NEUROSCIENCES APPLIED TO ACTION INTERPRETATION

Epistemological conflicting perspectives for infant social learning

Emiliano LORIA

In the last decades neurosciences provided so much important contributions to philosophy of mind that nowadays the latter is inconceivable without the former in every topic this philosophical branch deals with. The studies connected to action understanding provided great advances in the field of developmental psychology for what concerns social learning abilities grounded on imitation. All information received by the infants are transmitted through actions. It would be impossible to conceive infant imitation without action interpretation. According to Meltzoff's "like-me" hypothesis, imitation is possible in human infants already at birth in virtue of an identification mechanism with the adults supported by mirror neurons (MNs) based simulation system. However, if we split the types of actions in two general categories, instrumental and communicative actions, we will see, according to an alternative account, how infants modulate differently the comprehension of observed scenarios, depending on whether they are passive observers (in the case of instrumental actions) or actively involved (in the case of communicative actions). Such a recognition of action features seems to be evident through different degrees of motor activation, as ERP techniques applied to infants and young adults revealed. Neuroscientific evidences highlight the crucial role of brain areas connected to motor activation for action interpretation, but at the same time they allow both a bottom-up process and a top-down process interpretation whereby the motor activation is seen as a product of action understanding rather than its determining causal factor. The aim of the present study is to examine such epistemological conflicting perspectives underlying action interpretation, and their repercussions on different social learning theories.

1. Types of actions

In the last decades the studies connected to *action understanding* provided great benefits to pragmatics, and to the comprehension of communicative exchanges among humans. In 2013 Engels and colleagues recognized a *pragmatic turn* in cognitive science, and the consequent achievement of the «action-oriented paradigm»[1]. Action is not a synonymous of movement, its notion «implies that actions (i) are driven by goals and that they can reach these goals or fail to do so; (ii) often [actions] involve some

[1] A.K. ENGEL, A. MAYE, M. KURTHEN, P. KÖNIG, *Where's the action? The pragmatic turn in cognitive science*, in "Trends in Cognitive Sciences", 17, 5, 2013, p. 202.

degree of volitional control; (iii) require planning and decisions among alternatives; (iv) involve prediction or anticipation of an intended outcome; (v) are often, albeit not always, associated with a sense of agency, that is, the agent's conscious awareness of carrying out the particular action and of its goals»[2].

The action-features just outlined above involve an agent who, according to what Donald Davidson said, performs his action «in the light of his beliefs and desires»[3]. However, the purpose of the present study is focused on the phase of *action interpretation* made by infants who see and participate to actions displayed within the surrounding environment in order to achieve cultural information, and to learn social behaviours. From the birth newborns watch actions, and all information received and learnt by the infants are transmitted through actions. An action represents the bound to something recognizable to the infants in the environment, in order to be *understood* and *imitated* appropriately. So far, a good explicative strategy to read correctly the capacity and the target of childhood action interpretation has been to measure through the duration of gaze attention the infants' predictive power addressed to agents and to elements of the actions. Indeed, *action interpretation* entails *action prediction* in terms of expectations. But it is not enough.

Gergely Csibra and György Gergely proposed a criterion for infant understanding of intentional agency, according to which actions are identified by their intrinsic characteristics and by the perspective of observation by which the infant watches them, and the relative degree of her involvement[4]. In other words, action interpretation may be influenced by the role played by the infant observer. In virtue of the general scope of an action, it has been proposed to distinguish *instrumental actions* that are defined as goal-directed actions performed not for itself but to achieve an end[5], and that involve a passive observation by the infant (namely a *third person's view*), from *communicative actions* aimed to inform someone about something and/or about the very intention to

---

[2] Ivi, p. 203.

[3] D. DAVIDSON, *Essays on Actions and Events*, Oxford University Press, Oxford 1980, p. 84.

[4] G. CSIBRA, *Teleological and referential understanding of action in infancy*, in "Philosophical Transactions of the Royal Society of London B: Biological Sciences", 358, 2003, pp. 447-458; ID., *Recognizing communicative intentions in infancy*, in "Mind & Language", 25, 2, 2010, pp. 141–168; G. GERGELY, *Kinds of Agents. The origins of understanding instrumental and communicative agency*, in: U. GOSWAMI (eds.), *The Wiley-Blackwell Handbook of Childhood Cognitive Development*, II edition, Wiley-Blackwell, Oxford 2011, pp. 76-105.

[5] G. CSIBRA, *Teleological and referential understanding of action in infancy*, cit., p. 448.

communicate[6], and that involve directly the infants as the addressee of the communication (namely a *second person's view*)[7]. Recent researches show how such distinction has neuronal correlates, as we will see below[8].

2. Bottom-up processes for action understanding and infant social learning

Two conflicting accounts animate the debate around agency interpretation. On the one hand, there are those theorists who assert that motor mirroring mechanism sustains human action understanding, enabling observers to understand other individuals' actions[9]; on the other hand, other theorists suggest that motor mirroring is the result of action interpretation[10]. Indeed, although there is a general acknowledgment about the activation of the motor system in action processing, it has been alternatively proposed that such activation might be generated by top-down action interpretative processing, thus, motor mirroring is one of the outcomes of action understanding, and not necessarily the causal bottom-up factor as predicted by mirror-based account of action

[6] D. SPERBER, D. WILSON, *Relevance: Communication and Cognition*, II edition, Blackwell, Oxford 1995.

