TOPICS
TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science (2014) 1–11

ISSN:1756-8757 print / 1756-8765 online
DOI: 10.1111/tops.12099

The ITALK Project: A Developmental Robotics Approach to the Study of Individual, Social, and Linguistic Learning

Frank Broz*

ITALK Project

Received 31 March 2012; received in revised form 23 January 2014; accepted 17 February 2014

Abstract

This article presents results from a multidisciplinary research project on the integration and transfer of language knowledge into robots as an empirical paradigm for the study of language development in both humans and humanoid robots. Within the framework of human linguistic and cognitive development, we focus on how three central types of learning interact and co-develop: individual learning about one's own embodiment and the environment, social learning (learning from others), and learning of linguistic capability. Our primary concern is how these capabilities can scaffold each other's development in a continuous feedback cycle as their interactions yield increasingly sophisticated competencies in the agent's capacity to interact with others and manipulate its world. Experimental results are summarized in relation to milestones in human linguistic and cognitive development and show that the mutual scaffolding of social learning, individual learning, and linguistic capabilities creates the context, conditions, and requisites for learning in each domain. Challenges and insights identified as a result of this research program are discussed with regard to possible and actual contributions to cognitive science and language ontogeny. In conclusion, directions for future work are suggested that continue to develop this approach toward an integrated framework for understanding these mutually scaffolding processes as a basis for language development in humans and robots.

Keywords: Robotics; Development; Language action; Social interaction; Learning

1. Introduction

This article presents and discusses research on learning and language development conducted as part of the ITALK (Integration and Transfer of Action and Language

Correspondence should be sent to Frank Broz, Plymouth University, School of Computing and Mathematics, Drakes Circus, Plymouth, UK PL4 8AA. E-mail: frank.broz@plymouth.ac.uk

*Please see the Note section for a complete list of authors.

Knowledge in Robots) project, an interdisciplinary research project focused on the acquisition of complex behavioral, cognitive, and linguistic skills by artificial embodied agents through individual and social learning. The primary goal of the project is to produce new theoretical insights, explanations, and models of the integration of action, social, and linguistic skills that bootstrap cognitive development. Research methodologies are employed from human–robot interaction, computational modeling, artificial intelligence, neuroscience, psychology, and developmental linguistics in a constructive manner where situated interaction is always treated as the catalyst for cognitive and linguistic development.

The work conducted investigates the application of minimal autonomous constructivist approaches to learning in realistic scenarios. Cross-disciplinary research methodologies identify and investigate conditions in which mutual scaffolding can improve the development of linguistic and cognitive capabilities. Given the importance of embodied interaction in development, models are tested through interaction between humans and physical robots whenever possible (experiments performed in simulation will be identified as such). The robot used as the primary testbed for the project's experiments was the iCub, a high degree-of-freedom toddler-sized humanoid robot designed as a platform for research in cognitive systems and developmental robotics (Tsakarakis et al., 2007) (see Fig. 1).

There have been a number of related projects involving the use of robots to study language and/or development. For example, the JST ERATO Asada Project has a similar orientation in that it focuses on developmental robotics and treats learning as arising from embodied interaction. In an overview article on developmental robotics, Asada posits language learning as arising out of two-way patterns of imitation between infant and caregiver, though language is not the primary focus of that project (Asada et al., 2009). Yiannis and Meltzoff have proposed a developmental robotics model of social imitation learning that does not directly address the emergence of language (Demiris & Meltzoff, 2008). ITALK differentiates itself from the previously mentioned projects by focusing on language learning based on the mutual scaffolding of learning with other cognitive domains, specifically action and social interaction.

Fig. 1. An experiment participant interacts with the iCub robot.

2. Mutual scaffolding in development

A primary source of inspiration and guidance for this research program is the study of infant development. Following this focus, the article will be organized based on the developmental progress of a human infant, with research results relevant to particular capabilities presented and discussed in the order in which they typically occur. The ITALK project addresses development up to roughly 2 years in age, when the ability to communicate in short sentences emerges (Boysson-Bardies & DeBevoise, 2001). During this time period, an infant makes similarly important gains in its motor abilities and capability to function as a social agent (Vernon, von Hofsten, & Fadiga, 2011). Throughout this timeline, the interrelationship between individual action learning, social learning, and linguistic development is emphasized, with studies investigating the emergence of these capabilities in conjunction with one another rather than in isolation.

