The Challenges of Joint Attention

Frédéric Kaplan
Sony CSL Paris.
6 rue Amyot 75005 Paris France
kaplan@csl.sony.fr

Verena V. Hafner
Sony CSL Paris.
6 rue Amyot 75005 Paris France
hafner@csl.sony.fr

Abstract

This paper discusses the concept of joint attention and the different skills underlying its development. We argue that joint attention is much more than gaze following or simultaneous looking because it implies a shared intentional relation to the world. The current state-of-the-art in robotic and computational models of the different prerequisites of joint attention is discussed in relation with a developmental timeline drawn from results in child studies.

1. Introduction

Joint attention is probably one of the hardest problems to be solved by developmental robotics research. In a recent paper called "A constructive model for the development of joint attention", Nagai et al. describe "a constructive model that enables a robot to acquire the ability of joint attention" without a controlled environment nor external task evaluation (Nagai et al., 2003). Although this paper definitively makes an interesting contribution for understanding how a robot could learn to interpret human gaze in order to spot salient objects in its environment, we believe it does not cover all the aspects of joint attention.

We discuss in this paper the concept of joint attention and the different skills underlying its development. In the line of Tomasello's views (Tomasello, 1995), we argue that joint attention implies viewing the behaviour of other agents as intentionally-driven. In that sense, joint attention is much more than gaze following or simultaneous looking. Summarising results from developmental psychology, the paper presents a timeline showing at what age the different prerequisites for joint attention arise during the first two years in the life of a child. In relation with this developmental timeline, the paper discusses the current state-of-the-art in robotic and computational models of joint attention and identifies which issues remain to be addressed.

2. What is Joint Attention ?

2.1 Defining attention

Attention is the process whereby an agent concentrates on some features of the environment to the (relative) exclusion of others. This process can occur in two situations.

1. Passive attention: a salient event happens (e.g. loud noise) and automatically triggers the attention of the agent.

2. Active attention: the agent is involved in an intentionally directed process (e.g. climbing a mountain) and must actively select particular features of its environment.

The attentional behaviour is the externally perceivable behaviour that goes along with the attention process. To reach joint attention, agents must actively track and manipulate the attention of each other. Discussing the prerequisites of this coordination is the aim of this paper. But before that, we must specify what we mean by joint attention.

2.2 Defining joint attention

2.2.1 Joint attention is **not** simultaneous looking

Joint attention is often associated with a situation where two agents are looking at the same thing. We will now examine four cases of simultaneous looking which do not qualify for joint attention. For better illustration, we use examples of interaction between two robots (Figure 1).

Case 1a: Simultaneous looking triggered by a salient event (passive attention). The two robots are sitting in a room. Suddenly, one of their toys makes a squeaking noise. They both turn and look at it immediately.

Case 1b: Simultaneous looking triggered by a "pop-out" effect (passive attention). The robots found a box filled with balls. All the balls are blue, apart from one which is pink. Both robots are attracted by the pink ball.
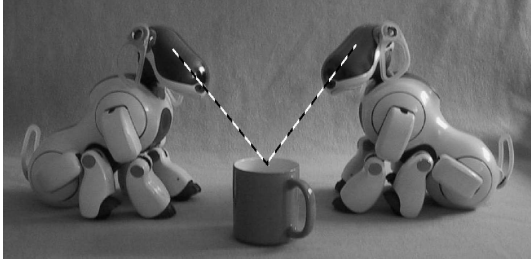
Figure 1: Two Sony AIBO robots are looking simultaneously at a coffee cup. Is this already joint attention?

Case 2: Coincidental simultaneous looking. The robots are looking for a toy to play with. At the same moment, they both see a pink ball on the floor. They pay attention to it without noticing each other. Each other's attention is not monitored.

Case 3: Gaze following. One robot is looking at a new toy. The other less experienced robot follows his gaze since it has learned that by doing that, it will often see something interesting. But attention is not joint, as the first robot is not paying attention to the behaviour of the other one.

