2018 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)
Tokyo, Japan, September 16-20, 2018

Ungrounding symbols in language development:
implications for modeling emergent symbolic communication in artificial systems

Joanna Rączaszek-Leonardi
Faculty of Psychology
University of Warsaw
Warsaw, Poland
raczasze@psych.uw.edu.pl

Terrence W. Deacon
Department of Anthropology
University of California at Berkeley
Berkeley, USA
deacon@berkeley.edu

***Abstract*— The relation of symbolic cognition to embodied and situated bodily dynamics remains one of the hardest problems in the contemporary cognitive sciences. In this paper we show that one of the possible factors contributing to this difficulty is the way the problem is posed. Basing on the theoretical frameworks of cognitive semiotics, ecological psychology and dynamical systems we point to an alternative way of formulating the problem and show how it suggests possible novel solutions. We illustrate the usefulness of this theoretical change in the domain of language development and draw conclusions for computational models of the emergence of symbols in natural cognition and communication as well as in artificial systems.**

Keywords*— Human-human and human-robot interaction and communication; Language acquisition; Epistemological Foundations and Philosophical Issues

I. Introduction

Despite years of theoretical, experimental and modeling work, the relation of symbolic cognition to embodied and situated bodily dynamics remains one of the hardest problems of the contemporary cognitive sciences. The problem permeates numerous cognitive domains, from most obvious ones such as explaining human natural language processing, interpretation of language in various domains of NLP to automatic linguistic description of visual scenes, linguistic control of robots' behavior and human-robot interaction. In all those domains, and many more, the gist of the problem is how to relate the seemingly abstract, conventional and formal entities, called symbols, to real physical, continuous dynamics of action, interaction and/or interpretation.

For many years, since at least the 1970s and early 80s, the problem was considered mainly from one specific perspective. Namely, it was asked how symbols are endowed with meaning, how can they refer to something other than other symbols. In other words, the problem was posed as a "symbol grounding problem" [1, 2, 3, 4, 5]. This manner of posing the question brings forward some aspects of the problem while obfuscating others, which may make important principles and processes difficult to recognize.

First the symbol grounding problem assumes that symbols exist and, fashioning them on the alleged properties of linguistic or mathematical symbols, accepts their nature as abstract, conventional and formally related to each other. Therefore, the very question presupposes them as ungrounded entities. It becomes thus more urgent to ask about their grounding than to ask about how the symbolic properties could have come about in the first place and which principles of cognition and which processes could have produced them. Thus in the "grounding problem", the very nature of the processes of abstraction, conventionalization and formalization are not inquired into.

Second, the search for the meaning of signs in most domains of the cognitive sciences seems to be limited to "symbolic meaning". This approach tends to overlook the fact that non-symbolic informational structures also play a vital role in regulating the relationships of the organism to its environment, relationships among organisms, and – crucially – the emergence the symbols themselves. So the problem is formulated as a "symbol grounding problem" and not a "sign grounding problem." This obfuscates the possible mediation of non-symbolic meaningful forms in the emergence of symbolic systems.

The goal of this paper is to change the way we pose the problem of the relation between symbols and dynamics. This way of formulating the question is inspired by a theoretical framework developed by Deacon (1997) and by research on language development, where it is particularly clear that initially all meaningful behaviors, linguistic forms included, appear in rich interactive, dynamical contexts. Just as any other stimuli or gestures, they are fully and causally embedded in rich meaningful dynamics, and, from very early on, have power of controlling these dynamics. In such contexts it becomes evident that the real problem is not the grounding of symbols but rather explaining the mystery of how such an embedded, embodied and situated use of signs can ever (at least partially) become liberated from the immediate reliance on the on-line events, thus, how they become, at least partially, "ungrounded". How do everyday interactions can ever give rise to the apparently abstract, conventional and formal symbols? And in the process how do they maintain their fundamental grounding and remain causal controls on interactive dynamics.

