

Workshop on Speech Act Theory  
ZAS, Berlin: 11-13 June 2015

# When are speech acts recognised?

Chris Cummins  
University of Edinburgh  
[c.r.cummins@gmail.com](mailto:c.r.cummins@gmail.com)



# Overview

- Querying some of the assumptions of recent theoretical/experimental work
- Canvassing perspectives on the architectural implications

# Question

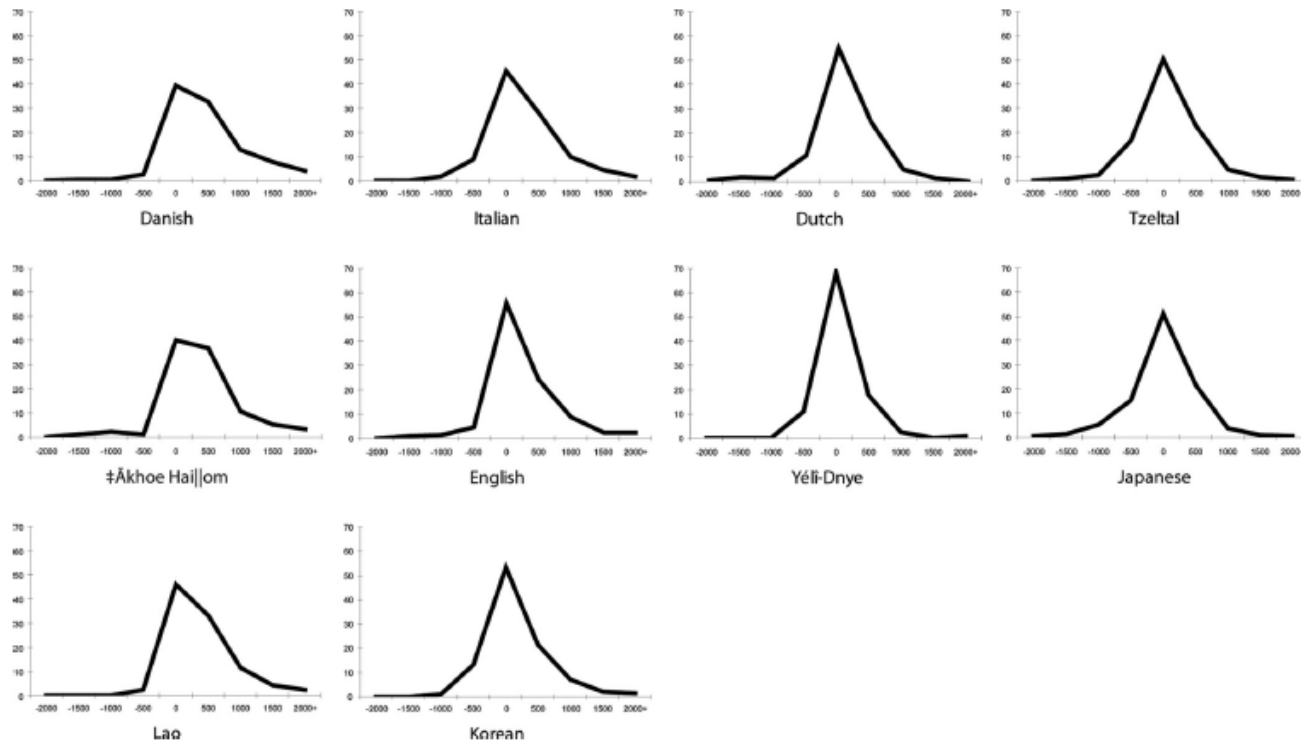
- “When are speech acts recognised?”...in very broad terms
  - Not so much the time-course to ms accuracy (although that would certainly be good to know)
  - Just broadly whether they’re recognised during the course of the utterance, or afterwards, or whether it depends

# The case for early recognition

- See e.g. Cummins and De Ruiter (2014)
  - Reasonably convincing evidence that speech acts / “dialogue act types” must be recognised before the end of the utterance...
  - ...but this gives rise to some potentially problematic consequences

