Time-to-smile, time-to-speak, time-to-resolve: timescales for shaping engagement in language

Joanna Rączaszek-Leonardi [a,*], Krzysztof Główka [a], Iris Nomikou [b], Nicole Rossmanith [a]

[a] Human Interactivity and Language Lab, Faculty of Psychology, University of Warsaw, Poland
[b] Department of Psychology, University of Portsmouth, UK

A B S T R A C T

This paper is a further step toward the construction of an ecologically valid framework for the emergence of symbolic communication in development. Building on recent advancements, which view language learning in terms of the increasingly skillful usage and navigation of interactive affordances within a re-enacted social world, we present a more detailed investigation of the timing of such affordances and the role of that timing in the progressive structuring of communication toward symbolic language. Seeking conceptual similarities with an action-based ecological framework, we extend the concepts of that framework to the context of early social interaction. We abandon a conception of language learning by infants as based on sound-reference mappings and envision this process as an attunement to *structures in time,* which entails the infant becoming increasingly able *to engage in those structures*, that is, to navigate and shape the relevant interactive dynamics skillfully. We illustrate attunements to several timing orders that are essential to the emergence of language: i) to participants' actions within routines, which leads to the emergence of social affordances; ii) to partners' vocalizations within an individuated layer, which supports perception of the systemic properties of language; and iii) to larger structures in time, shaped by energetic and affective envelopes, which facilitates discerning important semantic units. The goal of this work is to show how embodied interactions gain their linguistic character for the infant and how the structure of interactive engagements becomes increasingly complex and language-dependent without ever becoming fully ungrounded from interaction. Highlighting the importance of timing for learning to participate should improve our understanding of the progressive saturation of language with interactive structures for a child on the one hand and our understanding of the structuring of language as an interactive control on different timescales on the other.

© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction: time as time-to

Time is vital in ecological psychology's conception of action and perception in the world. Cognition does not consist of detecting punctate, momentary stimuli (Gibson, 1960) and deploying discrete, momentary reactions to them; it does not

proceed in sequences of separate steps, which are in principle reversible and hence timeless. Rather, it depends on a living, acting organism's adaptive coupling to relevant changes in the environment, changes which, with experience, become significant dynamic events in the world (Port and van Gelder, 1995; Kelso, 1995). Such events, with their durations, rhythms, orientations and forms of coupling, arguably constitute experienced temporality.

Importantly, information about structures-in-time comes from multiple sources. The rich structuring of the environment, which constrains the organism's actions, would not be possible without the organism's active movement (Thelen and Smith, 1994). Such movement, given its intrinsically temporal dimension, produces proprioceptive information for the organism and an awareness of one's own changing position (Gibson, 1979). Crucially, this movement provides the felt experience of self-movement (Sheets-Johnstone, 1999), which also has its temporal dimension (Stern, 2000), and changes in the world are perceived relative to this felt experience (Kelso & Fuchs, 2016; Popova and Rączaszek-Leonardi, 2020). The temporal characteristics of the surrounding world and our engagement with it are thus at least as important and informative for cognitive and linguistic functioning as their spatial characteristics. The structuring of significant events in time pertains not only to the short time spans involved in the organism's here-and-now but also to other time scales relevant to the "multiple projects" in which we are engaged (Merleau-Ponty, 1962/2005; Donald, 2007; Enfield, 2013) and of which the here-and-now is an integrating point.

Recognizing the historical, multitimescale and coconstitutive nature of the organism–world relation makes it evident that organisms do not merely interact with their environments but rather actively shape them via processes of niche construction (Laland et al., 1999; Costall, 1995; Reed, 1995). Organisms accomplish this task not only by changing the material landscape and creating artifacts but also, in the case of social species, by cocreating a realm of social routines, practices, norms and institutions. Intuitively, niche construction calls to mind spatial organizations in the form of material artifacts: nests, dwellings, buildings or cities; however, the social realm hinges on patterns and rhythms of engagement, which, in turn, may change our relation to the inanimate world (Vygotsky, 1978; Costall, 1995; De Barbaro, Johnson & Deák, 2013; Di Paolo, 2016). Thus, the social world, unlike the—perhaps more evident and graspable—world of material artifacts, relies crucially on structures in time of which the utterances that we produce in interactions are a prime example.

The importance of the social niche that is created by agents' active engagement is particularly evident when considering infants: newborns are largely dependent on their social environment to tend to their needs as well as to allow them to engage with and learn about the world. Thus, the most significant aspect of the surrounding "environment" for an infant is not the inanimate physical world, which must be perceptually categorized and reacted to, but the enacted social world, which requires skillful engagement. In contrast to the physical world, the social world is responsive to infants, and this world is formatted by culture to involve them. Infants, in turn, are well equipped to engage actively with their social partners (Reddy, 2018; Reddy and Morris, 2004). From the very beginning, when they are not yet capable of locomotion, grasping objects properly or controlling their posture, infants shape their environment to a tremendous extent, in some cultures even becoming the center of activities that would not exist without them. This shaping of the environment facilitates their engagement as fluent communication and interaction partners from an early age.

We use the term "engagement" instead of the omnipresent (both in physical and social contexts) term "interaction" to underscore the bidirectionality of this involvement, the symmetry of dependency between the organism and the environment (Reddy, 2018, 2019). This usage seems congruent with Gibson, who noted that "[t]he relation between the animal and its environment is not one of interaction in any sense of that word that I understand … It's one of, well, reciprocity's not too bad" (Gibson, 1982, p. 234). In the case of the social environment, this reciprocity emerges forcefully and enables a shift in researchers' approach to what learning is taken to be: from acquiring veridical descriptions of objects or events by an individual cognitive system to developing of the skills needed to engage in and sustain relations or ways of participating (Reddy, 2018, 2019; Rossmanith et al., 2014; Trevarthen et al., 2014).

In such a context of dynamic and structured engagements, we can trace the process of language development. The rhythms and structures of engagement are present in infants' worlds from the very beginnings of their experience. The capacity to engage socially by linguistic means emerges within such structured coactions. The study of early language development is particularly important for understanding how symbolic structures can emerge from an ecological perspective. This line of research strives to abandon the notion of a language system as a frozen, rule-based set of discrete forms that are isolated from real-time dynamics, which is common in some approaches within linguistics (see the critiques in, e.g., Rączaszek-Leonardi, 2012; Steffensen, 2011; Steffensen and Fill, 2014). Rather than to explicate how an abstract structure is learned and mapped onto something, the explanatory goal becomes to show how language might coalesce as a distinguishable layer in this dynamic realm, both on the level of individual cognition and on the level of a "language system" that is maintained via culturally instilled routines.

This goal requires tracing the emergence of systematicity that is grounded in dynamics and developing tools for capturing and analyzing that systematicity, with a primary focus on relations in an infant's environment that could be language-like. Learning to understand and engage with language at this early stage undoubtedly requires the infant to discern how utterances are linked to other elements of interaction. However, utterances are accompanied not only by nonverbal actions but also by other utterances. Language development is thus a progression toward such a "double grounding", both in coherent coaction and, simultaneously, in a system of utterances. Each of these two dependencies carries the potential for a partial "ungrounding" from the other; however, they also allow the relation between language and action to become much more complex than any mere word-object mapping (Rączaszek-Leonardi et al., 2018). This framework allows us to capture the

complexity of the relationships that is needed to account not only for the skillful use of language for the participation in interaction, which is grounded in first-person experience and affectivity but also for the ability of a discernible and trans-mittable symbolic structure to play a role in this process.