In what follows, Section II presents the theoretical framework for symbol ungrounding, which uses the well-known model of symbol emergence by Deacon [6]. In Section III, using the domain of language development, we demonstrate the usefulness of the model in 1) identifying the processes that make signs meaningful in development and 2) identifying the possible paths to ungrounding symbols. We provide multiple examples of both processes based on microanalyses of real parent-infant interactions [7, 8, 9],

This work was funded by the Beethoven UMO-2014/15/G/HS1/04536 grant to the first author.

paying particular attention to the structures present in the social environment of the child that scaffold development. Section IV briefly reviews some of the attempts at modeling language evolution with grounded symbols, points out how the models could be enriched by the microanalyses of early social interaction and postulates general desiderata for such models. In the Conclusions we summarize general guidelines for such models and note that the 'symbol ungrounding' approach and the reliance on microanalyses of language development data could be particularly valuable for the community of epigenetic robotics, which appreciates the ontogenetic processes shaping the experience of intelligent agents, and which often demonstrates ecological sensitivity to structures provided by both physical and social environment.

II. SEMIOTIC INFRASTRUCTURE OF SYMBOLIC REFERENCE

Deacon's 1997 model originated as a theory of the semiotic hierarchy that underlies symbolic reference. Following Peirce's semiotic theory, symbols are seen as requiring an infrastructure of simpler semiotic relationships. These more basic relationships include iconic signs, like pictures, in which the sign vehicle and its referent share formal properties; and indexical signs, like symptoms, in which the sign vehicle and its referent are physically or habitually linked. Because icons and indices share properties with what they refer to, they are in this sense "grounded" signs. In contrast, symbolic sign vehicles, typically lack properties shared with their referents and by virtue of this lack of grounding are able to be combined and manipulated in ways that makes possible nearly unrestricted referential relationships. According to Peirce, "Symbols grow. They come into being by development out of other signs" [10]. Therefore, it should be possible to trace the emergence of symbolic forms of reference from prior icons and indices, which have more obvious – isomorphic or causal - relationships with the social-pragmatic dynamics in which they are immersed.

Deacon initially analyzed this process as it was exemplified in a study involving two chimpanzees and a highly simplified 6-lexigram (computer keys with arbitrary marks on them) symbol system. In this well-known study by Savage-Rumbaugh and colleagues [11, 12] chimps had to learn to combine lexigrams for two food and two drink items in specific combinations with the appropriate delivery lexigram (glossed as 'pour' and 'give'). The chimps easily learned the indexical relation between a lexigram and the correlated food reward, but it was particularly difficult to get them to shift to using specific lexigram compositionality to refer to a specific food-action relation (e.g. pour juice). Only by foregrounding the lexigram-lexigram "agreement" rule and systematically extinguishing all other combinations was it possible to get the chimps to abandon simple indexicality and pay attention to the implicit abstract iconicity between lexigram-lexigram and food-action relations. Lexigrams that were initially grounded indexically to individual items in the chimps' world, thus became (also indexically) grounded in relation to other lexigrams. In this respect they became "doubly grounded" [13]. See Figure 1 for details of the stages of this process. This double grounding not only allows for single lexigrams to be used indexically, but also for the relations among them to also be referential – i.e., meaningfully

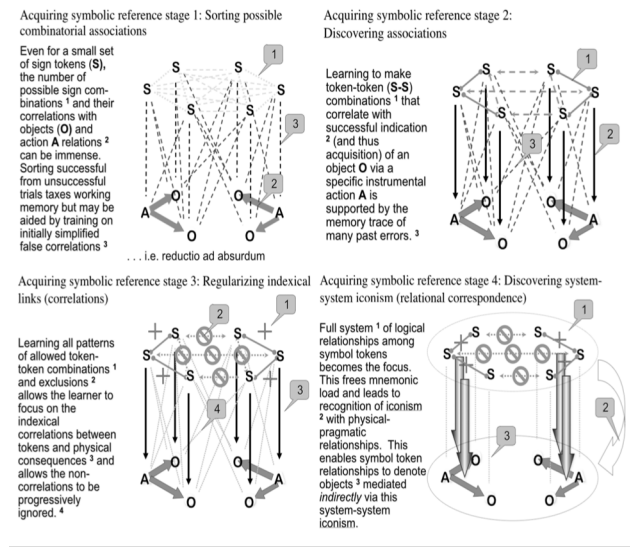

Figure 1. Stages of emergence of symbolic reference. From [9]

related to structured events in which these symbols are a part. In this way their syntactic relationships also became endowed with meaning. Although the original grounding of the lexigrams was not lost, their referential function was significantly transformed by the iconic and indexical relations between lexigrams. This system of relations allowed them to become partially "ungrounded" from these primary indexical relationships so that the abstract relationships between them could provide relevant referential clues.