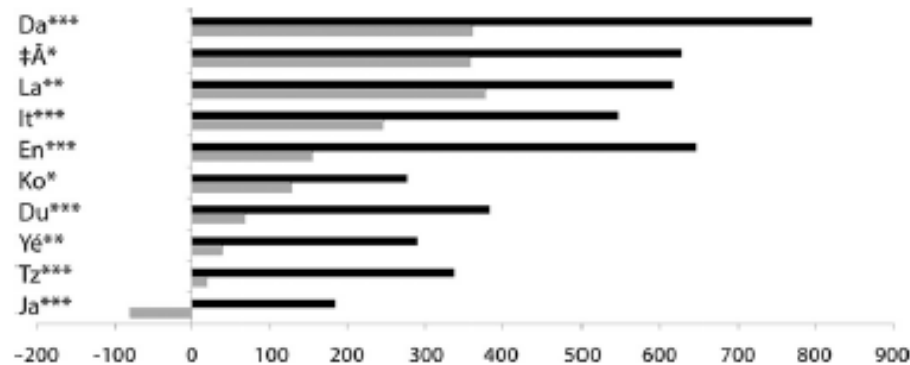
# Turn-taking

- Fast and fluent, in normal conversation
- Latencies typically <500ms (Stivers et al. 2009)



# Turn-taking

- Fast and fluent, in normal conversation
- Latencies typically <500ms (Stivers et al. 2009)
  - cf. Brown-Schmidt and Tanenhaus (2006) on speech planning
- Utterance planning must begin during previous utterance
- Moreover, turns are relevant to one another
  - For instance, respecting adjacency pairs (e.g. question-answer)



# Turn-taking

- Fast and fluent, in normal conversation
- Latencies typically <500ms (Stivers et al. 2009)
  - cf. Brown-Schmidt and Tanenhaus (2006) on speech planning
- Utterance planning must begin during previous utterance
- Moreover, turns are relevant to one another
  - For instance, respecting adjacency pairs (e.g. question-answer)
  - Also rapid and fluent in doing so
  - Unsurprising given Levinson's (1995) observation about the possible communicative effect of even a 500ms silence (in his example, unwillingness to comply with request)

# Implications

- Casts doubt on attractive accounts of how indirect speech acts are identified (Gordon and Lakoff 1971, Searle 1975)
  - Though these were already in dispute: Gazdar (1981) – no ‘literal meanings’; Levinson (1983) – preponderance of indirect acts
- Example: “Could you pass the salt?”
  - Analysable, rationally, as a pre-request
  - But “Could you?” (likewise “Would you?”) seems to be conventional
  - Expectation about how “Could you...?” is going to continue (at least at a speech act level)
  - Suggestive of a cue-based strategy



# Cues...to what?

- If we do use a cue-based strategy to recover speech acts, what good does that do us?
- e.g. the question-answer adjacency pair
  - In principle, potentially helpful to know that something is a question and requires an answer...
  - ...but in practice, is that any use without knowing the content of the question?
  - Not our goal just to produce some appropriate quota of adjacency pair transitions

# Limitations of question-answer



ZAS, Berlin, 12.06.2015

10/20

# Why recognise a question?

- Question-recognition helpful presumably if the hearer can exploit it to produce a timely and relevant answer
- Can they?
  - No obvious linguistic features associated with (wh-)question-answering, e.g. syntactic constructions (although perhaps more likely to be fragments)
  - Possibly for yes-no questions, in that space of possible/likely responses is heavily constrained
  - Similar story for some other speech acts, e.g. greetings
    - Can use a formulaic expression to buy time to come up with something more complete

# Strategies?

- As per *Yes, Prime Minister*
  - “That’s a very silly question”
  - “That’s a very good question. I’d like to thank you for asking it.”
  - “That’s a very interesting question, and there are nine points that I’d like to make in answer to it.”
  - “There’s a very full answer to that question, but it involves matters that are being discussed in confidence.”
  - “I think the more important question is this: ...”
- As per *Father Ted*
  - “Yes.”
  - “That would be an ecumenical matter.”

# Implications for production

- Suggests a setup that is not necessarily intention-first (this then being transcoded into a verbal message)
- Potentially a matter of having a few messages ready to go, and launching one if it's (reasonably) appropriate

# Implications for analysis

- What about the speech acts that aren't useful in the way they constrain the space of possible responses?
  - e.g. Searle's (1975) assertives and commissives, and some of his expressives and declaratives
  - Are there standard formulae for responding to swearing, or promising as opposed to threatening? (cf. Haigh et al. 2011)

If the results are written next week, I'll put you on the paper as an author.	Thanks!
	You don't have to do that. Really? You're kidding.
If the results are written later than next week, I'll take you off the project.	Wow – this is a shock.
	Don't worry, they'll be done. I'm sorry you feel that way.

