ARTICLE IN PRESS
Neural Networks ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

Introduction

Multimodal communication in animals, humans and robots: An introduction to perspectives in brain-inspired informatics

S. Wermter [a,*], M. Page [a], M. Knowles [a], V. Gallese [b], F. Pulvermüller [c], J. Taylor [d]

[a] Centre for Hybrid Intelligent Systems, Department of Computing, Engineering and Technology, University of Sunderland, St Peter's Way, Sunderland SR6 0DD, UK[1]
[b] Department of Neuroscience, Section of Physiology, University of Parma, 43100 Parma, Italy
[c] MRC Cognition & Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF, UK
[d] Department of Mathematics, Kings College, London, WC2R 2LS, UK

ARTICLE INFO

Keywords:
Multimodal communication
Neural networks
Robotics
Brain-inspired computing

ABSTRACT

Recent years have seen convergence in research on brain mechanisms and neurocomputational approaches, culminating in the creation of a new generation of robots whose artificial "brains" respect neuroscience principles and whose "cognitive" systems venture into higher cognitive domains such as planning and action sequencing, complex object and concept processing, and language. The present article gives an overview of selected projects in this general multidisciplinary field.

The work reviewed centres on research funded by the EU in the context of the New and Emergent Science and Technology, NEST, funding scheme highlighting the topic "What it means to be human". Examples of such projects include learning by imitation (Edici project), examining the origin of human rule-based reasoning (Far), studying the neural origins of language (Neurocom), exploring the evolutionary origins of the human mind (Pkb140404), researching into verbal and non-verbal communication (Refcom), using and interpreting signs (Sedsu), characterising human language by structural complexity (Chlasc), and representing abstract concepts (Abstract).

Each of the communication-centred research projects revealed individual insights; however, there had been little overall analysis of results and hypotheses. In the Specific Support Action Nestcom, we proposed to analyse some NEST projects focusing on the central question "What it means to communicate" and to review, understand and integrate the results of previous communication-related research, in order to develop and communicate multimodal experimental hypotheses for investigation by future projects. The present special issue includes a range of papers on the interplay between neuroinformatics, brain science and robotics in the general area of higher cognitive functions and multimodal communication. These papers extend talks given at the NESTCOM workshops, at ICANN (http://www.his.sunderland.ac.uk/nestcom/workshop/icann.html) in Porto and at the first meeting of the Federation of the European Societies of Neuropsychology in Edinburgh in 2008 (http://www.his.sunderland.ac.uk/nestcom/workshop/esn.html). We hope that the collection will give a vivid insight into current trends in the field.

1. Communication and neural organisation

Humans use different types of tools, or modalities, to communicate in their everyday lives, including spoken and written language, sign language, body gestures, facial expressions and also computational interfaces. For humans, the means to interact with each other is of vital importance in a social environment, but, communication does not occur solely amongst humans since animals and robots are also able to communicate using various modalities (Iba, Paredis, & Khosla, 2006; Rybski, Yoon, Stolarz, & Veloso, 2007; Spiliotopoulos, Andreoutsopoulos, & Spyropoulos, 2001).

The range of potential benefits from understanding how we communicate is wide. There are educational benefits from some projects of the EU NEST initiative, for example the Analogy, Calacei and Edici projects. Some work offers benefits for the cognitively impaired such as the CHLaSC (2008) and Edici Project (2008). Other projects offer the opportunity to improve the way we interact with technology, such as the Wayfinding Project (2008) which aims to produce better navigational aids. During this Nestcom project we did not aim at covering all aspects of communication,

* Corresponding author. Tel.: +44 191 515 3279; fax: +44 191 515 3553.
E-mail addresses: stefan.wermter@sunderland.ac.uk (S. Wermter), michael.knowles@sunderland.ac.uk (M. Knowles).
1 URL: http://www.his.sunderland.ac.uk.

0893-6080/$ – see front matter © 2009 Published by Elsevier Ltd
doi:10.1016/j.neunet.2009.01.004

however we have selected some representative research which addresses the underlying principles of what it means to be human and what it means to communicate where language certainly plays an important role. In the following sections, we will summarise some of the NEST project research.