This model is analogous to the situation of language development, in that the child initially interacts with caretakers by virtue of pragmatically grounded iconic and indexical means and eventually uses these signs as a scaffold in the acquisition of her first language. So by observing everyday pre-linguistic infant-caretaker interactions, and attending to their iconic and indexical functions, it should be possible to discern how the infant learns to communicate with the ungrounded sign vehicles of language. But the ungrounding process should be generic. So the study how an infant's non-symbolic use of gestures and words provides the scaffolding upon which ungrounded symbolic communication is built should also inform the design of systems capable of meaningful symbolic communication.

III. SEMIOTIC INFRASTRUCTURE IN LANGUAGE DEVELOPMENT

Language development gives us a particularly good opportunity to study the emergence of symbols [14, 15]. Early interactions provide especially vivid illustrations of the initial grounding of informational forms (signs), readily available for observations. We can thus appreciate the richness of the interactive context, i.e. the structuring provided by the caregiver and the infant for the events in which the first words appear. We can thus ask about the processes going on in the infants' heads and – adopting the focus of ecological psychology – we can ask "what their heads are inside of." Unlike the evolution of symbolic systems, lost in the past and difficult to study because of the lack of fossils, symbol ungrounding in development "happens all the time". This

gives researchers ample opportunity to study the processes behind this semiotic development: extending from utterances used as indices and icons controlling on-line interaction to symbolic communications and social conventions mediated by language.

A. Utterances as indices and icons in social physics

Posing the problem as symbol ungrounding suggests a reformulation of the questions of language acquisition: instead of asking how children ground the words they hear from adults or how they map utterances to objects events or states of the world, we ask how utterances that initially function as icons and indices come to acquire the properties of a symbolic system. Thus in the present approach utterances do not have (initially) the status of symbols. In fact, they do not have any status that would make them privileged with respect to any other actions that influence early interactions: gazes, gestures, smiles, non-linguistic vocalizations. The characteristics of symbols: arbitrariness, conventionality, formal structuring are not granted to symbols; rather the genesis of those properties is what requires explanation.

Let's illustrate this difference in the analysis with an example of conventional sign use by children, given by Elizabeth Bates and her colleagues [15]. When Carlotta, a 9 months-old infant (dressed in a red sweater) raises a fist in a combative gesture after an adult's utterance "Compagni!", a bystander (as well as the researcher) is prone to interpret this as the child's ability to recognize and use conventions: "*once a child begins using arbitrary signals – signals that he could not possibly have discovered without observing them in the social world – we have particularly clear evidence that he recognizes and uses conventions*."

A child, presumably, has formed an association between the utterance and the gesture and the association is arbitrary, because nothing in the sound 'Compagni!' naturally triggers or resembles the gesture that follows. This, however, is a bystander perspective, taking into account, at best, a slice of the process of how a sign effectively changes the infant's behavior. If we look at how the relation came about, we note that for a child *neither* the utterance *nor* the gesture are arbitrary. They are causally intertwined in a social routine, in which some actions enable predicting (and triggering) others because they are reenacted with a particular sequencing and timing. This is how the social world, constructed around the infant, works: in the "social physics" reenacted for a child, raising the fist after somebody shouted "Compagni!" is followed by a cascade of positive events, such as smiles, praises and being in the center of attention. And only because of this the gesture is performed.