[7] P. JACOB, G. GERGELY, *Reasoning about instrumental and communicative agency in human infancy*, in: J. B. BENSON, F. XU, T. KUSHNIR (eds.), *Rational Constructivism in Cognitive Development*, Academic Press, Cambridge 2012, pp. 59–94.

[8] V. SOUTHGATE, K. BEGUS, *Motor activation during the prediction of non-executable actions in infants*, in "Psychological Science", 24, 6, 2013, pp. 828-835; B. POMIECHOWSKA, G. CSIBRA, *Motor activation during action perception depends on action interpretation*, in "Neuropsychologia", 2017, pp. 1-28.

[9] M. JEANNEROD, *Neural simulation of action: A unifying mechanism for motor cognition*, in "Neuroimage" 14, 1 Pt 2, 2001, S103–S109; M. WILSON, G. KNOBLICH, *The case for motor involvement in perceiving conspecifics*, in "Psychological Bulletin", 131, 3, 2005, pp. 460–473. For a review C. SINIGAGLIA, G. RIZZOLATTI, *Through the looking glass: Self and others* in "Consciousness and cognition", 20, 2011, pp. 64-74. For more up-to-date researches see F. FERRERI, G. RIZZOLATTI (eds.), *Mirror neurons: fundamental discoveries, theoretical perspectives and clinical implications*, in "Philosophical Transaction of the Royal Society B", 369, 1644, 2014.

[10] G. CSIBRA, *Action mirroring and action understanding: An alternative account*, in: P. HAGGARD, Y. ROSETTI, M. KAWATO (eds.), *Sensorimotor foundations of higher cognition. Attention and performance*, XXII, Oxford University Press, Oxford 2007, pp. 435-459; G. HICKOK, *Do mirror neurons subserve action understanding?*, in "Neuroscience letters", 540, 2013, pp. 56-58; A.F. HAMILTON, *The mirror neuron system contributes to social responding*, in "Cortex", 49, 2013, pp. 2957-2959; L. SARTORI, S. BETTI, U. CASTIELLO, *When mirroring is not enough: that is, when only a complementary action will do (the trick)*, in "Neuroreport", 2013, 24, pp. 601-604.

understanding[11]. According to Rizzolatti and colleagues «an action is understood when its observation causes the motor system of the observer to 'resonate'»[12]. Such resonance allows the observer to understand the outcome, and thus finally the goal of the action thanks to the fact the observer knows «its outcomes when he does it»[13]. Visual analysis of the movements alone would be not enough for the comprehension of others' actions that requires instead a process of motor simulation. In brief, the very activation of MNs due to the observed actions would determine the recognition of others' motor intentions by a replication, or better a simulation (without execution) of their motor behaviour. The term *mirroring* refers to such sensorimotor simulation: unconscious and covert imitation of observed movements.

In general, we may say that MNs requests a rigid computation since they depend on perceptual inputs and automatically trigger a simulative response. MNs could serve as the mechanism underlying the human capacity to imitate observed novel actions performed by others. In particular, according to Rizzolatti and Craighero, MNs may generate a new *motor copy* from the observation of a *human* action through a *decomposition-recombination mechanism*: «During learning of new motor patterns by imitation the observed actions are decomposed into elementary motor acts that activate, via mirror mechanism, the corresponding motor representations [...]. Once these motor representations are activated, they are recombined, according to the observed model by the prefrontal cortex»[14].

For what concerns infant learning based on imitation of observed actions, Hunnius and Bekkering claim that when infants looks at others acting, they use the motor system to predict the outcome of the ongoing actions, coming in this way to a comprehension of others' actions «through the repeated observation of actions and the effects

---

[11] B. POMIECHOWSKA, G. CSIBRA, *Motor activation during action perception depends on action interpretation*, in "Neuropsychologia", 2017, p. 17.

[12] G. RIZZOLATTI, L. FOGASSI, V. GALLESE, *Neurophysiological mechanisms underlying the understanding and imitation of action*, in "Nature Review Neuroscience", 2, 2001, p. 661.

[13] V. GALLESE, C. KEYSERS, G. RIZZOLATTI, *A unifying view of the basis of social cognition*, in "Trends in Cognitive Sciences", 12, 2004, p. 496.

[14] G. RIZZOLATTI, L. CRAIGHERO, *The mirror-neuron system*, in "Annual Review Neuroscience", 27, 2004, pp. 182–183.

associated with them»[15]. The background assumption of such account consists in the fact that action experience and statistical regularities of others' behaviours are determining for the development of action understanding.

