

# Interaction and Experience in Enactive Intelligence and Humanoid Robotics

Chrystopher L. Nehaniv\*,

Frank Förster\*, Joe Saunders\*, Frank Broz\*<sup>†</sup>, Elena Antonova<sup>‡</sup>, Hatice Köse\*<sup>§</sup>, Caroline Lyon\*,  
Hagen Lehmann\*, Yo Sato\*, and Kerstin Dautenhahn\*

\*University of Hertfordshire, Adaptive Systems Research Group, U.K.

<sup>†</sup>Plymouth University, School of Computing and Mathematics U.K.

<sup>‡</sup>King's College London, Institute of Psychiatry, U.K.

<sup>§</sup>Istanbul Technical University, Computer Engineering, Turkey

**Abstract**—We overview how sensorimotor experience can be operationalized for interaction scenarios in which humanoid robots acquire skills and linguistic behaviours via enacting a “form-of-life” in interaction games (following Wittgenstein) with humans. The enactive paradigm is introduced which provides a powerful framework for the construction of complex adaptive systems, based on interaction, habit, and experience.

Enactive cognitive architectures (following insights of Varela, Thompson and Rosch) that we have developed support social learning and robot ontogeny by harnessing information-theoretic methods and raw uninterpreted sensorimotor experience to scaffold the acquisition of behaviours.

The success criterion here is validation by the robot engaging in ongoing human-robot interaction with naive participants who, over the course of iterated interactions, shape the robot's behavioural and linguistic development. Engagement in such interaction exhibiting aspects of purposeful, habitual recurring structure evidences the developed capability of the humanoid to enact language and interaction games as a successful participant.

## I. OVERVIEW

As part of the international RobotCub consortium that created the iCub humanoid as an open platform for cognitive systems research and in subsequent work, over the past 9 years in the Adaptive Systems Lab at the University of Hertfordshire, we have developed the notion of *extended sensorimotor experience* for robots interacting with their environment, including the social environment as a basis for the development of artificial intelligence via the acquisition of new skills and behaviours. Experience is operationalized as the temporally extended flow of information across sensory, motor and internal variables mediating interaction.<sup>1</sup> Such raw uninterpreted experience can be used with biologically plausible information-theoretic methods to self-structure sensory fields, and derive dynamical regularities relating motor actuation to what will be sensed, resulting in actively guided perception and action (e.g. the discovery of optical flow and the capacity for visually guided movement [1]).

<sup>1</sup>This is not the same as phenomenological experience. NB: Operationalized experience cannot be directly and unproblematically equated with phenomenological experience.

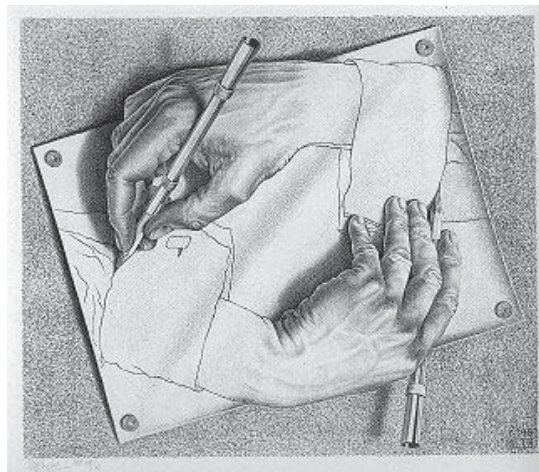


Fig. 1. *Drawing Hands* by M.C. Escher, 1948, Lithograph. (Fair use of low resolution image from [6] under U.S. copyright law.)

Without employing any notion of representation or static, truth-functional semantics, on several robots and humanoids we have harnessed enactive cognitive architectures based on interaction histories and shown that such temporally extended experiences can structure the acquisition of self-regulating skills such as turn-taking games in *social engagement* [2], [3], with the acquisition of new behaviours, and the dynamic switching between them adaptively depending on details of ongoing interactions with humans [4], [5].

Experience, thus operationalized, in turn shapes subsequent action and interaction, which in turn shapes the developmental trajectory of further experience (cf. M.C. Escher's image of hands drawing each other – Figure 1), as “paths are laid down in the walking”, for habits and skills as the robot explores its embodiment, and physical and social world.

This approach has also been employed in the acquisition of rudimentary linguistic capabilities in humanoids in the ITALK project [7] which harnesses mutual scaffolding in a circular feedback cycle between social learning, individual



Fig. 2. The University of Hertfordshire iCub, DeeChee (centre), acquires capability to use words meaningfully based on sensorimotor experience, linguistic and social interaction using an enactive cognitive architecture with Dr. J. Saunders (left) and Prof. C. L. Nehaniv (right). (Photo by Pete Stevens)

learning, and linguistic behaviours acquired in long-term interaction with humans. The approach synthesizes ideas of Wittgenstein on language as meaningful only in the context of use in interaction games [8] with the enactive approach of Varela, Thompson, and Rosch [9]. Furthermore, this is grounded by relying heavily on rigorous mathematical and information methods (esp. Shannon information theory [10]) applied to our operationalized notion of experience to achieve ontogeny through establishing, engaging in, and modifying *habitual activity*, via repeated interactions with humans, that lead to informational self-structuring of the robot's ongoing sensorimotor experience as a basis for unfolding action and interaction in a virtuous circle of development. (See references cited below for technical details.)

## II. ENACTIVE PARADIGM

### A. What is Enactivism?

Enactive cognitive science has been originally formulated by Francisco Varela, Evan Thompson, and Eleanor Rosch in their 1991 book *The Embodied Mind* [9]. Unsatisfied either with cognitivism or connectionism, Varela and colleagues sought to overcome the limitations of approaches that restrict consideration of the organism to just its brain on the one hand and that objectify the environment on the other.

