Social Interaction for Acquiring Word Meaning and Concepts of Negation in Robots

Joe Saunders, Frank Förster, Chrystopher L. Nehaniv
Adaptive Systems Research Group
University of Hertfordshire, College Lane, Hatfield, Herts, AL10 9AB, UK
Email: {j.1.saunders f.forster c.l.nehaniv }@herts.ac.uk

1 Extended Abstract

In this paper we will demonstrate how social behaviour and social cues derived from a human interacting with a humanoid robot can be used to allow the robot to learn labels of presented shapes. This work is based on the premise that social interaction is essential for effective language acquisition [6].

In human development, adults do not specifically attempt to 'teach' language to children but rather aim for joint understanding with the child. The interaction mechanism employed (often unconsciously) by the adult is to use a form of language called Infant Directed Speech (IDS) which is specifically tailored to the perceived level of linguistic skill of the young child [2]. IDS is characterised by adults speaking more slowly and often repeating what is said during the interaction. Attention is also drawn to topics via changes in pitch, pause duration, word placement, word duration and intonation. As joint understanding increases between the child and adult, fewer aspects of IDS are employed.

We report firstly, on a study with a humanoid robot (KasparII) where, through unrestricted speech and interaction with a human partner, the robot can attach meaning[1] to a series of individual shapes by exploiting some elements of IDS (see figure 1). Secondly, we will report on proposed developmental work with a humanoid robot (iCub) on foundational considerations for studies into the acquisition of human-like use and understanding of negation (see figure 2).

We will explain how human infants are thought to acquire word meaning and present results from the first study above to assess whether broadly equivalent mechanisms will allow a humanoid robot to do the same. This is achieved by firstly employing a rudimentary form of shared intentional reference and secondly making the assumption that regularities can be extracted from the human's speech stream and associating these with the robot's sensorimotor perceptions. In English (and many other languages) attention is drawn to salient words with prosodic markers and word placement. Often salient words are placed at the end of an utterance, partly reflecting the subject-verb-object structure of the language) [2, pages 62-66]. However there is also evidence that a similar phenomenon occurs in IDS in other languages (e.g. Turkish) even when to do so would appear to be ungrammatical for that language [1]. As well as utterance-final word placement, often the word is spoken with a longer duration than average during the utterance. We therefore use both the duration and placement of a word in the perceived speech matched with the sensorimotor stream experienced by the robot in order to bias the robot's learning ability. As well as being attentive to word markers, the robot must also be able to associate that word with important aspects of the sensorimotor perception and proprioceptions and associate them, even though they may not occur at the same time during the interaction. We will explain a method dynamically associating the relevant part of the sensorimotor stream with the referential focus of the robot.

We will also report on current work studying negation and its application in humanoid robotics. Negation has been considered primarily as a grammatical phenomenon [4]. However, literature on child language development shows that negation is much more varied and begins to emerge long before the production of grammatical utterances [2, pages 214-217]. The target of the work reported here is to modify an existing embodied language acquisition system in order to enable it to aquire negative utterances via unconstrained human-robot interactions. It is based on an analysis of early language development

[1] Following Wittgenstein [7], any derivation of 'meaning' must ultimately be evidenced by appropriate embodied action in language games (cf. [3]).

Figure 1: ***Sharing Reference with a Teacher in Context.*** *Kaspar2 interacts with the teacher, where both the robot and the human share attention on the shapes on the box, a case of rudimentary shared intentional reference.*

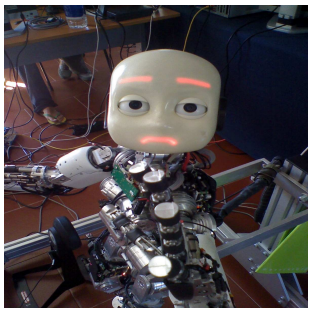

Figure 2: ***A Grumpy iCub.*** *We consider an alternative view of negation. In our research language is used as a means to manipulate the world rather than as a paradigm of propositional representation. We consider both affect and the need for memory.*

in humans [5]. We will present the varied classes of negation and show that three of five negation types are related to motivation/affect where their meaning is virtually void if motivation is disregarded. In addition we will demonstrate that certain types of negation suggest the need for memory and temporally extended representation as they refer to events in the past or to absent entities. The methodology being used to carry out this research will make use of the humanoid robot iCub, capable of facial expressions and gestures. An existing social learning architecture will be extended by adding short-term memory with a varying time horizon to language extraction as well as adding a motivation module in order to support sensorimotor-motivational grounding. We will use facial expression and gestures to indicate motivational state of the humanoid in order to provoke negative utterances from naïve human partners using unconstrained speech.

In both of the the studies above we will consider how to recognize important events in the social interactions, what the interaction between the human and robot yields in terms of meaning acquisition, and the limitations on these forms of social interaction.

Acknowledgment

The work described here was conducted within the EU Integrated Project ITalk ("Integration and Transfer of Action and Language in Robots") funded by the European Commission under contract number FP7-214668.

References

[1] R. N. Aslin, J. Z. Woodward, N. P. LaMendola, and T. G. Bever. Models of word segmentation in fluent maternal speech to infants. In J. Morgan and K. Demuth, editors, *Signal to Syntax*. Lawrence Erlbaum, 1996.

[2] E. V. Clark. *First Language Acquisition*. Cambridge University Press, Cambridge,UK, 2nd edition, 2009.

[3] C. L. Nehaniv. Meaning for observers and agents. In *IEEE International Symposium on Intelligent Control/Intelligent Systems and Semiotics (ISIC/ISAS'99)*, pages 435–440, 1999.

[4] C. L. Nehaniv, C. Lyon, and A. Cangelosi. Current Work and Open Problems: A Road-Map for Research into the Emergence of Communication and Language. In C. Lyon, C. L. Nehaniv, and A. Cangelosi, editors, *Emergence of Communication and Language*, pages 1–27. Springer, 2007.

[5] R. Pea. The development of negation in early child language. In D. Olson, editor, *The Social Foundations of Language Thought: Essays in Honor of Jerome S. Bruner*. W.W. Norton, 1980.

[6] J. Sachs, B. Bard, and M. L. Johnson. Language learning with restricted input: Case studies of two hearing children of deaf parents. *Applied Psycholinquistics*, 1:34–54, 1981.

[7] L. Wittgenstein. *Philosophical Investigations (Philosophische Untersuchungen)* – German with English translation by G.E.M. Anscombe. Basil Blackwell, 3rd edition, 1968. (first published 1953).