# The Necessity of Joint Attention for Language Evolution

Johan Kwisthout,
supervised by Dr. A. Dijkstra
and Dr. W. Haselager

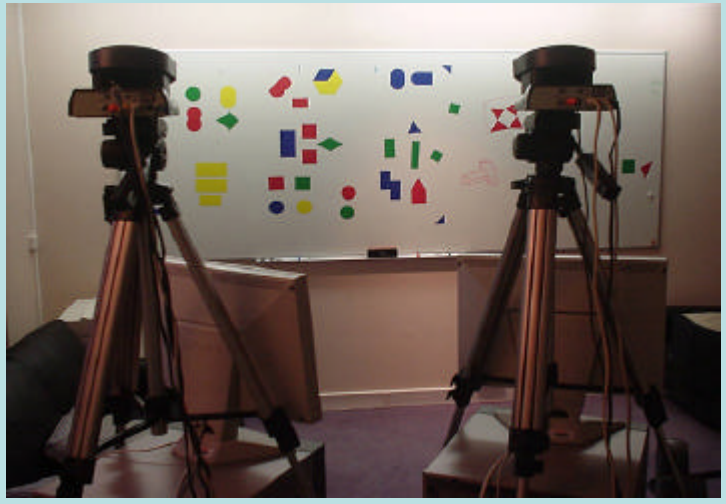
nici
Nijmegen Institute for Cognition and Information
University of Nijmegen

# 1. Introduction

The question how and why language has emerged during the evolution of mankind is, according to Christiansen and Kirby (2003), amongst the ‘hardest problem in science’. By its very nature, this question can only be answered using indirect evidence, which has been provided by scientists from research areas as diverse as archeology, linguistics, philosophy, biology, and mathematics. Recently, researchers in computer science and artificial intelligence joined them, using computational methods to simulate the emerging of language and other cognitive skills.

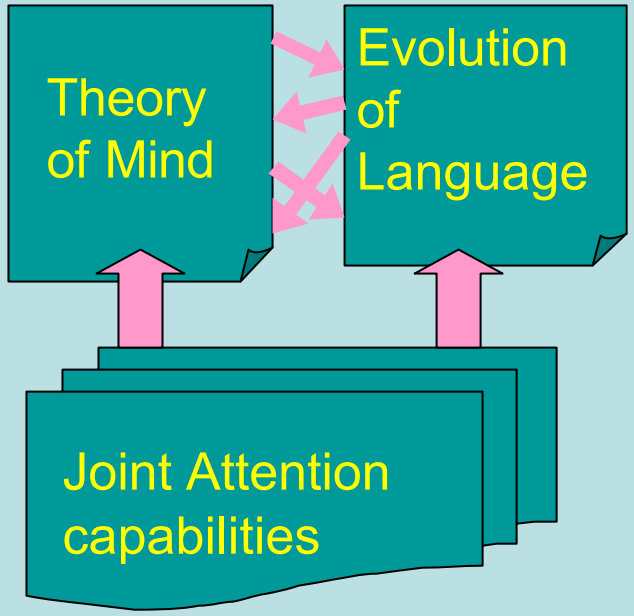
For example, Luc Steels (2001) employed robots to play so-called *language games* in the Talking-Heads experiments. In these language games robots were able to acquire a common lexicon of various properties related to geometrical objects on a whiteboard by using communicative actions, feedback and joint attention, where one of the agents tried to label the features of objects the other was ‘looking’ at.

Vogt and Coumans (2003) used agent simulations to compare various types of language games, which utilized feedback, joint attention, or just associative learning.