To achieve the transcendence of the dualistic dichotomies, the enactive approach emphasizes the interaction of the embodied organism with its environment. It disposes of the notions that the organism constructs symbolic representations of the (objective) world and manipulates these representations to guide its actions. Instead, it is the purposeful activity of the embodied organism that brings forth a world of meaning. Thus, at the heart of the enactive approach is an ongoing interplay between biology and phenomenology - dual, but not dualistic. Enactivism is also an approach to understanding and synthesizing living and life-like systems, including cognitive robots and understanding humans and animals constructively.

Several defining features of an enactive approach include:

- Embodiment (structural coupling with the world) is the basis of activity and interaction, including social interaction. The agent and environment mutually perturb each other's state and dynamics, even affecting their structure in the course of time.
- Experience and interaction are the central focus of study.
- Dynamics of interaction cut across agent/other/environment boundaries.
- There are no static entities, only processes. Phenomena can't be pinned down in an 'objective' manner, independent of any observer, but are co-dependently originating processes.
- At the basic level, the enactive paradigm presumes no reliance on propositional logic (a highly derived abstraction), nor reliance on representation.
- An enactive cognitive system functioning well becomes part of an ongoing world, or gives rise to a world through enaction (its activity). Varela, Thompson, and Rosch in *The Embodied Mind* [9], quoting a poet, describe this fact of how worlds of life and experience come to be in historical processes as "Laying down a path in walking".
- Social interaction and intersubjectivity: Others are part of the environment and interaction with them is as primary as with the rest of the environment.
- Developmental perspective: ontogeny is implicit since we are talking about dynamical systems that modify their structure. As a result, the role for habitual activity, long-term interaction, and continual change are central.
- Interaction Games and Language Games (cf. Wittgenstein [8]) : an enactive system is working if it participates in its world effectively. For example a robot initiates and maintains social engagement with humans, acquires the capacity to use some linguistic skills through interaction, and then uses language to manipulate the world (including its social world) and to participate in language games.
- Enactive cognitive systems actively create, maintain, and regulate their own ways of being in and manipulating the world (including the social world), rather than merely tuning parameters in some innate or otherwise pre-given models.

### B. What can the enactive approach tell us about human cognition that can't be explained by other theories?

Other theories miss the role of experience, or have to use a pre-given ontological framework to deal with it. They carve up the world in static ways, while the dynamics of interaction cuts across these forced/reified distinctions. Intertwining of perception-action is missed by other theories that pre-suppose they are separate stages.

This not only does not reflect how things are organized in nature, but it is also very expensive in terms of required resources and often completely infeasible. On the other hand, enactive methods can often achieve predictive and explanatory power missing from traditional models. Examples from cognitive science below illustrate this.

For human cognition, the enactive paradigm is yet to develop a detailed understanding of higher-order cognition, although some inroads are being made in psychiatry, child development and linguistic development.

### C. Benefits for Artificial Intelligence (AI) Research

The enactive viewpoint

- allows one to avoid modelling : agent, environment, and the sense-plan-act approach based on updating world models (e.g. of traditional cognitive science and AI).
- allows one to avoid shoehorning cognition into some form of representation manipulation and / or logical inference. (These are in no sense primary, and usually unnecessary.)
- can often lead one to constructive and effective mechanisms to achieve a working implementation of the phenomenon or behaviour under study, or lead to the identification of gaps in theoretical understanding that are revealed by an attempted implementation.

### D. What is the evidence supporting these ideas?

Much evidence showing the power of the enactive approach has been accumulating over the last decades.

Neuroscientific evidence of self-processes as transient modes of activation [11] supports an enactivist interpretation of these. Specifically, the work on the Default Mode Network (DMN) of the brain which appears to be associated with self-referential processing [12] shows that the DMN, mainly comprising midline regions of the brain, including medial prefrontal cortex (mPFC) and posterior cingulate (PC) is relatively more active during what Buckner and Carroll [13] have called self-projection: remembering the past, thinking about the future, or thinking about other minds (theory of mind). It is also more active during mind-wandering [14], narrative self-processing [15]. The same network is relatively less active when the individual is engaged in solving cognitive tasks or switch to modes of self-experience which are more sensory based (embodied) and do not have a rigid self-other boundary [15], [16]. Overall, the neuroscience evidences that there is interplay of transitory self-processes that functionally arise in a context-dependent manner depending on nature of human-environment interaction.

Operational or sensorimotor experience in meaningful interactions plays an important but non-exhaustive role in accounts of mind. Deep issues of the nature of experience and cognition are uncovered by adopting an enactive viewpoint. Operational mastery of sensorimotor laws in agent's experience and the particularities of enacting this mastery in the world have been argued by J.K. O'Regan and A. Noë to account for the particular character of that experience [17], [18]. In contrast to traditional approaches that regard sensing as a passive mechanism serving to build a up a world-model within the brain, the enactive sensorimotor approach is able to yield a much better explanatory and predictive understanding of, say, vision. Traditional approaches require elaborate mechanisms for the internal construction and maintenance of continually updated and corrected representations that should function

*despite* the structure of the sensory apparatus and embodiment (e.g. "correct" for the blind-spot in the human retina). By taking an enactive sensorimotor approach that, instead, *seeing (and perception generally) are about interacting with world*, one is able to better explain and make predictions concerning such phenomena as sensory substitution, change blindness, inattentive blindness, localization of touch, colour naming and colour perception, and other phenomena of perception [18].