# Similar computational question

- Traum (1999) on the usefulness of speech acts (dialogue act types) as a level of analysis in dialogue systems
  - Potentially valuable in complex systems, and a solution to some issues around scalability
  - Less useful in simple systems
    - My example: vending machine
    - No use in being able to distinguish “requests” as a type: everything (apart from side-sequences) is like this, no common denominator to request-responses

# How many speech acts?

- And how are they defined so as to be distinguishable from one another?
- Especially difficult if we want psychological reality...
  - Consider e.g. project of demonstrating that all Searle's candidate speech acts were distinguishable in processing
  - But without establishing a tagset, potentially a waste of time to go ahead and tag corpora



# Other classification possibility

- If there are speech acts that are recognised early and inform processing, while others are not, do they really belong in the same classification system?
- Could consider the first alongside observed behaviours that promote particular responses
  - Sneezing
  - Using an erroneous form
  - Using a standard form of words, e.g. in a religious context
  - (None of these seem to be 'dialogue acts' in the usual sense)

# (Some of) the data so far

- Gisladdottir et al. (2012) – accurate classification of speech acts of three kinds (and early differences in reading)
- Gisladdottir et al. (2015) – corresponding early ERP effects (frontal positivities at 200ms)
- Neural correlate of speech act recognition? Or of other processes specific to these particular occurrences and their discourse consequences?

Condition	Context	Target Sentence
Answer	Hoe ga je voor het ticket betalen? <i>How are you going to pay for the ticket?</i>	Ik heb een creditcard. <i>I have a credit card.</i>
Declination	Ik kan je wat geld lenen voor het ticket. <i>I can lend you money for the ticket.</i>	Ik heb een creditcard. <i>I have a credit card.</i>
Pre-Offer	Ik heb geen geld om het ticket te betalen. <i>I don't have any money to pay for the ticket.</i>	Ik heb een creditcard. <i>I have a credit card.</i>

# What might we need?

- Perhaps more data will allay this concern, and make it abundantly clear that the online processing really corresponds to what we call “speech act recognition”
- Perhaps it would be helpful first to have more detailed theories about how speech act recognition fits within the whole process of interaction
  - Asking whether it’s *always* important, and if not, what factors bear upon that

# References

- Brown-Schmidt, S., & Tanenhaus, M. K. (2006). Watching the eyes when talking about size: an investigation of message formulation and utterance planning. *Journal of Memory and Language*, 54: 592-609.
- Cummins, C., & De Ruiter, J. P. (2014). Computational approaches to the pragmatics problem. *Language and Linguistics Compass*, 8: 133-143.
- Gazdar, G. (1981). Speech act assignment. In A. K. Joshi, B. L. Webber & I. A. Sag (eds.), *Elements of Discourse Understanding*. Cambridge: Cambridge University Press. 64-83.
- Gisladottir, R. S., Chwilla, D. J., Schriefers, H., & Levinson, S. C. (2012). Speech act recognition in conversation: experimental evidence. In N. Miyake et al. (Eds.), *Proceedings of the 34th Annual Meeting of the Cognitive Science Society*. 1596-1601.
- Gisladottir, R. S., Chwilla, D. J., & Levinson, S. C. (2015). Conversation electrified: ERP correlates of speech act recognition in underspecified utterances. *PLOS ONE*, 10(3): e0120068.
- Gordon, D., & Lakoff, G. (1971). Conversational postulates. *Seventh Regional Meeting of the Chicago Linguistic Society*, 63-84.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Levinson, S. C. (1995). Interactional biases in human thinking. In E. N. Goody (ed.), *Social intelligence and interaction*. Cambridge: Cambridge University Press. 221-260.
- Searle, J. R. (1975). Indirect speech acts. In P. Cole & J. Morgan (eds.), *Syntax and Semantics, Vol. 3: Speech Acts*. New York: Academic Press. 59-82.
- Stivers, T., et al. (2009). Universals and cultural variation in turn-taking in conversation. *PNAS*, 106: 10587-10592.
- Traum, D. R. (1999). Speech acts for dialogue agents. In M. Wooldridge & A. Rao (eds.), *Foundations of Rational Agency*. Dordrecht: Kluwer Academic Publishers. 169-201.