Case 4: Coordinated gaze on an object. Both robots are looking at a toy bunny, and are also aware that the other one is looking, too. From an outside observer's point of view, this situation looks like joint attention. However, one robot is attending to the bunny in order to play with it, the other one is purely attracted by its colour. They are therefore not attending to the same aspect of the object.

These different cases of simultaneous looking are summarised in table 1. Joint attention is an active bilateral process which involves attention alternation, but it can only be fully understood if we assume that it is realized by intentional agents.

2.2.2 Joint attention as a shared intentional relation to the world

Active attention occurs when an agent is involved in an intentionally directed process. This means that the agent tries to achieve a particular desirable situation that constitutes its aim or goal (e.g. being on top of a mountain, reducing hunger, following someone, learning something). To realize this aim, the agent focuses selectively on particular perceptual features. In that sense, attention is intentionally directed perception (Tomasello, 1995).

The only way for an agent to read the intention of another agent is by watching its behaviour. Here are a few examples:

Example 1: Intention detection through general behaviour. One robot sees another robot walking towards the charging station. He infers that his battery is low and that he needs to recharge. In this case, the observer did not need to track the other one's attention to understand the underlying intention.

Example 2: Intention detection through attentional behaviour. One robot is looking attentively at the closed door. The other robot infers that it attends to the door because it wants to go outside. Here, tracking the attentional behaviour is relevant to understand what the other robot attends to.

To reach joint attention an agent must understand, monitor and direct the intentions underlying the attentional behaviour of the other agent. Attention can only be reached if both agents are aware of this coordination of "perspectives" towards the world (Hobson, 2002).

In the same way that attention cannot be reduced to visual orientation, joint attention is much more than a geometrical phenomenon. It needs to be understood as a crucial step in the development of social cognition.

2.3 The prerequisites of joint attention

Reaching joint attention implies at least four kinds of prerequisites.

- **Attention Detection.** An agent must be able to track the attentional behaviour of other agents. This may imply being able to follow the gaze of another agent.

- **Attention Manipulation.** Agents must be able to manipulate the attentional behaviour of other agents. The use of pointing gestures or words can be used in that respect.

- **Social coordination.** Agents must be able to engage in coordinated interaction with other agents. This implies mastering social techniques such as turn-taking, role-switching and ritualised games.

- **Intentional stance.** Agents must view themselves and others as intentional agents. They must understand that others have intentions possibly different from their own. By taking the intentional stance, agents interpret and predict the behaviour of other agents assuming it is goal-directed (Dennett, 1987).

The rest of the paper examines data drawn from developmental psychology on the development of

Table 1: Different cases of simultaneous looking

Case	Active/Passive	Attention detection	Unilateral/Bilateral
Case 1: Simultaneous Looking triggered by a salient event or a "pop-out" effect	Passive	No	-
Case 2: Coincidental simultaneous looking	Active	No	-
Case 3: Gaze following	Active	Yes	Unilateral
Case 4: Coordinated gaze on same object	Active	Yes	Bilateral

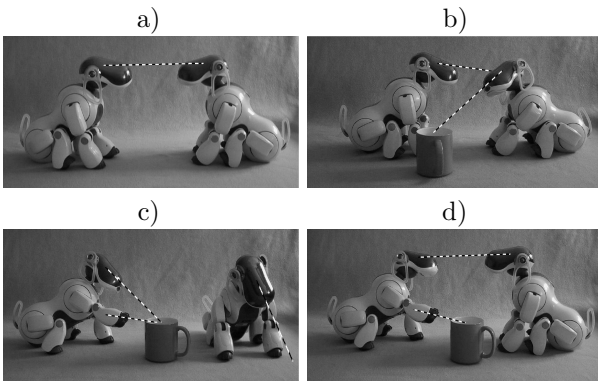a)   b)

c)   d)

Figure 2: Demonstration of different situations preceding joint attention during development. a) Mutual Gaze. Both robots are attending to each other's gaze simultaneously. b) Gaze Following. One of the robots is paying attention to an object, the other one watches its eyes in order to detect where it is looking. c) Imperative Pointing. Pointing to an object regardless whether another person or robot is attending. d) Declarative Pointing. Pointing to an object to create shared attention.

these capabilities and discusses existing robotic and computational models for each of them.