Activity of the organism in interaction within its environment, rather than a brain-limited process, may be a better account for mind than the crude computer-science metaphor-laden notions prevalent in much of cognitive science, identifying the mind either with the brain or as "software" that is "running" on the brain. Alternatives evidencing the explanatory power of enactive paradigm in various fields include, for instance, D. Noble's *The Music of Life* [19] for systems biology, physiology and remarks on skills (such as musicianship); versions of the extended mind hypothesis (e.g., Noë [20]) for cognition and conscious experience; or notions of Maturana and Varela equating *doing* and *knowing* [21]; as well as studies in intersubjectivity and social cognition [22] (cf. also early studies by Trevarthen [23] and Kaye [24]).

Constructive enactive approaches lend themselves to create artifacts that achieve or model behaviours and processes of interest. Systems can be built using these ideas that work, and avoid representation, truth-functional propositional logic and semantics, any need to construct an ontology or use explicit internal modelling. For example these are properties of the simple, robust behaviour-based robots of Rodney Brooks [25], [26], which use none of these. Closely engaging with the environment including the social environment enacts interaction and experience in a 'strange loop' influencing or even producing the structures and dynamics of the agent, environment and others in environment. Such approaches are now being scaled up to more human-like cognition, such as in the EU-funded RobotCub project (2004-2010), which explicitly sought to follow an enactive approach to the ontogeny of cognition from its inception [27], [28], informed awareness of constructive minimal social robotics [29], [30] and by notions of achieving different kinds and degrees of embodiment [31]. The RobotCub Consortium, in which the University of Hertfordshire as AI partner focused on human-robot interaction and ontogeny of behaviour, created the humanoid iCub. In that project and subsequent work on FP7 ITALK project (2008-2012) working with our iCub, DeeChee, such enactive methods were applied to achieve maintenance and regulation of social interaction with human participants, acquisition and scaffolding of new skills via social learning [5], [32]; acquisition of phonetic word-forms [33]; acquisition of linguistic skill (analogous to two-word stage in human children: [34], [35], etc.); and acquisition of some forms of linguistic negation [36], [37]. Moreover, there is growing evidence that enacting appropriate contingent, nonverbal engagement in social learning scenarios scaffolds the interaction in such a way that learning can proceed better [38], [39].

### E. Varieties of Enactivism

At present, the enactive paradigm brings together radical insights from philosophy, behaviour-based artificial intelligence, experiential intelligence, embodiment, sensorimotor theory, and primary intersubjectivity, development and social interaction. It exists in various forms. Origins, maintenance, and regulation of interactive processes, dependently originated boundaries and habit are central. Radical enactivism completely rejects the idea that propositional content is necessary for basic cognition [40]. Closely related are approaches that are exploring *interaction* as a basis for computation [41]. The enactive viewpoint encompasses an powerful framework in which to understand cognition, behaviour and intelligent activity, using interaction and experience, but also unifies all of this with the study of biology (see, e.g. [42]), and hence can serve as a useful framework for artificial life and a constructive generalized biology.

### III. THE EXPERIENCE OF A ROBOT

We operationalize experience as the flow of activation of sensorimotor and internal variables [3], [43], [44] of an agent interacting with its environment, including other agents around it.

For any pair of jointly distributed random variables (sensors)  $X$  and  $Y$  of the robot, the *conditional entropy*  $H(X|Y)$  of  $X$  given  $Y$  is the amount of uncertainty (in bits) that remains about the value of  $X$  given that the value of  $Y$  is known, and is given by

$$H(X|Y) = - \sum_{x \in X} \sum_{y \in Y} p(x, y) \log_2 \frac{p(x, y)}{p(y)},$$

where  $p(x, y)$  is the joint distribution of  $X$  and  $Y$ .<sup>2</sup> The *information distance* between  $X$  and  $Y$  is then defined as

$$d(X, Y) = H(X|Y) + H(Y|X),$$

and is mathematically a metric satisfying the triangle inequality, symmetric, etc., giving a geometry on information sources [45].

This metric has been used to do sensorimotor reconstruction to enable robots to construct sensory fields and acquire the ability to control their own actuators to achieve desired effects in sensing via the discovery of sensorimotor laws, such as optical flow due to turning the neck and head [1] without prior knowledge of the sensors and effectors.

Given the above definition we can operationalize an agent's experience from time  $t$  over a temporal horizon of  $h$  time units as  $E(t, h) = (X_{t,h}^1, \dots, X_{t,h}^N)$  where  $\{X_1, \dots, X_N\}$  is the set of all sensorimotor (or other) variables available to the agent and each  $X_{t,h}^i$  is the random variable estimated from the values of  $X_i$  over a temporal window of length  $h$  beginning at time

<sup>2</sup>By convention, any terms mentioning a zero probability event are not included in the summation in computing entropy. For instance, this prevents any problem taking logarithms if  $p(y)$  or  $p(x, y)$  is zero in the formula given here for conditional entropy.

$t$  ( $1 \leq i \leq N$ ). We can then define a metric, the *experience metric*  $D$ , on experiences of temporal horizon  $h$ , as

$$D(E, E') = \sum_{k=1}^N d(X_{t,h}^k, X_{t',h}^k),$$

where  $E = E(t, h)$  and  $E' = E(t', h)$  are experiences of an agent at time  $t$  and  $t'$  over horizon  $h$  and  $d$  is the information distance. Alternatively, one can define a cross-modal experience metric  $D'$ , by taking all the information distances between all the sensorimotor variables and summing these [46]. This puts (operationalized) experiences in a geometric space.

Using this metric an agent can compare its current and previous experiences, and select actions (probabilistically) in the light of their previous results in similar contexts, or when a skill or behaviour has been mastered, the agent can explore at the boundary of its capabilities (cf. Vygotsky's notion of "zone of proximal development" [47]). By using temporally extended experiences to guide action and interaction, we have grounded episodic intelligence in artificially constructed enactive agents that grow, develop, and adapt their cognitive structures with a broader temporal horizon.