In the present paper, we aim to discuss this view in further detail by clarifying and identifying several more concrete key processes, which have been described previously only on an abstract level (Rączaszek-Leonardi et al., 2018). We claim that tracing the timings of those processes highlights another dimension relevant to explaining an infant's increasing ability to engage in a variety of interactions. Importantly, this investigation opens up the possibility of conducting research on language use and development by employing methods created within the ecological psychology framework. This is because such investigation concerns the processes of attuning action–perception cycles for social action and coaction, which, according to the perspective discussed above, is a matter of coupling or engagement rather than one of reaction to stimuli.

This characterization of the relevant timings is conceptually parallel to Lee's "time-to-contact" (Lee, 1976), in which the rate of changes in the dynamic optic array specifies the timing of an object's approach and possible collision, information crucial for the observer's active and adaptive engagement. However (at least for now), we do not have any ambition to determine an invariance that can reliably specify infants' behavior—in a rich and varied social environment, this task would be exceedingly difficult, if not impossible. Various "times-to" are specified by individualized and routinized strands of behavior, and they specify actions (such as, in Lee's case, "brake" or "duck" or "shout") rather than an objective phase of an event ("contact"). However, our focus on the relational character of the specifying variables and the holistic treatment of an organism in "reciprocity" with the environment remain similar to the aims of that approach. In our case of language development, these factors specify the "proper" rhythms for engaging in coactions, including engaging with linguistic means.

We chose several particular processes to illustrate important time dependencies: first, the dynamics of coaction, which engages the infant from the beginning and facilitates the shaping of social affordances (*Time-to-smile*); second, the way in which language shapes and is shaped by these dynamics (*Time-to-speak*); and third, the fact that coactions occur at different "frequencies of engagement", which allows for attunement to adjacent events within coaction and language as well as within larger interactive structures (*Time-to-resolve*). For each process, we describe its structure and then illustrate that structure via microanalyses of naturalistic mother-infant interactions and, where possible, present quantitative analyses that capture more general principles. Our aim in this process is to draw attention to the complementarity of methodologies that is necessary for studying such early communicative engagements (De Jaegher and Di Paolo, 2007; Rączaszek-Leonardi et al., 2013; Rossmanith et al., 2014): repeated engagement with recorded interactions leads to relational coding (see, e.g., Nomikou et al., 2016), which can then enter dynamic systems analyses, which are suitable for tracing changes in complex relations over time (Cox and van Dijk, 2013; Van Geert and van Dijk, 2021; Leonardi et al., 2016). We conclude with a perspective on the issue of linguistic systematicity that emerges from our work.

2. Time-to-smile: the affordance dance of early interactions

Even a cursory observation of early interactions with infants makes it obvious that the infant is a person who is treated as an agent and who participates in activities with others from the first minutes of her life. The arrival of the infant changes the flow of social events, making the infant the focal point not only for the mother but also frequently for larger social systems (Rogoff, 2003; Reed, 1995), which, in turn, are more or less "prepared" for such an arrival in terms of routines, rites and institutions. People in the infant's vicinity engage with her: her movements are immediately picked up by other participants, who change their behavior in such a way as to make those movements sensible contributions. In our earlier work (Rączaszek-Leonardi et al., 2013), we focused on the content of such events, showing, for example, how a single, brief gaze from an infant evokes a cascade of events from a mother who ceases her actions and changes them to accommodate the gaze as an intentional contribution to coaction.

Such mundane examples show how elaborate these structured behaviors can be and how, in some sense, they may seem "superfluous" with respect to the simple functional actions required to meet the infant's daily physical needs. The process of changing a diaper or dressing a baby becomes an engaged "conversation" featuring turn-taking, questions, teasing, and play. Namely, such interactive episodes are oriented not only toward realizing the practical goals at hand but, perhaps more importantly, enabling the infant's own values, such as safety, curiosity, and agency, to become aligned with values that are crucial for human interactions, such as care, collaboration, respect, attention and mutual agency (Hodges and Baron, 1992; Hodges, 2007; Hodges and Rączaszek-Leonardi, 2022; Rączaszek-Leonardi and Nomikou, 2015; Valsiner, 1987). Such exchanges are culturally prescribed as series of moves that have a specific order and turn-taking structure and they are affectively loaded, which is crucial for instilling and preserving these key values (Colombetti, 2018; Reddy, 2019; Rossmanith et al., 2014).

The enacted social niche is as real, pervasive, reliable, and "physical" as the nonsocial physical world; however, for an infant, the former is arguably more pertinent than the latter. In contrast with "mere physics", such "social physics"[1] is historically constructed and geared toward actively engaging the child, both in the sense that it meets the needs of an agentive participant and in that this process is shaped by infants to a substantial degree, since infants are often the instigators and centers of activities and since their engagement is a vital constituent of this world (Reddy, 2019). This description resonates with the enactive view of social cognition, in which engagement itself is viewed as an emergent, identifiable level of organization, with its own history and ability to generate meaning via a participatory process enacted by individuals (De Jaegher and Di Paolo, 2007)[2]. Living within such a structured world attunes the infant's actions and perceptions to the behavior of others and to interaction–level relations, which are understood as opportunities for the infant's own skillful engagements (Di Paolo et al., 2008).

The concept of affordance, which is central for ecological psychology, was initially used mainly with respect to the inanimate environment but is suitably general to capture this dynamic social realm (Marsh et al., 2009; Van Dijk and Rietveld, 2017). Movements that occur in interaction, since their specific shapes and timings are directly perceived, afford certain actions at particular moments. We can regard participatory sense-making as an affordance dance: the caregiver initially uses the infant's movements as affordances for their own actions (to let the infant experience their potency and ability to move others) and, in a complementary process, nudges the infant into "properly using" the caregiver's actions as affordances, which simultaneously enable and narrow down the spectrum of possible (culturally acceptable) behaviors (Rączaszek-Leonardi et al., 2013).

The relationality of the term "affordance" (Heft, 2001; Chemero, 2003) entails that perceiving and acting happen in relation to a child's bodily abilities and to intentional engagements or projects that are currently realized by individuals and by the dyad, thus in terms of the contributions one can make to the interaction. The perception of others in such encounters, which constitute the greater part of the infant's world, is thus rarely a passive, 3rd person perception but rather an engaged and participatory being-with that involves one's own affective experience, crucially including time experience (Stern, 2000). Infants thus become attuned to complex relational and dynamic structures, including the emergent level of interaction itself, such as cocreated rhythms, rather than learn to respond to some simple features of the world. Such an attunement results in an increasingly skillful, affectively regulated form of participation, which is an important aspect of world knowledge.

In this paper, we draw attention to the fact that these engagements, in addition to being prescribed sequences of movements, have a very precise structure of timings: their durations, the intensity dynamics (changes-in-time) of the contributions, and the pauses or overlaps involved are quite specific and actively maintained. These timings are present in infants' experience and are directly grasped and appreciated in their complexity (Stern, 2000). Actions and their couplings have specific "frequencies", which are biologically constrained by our bodily functioning, e.g., by rhythms such as breathing or the natural force dynamics of our movements, and which are culturally shaped from our earliest moments (even in the uterus, infants feel the rhythms of their mothers' movements or speech) by choosing certain organizations from a metastable potential landscape over others, stabilizing them by repetition and shared affectivity. The wavelengths and intensities of the enacted events are thus of no less importance for specifying actions than are the wavelengths of light reflected from surfaces for structuring our engagement with the world.

