



# Language and the emergence of Theory of Mind

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

CENTRE for  
**LANGUAGE EVOLUTION**  
[www.lel.ed.ac.uk/cle](http://www.lel.ed.ac.uk/cle)

# Overview

- Theory of Mind (ToM) and its interplay with the emergence of language, at various levels
- Outline of recent computational work on the acquisition of the lexicon (Marieke Woensdregt's PhD research)
- Other involvement of ToM in language
- Open questions about the 'mutual scaffolding' of ToM and language



# Theory of Mind

- Generally, the ability to attribute beliefs, desires and intentions to others (and to ourselves)
  - cf. Iris Murdoch – “Love is the extremely difficult realisation that something other than oneself is real”...
- An enormously powerful means of explaining a great deal of superficially mysterious behaviour in our conspecifics

# Projecting ToM onto other things



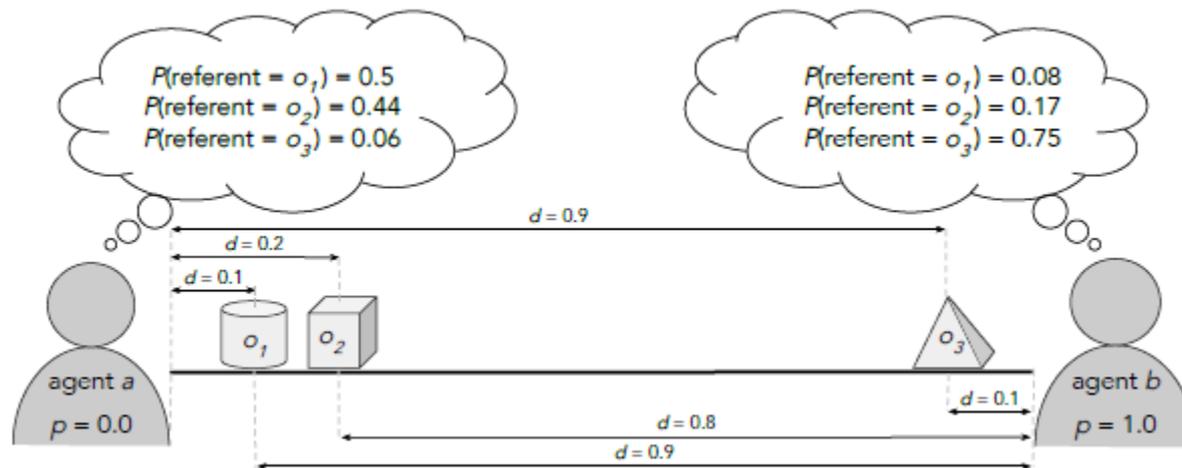
[https://www.youtube.com/watch?v=3\\_9aoCvFBNO](https://www.youtube.com/watch?v=3_9aoCvFBNO)

# ToM in Gricean pragmatics

- Grice's (1957) conception of meaning<sub>NN</sub>:
  - “for x to have meant<sub>NN</sub> anything, not merely must it have been “uttered” with the intention of inducing a certain belief but also the utterer must have intended the “audience” to recognize the intention behind the utterance.”
- In practice, a fair chunk of what we appear to communicate doesn't seem to rely on this kind of intention recognition
  - Also, ToM impairment (e.g. in ASD) doesn't preclude advanced language use or (always) recovery of implicature, metaphor, irony...
  - Suggests either that the analysis of these kinds of meanings is wrong, or that of ASD is wrong
- However, it is clearly present, and pretty widespread

# Woensdregt: lexical acquisition

- Marieke's PhD thesis concerns the acquisition of recurrent mappings between symbols and referents (the lexicon)
- Computational modelling of a simple scenario (ibid., p.88):



# Woensdrecht: lexical acquisition

- Marieke's PhD thesis concerns the acquisition of recurrent mappings between symbols and referents (the lexicon)
- Computational modelling of a simple scenario (ibid., p.88)
- Construing the task of lexical acquisition as involving **joint inference** of the lexicon and the speaker's perspective
  - Knowledge of one bootstraps the other, for obvious reasons

# Questions explored

- How does the ability to recognise a difference between the speaker's perspective and one's own help?
  - Does it lead to faster and more accurate acquisition of the lexicon?
  - Does it matter whether we start with a bias towards attributing the speaker our own perspective?
- Is it beneficial for acquisition if speakers are pragmatic in their production, in using less ambiguous forms?
  - Do we benefit from having higher-order perspective-taking capabilities?
  - Does this influence the shape of the lexicon that emerges over cultural-evolutionary time?