The psychologist Andy Meltzoff claims more precisely that imitation of conspecifics' actions is based on an *identification* with the (adult) agent[16]. Meltzoff's proposal is termed *"like-me" hypothesis* and it found a strong support in MNs simulation account. According to this theory, at birth children are equipped of «an imitative brain»; the cultural context and social interactions in which infants are immersed together with psychological agents help the early maturation of *an intentional self*, that decodes others' actions through the "like-me" process. The notion "like-me" means that infants see, or rather they recognize others like themselves. «*Human* acts are *especially relevant* to infants because they *look like* the infant *feels* himself to be and because they are events that infants can *intend*»[17]. Meltzoff hypothesizes that imitative process is due to human neural mirroring mechanism, which allows infants to imitate at birth, but not to infer intentions or understand "perception" in others. The intentionality is a later achievement. It has been supposed that the putative infant's self grows up in the second half-year of life accompanied by a sense of intentional agency that is used as a framework for interpreting the intentional actions of others. Briefly, «the self serves the function to understand the actions, goals, and psychological states of others»[18]. The "like-me" hypothesis suggests that the infants map the other from the self; in shorthand, infants recognise when an agent acts as they do: «Through everyday experience infants

[15] S. HUNNIUS, H. BEKKERING, *What are you doing? How active and observational experience shape infants' action understanding*, in "Philosophical Transaction of the Royal Society B", 369, 1644, 2014, p. 1.

[16] A.N. MELTZOFF, *Imitation and other minds: The like-me hypothesis*, in: S. HURLEY, N. CHATER (eds.), *Perspectives on imitation: From neuroscience to social science*, MIT Press, Cambridge 2005, pp. 55-77. Similarly to Meltzoff's account is the model provided by Tomasello and colleagues according to which imitative learning relies on infants' capacity to identify with adults. However, such identification takes place not only at the level of observed motor behaviours, but it origins from psychological motivation to share mental states of others that foster more deeply levels of identification (M. TOMASELLO, M. CARPENTER, J. CALL, T. BEHNE, H. MOLL, *Understanding and sharing intentions: The origins of cultural cognition*, in "Behavioral and Brain Sciences", 28, 5, 2005, pp. 675-691).

[17] Ivi, p. 74.

[18] A.N. MELTZOFF, *'Like me': a foundation for social cognition*, in "Developmental Science", 10, 1, 2007, p. 126. See also A.N. MELTZOFF, *The 'like me' framework for recognizing and becoming an intentional agent*, in "Acta Psychologica", 124, 1, 2007, pp. 26–43.

map the relation between their own bodily states and mental experiences»[19]. In virtue of action observations, infants are able to project onto others what they felt bodily and registered, but they can do such projection only if they *see* an *equivalence* between their acts and those of others[20]. Infant reasoning can be expressed under the form: "your face is similar to my face; my hand is similar to your hand" and so on; thus, infant can imitate human acts because she identifies the corresponding body parts.

However, the hypothesis grounded on *motor copy* phenomenon entails that without simulation of observed movements in motor system, individuals should have difficulty interpreting observed actions. Vannuscorps and Caramazza challenged this view across eight sensitive experiments in which «individuals born with absent or severely shortened upper limbs (upper limb dysplasia), despite some variability, could perceive, anticipate, predict, comprehend, and memorize upper limb actions, which they cannot simulate, as efficiently as typically developed participants»[21]. The authors point out that their results are based on the same experimental materials and procedures used in those studies that have been interpreted as the best evidences in favour of motor simulation theories[22]. Therefore, future research should investigate deeply how (visuo-) perceptual and cognitive system encode information of body part movements, and in general, support interpretation of actions[23].

3. Top-down processing for action interpretation: teleological stance

Now I would like to turn the attention towards an alternative account that has been judged more parsimonious under several points of view; first of all because it doesn't predict the precocious development of a strong self enable to make introspection and

---

[19] Ivi, p. 56.

[20] P.J. MARSHALL, A.N. MELTZOFF, *Neural mirroring mechanism and imitation in human infants*, in "Philosophical Transaction of the Royal Society B", 369,1644, 2014, p. 2.

[21] G. VANNUSCORPS, A. CAMARAZZA, *Typical action perception and interpretation without motor simulation*, in "Proceedings of the National Academy of Sciences", 113, 2016, p. 89.

[22] S. BOSBACH, J. COLE, W. PRINZ, W, G. KNOBLICH, *Inferring another's expectation from action: The role of peripheral sensation*, in "Nature Neuroscience", 8, 10, 2005, pp. 1295–1297; M. WILSON, J. LANCASTER, K. EMMOREY, *Representational momentum for the human body: Awkwardness matters, experience does not*, in "Cognition", 116, 2, 2010, pp. 242–250

[23] An attempt in this direction is constituted by C. BECCHIO, A. KOULAB, C. ANSUINIA, A. CAVALLO, *Seeing mental states: An experimental strategy for measuring the observability of other minds*, in "Physics of Life Review", in press.

projections of mental states to others. In particular, it has been proposed by György Gergely and Gergely Csibra that infants are able to «represent, explain and predict goal directed actions by applying [...] the "teleological stance"»[24]. The notion *teleological stance* is to intend an interpretational strategy that construes events in terms of purposes. As Csibra acknowledged, it is akin to *intentional stance* descripted by the philosopher Daniel Dennett[25] in virtue of two common features: a) it is a bias and not an explicit inferential system; b) it predicts an ascription of *rationality principle* made by the infants to agent's mental state (in the case of mentalistic stance), and to actions itself in terms of *efficiency* (in the case of teleological stance). Therefore, the two stances are different because the teleological stance «does not attribute mental states to the agents»[26]. Teleology provides an explanatory relation oriented to the scopes and not to the causes among three elements of observed and future reality: «the *action*, the (future) *goal state*, and the current *situational constraints*»[27]. In brief, teleological stance entails: 1) representation of the goal, 2) representation of the physical constraints present in the action's scenario, and 3) representation of the means for achieving the goal (given the environmental constraints). These infant's representations are supported by the principle of efficiency strongly bound to the features of the action itself. As Susan Carey well summarized, there is not «any relation between an agent and a desired state explicitly represented. Rather, it is the action itself that is represented as goal-directed»[28].