### IV. ONTOGENY: DEVELOPMENT OF BEHAVIOURS AND SELECTION BASED ON SOCIAL CUES

Using an enactive cognitive architecture based on operationalized experience, several generations of the *Interaction History Architecture* have been developed by N. Assif Mirza and by Frank Broz working in our group at the University of Hertfordshire.<sup>3</sup> Figure 3 shows a schema for the basic Interaction History Architecture.

Using social cues such as turn-taking in various modalities, detection of social engagement, and feedback from vocal, face- and gaze interaction to help acquire and select action, these cognitive architectures have allowed robots to acquire behaviours such as predictive gaze and peek-a-boo [2]. Moreover, on the basis of social feedback our humanoids have first developed distinct interactive drumming behaviour and peek-a-boo interactive behaviours, and then were able to autonomously switch between these behaviours, depending on social cues and engagement with a human interaction partner [4], [5].<sup>4</sup> In our experiments, notably, peek-a-boo turning-taking behaviour could not be stably acquired if action selection is random, i.e. not based on prior experience, or if the temporal horizon of experiences is too small or too large [2], [5], and the same is true for interactive drumming which additionally required short-term memory of recent social engagement to be acquired [5]. This milestone work in embodied artificial intelligence using human-robot interaction exhibits for the first time together both (1) autonomous behaviour ontogeny and (2) switching between acquired behaviors in response to social contingency [4].

<sup>3</sup>The software developed and used in this work is available open-source at <http://sourceforge.net/projects/robotcub/>.

<sup>4</sup>See [http://eris.liralab.it/misc/icubvideos/ihaNew\\_short2\\_web.mov](http://eris.liralab.it/misc/icubvideos/ihaNew_short2_web.mov) for demo video.

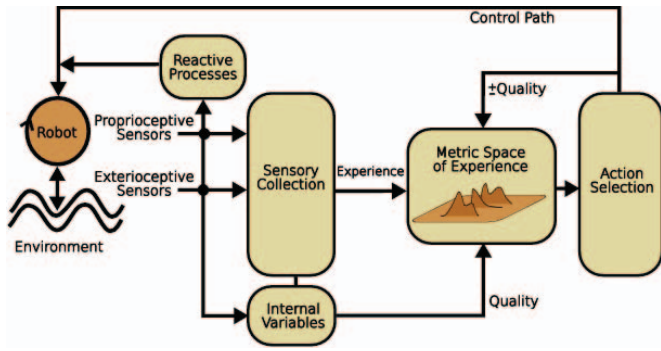


Fig. 3. Interaction History Architecture (IHA), a basic form of enactive cognitive architecture, selecting actions on the basis of prior experiences (see text) and as a result modifying subsequent experience. The node labelled ‘robot’ comprises actuation whereby the ‘strange loop’ between the robot’s experience and action interfaces and interacts with the environment, manipulating or shaping the subsequent environment, including people in it, as well as subsequent interaction. NB: This basic form of IHA shows a continual ongoing developmental process, giving rise to a long-term trajectory through recurring habits, experience and dynamics. No explicit modelling, semantics, or representation are employed. There is no separation of development into artificial training and testing phases; instead, learning, experiencing and development are intertwined and ongoing. Actions or behaviours are selected probabilistically based on the agent’s own assessment of their quality and environmental feedback (including social engagement), but also low-rate background chance random selection of actions leads to iterated cycles of exploration at the boundary of current behaviours, mastered skills, and activities. (Schematic drawing of the basic architecture from (Mirza et al. [3]) – for the most recent Extended Interaction History Architecture (EIHA), see [5].)

## V. LANGUAGE GAMES AND INTERACTION GAMES

Linguistic development takes place in children over years, involving long histories of interaction in various contexts in which regularities and habits are engaged in and re-engaged despite many changes in manner that involves deep familiarity with interaction partners, contexts, and practical needs. For example, for a child, contexts such as bedtime, or eating together as a family (e.g., breakfast), bath-time, car journey, play with siblings or caretakers; making a puzzle; playing monopoly with a sibling; helping with a chore; arguing with siblings for a toy, or with parents for a reward; playing peek-a-boo; drumming and singing together; etc., may provide the framework in which language development occurs, together with the development of action and social skills. To achieve developmental learning of action and language in habitual contexts, we can make use of philosopher Ludwig Wittgenstein’s insight on language deployed in embodied contexts and his notion of *language games*, which we have generalized to *interaction games* to emphasize also the role on non-verbal action, behaviour and manipulation in social contexts [2]–[5], [35], [44], [48], [49].

In the growth of intersubjective intelligence, as with human children, a typical scenario involves the contact of a learning agent with the context of a new interaction game (i.e., new to the learner) which other persons or social agents are already engaged in. Behaviourally, this “form of life” is manifested by their practice of some unknown set of skills together with some linguistic behaviour. Participants in a language manipulate the

world of the game and others in it, via actions and utterances. Entry into the interaction game is entry into a form of life based on action, language and interaction, e.g. workers on a construction site might yell ‘slab’, ‘board’, ‘hammer’ to one another, or bring and use these objects and tools, deploying their skills, utterances, and the items to regulate their activity and perhaps achieve a common purpose such as constructing a house (cf. [8]), thus enacting an interaction game.

The acquisition of these behaviours by the learner agent must proceed via stages of coarse-level to finer level understanding both of the behavioural goals of the participants, and of the linguistic means deployed in achieving them. Similarly, a small child who enters the world of, e.g., ‘breakfast’ and learns what others at the table or doing and how to get what it wants in this context by integrating language and action, needs to master behavioural and linguistic skills, which may start at a level of coarse understanding of the purpose of the game (eating breakfast with the family) and holistic linguistic phrases (‘hungry’, ‘gimme’), that develops in the course of long-term, habitual interaction in fine-grained mastery of the skills and behaviours (eating cereal with spoon, drinking from a cup in one’s own hands without spilling, etc.) as well as compositional language (‘I don’t want eggs!’, ‘Please pass the milk’, etc.). This movement from the coarse- to fine-graining in acquisition of habits, skills and behaviours via recurrent interaction in interaction games is a current topic of our research. Moreover, transfer, accommodation, and integration of action, language, and skills from other contexts plays a key role in this ontogeny.