3.   Developmental Timeline

It is interesting to observe at what age the different skills and prerequisites for joint attention arise in young children during their development. Table 2 presents these skills in the temporal order in which they occur first between three and 24 months when joint attention is fully developed.

For better illustration, some situations on attention detection and attention manipulation are displayed in figure 2 using Sony AIBOs.

3.1   Attention detection

In the first month of their lives, babies progressively bootstraps the capability to pay attention to a growing number of things in their environment: their own body, external objects, animate beings, etc. In this developmental process, they start paying attention to the attentional behaviour of other agents.

T1.1 Mutual gaze.   (Figure 2a) Mutual gaze between an adult and a child occurs first around the age of three months. At this age, the baby shows a strong preference towards face-like patterns and is capable of recognising and maintaining eye contact. This sensibility of eye contact is also reported in the behaviour of many animals, in particular in primates (Cheney and Seyfarth, 1990). Mutual gaze is a special case of attentional behaviour since it does not involve any objects or persons apart from the two involved.

T1.2-5 Gaze following.   (Figure 2b) At the age of six months, the first true incident of attention detection starts. The child is able to attend to an object in the correct side of the room depending on where the adult is looking at (T1.2). The angle error between the attended object of the adult and the attended object of the infant can be as large as 60 degrees (Butterworth, 1995). Only at the age of nine months can the gaze direction of the adult be accurately detected, however, always the first object within the line of sight is chosen (T1.3). The correct object can be attended to by the age of twelve months (Butterworth and Jarrett, 1991) taking into account vergence and probably context (T1.4). By this age, only objects which are in the field of view of the child are being considered, even though the child is already turning to sounds coming from behind (Butterworth, 1995, Butterworth and Cochran, 1980). Only at 18 months, children start following the gaze of an adult to objects outside their field of view (T1.5). If directing the gaze towards an object is supported by also pointing towards that object, the accuracy of attending to the correct object increases in infants older than twelve months (Butterworth, 1995). Before that age, pointing is not understood by the child and does not make any difference to the child's attention.

3.2   Attention manipulation

Skills which fall into the category of attention manipulation are the act of pointing at something and the use of language.

Table 2: Developmental timelines of the prerequisites for joint attention

Age from:	Attention detection	Attention manipulation	Social coordination	Intentional stance
0-3 m	**T1.1 Mutual gaze** - Eye contact detection		**T3.1 Protoconversations**: Simple rhythmic interaction including turn-taking mediated by the caregiver.	**T4.1 Early identification with other persons**
4 m			**T3.2** Possibility of breaking interactions	
6 m	**T1.2** Discrimination between **left and right position** of head and gaze		**T3.3 Shared games**: Conventional routines established between the child and the caregiver	
9 m	**T1.3 Gaze angle detection** - fixation on the first salient object encountered	**T2.1 Imperative Pointing**: Drawing attention as a request for reaching an object (attention not monitored)	**T3.4 Simple immediate imitation**: The child commonly imitates a movement performed by the caregiver.	**T4.2** First goal-directed behaviours
12 m	**T1.4 Gaze angle detection** - fixation on any salient object encountered - Accuracy increased in the presence of a pointing gesture	**T2.2 Declarative Pointing**: Drawing attention using gestures		**T4.3** Systematic separation between **goals and means** - Beginning of the intentional stance
13 m		**T2.3 Declarative/Referential words**: Drawing attention using a word		
18 m	**T1.5 Gaze following** toward object outside the field of view - Full object permanence	**T2.4 First predications**: Drawing attention using non-linguistic gesture for the topic and a word to specify which aspect of the object should be attended	**T3.5 Complex imitative games** Social exchanges using imitation including conventional routines and role-switching	**T4.4 Intentional stance** developed. Behaviour of others viewed as goal-directed.
24 m		**T2.5 Conversations**: Both topic and aspect can now be specified linguistically		

T2.1 Imperative pointing. (Figure 2c) The first occurrence of pointing, imperative pointing, starts first at the age of nine months (Baron-Cohen, 1997). Imperative pointing might be an extension of grasping an object, and it also occurs when nobody who could pay attention is present in the room. This means that the attention is not monitored. Imperative pointing is the request for a certain object.

