

JOANNA RĄCZASZEK-LEONARDI<sup>1</sup>, NICOLE ROSSMANITH<sup>1</sup>,  
IRIS NOMIKOU<sup>2</sup>, KATHARINA J. ROHLFING<sup>3</sup>

<sup>1</sup> University of Warsaw, Warsaw, Poland

<sup>2</sup> University of Portsmouth, Portsmouth, United Kingdom

<sup>3</sup> Paderborn University, Paderborn, Germany

## LEVELS OF COORDINATION IN EARLY SEMANTIC DEVELOPMENT

In this paper, we show that early interaction can be seen as comprising of strands of coordinated activity on multiple levels and timescales. In tracing the development of such multilayered organization from an embodied and situated perspective, we underscore the role of the reliable presence of the structured environment, an enacted niche, supporting the segregation and integration of participatory interaction strands. This perspective allows us to study the development of social coordination not only in terms of development of individual skills but, crucially, as a change of participatory emergent patterns, a transformation in engagement. We illustrate this approach with some results from the collaborative research project on Early Semantic Development (EASE). Using qualitative microanalysis combined with quantitative dynamical time series analyses, we were able to demonstrate several layers of such organization: from local forms of coordination, such as basic informational coupling within a modality, and the emergence of specific social affordances, to more global co-action structures such as affect imbued ‘action arcs’ – dynamic action contours with a beginning, build-up, climax and resolution, co-enacted by participants. Pointing to future work, we underscore the potential of these global structures to contribute to the emergence of more complex interactions, such as composite activities within ‘pragmatic frames’, narratives, or language.

*Keywords:* language development, interaction dynamics, interpersonal coordination, semantic development

## Introduction: Sense-Making is Participatory, Multilevel and Multi-Scale

The last several decades of research on the first months of development showed us, on the one hand, how active and prepared for co-action infants are, participating in interactions from the earliest moments of life. On the other hand, we realized how strongly the niche, or *Umwelt*, of the child is structured towards this participation. The key aspect of it is the social world, enacting for the child a sort of ‘social physics’<sup>1</sup>, with its specific timing and causal structures, where perceivable events specify the actions to be performed and where own actions systematically trigger events.

The Early Semantic Development (EASE) project, realized collaboratively by the University of Warsaw, the Paderborn University and the University of Portsmouth, was geared towards understanding the semantic, sense-making skills that develop within active interaction participants, and towards identifying the structures of the natural, socially enacted niche that prompt and sustain this development. The main purpose was to understand the rich, interactive and meaningful ground that is available for linguistic development.

This project was based on a significant body of work on early development in interaction, “development of communication itself” (Kaye, 1979), which has been carried out at least since the 1970s. Firstly, this research poignantly demonstrated that sense-making, both in language and action, is participatory and distributed (Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979; Bateson, 1975, 1979; Brown, 1973; Bruner & Watson, 1983; Kaye, 1979; Kaye & Fogel, 1980; Newson & Newson, 1975), and thus, accounting for it requires not only studying individual cognitive processes but also the collective ones (for more recent accounts see De Jaegher & Di Paolo, 2007, 2008; Fogel, 1993). Secondly, it pointed to sense-making occurring on multiple levels as participants make eye-contact and affect each other taking turns, negotiate and establish (shared) routines, learn local dependencies as well as more complex hierarchically organized actions, conceive of and co-ordinate activities in terms of “(shared) goals and sub-goals”, or as “narratives”.

What has changed since the 70s is that we are now lucky to have an enriched conceptual and methodological toolkit at our disposal to address naturalistic interactions. An important theoretical development is that the dynamical systems approach to development, which explicitly emerged at the end of the last century (Fogel & Thelen, 1987; van Geert, 1994; Thelen & Bates, 2003; Thelen & Smith, 1994), recently seems to be coalescing into a robust conceptual

<sup>1</sup> We use the term ‘social physics’ to point to the near inevitability and nearly law-like dependencies, which comprise the social environment enacted around and with the child. This is a notion akin to Brette’s ‘subjective physics’ (Brette, 2019), crucially enriched by a social dimension. We agree with one of the reviewers that the term ‘physics’ might seem to strip the notion of the value-realizing aspect. This is not our intention, as both social and subjective physics are shaped within a value-realizing eco-system (for the account of values as eco-system demands see e.g., Hodges and Baron, 1992; Hodges, 2009, Hodges & Raćzaszek-Leonardi, submitted).

framework, implicating new methods for data gathering and analysis. This framework, due to its approach to complexity, is well suited to account for multi-level phenomena, such as individual processes nested within collective ones, which gives the notion of ‘emergence’ a more graspable and measurable form. It is also adequate for showing how multiple modalities integrate into interactive patterns on different timescales. In parallel, new developments in data collection methods help satisfy the requirement for particular kinds of data within such a framework. Dense multimodal time series can be obtained from participants and on systemic levels, reflecting dynamic coupling rather than ‘responses to stimuli.’ Cheap and easy-to-use (though still time intensive) video recording technology, open source transcription and analysis tools such as ELAN (Wittenburg, Brugman, Russel, Klassmann, & Slocetjes, 2006) and PRAAT (Boersma, 2001), as well as automatic motion capture tools such as TLD (Kalal, Mikolajczyk, & Matas, 2010), allow to take a detailed look at multiple strands of action in parallel, in several modalities and in good temporal resolution.

Obviously, the range, complexity, and the depth of questions that the communicative organization of early behavior generates is immense, and thus, not surprisingly, the EASE project brought more new questions than definite explanations for the processes under investigation. In this paper, we would like to give a brief overview of the findings and focus on some insights about the organization of early interaction that we gained through our research. We hope to show that at least some of the important novel questions present in the field can now be posed in a more definite and clearer way, which will inform future research.

Our ecological and dynamical perspective has led us to consider the ways in which participants are in close contact with each other, providing social niches for each other, mutually shaping their actions on-line and ‘educating’ or ‘formatting’ them for future interactions. We will thus start with the findings regarding basic, modality-specific informational coupling (through gaze and vocalizations), showing developmental progression towards temporal coordination, which allows for a greater precision of the mutual influence. Progressing towards more specific mutual control in interaction, our analyses reveal the structure of enacted participatory activities, in which particular multimodal local contingencies emerge. These activities educate attention and action for picking up those moves in interaction that become interactional affordances and which, in turn, provide interactional affordances for others. In contrast to many traditional approaches to explaining the organization of activities, here we consider specific goals or tasks to constitute only a subset of the shaping forces, with global values of agentivity, mutual respect and good coordination being equally, if not more, important. The emergent value system embodied in particular contingencies and affordances, makes it possible to establish certain behaviors as particularly valued. We will

illustrate how this process is engaged in the case of language, where a coherent dialogical vocal modality is shaped by tightly relating vocalizations and utterances and placing them within a coordinated whole of an ongoing activity.

The above research mainly focused on levels of interaction organization pertaining to sequencing and timing of the particular adjacent actions. In several contexts, we acutely felt the necessity and urgency to address more complex interactive levels. For example, local dependencies felt hopelessly insufficient when accounting for the developing ability of infants to differentiate the degree of their participation across sequences of actions: the sensitivity to local cues when playing peek-a-boo, for example, seemed to be moderated by what has happened several ‘steps in a game’ before (Nomikou, Leonardi, Radkowska, Raçaszek-Leonardi, & Rohlfing, 2017). Similarly, research on language learning, closely connected to the EASE project (Heller & Rohlfing, 2017; Rohlfing, Wrede, Vollmer, & Oudeyer, 2016), demonstrated the necessity of going beyond simple object–word associations and understanding the role of a word in larger, shared ‘pragmatic frames’, which develop through a history of joint practice. Equally importantly, our insights into the process of the emergence of symbolic communication in early development (Raçaszek-Leonardi, Nomikou, Rohlfing, & Deacon, 2018) rely on recognizing the role of the systemic level of utterances and how those relate to more complex action organization and control.