## VI. ENACTIVE ACQUISITION OF LINGUISTIC BEHAVIOURS BY A CHILD-LIKE HUMANOID ROBOT INTERACTING WITH NAIVE PARTICIPANTS

Adopting Wittgenstein’s notion of language games, whereby interaction partners engage in a “form of life”, i.e. use interaction and speech to enact joint embodied activity, we have applied minimal social robotics principles and a social learning architecture to allow our childlike humanoid iCub robot DeeChee to acquire rudimentary linguistic capabilities. This approach takes intersubjective interaction as central and does not assume any a priori division of the dynamics via some agent-partner-environment ontology, rather it allows language and interaction game dynamics to emerge cutting across these often unhelpful categories. Nonverbal, prosodic, and social cues, embodied physical presence and sensorimotor experience, and regulation of interaction, together with manipulation of objects in context, help to scaffold the acquisition of linguistic capabilities by the robot interacting with naive participants in embodied learning scenarios. Success of the system in this approach equates to successful engagement by the humanoid in language games with human participants. Using this enactive social learning approach, DeeChee has acquired (1) the phonetic structure of salient words referring to objects whose names are being taught by participants in real-time interaction; and over a series of sessions of interaction: (2) capability to correctly use learned lexical items to refer

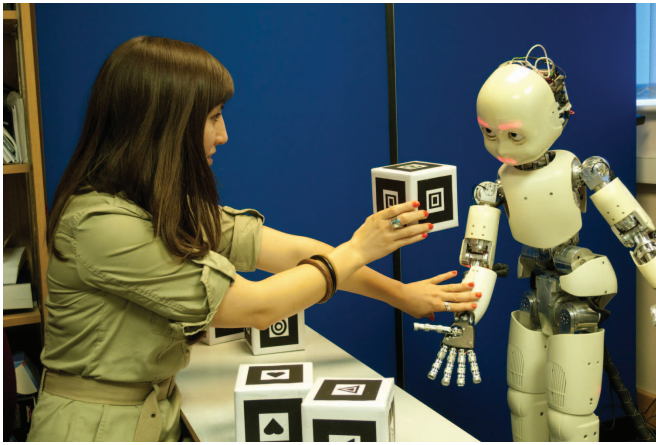


Fig. 4. The interaction partner prohibits the robot DeeChee from reaching the forbidden object physically and verbally, allowing the learning robot to associate this experience with the spoken “No”. (Photo by Frank Förster, who led the work on linguistic negation).

to manipulated objects and their properties, as validated by grounded usage in context with participants (and also by assessment of internal sensorimotor associations); (3) one-, two- and sometimes multi-word utterances used meaningfully in the context of this teaching scenario; and also (4) capability to use a variety of forms of linguistic negation.

#### A. Grounding Meaningful Utterances

Another cognitive architecture ROSSUM, developed by Joe Saunders and later Frank Förster in our lab, uses information gain in sensorimotor experience and also prosodic cues from speech of naive participants to learn and contextually select motor actions by the robot that linguistically manipulate the interaction, i.e. the utterance of words heard previously in interaction games. This gives rise to the grounded learning and meaningful use of words in the context of recurring interactions [34], [50].

#### B. Two-Word Stage in Linguistic Development

Extension of this work to allow the utterances of more than one word by Saunders et al. [35] leads to the meaningful use of two- or multiword utterances appropriate to experiential context that have a protogrammatical structure, similar to the two-word stage in linguistic development of human children.

#### C. Emergence of Linguistic Negation

Longstanding hypotheses on human development (e.g. Kaye [24]; cf. Tomasello [51]) assert that human infants develop into communicating languaging individuals in their first two years in large part due to being treated as intersubjective, intentional, communicative agents, and that this in fact is needed for them to become such agents, even though much of their behaviour may lack these characteristics at the start. However, through habitual interaction in which they are treated this way, infants do in fact become communicative persons. Following Wittgenstein [8], language usage reflecting a form

of practice in life arises in interaction games, and such appropriate usage of language in embodied contexts, rather than mere truth-functional semantics giving so-called ‘meaning’, gives utterances their meaning in human experience. In the work here, focusing on enaction and interaction (cf. Varela et al. [9]), we have an existence proof that development of the use of linguistic negation is possible in a humanoid which is successful at least in some limited contexts of interaction games in Wittgenstein’s sense of engaging. Results from a rejective scenario show how expressing rejective affect-motivational expressive behaviour in a child-like humanoid can evoke various forms of negation from naive human interaction partners in embodied interactions with the robot. Association of prosodically salient words from the participants’ speech with the experience (operationally defined as the flow of values over the sensorimotor-motivational variables of the robot [43], [44]) and triggering the utterance of these words in the humanoid’s speech under similar experience leads in the dyadic interactions to acquisition and linguistic expression of various forms of linguistic negation by the robot.