T2.2 Declarative pointing. (Figure 2d) At twelve months, shortly before the use of linguistic symbols, pointing starts to become declarative. It is used to draw someone's attention to something which might also be outside of reach for the adult, such as objects like the sun or an aeroplane. One could think that this pointing behaviour results from an imitation of the gestures of the adult. However, no relation between the production of pointing and the comprehension of pointing has been found in studies with young children (Desrochers et al., 1995).

T2.3 Declarative/Referential words. After drawing attention using gestures, the child starts to use single words to draw attention to objects or persons around the age of 13 months.

T2.4-5 First predications. First predication follows at about 18 months, and already requires building of a simple context representation. At this age, the child specifies the subject of interaction by pointing and then adds a comment linguistically in order to draw the attention of the adult towards a particular aspect of it (e.g. "big") (T2.4). By the age of 24 months, both the topic and the comment start to be expressed linguistically (e.g. "big dog") (T2.5).

3.3 Social coordination

T3.1-2 Protoconversation. Six-week old children are already communicating extensively face-to-face with their caregiver. These first simple rhythmic interactions are crucial for the development of social know-how (Trevarthen, 1979). Newson argues that these early social responses are treated by the adult as normal social behaviour (Newson, 1979). For instance when the child does something that can be interpreted as role switching or change in the course of the "dialog", the adult adapts in order to make it become meaningful. In such conditions, these proto-dialogs exhibit already simple turn-taking behaviours. As the adult scaffolds these interaction into structured dialogs, children learn to predict the social effects of their behaviour (Schaffer, 1977). By the age of four months, children are able to break their caregiver's gaze in order to look at other things in the world (Siegel, 1999) (T.3.2). This opens to the possibility of more complex interactions.

T3.3 Shared games. Each caregiver develops his or her own set of conventional games. By the age of six months, a child manages to master an important number of them. These ritualised structures play a crucial role for defining roles and imposing consistency and predictability in social exchanges (Kaye, 1982).

T3.4-5 Imitative games. A common interaction routine consists in the immediate imitation by the child of a movement produced by the caregiver (T3.4). This skill, already present in the very young infant, gradually develops and is used commonly around nine months. Nadel has emphasised the role of such immediate imitations for bootstrapping social exchanges in particular for turn-taking, role switching and in order to share topic (Nadel, 2002). Around 18 months, it starts to be used inside complex games (T3.5).

3.4 Intentional stance

Tomasello argues that a crucial behavioural transition occurs around twelve months (Tomasello, 1995). Before one year, children begin following and directing the attention of other persons, but do not view them as intentional agents. At the beginning of the second year of their life, they demonstrate a qualitative change in the nature of their behaviour. Complex social skills such as social referencing, imitative learning or symbolic communication with gestures appear almost simultaneously (see table 2). This synchrony suggests that a radical shift has occurred in children's awareness of their environment: they have taken the intentional stance. There is a vast range of theories on how to interpret this shift from totally nativistic to totally cultural hypotheses. For instance, Trevarthen argues that children view other persons as intentional agents from birth, independently from any prior experience (Trevarthen, 1979). On the opposite side, Kaye believes that children construct the notion of intentional agents totally from experience. During the first year of their life, an important part of children's experiences are mediated by the parents. The fact that parents are always treating children as intentional agents even before they are such may also play an important role for their development of the intentional stance ("parents create persons") (Kaye, 1982). Some important milestones occur in this developmental route towards the intentional stance.

T4.1 Early identification. Early identification with other persons, taking the form of simple imitative behaviours, has been observed in the first months of life. To explain these experiments, some totally or partially nativist theories

have been put forward (Meltzoff and Gopnick, 1993, Moore and Corkum, 1994). Whatever their innate basis is, these neonatal forms of imitation make children exposed to situations in which their intention and the one of the adult happen to converge. They may play a role for the progressive distinction by the child of the first and third person perspectives.