A selection of NEST projects have researched the structure of the brain as related to communication, considering the operation of the brain at a regional level and the structures and arrangements at the individual neuron level. Recent research has moved our understanding of the regional division of cognitive and communicative functions. The brain is now considered to be far less modular than it once was. Neuroscience principles operative in neuroanatomically structured network simulations can now be used to explain how local modules in specific cortical areas bind into distributed neuronal circuits that give rise to action–perception association and mirror neuron activity (Knoblauch & Palm, 2002; Wennekers, Garagnani, & Pulvermüller, 2006). These action–perception circuits may be of utmost importance for language and communication (Garagnani, Wennekers, & Pulvermüller, 2008; Pulvermüller, 1999; Wennekers et al., 2006). We find an example of this in the processing of language. For example action words such as 'kick' fire neurons in the part of the brain that controls the motor action associated with the word i.e. the foot in this case (Pulvermüller, 2005). However, the structure in the brain is far from random in its organisation and there is some evidence for an underlying language system with different 'modules' performing various functions (Friederici, 2002).

Several NEST projects' studies have isolated particular brain regions which are responsible for certain cognitive traits. The Far Project (2008) has investigated rule-based learning. It established that the frontal lobe and basal ganglia are responsible for rule-based learning, while similarity-based learning occurs in frontal and temporal areas (Opitz & Friederici, 2004). The project also worked on developing a deeper understanding of these modes of learning. The Neurocom project is aimed at mapping the neural substrate of the brain responsible for speech, gestures and species specific communicative articulations (e.g., monkey calls), with the aim of developing new knowledge of the functional architecture of the brain. It is already known that the Superior Temporal Sulcus (STS) contains voice-selective areas and is generally more active in response to speech sounds than non-speech stimuli (Belin, Zatorre, Lafaille, Ahad, & Pike, 2000), and that certain regions of the STS react more strongly to familiar voices than to those of strangers (Kriegstein & Giraud, 2004). The Neurocom project also investigated the Visual Word Form Area (VWFA) which is an extended strip of the left fusiform gyrus. This region is activated by visual words but only when they are presented in a written context and not when they are spoken (Cohen, Lehericy, Chochon, Rivaud, & Dehaene, 2002; Dehaene et al., 2004). Some researchers have denied a general role of this area in word form processing and argued that "VWFA" may contribute to semantic processing (Price & Devlin, 2003).

The Calacei project studied areas of the brain which are key to language acquisition and their development in infants. The Wayfinding Project (2008) also includes objectives for isolating the regions of the brain concerned with spatial cognition. Their research has also investigated the suggestion that spatial processing occurs in the hippocampus and parahippocampus, fronto-parietal cortex, anterior cingulate, and the vestibular system (Van Asselen, 2005).

2. Multimodal neural communication

The role of the superior colliculus has been the subject of recent research since it is the first midbrain area where multimodal neural integration occurs. Research in this area has suggested that it is responsible for multisensory integration of the senses, specifically with respect to orientation and attention (Stein & Meredith, 1993; Sullivan & Konishi, 1986). Besides the midbrain and superior colliculus, recent research into cortical neural structures within the brain has uncovered the importance of networks of the so-called mirror neurons. The mirror neuron system is discussed in the work by Gallese (2005), which describes the firing of mouth-related mirror neurons in monkeys. Witnessing actions which belong to or are similar to the observers' motor repertoire, in this case monkey-style lip-smacking and biting, causes firing in the motor regions of the brain whereas actions unrelated to this repertoire lead to firing in visual areas of the brain. Mirror neurons and the capacity for imitation seem to play a part in the learning of language (Rizzolatti & Craighero, 2004) and in the neural coding of abstract action concepts (Gallese & Lakoff, 2005; Kohler et al., 2002). Mirror neurons seem to play a part in the human capacity to understand or develop an insight into the goals or motivations when observing others performing a task, rather than passively watching (Gallese, Keysers, & Rizzollati, 2004).

This capacity for action understanding and even empathy on a higher cognitive level involving mirror neurons may offer an insight into what distinguishes humans from other species. The overall goal of the Mirrorbot Project (2008) was the development of biomimetic multimodal learning and language instruction in a robot. Based on cognitive neuroscience evidence of cell assemblies and mirror neurons, neural architectures were designed and trained to perform basic actions. Neuroscience experiments were also carried out to aid understanding on how the brain processes words and how the mirror neuron system aids the recognition and understanding of actions (Wermter & Elshaw, 2003; Wermter, Weber, Elshaw, Gallese, & Pulvermüller, 2005a, 2005b). The Edici Project (2008) was focused on studying imitation, including neurological effects of perceptual-motor experience and a review of comparative neuron-anatomy (Gillmeister et al., 2008; Gillmeister, Catmur, Liepelt, Brass, & Heyes, 2008).

