non-literal" interpretations, and these potentially conventionalised items did indeed appear to admit more homogeneous interpretations. Items such as "Can you find out the name of this song?" and "Can you fix my bike?" were widely interpreted as requests, although the formally similar "Can you email your boss?" was variously interpreted as either a request (5) or as a suggestion (9).

(8) B: Can you email your boss? A: When I get the chance, sure.
(9) A: What should I do about this situation?
B: Can you email your boss? A: Yeah that's a good idea.

69% of interpretations were top-down, with interrogative sentences attracting the highest rate of top-down interpretations (including requests, offers, and implied answers to previous questions). An interpretation is more likely to be top-down when the preceding context is filled (71% top-down) than when it is left empty (60%). Whether the following context is filled doesn't make a difference in the proportions of top-down readings.

Table 1 summary of results

Type of item	number of items	% preceding cntxt filled	% following cntxt filled	number of speed acts interpreted	% top-down readings
positive declarative	14	84%	75%	4.21	62%
negative declarative	6	82%	79%	3.83	53%
interrogative	14	67%	90%	3.07	83%

Overall, our results show that when presented with utterances context out of the blue, both approaches can be used. More often, participants take the 'top-down' approach: attribute rich overarching conversational goals and thereby derive additional pragmatic interpretations of utterances. Inferring from our data to speech-act recognition in natural settings (when at least some context is available), it is likely that top-down considerations play a more primary role interpreting the speech act of an utterance and its discourse relevance to the broader context. Our results also demonstrates that certain utterances are indeed ambiguous when presented out of context, to an extent that makes them suitable for further experimental investigation of the online interplay between top-down and bottom-up interpretative processes.

References: Ginzburg, J. (2012). The interactive stance. CSLI: Center. Oxford University Press. Roberts, C. (2012). Information structure: toward an integrated theory of formal pragmatics. *Semantics and Pragmatics*, 5(7):1-19. Roberts, C. (to appear). Speech acts in discourse context. To appear in D. Fogal, D. Harris and Matt Moss (eds.), *New Work on Speech Acts*. Oxford University Press.