To illustrate this fact, consider the following example, which we have analyzed in detail in Rączaszek-Leonardi et al. (2013): a mother is dressing her baby and putting on his knitted shoes. Holding the shoes, she freezes her activity and waits. The infant kicks his legs vigorously; the mother takes this movement as a decision, grasping the leg which is moved first and commenting as she puts on the shoe: "*this one? this one first?*". She then ceases her actions once again, holding the second shoe within the infant's field of vision and now intently shifting her gaze between the face of the infant and his other leg. When no action is taken by the infant, she taps his other leg lightly, saying *"I need this foot here"* before grasping the foot to put on the other shoe. We argued that by repeating this exact sequence of actions and stopping at certain moments, the mother is creating a "slot" for the infant's action, i.e., a time-to-do something in particular.

Below, we present a microanalysis of another example: the case of specifying, through coaction, that it is a time-to-look.

Example 1. A 6-month-old girl can sit up on her own. In such a position, she is naturally inclined to look down, making it more difficult to establish eye contact (Fig. 1a). The infant expresses interest in the mother's hands (Fig. 1b). To redirect her infant's attention, the mother opens the sequence with a greeting, *"hey you"*, while synchronously caressing the infant's waist (Fig. 1c), using a modality that seems to be close to the direction of the infant's gaze. The mother is not successful at drawing

---

the infant's attention. She then tries again, saying *"look here little mouse"* and adding a tactile sensation to the interaction by simultaneously grasping the infant's arms, which are already extended (Fig. 1d).
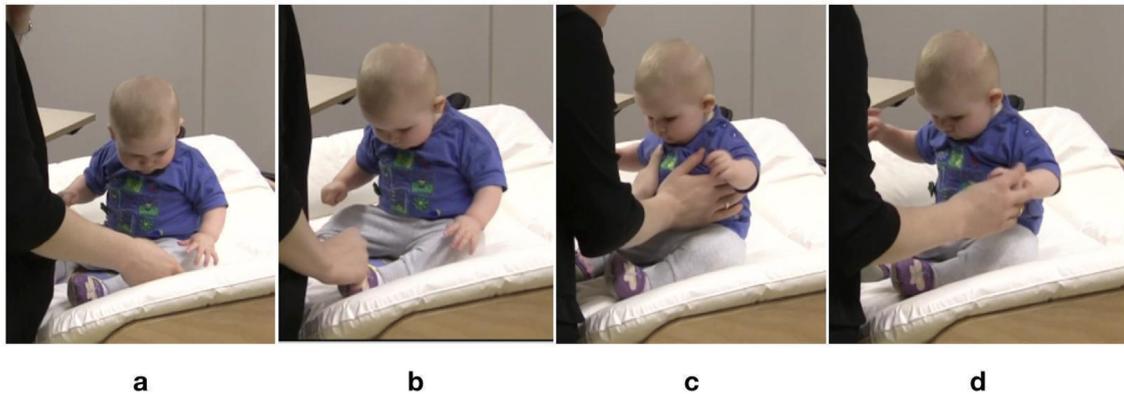

Fig. 1. Images a–d correspond to Example 1. The mother engages the infant to look up.

The infant does not look up. Next, the mother puts her hands together and positions them in the infant's field of vision while expressing the directive *"look up"* (Fig. 2a). She moves her fingers upward, and the infant follows the movement of the mother's fingers and looks up. The mother approaches the infant with an extended finger and taps the infant's nose (Fig. 2b and c). Having drawn her infant's attention, the mother quickly lowers her body and positions her face at the level of the infant's gaze while simultaneously saying *"here is mommy"* (Fig. 2d).
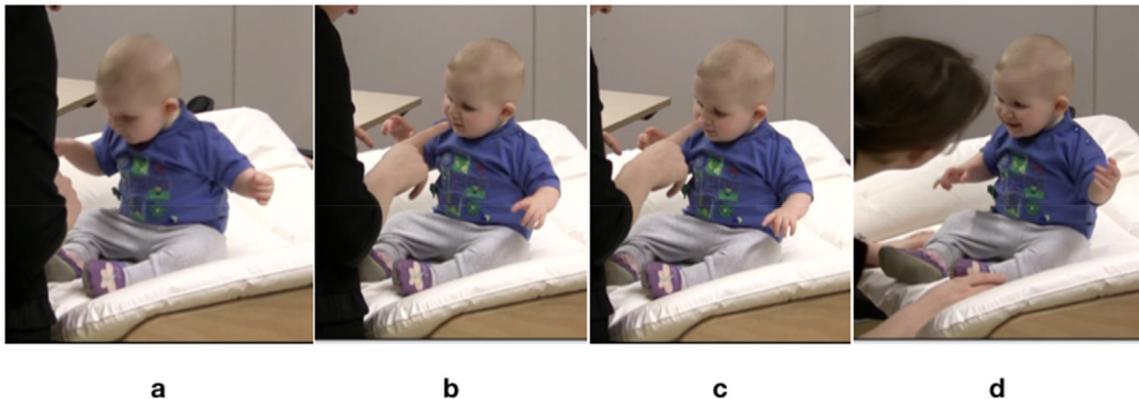

Fig. 2. Images a–d correspond to Example 1. The mother draws the infant attention to her hand then enters the infant's gaze space to enable eye contact.

It was time-to-look-up for the child and to engage in eye contact after the mother's call. After opening the episode with a greeting, the mother waited for the child to look up at her. When this action did not occur, the mother then pursued the desired interaction by touching the infant's waist and hands, repeatedly asking the infant to look up, making playful finger movements in front of the infant's face, tapping her nose and ultimately lowering herself into the line of the infant's gaze. After a great deal of nudging, the infant did look up, and that gaze was embraced as a contribution; it was only then that the interaction proceeded.

Infants are invited to take part in this affordance dance as constitutive participants with specific, age-appropriate roles. However, unlike in the case of a dance, it is much harder to decline participation: the reason we call this form of social organization "physics" is the pervasiveness and unavoidability of the participation structures that are imposed by it. Failing to adjust and participate in an enacted social event, failing to deploy a specific contribution with a specific intensity at a specific time, has certain consequences. Disruption of the emotional and action-based flow and various repair attempts occur to nudge an infant to exhibit "proper" behavior. The expected sequence and timing, especially when achieved in such an arduous way, is rewarded with a—once again, carefully timed—smile and vocalization. These sequences and timings are quickly learned and relied upon: violations of an infant's expectation of social responsivity lead to visible distress (e.g., Tronick et al.,

1978; Murray and Trevarthen, 1985), and at this phase of development, this distress is perhaps even greater than that caused by nonfunctional encounters with the physical world.

In developmental progression, the timings of such participations become more defined and better coordinated. As we showed using dynamic time-series analyses (cross recurrence analysis, Webber and Zbilut, 2005) of the gazes of infants and mothers (Nomikou et al., 2016), steady gazes toward each other at 3 months of age develop by the age of 6 months into a different structure, reflecting a more precise informational coupling. We observed a sharper coordination of mutual gaze, with shorter looks at each other that were nevertheless more tightly coordinated in time. This development reflects how one's own attention becomes constrained by the attention of one's partner in interaction and how that attention becomes mutual in the context of other activities. Using frequent pattern mining, we have also shown examples of such coalescing sequences of contributions across modalities, and cross-recurrence analysis allowed us to investigate their timing structure (Rohlfing et al., 2019). For example, the sequence featuring a smile from the mother followed by a smile from the infant stands out as a frequent occurrence; gazing at one's partner increases the frequency of the partner smiling within a certain time-frame (for 3-month-old infants, the increased probability of smiling peaks within 2 s).