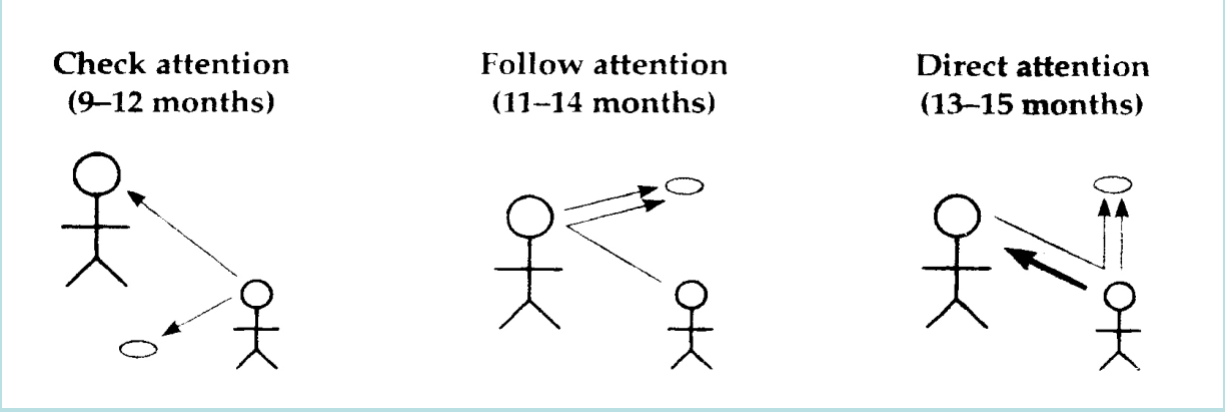


*The Talking Heads experiment*

One important topic in language evolution is the relation between *language*, *Theory of Mind*, and *joint attention* in the evolution of mankind. Malle (2002) and Reboul (2004) proposed an *escalator* model, where joint attentional skills are the precursor for both Theory of Mind and language, which build on each other's advances to further develop.




In this study we investigated whether language games can provide any evidence for the hypothesis that joint attention is a prerequisite for basic language evolution, and that further advances need some Theory of Mind. Furthermore, we investigated *which* joint attentional capability is crucial: the ability to *share*, *follow*, or *direct* attention, since these capabilities are typically acquired at different moments during a child's development and require different cognitive skills (Carpenter, Nagell, & Tomasello, 1998).




*(picture from Tomasello, 1999)*

# 2. Method

In the simulations, we used *guessing games*, *observational games*, and *selfish games* (all described in Vogt & Coumans, 2003), and *enhanced* observational games (which model various stages of joint attentional capabilities) in different settings to investigate the influence of joint attention. While the guessing game had *feedback* as its primary mechanism, the observational game utilized *sharing attention*. The selfish game had neither feedback nor joint attention, and relied on associative learning.

| Game | Joint Attention | Level |
|---|---|---|
| Guessing Game | N/A | N/A |
| Selfish Game | None | 0 |
| Observational Game | Share Attention | 1 |
| OG/follow | Follow Attention | 2 |
| OG/direct | Direct Attention | 3 |
| OG/both | Follow & Direct | 4 |

*Language games and joint attention level*

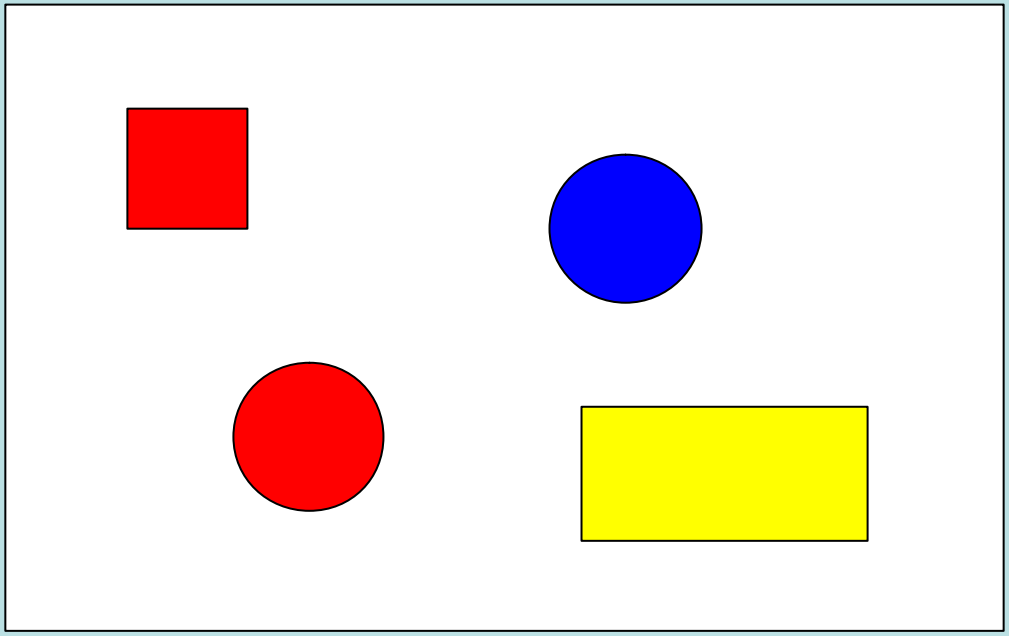
In all games, the agents had to relate utterings to features (like colors), where the association between uttering and features was used as a *weight measure*.

These games were played with varying number of agents, number of features and level of ambiguity of the environment that the agents percepted. After 100 language games, the coherence rate of the evolved lexicon among the agents was calculated and used as an independent variable.

We enhanced the observational game to allow agents to employ additional methods to further specify the feature that was object of the game. For example, if the feature that was meant in the game was ‘red color’, and the object that was communicated was a red circle, the has no further information whether the feature that was meant was **red** or **circle**, and ambiguity would arise (see picture).