3.1 Experiments in favour of teleological interpretation of instrumental actions

In the last ten years Victoria Southgate and cooperators provided the most decisive contributions to teleological account sustaining the proposal that instrumental action interpretation is not driven directly by motor activation. Southgate and colleagues analyzed the application of efficiency principle in 6-8 month-old infants who observed biomechanical impossible events, measuring the number of action steps the agent (a

[24] G. GERGELY, G. CSIBRA, *Teleological reasoning in infancy: the naïve theory of rational action*, in "Trends in Cognitive Science", 7, 7, 2003, p. 289.
[25] D. DENNETT, *The Intentional Stance*, MIT Press, Cambridge 1987.
[26] G. CSIBRA, *Teleological and referential understanding of action in infancy*, cit., p. 448.
[27] G. Gergely G., G. Csibra, Op. cit., p. 289.
[28] S. CAREY, *The origins of concepts*, Oxford UP, Oxford 2009.

human arm showed in a videotape) performed for achieving its goal. The assumption was that the less number of steps it took to achieve a goal, the more efficient was the action. The aim was to test whether infants recognized actions as goal-directed on the basis of their experience or «on the recognition of a specific event structure»[29]. The researchers showed how infants extended goal attribution also to a human arm that reached an object overcoming obstacles with snake-movements. They concluded that the goal cannot be understood in virtue of infants' experience. If the actions were understood for the fact that young observers directly mirror «the observed action onto their own motor system», it would follow that «only action that observers can themselves perform w[ould] be able to be simulated»[30].

Few years later, together with Mikolaj Hernik, Southgate demonstrated that 9 month-old infants comprehend the structure of an oriented-goal action regardless of the agent's preferences. Indeed, they showed that 9-month-olds expected the agent (represented by a red cube) to continue acting towards the previous goal (reaching a blue cylinder behind a barrier) even if additional choice-options (another brown cylinder) became available. The results provided by Hernik and Southgate indicated that «there was no preference-related evidence. [Therefore, they] concluded that infants do not need to know about the agent's preferences in order to form expectations about its goal-directed actions»[31]. The use of objects instead of human agents may further indicate that there is no possible identification and simulation processing. This theoretical and

---

[29] V. SOUTHGATE V., M.H. JOHNSON, G. CSIBRA, *Infants attribute goals even to biomechanically impossible actions*, in "Cognition", 107, 2008, p. 1059.

[30] Ivi, p. 1060.

[31] M. HERNIK, V. SOUTHGATE, *Nine-months-old infants do not need to know what the agent prefers in order to reason about its goals: on the role of preference and persistence in infants' goal-attribution*, in "Developmental Science", 15, 5, 2012, p. 714. Recently, also Liu and Spelke accepted the evidences according to which very young infants expect agents to go directly to their targets when no obstacles stop their paths, or anyway to move along the least costly path given environmental constraints. On these bases, they tested sixties 6-month-olds across three experiments with novel, curvilinear action trajectory, and they found that «infants expected minimally costly action when presented with a novel constraint, and extended this expectation to agents who had previously acted inefficiently» (S. LIU, E.S. SPELKE, *Six-month-old infants expect agents to minimize the cost of their actions*, in "Cognition", 160, 2017, p. 35). See also Scott and Baillargeon, who tested the application of rationality principle too, but without involving infrequent, impossible or odd actions (R.M. SCOTT, R. BAILLARGEON, *Do infants really expect agents to act efficiently? A critical test of the Rationality Principle*, in "Psychological Science", 24, 4, 2013, pp. 466-474).

methodological approach goes against Woodward's interpretation according to which young infants can ascribe goals only to those agents who appear perceptually very similar to their own bodily aspect, and whose body-movements can be mapped onto infants' own motor repertoire[32].

Using neural indication of sensorimotor-cortex activation measured by electroencephalography (EEG), Southgate and Begus demonstrated «that 9-month-old infants recruit their motor system whenever a context suggests an impending action, but that this recruitment is not dependent on being able to match the observed action with a corresponding motor representation»[33]. Sensorimotor cortex is activated during action execution and observation both in adulthood and in childhood, as it has been discovered measuring alpha rhythm (or mu suppression), whose oscillations frequency are around 8-13 Hz in adults and 6-9 Hz in infants[34]. Southgate and Begus' results support the hypothesis that motor activation must be interpreted as the *result* of goal identification, rather than the *cause*.

4. Communicative action interpretation

Communicative actions are grounded on mutual interaction conveyed by particular signals. From birth human beings are sensitive to the presence of eyes. Farroni, Csibra, Simion and Johnson showed how newborns looked longer at faces that engaged them in mutual gaze when compared to averted gaze. In particular, they found that when neonates have the possibility to choose between photographs of faces looking directly at them or looking in another direction, 3-day-olds prefer to look at the face that appears to establish eye contact with them. According to Farroni and colleagues the preferential attention toward direct-gaze faces provide a clear evidence that «human newborns are born prepared to detect socially relevant information»[35].