From the viewpoint of the participant or an adult speaker, the robot engages in uses of negation which appear to have the function of not only rejective negation but also other forms (even truth-functional negation). Participants appear to understand and respond to the robot’s behaviour and utterances as intentional, and appear to engage with the linguistically developing robot in a manner presuming and maintaining intersubjective engagement. Figures 5, 6, and 7 give excerpts in conversation analysis notation (following a situationally extended version of Jefferson notation [52]) of an interaction between the humanoid DeeChee and a (naive) human in the 5th session of teaching DeeChee the names of shapes. Participants were unaware that negation acquisition was being studied. Acquisition by the robot of words and contextual sensorimotor experiential triggers for their usage proceeded along independent trajectories for the various participants. DeeChee exhibits valanced motivational responses to the various objects. The human’s utterances include many forms of negatives, e.g. truth-functional negation in Figure 6. Extracts show the robot uttering negative words, interpreted in the interaction game, e.g. as motivation-dependent denial (Figure 5) and truth-functional negation (Figure 7), among many other classified types of negation. The “head-shaking” in Fig. 5 is an emergent consequence of an interaction between the robot turning to avoid the object toward which it has negative motivational valence and the human continually moving the object into the robot’s visual field.

In several example developmental trajectories of the robot in embodied interaction with different human participants in the rejective negation scenario, this leads in just a few sessions to the expression by the child-like humanoid of an array of types of what functions as and is construed as linguistic negation in the interaction. Strikingly, in these experiments the robot is responding on the basis of its current and prior sensorimotor-motivational experience and its prior experience of the interaction game with the participant; however it did

```

119 what about:
120 (1.75) ((P picks up moon box))
121 the: moon? ((R starts frowning, turns head away))
122 (1.21) .pt (.21)
123 the cr'escent th[ere]
124 R [no]
125 (.4)
126 P _no_?
127 (.19) .hh (.12)
128 >no you don't like that one ei'ther< ((R 'shakes head''))
129 (2.27)
130 no:? (.) shaking your head

```

Fig. 5. Negation Experiments: Motivation-Dependent Denial exhibited by the child-like humanoid iCub robot DeeChee. R marks the robot's utterances (line 124) and P the human participant's when there is a change of speaker.

```

5 see (.) can you remember= ((P picks up triangle box, looks at R)
6 =th' _is_ one ((R focuses on box))
7 (.2)
8 remember the tr'angles? ((R looks at P))
9 (1.38)
10 can you remember the triangles there= ((R looks at box))
11 =Deechee:
12 (1.73)
13 see them th[ere]? (.) you've got ((R looks at P))
14 (.0.26)
15 .hh thr'ee= ((R looks at box))
16 =triangles there:
17 (0.1)
18 inside [each oth'er ((R looks at P))
19 R [there]
20 (0.87)
21 P square:= ((R looks at box))
22 =(.) 'no it's not a square ((R looks at P))
23 (0.52)
24 .hh it's a: tr'angles ((R looks at box))

```

Fig. 6. Negation Experiments: Truth-Functional Denial exhibited by the Human engaging in a teaching scenario interaction game with DeeChee. R marks the robot's utterance (line 19) and P the human participant's when there is a change of speaker.

not hear the speech of the participant in real-time, but only learned from the participant's speech associated to the rest of its sensorimotor experience between sessions. Further studies characterize these trajectories in more detail than outlined above, and also compare the effect of additional exposure to prohibitive negation (see Figure 4) on the distribution and dynamics of usage of negation in human-robot interaction learning scenarios [37].

## VII. CONCLUSIONS AND OUTLOOK

We have employed experiential intelligence in enactive cognitive architectures for the acquisition via information-theoretic self-structuring of sensory fields and self-organization of sensorimotor control, as well as for socially regulated and embedded behaviours, including skills acquired by imitation, development of turn-taking and autonomous selection of behaviour in response to social contingency and engagement in habitual learning via long-term human-robot interactions. We have shown how aspects of language as motor activity in such contexts, including grounded reference, two-word utterance, and negation can be acquired using a radically enactive approach. More general learning of these *interaction games* by making use of teleological understanding [53] in contexts of habitual purposeful engagement is the usual context of child language development, and also the arena of meaningful language use. We aim to further emulate this development in humanoid robots by continuing to harness information self-structuring and insights of the enactive paradigm.

```

92 (or) what! about the heart ((P picks up heart box and presents it to R, R starts frowning))
93 (.6) ((R turns head to the right))
94 no? (.) you don't like the heart
95 (1.33)
96 no?
97 (1.12)
98 not a big fan
99 (.4)
100 R that [[sound similar to 'heart']]
101 (.73)
102 P heart (.) you will? done!
103 (.6) :hh (.2)
104 it's ah: >I think that's what you 'said anyway<
105 (.37)
106 'I'll take[ that as a
107 R [no]
108 (.12)
109 P _no_ ((R looks back ..
110 (.15) ((towards ..
111 no (.) it's not what you said ((stable ...
112 (.94) .pt (1.02) ((and 'shakes head'))
113 don't like [the heart]=
114 R [no]
115 P (.) no?
116 (.15)
117 ok ((R looks back at P, still frowning))

```

Fig. 7. Negation Experiments: Truth-Functional Denial (line 107) and Negative Agreement (line 114) exhibited by the humanoid robot DeeChee. R marks the robot's utterances and P the human participant's when there is a change of speaker.

## ACKNOWLEDGMENTS

This work was supported by the European Union Integrated Project RobotCub (Robotic Open-architecture Technology for Cognition, Understanding, and Behaviours), funded by the EC through the E5 Unit (Cognition) of FP6-IST under Contract FP6-004370, under project grant FP7-214668 for ITALK: Integration and Transfer of Action and Language Knowledge in Robots, and by the EU FET FP7 BIOMICS project, contract number CNECT-318202. Their support is gratefully acknowledged. We thank the three anonymous reviewers for their constructive comments that helped us to improve this article.