Such a multimodal, turn-taking structure is reflective of the mutuality of the contributions and agency of both participants, which also allows for emotional reciprocity over subsequent turns (Stevanovic and Peräkylä, 2015). The timing allows the partners to maintain emotional flow, providing a sense of togetherness and of coconstructing a joint event (Jaffe et al., 2001; Kaiser and Butler, 2021). Considering the coupling of behaviors not as a chain of associations of punctate behaviors but as coordinations within multiple rhythms reveals actions to be both responses and proposals, i.e., both reactions to earlier events and "proactions", anticipations and evocations of subsequent actions, both one's own and—since the environment is alive and engaged—those of others.

The picture of stimuli-driven, start-and-stop cognition gives way to the image of cyclical engagements with partners, where the kinds and timings of behaviors depend not only on adjacent "causes" but also on the role of a particular move within multiple cycles. The presence of certain actions in such cycles and their timings are not only manifestations of an orientation toward specific goals or practical needs but also reflect and sustain certain ongoing relations among participants, which are based on values that are important for us as biological and cultural agents (Hodges and Baron, 1992; Hodges and Rączaszek-Leonardi, 2022; Rączaszek-Leonardi and Nomikou, 2015). The "when", i.e., the timings and the durations of behaviors, are thus as important as their form in specifying their meaning. A seemingly "identical" action may have different meanings and thus afford different contributions depending on its timing; for example, a gaze can be seen as either reproachful or appreciative attention, an elongation of a vowel in a word may cause that word to take on the opposite meaning, and a short hesitation may change the meaning of an exchange (De Jaegher and Di Paolo, 2007). We learn to be and act with others by attuning ourselves to such timings, our body becoming skillful and our feelings sensitive to the timings and forms of this intricate dance, in which events not only follow each other but also form structured and nested hierarchies.

We can now explicate the role of timing in shaping sensitivity to specific strands, modalities or layers of interaction, which results in an individuation of the processes that are crucial for language development, thereby stabilizing and strengthening some relations while possibly weakening others. We can examine the placement of utterances within coaction and, at the same time, trace the segregation of the vocal modality as a relatively and increasingly independent layer of coaction with its own intrinsic timing, within which the systemic properties of language may be highlighted for an infant.

3. Time-to-speak: vocalizations in coactions and in utterance–utterance relations

Embracing the ecological approach, which views the process of language learning as one of becoming skillful at managing social affordances rather than mapping utterances to referents, poses a difficult problem: how might a language that can be described as a "system" arise from such grounded coaction and communication while still remaining an efficient resource for interaction (Kempson et al., 2016) within a given context? Thus, rather than in grounding an already assumed symbolic system, we are interested in the peculiar character of particular types of behaviors within this multimodal and multitimescale dance –vocalizations—which seem to gain regularity with respect to other behaviors.

In the typical development of the infant-parent dyad, the caregiver's language and the infant's vocalizations are omnipresent in interactions from the earliest moments. Language takes place in participatory action frames, accompanying other moves and occupying a specific place and time with respect to them. It is present within such an intricate action dialog long before what we usually regard as "language use", dialog, or "linguistic interaction" occurs. Thus, like the nonverbal behaviors discussed in the previous examples, vocalizations change the course of interaction by starting a cascade of actions on the part of parents, who embrace such vocalizations as meaningful contributions. Thereby, parents foster the infant's agency and reinforce the culturally appropriate ways in which interactions are to proceed.

The vocal modality is particularly well equipped to support time structuring in coaction. This modality is relatively independent from other movements that occur in interactions, although it always remains coupled to them, e.g., through breathing rhythms or gesturing. Utterances provide additional variability for the stream of coaction; they can be shaped in accordance with multiple demands and values in terms of their timings and qualities such as intensity, pitch, and form. Variability in these dimensions adds an important expressive layer to emotions. Vocalizations are located at specific moments, making these moments salient and directing the attention of the infant in particular ways (Nomikou and Rohlfing, 2011). They punctuate, extend and solidify the rhythms of coaction, thus serving to control the courses of action. In addition, they themselves are controlled: in the social world, it is proper to vocalize at certain moments of

interaction and to respond to certain vocalizations in specific ways and at specific times (both in terms of specific phonological forms and pitch contours). Due to their expressivity and flexibility, utterances can mimic the actions that they accompany and can nudge or anticipate other actions. In semiotic terms, utterances become icons and indices of the interactants' moves, i.e., signs that are exploited for better coordination, even though their symbolic content may not be grasped by the infant.

However, when parents use language to emphasize, accompany and anticipate actions, they rarely use words in isolation. The caregiver's speech has its own, largely independent rhythm, which itself has been shaped by innumerable encounters occurring within the culture, even if that rhythm can be adapted slightly to the infant's sensitivities (motherese). Parents embrace infants' contributions to intricate dialogs, accommodating both the rhythms of coaction with an infant and those of everyday adult speech within a single interaction. Incidentally, this accommodation is not a difficult feat because adult language, rather than being an outcome of a "generator" residing in an individual mind, can be viewed as structured for participation (Schegloff et al., 1996; Gregoromichelaki and Kempson, 2013; Jennings and Thompson, 2012). In adults, language can also be "action-dictated", as in the case of coordinated physical work; however, utterances have a strong internal structuring, which allows for partial situational independence.

Thus, the double grounding of language advocated by Rączaszek-Leonardi et al. (2018) becomes manifest in two time orders: the timing of the course of action, of which language is one part, and the timing of the rhythms of speech. To understand how language emerges in interaction and how it becomes partially ungrounded throughout the developmental time scale, we must understand the relation between these two time orders, thereby grasping these two *processes* and the potential of both to contribute to meaning. In early interactions, the two time scales seem to be relatively disjoint from one another: language accompanies coaction—as icons and indices—on the time scale of daily routines, sometimes in the form of one-word utterances, but it is simultaneously present as a structured system of utterances in parents' speech. How these two kinds of structures (the structure of coaction and the structure of language) align seems to be understudied, but this alignment is arguably a decisive step in becoming skillful at using language as a symbol system. According to the double grounding model (Rączaszek-Leonardi et al., 2018), the relations detected within one time order impinge on the relations present in the other.

The way in which these two time orders become interwoven with one another can be seen in Example 2 below, where the "normal" rhythms of speech must be aligned with respect to onomatopoeic vocal markings of actions.

Example 2. The following situation is the closing sequence of an interaction. The mother has finished changing a diaper and is lifting the infant from the changing table. She stretches her arms in preparation to lift the infant, saying *"So then, come here"*. She extends her arms to pick him up, taking a loud inbreath as she positions her hands around his body, and produces a prolonged "oh" (expressed as a breathy grunt), which is elongated to mark the entire duration of movement of the infant body from the table to her arms. Once the infant is in her arms, she says *"big boy"*.

This example shows how the timing of the vocal modality is modified to match the timing of the coaction, providing a unified experience for the infant (Nomikou, under review). Additionally, the infant is a participant in this action, as the effect of his body weight regulates the quality of the mother's vocalization (the grunt-like *"oh"*), and this weight is also a topic of the mother's comment that he is a *"big boy"* and therefore heavy to lift from the changing table.

