Allow me to pick your brain: how language and thought can be shared between robots

Tony Belpaeme

University of Plymouth, UK

tony.belpaeme@plymouth.ac.uk

Language and thought are mutually supportive

In recent years, through experimental studies, it has become clear that language has a definite influence on cognition, and specifically on the use and interpretation of conceptual knowledge (cf. Davidoff, Davies and Roberson 1999; Roberson, Davidoff, Davies and Shapiro 2005; Gilbert, Regier, Kay and Ivry 2006). Many, including me, support a view where language not only has an impact on cognition, but where thought and language are mutually supportive.

The focus of the talk will be the interplay between language and category acquisition. We start from the premise that intelligent systems will need representations which resonate with human representations in order for human-machine interaction to succeed (Steels and Belpaeme 2005; Belpaeme and Cowley 2007). Recent psychological evidence has demonstrated that thought and language are mutually supportive and impact on each other. Category acquisition is facilitated by linguistic interaction, both in infants (Xu 2002; O'Hanlon and Roberson 2007) and adults (Lupyan 2006). This resonates with memetics in the sense that knowledge, present in categories and concepts, is propagated through culture. We believe that cognitive systems should be sensitive to culture and language (with language being the prime medium for memetic propagation) in order to acquire human-like semantic representations.

We will argue that language is not contained within the individual, but is the amalgamated result of producing and interpreting communicative expressions in a population of language users. No single individual has access to the entire language, but has its own interpretation of the population's language. By implication, if we accept that language has an influence on concept formation, concepts are subject to the same constraints: each individual holds an interpretation of a concept. These concepts are private, but are similar enough to concepts of other individuals to allow communication; concepts of individuals are *coordinated* (Steels and Belpaeme 2005). There are several mechanisms through which concepts can become coordinated. Some concepts might be native and as such do not need to be acquired and coordinated. Others, such as perceptual categories, can be acquired by the individual through interacting with the environment. However, acquiring concepts through interaction with the environment does not guarantee the concepts to be sufficiently coordinated to allow communication (for an example on colour see Belpaeme and Bleys 2004). A third mechanism is cultural acquisition of concepts: here concepts are acquired by interaction with one's peers, with language being the most important medium of transmitting conceptual knowledge. It is obvious why this is necessary for abstract concepts, such as DEMOCRACY, but it has been argued that language is also crucial in acquiring perceptual concepts, such as RED, EMPTY or ANGRY. Among others, language helps a developing individual to indirectly access concepts of others — obviously, one does not have direct access to others' concepts. It also allows the learner to access different hypothesis maintained by a range of individuals. This aids learning, and specifically generalisation during learning (cf. machine learning techniques where a large training set supports learning). Finally, language also serves to delimit or constrain conceptual representations, and during concept development linguistic labels act as anchors that prevent concepts from drifting in a semantic sea (cf. Belpaeme and Bleys 2006).

Robots acquiring human concepts

If robots are to interact, communicate and reason with us, they will need conceptual representations which resonate with ours. A straightforward approach is to specify the semantics of internal representation in a machine, but as robots are expected to accrue large amounts of information it will be unlikely that conceptual information can provided through pre-programming semantics.

It has been suggested that a developmental approach could provide a successful approach to achieving artificial intelligence (e.g. Weng, McClelland, Pentland, Sporns, Stockman, Sur and Thelen 2001), and we would like to argue that a developmental trajectory will be needed for robots to acquire grounding language.

When building a robot that acquires the meaning of words, several elements can be shortcut. We are for example not interested in auditory perception or vocal production; instead we will focus on elements that are part of a semiotic schema: perception of external stimuli, mental representations, and the association between mental representations and linguistic labels and structures.

A first requirement is an understanding of learning mechanisms involved. Infants employ a number of biases, constraints or preconceptions which aid in language and meaning acquisition. Several of these have already been uncovered in developmental psychology and linguistics, such as mutual exclusivity constraint (Markman 1989) or novel name-nameless category

(N3C) constraint (Mervis and Bertrand 1994). Some constraints require a deeper understanding to allow an implementation on a robot. Others are not understood well enough: for example, the amount of feedback given by the teacher to the learner. The balance between receiving no feedback, receiving implicit feedback or receiving explicit feedback is still not well understood and will be crucial in constructing artificial systems which acquire semantic through human-robot interaction.