## REFERENCES

- [1] L. A. Olsson, C. L. Nehaniv, and D. Polani, "From unknown sensors and actuators to actions grounded in sensorimotor perceptions," *Connection Science*, vol. 18, no. 2, pp. 121–144, 2006.
- [2] N. A. Mirza, C. Nehaniv, R. te Boekhorst, and K. Dautenhahn, "Peek-aboo: Effect of experience length on the interaction history driven ontogeny of a robot," in *Proceedings the of 6th International Conference on Epigenetic Robotics*, 2006.
- [3] N. A. Mirza, C. L. Nehaniv, K. Dautenhahn, and R. te Boekhorst, "Developing social action capabilities in a humanoid robot using an interaction history architecture," in *Humanoid Robots, 2008. Humanoids 2008. 8th IEEE-RAS International Conference on*. IEEE, 2008, pp. 609–616.
- [4] F. Broz, H. Kose-Bagci, C. Nehaniv, and K. Dautenhahn, "Learning behavior for a social interaction game with a childlike humanoid robot," *Social Learning in Interactive Scenarios Workshop, IEEE Humanoids, 2009*, 2009.
- [5] F. Broz, C. L. Nehaniv, H. Kose-Bagci, and K. Dautenhahn, "Interaction histories and short term memory: Enactive development of turn-taking behaviors in a childlike humanoid robot," *Preprint: arXiv:1202.5600v1 [cs.AI]*, 25 February 2012, 2012.
- [6] [Online]. Available: <http://en.wikipedia.org/wiki/File:DrawingHands.jpg>
- [7] A. Cangelosi, G. Metta, G. Sagerer, S. Nolfi, C. Nehaniv, K. Fischer, J. Tani, T. Belpaeme, G. Sandini, F. Nori *et al.*, "Integration of action and language knowledge: A roadmap for developmental robotics," *Autonomous Mental Development, IEEE Transactions on*, vol. 2, no. 3, pp. 167–195, 2010.
- [8] L. Wittgenstein, *Philosophical Investigations*. Blackwell Publishing, 2001 [1953].
- [9] F. Varela, E. Thompson, and E. Rosch, *The Embodied Mind*. MIT Press, 1991.
- [10] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, July, October 1948.