T4.2-4 Separation between goals and means. Piaget observes that children first start to display goal directed behaviour around nine months (Piaget, 1952). They may for instance remove an obstacle in order to reach a particular place. This means that they start to differentiate goals and means in their own behaviour. They start to view their own behaviour as intentionally-driven. This kind of goal-directed behaviour becomes common around twelve months (Frye, 1991). Extensions of this discrimination for the interpretation of the behaviour of other agents may occur as a consequence of this first finding (Tomasello, 1995). Experimental evidence that infants understand other's goals and intentions appears at 18 months (Meltzoff, 1995).

4. Robotic and Computational Models

The precise developmental route that leads to mastering the necessary skills for joint attention is largely unknown. Robots are ideal tools to model the development of joint attention. Their embodiment in the real world allows for interactions between robots as well as interaction between humans and robots. Experiments are - in contrast to observing the behaviour of children - repeatable and different aspects can be easily separated.

In this section, we shortly review the state-of-the-art research in robotics concerning joint attention and its various prerequisites. No system has yet achieved true joint attention between a robot and a human or between two robots in the sense we defined it in the previous sections. Several crucial steps have started to be investigated, but important parts of this developmental puzzle are still unexplored.

4.1 Models for attention detection and attention manipulation

Table 2 shows that the child manages to make progress in detecting and manipulating the attention of the adult through a series of steps of increasing complexity. Some of these skills have already been designed by hand on a robot. Imai et al.'s robot 'Robovie' (Imai et al., 2001) is able to attract a human's attention by pointing at an object and establishing mutual gaze. Kozima et al. (Kozima and Yano, 2001) have designed the robot called 'Infanoid' that can track human faces and ob-

jects with salient colour (T1.1), point to and reach for objects (T2.1), and gaze alternatively between faces and objects (T1.2-4).

Scassellati describes how he intends to accomplish joint attention between the robot and a human, but he mostly concentrates on issues related to attention detection (Scassellati, 1999). So far, only the eye contact has been implemented on the robot 'Cog'. Applied techniques are face detection using ratio templates and eye extraction (T1.1).

Some researchers tackle the development of attention detection, as opposed to simply designing a system capable of doing it. Carlson and Triesch (Carlson and Triesch, 2003) presented a computational model of the emergence of gaze following based on reinforcement learning. They identify a basic set of mechanisms sufficient for the development of this skill. In Nagai et al. (Nagai et al., 2002, Nagai et al., 2003), a learning module learns the correlation between the gaze of a human and an object in the visual field at a certain position. The robot progressively learns to use the human gaze in order to find objects more rapidly. This corresponds to the acquisition of gaze following (T1.2-5).

Several issues concerning the development of attention detection and manipulation have not been addressed yet. How can pointing emerge from grasping behaviour (T2.1)? How does declarative pointing appear (T2.2)? By which process can words replace gestures for drawing attention (T2.3)? On which basis does predication appear (T2.4, T2.5)?

4.2 Models for the emergence of social coordination

Several robotic experiments have emphasised the importance of structured interactions (T3.3) for the development of higher social skills like language acquisition (Breazeal, 2002, Steels and Kaplan, 2001, Steels et al., 2002), but a limited number of works has addressed the problem of how shared interaction routines necessary for coordinating behaviour in joint attention may develop.

Ikegami and Izuka (Ikegami and Iizuka, 2003) use robots in a simulated environment to study turn-taking. Their experiment demonstrates the evolution of a turn-taking behaviour for two robots when a fitness function explicitly favours such a behaviour (T3.1). Andry et al. (Andry et al., 2001) report several experiments where a robot demonstrates immediate imitation for simple motor skills (T3.4) and discuss how simple architectures could account for the emergence of rhythmic interactions (T.3.1) including the possibility of breaking rhythm (T3.2). Imitation has recently been an important topic of investigation (Dautenhahn and Nehaniv, 2002) but only a few works investigate its role for social coordination.