Careful observations of early language development, in which language is granted a controlling and not merely a descriptive (mapping) role, makes it more amenable to pragmatic and ecologically valid analysis. A child "tunes-in" to the utterances as affordances, which control individual and collective behavior. This is how 'words' can become 'messages' in the first place, i.e., as signs sustained in their causal roles due to reenacted social routines [16]. Tuning-in, congruently with the tenets of ecological psychology, consists in changes in the way utterances are perceived, as specifying action and co-action in a social world. Before we turn to the ungrounding process, let's consider how early utterances are embedded within the reenacted social environment.

Utterances may play the role of indices for specific interactional behaviors. Examples abound in early interactions. Mothers use vocalizations to draw the attention of a child, to forecast events and behaviors, and to evoke particular responses. A greeting ('hello!'), a child's name, an imperative (e.g., 'look'!) a farewell ('bye-bye!') are each typical examples of utterances that first function as simple indices of subsequent moves in interaction. It is important to note that they cannot be usefully described as indices for 'referents out there' but are rather elements of coordinative events: it is more important that they evoke actions than images in the head of the child. An interpretation of a sign "Grego!" is not "ah, mother is referring to me" but rather "I have to look at my mom". Such a sign functions as an index in a series of events that allows prediction of what is likely to happens next or what is expected from the child. Soon it becomes used by the child as well, not only to 'refer' in the sense of indicating or describing the world but rather to control the flow of interaction.

Utterances can also serve as icons: the context and the time at which they are produced is aligned in an isomorphic way with important dimensions of interactive events. Parents use specific prosody, length of utterances, amplitude modulation that is coordinated and thus helps predict properties of events. A mother picks up a child with a long "ooooooppallaaa", coordinated with the length of the upward movement; says 'peek-a-boo' coordinated with the surprising suddenness of appearance; says roll, roll, roll in a 'rolling' way when turning the infant [9]; says 'tap, tap, tap' when tapping on his belly.

Semiotic analysis of such interactive uses of utterances makes it clear that linguistic forms need not be engaged in communication only as symbols. Long before they are used as a full-blown symbolic system (i.e. natural language) they function as interaction coordinators and controls, managing attention and joint attention, establishing rhythms, aiding the partition of events, and synchronizing emotions. Two things are crucial to note. First, because of the immersion in multimodal on-line interactions, utterances can be quite precisely grounded, i.e., infants are quickly tuned-in to them as interaction controls. As Jerome Bruner says, in contrast to corrections of grammar, "*speech acts, on the contrary, get not only immediate feedback but also correction*." [17, p. 37-38]. Second, even though later the linguistic forms enter in complex relations with other linguistic forms, the initial grounding does not vanish but continues to provide for the language's pragmatic coordinative role, albeit transformed by the possibility of more complex utterances.

B. Ungrounding through relations to other signs

The observation that a child's first "words" are grounded in this manner in the on-going interactions makes the task of explaining how they gain symbolic properties seemingly more difficult. The model presented above suggests that a possible aid to "ungrounding" is through grounding in other signs. Caregivers rarely talk in mono-words, rather providing structured utterances. Soon the (indexical and iconic) relations of words to other words become apparent to a child, constituting a second kind of grounding: within the vocal modality. This allows for reliable predictions about what

follows within an utterance besides the predictions of what follows within a coordinative event.

We observe several processes that lead to privileging the vocal modality in communication and to its emergence from the multimodal stream of events. The structure of turn taking (present in actions in general) is seen in vocal interactions from very early on [18, 19]. Additionally, research shows that mothers' responses to infant's vocalizations differ depending on the quality of the vocalizations: i.e., the more an infant's vocalization resembles language (e.g. in its syllabic form) the greater the probability that a) the vocalization will be responded to and b) that the response will be language-like. Importantly, if the mother responds with speech it is also more probable that the next vocalization of the child will be language-like [20]. It was also noted that a greater propensity in mothers to respond to language-like vocalizations of infants with verbal responses is correlated with better language development as measured several months later [21]. These observations demonstrate that the infant is embedded in a highly structured behavioral and social niche, enacted by adults and which provides semiotically grounded scaffolding for the emergence of symbolic language.