If the ‘hearer’ would be able to engage in *follow attention*, it could ask for another object, and, when shown the red square, could conclude the utterance was related to the feature **red**.

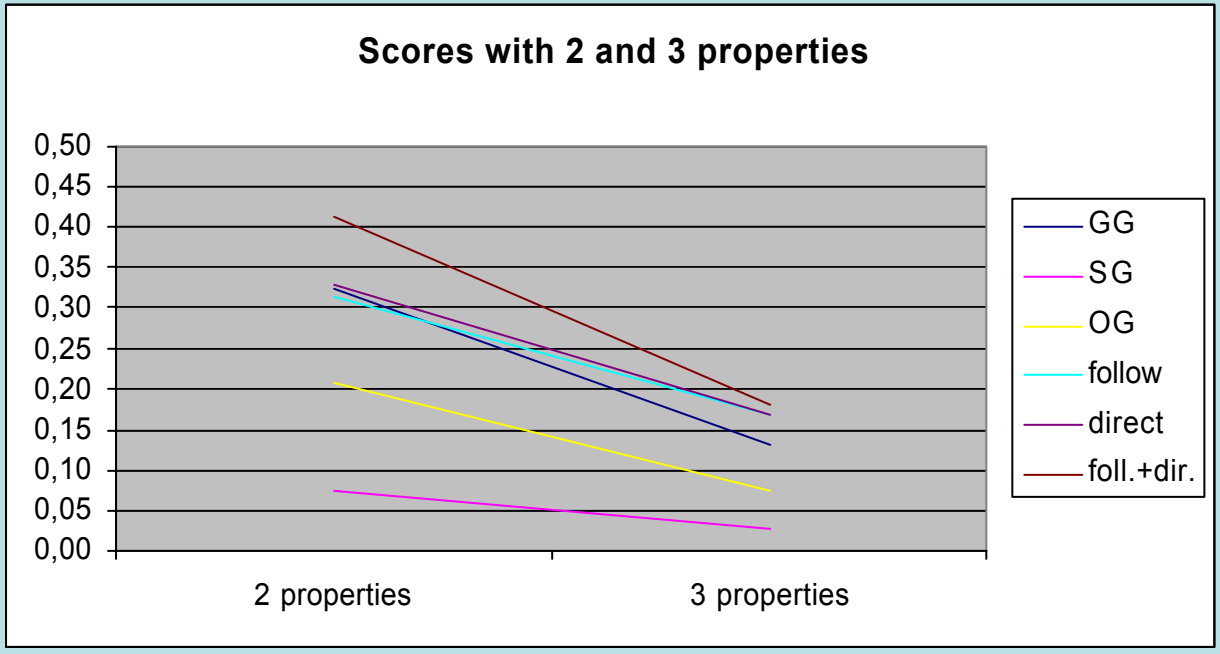
If the ‘hearer’ would be able to employ *direct attention*, it could communicate the blue circle and, upon denial, would be able to relate the utterance to **red** as well. Both techniques could be employed together in even more complicated settings to limit the number of features that are consistent with the agent’s knowledge.