While it is unclear whether mutual scaffolding is a necessary prerequisite for linguistic development, there is evidence from child development that it plays a major role in typical development and that impairment of one or more of the component cognitive abilities may be associated with developmental delays in another. Language and motor impairments often co-occur in children, suggesting that links exist between the development of these capabilities (Hill, 2001). Children with autism (a disorder that results in a lack of social attention) often experience language impairment as well, and the level of impairment can be predicted from the child's ability to perform joint attention and other social behaviors (Dawson et al., 2004). While children with impairments in these areas often do acquire full language use, the prevalence of co-occurring delays or impairments highlights the coupled nature of these capabilities and provides a motivation for modeling language acquisition as a process that is embedded in a physical embodiment and social context.

The purpose of this article is to highlight the accomplishments of this project and present them in relation to one another. Therefore, rather than exploring individual results in depth, an overview will be provided with references to relevant papers. While the results arising from this research effort are numerous, there are a few major results that serve to exemplify the approach of the project. Given the constructive nature of this work, one of the major outcomes was the production of integrated systems that combined the work of multiple research endeavors. Most notably, the final project demonstration involved the integration of multiple systems to enable action and language learning by an iCub. Details of the demonstration system are presented after the separate systems and research results that contribute to it have been introduced. The final section addresses the important role that studies involving human–robot interaction play in this work and identifies insights gained as well as areas for further research.

2.1. *One to six months*

In the first 6 months of life, infants listen to language and imitate actions. Through their close relationships with adult humans, they begin to recognize the structure of intentional behavior.

Infants become capable of recognizing syllables in utterances between the age of 1 and 5 months and able to pick out words from continuous speech between 5 and 7 months (Boysson-Bardies & DeBevoise, 2001). Computational modeling was used to explore possible mechanisms which may govern this process, investigating the feasibility of speech segmentation based on only the audio modality (Sato, Ze, & van Dijk, 2013). Another alternative model of infant speech segmentation makes more holistic use of information, relying on both acoustic and visual cues. Hirsh-Pasek and Golinkoff's concept of "acoustic packaging" proposes that language serves to give structure to the stream of actions that an infant perceives (Hirsh-Pasek & Golinkoff, 1999). The acoustic packaging system, a computational implementation of this concept, was used to analyze adult–child and adult–adult interaction data and identified differences between these groups that indicate a more structured presentation of action in adult–child interaction (Schillingmann, Wrede, & Rohlfing, 2009). This result demonstrates how adults modify their action and language in order to facilitate learning by infants, a topic which will be returned to in later sections.

From around 6 months of age, infants begin to imitate goal-directed acts on objects more often than gestures (Uzgiris, 1999). The ability to act toward goals is an important step in the expansion of an infant's motor repertoire and overall development. The Passive Motion Paradigm (PMP) framework was developed to investigate how goal-oriented action can be produced for robots without explicit prior representation (Mohan, Morasso, Metta, & Sandini, 2009). Actions emerge due to implicit interactions between the acquired internal body schema of a robot and its intended goal, through a learning process that combines imitation learning from demonstration by a human teacher, practice, and motor imagery (Mohan, Morasso, Zenzeri, & Metta, 2010).

2.2. *Seven to twelve months*

Within the first year of life, infants begin producing their first words. They also acquire word meanings through social interaction and physical interactions with their environment.

Infants begin to babble around the age of 7 months, imitating and practicing sound production that leads to the ability to pronounce words (Vihman, DePaolis, & Keren-Portnoy, 2009). In a series of human–robot interaction experiments, a computational model of the transition from babbling to word form acquisition was evaluated through proto-conversations between the iCub and a human (Lyon, Nehaniv, & Saunders, 2012). In this system, chance occurrences of words in random syllabic babble spoken by the robot are reinforced by positive feedback from the interaction partner, biasing the robot towards repetition.