Through the timed placement within the action, the parent's language thus gradually stabilizes *both* i) its special role in interaction as an additional constraint on coaction, given that it is coordinating it at important points, which ensures good flows of movement and affect, and ii) the "system" of relations among utterances, which, although it is always immersed in such interaction, gradually coalesces as a particular integrated modality or layer for the infant (Rączaszek-Leonardi et al., 2018). The time-to-speak is regulated by these two time orders: the course of coaction and the organization of the vocal layer. A key step in this process is that infants learn that the vocal layer of their parents' interactions is also something in which they can participate.

Infants seem to be able to create this vocal layer of interaction jointly from the beginning, placing their own vocalizations at the 'right' time and place and with the 'right' pitch relative to the vocalizations of the adult (Malloch and Trevarthen, 2009; Gratier et al., 2015). The infant's vocal behavior further becomes attuned to the interaction when the caregiver, sensitive to the infant's vocalizations, incorporates them into meaningful events, this time within the vocal modality. Time-to-speak within this stream of vocal behavior gradually becomes more regular and develops a clear turn-taking structure and a pace of its own, which becomes increasingly faster than the pace of coaction. To see how a dialogical pattern emerges in the vocal layer, consider the following example interaction:

Example 3. A mother is changing an infant's diaper, talking to her and inviting the infant to contribute by saying *"tell me"*. In line 1, before the mother has concluded her utterance, the infant vocalizes. The mother immediately interrupts her utterance, giving way to the infant's vocalization. She invites the infant to vocalize again (line 3). Next, the mother continues to speak, and the infant produces multiple vocalizations. What is interesting in this context is the fact that the mother treats the infant vocalizations as important contributions to the interaction. We see this importance in line 3, where the mother starts her turn with a *"yes"*, as if she were agreeing with what the infant just said. She accompanies her reply with a notable nod of the head. Next, in line 9, she starts her turn with *"what?"*, leaving a pause after her utterance and treating the infant's vocalization as noteworthy or surprising. Finally, in line 13, the infant produces a two-syllable vocalization that sounds like *"ha-he"*. We see that the mother immediately picks up this vocalization and treats this expression as referring to the infant's hair (German:

Haare) and thus as a proper contribution to the discussion. She reciprocates the infant's vocalization in line 14 and marks this reciprocation with a simultaneous touch of the infant's head (hair).

```
01   M:    Erzä-
           Te-
02   I:    ((vocalization))
03   M:    Ja erzähl nochmal
           Yes tell again
04   I:    ((vocalization))
05   M:    Waren wir heute bei Georg?
           Did we go today to George ?
06   I:    ((vocalization))((vocalization))
07   M:    Der ist viel kleiner als du?
           He is much smaller than you?
08   I:    ((vocalization))((vocalization))
09   M:    Wie jetzt? (0.3s) Der hat schwarze Haare?
           What? (0.3 s) he has black hair?
10   (2.9 s)
11   M:    Jetzt musst du überlegen, ne?
           Now you have to think, right?
12   I:    ((vocalization))
13   (1.3 s)
14   M:    Haare (0.2 s) Haare hier oben
           Hair (0.2s) Hair here at the top
15   I:    ((vocalization))
16   M:    Ach (0.2s) ja genau du hast blonde
           Ah (0.2s) yes exactly you have blond
17   (1.5)
18   M:    ((laughs))
```

A further aspect that becomes clear from this short sequence is the way in which the timing patterns that are conducive to the emergence of dialog can be discerned when considering the parent's contributions alone. Unlike monologues or expressions of formulated propositions, these contributions seem to be half-dialogs that are full of carefully placed pauses—either to accommodate the infant's movements, gazes, facial gestures or vocalizations or to leave room (by pausing) for a response from the infant. This is evident in the design of the mother's turns, which are formulated both syntactically and prosodically as questions (lines 5, 7, 9 and 11), thus creating a slot for a response. The vocalization provided by the infant is embedded in the storytelling of the mother and contributes to the progression of the narrative. Interestingly, when a response from the infant is delayed, as in line 10, the mother focuses on this disruption of the dialogic structure by attributing hesitation to the infant (line 11) as a result of the difficulty of the question asked in the previous turn.

A history of interactions structured in this way establishes an anticipatory cycle of mutual participation, which can be seen cross-culturally (Levinson, 2006; Stivers et al., 2009) and characterized by a cyclic "readiness to speak" (Wilson and Wilson, 2005), albeit with some degree of cultural specificity in terms of rules for overlapping speech or the length of pauses. Within this milieu, controlling one's own vocalizations is increasingly coupled to one's interaction partners and not merely to one's own needs, movements and rhythms. Importantly, a history of interactions improves not only the timings of the turn-taking rhythms involved in human vocal expression but also the affective support underlying this process. This increased tuning becomes manifest, for example, in the feeling of uneasiness that arises when pauses extend beyond a specific length, irritation when somebody is interrupted, or the positive affect that develops via the smooth flow of a proto-conversation or, later, a conversation.

Learning to engage with language means becoming skillful at using and providing affordances to constrain or control engagements. Utterances are not used with a clear vision, "plan" or "model" of the listener's reaction (such a picture is the result of considering cognition in terms of start–stop processes that are driven by "stimuli", or, more recently, in terms of more dynamic, predictive-processing models). Rather, language creates affordances for the realization of values (which is sometimes, but not always, goal-directed) and for engagements to occur and be sustained (Gregoromichelaki and Kempson, 2013; Van Dijk and Withagen, 2016); it provides enabling constraints, which can be used creatively by the other in the context of the current situated interaction. This process is based on an anticipation, sometimes vague and unconscious, of a general direction of the interaction and not, or at least not always, on the ability to anticipate upcoming concrete percepts (Westermann & Jones, 2021).

Studies have shown the degree to which vocalizations are time-coordinated (Warlaumont et al., 2014) as well as how the mother's utterances are dependent on the quality or language-likeness of the infant's vocalization (Warlaumont et al., 2014, 2016; Nomikou et al., in preparation). The more language-like the vocalization is, the greater the probability that the mother reacts with her own utterance within the specific time window. Our own study using cross-recurrence time-series analyses shed light on the ways in which time relationships in mother-infant vocal interaction unfold over the course of the infant's development: a pattern that is visible at 3 months of age, such that mothers closely (within 2 s) and reliably respond to the infant's vocalization, gives way to a more symmetric and clear turn-taking structure at 6 or 8 months of age (Leonardi et al., 2016). These timings facilitate individuation of the vocal modality as a discernible layer of interaction. The infant learns that utterances follow each other, both within one turn and across turns, resulting in a joint activity of speaking or conversing.

Within such coherent patterns, novel relations among utterances can be perceived. Some preliminary results have seemed to show a link between the clarity of such patterns and later language development (Nomikou et al., in preparation).

3.1. Relating the relations

It is within such dynamically structured settings of carefully timed and affective interactions that more elaborate forms of language appear as complex controls. The vocal modality emerges as distinct (but never separate) from the ongoing rhythms of coaction. This modality solidifies both because the infant becomes increasingly agentive and skillful with respect to vocal exchanges and because the parent "injects" adult speech into the interaction. The vocal modality, however, remains in a controlling relation with respect to ongoing actions, and now complex structures of speech may also influence complex coactions. Once again, timing is crucial in this context, as language can be aligned with the "grammar of action", making the relata and the relationship salient; however, this alignment does not take the form of a correspondence or mapping but rather of language guiding or steering the interaction (Bickhard, 2015).