It is important to recognise that communication does not solely occur over a single modality, but several modalities are often interlinked to form the communication as a holistic entity. This is known as cross-modal or multimodal communication. Links also exist between communicative areas and other systems. Pulvemüller's group were among the first to document the tight functional links between language and the motor system, especially in the case of action-related verbs (Hauk, Johnsrude, & Pulvermüller, 2004; Preissl, Pulvermüller, Lutzenberger, & Birbaumer, 1995; Pulvermüller, 2001, 2005; Pulvermüller, Lutzenberger, & Preissl, 1999; Pulvermüller, Preissl, Lutzenberger, & Birbaumer, 1996; Shtyrov, Hauk, & Pulvermüller, 2004). Giovanni Buccino [Nestcom, Mirrorbot Project (2008)] confirmed links between language and motor regions of the brain, with certain 'action' words modulating areas of the brain concerned with performing those actions (Buccino et al., 2005). Hauk, Shtyrov, and Pulvermüller (2006) extended this work further by subjecting subjects to the sounds of finger clicking and tongue movements and similarly observing a response in related regions of the brain.

A computational architecture which simulates the neural integration of language, vision and motor control is described in Wermter et al. (2004). The Mirrorbot Project (2008) has explored applying a mirror neuron system to develop a biomimetic multimodal learning and language instruction system, using a robotic platform (Elshaw, Weber, Zochios, & Wermter, 2004). Neural architectures were designed and trained to perform basic actions, with neuroscience experiments performed to aid understanding on how the brain processes words and how the mirror neuron system aids the recognition and understanding of actions.

Please cite this article in press as: Wermter, S., et al. Multimodal communication in animals, humans and robots: An introduction to perspectives in brain-inspired informatics. Neural Networks (2009), doi:10.1016/j.neunet.2009.01.004

Links have also been established between communicative processes and other cognitive areas. For instance, the Wayfinding Project (2008) examined the links between linguistic communication and spatial cognition (Noordzij, 2004). Perhaps the most widespread example of cross-modal communication is the link between speech and gesture.

3. Verbal communication

The ability of humans to develop language skills is one of the foremost aspects of distinguishing humans from non-humans. The development of language in children is one of the key stages in infant development. Some research groups compared pre-linguistic infants with non-linguistic species, emphasising the importance of language skills in human development. One of the key aims of the Far *(From Associations to rules in the development of Concepts)* project was to establish whether the use of language is responsible for differences between humans and non-humans. In order to achieve their aim, they investigated the differences between adult humans and other animals, including the ability to use language to communicate, the use of logic and mathematics to reason and the ability to abstract relations that go beyond perceptual similarity. The Far Project (2008) wanted to answer the question, "*When* and *how* does this rule-based system come into being?" To analyse this problem, they proposed to study the transition from associative to rule-based cognition within the domain of concept learning. Concepts are said to be the primary cognitive means by which we organise things in the world. It was thought that the differences in the way that concepts are formed may go a long way in explaining the greater evolutionary success that some species have had over others. The relationship between associative learning and the rule concept is also discussed in the paper by Pulvermüller and Knoblauch (in this issue).

Other NEST research into the development of language skills includes the Calacei (*Universal and Specific Properties of a uniquely Human Competence*) project, which investigated the cerebral processes involved in speech, and investigating the mechanisms involved in infants acquiring semantic and syntactic elements of a language (Friederici, Friedrich, & Christophe, 2007; Gervain, Macagno, Cogoi, Peña, & Mehler, 2008). Another relevant project is the Edici (*Evolution, development and intentional control of imitation*) project, which investigates the hypothesis that the human capacity for imitation provides the foundation for language acquisition, skill learning, socialisation and enculturation. It is thought that imitation is innate, however the Edici Project (2008) argued that this previous conception should be replaced with a model of imitation that instead should incorporate evolutionary, developmental and cultural inputs (Press, Cook, Dickinson, & Heyes, 2008; Voelkl & Huber, 2007).

