

THE INFLUENCE OF LANGUAGE ON CONCEPT FORMATION IN ARTIFICIAL AGENTS

JOACHIM DE GREEFF, FRÉDÉRIC DELAUNAY AND TONY BELPAEME

*School of Computing and Mathematics, University of Plymouth,
Plymouth PL4 8AA - United Kingdom
Joachim.deGreeff@plymouth.ac.uk*

For any cognitive system to behave properly within a dynamic, ever changing natural environment, the ability to work with concepts is quite important. Humans display an ability to talk about, reason over, infer and generate new concepts with remarkable ease, but for artificial agents this is far from trivial. In the case of human concept learning, it has been acknowledged by several authors that the formation of new concepts is heavily influenced by language. As such, the words used to describe a stimulus govern the way in which this data will be integrated into existing conceptual structures.

Several studies have shown that young children, in addition to learning directly from sensory exploration, rely on linguistic labels to acquire new concepts. Xu (Xu, 2002), for example, demonstrated how linguistic labels help 9-month old infants to establish a representation for different objects. Learning without linguistic labels, or with the presence of tones, sounds or emotional expressions is not as effective. Plunkett (Plunkett, Hu, & Cohen, 2008) came to the same conclusion in a controlled experiment in which they demonstrated how category formation in 10-month old infants is influenced by linguistic labels. Linguistic labels also have an effect on category learning in adults; adults who learn a new category did so significantly faster and showed more robust category recall when the learning experience was accompanied by novel linguistic labels (Lupyan, Rakison, & McClelland, 2007). This shows that linguistic labels facilitate category acquisition, both in pre-linguistic infants and adults. These insights tie in with linguistic relativism, which gained renewed attention as a series of experiments demonstrated how perception of stimuli and use of categories is influenced by language (Gilbert, Regier, Kay, & Ivry, 2006) (Majid, Bowerman, Kita, Haun, & Levinson, 2004).

In order to capture these insights, we developed a computational model in which a learning agent is able to learn new concepts through linguistic interaction with a teacher. To do so, we adapted interaction based on Language Games (Steels & Belpaeme, 2005) to a teacher-learner scenario. This allows for the usage of language as a steering mechanism in the acquisition of conceptual knowledge and as-

sociated meaning. A learning agent engages into a series of Language Games with a teacher, and thus gradually builds a repertoire of word-meaning mappings. To represent conceptual knowledge, we use a Conceptual Space (Gärdenfors, 2000), which consists of a geometrical representation within a number of quality dimensions with a metric, allowing for similarity measurement. Within a Conceptual Space, concepts can be stored as prototypes and associated with a lexicon of word labels. Newly perceived stimuli can be matched to existing conceptual prototypes and an appropriate linguistic expression can be found.

We propose that the model exhibits properties comparable to how young children learn new concepts; namely *language driven acquisition*, *fast mapping* and *overgeneralization*. The model is based on previous work as reported in (Greeff, Delaunay, & Belpaeme, 2009), in which we studied the effect of adding interactive features to the learning process. We then augmented the model with a Spreading Activation layer (Rumelhart, McClelland, & PDP Research Group, 1986) which allows for association between concepts, even when they are not perceptually similar. Typically the colour domain is used as a test case, but the model characteristics are general enough to be applied in any domain. We argue that our model is a feasible way of acquiring conceptual knowledge in a linguistic relativism spirit.

References

- Gärdenfors, P. (2000). *Conceptual spaces: The geometry of thought*. Cambridge, MA: the MIT Press.
- Gilbert, A., Regier, T., Kay, P., & Ivry, R. (2006). Whorf hypothesis is supported in the right visual field but not the left. In *Proceedings of the national academy of sciences* (p. 489-494).
- Greeff, J. de, Delaunay, F., & Belpaeme, T. (2009). Human-robot interaction in concept acquisition: a computational model. *ICDL, 0*, 1-6.
- Lupyan, G., Rakison, D. H., & McClelland, J. L. (2007). Language is not just for talking - redundant labels facilitate learning of novel categories. *Psychological Science, 18*(12), 1077-1083.
- Majid, A., Bowerman, M., Kita, S., Haun, D. B. M., & Levinson, S. C. (2004). Can language restructure cognition? the case for space. *Trends in Cognitive Sciences, 8*(3), 108 - 114.
- Plunkett, K., Hu, J., & Cohen, L. B. (2008). Labels can override perceptual categories in early infancy. *Cognition, 106*(2), 665-681.
- Rumelhart, D. E., McClelland, J. L., & PDP Research Group the. (1986). *Parallel distributed processing, volumes 1 and 2*. Cambridge: MIT Press.
- Steels, L., & Belpaeme, T. (2005). Coordinating perceptually grounded categories through language: A case study for colour. *Behavioral and Brain Sciences, 28*(4), 469-89. (Target Paper, discussion 489-529)
- Xu, F. (2002). The role of language in acquiring object kind concepts in infancy. *Cognition, 85*(3), 223 - 250.