Because of caretakers' differential responsiveness to an infant's language-like vocalizations, the relations among words within utterances become more salient to the infant. At the same time, utterances remain functional in their pragmatic contexts. In this way, intensive multimodal interactions continue to provide the embedding context for higher-order relations. This is crucial: the relations themselves become meaningful by controlling the interactive events, which also can become more complex. As Bruner points out [17], the first context of the use of complex two-word affordances by children is often the request format, where the roles of actors and objects in the real world are quite evident for a child. The socially-causal relations between these actors and objects in the interactive situations may aid in understanding the relations (semantic bootstrapping), while the perception and control of the events might, with time, be ordered by these inter-sign relationships (syntactic bootstrapping).

Apart from regular everyday interactions, playing with infants is also structured in a way that may aid understanding how the relations among signs, in turn, relate to the relations and events in interaction. Types of games, in which language is in a pre-specified way connected to a series of events and movements (or motions, like motions in a game) performed with a child are particularly good examples: e.g. enumerating while touching fingers, enacting simple narratives, when touching and moving the baby (such as in the games "This little finger went….", "Questo è l'occhio bello…", and so on). These are events in which linguistic structures do not "map onto" interactive structures but rather help to control/predict them. While relations among grounded signs may lead to simple associations resulting in generalizations [5], it is important to note that the grounding is more comprehensive and complex with respect to symbols since symbol-symbol relations are themselves are grounded. At the same time, perhaps paradoxically, grounding in other utterances across situations provides a mechanism for liberation from the immediate context. Relations among elements of utterances can bring attention to dimensions that might not be immediately perceptually present. Thus, no wonder the critique for early symbolic models for their solipsism was based on pointing out that one cannot get semantics by grounding symbols just in other symbols [2, 3]. Grounding symbols in other symbols cannot provide semantic grounding because it leads to the (always partial) ungrounding of language from the immediate context. According to Deacon's model, the systemic property constituted by the relations among signs makes those signs symbols. In development, patterns of words co-occurrences (systemicity) are provided by the adult's utterances and by enactments of early dialogue around the vocalization of the child. Grounded first in the relations among controlling events the structures transfer the control to novel situations.

IV. IMPLICATIONS FOR MODELING THE EMERGENCE OF SYMBOLIC REFERENCE

The aim of this section is to formulate guidelines for computational models of the emergence of symbols from non-symbolic meaningful forms, i.e. for simulating the ungrounding process. Many of the current ingenious and successful models for clarifying aspects of symbolic functioning, have nevertheless usually been concerned with grounding the symbolic forms. Additionally, they have often been concerned primarily with an evolutionary timescale. This perspective renders many critical aspects of the process of language emergence inaccessible to observation. In the work presented above, Deacon's model has been applied to understanding language emergence in development. This fills in the theoretical frame with real-life examples. The developmental timescale for the emergence of symbolic communication undoubtedly differs from the evolutionary one (the most prominent difference is, of course, the co-occurrence of symbolic sign use in linguistic structures provided by the adults). However, the developmental time scale makes some elements of the model more amenable for study, and the semiotic principles, as noted earlier, should be generic and relevant to research on other timescales.

There is not space to review the extensive modeling work on the emergence of symbolic communication, even though some of it is directly relevant to the present work [4, 5, 13, 22-29]. These models are based on a variety of architectures and diverse learning algorithms, and they aim at explaining various aspects of the emergence of structured communication systems. However, in general most of the models of symbol acquisition by cognitive systems take the prior existence of symbols for granted.

Some models remain at the purely symbolic level, without any concern for the grounding problem. This does not mean that they aren't informative. Consider, for example, the study by Smith, Brighton and Kirby who demonstrated that compositional systems are more stable in the face of bottlenecks in cultural transmission [30]. Other models that include semantic aspects, explain the necessity of grammar by invoking the semantic complexity of the content conveyed by symbols [25, 31]. Yet others ground symbolic reference more thoroughly in the actions of agents in the environment, by coupling symbolic functioning to evolutionary fitness [3, 4, 32] or success in on-line interactions [22, 23]. However even in these pragmatically oriented models, grounding is assumed

to be a mapping relation, either a simple one, from objects in the agent's environment to signs [32] or more complex, mediated by generalized conceptual representations [4, 5] or by internal structured representations of the environment or action plans [23]. Reformulating the problem as ungrounding, along the lines presented above, as well as capitalizing on the controlling role of signs (including symbols) provide two general tenets guiding the future modeling work. Glimpses of similar approaches can already be discerned in existing studies and will be very helpful in the elaboration of our models.