One of the PhD projects at Portsmouth (Rossmanith, 2017) was carried further within the EASE project to help set up the necessary stepping stones towards these larger units of organization. The work tackled the problem of the development of ‘joint practice’ in a more holistic way and from a complementary perspective: facing the richness, complexity, and messiness of everyday natural interaction, the central question became how participants manage to jointly practice these complex activities from early on. This question invites us to look at learning and development as “transformation in engagement” (Rossmanith et al., in preparation), where both, participants’ skills and the enacted niche, get increasingly organized over time, giving rise to larger scale and more complex differentiated forms of interaction and coordination, including narratives and language (also see Heller & Rohlfing, 2017). In this paper, we can show only a preview of the multiple steps towards such complex organizations. We will underscore the role of the jointly enacted energetic and emotional contours (‘action arcs’).

This paper will be organized according to the complexity of the levels of interaction organization, which our work addressed. We will thus focus on 1) modality-specific informational coupling, in our case, through gaze and vocalization; 2) the emergence of particular multimodal local contingencies, which educate attention and action for picking up and shaping interactional affordances; 3) showing that even such local contingencies are shaped in ways that realize global values. We will also point out 4) how, having points 1-3 in

place, segregating functional strands of co-action becomes possible by making certain behaviors particularly valued, and placed within another important organization, in our case, a coherent dialogical vocal modality, which, as we claim, is crucial for development of language as a symbolic system. Finally, we devote a bit more space to the issue of 5) organization of larger units of interaction, showing how natural ‘action arcs’ may aid in shaping longer sequences of action and aid in constituting more complex ‘wholes’, that subsequently enable novel forms of engagement. We still feel these insights are only the beginning of a long journey linking the levels towards greater complexity. However, as we underscore in Conclusions, seeking important shaping forces in engagements within co-action points to encouraging directions for future research, with a goal of integrating the individual cognitive perspective with the constitutive role of multilevel participation.

## **Early Orchestration of Modality-Specific Informational Coupling**

Synchronizing and coordination in living dynamical systems require informational coupling<sup>2</sup>. In human interaction, vision plays a crucial role in this coupling (Stern, 1971, 1974; Reddy, 2008; Reddy, Markova, & Wallot, 2013), becoming a key modality very early on for engagement, in-forming partners’ activities within interaction. Other modalities also carry information about the partner and his/her engagement in a system, thus constructing interaction depends, among others, on weaving multimodal entrainment of the partners. As Mary Catherine Bateson wrote, describing early dyadic mother-infant proto-conversation, “we see infants and mothers gazing at each other, as each smiles and vocalizes, apparently with pleasure and a sort of delighted courtesy. As with adult conversation, there is near-constant communication in one modality (visual) and intermittent, alternating communication in another” (Bateson, 1975, p. 101).

Extending existing research (e.g., Bakeman & Adamson, 1984; Perra & Gattis, 2012), we complemented the use of the aggregate measures with novel analyses, which help elucidate the relative timing of events with greater detail. Our goal was to show how this informational coupling changes with age and how it varies with modality. We first coded gaze behavior of the infants and mothers in three categories: gaze at the face of the partner, gaze away, and gaze at objects during a regularly practiced activity (diaper change, Bielefeld corpus, Nomikou & Rohlfing, 2011) at 3 time points: when children were 3, 6, and 8 months old (17 dyads). Next, we ran cross-recurrence analysis (Webber Jr & Zbilut, 2005), which illustrates how often similar behaviors occur in

<sup>2</sup> Informational coupling is a notion used in dynamical systems approaches to human behavior. As described, e.g., in Kelso (1994, 1998), in contrast to coupling in non-living systems, living systems coordinate via meaningful informational quantities, which are usually selected relational variables. They become perceivable “differences” (relations) that are effective in co-regulating (inter)action. Here we refer to such ability of coupling within a selected medium (here, modality).

partners within the dyad, and constructed so-called diagonal recurrence profiles. Diagonal recurrence profiles illustrate how much given behaviors from the dyad members match at the same time (lag 0) and at other lags with respect to each other (Figure 1). From the analyses of gaze behaviors, we saw (Figure 1a) that at 3 months, there is much more recurrence in general (there is more steady ‘gazing-at’ both by the mother and by the infant, than in the months 6 and 8). However, when we normalized the profiles, we saw (Figure 1b) that at 6 and 8 months of life, the gazing-at behavior—even if less frequent overall—is coupled much more tightly, as the peaks at lag 0 are significantly higher and sharper than at other lags. Additionally, we can see a pattern at 8 months, when the profile gets less symmetric, which might point to a reactive behavior in a child to the mother’s gaze, possibly cued by both the gaze and the vocalizations of the mother. This paves the way to effective coupling, through shorter eye gaze contacts, which become conventionalized (Rohlfing & Nomikou, 2014). The results of our study point to the infants’ participation from the beginning: crucially, according to the profiles, infants are not only motivated to participate by their mothers’ behavior but they also seem to play an active role and to engage their mothers.

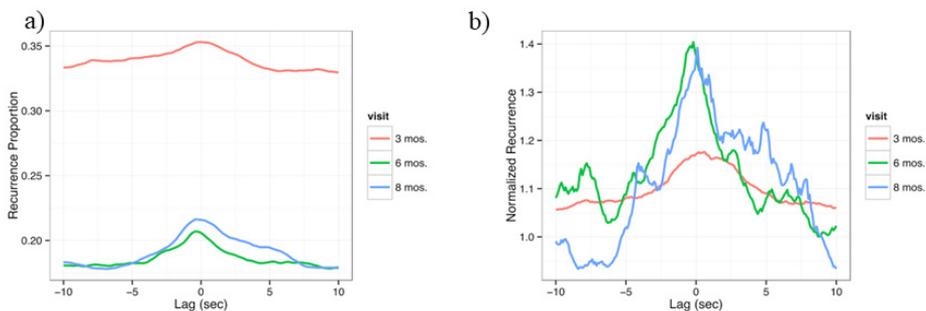


Figure 1. Diagonal profiles showing synchronization of gaze in mother-infant dyads. a) Non-normalized and b) Normalized. The left side of the profile indicates the infant leading the interaction (Figure adapted from Nomikou et al., 2016)

An interesting contrast to this gaze behavior is demonstrated in the vocal modality, in which, as in Bateson’s quote, a different pattern emerges. We know from other research that mothers create pauses in their speech and behavior to accommodate the infant’s actions and that vocal modality, due to its general and universal turn-taking structure (Stivers et al., 2009), should be taking this organization early on (see also Abney, Warlaumont, Oller, Wallot, & Kello, 2017; Warlaumont, Richards, Gilkerson, & Oller, 2014). We coded all vocalizations of mothers and infants — excluding vegetative sounds, again in the same situations and at the same time-points: 3, 6 and 8 months of age. In Figure 2, using the same cross-recurrence diagonal profile method, we show the longitudinal pattern of this coordination. In the

third month of life (red), we see almost no pattern of turn-taking (which should show as a valley at lag 0, meaning that when one partner vocalizes the other does not). We see however a peak on the left, reflecting mothers frequently following infants' vocalizations. This pattern seems to give way to a clearer turn-taking and less pronounced following by the mother, visible in Figure 2 (blue and green lines), as a) a deepening valley at lag 0 and b) more symmetrical dialogical profile, which shows no distinct leader-follower (there is a significant decrease in the tendency for mother's following, i.e., the peak on the left diminishes; for details see Leonardi, Nomikou, Rohlfing, & Rączaszek-Leonardi, 2016).