---

[32] A. WOODWARD, *Infants selectively encoded the goal object of an actor's reach*, in "Cognition", 69, 1998, pp. 1-34.

[33] V. SOUTHGATE, K. BEGUS, *Motor activation during the prediction of non-executable actions in infants*, in "Psychological Science", 24, 6, 2013, p. 828.

[34] Cfr. J.A. PINEDA, *The functional significance of mu rhythms: translating "seeing" and "hearing" into "doing"*, in "Brain Research Reviews", 50, 1, 2005, pp. 57-68.

[35] T. FARRONI, G. CSIBRA, F. SIMION, M.H. JOHNSON, *Eye contact detection in humans from birth*, in "PNAS", 99, 14, 2002, p. 9602.

In this way infants are guided towards the object-target by the adult's gaze. Following eye gaze and shifting attention to the direction of an adult's eye gaze represent fundamental abilities for the maturation of joint attention, which are crucial for imitative learning in infants, and determining also for other cognitive processes like language development[36]. In former times, it has been sustained that only after 9 months of age infants are able to engage in joint attention[37]. More recent studies challenged this view. At this regard, Hoehl and colleagues tested seventeen 4-month-old infants showing them «static photographs of faces with eye gaze averted to the left or right side, with one object presented near to the face»[38]. The results of their experiment suggest that infants process object-directed eye gaze faster than non-object-directed gaze. In few words, the 4-month-olds processing of objects differently depends on whether the objects were cued by the direction of adult's eye gaze. Information transmitted through these kind of social interactions are better processed by so young infants than information provided by non-social sources. This study induces to think that infants see and look for other people around them not only as the source of nutritive and emotional cares, but also as the source of information about the surrounding environment, and as the guides who allow them to learn about natural and artefact objects. We are assuming, in this way, a sort of *innate epistemic attitude* that projects infants toward the world like spontaneous searchers, or hunters of information and knowledge.

---

[36] R. BROOKS, A.N. MELTZOFF, *The development of gaze following and its relation to language*, in "Developmental Science", 8, 2005, pp. 535-543; T. STRIANO T., STAHL, *Sensitivity to triadic attention in early infancy*, in "Developmental Science", 8, 2005, pp. 333-343; T. STRIANO, X. CHEN, A. CLEVELAND, S. BRADSHAW, *Joint attention social cues influence infant learning*, in "European Journal of Developmental Psychology", 3, 3, 2006, pp. 289-299; M. TOMASELLO, *Joint attention as social cognition*, in: C. MOORE, P. DUNHAM (eds.), *Joint attention: Its origin and role in development*, Lawrence Erlbaum, Hillsdale 1995, pp. 103-130.

[37] M. CARPENTER, K. NAGELL, M. TOMASELLO, *Social cognition, joint attention, and communicative competence from 9 to 15 months of age*, in "Monographs of the Society for Research in Child Development", 63, 1998, pp. 1–174.

[38] S. HOEHL, V. REID, J. MOONEY, T. STRIANO, *What are you looking at? Infants' neural processing of an adult's object-directed eye gaze*, in "Developmental Science", 11, 1, 2008, p. 11.

4.1. ERP technique for investigating the processing of referent objects in 5-month-old infants

Event Related Potentials (ERP) are the most appropriate techniques for investigating cognitive processes in infants because they result less invasive than fMRI. ERP allows to measure processing information between a stimulus inducted and a response, and it is one of the most reliable method adopted in cognitive neuroscience to study neuronal correlates of perceptual and cognitive activity. ERP is constituted by underlying components which are related to waveform mark that has a series of positive and negative voltage deflection. Parise and colleagues focused on measuring the so called "middle latency negative central component (Nc)"[39] with the intent to investigate the neural effect of joint attention on object processing during live interactions between adults and 5-month-old infants, who were divided in two blocks. Each block included a pretest phase in which a female experimenter uttered short phrases like: "Oh nice!", "So many colours", with friendly face expressions, smiles, and positive tone of voice (that are all social communicative signals), while she was sitting in front of the infant and turned her head alternately from the baby to the screen, where three different objects were presented. The only difference between the two blocks was in the kind of familiarisation, i.e. the nature of eye contact that the experimenter engaged with the infant during the pretest phase: in "joint attention condition" there was a mutual eye contact between the experimenter and the infant, while in "no-joint attention condition" the experimenter did not look directly at eyes' infant. The interpretation provided highlights the strong influence on infant referent object processing of a period of mutual eye contact. Parise and his co-operators noticed that there is a great similarity in Nc ERP component between 5 and 9-month-olds, suggesting «the possibility that the neural systems subserving the extraction of information during social interactions are the same at both ages»[40].

---

[39] Nc appears approximately 300-700 ms after stimulus onset and it is most prominent at fronto-central electrodes (S.J. WEBB, J.D. LONG, C.A. NELSON, *A longitudinal investigation of visual event-related potentials in the first year of life*, in "Developmental Science", 8, 2005, pp. 605-616). E. PARISE, V.M. REID, M. STETS, T. STRIANO, *Direct eye contact influences the neural processing of objects in 5-month-old infants*, in "Social Neuroscience", 3, 2, 2008, pp. 141-150.