For example, in a recent model [29], agents use the signaling of other agents to directly control their actions. Importantly, in some scenarios, agents, besides using the signals of others to compute their movement trajectories, include them in the computing of their own signals "in response." This results in a dialogical, communicative behavior, which may lead to cooperation. This is a very promising direction, however it is not clear how a signal from the other agent differs from other aspects of the environment that are used to compute trajectories and further signals (i.e., it is not clear why they are called "symbols" and not "indices") and also how syntactic structures may emerge as used by one agent (not only in a turn-taking mode (but see e.g. [33] on the role of dialogicity in the emergence of grammatical structures).

Also very helpful in this context is research that capitalizes on the relations among the signs in modeling the emergence of symbolic reference. An example is 'symbolic theft' in which grounding of abstract dimensions can be achieved by associating the names for abstract categories with already grounded ones [5]. This can explain the enlargement of a symbol system, though not its emergence. But it also demonstrates how the more concrete (in our framework iconic and indexical) controls can become ungrounded through selecting and creating important dimensions not obviously present in the input, by only in other signs.

The above strands of modeling work can benefit both from the change of perspective we propose here, as well as from the emphasis on language development, which makes steps in the ungrounding process more obvious. This aids recognizing that being immersed in co-action with others provides the complex semiotic infrastructure on which symbolic systems rely. The indexical and iconic involvement of signs in the control of interactive situations constitutes a vital part of the model. Without accounting for the direct, Gibsonian-like involvement of signs grounded in the social physics as controls, or enabling constraints, symbolic reference appears unattainable

Recognizing the "double grounding", i.e. indexical and iconic grounding of signs both in coactions in the world and in other linguistic forms is another key requirement of the model. As noted above, it is the relation between/among signs that provides a novel form of control in pragmatic social interaction. The fact that grammar can reduce the computational complexity of semantic interpretation [23] stems from the fact that grammar imposes constraints on the relations between referents. This realization might be helpful in the development of the models such as [29].

Epigenetic robotics seems to be a particularly good environment for developing models of the emergence of symbols, as guided by these principles. Robots are immersed in some kind of structured physics, in which signs may function as icons and indices, thus events can be predictably, informationally connected for them. Relations among signs reflect also these informational connections and generalize them to other relations. Agents are immersed in their environments as actors, therefore their primary attitude towards reality is not the description or representation but control. Most importantly, the environment is constituted by other actors, thus the criteria for this control are primarily pragmatic and coordinative. Symbolic systems emerge in dialogical scenarios of mutual control and coordination within joint activities. Congruently with Vygotsky: "*A sign is always originally a means used for social purposes, a means of influencing others, and only later becomes a means of influencing oneself.* " [34, p. 157].

V. CONCLUSION

Instead of following the usual approaches to symbol grounding — i.e., starting from ungrounded symbols and trying to link them to dynamic events — we frame the problem differently. We ask: how does an infant learn to communicate with ungrounded sign vehicles (symbols) that are amenable to conventionalization and formal relations, beginning with only initially grounded signs. We think that an answer to the problem of the emergence of symbols requires answering questions about how events in interaction become understood as icons and indices and how these become symbols.

We employed a model proposed by Deacon [9], which shows that one important path to developing ungrounding symbols relies on their systemicity, i.e. grounding of signs not only in events but also in other signs. We showed the developmental realization of such a process, where linguistic signs are first icons and indices in the infant's "social physics" making it predictable and controllable. Subsequently, through establishing relations to other signs the control can become qualitatively different, guided by transmittable relations among linguistic forms. Finally, we described what features of our computational models, are likely necessary to model the ungrounding process. Exploring this will be the next step in our work. Summarizing the features of such models, they should:

- Be informed by developmental processes, where the data on coordinative processes constituting the meaning of the utterances are readily available.
- Pay attention to available patterns created by social physics of the agents, i.e. their active involvement in complex events. In simulated environments this could be achieved by immersion of agents in collaborative tasks.
- Allow not only for agents' action (pragmatic goals) but social-coordinative action, allowing for symbolic systems to emerge the „Vygotskyan way".
- Capitalize on the physicality of signs: signs must be physical entities with physical structure, present publicly in the environment, and amenable to re-presentation by the agents to each other. In this way, they can remain causal in "social physics".