Telepathic robots

While humans are not able to access each others' brains to extract knowledge and have to rely on communication to do so, robots do not have this limitation. A robot could easily request information from others robots through channels other than gestural or linguistic channels. If a robot does not know what "dog" is, it could request a human to explain what dog is or it could request for it to be shown a dog. However, it might be more effective to request the meaning of dog from other robots connected through the internet. Such "telepathic" concept acquisition poses a number of opportunities and challenges.

Concept acquisition could be sped up enormously: while children take years to crystallise certain concepts (for example, colour categories are only a mastered around the age of 3 or 4), a robot could download the meaning of certain words from another robot or from several robots, unconstrained by the noisy media of linguistic communication and perception. A robot which did not know the meaning of "dog" could, without ever being shown a dog, learn what a dog is from another robot that has sensory access to dogs.

However, this poses a number of challenges for the learning algorithms. They should be geared towards acquiring and adapting semantic representations according to linguistic interaction with humans, but at the same time should allow semantic information to be injected from other robots. Another issue concerns embodiment: the conceptual representations of a robot will be tightly coupled to its embodiment, i.e. its sensors and actuators. However, how does one integrate concepts from a robot with a different embodiment? Embodiment can vary radically, such a robot having a different type of camera or a different mode of moving through its environment, but even robots which are identical will have slight variations in embodiment due to noise on sensors and actuators which might have far reaching implications for its semantic concepts.

We wish to study these and related issues in a number of project, most notably the ITALK project sponsored by Europe under the 7th framework programme; of which a brief overview will be given.

References

Belpaeme, T. and Cowley, S. J. 2007. Extended Symbol Grounding. *Interaction studies: Social behaviour and communication in biological and artificial systems*, **8**, (1) 1-6.

O'Hanlon, C. G. and Roberson, D. 2007. What constrains children's learning of novel shape terms? *Journal of Experimental Child Psychology*, **97**, 138-148.

Gilbert, A., Regier, T., Kay, P. and Ivry, R. 2006. Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences*, **103**, (2) 489-494.

Lupyan, G. 2006. Labels facilitate learning of novel categories. In: Cangelosi, A., Smith, A.D.M. and Smith, K. (Eds). *The Evolution of Language.* World Scientific. pp 190-197.

Belpaeme, T. and Bleys, J. 2006. The subtle interplay between language and category acquisition and how it explains the universality of colour categories. In: Cangelosi, A., Smith, A.D.M. and Smith, K. (Eds). *The Evolution of Language.* World Scientific. pp 395-396.

Roberson, D., Davidoff, J., Davies, I. R. L. and Shapiro, L. R. 2005. Color categories: Evidence for the cultural relativity hypothesis. *Cognitive Psychology*, **50**, (4) 378-411.

Steels, L. and Belpaeme, T. 2005. Coordinating Perceptually Grounded Categories through Language. A Case Study for Colour. *Behavioral and Brain Sciences*, **24**, (8) 469-529.

Belpaeme, T. and Bleys, J. 2004. Does structure in the environment influence our conceptualization? Proceedings of the Evolution of Language Conference 2004, Leipzig, Germany.

Xu, F. 2002. The role of language in acquiring object kind concepts in infancy. *Cognition*, **85**, 223-250.

Weng, J. Y., McClelland, J., Pentland, A., Sporns, O., Stockman, I., Sur, M. and Thelen, E. 2001. Artificial intelligence - Autonomous mental development by robots and animals. *Science*, **291**, (5504) 599-600.

Davidoff, J., Davies, I. and Roberson, D. 1999. Colour categories in a stone-age tribe. *Nature*, **398**, 203-204.

Mervis, C. B. and Bertrand, J. 1994. Acquisition of the Novel Name--Nameless Category (N3C) Principle. *Child Development*, **65**, (6) 1646-1662.

Markman, E. M. 1989. *Categorization and naming in children: problems of induction.* The MIT Press, Cambridge, MA.