# Questions unexplored

- Model is clearly a substantial simplification
- Doesn't permit us to address context-dependent reference (deixis) such as the use of pronouns
  - Referring expressions in this model can be ambiguous, but have stable denotations across speakers/scenarios
- Assumes shared (visual) access to candidate referents, and systematic differences between the two perspectives in play
  - In reality, speakers can name things in privileged ground, or that aren't present, or don't have extensions
  - Saliency of referent from one perspective doesn't predict its saliency (or lack thereof) from another – e.g. “*I'm happy*”

# Joint inference

- Even so, model captures part of this important idea: that we are inferring lexicon and perspective at the same time
  - Actually quite a general observation – I want to learn *what is the case* rather than just *what you mean to tell me about what is the case*
- We can't rely purely upon cooccurrence between words and our experience
  - At least, we would need to explain away a lot of problematic usages, as far as our hypotheses were concerned
  - Helpful to be able to appeal to the idea that the words denote salient entities in the mind of the speaker

# ToM in language acquisition

- Putatively relevant, ontogenetically, at a number of levels
  - Acquisition of metacognitive verbs (*think, know*) as a consequence and cause of ToM development (Astington and Olson 1990, i.a.)
  - Mutual exclusivity inferences (Markman and Wachtel 1988)  
(although these are simpler in many respects...)

Show me the dax



# ToM in language acquisition

- Putatively relevant, ontogenetically, at a number of levels
  - Acquisition of metacognitive verbs (*think, know*) as a consequence and cause of ToM development (Astington and Olson 1990, i.a.)
  - Mutual exclusivity inferences (Markman and Wachtel 1988)  
(although these are simpler in many respects...)
  - And, at a higher level, speech acts – “how to do things with words”
    - Example: how do we learn that asking a question is a good way to find out information that we want to know?
    - Perhaps possible without ToM, but more difficult?

# Language in development of ToM?

- Given that language benefits so widely from ToM, does it create an ecological niche for ToM's emergence?
  - Will language (or its precursors) create evolutionary pressure for ToM, or does that pressure arise elsewhere, with language merely benefiting from it much later on?
  - Easy to see how other pressures might motivate some form of ToM, e.g. competition over food resources – I would like to be able to anticipate the movements of others and thus adjust my own...

# Scenario: alarm calls

- Consider a (simplified) putty-nosed monkey (as studied by Zuberbühler et al.)
- Two main alarm calls:
  - “pyow” (ground-based predator)
  - “hack” (air-based predator)
- Trigger responses, but needn’t be intentional in origin, nor under conscious control



# Advanced alarm calls

- Suppose that these calls reach a stage of being conceived of as alarm calls by the callers
  - Don't need communicative intention or ToM: might just have noticed that these behaviours make conspecifics run away, and have an independent interest in causing that
  - Could even have deceptive calls without ToM, as argued for some birds: only need the caller to notice that the call brings about that effect, and to desire that outcome
- Why might ToM help, then?
  - Perhaps I can avoid the risk of making alarm calls if I know that you have already seen the danger
  - I could avoid being conned by deceptive calls, if there are any

# Summary

- ToM extremely useful in complex communicative settings
- Potentially useful in rather simpler settings
- Conceivable that language could promote the development of ToM in evolutionary time, and perhaps vice versa
- However, other aspects of social interaction might have been involved in its emergence

# An open question (I think)

- ToM an advantage, all things being equal
  - Certainly to humans, given our cognitive capacity (although in practice we can't track all that much of others' mental states)
- But an implied trade-off
  - Presumably fully-fledged ToM is cognitively costly
  - Emergence of this capability seems like an implausible evolutionary step (and ToM is itself a remarkable hypothesis for a simple animate entity to form about the world)
  - Seems a bit of a “double circulation” step, considering how resources would have to be divided
- What do the component parts of this look like?