[40] Ivi, p. 148. This is congruent also with Striano and colleagues who found that in 9-month-old infants Nc increased in amplitude during the processing of objects during joint attention interaction

4.2. Infants recognise to be the addressee of the message

The question that arises is whether 5-month-olds at least are able to understand that targeting an object through gaze shift, and also other accompanying ostensive stimuli, implies a message *for* themselves, i.e. whether 5-month-olds, or younger infants, feel being addressed of the message. Csibra and Volein's finding shows how infants of just 8 months of age *expect a referent object for a gaze shift*. In their experiment infants expected an object to be at a location indicated by someone's gaze even if the object is hidden by an occluder. This has an important entailment, because it could mean that infants can infer that the mutual eye contact and the following gaze shift imply a message transmitted by the communicator and such a message is the referent[41]. Around the fifth month of age, in addition to following gaze shift, infants start to learn other ostensive signals such as the own name[42]. Parise, Friederici and Striano used ERP methodology to compare neural response in infants when they hear their own name or stranger's name whereas they look at novel objects. The results of this finding indicate that «hearing her own name prepares the infant to receive new relevant information»[43].

In order to overcome experimental constraints that may influence and alter the results, Lloy-Fox and colleagues settled their experiments in more ecological environment[44]. They created two naturalistic social interactions in which ostensive cues (eye contact and IDS) were presented live and for a long duration. The researchers used fNIRS[45] (Near-Infrared Spectroscopy) techniques for recording cortical responses

---

with adults compared to a non-joint attention condition (T. STRIANO, V.M. REID, S. HOEHL, *Neural mechanisms of joint attention in infancy*, in "European Journal of Neuroscience", 23, 2006, pp. 2819-2823).

[41] G. CSIBRA, A. VOLEIN, *Infants can infer the presence of hidden objects from referential gaze information*, in "British Journal of Developmental Psychology", 26, 2008, pp. 1–11.

[42] D.R. MANDEL, P.W. JUSCZYK, D.B. PISONI, *Infants' recognition of the sound patterns of their own names*, in "Psychological Science", 6, 1995, pp. 314–317.

[43] E. PARISE, A.D. FRIEDERICI, T. STRIANO, *"Did you call me?" 5-month-old infants own name guides their attention*, in "PLoS One", 5, 2010, e14208.

[44] S. LLOYD-FOX, B. SZÉPLAKI-KÖLLŐD, J. YIN, G. CSIBRA, *Are you talking to me? Neural activations in 6-month-old infants in response to being addressed during natural interactions*, in "Cortex", 70, 2015, pp. 35-48.

[45] «The NIRS method relies on the optical determination of changes in hemoglobin concentrations in cerebral cortex which result from increased regional cerebral blood flow» (T. GROSSMANN, E.

to the communicative stimuli presented simultaneously. The study was conducted with pairs of infants seated on their parents' lap. Six-month-old infants interacted with a female experimenter who exchanged mutual eye contact with each baby for 15 seconds, meanwhile she uttered nursery rhymes (in infant-direct-speech (IDS) modality) accompanied by hand movements. In another experimental condition, she interacted in a one-one communication in which the combination of the infant direct gaze (IDG) and IDS was compared with IDG and the adult direct speech (ADS) modality. Confronting the results of the two conditions, Lloyd-Fox and colleagues found that direct gaze performed by the experimenter «increased neural responses to the multimodal communicative actions (speech plus gestures)»[46]. Therefore, they hypothesized that the detection of direct gaze affected significantly the processing of the accompanying communicative signals, i.e. speech and hand movements, recording a strong activation in inferior frontal and temporal regions in both hemispheres. This occurred only when the infants were the *addresses* of the direct experimenter's gaze, otherwise when it was not directed to them or when the infants were performed in ADS modality, they didn't process IDS and hand gestures, at least in the brain regions analysed. The authors invite to remain prudent about the correct interpretation to provide in order to explain neural activation in the cortical areas highlighted by their finding. They doubt that the underlying mechanism is identifiable with mindreading system, because they did not find remarkable modulations of the prefrontal cortex (PFC) in response to ostensive cues. This fact is likely due to the methodology adopted by Lloyd-Fox and colleagues, that is the most recommendable because it reproduces ecological and naturalistic conditions.

Their results appear to contradict, but only in part, the conclusions provided by Parise and Csibra[47] to their findings on multimodal ostensive signals. Indeed, Parise and Csibra tested 5-month-olds using ERPs and gamma-band event-related oscillations with static visual stimuli on a computer screen (female face with close eyes, open eyes with direct gaze, open eyes with averted gaze) combined with IDS/ADS (but using only

---

PARISE, A.D. FRIEDERICI, *The detection of communicative signals directed at the self in infant prefrontal cortex*, in "Frontiers in Human Neuroscience", 4, 12, 2010, p. 2).

[46] Ivi, p. 7.

[47] E. PARISE, G. CSIBRA, *Neural responses to multimodal ostensive signals in 5-monthold infants*, in "PloS one", 8, 2013, e72360.

one word in the two different intonations). They found same activations in the frontal cortex independently on the presence of one stimulus that elicited an equal response to multiple ostensive signals[48]. However, Parise and Csibra underline, and Lloyd-Fox and colleagues confirm, that the main element that we can figure out in virtue of the statistical results provided consists in what is termed «*obligatory response*» shown by infants addressed by ostensive communication. In authors' terms: «ostensive signals obligatorily indicate to young infants that communication is directed to them»[49].