Nomikou and Rohlfing (2011) provided an example of a way in which the verbal layer becomes aligned with multistep actions. The mother is putting a diaper on the infant, which involves a three-step sequence of actions: wrapping the diaper around the infant, fastening one side, and then fastening the other side. These authors showed how the mother vocally demarcates each intermediate step in her action sequence with pauses and by saying *"now we pack everything in"* (step one), *"first, one side"* (step two), and *"then the other side, huh?"* (step three). She also extends the length of certain words, thus highlighting the duration of each movement and allowing for synchrony at the conclusion of each action step.

Similarly, in Example 4 below, the mother lifts the infant from the changing table after completing a diaper change. She performs this action over the course of two steps. One step is to lift the infant to a sitting position (Fig. 3a and b), and then she takes the infant by the arms and lifts him to a standing position (Fig. 3c and d) and subsequently pulls him into her arms. She accompanies this action by saying *"oh! stand up"*.

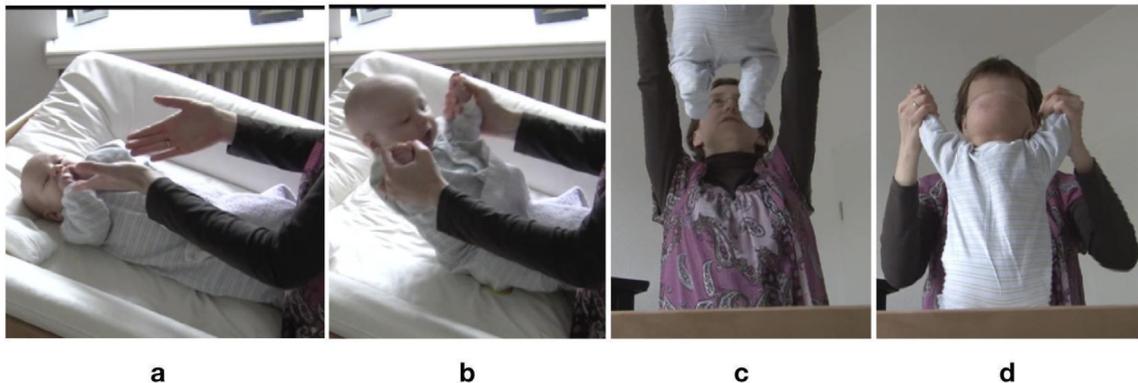**a**　　　　**b**　　　　**c**　　　　**d**

Fig. 3. Images a–d correspond to Example 4. They illustrate the mother lifting the infant from the changing table.

The above description of the utterances used in diverse streams of interactions indicates that at least some of the systematicity of language is due to the structuring of coactions, which the utterances accompany and, over time, control. Other pressures structuring language may originate on different time scales and may include learnability or being selected for in a given culture as effective controls (and therefore enacted by the encultured parent). Thus, the task of answering the question "What is the language system and where is it to be found?" is by no means straightforward. Examining this question from an ecological perspective, however, helps identify various pressures and is crucial for an interaction-oriented and integrative theory of language emergence and functioning (see the discussion in Section 5).

In the next section, we aim to highlight the way in which structures in time, as realized both by coaction and by language, extend beyond local dependencies.

4. Time-to-build-up, time-to-resolve: larger structures for participation

In the previous sections, we showed how interactive structures cohere locally, i.e., through the enactment of specifically timed sequences of moves and vocalizations. However, the social world is engaging on various timescales and instills meaningful patterns beyond the level of local adjacencies (Donald, 2007; Enfield, 2013; Rossmanith et al., 2014). Close examination of our activities in the world reveals, as if by a kind of Fourier analysis, a variety of "projects", waves of activities of various periods and phase relations that are interwoven within each interaction. Qualitative analyses may help discern at least some of the experientially important threads, but the overall quantitative nature of the dynamic relationships among them can also be determined (Van Orden et al., 2003). Dynamic analyses show general signs of such a complex time structure for interaction (Abney et al., 2014; Kelty-Stephen, 2017; Wijnants et al., 2012; Zapata-Fonseca et al., 2016). Rather than a set of

stimuli, responses, and algorithms for mapping one onto the others, the description of such a dance seems to require a musical score (Papoušek and Papoušek, 1981; Malloch and Trevarthen, 2009; Moreno-Núñez and Alessandroni, 2021).

Here, we would like to note that interactive events, even though they are situated and task dependent, seem to have an intrinsic timing (Fowler, 1980), which often spans several actions or action steps. Extending over immediate contingencies, such events mark "larger" structures, which can be nested within each other, weaving more complex interactive patterns. Their intensity contours and the internal order of actions stem from and provide orienting forces within routines. Since we are involved in enacting them through both our own bodily rhythms and our actions, i.e., since they are "movement-first" in a sense (Rączaszek-Leonardi et al., 2013), they ground our experience of the world in an experience of movement or in "our original capacity to think in movement" (Sheets-Johnstone, 1999, p. 434), which obviously has a temporal dimension. Each such interactive event, with its intricate time structure is characterized by a particular order of contributions and its own intensity or vitality contour (Stern, 1999).

Called "primary narrative structures" (Delafield-Butt and Trevarthen, 2015) or "action arcs" (Rossmanith et al., 2014, p. 8), such patterns are characterized by "a beginning, build-up, climax, and resolution" (ibidem, p. 8), which are experienced from the first-person perspective and shared with one's partner in the interaction within a coconstructed event. Both the excitement of the climax, i.e., the peak of the arc, and the relaxation of the resolution are mutually recognized and are often marked by specific actions, such as a wide opening of the eyes or a sharp intake of breath in the case of climaxes and, e.g., a smile and reciprocal acknowledgment in the context of resolution. Such energetic contours have the power of punctuating and relatively discretizing "events" in a continuous flow, establishing a specific, albeit culturally sanctioned, "narrative frequency" for a dyad.

Action arcs are thus arcs in time that provide tension, produce differences in arousal and affect, and structure events. The arc's tension is created and released with a specific timing. As shown in Rossmanith et al. (2014), in shared book-reading routines, "(t)his shaping of action arcs is found across all kinds of actions and at different levels and multiple timescales within an activity, nested into one another" (p. 8). Action arcs may confer on recursion and embeddings a distinct time dimension, which is possibly relevant both for structuring routines and for structuring language.

Example 5. This example starts with a mother and an infant just prior to putting a new diaper on the infant. The mother announces the upcoming action while holding the diaper in her hand: *"So, now we put the new diaper on"*. At that moment, the infant lifts both his legs into the air. The mother notices this movement, and instead of lifting the infant's body to position the diaper under it, she structures the activity as one that they accomplish together. *"Are you already lifting your bottom by yourself? Are you helping me?"* What follows is a series of short utterances prompting the infant to lift up his bottom: *"Lift your bottom ((intake of breath)) bottom up ((laughs)) bottom up; yes; bottom up, come on"*. These utterances are produced in very quick succession, each building on the previous utterance and creating an escalation that is captured in the rising pitch and intensity of the mother's voice and by the lack of pauses that unites the entire sequence in a bigger structure. Interestingly, the mother pauses in the middle of the buildup, and the infant joins in and vocalizes twice at this point (lines 1 and 3) between the mother's turns. His vocalizations are noticed and reciprocated by the mother (lines 2 and 4) and embedded into the developing action arc as the mother continues:

```
01   I:    ((vocalization))
02   M:    ((imitation)) bottom up
03   I:    ((vocalization))
04   M:    ((imitation)) bottom up
```

Upon reaching its peak, the sequence is resolved with a laugh from the mother; the diaper has been positioned under the infant's body, and the interaction progresses.