4. Visual communication

Visual information plays a crucial rule in our ability to communicate and a vast amount of information about our environment can be seen. The use of visual information in communication enables us to add further knowledge to our current understanding of a situation, such as interpreting signals from others or the emotional state of the person involved in a conversation. During a conversation a person typically generates visual information such as facial expressions and gestures/body language. Visual communication can also include written language, sign language and image-based communication. One of the key benefits of communicating information visually is the ease of conveying spatial information. One of the early objectives of the Wayfinding (*Finding your way in the world — on the neurocognitive basis of spatial memory and orientation in humans*) project was to study the development of the capacity for mental imagery and its formation in infants and how this interacts with spatial cognition. The ability to form mental images is considered by this project team to be uniquely human (Salvesen, 1965).

The amount of information that is obtained and inferred by the human vision system and related cognitive processes about the world around us is vast, and could not be easily transmitted by any other means. The Far Project (2008) has already discovered that children as young as 8 years can form categories of objects based on abstract spatial rules. Another advantage of vision as a mode of receiving information is that it is possible to acquire many different 'channels' of information. Examples for such channels are colour and form. The Mirrorbot Project (2008) focused on the processing of colour as well as form information, studying brain activity related to the processing of colour and form concepts and found that these concepts are handled in a similar manner to certain abstract concepts and are linked to certain semantic models and motor responses (Fay, Kaufmann, Knoblauch, Markert, & Palm, 2005; Moscoso Del Prado Martin, Hauk, & Pulvermüller, 2006; Pulvermüller & Hauk, 2006).

Another form of visual information is referential communication, investigated by the Refcom (*Origins of Referential Communication*) project. This project tried to understand the origins of referential communication, by observational and experimental studies in the wild and in captivity. The main endeavour of the Refcom Project (2008) is to produce a model of how and why a human's uniqueness of referential communication evolved. Previous research had already indicated that other species have also been able to evolve referential signals. The devised experiments involved observational and experimental studies in order to gather evidence on referential signals in primates and non-primates. They explored the semantic flexibility of referential calls in monkeys, dolphins, and parrots, and the use of referential calls and referential gestures in great apes and canids.

Human beings use gestures in a variety of ways. Sometimes this is deliberate, particularly if there is some form of spatial information such as a direction to be conveyed. Sometimes the gesture is less deliberate, such as a shrug of the shoulders or a nod of the head to convey uncertainty or an affirmative, or in some cases, a negative response respectively. Research on the Edici Project (2008) into communicative referential cues and language acquisition in infants suggests that gesture is a key facet. This is also supported by the presence of links between the neural regions involved in language and motor regions (Buccino et al., 2005) and work performed as part of the HandtoMouth project investigating the part that gesture played in the evolution of language.

5. Papers in this special issue

Following this context of some Nest projects we will now give an overview of the papers in this special issue. The first paper is "Emergence of structured interactions: from a theoretical model to pragmatic robotics" by Revel and Andry (in this issue). The authors consider neural architectures for socially interacting robots and propose two specific architectures for 'turn-taking' and 'synchrony'. It is argued that such a system will help facilitate the design of social abilities in robots, enabling intuitive human–robot interactions, and also robot–robot interactions.

Alexandre addresses the "Cortical basis of communication: local computation, coordination, attention". The properties of information processing in the repeated circuits in the cortex are considered and a generic architecture is described in the role of directing visual attention in a communicative scenario. Alexandre argues that implementing the proposed architecture could result in an artificial agent able to exhibit some fundamental communication capabilities as observed in an evolved living being.

Please cite this article in press as: Wermter, S., et al. Multimodal communication in animals, humans and robots: An introduction to perspectives in brain-inspired informatics. Neural Networks (2009), doi:10.1016/j.neunet.2009.01.004

The integration of sensory modalities is discussed by Markert, Kaufmann, Kara Kayikci, and Palm (in this issue) in "Neural associative memories for the integration of language, vision and action in an autonomous agent". The role of visual attention is addressed, in connection with visual object recognition and language understanding for action selection using associative memories. The produced model is able to deal with ambiguities and it is implemented on a robot to demonstrate the ability of semantic understanding through the robot performing corresponding actions.

Pitti, Alirezaei and Kuniyoshi also consider multiple modalities, in this case the integration between vision and somatosensory maps in their article "Cross-modal and scale-free action representations through enaction". Described is a neural network for action representation and recognition which uses spiking neurons to replicate some of the properties of the mirror neuron system. The ability to grasp is represented through multimodal integration of vision and somatosensroy maps and the system is also able to respond to observed actions performed by another person.