Once understood that the message is addressed to them, infants are biased to trigger referential expectation that represents a crucial and preparatory moment to approach the informative content of the message. A recent finding seems to highlight the remarkable sensitivity demonstrated by just 4-month-olds on the comprehension of referential nature of human speech in combination with direct eye gaze. It is the case of Marno and colleagues' study that tested 4-month-old infants to investigate the presence or the absence of referential expectations in infants hearing human speech compared to other auditory stimuli (and silence too) in presence of both direct eye contact and object-directed gaze of the speaker[50]. Their striking results highlighted that when infants were looking at a female face, who uttered ostensive speech, they appeared well prepared to find some visual referents of the words, as indicated by their faster orienting towards the visual objects she targeted at the end of the speech. This study confirms that very young infants recognise both the *ostensive value of speech*, when it is accompanying with other stronger ostensive signal like mutual eye contact, both the *informative value of speech* independently on the understanding of the meaning of the words.

---

[48] This may be in contrast with the peculiar status of eye detection that modulates and enhances the reception of the other ostensive stimuli, but we can reasonably argue that this fact could be due to the modalities of stimuli presentation which were rapid, and a long interaction with the sources of ostensive cues did not occur.

[49] E. Parise, G. Csibra, *Neural responses to multimodal ostensive signals in 5-monthold infants,* cit., e72360 (p.1).

[50] H. Marno, T. Farroni, Y.V. Dos Santos, M. Ekramnia, M. Nespor, J. Mehler, *Can you see what I am talking about? Human speech triggers referential expectation in four-month-old infants*, in "Scientific Report", 5, 13594, 2015, pp. 1-10.

4.3. Social learning mechanism based on ostensive communication

The recognition of a teaching context in virtue of ostensive signals encoding constitutes the basis of the social learning theory proposed more than ten years ago by Csibra and Gergely. They suppose the existence of the so called "natural pedagogy", a specialized human-specific cognitive system dedicated to cultural learning, grounded on ostensive (verbal and non-verbal) communication expressed by eye-contact, smiles, particular vocalizations, IDS, contingent reactivity, deictic gestures, joint attention[51]. In this view «preverbal human infants are prepared to receive culturally relevant knowledge from benevolent adults who are, in turn, spontaneously inclined to provide it»[52]. The knowledge domains transferred pedagogically regard objects functions, social norms, first simple words, practical know-how. The very ostensive manifestations allow infants the recognition of a potential teaching context. The encoding capacities predispose infants to achieve information from the communicative source of knowledge which by ostensive signals orient infants towards the relevant aspects of the referent. The pedagogical stance construed by the adult-infant communicative relation triggers three biases in infants' mind, or in other terms, three different inferential processes (called "assumptions") about the referential object of the informative transmission.

The crux of natural pedagogy theory is the fact to be object-centered account in virtue of the infant capacity to encode ostensive signals, i.e. to catch their referential nature without inferring psychological intentions about the informative source. Therefore, according to natural pedagogy theory, infants don't need to do *any identification* with the adult teacher[53], because ostensive communication triggers in infants' mind an interpretational switch oriented towards the relevant aspects of the

---

[51] G. CSIBRA, G. GERGELY, *Social learning and social cognition: The case for pedagogy*, in: Y. MUNAKATA, M.H. JOHNSON (eds.), *Processes of change in brain and cognitive development. Attention and performance*, XXI, Oxford University Press, Oxford 2006, pp. 249-274; G. GERGELY, G. CSIBRA, *Sylvia's recipe: The role of imitation and pedagogy in the transmission of cultural knowledge*, in: N.J. ENFIELD, S.C. LEVENSON (eds.), *Roots of Human Sociality: Culture, Cognition, and Human Interaction*, Berg Publishers, Oxford 2006, pp. 229-255.

[52] G. CSIBRA, G. GERGELY, *Natural pedagogy as evolutionary adaptation*, in "Philosophical Transaction of the Royal Society B", 346, 2011, p. 1154.

[53] G. GERGELY, H. BEKKERING, I. KIRALY, *Rational imitation in preverbal infants*, in "Nature", 415, 2002, p. 755; K. EGYED, I. KIRÁLY, G. GERGELY, *Communicating shared knowledge in infancy*, in "Psychological Science", 24, 7, 2013, pp. 1348–135.

referent. On the basis of such early communicative action interpretation skill, infants can learn many aspects of the surrounding social world through others without learning about their minds[54].

5. Comparing neuronal correlates for instrumental and communicative actions

Communicative and instrumental actions seem to be processed by two different cognitive channels supporting the corresponding interpretation. The question concerns not only the quality and the shape of body movements observed, e.g. raising an arm to touch, or push or to point something with the finger representing three different physical, perceptual and intentional acts, but also the contextual multimodal features in which action observed occurs.

Pomiechowska and Csibra compared «the levels of motor activity across instrumental and communicative actions», assuming that «the selected interpretation should directly modulate the level of motor activation»[55]. They expected a different degree of activation (revealed by different mu suppression marks) depending on the kind of action observed. The participants at the experiments (mean age 22 years with a range from 18 to 27 years) were presented with speech (e.g.: "Look") or a matched pure tone previous to the onset of the action. The idea was that the presence of many ostensive signals would clearly «bias participants to suppress the instrumental interpretation of the observed acts» and foster them to construe the action as referential.