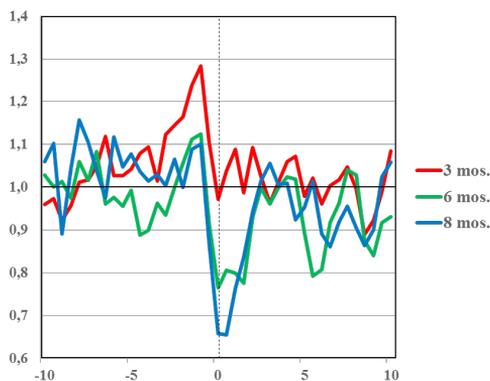


Figure 2. Diagonal profiles showing synchronization of vocalization in mother-infant dyad. The left side of the profile indicates the infant leading the interaction

This level of organization of modality-specific informational coupling, while no doubt to some degree activity-dependent, may nevertheless be expected to show this increasing mutuality patterns in development, which underlies not necessarily tighter, but rather more efficient coordination.

It should be noted that our methods so far did not allow for detection of multimodal coordination, which might be particularly interesting in the child's early development. While infants actively engage in gaze behavior, their vocal behavior is less balanced as their mothers tend to follow them more than they are able to follow their mothers. It is plausible to argue that at this early age, vocal turn-taking might be highly embedded within bodily actions — a point that was raised by Nomikou and Rohlfing (2011) and subsequently developed in Rohlfing, Leonardi, Nomikou, Rączaszek-Leonardi and Hüllermeier (2019).

## Multimodal Local Contingencies: How Does a Movement Become a Message?

This ubiquitous coupling allows for further, more task-and action-specific patterns to emerge. In our work, we strived to show how this process might be due to education of attention and perception (Gibson & Pick, 2000; Zukow-Goldring, 1997, 2012), where, with time, infants are guided to attend to the actions of

others as opportunities for their own actions, as well as to learn to use their own actions to influence the partner in interaction. Understanding such interaction-controlling, embodied skills, we assumed, will add significant explanatory value to the putative internal individual cognitive constructs developing at this time.

We performed a set of qualitative microanalyses of several short interactions within co-actions: dressing a child, interaction with objects, or interactional closings. We employed the theoretical framework of Gibsonian perception-for-action, which rather than cognitive ‘event understanding’ prioritizes development of the sensitivity to action possibilities (affordances) that open in specific situations. This framework has been shown by Heft (1989) to be suitable to account for intentional action development, when intentionality is defined (after Merleau-Ponty, 1965), as embedding actions in projects realized in a particular environment. Affordances are thus learned not just as body-specific relations to the environment but as body-in-action-specific relations. In early interactions, these actions are mostly social and participatory. Bringing together these theoretical considerations with our empirical mother-infant interaction data, we have shown how, in early interactions, affordances can indeed be learned ‘movement first’, with caregivers enactments giving a social systematicity, immersing a child in joint actions and goal-realizations from early on (Raćzaszek-Leonardi, Nomikou, & Rohlfing, 2013; Raćzaszek-Leonardi, 2016). Moves in interactions become meaningful as controls of mutual behaviors, which guide the unfolding social events.

We showed, for example, how an initially random move of the baby changes the mother’s behavior: a simple turn of head of the infant 1) constrains the direction of the mother’s gaze, 2) prompts her to add elements to her sequences of movements, 3) changes the shapes, timings and trajectories of her movements, 4) changes the topic of her speech, etc. We can say that a head turn (or a gaze shift, which it implies) may become an interactional affordance both for the mother and for the baby. The mother, due mostly to her cultural background, picks up this behavior as an affordance and behaves accordingly: when in interaction, it is culturally ‘proper’ to look where the interlocutor looks. By enacting this behavioral pattern, she makes the gaze of the baby intentional through embedding it in a realized project, recreating and framing the event so that the gaze makes (interactive) sense. In this way, the infant learns about the social consequences of her or his gaze. A history of such interactions, within the reliable ‘social physics’ recreated around the child, educates attention to important aspects of situations linked to gazing and leads to the accruing of the gaze’s ‘power’ to control interpersonal situations. Social affordances are thus established as means to mutually constrain and control behavior.

Even though ‘educating attention’ (Nomikou, Rohlfing, & Szufnarowska, 2013) focuses on the processes within individuals, it should be clear that the individual skills are not only in the service of the interactional engagement but that their manifestation is impossible without such engagement. Interaction, in

one way or the other, has to be there for a gaze to exert its power. The gaze may have effect only within enacted interactive structures and, therefore, representing these structures in the forms of individual mental schemas might not be necessary.

### **Local Contingencies are Value-Realizing**

The above picture of shaping mother–infant interactions and infants’ doings in the world might lead one to think that it can be attained in a rather mechanistic way, through coupling and learning local contingencies in a goal of instrumental control of a partner in interaction. In later work within EASE, we sought to address this critique and pointed out that even at this relatively simple local behavioral contingency level, specific features of these contingencies, especially the timing and imbuing with affect, the visible “pleasure and a sort of delighted courtesy” from Bateson’s quote above, seem to be driven by more global, value-realizing considerations (Rączaszek-Leonardi & Nomikou, 2015).

Like a structure of a fractal, each move in interaction, its particular shape, trajectory, timing of execution, and accompanying emotionality reflects a set of co-present, nested and parallel (heterarchical, see e.g., Hodges & Baron, 1992) values being realized. One cannot thus think about development as learning social affordances just through learning simple local behavioral contingencies, both in the case of learning what others’ behaviors afford (“when somebody looks somewhere, I follow the gaze”) or learning that my movements are affordances for others (“when I look somewhere, I can control their gaze and even move objects through their behavior”). These contingencies are indeed learned but each move is in the grander service: that of realizing multiple values in interaction.

A mother, whose 6-month old child is more interested in her hands than face, after calling the infant’s name, expects his prompt look at her face. When this does not happen after a couple of such callings, she deploys a set of variegated and creative actions that would let her regain the control over the infant’s attention. Since the infant is looking at her hands, she starts to draw circles in the air for the child to follow, then narrowing the circles to touch the infant’s nose and, at this moment, when the attention of the child is thus captured, she leans into this ‘region of attention’ with her face. This creative dance is not in the service of a particular joint action that has to be executed, such as nappy changing or dressing. It is in the service of showing the importance of a set of values. In this case, the values of engagement and respecting each other in interaction are probably the most prominent. They are realized through culturally sanctioned sequences (I call your name; you look at me) and timings of those sequences. This way, specific behavioral patterns of joint enactment emerge. The variety of movements and complex multimodal behaviors employed in a given interaction are geared towards the enactment of such sequences within a desired timing.

Such values as being together, respecting each other and each other's agentivity in a shared project, valuing developmentally-appropriate types of behaviors (e.g., certain types of vocalizations over others, see below) provide guidelines for understanding why only certain, selected affordances are picked out and molded to be interaction controls. Thinking in terms of values allows for better understanding of the details of local contingencies (e.g., the amount of pitch excursion, particular timing of events, etc.) and their flexibility in various contexts, which allow for creativity as long as the moves preserve the values (Rączaszek-Leonardi & Nomikou, 2015; Hodges & Rączaszek-Leonardi, submitted).

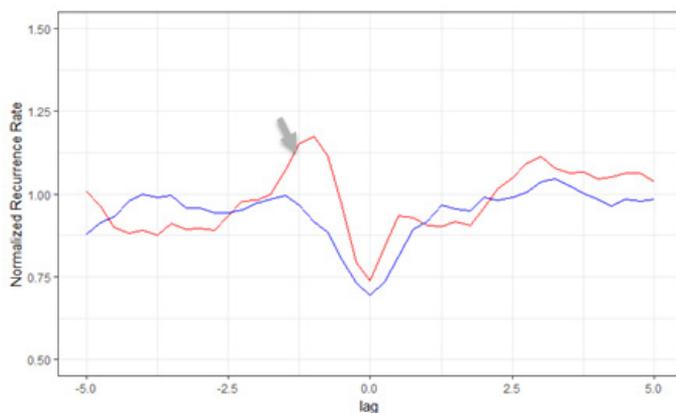
### **Segregating Functional Strands: vocal-vocal coupling privilege**

The sections above give a glimpse of some of the rich structure present in early interactive behaviors and the 'social physics' that helps shape this structure. We could see a qualitative shift in the informational coupling in visual and vocal modalities, which makes coordination more efficient in development towards a global value-realizing system. Importantly, we could identify the means, which are employed for segregation of functionally important patterns of behavior, within or across modalities. The patterns reflect relations among behaviors and contexts, making behaviors mutually constraining in constellations of actions, which preserve culturally appropriate properties and values. The example we give here concerns the emergent vocal modality; it was a focus in our studies on early semantic development as it paves the way to understanding how the systemicity of language may aid in the emergence of symbolic linguistic communication (see below and Rączaszek-Leonardi et al., 2018 for more details on emergent systemicity).