*Geometrical shapes on a whiteboard*

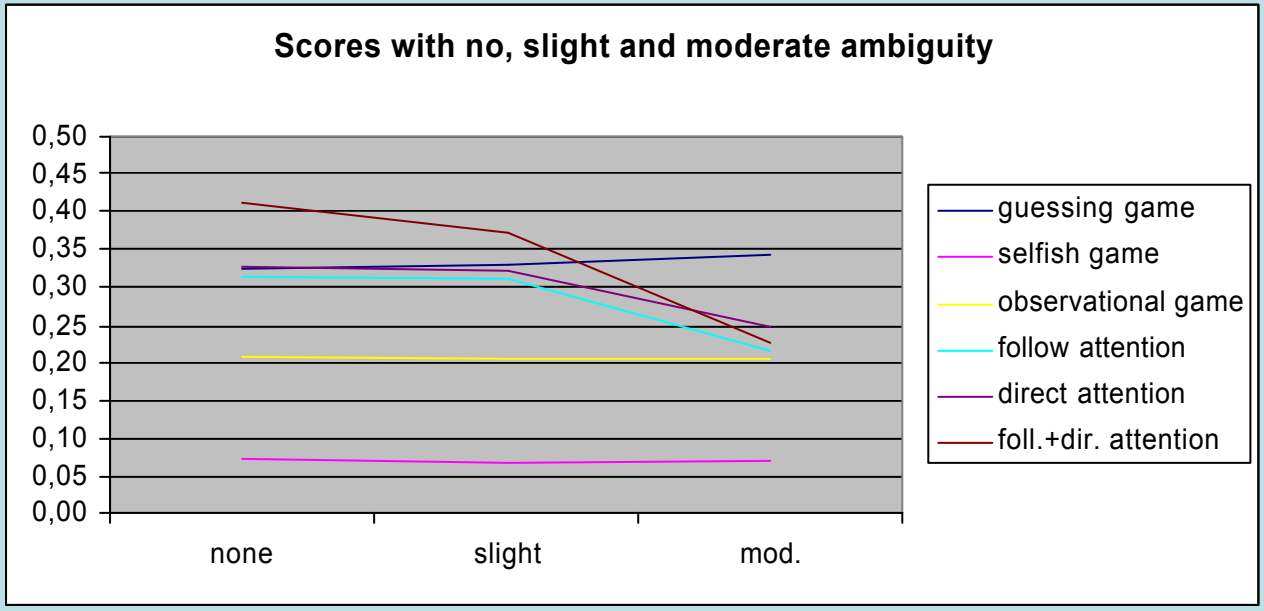
# 3. Results

In simulations with 2 agents and 2 attributes (shape and color) with both 3 values, the combined use of both follow and direct attention (.41) caused higher coherence than feedback (.33), only follow attention (.31) and only direct attention (.33). The traditional observational game gave a result of .20 and the selfish game scored .07. While no significant interaction could be found between game type and number of agents, there was a significant interaction ($F(5, 108) = 4.46$, $p < .001$) between game type and number of properties. The differences in the coherence rate become smaller, notably scores of the guessing game and follow+direct game fall sharply.




| game type | 2 prop. | 3 prop. |
|---|---|---|
| guessing game | 0,33 | 0,13 |
| selfish game | 0,07 | 0,03 |
| observational game | 0,21 | 0,07 |
| follow attention | 0,31 | 0,17 |
| direct attention | 0,33 | 0,17 |
| foll.+dir. Attention | 0,41 | 0,18 |

There is also a significant interaction ($F(10, 162) = 2.31$, $p < .05$) between game type and level of ambiguity. While the coherence rate in the traditional games was hardly influenced by ambiguity in the environment, the joint attentional enhancements dropped to the level of the observational game when there was moderate ambiguity in the environment.




| game type | none | slight | mod. |
|---|---|---|---|
| guessing game | 0,33 | 0,33 | 0,34 |
| selfish game | 0,07 | 0,07 | 0,07 |
| observational game | 0,21 | 0,21 | 0,20 |
| follow attention | 0,31 | 0,31 | 0,22 |
| direct attention | 0,33 | 0,32 | 0,25 |
| foll.+dir. attention | 0,41 | 0,37 | 0,23 |

# 4. Conclusion

Although computational simulation of artificial language grounding is a rather abstract research method that leaves out many – possibly important – details in the simulation model, the use of abstract language games gives researchers the opportunity to formulate and simulate hypotheses that are difficult to test otherwise.

This study is ongoing, but the results so far provide little evidence for the hypothesis that joint attention is a crucial prerequisite for language evolution. While strategies to narrow down the possible alternatives (i.e., to disambigue the topic of the language game) help in very easy games, the use of methods that mimic the capability of *follow* and *direct attention* are less successful when more features play a role, and are more or less useless in ambiguous situations.

From these results, it seems unlikely that the ability to engage in joint attention is (by itself) a crucial factor in the development of language. It seems that a *combination of strategies* (e.g., feedback, imitation, joint attention) is needed to give better results.

Nevertheless, care should be taken to over-generalize results from *artificial language grounding* to *language development* or even *language evolution*.

# References


**Carpenter, M., Nagell, K., & Tomasello, M. (1998).** Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, 63(4).

**Christiansen, M. and Kirby, S. (2003).** Language Evolution. Oxford University Press.

**Malle, B.F. (2002).** The relation between language and theory of mind in development and evolution. In T. Givón & B. F. Malle (Eds.), *The evolution of language out of pre-language*. Amsterdam: Benjamins.

**Reboul, A. (2004).** Evolution of Language from Theory of Mind or Coevolution of Language and Theory of Mind? *Issues in Coevolution of Language and Theory of Mind.*

**Steels, L. (2001).** Language games for autonomous robots. *IEEE Intelligent Systems*, 16(5).

**Tomasello, M. (1999).** The Cultural Origins of Human Cognition. Harvard University Press.

**Vogt, P., and Coumans, H. (2003).** Investigating social interaction strategies for bootstrapping lexicon development. *Journal of Artificial Societies and Social Simulation*, 6(1).



This research was conducted as part of an MSc program at Radboud University Nijmegen.


Contact first author at johank@cs.uu.nl or http://www.cs.uu.nl/staff/johank.html/