From the age of 7 to 10 months, infants start recognizing words in context, learning word meanings by between 10 and 12 months and speaking their first words around 12 months (Boysson-Bardies & DeBevoise, 2001). The ROSSUM architecture empirically models the learning of word meanings though social interaction. Heuristics based on the characteristics of motherese allow a robot to make associations between words and features of its environment or its own embodiment through interaction with a human teacher (Saunders, Nehaniv & Lyon, 2011). Motherese is a term for a style of speaking typically employed by caregivers when talking with a child (Fernald & Kuhl, 1987). Participants' use of motherese became more pronounced in successive interaction sessions with the robot as they became more aware of its limited abilities. This result provides another example of how adults socially scaffold language learning by modifying their behavior based on the perceived needs of their listener. The ROSSUM architecture has also been used to model the acquisition of linguistic negation (Förster, 2013). Infants learn the pragmatics of words from their caretakers, and rejective "no"s are typically among the group of first words uttered by toddlers (Pea, 1980). ERA (Epigenetic Robot Archictecture) is another approach to modeling the learning of word meanings, a neural cognitive architecture that focuses on the influence of embodiment and spatial information on learning (Morse, DeGreeff, Belpeame, & Cangelosi, 2010). In a series of experiments on the iCub, results from children's word learning experiments were replicated, demonstrating the impact of spatial information, embodiment, and proprioceptive information (Morse, Belpaeme, Cangelosi, & Smith, 2010; Morse, Baxter, Belpaeme, Smith, & Cangelosi, 2011).

Gaze plays an important role in organizing episodes of learning. Infants prefer to look at faces that engage them in mutual gaze, and by the age of 10 to 12 months, they become adept at gaze following (Vernon et al., 2011). Studying mutual gaze during conversation can identify factors that influence this behavior and aid in the design of socially appropriate robot gaze controllers (Broz, Lehmann Nehaniv, & Dautenhahn, 2012a,b). Gaze makes information about a caregiver's attention available to an infant, and the caregiver monitors the infant's attention in turn. Analyses of human–human interactions reveal that parents monitor their child's attention more closely in younger than in older infants (Pitsch et al., 2009). In tutoring situations with pre-lexical infants, feedback consists primarily of gazing behavior (Vollmer, Lohan, Fritsch, Wrede, & Rohlfing, 2010). Changes in the type of feedback given by children at more advanced stages of development will be discussed in subsequent sections.

2.3. Twelve to twenty months

In the second year of life, children learn about the compositionality of language, beginning to recognize grammatical constructs and produce their own multiword utterances. They simultaneously become more skilled at composing actions into goal-directed behaviors. As their motor and linguistic skills improve, so does their level of agency during interaction.

Between the ages of 12 and 16 months, infants begin to understand simple sentences (Boysson-Bardies & DeBevoise, 2001). Motor capabilities also increase, and by the age of about 13 months, infants are able to smoothly compose general compound actions such as reaching and grasping (Vernon et al., 2011). A series of simulation experiments investigated the development of language comprehension, including compositional semantics grounded in an agent's physical capabilities and environment. Neural controllers were trained for a task that involved the simultaneous learning of a set of physical skills and the ability to respond to imperative utterances to perform those skills (Tuci, Ferrauto, Zeschel, Massera, & Nolfi, 2011). The acquired controllers generalize in both language comprehension and behavior generation, able to respond to new, never-experienced utterances with never-produced behaviors without further training. The previously mentioned PMP framework also addresses action compositionality though its learning architecture that reuses existing motion knowledge while learning novel motor skills (Mohan et al., 2011).

Sometime between 16 and 20 months, children begin expressing two-word utterances and are able to distinguish word categories (Boysson-Bardies & DeBevoise, 2001). The ROSSUM architecture was extended to model the two-word stage through the learning of meaningful adjective–noun combinations (Lehmann, Förster, & Nehaniv, 2012). Words were grouped using heuristics based on typical English word order. The perceptual associations learned for these categories were related to the words' grammatical functions.

Children of this age in tutoring situations use more modalities to give feedback and begin to anticipate next actions with their gaze direction (Vollmer et al., 2010). Infants begin to exhibit gaze that anticipates the goal of another's action once they have learned to perform that action themselves (Falck-Ytter, Gredebaeck, & von Hofsten, 2006). Experiments found that observing actions of a humanoid robot also evokes gaze proactivity, suggesting there is a mapping between the observed action and the observer's motor repertoire based in the mirror neuron system (Sciutti et al., 2013). Experiments on motor resonance have found that biologically plausible motion by humanoid robots can produce the same type of motor resonance effects as human motion (Sciutti et al., 2012). These results suggest that people do process robot action similarly to human action when it is human-like, so robots should be able to elicit similar responses during interaction.