Most of the work remains to be done for this aspect of joint attention. What kind of reward structure must be present so that interaction and entrainment spontaneously emerge (T3.1)? What dynamics lead to the formation of turns during the interaction (T3.1)? How is the structure of new games captured (T3.3)?

4.3 Models for the emergence of the intentional stance

How can a robot start to view the behaviour of another robot as intentional? In research on imitation, some authors have investigated the problem of "what to imitate" in the observed behaviour of another agent (Alissandrakis et al., 2000). They address the issue on how to decompose and recreate an observed behaviour but not how to view it as intentional. The issues of goals and intentional agents are central to research on agent architecture (Dignum and Conte, 1998), but their models do not give insight on the developmental and cognitive mechanisms that lead to the notion of intentionally-directed behaviour. Taking inspiration from animal training techniques, Kaplan et al. showed how a robot could try to model its user's expectations and adapts in order to perform a particular desired behaviour while keeping its general behavioural autonomy (Kaplan et al., 2002). However the robot did not develop its intentional stance by itself.

The development of the intentional stance is probably the most challenging prerequisite that research on joint attention has to investigate. What are the mechanisms or dynamics that enable an agent to identify itself with other agents of the same kind (T4.1)? How can a robot discover the goal-means distinction if these notions are not already explicit in its internal architecture (T4.2-3)? How can it apply this insight to interpret the behaviour of other agents (T4.4)?

5. Conclusions

The development of joint attention between a human and a robot or between two robots depends on the successive appearance of a number of underlying skills. The aim of the present paper is to identify the challenges and to pinpoint what kinds of results are still to be obtained in order to succeed in this goal.

It appears from this survey that one of the most underinvestigated aspects of this problem is the modelling of the mechanisms responsible for the emergence of the intentional stance. Understanding this crucial step in child development would open up the way to the creation of robots with a qualitatively different kind of awareness, making the problems of imitation and social learning easier and ultimately leading to the development of true joint attention.

We believe that this developmental route needs to be understood and that robots are ideal tools to investigate it.

References

Alissandrakis, A., Nehaniv, C. L., and Dautenhahn, K. (2000). Learning how to do things with imitation. In *AAAI Fall Symposium on Learning How to Do Things*, pages 1–6. American Association for Artificial Intelligence.

Andry, P., Gaussier, P., Moga, S., Banquet, J., and Nadel, J. (2001). Learning and communication in imitation: an autonomous robot perspective. *IEEE Transaction on Systems, Man and Cybernetics, Part A : Systems and Humans*, 31(5):431–444.

Baron-Cohen, S. (1997). *Mindblindness: an essay on autism and theory of mind.* MIT Press, Boston, MA, USA.

Breazeal, C. (2002). *Designing sociable robots.* Bradford book - MIT Press.

Butterworth, G. (1995). Origins of mind in perception and action. In Moore, C. and Dunham, P., (Eds.), *Joint attention : its origins and role in development.* Lawrence Erlbaum Associates.

Butterworth, G. and Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, 9:55–72.

Butterworth, G. E. and Cochran, E. (1980). Towards a mechanism of joint visual attention in human infancy. *International Journal of Behavioural Development*, 3:253–272.

Carlson, E. and Triesch, J. (2003). A computational model of the emergence of gaze following. In *Proceedings of the 8th Neural Computation Workshop (NCPW8)*.

Cheney, D. and Seyfarth, R. M. (1990). *How monkeys see the world.* University of Chicago Press.

Dautenhahn, K. and Nehaniv, C. (2002). *Imitation in animals and artifacts.* MIT Press.

Dennett, D. (1987). *The intentional stance.* MIT Press, 5th edition.

Desrochers, S., Morisette, P., and Ricard, M. (1995). Two perspectives on pointing in infancy. In Moore, C. and Dunham, P., (Eds.), *Joint Attention*, pages 85–101. Lawrence Erlbaum Associates.