Speech, or more generally, the vocal modality, is present at each point of development. In our work, we show how much structure is brought into interaction by the vocal modality, which is intertwined into multimodal strands of behaviors. However, it is also important to note that from very early on, there is a propensity to differentiate the vocal modality as a patterned strand of actions. This is visible in the qualitative microanalysis of the situations in which partners of interactions vocalize (e.g., Rączaszek-Leonardi & Nomikou, 2015): in the earliest months, infants' vocalizations are met with attention and gaze of the mother, ceasing own vocalization to avoid overlap, and often with a vocal response. What is more interesting, in many cases we observed that vocalizations that are language-like are 'valued' more, i.e., draw more focused attention, connected also to ceasing parents' vocalizations, and are responded to more amply and reliably than other, non-language-like vocalizations, creating a "social feedback loop" (Warlaumont et al., 2014).

Quantitative analyses, both within the vocal modality and multimodal, seem to confirm these observations. Above, we have already showed that vocalizations tend to be organized in turns. In a subsequent study relating

various modalities, we have demonstrated, using both RQA and frequent pattern mining, that vocalization is present in some frequent patterns of intersubjective behavior including also eye gaze coordination (Rohlfing et al., 2019). In another study, using the same corpus, we coded vocalizations not only for their timing but also for their closeness to language sounds. Diagonal profiles, similar to the ones presented in Section 2 above, showed that indeed, already at 6 months of infants' age, the propensity of the mother to vocalize after language-like vocalizations of the infant is greater than to vocalize after non-language-like vocalizations (Figure 3), confirming earlier findings for older children (Warlaumont et al., 2014). In the preliminary analyses, this pattern of dyadic behavior correlated with later linguistic development (Radkowska, Nomikou, Leonardi, Rohlfing, & Rączaszek-Leonardi, 2017).



*Figure 3.* Diagonal cross-recurrence profile of mother-infant vocalizations. At 6 months of infants' age, mothers follow (peak on the left, with a delay roughly within 1.5 s) more reliably those vocalization of infants which are more speech-like (red line) than non-speech-like (blue)

Again, just as in the previously presented processes of education through participation, such dyadic behavior gives a developing child an ample opportunity to shape anticipations concerning the next occurrences in interactions and cues for own behavior. In this case, the systemicity of vocal-vocal exchanges is underscored which, in turn, facilitates learning the systemic nature of language, i.e., noticing that particular utterances tend to follow each other (e.g., in adjacency pairs), or form longer patterns. This is claimed to be vital for achieving more complex interaction control — not only through single utterances but through relations between them. In turn, admitting that relations may have such a constraining role in interaction is a step towards the emergence of symbolic systems out of simpler modes of signification (for a model, see Deacon, 1997; Rączaszek-Leonardi et al., 2018).

## Shared Projects: Organization of Larger Units of Interaction

Exploring how complex forms of organization such as language could emerge from jointly performing everyday activities, researchers' attention is quickly directed beyond the level of local contingencies, to more complex forms of co-ordination and the organization of an activity as a whole. Everyday activities such as putting on socks, changing a diaper, playing peek-a-boo, or telling stories turn out to be considerably complex: featuring several actors and objects, having their intrinsic time dynamics, and comprising multiple events, often nested and hierarchically ordered, with some invariant parts, which however can be realised in various ways by one or the other participant. Neither the local affordance-learning level, nor the pervasive constraints from the realized values explain the skill of participation in a shared project, or the ability to realize that a jointly performed activity is a larger project.

The issue of larger scale action organization and how it is achieved and managed by participants has repeatedly surfaced in philosophy, psychology, cognitive, as well as social sciences, and has been addressed under various headings: "forms of life" (Wittgenstein, 2009/1953), "frames" (G. Bateson, 1955; Goffman, 1986/1974; Minsky, 1974), "scripts" (Schank & Abelson, 1977), "formats" (Bruner, 1985) (for a more detailed discussion, see Rohlfing et al., 2016; Rossmanith, 2017; Rossmanith et al., in preparation). In many approaches within contemporary developmental psychology, this level of organization is typically accounted for in terms of individual capacities and mental representations of knowledge about events.

However, in accordance with insights from recent research, including ours, we focus on factors which are complementary to the individual capacities: the role of joint practice and the ubiquitously and reliably present social niche for the emergence of larger scale complex interactions. Supported by what is known so far about the rich, multilevel organization of interaction, we ask: how do the forms of coordination described in the previous sections contribute to higher level action organization and what other aspects of coordination and information are present in the repeated co-enactment of social practices, which may help orient and properly participate in complex action structures?

In the EASE project, the issue of understanding and organizing larger meaningful units of interaction opened up in several contexts. One of them was the study of infants' agentivity in interaction (Nomikou et al., 2017). We observed, as described above, how, on a global level, the value of agentivity of a child as an independent and responsible human being is instilled, which locally takes a form of a parent encompassing actions of a child so as to make them agentive and responsible. Yet what we observed is that the dependencies are also 'long-distance', concerning events and moves other than immediate contingencies. There are particular moments in interactions at which the agentivity is more called for than at other moments and the ability to evoke it

seems to depend on something else other than the performance of local cues. This prompts to seek structures of participation that are larger than a sequence of 2-3 consecutive actions (in which only neighboring elements count). In Nomikou et al. (2017), we empirically analyzed the interpersonal processes that lead to boosting the active participation in a child. We analyzed a longitudinal video corpus of mother-infant dyads playing peek-a-boo (Szufnarowska & Rohlfing, 2014), investigating the dependencies of behavior at different phases of the game. We observed that in dyads, in which mothers waited longer for their children's actions at particular moments in a game and put more effort and time into preparatory stages by setting up the right engagement for the whole activity, children participated more actively at other points in the game, sometimes not immediately adjacent, for example by covering and uncovering their eyes by themselves.

The level of the 'structured event', or larger organized activity, transpired strongly also in the research associated with the EASE project, which focused on language learning in interaction. Rohlfing et al. (2016) presented an alternative to thinking about word learning (here object naming) as forming "associations" or mappings based on simple contingencies. Our colleagues point to the organization of naming events as more complex interpersonal formats, termed "pragmatic frames", which are set up in word-learning situations. They are multimodal, carefully timed sequences of actions performed within a dyad in order to achieve a common goal. Thus, learning occurs in a unit much more complex and longer than adjacent contingency: it is a multimodal, joint (dialogical) "learning unit" co-constructed with a partner (Rohlfing et al., 2016, p. 1; Heller & Rohlfing, 2017). Especially, in young children, what gathers individual events and behaviors together to such a unit are larger dynamic shared action structures (see below). They may also facilitate the understanding of "shared goals" in the sense that their inner structure follows an energetic contour, according to which a clear change in the environment or a clear change in the emotional way of co-construction (Rossmannith, Costall, Reichelt, López, & Reddy, 2014; Rossmannith et al, in preparation) can be expected at the end.

Equally importantly, our developing framework for emergent symbolic communication relies on coalescing systems of action on the one hand and patterns in vocal modality on the other. This is not a place to describe the model in detail (for this, see Rączaszek-Leonardi et al., 2018). In brief, it is based on relating complex utterances to complex interactive events in such a way that the controlling power of relations, and not only of single acts or utterances, are noted and used. This, we claimed, is facilitated both by the emergent vocal modality (see the section above) and by early routines (including games), which also, necessarily, comprise coherent strands of multimodal actions, involving both participants and (often) objects in directed and carefully timed interactive

events. Understanding how participation in such interactive forms is guided by increasingly complex language is, we claim, a key step to understanding the emergence of language as a symbolic system.