2.4. Twenty to twenty-four months

By 24 months of age, children are able to construct two- and three-word sentences (Boysson-Bardies & DeBevoise, 2001). They have integrated the skills of acting on objects in their environment and talking to interaction partners about those actions. In children aged 2 years and up, the type of feedback they give to caretakers changes to reflect their greater action and language comprehension of action. They take an increasingly active role in learning, structuring their feedback by timing their verbal and bodily action in relation to the adult's presentation (Vollmer et al., 2010). Parents, in turn, use more speech (instead of or in addition to motion) to attract older children's attention (Vollmer et al., 2009). From these analyses of tutoring in parent–child dyads, a tutor spotter, a continuous measurement scale based on interaction features, was created. It was used to analyze human tutoring behavior toward a simulated robot and the iCub robot. Simple saliency-based robot gaze behavior was not found to produce a contingent interaction and did not induce tutoring behavior in the human (Lohan et al., 2012).

In response, contingent robot gazing and reaching behavior was implemented. Conversation analysis applied to interactions between a human and a robot exhibiting these behaviors demonstrate how the robot's behavior influences the human's tutoring style (Pitsch et al., 2012). These behaviors may improve the learning ability of interactive systems because the information presented is more likely to be at a more appropriate level of complexity. Humans are especially sensitive to the robot's behavior during the first moments of interaction, which creates their expectations of the robot's abilities. Additional results on the influence of human expectations on language use toward robots will be discussed in Section 4.

3. Integrated language learning systems

One of the significant outcomes of this construction-oriented research program was the production of several working systems that were a result of collaboration between partners. For example, the tutor spotter system and the ROSSUM learning architecture were integrated on an iCub robot for an experiment on socially directed robot learning (Lohan, Rohlfing, Saunders, Nehaniv, & Wrede, 2012). In this study, it was investigated how different robot feedback might induce tutoring behavior in a human that could help support robot language acquisition. This type of integrated systems allowed partners to conduct research on the interrelationship between motor, social, and language learning through interaction with physical robots rather than modeling individual aspects in simulation.

The major integration effort for the ITALK project was the production of a robot learning demonstration that merged multiple systems designed over the course of the project in order to enable grounded robot language and action learning up to the point of multiword utterances. This was achieved through the integration of the PMP motion control framework, ERA learning architecture, acoustic packaging system, and compositional grammar learning. In the demonstration, the robot displays language and action capabilities during interaction with a human tutor at three stages of development that are analogous to a human infant's developmental trajectory and learns new competencies at each stage. At the first stage, the robot performs simple actions on objects and learns word-object associations from the human. During the second stage, the robot learns word-action and word-attribute associations from the tutor's verbal descriptions of actions it performs on objects. And during the third stage, the robot demonstrates simple grammar learning based on the speech it has been exposed to and learns to compose complex actions from simple ones. Video documentation of the system in action can be found at http://youtube/5l4LHD2lYJk.

4. How humans talk to robots (and why)

An important goal of the ITALK project is to explore computational models that bootstrap language learning through interaction, both sensorimotor interaction with the environment and social interaction with humans. It is essential to know how people use language with robots, because these interactions constitute the input robots bootstrap their language knowledge from. Caretakers have been found to produce adaptations that potentially facilitate language acquisition for child language learners. Will people interact with robots in similar ways as with young children, making these facilitative cues available to robot language learners?

Extrapolating from previous work in human–computer interaction suggests that people may involuntarily and automatically interact with robots like they interact with humans (Reeves & Nass, 1996). Instead, comparisons between child-directed and robot-directed speech in this project revealed interaction to be mainly based on participants' perceptions of the robot's affordances (Fischer, 2011a). Users' preconceptions also influence their interactions with robots considerably, guiding them to a particular understanding of the situation (Fischer, 2011b). A robot's appearance influences the way people talk to it, but while humanoid appearance may suggest human-like capabilities, people do not necessarily talk to a robot as they do to a child (Fischer, Foth Rohlfing, & Wrede, 2011a,b). Instead, they behave based on what they understand the functionality of the robot to be, informed by a more global partner model (Fischer, 2011c). Infant-like robot behavior is a far stronger influence on human interaction behavior than appearance alone. Therefore, feedback plays a considerable role in the input the robot receives for language learning.