Representations are the central theme for "Implicit and Explicit Representations" by Rougier. The symbol grounding problem is addressed and separate models for implicit and explicit representations are presented with visual attention again being a major theme through a study of the emergence of visual attention in the presented models. It is claimed that the models described are able to manipulate represented information in a scenario, without having explicit knowledge of the underlying nature of the information.

Pulvermüller and Knoblauch (in this issue) describe under which conditions discrete rule representations emerge in brain-inspired neural network models. Networks include ample auto and heterossociative links and pre-established sequence detector units as they are found in animal brains. The sequence detectors would be sensitive to elementary word sequences. Given these conditions are met, aggregation of sequence detectors driven by biologically realistic Hebbian associative learning is a consequence of the combinatorial information immanent to word stings in text corpora. Pulvermüller and Knoblauch argue that such sequence detector aggregates may carry the function of combinatorial rules as they are relevant for certain types of syntactic binding. The neuronal assemblies implementing rules through the aggregation of sequence detectors are discussed and evaluated in a language processing scenario.

Murray, Erwin, and Wermter (in this issue) in their paper "Robotic Sound Source Localisation Architecture using Cross-Correlation and Recurrent Neural Networks" present a robust and accurate sound source model for localising and tracking an acoustic source of interest in acoustically cluttered environments, which is to be applied to a mobile service robot. The implemented model has a hybrid architecture, applying cross-correlation and recurrent neural networks; the system was developed based on inspiration from the central auditory system (CAS) of the mammalian brain. Not only does the hybrid architecture perform well in restricted testing environments, but also in tougher real-world conditions.

Gómez (2007), in his paper "Embodying meaning: insights from primates, autism, and Brentano", discusses how meaning can be extremely complex and have many different notions. It is described how there can be simple instances of meaning, such as "clouds mean rain" that are based on the recognition of natural physical connections. There can also be more complex instances, for example, where a person could be implying the same meaning by using a pointing gesture towards a group of dark clouds about to pass overhead. Not only could this refer to the possibility of rain, but have further implications such as the cancelling of going outdoors for a particular activity.

6. Conclusions

We hope that this special issue may inspire readers to perform further interdisciplinary research on multimodal communication based on neural substrates. Many projects have demonstrated the benefits of using biologically inspired techniques both in low level processing of sensory data and high level cognitive processing, see for example Wermter, Palm, and Elshaw (2005). We would also like to thank all the authors who have provided us with exciting and forefront contributions to this special issue and those who have contributed to the Nestcom project.

References

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, *403*(6767), 309–312.

Buccino, G., Riggio, T. L., Melli, G., Binkofski, F., Gallese, V., & Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: A combined TMS and behavioral study. *Cognitive Brain Research*, *24*(3), 355–363.

Cohen, L., Lehericy, S., Chochon, F., Rivaud, S., & Dehaene, S. (2002). The visual word form area: Spatial and temporal characterisation of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*, (123), 291–307.

CHLaSC. Online Available at: http://www.zas.gwz-berlin.de/chlasc/index.html Accessed 20.10.08.

Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, JB., Le Bihan, D., et al. (2004). Letter binding and invariant recognition of masked words: Behavioural and neuroimaging evidence. *Psychological Science*, *15*(5), 307–313.

Edici Project. Online Available at: http://www.univie.ac.at/edici/ Accessed 20.10.08.

Elshaw, M., Weber, C., Zochios, A., & Wermter, S. (2004). An associator network approach to robot learning by imitation through vision, motor control and language. In *Proceedings of the international joint conference on neural networks*, (pp. 591–596).

Far Project. Online Available at: http://www.cbcd.bbk.ac.uk/research/projects/far/FAR Accessed 20.10.08.

Fay, R., Kaufmann, U., Knoblauch, A., Markert, H., & Palm, G. (2005). Combining visual attention, object recognition and associative information processing in a neurobiotic system. In S. Wermter, G. Palm, & M. Elshaw (Eds.), *Biomimetic neural learning for intelligent robots* (pp. 118–143). Heidelberg, Germany: Springer Verlag.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*(2), 78–84.

Friederici, A. D., Friedrich, M., & Christophe, A. (2007). Brain responses in 4-month-old infants are already language specific. *Current Biology*, *17*(14), 1208–1211.

Gallese, V., Keysers, C., & Rizzollati, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Science*, *8*(9), 396–403.

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in reason and language. *Cognitive Neuropsychology*, *22*, 455–479.