## Action Arcs as Orienting Structures with Intrinsic Dynamics

One piece of the puzzle which may contribute to answering how caregivers and infants coordinate large scale activities and how infants learn to be active co-creators of complex actions was presented in a naturalistic longitudinal study performed at Portsmouth (Rossmanith, et. al, 2014; Rossmanith, 2017; Rossmanith et al., in preparation). The study investigated how infants and caregivers multi-modally coordinate the practice of everyday social object routines and how this changed over the course of the first year. What stood out in the study was that caregivers, besides structuring actions by rhythmically patterning them and marking relevant events by making them salient<sup>3</sup> (for examples see Nomikou & Rohlfing, 2011, Zukow-Goldring, 1997), give actions particular energetic contours, shaping them into dynamic action arcs with a beginning, build up, climax and resolution.

Using ELAN and PRAAT to visualize and analyze different strands of action in parallel, action arcs became most readily visible in the caregiver's intonation contours with intensity and pitch curves literally forming arcs within and/or across utterances and vocalizations. A closer look showed that large scale actions such as shifts in posture also fitted and contributed to the arc shape, with important events, actions or words typically placed at the peak of action arcs. Infants' participation with distinct actions could often be found at the peak of an arc or at the end of it (between arcs).

For the purposes of this summary, two brief examples of everyday infant-caregiver activities at 3 months may illustrate the different ways action arcs were realized at multiple nested levels and time-scales from sub-actions to the overarching activity. Example 1 (Figure 4) demonstrates a simpler stand-alone arc, where a literal arching of pitch and intensity is visible.

*Example 1: 'Single page turn' (as a stand-alone activity). After setting the stage by drawing attention through the surprise exclamation ".h!" and announcing the page-turning action with the question, "What's on the next page?", the mother develops the action arc: by leaning forward and repeating the question followed by two more ".h!" surprise exclamations of increasing intensity and pitch, she builds up tension which is mirrored in the growing arousal of the infant, indicated by her increasing movement, body tension, and facial expression, culminating in her mouth dropping open and a sharp intake of breath just before the climax. After a short hesitation—drawing forth the tension still*

<sup>3</sup> In the sense of creating a local informational pop-out effect, e.g., through local high intensity (compared to neighboring actions) or the convergence of multiple modalities into cross-modal invariances (for more details see Rossmanith et al., in preparation).

further—a sudden quick page turn releases the tension, and the arc levels off and comes to a close in a soft, whispered “There we go” from the mother, coinciding with the infant relaxing and closing her mouth again (Rossmanith et al., 2014).

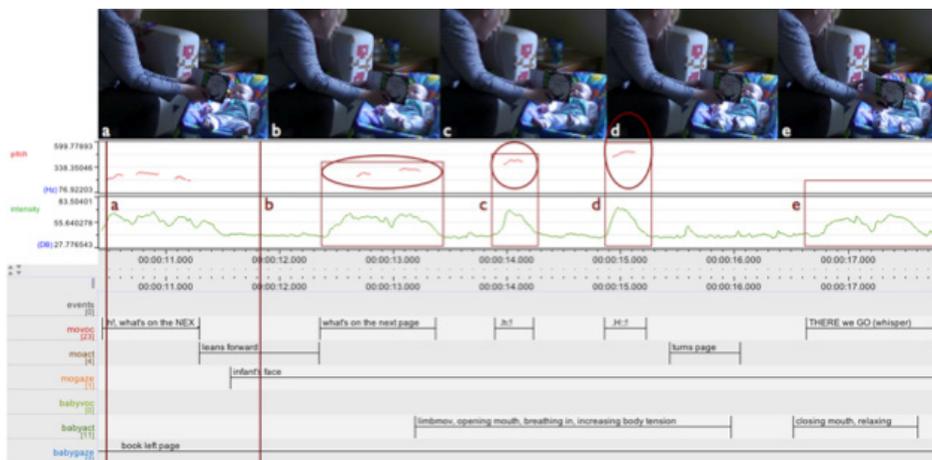


Figure 4. ELAN analysis details of page-turning, showing vocalizations with pitch (red) and intensity (green) curves, as well as the partners' actions and gaze direction. Letters (a–e) map upper row stills to ELAN time

Thus, in this example, the mother shapes a single page turn (usually a prototypical sub-part of book sharing) into an exciting stand-alone activity: here the mother's pitch and intensity contours literally form an arc shape, in concert with her postural shifts. This enables preparing, contextualizing, and marking the crucial page-turning action, and—as the arc is reflected in the infant's responses—co-orienting the participants.

Example 2 (Figure 5) shows a nested structure, where a 3-month-old infant meaningfully participates by contributing a task-functional 'bottom lift' within an arc-shaped sub-routine of a larger action, which is the nappy change. In this ritualized sub-routine we see a) the mother's pitch and intensity contours forming an arc in a 'relative' way, i.e., always returning to and starting from a similar baseline but with the contours of single utterances resembling the respective parts of an action arc (ascending, drawn out, peaking, falling/rising, falling) and falling really low only at the end of the whole activity. Note also b) the infant's distinct action co-constituting an action arc's peak.

*Example 2: Bottom lift in nappy change (nested sub-routine). The mother opens the arc by undoing the first strap of the nappy with a tearing noise vocally marked by a sharp intake of breath “.h!”. After a short detour, “What are you doing?”, responding to and waiting out the infant's unilateral kicking (not belonging to nappy change) (a), the*

*mother starts building up the arc undoing the second strap, establishing eye-contact, inviting the infant in a whisper “you’re REAdy?”, then repeating the invitation a second time with increased intensity “Are you REAdy?”, further creating tension with sequential repetition, the flat drawn out pitch contour and the sudden slight increase in pitch at the end (c). At this point of heightened tension, the infant performs a context-specific and task-relevant co-ordinated whole-body action: tucking up her legs left and right of her body, (head and chin being slightly pulled back), then pulling them closer to her chest—right hand on her right knee—thus lifting her bottom off the mat. The partners’ co-ordinated actions—the infant lifting her bottom and the mother folding down the nappy and marking the co-ordinated action with another “.h!” – occur at and constitute the peak of the action arc (e). After that, the mother responds to and marks the successful completion of the infant’s part in the activity with two emphatic “Thank you!”s. Their falling-rising pitch contours reflect the temporary closure of the current action (h-i), while already segueing into the next phases of the multi-step nappy-change activity, before the mother will finally conclude the overarching action arc running through the activity with an emphatic “ALL done”, spoken twice while maintaining eye-contact with the infant, each with a marked fall across the whole pitch range, and accompanied by two synchronous gentle chest rubs (j-k) – thus creating specific cross-modal invariances, with a specific widely falling contour constituting a memorable “closure event” (for more details, see Rossmanith, 2017).*

## **Action Arcs as General Action Structures and Bridges to More Complex Forms of Interaction**

Such dynamic arc-form has repeatedly caught the attention of researchers in different fields within developmental psychology, prominently and particularly in the context of proto-conversation: it shines through in Brazelton, Koslowski and Main’s (e.g., 1974) early descriptions of attention and withdrawal cycles, and is explicitly discussed as “vitality contours” (Stern, 2010a, 2010b), early “musicality” (Malloch & Trevarthen, 2009), and—carrying further Bruner’s legacy (e.g., Bruner, 1990) – early “narrative forms” (Delafield-Butt & Trevarthen, 2015). The arc shape has traditionally been used to characterize narrative and dramatic form (from Aristotle’s poetics onward); by conceptualizing arc shaped intentional actions as “(early) narratives”, Delafield-Butt and Trevarthen point the reader’s attention to their similar form and to the potential roots of narrative in action.