Designing appropriate behavior thus constitutes an important challenge for future work on the acquisition of language by robots. The study of human interaction helps to identify what behaviors among the complex patterns that characterize embodied verbal interactions a robot should perform to encourage humans to conceive of it as a potential language learner. The human–human and human–robot interaction experiments conducted during the ITALK project examined human behaviors that facilitate face-to-face communication and language learning, investigated under what conditions humans will engage in such facilitative behaviors with a robot, and suggested further directions for the design of robot behavior to facilitate language learning through interaction with humans.

5. Conclusion

This article provides an overview of the products of a research program focused on investigating the mutual scaffolding of individual, social, and language learning in the development of human children by employing a constructive approach to modeling and studying linguistic development using a humanoid robot platform. This embodiment-oriented and ecologically situated methodology is quite different from focusing on specific aspects of development in isolation, where they are artificially treated as disconnected phenomena. We use language to enact our relationships with the world around us, both in humans and in robots. Results from human experiments conducted during this project informed further experimentation using computational models. The extensive use of human–robot interaction experiments to study language acquisition has led to new discoveries about the role of robot behavior in eliciting human behavior that can scaffold learning. The results of the ITALK project provide a basis for further work on more advanced language development expanding on this paradigm, including work that uses robots to test theories from child development and uses computational models to make predictions about human language learning.

Acknowledgments

The work described in this paper was conducted within the EU Integrated Project ITalk funded by the European Commission under contract number FP7-214668.

Note

Chrystopher L. Nehaniv, Tony Belpaeme, Ambra Bisio, Kerstin Dautenhahn, Luciano Fadiga, Tomassino Ferrauto, Kerstin Fischer, Frank Förster, Onofrio Gigliotta, Sascha Griffiths, Hagen Lehmann, Katrin S. Lohan, Caroline Lyon, Davide Marocco, Gianluca Massera, Giorgio Metta, Vishwanathan Mohan, Anthony Morse, Stefano Nolfi, Francesco Nori, Martin Peniak, Karola Pitsch, Katharina J. Rohlfing, Gerhard Sagerer, Yo Sato, Joe Saunders, Lars Schillingmann, Alessandra Sciutti, Vadim Tikhanoff, Britta Wrede, Arne Zeschel, and Angelo Cangelosi.

References

Asada, M., Hosoda, K., Kuniyoshi, Y., Ishiguro, H., Inui, T., Yoshikawa, Y., & Yoshida, C. (2009). Cognitive developmental robotics: A survey. *IEEE Transactions on Autonomous Mental Development*, *1*(1), 12–34.

Boysson-Bardies, B., & DeBevoise, M. (2001). *How language comes to children: From birth to two years*. Cambridge, MA: MIT Press.

Broz, F., Lehmann, H., Nehaniv, C. L., & Dautenhahn, K. (2012a). Automated analysis of mutual gaze in human conversational pairs. In Y. Nakano, C. Conati, & T. Bader (Eds.), *Eye gaze in intelligent human computer interaction* (pp. 41–60). New York: Springer.

Broz, F., Lehmann, H., Nehaniv, C. L., & Dautenhahn, K. (2012b). Mutual gaze, personality, and familiarity: Dual eye-tracking during conversation. In *Proceedings of IEEE International Symposium on Robot and Human Interactive Communication (Ro-Man)* (pp. 858–864). Paris, France: September 9-13.

Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., & Liaw, J. (2004). Early social attention impairments in autism: Social orienting, joint attention, and attention to distress. *Developmental Psychology*, *40*(2), 271–283.

Demiris, Y., & Meltzoff, A. (2008). The robot in the crib: A developmental analysis of imitation skills in infants and robots. *Infant and Child Development*, *17*(1), 43–53.

Falck-Ytter, T., Gredebaeck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience*, *9*(7), 878–879.

Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, *10*(3), 279–293.

Fischer, K. (2011a). How people talk with robots—Dialog as uncertainty reduction. *AI Magazine*, *32*(4), 31–38.

Fischer, K. (2011b). Interpersonal variation in understanding robots as social actors. In *Proceedings of the 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 53–60). ACM, New York, NY, USA.