Dignum, F. and Conte, R. (1998). Intentional agents and goal formation. *Intelligent Agents IV: Agent Theories, Architectures, and Languages.*

Frye, D. (1991). The origins of intention in infancy. In Frye, D. and Moore, C., (Eds.), *Children's theories of mind*, pages 15–38. Lawrence Erlbaum Associates, Hillsdale, NJ.

Hobson, P. (2002). *The craddle of thought.* MacMillan.

Ikegami, T. and Iizuka, H. (2003). Joint attention and dynamics repertoire in coupled dynamical recognizers. In *AISB 03: the Second International Symposium on Imitation in Animals and Artifacts*, pages 125–130.

Imai, M., Ono, T., and Ishiguro, H. (2001). Physical relation and expression: Joint attention for human-robot interaction. In *Proceedings of the 10th IEEE International Workshop on Robot and Human Communication.*

Kaplan, F., Oudeyer, P.-Y., Kubinyi, E., and Miklosi, A. (2002). Robotic clicker training. *Robotics and Autonomous Systems*, 38(3-4):197–206.

Kaye, K. (1982). *The mental and social life of babies.* University of Chicago Press, Chicago.

Kozima, H. and Yano, H. (2001). A robot that learns to communicate with human caregivers. In *First International Workshop on Epigenetic Robotics (Lund, Sweden).*

Meltzoff, A. (1995). Understanding the intentions of others : Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, 31:838–850.

Meltzoff, A. and Gopnick, A. (1993). The role of imitation in understanding persons and developing a theory of mind. In S. Baron-Cohen, H. T.-F. and D.Cohen, (Eds.), *Understanding other minds*, pages 335–366. Oxford University Press, Oxford, England.

Moore, C. and Corkum, V. (1994). Social understanding at the end of the first year of life. *Developmental Review*, 14:349–372.

Nadel, J. (2002). Imitation and imitation recognition : functional use of imitation in preverbal infants and nonverbal children with autism. In A.Meltzoff and Prinz, W., (Eds.), *The imitative mind : development, evolution and brain bases*, pages 42–62. Cambridge University Press.

Nagai, Y., Asada, M., and Hosoda, K. (2002). A developmental approach accelerates learning of joint attention. In *Proceedings of the second international conference of development and learning.*

Nagai, Y., Hosoda, K., Morita, A., and Asada, M. (2003). A constructive model for the development of joint attention. *Connection Science*, 15(4):211–229.

Newson, J. (1979). *The growth of shared understandings between infant and caregiver*, pages 207–222. Cambridge University Press.

Piaget, J. (1952). *The origins of intelligence in children.* Norton, New York.

Scassellati, B. (1999). Imitation and mechanisms of joint attention: A developmental structure for building social skills on a humanoid robot. In *Computation for metaphors, analogy and agents*, Vol 1562 of Springer Lecture Notes in Artificial Intelligence. Springer Verlag.

Schaffer, H. (1977). Early interactive development in studies of mother-infant interaction. In Press, A., (Ed.), *Proceedings of Loch Lomonds Symposium*, pages 3–18, New York.

Siegel, D. (1999). *The developing mind : Toward a neurobiology of interpersonal experience.* The Guildford press, New York, NY.

Steels, L. and Kaplan, F. (2001). Aibo's first words: The social learning of language and meaning. *Evolution of Communication*, 4(1):3–32.

Steels, L., Kaplan, F., McIntyre, A., and Van Looveren, J. (2002). Crucial factors in the origins of word-meaning. In Wray, A., (Ed.), *The Transition to Language*, chapter 12, pages 252–271. Oxford University Press, Oxford, UK.

Tomasello (1995). Joint attention as social cognition. In Moore, C. and Dunham, P., (Eds.), *Joint attention : its origins and role in development.* Lawrence Erlbaum Associates.

Trevarthen, C. (1979). Instincts for human understanding and for cultural cooperation: Development in infancy. In von Cranach, M., Foppa, K., Lepenes, W., and Ploog, D., (Eds.), *Human ethology: Claims and limits of a new discipline.* Cambridge University Press.