Building on this, we seek to make a different, complementary move and further explore and differentiate the relationship between (inter)action and narrative. Here, we consider what we will descriptively call ‘action arcs’

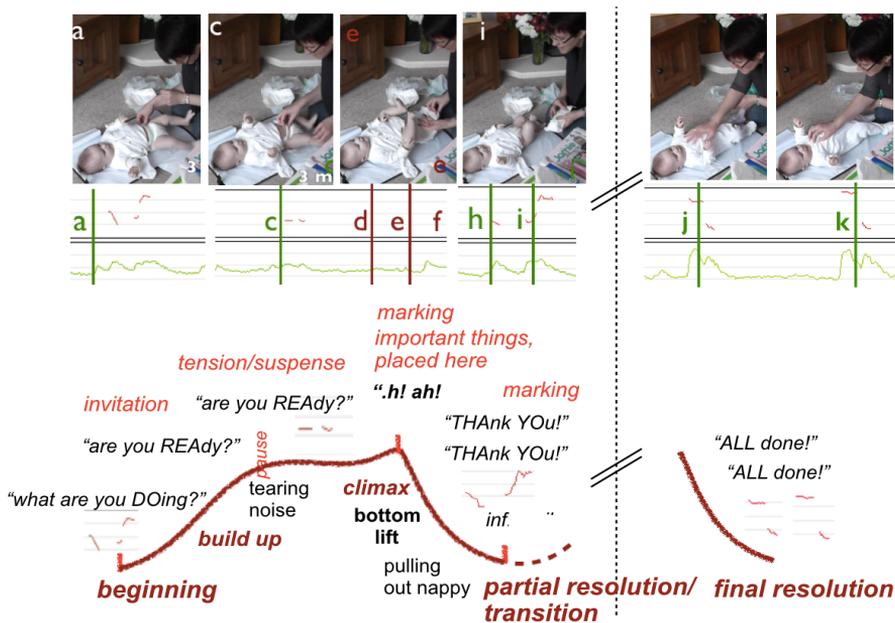


Figure 5. Sub-arc nested in a larger nappy change activity as co-composed by the participants' actions, illustrated through selected details from ELAN analysis: snapshots showing actions, and corresponding pitch (red) and intensity (green) contours of the mother's vocalizations. The inset on the right illustrates the closing of the global arc

as more general, fundamental to action, resonating with Susanne Langer's (1967) broad understanding of the dynamic arc form as a means to capture life processes ("acts") in general. In contrast, here we count "narrative" (including aspects of a narrower, language-related reading, see e.g., Bruner, 1990; Bokus, 1992) to a set of more complex forms of interaction which – aside from exhibiting a general action arc form – typically feature some additional characteristics, such as a composite hierarchical structure, being 'about' something, etc. Trying to put what such complex actions have in common in more general action related terms: they tend to stand out from a particular action context, as self-sustained "objects", arguably defined through the relations of their subparts and to each other. This "object-character" makes them publicly and flexibly available for joint engagement and allows them to be put in relation to the context and particular actions.

These theoretical distinctions and framing allow us to ask and investigate from a new angle, and within one framework, how fundamental forms of action coordination are related to more complex forms of interaction traditionally associated with 'higher cognition', such as language. The concept of action arcs might serve as a bridge to understand the relationships among forms of coordination and interaction and their development.

That action arcs may indeed play a functional role and be utilized by interaction partners themselves as a scaffolding bridge structure in interaction and development, is suggested by new insights from our empirical data (Rossmanith corpus 2011-2013), including: a) the ubiquitous use of action arcs beyond protoconversation, b) infants' surprisingly early and extensive functional contributions (even beyond dyadic infant-directed routines), and c) the quickly increasing complexity of coordination over the first year. These suggest that early joint enacting of action arcs – rather than being merely “behavioral turn taking” and “expression/exchange of affect” having nothing to do with “cognition” yet (Tomasello, Carpenter, Call, Behne, & Moll, 2005) – may indeed play a functional role enabling infants to come to understand action structures and participate in joint activities from the beginning and may, in particular, facilitate the development towards the mastery and co-creation of increasingly complex activities, including, at some point, narratives and language.

Considering how action arcs might facilitate development of more complex forms of action-coordination, two characteristics of these dynamic arc structures seem particularly relevant: 1) Action arcs are pervasive in any form of action (down to breathing, walking, grasping etc.), so their general time-space-intensity dynamics are familiar to even young infants; hence, they may serve as an *intuitive embodied orientation structure*. 2) In a social context, action arcs create ‘sharedness’ in multiple ways: they are co-constructed, they co-affect and constrain the participants; gaining an existence and dynamics of their own, they come to constitute a ‘third thing’, publicly available to jointly relate to – thus *providing a medium, container, and object* for the development of action co-ordination and joint learning.

These two characteristics are utilized in what we widely observed in our study and might be key to how action arcs become effective as a developmental bridge (for details, see Rossmanith 2017, Rossmanith et al., in preparation): caregivers constantly mapped complex activities involving hierarchical, multi-step, goal-directed actions onto distinctively shaped (often exaggerated) familiar action arcs. These actions may contain goals and sub-goals and conventional communicative markers and signals (requests such as “Can you...?” or closing acknowledgements like “Thank you!”), neither of which have their conventional meaning for the infant yet. The action arcs – due to their particular characteristics described above – provide an embodied, meaningful, and to some extent, predictable shared global orientation structure: literally forming an ‘arc’ of tension over a sequence of actions, they provide a directly felt experience of the ‘place’ of a particular action in the larger whole.

Thus, action arcs may function as a bridge in a two-fold way: first, within a particular activity, they link (a) local small scale events to larger scale, global activity structures, and (b) more analogue, sensual aspects of the interaction (e.g., flowing intensity dynamics) to more discrete, conventionalized,

conceptual ones (e.g., a ‘sub-task’, ‘marker/signal’, or ‘word’), helping to bind them together in a meaningful whole. Second, from a developmental perspective, action arcs—by promoting socially shared action parsing and orientation, and simultaneously creating shared action-building-blocks to jointly relate to—may serve as a scaffolding bridge towards the understanding and mastery of more complex activities, featuring complex hierarchical organization, role distribution, etc. For more details, see Rossmanith et al. (in preparation), where we (taking a close look at a broader sample of everyday interactions) sketch a developmental trajectory step by step, tracing qualitative changes in coordination from 3 to 12 months of infants’ age and discussing the potential functional roles of action arcs and the ways to study them in more detail.

For now, pointing to these more global aspects of action organization makes us aware of the rich in-formative context available to the developing engaged actor. This may facilitate the conceptual passage from basic forms and local parts of coaction, captured by concepts such as entrainment, local contingencies, etc., to ‘higher order’ action organization in complex tasks, narratives or language, typically framed in cognitive terms.

## Conclusions

In this short paper, it was possible only to give a glimpse at the levels of organization present in each, however simple, interaction. While it was impossible to do justice to the intricate nature and the variety of learning and developmental processes that aid in coalescence of those levels, we hope to have carried across the following points:

1. Interaction and participation in joint practice occur from early on with intricate action organization on multiple levels. We showed examples of how multiple, intertwined and co-dependent forms of coordination on these levels can be made visible and studied, using methods which render the emergent systematicities observable: from modality-specific coupling, to local emergence of interactional affordances by educating attention to enacted cues, to larger interactional synergies supported by action arcs and the composite internal action structure.
2. Taking a closer look at these forms of organization and co-ordination as they develop suggests that the differentiated and complexified participation is not only the result of individually acquired knowledge but that interaction itself and the jointly created action structures play a constitutive role in the process of emergence and maintenance of interactive forms.