Fischer, K. (2011c). *Recipient design, alignment, interaction: The role of the addressee in so-called "simplified registers."* Habilitation thesis, University of Bremen.

Fischer, K., Foth, K., Rohlfing, K., & Wrede, B. (2011a). Is speaking to a simulated robot like speaking to a child? In *Proceedings of 1st Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics ICDL-EPIROB*. vol: 2, IEEE. Frankfurt am Main, 24-27 August.

Fischer, K., Foth, K., Rohlfing, K., & Wrede, B. (2011b). Mindful tutors—linguistic choice and action demonstration in speech to infants and to a simulated robot. *Interaction Studies*, *12*(1), 134–161.

Förster, F. (2013). *Robots that say 'no': Acquisition of linguistic behaviour in interaction games with humans* (Tech. Rep.). Hatfield, UK: University of Hertfordshire.

Hill, E. L. (2001). Non-specific nature of specific language impairment: a review of the literature with regard to concomitant motor impairments. *International Journal of Language & Communication Disorders*, *36* (2), 149–171.

Hirsh-Pasek, K., & Golinkoff, R. (1999). *The origins of grammar: Evidence from early language comprehension*. Cambridge, MA: MIT Press.

Lehmann, J. S. H., Förster, F., & Nehaniv, C. L. (2012). Robot acquisition of lexical meaning - moving towards the two-word stage. In *Proceedings of 2nd Joint IEEE Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)* IEEE. San Diego, CA, 7-9 November.

Lohan, K. S., Rohlfing, K. J., Saunders, J., Nehaniv, C. L., & Wrede, B. (2012). Contingency scaffolds language learning. In *Proceedings of 2nd Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics (ICDL-EpiRob)* IEEE. San Diego, CA, 7-9 November.

Lohan, K. S., Rohlfing, K. J., Pitsch, K., Saunders, J., Lehmann, H., Nehaniv, C. L., & Wrede, B. (2012). Tutor spotter: Proposing a feature set and evaluating it in a robotic system. *International Journal of Social Robotics*, Special Issue on Expectations, Intentions, and Actions, *4*(2), 131–146.

Lyon, C., Nehaniv, C. L., & Saunders, J. (2012). Interactive language learning by robots: The transition from babbling to word forms. *PLoS ONE*, *7*(6), e38236.

Mohan, V., Morasso, P., Metta, G., & Sandini, G. (2009). A biomimetic, force-field based computational model for motion planning and bimanual coordination in humanoid robots. *Autonomous Robots*, *27*(3), 291–301.

Mohan, V., Morasso, P., Zenzeri, J., & Metta, G. (2010). Teaching humanoids to imitate 'shapes' of movements. In K. Diamantaras, W. Duch, & L. S. Iliadis (Eds.), *Proceedings of the 20th International Conference on Artificial Neural Networks: Part II (ICANN)* (pp. 234–244). Berlin: Springer-Verlag.

Mohan, V., Morasso, P., Zenzeri, J., Metta, G., Chakravarthy, V., & Sandini, G. (2011). Teaching a humanoid robot to draw "shapes." *Autonomous Robots*, *31*(1), 21–53.

Morse, A. F., DeGreeff, J., Belpeame, T., & Cangelosi, A. (2010). Epigenetic robotics architecture (ERA). *IEEE Transactions on Autonomous Mental Development*, *2*(4), 325–339.

Morse, A. F., Belpaeme, T., Cangelosi, A., & Smith, L. B. (2010). Thinking with your body: Modelling spatial biases in categorization using a real humanoid robot. In Ohlsson, S. and Catrambone, R. (Ed.), *Proceedings of the 2010 Annual Meeting of the Cognitive Science Society* (pp. 1362–1368). Curran Associates: Red Hook, NY.

Morse, P., Baxter, A., Belpaeme, T., Smith, L. B., & Cangelosi, A. (2011). The power of words (and space). In *Proceedings of the 1st Joint International Conference on Developmental Learning and on Epigenetic Robotics (ICDL-EpiRob).* vol: 2, IEEE. Frankfurt am Main, 24-27 August.

Pea, R. (1980). The Development Of Negation In Early Child Language. In D. Olson (Ed.), *The social foundations of language & thought: Essays in honor of Jerome S. Bruner* (pp. 156–186). New York, NY: W.W. Norton.