Such contours support the binding of coactions into sensible wholes both with respect to the action and the vocal layer, which can be tightly coupled. Local dependencies of utterances are strengthened not only by the frequency of immediate adjacencies (collocations in infant-directed speech) or by the real experience of the controlling power of such utterances with respect to the interaction but also when they are enveloped within those arcs of larger intensity. The vocal layer, with its relative modal freedom from the physical actions of the body as well as its own intrinsic timings—syllabic, phrasal, sentential, narrative—has the potential of at least highlighting and at best instilling such intensity envelopes, thus facilitating a narrative interpretation of even the most accidental and simple actions.

As Rossmanith et al. (2014) noted, such time- and affect-structured interactions may develop into events that can be the topic of other interactions, constituting a sort of primary "aboutness". What the present work has indicated clearly is that such primary "aboutness" pertains not to a static object in the world but rather to this affective, lived-through, participatory structure, which is constituted by its internal timing. Whether such systemic hierarchies are ways of achieving relative discretization and how they can be coconstructed in action and language lead to a variety of research questions that have only been briefly mentioned in this paper.

5. Summary and discussion

In summary, in this paper, following the ecological turn (Steffensen and Fill, 2014; Rączaszek-Leonardi, 2016), we claim that language development is more fruitfully understood not as "language learning", in which the detached perception of

objects in the world and symbols of language are identifiable, reified parts that can be mapped onto each other, but rather as participation in languaging, i.e., skilled engagement with others within events at various timescales, which is scaffolded by routines and affective involvement. Focusing on a time analysis of the relevant strands of coaction reveals the processual nature of language and interaction and simultaneously provides an important dimension and additional measurables, which may allow us to improve our understanding of the structure of the enacted environment. An important task for researchers is thus that of identifying the timings and affective and behavioral intensities that weave such coherent participation in the context of language.

In this work, we have illustrated several key structures in time in early interaction, which enable language emergence as a skill of coordinating with others and can facilitate language learning. We have described three different "times-to", i.e., the timings of events oriented toward specific points, as in the case of "time-to-contact" (Lee, 1976), which are nevertheless perceived in terms of the agent's own required actions at that specific moment. These "times-to" illustrate the presence of a variety of dynamic structures in the environment of a child that can support the infant's ability to learn how to engage. The examples presented showed that timing is crucial for structuring coaction in general (time-to-smile) as well as language (time-to-speak) and that sequential ordering and timings in the latter case support individuation of the vocal modality within the infant's perception-action. This individuation, in turn, facilitates learning the systematic relationships between utterances and their constraining role. The third timing (time-to-resolve) indicates the variety of timescales that are important for structuring participatory events or "frequencies of engagement", illustrating the presence of "action arcs" that span several contributions, which provide orientation, energetic contours and possibly hierarchical organization to these larger interactive structures (Rossmanith et al., 2014). Such envelopes for sensible and sensed structures integrate them as events, larger units in our action and experience, to which we can relate within more complex meaning structures.

The purpose of this work was to demonstrate the variety of the intrinsic timings and enacted structures in the environment in which the child participates. The biological time scale has prepared us for social engagement, structuring our bodies in a particular way: our posture is straight, our faces are flat and able to change expression with multiple degrees of freedom, and our eyes have white sclera, which helps express the direction of our gaze. However, our body is also active and exhibits specific rhythms, such as breathing, heartbeats, blinking, twitching, and movements. Via this structured and rhythmic body, we enter the environment, which is a social environment, enacting rhythmic behaviors involving us, some of which are already familiar and others of which are novel. The basic rhythm of human speech is already heard in the uterus, and others breathe, speak and move with similar rhythms. Through participation in social events, novel rhythms can be created, and novel events can occur on differential waves of various frequencies.

The examples of timings on which we focused in this work are not meant to exhaust the range of rhythms that are important for human interaction. We chose these examples because they pertain to processes that seem to be central for the emergence of the systematicity and complexity of language in the context of development (Rączaszek-Leonardi et al., 2018). The ecological psychology framework is helpful for capturing the continuity, temporal nature and active character of reciprocal engagement with the dynamically changing social world. This world is the world of affordances, in which learning is understood as an attunement to social events in terms of the agent's own increasingly skilled action. This perspective allows for a better characterization of the structure of engagement, in which context we can trace the structure of language as it coalesces as a system and leads to a better understanding of the relationship between the structure of coordination and the structure of language.

The basic rhythms of interaction entrain the child's action, attention and perception. Structures are composed of various actions; however, importantly, they have a specific timing, to which the infant is attuned. The baby learns not only that she should look and smile following the mother's utterance but also when it is time to smile; the infant learns not only that he should vocalize following the mother's vocalization but also when it is good to do so. This entrainment is effective with respect to intermodal behaviors, but it simultaneously provides a way to discretize or individuate coherent strands or subsystems. The frequency of responding to utterances with utterances within a specific timeframe characterizes this modality as a verbal modality, within which patterns of the language "system" (such as word contingencies) can be discerned and made subject to learning. Relations among words, which regulate the relations among interactive events, are arguably the birth of the symbolic nature of language (Deacon, 1997; Rączaszek-Leonardi et al., 2018).

Importantly, structures enacted in time are also affective structures. Affordances are value-laden (Hodges and Baron, 1992); perceptions of others' contributions as opportunities for engagement are founded on the memory of felt bodily movements and affective states. Action arcs instantiate the changing intensity of affect. Over the course of multiple encounters, the child learns when it is a good time to begin a new event, when it is a good time to increase intensity and when it is a good time to resolve the event. These narrative structures are carefully created through social feedback and they are experienced or lived-through together with a caregiver in a shared affect (Popova, 2015; Popova and Cuffari, 2018).

The skill of using language in interaction depends on recognizing the rich and stable system of relations among utterances, which is facilitated by the parents' speech present from the earliest moments. However, this skill is never separate from the skill of using those complex structures to modify ongoing interactions. Both skills are needed, and the search for an integrated approach to language on ecological grounds would be incomplete without recognizing the presence of well-timed speech as a particular, salient, intricate and reproducible structure within our environment and without characterizing its role as providing complex forms that can guide interactions. Integrating the approaches will hopefully also entail integrating the methods. Incorporating the time dimension into this analysis may better show the (perhaps surprising) compatibility

between the dynamically oriented and structurally oriented approaches, which cater to these complementary needs. In the remainder of this section, we try to take a step in this direction.

5.1. From time-to-speak to a language "system"

As mentioned previously, in Rączaszek-Leonardi et al. (2018) we proposed a model of symbolic language development encompassing the individuation and integration processes, which we enriched in this work by detailing the crucial time dependencies involved. In that model, which drew on an earlier idea of Deacon (1997), the processes of "double grounding" (both in coaction and in other utterances within the vocal layer[3]) are necessary for language to become symbolic. In this way, interaction control can be based on relations (among utterances and among experiences and actions), thus satisfying the tenets of ecological psychology (Heft, 2001) and agreeing with processual models of cognition (Bickhard, 2009).

This model, however, was criticized for contradicting its ecological aspirations and therefore newer strands of ecolinguistics (Steffensen and Harvey, 2018). First, these authors argued, the model in question overrates the relevance of relations within the vocal layer for controlling parent-infant interaction when compared to all other bodily accompanying signs. Second, Steffensen and Harvey noted that by seeking an explanation for some common intuitions regarding language, the model legitimizes structuralist terminology and incorporates the resulting concepts into the theory once again, including notions such as 'reference' or 'arbitrariness of the form'. Foregrounding the time dimension should make the double grounding model clearer and thus less susceptible to that criticism.