3. Joint practice with a responsive and competent member of culture and the resulting action structures provide a continuously co-created social niche which a) through its ubiquitous and reliable presence (hence, ‘social physics’) renders skillful participation (and indeed accounting for it) less representation-hungry, and b) asks to be further investigated as a formative force over the history of parent-infant interactions, potentially contributing to both inter- and intra-individual development.
4. Conceiving of development along such lines as “transformation in and of engagement” thus opens a new path to understanding the progression towards more complex forms of value-realizing interaction and participation, including nested, hierarchical action structures, role distribution, reference, narrative, and language.

Our future work is geared at further exploration and better understanding of the social enacted niche. We are especially interested in how multiple modalities serve emergent interactional synergies (Rohlfing et al., 2019), how the transformation of engagement is enabled by nonverbal behavior and how it further changes with language. We seek to further describe and elucidate how different forms of co-ordination play together (including the functional role of action arcs), how this differs across activity contexts (Rossmanith, in preparation), and changes over time, giving rise to increasingly complex forms of interaction and participation (Rossmanith et al., in preparation). In addition, we are interested in how the ability to differentiate strands of coherent co-actions helps predict further development, for example, how the emergence of dialogical, vocal modality supports linguistic and cognitive development.

Neither last nor least, a natural progression is to investigate how EASE (Early Semantic Development) forms a foundation of LASE (Later Semantic Development), which includes the mastery of grammar, abstractness and compositionality of language. Showing that language is realized through particularly placed activities within participative and structured interaction facilitates solution to some stubborn problems in linguistic and psycholinguistic theories. What may be attractive in this approach for the study of language development is a certain ‘reversal’ of thinking about the role and the nature of linguistic forms in interaction. Language is seen not as an independent content-bearing module, but as a system of constraints (Rączaszek-Leonardi, 2012, 2016), which directs and controls the interaction flow. Knowing more about the shape and structure of the interactive dynamics which is the substrate of the control processes will, we hope, pave the way to understanding how, for example, grammatical structures are linked to the needs for complexity of this control. Filling in the framework presented in Rączaszek-Leonardi et al. (2018), we will thus employ observational, experimental and

computational modeling methodologies to demonstrate the constitutive reliance of emergent communicative structures on multilevel participation: both on its history in a given dyad and on ubiquitous presence of involvement within collaborative projects unfolding over multiple timescales and levels of social life.

The final remark concerns the ways of studying such multilevel and multisystem emergent organizations. Reversal of thinking about constructive cognitive processes, not as content rich, representational databases but rather as constraints on ongoing, complex, multilevel and distributed (collective) dynamics, puts on the researchers' shoulders the burden for recognizing and describing both the underlying dynamics and the relevant constraints. The role of qualitative and microanalytic research in this endeavor is difficult to overstate. How it should be linked to more quantitative methods comes forward as a key challenge. The background assumptions of the processual and interaction-dominant nature of cognition and development make dynamical systems methodology particularly helpful. Such methods supply tools for studying dynamical and longitudinal differentiation and integration processes and for linking the levels across timescales and—most importantly—across systems: from individual to dyadic to collective.

## Acknowledgement

This work was funded by the Beethoven UMO-2014/15/G/HS1/04536 grant of the Polish NCN and German DFG to the authors.

## References

- Abney, D. H., Warlaumont, A. S., Oller, D. K., Wallot, S., & Kello, C. T. (2017). Multiple coordination patterns in infant and adult vocalizations. *Infancy*, 22(4), 514–539.
- Bakeman, R., & Adamson, L. B. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Development*, 55, 1278–1289.
- Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The emergence of symbols: cognition and communication in infancy*. New York: Academic Press.
- Bateson, G. (1955). A theory of play and fantasy. *Psychiatric Research Reports*, 2, 39–51.
- Bateson, M. C. (1975). Mother-infant exchanges: the epigenesis of conversational interaction. *Annals of the New York Academy of Sciences*, 263(1), 101–113. <https://doi.org/10.1111/j.1749-6632.1975.tb41575.x>
- Bateson, M. C. (1979). The epigenesis of conversational interaction: a personal account of research development. In M. Bullowa (Ed.), *Before speech. The beginning of interpersonal communication*. (pp. 63–77). London, New York: Cambridge University Press.

- Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glott International* 5, 9/10, 341-345.
- Bokus, B. (1992). Peer co-narration: Changes in structure of preschoolers' participation. *Journal of Narrative and Life History*, 2(3), 253–275.
- Brazelton, T. B., Koslowski, B., & Main, M. (1974). The origins of reciprocity: The early mother-infant interaction. In M. Lewis & L. A. Rosenblum (Eds.), *The effect of the infant on its caregiver* (pp. xxiv, 264). Oxford, England: Wiley-Interscience.
- Brette, R. (2019). Is coding a relevant metaphor for the brain? *Behavioral and Brain Sciences*, 1-44. doi:10.1017/S0140525X19000049
- Brown, R. (1973). *A first language: The early stages*. Harvard: Harvard University Press.
- Bruner, J. S. (1985). The role of interaction formats in language acquisition. *Language and Social Situations*, 31–46.
- Bruner, J. S. (1990). *Acts of meaning*. Cambridge, Mass.: Harvard University Press.
- Bruner, J. S., & Watson, R. (1983). *Child's talk: learning to use language*. New York: W.W. Norton.
- De Jaegher, H., & Di Paolo, E. A. (2007). Participatory sense-making. *Phenomenology and the Cognitive Sciences*. <https://doi.org/10.1007/s11097-007-9076-9>
- De Jaegher, H., & Di Paolo, E. A. (2008). Making sense in participation: An enactive approach to social cognition. In F. Morganti, A. Carassa, & G. Riva (Eds.), *Enacting Intersubjectivity* (Vol. 10, pp. 33–47). IOS Press. Retrieved from <http://www.sussex.ac.uk/Users/ezequiel/01DeJaegher-DiPaolo.pdf>
- Deacon, T. W. (1997). *The Symbolic Species: The Co-evolution of Language and the Brain*. W. W. Norton & Co., New York.
- Delafield-Butt, J. T., & Trevarthen, C. (2015). The ontogenesis of narrative: from moving to meaning. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01157>
- Fogel, A. (1993). *Developing through relationships: origins of communication, self, and culture*. Chicago: University of Chicago Press.
- Fogel, A., & Thelen, E. (1987). Development of early expressive and communicative action: Reinterpreting the evidence from a dynamic systems perspective. *Developmental Psychology*, 23(6), 747–761. <https://doi.org/10.1037/0012-1649.23.6.747>
- Gibson, E. J., & Pick, A. D. (2000). *An ecological approach to perceptual learning and development*. Oxford University Press, USA.
- Goffman, E. (1986). *Frame analysis*. Boston: Northeastern University Press.
- Heft, H. (1989). Affordances and the body: An intentional analysis of Gibson's ecological approach to visual perception. *Journal for the Theory of Social Behaviour*, 19(1), 1–30.