Pitsch, K., Vollmer, A.L., Fritsch, J., Wrede, B., Rohlfing, K., & Sagerer, G. (2009). On the loop of action modification and the recipient's gaze in adult-child interaction. In Jarmolowicz-Nowikow E. Juszczyk K.

Malisz Z. and Szczyszek M. (Ed.), *Proceedings of Gesture and Speech in Interaction (GESPIN)* (pp. 32. Poznan, Poland: 24-26 September.

Pitsch, K., Lohan, K. S., Rohlfing, K., Saunders, J., Nehaniv, C. L., & Wrede, B. (2012). Better be reactive at the beginning. implications of the first seconds of an encounter for the tutoring style in human-robot-interaction. In *Proceedings of IEEE international symposium on robot and human interactive communication (Ro-Man)* (pp. 974–981). IEEE. Paris, France, 9-13 September.

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. New York: Cambridge University Press.

Sato, Y., Ze, J., & van Dijk, S. (2013). Recurrence-based learning of word-like units: learning by 'echoing' in human-robot interaction. In L. Gogate (Ed.), *Theoretical and computational models of word learning: Implications for psychology and artificial intelligence* (pp. 327–349). Hershey, PA: IGI Global.

Saunders, J., Nehaniv, C. L., & Lyon, C. (2011). The acquisition of word semantics by a humanoid robot via interaction with a human tutor. In K. Dautenhahn, & J. Saunders (Eds.), *New frontiers in human-robot interaction* (pp. 211–234). Amsterdam, The Netherlands: John Benjamins.

Schillingmann, L., Wrede, B., & Rohlfing, K. J. (2009). A computational model of acoustic packaging. *IEEE Transactions on Autonomous Mental Development*, *1*(4), 226–237.

Sciutti, A., Bisio, A., Nori, F., Metta, G., Fadiga, L., Pozzo, T., & Sandini, G. (2012). Measuring human-robot interaction through motor resonance. *International Journal of Social Robotics*, *4*(3), 223–234.

Sciutti, A., Bisio, A., Nori, F., Metta, G., Fadiga, L., & Sandini, G. (2013). Robots can be perceived as goal-oriented agents. *Interaction Studies, Special Issue on Gaze in Human-Robot Communication*, *14*(3), 329–350.

Tsakarakis, N., Metta, G., Sandini, G., Vernon, D., Beira, R., Becchi, F., & Caldwell, D. (2007). iCub—The design and realization of an open humanoid platform for cognitive and neuroscience research. *Journal of Advanced Robotics, Special Issue on Robotic platforms for Research in Neuroscience*, *21*(10), 1151–1175.

Tuci, E., Ferrauto, T., Zeschel, A., Massera, G., & Nolfi, S. (2011). An experiment on behaviour generalisation and the emergence of linguistic compositionality in evolving robots. *IEEE Transactions on Autonomous Mental Development*, *3*(2), 176–189.

Uzgiris, I. (1999). Imitation as activity: its developmental aspects. In J. Nadel & G. Butterworth (Eds.), *Imitation in infancy* (pp. 186–206). Cambridge, UK: Cambridge University Press.

Vernon, D., von Hofsten, C., & Fadiga, L. (2011). *A roadmap for cognitive development in humanoid robots* (1st ed., Vol. 11). Berlin: Springer-Verlag.

Vihman, M., DePaolis, R., & Keren-Portnoy, T. (2009). A dynamic systems approach to babbling and words. In E. Bavin (Ed.), *Handbook of child language* (pp. 163–182). Cambridge, UK: Cambridge University Press.

Vollmer, A. L., Lohan, K. S., Fischer, K., Nagai, Y., Pitsch, K., Rohlfing, K. J., & Wrede, B. (2009). People modify their tutoring behavior in robot-directed interaction for action learning. In *Proceedings of the IEEE 8th international conference on development and learning (ICDL)* IEEE. Shanghai, 5-7 June.

Vollmer, A. L., Lohan, K. S., Fritsch, J., Wrede, B., & Rohlfing, K. J. (2010). Developing feedback: How children of different age contribute to a tutoring interaction with adults. In *Proceedings of the IEEE 9th International Conference on Development and Learning (IEEE)* (pp. 76–81). IEEE. Ann Arbor, MI. 18-21 August.