Gallese, V. (2005). The intentional attunement hypothesis—the mirror neuron system and its role in interpersonal relations. In S. Wermter, G. Palm, & M. Elshaw (Eds.), *Biomimetic neural learning for intelligent robots* (pp. 19–30). Berlin: Springer.

Garagnani, M., Wennekers, T., & Pulvermüller, F. (2008). A neuroanatomically-grounded Hebbian learning model of attention-language interactions in the human brain. *European Journal of Neuroscience*, *27*(2), 492–513.

Gervain, J., Macagno, F., Cogoi, S., Peña, M., & Mehler, J. (2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences*, *105*(37), 14222–14227.

Gillmeister, H., Catmur, C., Bird, G., Liepelt, R., Brass, M., & Heyes, C. (2008). Body part priming in action imitation and the mirror neuron system depends on correlated sensorimotor experience. In *Poster presented at the Evolution, Development and Intentional Control of Imitation Workshop*.

Gillmeister, H., Catmur, C., Liepelt, R., Brass, M., & Heyes, C. M. (2008). Experience-based priming of body parts: A study of action imitation. *Brain Research*, *1217C*, 157–170.

Gómez, J. C. (2007). Pointing behaviors in apes and human infants: A balanced interpretation. *Child Development*, *78*(3), 729–734.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in the motor and premotor cortex. *Neuron*, *41*, 301–307.

Hauk, O., Shtyrov, Y., & Pulvermüller, F. (2006). The sounds of actions as reflected by mismatch negativity: Rapid activation of sensory-motor networks by sounds associated with finger and tongue movements. *European Journal of Neuroscience*, *23*, 811–821.

Iba, S., Paredis, C. J. J., & Khosla, P. K. (2006). Interactive multi-modal robot programming. In M. H. Ang, & O. Khatib (Eds.), *STAR*: *21. : Experimental Robotics IX* (pp. 503–512). Berlin, Heidelberg: Springer.

Knoblauch, A., & Palm, G. (2002). Scene segmentation by spike synchronization in reciprocally connected visual areas. II. Global assemblies and synchronization on larger space and time scales. *Biological Cybernetics (Berlin)*, *87*(3), 168–184.

Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L, Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: Action representation in mirror neurons. *Science*, *297*(5582), 846–848.

Please cite this article in press as: Wermter, S., et al. Multimodal communication in animals, humans and robots: An introduction to perspectives in brain-inspired informatics. Neural Networks (2009), doi:10.1016/j.neunet.2009.01.004

Kriegstein, K. V., & Giraud, A. L. (2004). Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *Neuroimage*, *22*(2), 948–955.

Markert, H., Kaufmann, U., Kara Kayikci, Z., & Palm, G. (2009). *Neural associative memories for the integration of language, vision and action in an autonomous agent*. *Neural Networks*, in this issue (doi:10.1016/j.neunet.2009.01.011).

Mirrorbot Project. Online Available at: http://www.his.sunderland.ac.uk/mirrorbot/ Accessed 20.10.08.

Moscoso Del Prado Martin, F., Hauk, O., & Pulvermüller, F. (2006). Category specificity in the processing of color-related and form-related words: An ERP study. *Neuroimage*, *29*(1), 29–37.

Murray, J. C, Erwin, H. R, & Wermter, S. (2008). *A hybrid architecture using cross-correlation and recurrent neural networks for robotic sound-source localisation*, Elsevier (in this issue).

Noordzij, M. L. (2004). Communicating spatial information from verbal descriptions. In *Thesis submitted to Helmholtz Institute*. Utrecht University. http://wayfinding.fss.uu.nl/.

Opitz, B., & Friederici, A. D. (2004). Brain correlates of language learning: The neuronal dissociation of rule-based versus similarity-based learning. *Journal of Neuroscience (New York, NY)*, *24*(39), 8436–8440.

Preissl, H., Pulvermüller, F., Lutzenberger, W., & Birbaumer, N. (1995). Evoked potentials distinguish nouns from verbs. *Neuroscience Letters*, *197*, 81–83.

Price, C. J., & Devlin, J. T. (2003). The myth of the visual word form area. *Neuroimage*, *19*(3), 473–481.

Press, C., Cook, R., Dickinson, A., & Heyes, C. M. (2008). Does learning to imitate depend on contingency. In *International Workshop on the Evolution, Development and Intentional Control of Imitation*.