> Under this assumption, identical grasping actions should be interpreted differently depending on the preceding sound stimuli, thus affecting levels of motor activation. Specifically, if the presence of speech changes the interpretation of grasping from instrumental to referential, less mu suppression should be recorded in the presence of speech than in the presence of a pure tone.[56]

In video clips the participants could see a human hand that grasped an object, reached for an object without grasping it, and did point to an object with the extended

---

[54] G. GERGELY, *Learning "about" versus learning "from" other minds: Natural pedagogy and its implications*, in: P. CARRUTHERS, S. LAURENCE, S. STICH (eds.), *The innate mind. Vol. 3: foundations and the future*, Oxford UP, Oxford 2007, p. 193.

[55] B. POMIECHOWSKA, G. CSIBRA, *Motor activation during the prediction of non-executable actions in infants,* cit., p. 6.

[56] Ivi, p. 8.

index finger. The study combined communicative and non-communicative sounds (pure tone vs. speech) during the different kinds of actions. In a particular experimental condition (Experiment 1), «one type of action was paired with one type of sound (pure tone – grasping, pure tone – reaching, pure tone – pointing, speech – grasping, speech – reaching, speech – pointing)». While in another experimental condition (Experiment 2) grasping, reaching, and pointing actions were paired only by sound[57]. The results indicate that:

> When the context suggested a referential interpretation of the observed action, either due to the semantics of the witnessed gestures (i.e., pointing) or to the presence of speech, there was no sign of significant motor activation. This pattern of results suggests that *action interpretation is not dependent on the observer's motor system* and that the presence of subsequent motor activation is conditioned by this interpretation: *only conceiving of an action as instrumental*, but not as referential, leads to *the recruitment of sensorimotor cortices* during action observation [58]

Another very important result of this study consists in the fact that the presence of communicative speech signal before the onset of the instrumental action of grasping modifies the interpretation of the observed action. Therefore, the communicative context influences the expectation for referential signals rather than for goals oriented action. In this way, the object exchanges its role: from the *target* of an instrumental action it becomes the *referent* of a communicative act. According to these results, Pomiechowska and Csibra suggest that also with goal-oriented action motor activation seems to be «the result of, rather than a contributor to, goal identification»[59].

We must remember that both instrumental and communicative acts are oriented toward an object. Therefore, it is reasonable to wonder if also object representation is influenced by top-down action interpretation itself. According to Yoon and colleagues preverbal infants (9-month-old) remember the location of a toy better in an instrumental context than in a referential context, in which they tend to store the information about its appearance. Indeed, they found «that in a communicative context, infants devoted their limited memory resources to encoding the identity of novel objects at the expense of encoding their location, which is preferentially retained in non-communicative

[57] Ivi, p. 11.
[58] Ivi, p. 16 [my italics].
[59] Ivi, p. 20.

contexts»[60]. As shown also by Wurm and Schubotz, contextual information modulates action recognition at different levels of processing. They investigated the ability to recognise hardly identifiable actions that occurred in compatible, incompatible, and neutral contextual settings. Their findings suggest that contextual information is «effectively exploited during action observation, in particular when visual information about the action itself is sparse»[61].

We may conclude that infants begin to interpret communicative actions in virtue of encoding systems that allow them to understand the referential nature of intentional communicative signals, but at the same time the modality and the context in which communicative acts occur influence the very interpretation. Therefore, in the second person perspective, the infant's approach to the referent is strongly biased by the style and the features of communication. Yoon and colleagues' experiment and Pomiechowska and Csibra's study combined together seem to strengthen the view whereby the intrinsic instrumental and communicative acts features, and the degree of infant's attendance (passive in the case of third person view, active in second person view) influence directly action understanding that even supports two different neural pathways. Even if we accept that action interpretation is not driven by action-simulation, the question remains open about which neural substrates underlying top-down processes in action interpretation. Indeed, it has been not clarified yet how multimodal perceptual components are processed to provide an adequate action interpretation without the immediate causal involvement of MNs.

---

[60] J.M. YOON, M.H. JOHNSON, G. CSIBRA, *Communication-induced memory biases in preverbal infants*, in "Proceedings of the National Academy of Sciences", 105, 2008, p. 13690.

[61] M.F. WURM, R.I. SCHUBOTZ, *What's she doing in the kitchen? Context helps when actions are hard to recognize*, in "Psychonomic Bulletin & Review", 24, 2, 2016, p. 503.

6. Conclusions

Up to now, at epistemological level, we cannot claim with certainty which is the direction of the causal arrow for action interpretation processing and relative behavioural responses (among which we can include immediate and delay imitation processing, which is the hinge of learning development). Do bottom-up processes (likely interlocked with MNs) determine action understanding? Or are top-down processes guided by the context and the structure of the actions to shape such a comprehension and to activate motor response? Neuroscientific techniques have demonstrated to be flexible and comply human ecological environment and infant needs. A more robust theoretical framework underlying action interpretation will allow to overcome the limits and the critical points of social learning theories that see imitation as the fundamental process able to guarantee cultural transmission among generations, highlighting the social cognitive skills of infants from birth. Neurosciences will be needful to provide adequate support for whatever theoretical perspective will emerge more clearly in the future researches, and maybe it will be possible to conceive a model that might include and integrate both the accounts in a virtuous circularity.