- [11] E. Antonova and C. L. Nehaniv, "Do enactive cognitive robots need a self? Lessons from neuroscience," in *Foundations of Enactive Cognitive Science, 27-28 February 2012 – Cumberland Lodge, The Great Park, Windsor, U.K.*, 2012.
- [12] M. E. Raichle, A. M. MacLeod, A. Z. Snyder, W. J. Powers, D. A. Gusnard, and G. L. Shulman, "A default mode of brain function," *Proceedings of the National Academy of Sciences U.S.A.*, vol. 98, no. 2, pp. 676–682, 2001.
- [13] R. L. Buckner and D. C. Carroll, "Self-projection and the brain," *Trends in Cognitive Science*, vol. 11, no. 2, pp. 49–57, 2007.
- [14] M. Mason, M. Norton, J. Van Horn, D. Wegner, S. Grafton, and C. Macrae, "Wandering minds: the default network and stimulus-independent thought," *Science*, vol. 315, no. 5810, pp. 393–395, 2007.
- [15] N. A. S. Farb, Z. V. Segal, H. Mayberg, J. Bean, D. McKeon, Z. Fatima, and A. K. Anderson, "Attending to the present: Mindfulness meditation reveals distinct neural modes of self-reference," *Social Cognition and Affective Neuroscience*, vol. 2, pp. 313–22, 2007.
- [16] J. A. Brewer, P. D. Worhunsky, J. R. Gray, Y. Tang, J. Weber, and H. Kober, "Meditation training is associated with differences in default mode network activity and connectivity," *Proceedings of the National Academy of Sciences U.S.A.*, vol. 108, no. 50, pp. 20 254–20 259, 2011.
- [17] J. K. O'Regan, A. Noë *et al.*, "A sensorimotor account of vision and visual consciousness," *Behavioral and Brain Sciences*, vol. 24, no. 5, pp. 939–1031, 2001.
- [18] J. K. O'Regan, *Why Red Doesn't Sound Like a Bell: Understanding the feel of consciousness*. Oxford University Press, USA, 2011.
- [19] D. Noble, *The Music of Life: Biology beyond Genes*. Oxford University Press, USA, 2008.
- [20] A. Noë, *Out of Our Heads: Why You are Not Your Brain and Other Lessons from the Biology of Consciousness*. Hill and Wang, 2009.
- [21] H. R. Maturana and F. J. Varela, *The Tree of Knowledge: The Biological Roots of Human Understanding*, 3rd ed. Shambala, 1992.
- [22] E. T. Thompson, *Between Ourselves: Second-Person Approaches to the Study of Consciousness*. Imprint Academic, 2001.
- [23] C. Trevarthen, "Communication and cooperation in early infancy: A description of primary intersubjectivity," *Before Speech: The beginning of interpersonal communication*, pp. 321–347, 1979.
- [24] K. Kaye, *The Mental and Social Life of Babies: How Parents Create Persons*. University of Chicago Press, 1982.
- [25] R. A. Brooks, "A robust layered control system for a mobile robot," *IEEE Journal of Robotics and Automation*, vol. 2, no. 1, pp. 14–23, March 1986.
- [26] R. Brooks, C. Breazeal, M. Marjanović, B. Scassellati, and M. Williamson, "The Cog project: Building a humanoid robot," in *Computation for Metaphors, Analogy, and Agents*, C. L. Nehaniv, Ed. Springer Lecture Notes in Computer Science, 1999, vol. 1562, pp. 52–87.
- [27] D. Vernon, G. Metta, and G. Sandini, "The iCub cognitive architecture: Interactive development in a humanoid robot," in *IEEE 6th International Conference on Development and Learning, 2007 (ICDL 2007)*. IEEE, 2007, pp. 122–127.
- [28] K. Dautenhahn and A. Billard, "Studying robot social cognition within a developmental psychology framework," in *Advanced Mobile Robots, 1999.(Eurobot'99) 1999 Third European Workshop on*. IEEE, 1999, pp. 187–194.
- [29] K. Dautenhahn, "Getting to know each other - artificial social intelligence for autonomous robots," *Robotics and Autonomous Systems*, vol. 16, no. 2, pp. 333–356, 1995.
- [30] —, "I could be you: The phenomenological dimension of social understanding," *Cybernetics & Systems*, vol. 28, no. 5, pp. 417–453, 1997.
- [31] K. Dautenhahn, B. Ogden, and T. Quick, "From embodied to socially embedded agents—implications for interaction-aware robots," *Cognitive Systems Research*, vol. 3, no. 3, pp. 397–428, 2002.
- [32] J. Saunders, C. L. Nehaniv, K. Dautenhahn, and A. Alissandrakis, "Self-imitation and environmental scaffolding for robot teaching," *International Journal of Advanced Robotics Systems*, vol. 4, no. 1, pp. 109–124, 2007.
- [33] C. Lyon, C. L. Nehaniv, and J. Saunders, "Interactive language learning by robots: The transition from babbling to word forms," *PLoS ONE*, vol. 7, no. 6, p. e38236, 2012.
- [34] J. Saunders, H. Lehmann, Y. Sato, and C. L. Nehaniv, "Towards using prosody to scaffold lexical meaning in robots," in *Development and Learning (ICDL), 2011 IEEE International Conference on*. IEEE, 2011.
- [35] J. Saunders, H. Lehmann, F. Foerster, and C. L. Nehaniv, "Robot acquisition of lexical meaning: Moving towards the two-word stage," in *Proceedings of 2012 IEEE Intl Conf on Development and Learning and Epigenetic Robotics (ICDL 2012)*. IEEE, 2012.
- [36] F. Förster, C. L. Nehaniv, and J. Saunders, "Robots that say 'No'," in *Proceedings of the 10th European conference on Advances in Artificial Life: Darwin meets von Neumann - Volume Part II*, ser. ECAL'09. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 158–166.
- [37] F. Förster, "Robots that Say 'No': Acquisition of Linguistic Behaviour in Interaction Games with Humans," Ph.D. dissertation, University of Hertfordshire, 2013 (in prep), 2013.
- [38] K. Fischer, J. Saunders, H. Lehmann, C. Nehaniv, K. Lohan, K. Pitsch, K. Rohlfing, and B. Wrede, "Contingency allows the robot to spot the tutor and to learn from interaction," in *Procs of 2011 IEEE Int Conf on Development and Learning*. IEEE, 2011.
- [39] K. S. Lohan, K. Rohlfing, J. Saunders, C. Nehaniv, and B. Wrede, "Contingency scaffolds language learning," in *Development and Learning and Epigenetic Robotics (ICDL), 2012 IEEE International Conference on*. IEEE, 2012.
- [40] E. Myin and D. D. Hutto, *Radicalizing Enactivism: Basic Minds without Content*. Cambridge, MA: The MIT Press, 2013.
- [41] P. Dini, C. L. Nehaniv, A. Egri-Nagy, M. J. Schilstra, D. Schreckling, J. Posegga, G. Horváth, and A. J. Munro, "BIOMICS project: Biological and mathematical basis of interaction computing," *International Journal of Unconventional Computing*, vol. 8, no. 4, pp. 283–287, 2012.
- [42] F. J. Varela, *Principles of Biological Autonomy*. North Holland New York, 1979.
- [43] C. L. Nehaniv, "Sensorimotor experience and its metrics: Informational geometry and the temporal horizon," in *Evolutionary Computation, 2005. The 2005 IEEE Congress on*, vol. 1. IEEE, 2005, pp. 142–149.
- [44] N. A. Mirza, C. L. Nehaniv, K. Dautenhahn, and R. te Boekhorst, "Grounded sensorimotor interaction histories in an information theoretic metric space for robot ontogeny," *Adaptive Behavior*, vol. 15, no. 2, pp. 167–187, 2007.
- [45] J. Crutchfield, "Information and its metric," in *Nonlinear Structures in Physical Systems – Pattern Formation, Chaos, and Waves*, L. Lam and H. C. Morris, Eds. Springer-Verlag, 1990, pp. 119–130.
- [46] C. L. Nehaniv, N. A. Mirza, and L. Olsson, "Development via information self-structuring of sensorimotor experience and interaction," in *50 years of artificial intelligence*, M. Lungarella, R. Pfeifer, F. Iida, and J. Bongard, Eds. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 87–98.
- [47] L. S. Vygotsky, *Mind in Society*. Harvard University Press, 1978.
- [48] C. L. Nehaniv, "Meaning for observers and agents," in *Intelligent Control/Intelligent Systems and Semiotics, 1999. Proceedings of the 1999 IEEE International Symposium on*. IEEE, 1999, pp. 435–440.
- [49] C. L. Nehaniv, K. Dautenhahn, and M. J. Loomes, "Constructive biology and approaches to temporal grounding in post-reactive robotics," in *Proceedings of the SPIE*, 1999.
- [50] J. Saunders, C. L. Nehaniv, and C. Lyon, "The acquisition of word semantics by a humanoid robot via interaction with a human tutor," *New Frontiers in Human-Robot Interaction*, vol. 2, 2011.
- [51] M. Tomasello, *Origins of Human Communication*. Cambridge, MA: MIT Press, 2008.
- [52] P. ten Have, *Doing Conversation Analysis - A Practical Guide*. London: Sage Publications, 2007.
- [53] G. Gergely and G. Csibra, "Teleological reasoning in infancy: the naive theory of rational action," *Trends in Cognitive Science*, vol. 7, no. 7, 2003.