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences*, *22*, 253–336.

Pulvermüller, F. (2001). Brain reflections of words and their meaning. *Trends in Cognitive Sciences*, *5*(12), 517–524.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, *6*(7), 576–582.

Pulvermüller, F., & Hauk, O. (2006). Category-specific processing of color and form words in left fronto-temporal cortex. *Cerebral Cortex*, *16*(8), 1193–1201.

Pulvermüller, F., & Knoblauch, A. (2009). Discrete rule formation in brain-inspired neural networks. *Neural Networks* (in this issue).

Pulvermüller, F., Lutzenberger, W., & Preissl, H. (1999). Nouns and verbs in the intact brain: Evidence from event-related potentials and high-frequency cortical responses. *Cerebral Cortex*, *9*, 498–508.

Pulvermüller, F., Preissl, H., Lutzenberger, W., & Birbaumer, N. (1996). Brain rhythms of language: Nouns versus verbs. *European Journal of Neuroscience*, *8*, 937–941.

Refcom Project. Online Available at: http://psy.st-andrews.ac.uk/research/refcom/ Accessed 20.10.08.

Revel, A., & Andry, (2008). Emergence of structured interactions: From a theoretical model to pragmatic robotics. *Neural Networks*, in this issue (doi:10.1016/j.neunet.2009.01.005).

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–192.

Rybski, P. E., Yoon, K., Stolarz, J., & Veloso, M. M. (2007). Interactive robot task training through dialog and demonstration. In *HRI'07*.

Salvesen, C. (1965). *The landscape of memory*. London: Arnold.

Shtyrov, Y., Hauk, O., & Pulvermüller, F. (2004). Distributed neuronal networks for encoding category-specific semantic information: The mismatch negativity to action words. *European Journal of Neuroscience*, *19*(4), 1083–1092.

Spiliotopoulos, D., Andreoutsopoulos, I., & Spyropoulos, C. D. (2001). Human-robot interaction based on spoken natural language dialogue. In *Proc. Euro. Workshop on Service and Humanoid Robotics*.

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. MIT Press.

Sullivan, W. E., & Konishi, M. (1986). Neural map of interaural phase difference in the owl's brainstem. *Proceedings of the National Academy of Sciences of the United States of America*, *83*(21), 8400–8404.

Voelkl, B., & Huber, L. (2007). Imitation as faithful copying of a novel technique in marmoset monkeys. PLoS ONE 2(7) Available at http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1905941 Accessed 14.01.08.

Wayfinding Project. Online Available at: http://wayfinding.fss.uu.nl/ Accessed 20.10.08.

Wennekers, T., Garagnani, M., & Pulvermüller, F. (2006). Language models based on Hebbian cell assemblies. *Journal de Physiologie (Paris)*, *100*, 16–30.

Wermter, S., & Elshaw, M. (2003). Learning robot actions based on self-organising language memory. *Neural Networks*, *16*(5), 661–669.

Wermter, S., Weber, C., Elshaw, M., Panchev, C., Erwin, H., & Pulvermüller, F. (2004). Towards multimodal neural robot learning. *Robotics and Autonomous Systems Journal*, *47*(2–3), 171–175.

Wermter, S., Palm, G., & Elshaw, M. (Eds.) (2005). *Biomimetic neural learning for intelligent robots*. Heidelberg, Germany: Springer.

Wermter, S., Weber, C., Elshaw, M., Gallese, V., & Pulvermüller, F. (2005a). Grounding neural robot language in action. *Biomimetic Neural Learning for Intelligent Robots: Intelligent Systems, Cognitive Robotics, and Neuroscience*, *3575*, 162–181.

Wermter, S., Weber, C., Elshaw, M., Gallese, V., & Pulvermüller, F. (2005b). Neural grounding of robot language in action. In S. Wermter, G. Palm, & M. Elshaw (Eds.), *Biomimetic neural learning for intelligent robots* (pp. 162–181). Berlin: Springer.

Van Asselen, M. (2005). The neurocognitive basis of spatial memory. In *Thesis submitted to Helmholz Institute*. Utrecht University. http://wayfinding.fss.uu.nl/ Accessed 5.01.09.

Please cite this article in press as: Wermter, S., et al. Multimodal communication in animals, humans and robots: An introduction to perspectives in brain-inspired informatics. Neural Networks (2009), doi:10.1016/j.neunet.2009.01.004