- Allow for signs to be predictably linked among themselves. Symbolic signs do not just co-exist but 1) they are usually in systematic sequences, which, for example, make one an index for another and 2) they co-exist as controls, grounded in events, transferring relations between episodes of control.

The field of epigenetic robotics (if indeed epigenetic and indeed robotic ,i.e., developmentally and pragmatically related to the environment) seems like a promising environment for exploring the emergence and evolution of symbolic communication. But, the field could benefit from including the "ungrounding" process in the design of artificial systems. Cybernetic relations between meaningful forms and the behavioral interactive dynamics in an environment help to demonstrate how symbols relate to dynamics and to foreground the pragmatic aspects, which are transparent for the participants and thus often taken for granted and difficult to study.

REFERENCES

[1] Dreyfus, H. (1972). *What Computers Can't Do*. New York: Harper and Row.

[2] Searle, J. R. (1980). Minds, brains and programs. *Behavioral and Brain Sciences 3,* 417-424.

[3] Harnad S. (1990). The Symbol Grounding Problem. *Physica D* 42: 335-346

[4] Cangelosi A. (1999). Modeling the evolution of communication: From stimulus associations to grounded symbolic associations. In D. Floreano et al. (Eds.), Proceedings of ECAL99 European Conference on Artificial Life, Berlin: Springer-Verlag, 654-663

[5] Cangelosi A., Greco A., & Harnad S. (2000). From robotic toil to symbolic theft: Grounding transfer from entry-level to higher-level categories. Connection Science, 12(2), 143-162

[6] Deacon, T. W. (1997). *The Symbolic Species: The Co-evolution of Language and the Brain.* New York: W.W. Norton & Company.

[7] Nomikou, I., & Rohlfing, K. J. (2011). Language does something: Body action and language in maternal input to three-month-olds. *IEEE Transactions on Autonomous Mental Development 3*(2):113 - 128

[8] Szufnarowska J., Rohlfing K. J. (2014). Enfolding interaction with two-month-olds. In: *Proceedings of the 16th European Conference on Developmental Psychology*, Lausanne, Switzerland. Bologna: Monduzzi Editore, 213–218.

[9] Rączaszek-Leonardi, J., Nomikou, I., Rohlfing, K. J. & Deacon, T. W. (2018). Language Development From an Ecological Perspective: Ecologically Valid Ways to Abstract Symbols. *Ecological Psychology, 30:1*, 39-73, DOI: 10.1080/10407413.2017.1410387

[10] Peirce, Charles Sanders (1931) *Collected Papers of Charles Sander Pierce. Vol. II Elements of Logic.* C. Hartshorn and P. Weiss (eds.) Cambridge, MA: Harvard University Press

[11] Savage-Rumbaugh S. & Rumbaugh D.M. (1978). Symbolization, language, and Chimpanzees: A theoretical reevaluation on Initial language acquisition processes in four Young Pan troglodytes. Brain and Language, 6: 265-300.

[12] Savage-Rumbaugh, E. S., Rumbaugh, D. M., Smith, S. T., & Lawson, J. (1980). Reference: The linguistic essential. Science, 210 (4472), 922-925.

[13] Cangelosi, A . 2001. "Evolution of communication and language using signals, symbols and words". IEEE Transactions on Evolutionary Computation 5(2): 93–101.

[14] Piaget, J. (1945/962). *Play, Dreams, and Imitation in Childhood*. New York: W. W. Norton & Company. (orig: "La formation du symbole chez l'enfant: Imitation, jeu et rêve, Image et représentation")

[15] Bates, E., with L. Benigni, I. Bretherton, L. Camaioni, & V. Volterra. (1979). *The emergence of symbols: Cognition and communication in infancy.* New York: Academic Press.