- Heller, V., & Rohlfsing, K. J. (2017). Reference as an interactive achievement: sequential and longitudinal analyses of labeling interactions in shared book reading and free play. *Frontiers in Psychology*, 8, 139.
- Hodges, B. H. (2009). Ecological pragmatics: Values, dialogical arrays, complexity, and caring. *Pragmatics & Cognition*, 17(3), 628–652.
- Hodges, B. H., & Baron, R. M. (1992). Values as constraints on affordances: Perceiving and acting properly. *Journal for the Theory of Social Behaviour*, 22(3), 263–294. <https://doi.org/10.1111/j.1468-5914.1992.tb00220.x>
- Hodges, B. H., & Rączaszek-Leonardi, J. (submitted). *The necessity of values: Values-realizing theory in action and interaction*.
- Kalal, Z., Mikolajczyk, K., & Matas, J. (2010). Face-tld: Tracking-learning-detection applied to faces. In *Image Processing (ICIP), 2010 17th IEEE International Conference on* (pp. 3789–3792). IEEE.
- Kaye, K. (1979). Thickening thin data: the maternal role in developing communication and language. In M. Bullowa (Ed.), *Before speech. The beginning of interpersonal communication* (pp. 191–206). Cambridge; New York: Cambridge University Press.
- Kaye, K., & Fogel, A. (1980). The temporal structure of face-to-face communication between mothers and infants. *Developmental Psychology*, 16(5), 454–464. doi:10.1037/0012-1649.16.5.454
- Kelso, J. A. S. (1994). The informational character of self-organized coordination dynamics. *Human Movement Science*, 13, 393–413.
- Kelso, J. A. S. (1998). From Bernstein's physiology of activity to coordination dynamics. *Progress in Motor Control*, 1, 203–219.
- Langer, S. K. (1967). *Mind: Essay on human feeling*. Baltimore.
- Leonardi, G., Nomikou, I., Rohlfsing, 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 IC DL-EpiRob*, Cergy-Pontoise, pp. 288-293.
- Malloch, S., & Trevarthen, C. (2009). Musicality: Communicating the vitality and interests of life. In S. Malloch & C. Trevarthen (Eds.), *Communicative musicality: Exploring the basis of human companionship* (pp. 1–11). Retrieved from <http://www.oup.co.uk/pdf/0-19-856628-X.pdf>
- Merleau-Ponty, M. (1965). *Phenomenology of perception*. London; New York: Routledge & K. Paul; Humanities Press.
- Minsky, M. (1974). *A framework for representing knowledge* (Laboratory Memo 306). Cambridge, MA: MIT. Retrieved from <https://web.media.mit.edu/~minsky/papers/frames/frames.html>
- Newson, J., & Newson, E. (1975). Intersubjectivity and the transmission of culture: On the social origins of symbolic functioning. *Bulletin of the British Psychological Society*. Retrieved from <http://psycnet.apa.org/psycinfo/1979-03234-001>

- Nomikou, I., Leonardi, G., Radkowska, A., Rączaszek-Leonardi, J., & Rohlfing, K. J. (2017). Taking up an active role: emerging participation in early mother–infant interaction during peekaboo routines. *Frontiers in Psychology*, *8*, 1656.
- Nomikou, I., Leonardi G., Rohlfing K, Rączaszek-Leonardi, J. (2016). Constructing interaction: the development of gaze dynamics. *Infant and Child Development*, *25*(3), 277–295, DOI: 10.1002/icd.1975
- 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. <https://doi.org/10.1109/TAMD.2011.2140113>
- Nomikou, I., Rohlfing, K. J., & Szufnarowska, J. (2013). Educating attention: Recruiting, maintaining, and framing eye contact in early natural mother–infant interactions. *Interaction Studies*, *14*(2), 240–267. <https://doi.org/10.1075/is.14.2.05nom>
- Perra, O., & Gattis, M. (2012). Attention engagement in early infancy. *Infant behavior and development*, *35*(4), 635–644.
- Rączaszek-Leonardi, J. (2012). Language as a system of replicable constraints. In *Laws, language and life* (pp. 295–333). Springer.
- 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
- Rączaszek-Leonardi, J., & Nomikou, I. (2015). Beyond mechanistic interaction: value-based constraints on meaning in language. *Frontiers in Psychology*, *6*, 1579.
- Rączaszek-Leonardi, J., Nomikou, I., & Rohlfing, K. (2013). Young children’s dialogical actions: The beginnings of purposeful intersubjectivity. *IEEE Transactions in Autonomous Mental Development*, *5*(3), 210–221.
- 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.
- Radkowska, A., Nomikou, I., Leonardi, G., Rohlfing, K. J., & Rączaszek-Leonardi, J. (2017). *Scaffolding vocal development: maternal responsiveness to infant speech-like vocalizations at three, six and eight months*. Poster presented at the 14th International Congress for the Study of Child Language.
- Reddy, V. (2008). *How infants know minds*. Cambridge, MA: Harvard University Press.
- Reddy, V., Markova, G., & Wallot, S. (2013). Anticipatory adjustments to being picked up in infancy. *PLoS ONE*, *8*(6), e65289. <https://doi.org/10.1371/journal.pone.0065289>

- Rohlfing, K. J., Leonardi, G., Nomikou, I., Rączaszek-Leonardi, J., & Hüllermeier, E. (2019). Multimodal turn-taking: Motivations, methodological challenges, and novel approaches. *IEEE Transactions on Cognitive and Developmental Systems*, 1–1. <https://doi.org/10.1109/TCDS.2019.2892991>
- Rohlfing, K. J., & Nomikou, I. (2014). Intermodal synchrony as a form of maternal responsiveness: association with language development. *Language, Interaction and Acquisition*, 5(1), 117–136.
- Rohlfing, K. J., Wrede, B., Vollmer, A. -L., & Oudeyer, P.-Y. (2016). An Alternative to Mapping a Word onto a Concept in Language Acquisition: Pragmatic Frames. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.00470>
- Rossmannith, N. (2017). *Culture in the making: Jointly structuring shared spaces of meaning and action in infant-caregiver-object interactions over the first year of life*. University of Portsmouth.
- Rossmannith N. (in preparation). *Varieties of triads: different patterns of joint attention-and-action-coordination in different ecological infant-caregiver activity contexts and their development over the first year*.
- Rossmannith, N., Costall, A., Reichelt, A. F., López, B., & Reddy, V. (2014). Jointly structuring triadic spaces of meaning and action: book sharing from 3 months on. *Frontiers in Psychology*, 5, 1390. <https://doi.org/10.3389/fpsyg.2014.01390>
- Rossmannith, N., et al. (in preparation). *From jointly enacting affect-imbued action arcs to co-creating systemically structured sense-and-action scapes: proposing a pathway towards increasing participation, cooperation, and symbol use*.
- Schank, R. C. & Abelson, R. P. (1977). *Scripts, Plans, Goals and Understanding: an Inquiry into Human Knowledge Structures*. L. Erlbaum, Hillsdale, NJ.
- Stern, D. N. (1971). A micro-analysis of mother-infant interaction: Behavior regulating social contact between a mother and her 3-1/2-month-old twins. *Journal of the American Academy of Child Psychiatry*, 10(3), 501–517. [https://doi.org/10.1016/S0002-7138\(09\)61752-0](https://doi.org/10.1016/S0002-7138(09)61752-0)
- Stern, D. N. (1974). Mother and infant at play: The dyadic interaction involving facial, vocal, and gaze behaviors. In *The effect of the infant on its caregiver* (pp. xxiv, 264). Oxford, England: Wiley-Interscience.
- Stern, D. N. (2010a). *Forms of vitality: exploring dynamic experience in psychology, the arts, psychotherapy, and development*. Oxford; New York: Oxford University Press.
- Stern, D. N. (2010b). The issue of vitality. *Nordic Journal of Music Therapy*, 19(2), 88–102. <https://doi.org/10.1080/08098131.2010.497634>
- Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., ... & Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, pnas–0903616106.

- Szufnarowska, J., & Rohlfig, K. J. (2014). Enfolding interaction with two-month-olds. In *Proceedings of the 16th European Conference on Developmental Psychology* (pp. 213–218). Monduzzi Editore Lausanne, Switzerland, Bologna.
- Thelen, E., & Bates, E. (2003). Connectionism and dynamic systems: Are they really different? *Developmental Science*, 6(4), 378–391.
- Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, Mass.; London: MIT Press.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28(05), 675–691. <https://doi.org/10.1017/S0140525X05000129>
- Van Geert, P. (1994). *Dynamic systems of development: Change between complexity and chaos*. Harvester Wheatsheaf.
- 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.
- Webber Jr, C. L., & Zbilut, J. P. (2005). Recurrence quantification analysis of nonlinear dynamical systems. *Tutorials in Contemporary Nonlinear Methods for the Behavioral Sciences*, 26–94.
- Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: a professional framework for multimodality research. In *5th International Conference on Language Resources and Evaluation (LREC 2006)* (pp. 1556–1559).
- Wittgenstein, L. (1953/2009). *Philosophical Investigations* (Revised 4th ed.). Oxford: Wiley-Blackwell.