[16] Rączaszek-Leonardi, J. (2016). How does a word become a message? An illustration on a developmental time-scale. New Ideas in Psychology, 42, 46-55. doi:10.1016/j.newideapsych.2015.08.001

[17] Bruner, J. S. (with Watson, R.). (1983). *Child's talk: Learning to use language.* Oxford, UK: Oxford University Press.

[18] Trevarthen C. (1979). Communication and cooperation in early infancy: a description of primary intersubjectivity. In: Before Speech: The Beginning of Interpersonal Communication ed. Bullowa M., editor. Cambridge: Cambridge University Press, 321–347.

[19] Leonardi, G., Nomikou, I., Rohlfing, K. J. & Rączaszek-Leonardi, J. (2016). Vocal interactions at the dawn of communication: The emergence of mutuality and complementarity in mother-infant interaction. In: *Proceedings of the IEEE ICDL-EpiRob*, Cergy-Pontoise, pp. 288-293.

[20] Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological Science , 25 (7)*, 1314–1324. doi:10.1177/0956797614531023

[21] Radkowska, A., Nomikou, I., Leonardi, G., Rohlfing, K. J. & Rączaszek-Leonardi J. (2017). Scaffolding vocal development: maternal responsiveness to early speechlike vocalizations in three, six and eight month olds. Poster presented at IASCL.

[22] Steels, L. (2000) The Emergence of Grammar in Communicating Autonomous Robotic Agents. In Horn, Werner, editor, ECAI2000, pages 764—769.

[23] Steels, L. (2005) What Triggers the Emergence of Grammar? In AISB'05: Proceedings of the Second International Symposium on the Emergence and Evolution of Linguistic Communication (EELC'05), pages 143--150.

[24] Batali J. (1994). Innate biases and critical periods: Combining evolution and learning in the acquisition of syntax. In R. Brooks & P. Maes (eds), Artificial Life IV, Cambridge, MA: MIT Press, 160-171.

[25] Batali, J. (1998). Computational simulation of the emergence of grammar. In J. R. Hurford, M. Studdert-Kennedy, & C. Knight (Eds.), Approaches to the evolution of language. Cambridge, UK: Cambridge University Press.

[26] Hutchins E. & Hazelhurst B. (1995). How to invent a lexicon. The development of shared symbols in interaction, In N. Gilbert e R. Conte (Eds.) Artificial societies: The computer simulation of social life, London: UCL Press.

[27] Hashimoto, T. and T. Ikegami (1996) Emergence of net-grammar in communicating agents. BioSystems 38 (1996) 1-14.

[28] Leijnen, S. (2012). Emerging symbols. In T. Schilhab, F. Stjernfeld & T. Deacon (Eds.) The Symbolic Species Evolved. Biosemiotics 6. Dordrecht: Springer, pp. 253-262.

[29] Grouchy, P., D'Eleuterio, G.M., Christiansen, M.H., & Lipson, H. (2016). On The Evolutionary Origin of Symbolic Communication. Scientific Reports, 6, 34615.

[30] Smith, K., Brighton, H., & Kirby, S. (2003). Complex systems in language evolution: The cultural emergence of compositional structure. Advances in Complex Systems, 6(4), 537–558.

[31] Schönemann, P. T. (1999). Syntax as emergent characteristic of the evolution of semantic complexity. Minds and Machines, 9, 309–346.

[32] Cangelosi, A., Parisi D. (1998). The emergence of a "language" in an evolving population of neural networks. Connection Science, 10(2), 83-97.

[33] Jennings, R.E., & Thompson, J.J. (2012). The Biology of Language and the Epigenesis of Recursive Embedding. Interaction Studies, 13(1), 80–102.

[34] Vygotsky, L. (1931/1981). The Genesis of Higher Mental Functions. In James Wertsch, The Concept of Activity in Soviet Psychology. Armonk, NY: M.E. Sharpe.