To reiterate, the key moment in the double grounding model of symbolic communication development (Rączaszek-Leonardi et al., 2018) is when the constraining role in the interaction can be fulfilled by a relation between utterances and not merely by a single word or a holophrase. From this moment onward, a variety of relations within the vocal layer can become complex controls in interaction. Such relations are stabilized within an environment in which action–action relations are pervasive, as indicated by the examples discussed throughout this paper. The fact that the early interaction environment is already action-saturated seems to be crucial for languaging from the perspective of both integrationist linguistics and ecolinguistics. As Steffensen and Harvey (2018, p. 11) noted, "[inter-sign relations] are in fact ubiquitous across all interbodily coordination". Vocal layer structuring clearly needs and builds on such interbodily coordination; learning a language means knowing how to engage in interaction using the controlling power of utterances (Rączaszek-Leonardi, 2012, 2016).

Since it is the relations among utterances that carry out the work at some point, it is these relations themselves, rather than holistic forms, that should be looked for and perceived by the infant. The emergence of the vocal modality facilitates the ability to perceive these relations. The caregiver's behavior reinforces the individuation and consolidation of the vocal layer, in which relations among utterances are emphasized. As shown in the previous sections, by specific timings of their speech with respect to the vocalizations of the infant, parents insist on promoting and remaining within that layer; they reinforce vocalizations affectively and embrace them to form larger "narratives". Importantly, parents do not refrain from speaking in an adult manner, but they use this mode of speech to engage in dialog with the child and to comment on the relevant coactions. This form of engagement solidifies a more modal layer of interaction, which features its own culturally sanctioned rhythms and timings as a relatively isolated behavioral strand, which nevertheless remains in constant contact with other actions.

Let us emphasize the fact that the vocal layer is distinguished not because it is, a priori, "pre-theoretically and phenomenologically identified as being 'languagy'" (Steffensen and Harvey, 2018, p. 9). Rather, the examples and studies discussed in this work demonstrate how the vocal modality undergoes a process of individuation over the course of the infant's development, which then allows the systematic relations to be detected more readily. Attending to the time dimension allows this segregation to become particularly tangible, but it can also be seen in the preferred sequences of actions (vocalizations after vocalizations) and in instances in which "language-like" vocalizations or sequences are affectively rewarded.

The essentially relational character of language, which can be perceived within the individuated vocal layer, is what justifies viewing language as a system, which may initially be a concern to dynamically oriented researchers because of the strong association between this notion and qualities such as separateness, static structure, completeness or closeness. However, if the notion of a "system" is to be used with respect to a vocal layer, this term must be employed in Von Bertalanffy's (1968) sense of an open, flexible organization of parts that is individuated by their relative coherence, including characteristic timings, rather than in the sense of a petrified and isolated structure, which is (often unjustly; see Stawarska, 2015) attributed to a structuralist vision of language. The notions traditionally connected to structuralist views, such as, for example, a definite reference, do not make much sense with respect to such an open system of controls. The outcome of a controlling process by a linguistic structure is always underdetermined and crucially relies on the environment and the pervasively present social projects in which the speakers are engaged. The apparent arbitrariness of the linguistic form, in turn, depends strongly on its historicity and conventionality, i.e., on the immense work and energy required for shaping a reliable constraint within a dynamical system. Finally, language systematicity, which is often taken for granted in standard linguistics under the term of

---

[3] We would like to point out here that the vocal modality is obviously not the only modality within which such discretization/individuation processes may occur. It would be fascinating to trace how gestures become signs in sign language in a similar way, coordinated with an ongoing coaction but simultaneously establishing regularities within the sign system. Our model and the attention that we give to timings can perhaps enrich predictions concerning how this process impinges on a typical gesturing system (for background, see, e.g., Goldin-Meadow and Brentari, 2017).

compositionality, cannot be safely assumed to be a straightforward feature; rather, it requires the identification of multiple sources of structuring (Korbak et al. 2021; Mordatch and Abbeel, 2017; Zubek et al., in preparation). Thus, the meaning of an utterance is bound to depend on its components in a nontrivial way (Korbak et al., 2020; Recanati, 2012; Steinert-Threlkeld, 2020).

Whatever theoretical perspective on language one espouses, it seems counterintuitive to deny a certain degree of systematicity in the vocal modality, even if one is skeptical about the traditional conceptualizations of this systematicity as rule-based and of systematic meaning as *compositional* as Steffensen and Harvey are when writing that "language is not compositional by nature, but can be described as compositional if one is literate and adopts a particular theoretical perspective" (Steffensen and Harvey, 2018, p. 11). Human capability to create novel utterances with (broadly and flexibly) predictable effects on interaction, without the necessity to invent new words, shows that language offers incredibly powerful and precise controls. Yet is it crucial that the systematicity[4] of language pertains to the *constraints* on multitimescale interactive dynamics and not to specific *contents* (for similar accounts see, e.g., Bickhard, 2009, 2015). Notably, relatively stable patterning of the vocal modality, which aids replicability and learnability of language, does not preclude the situational flexibility of meaning and thus effectiveness in any way (Barwise and Perry, 1983). The task of uncovering the factors that lead an infant to develop sensitivity to the inner structure of utterances is thus a crucial element in any theory of language development that aims to integrate the work of dynamicists with more formally oriented linguistic models.

We do not intend to suggest that a language can be understood as a single population-level system, since utterances and their meanings are always emergent, constructed anew within the current interaction (Di Paolo et al., 2018). As emphasized in this work, in interaction, utterances are immersed in two time orders: the order of action and that of the vocal layer. However, the latter, as it coalesces between parent and infant, is not idiosyncratic to a particular dyad; it is saturated with utterance patterns that are widespread in the population, patterns that are embodied in the parent who imposes them as constraints on the interaction and who promotes the proper use of those patterns by the infant. Through these constraints, the infant's behavior and speech are shaped by her parent's more culturally stabilized language. Importantly, complex linguistic expressions emerge as effective complex controls, and while the combinatorial power of language may have been idealized by some linguists, it is not purely phenomenological either.

Structures in the vocal modality, which the child can appreciate and use, become increasingly complex, but they always remain closely related to the flow of the interaction. A change in the communicative situation, such as the appearance of another person or a sudden turn in the interaction, will change the need for interaction controls. By definition, speech cannot be reduced to a timeless structure and abstracted from its context. Potentially, however, the form of such utterances may be replicated (e.g., recorded or reprinted in the case of written language), so, in a sense, it can be "uprooted". The regularities of such forms can be analyzed and—by some—forced into the notion of a separate system of forms that is more similar to the notion of a "formal system", which, as mentioned above, is very different from the way in which the notion of a "system" is understood in the context of dynamic approaches.

From such a slice-in-time perspective, it can often be forgotten that utterances and the relations among them function as controls in dynamic interactions. However, this fact does not mean that such a slice in time is not informative with respect to the "regularities of control" that characterize language use, or for that matter that such a slice is not worth studying and modeling, as shown, for example, by recent high-dimensional embedding models, which claim to be models of semantics. Their success, e.g., in automatic translation, testifies to the amount of information that can be recovered from the relations among some utterances concerning the relations among other utterances; however, such models do not provide much information regarding how such structures arise and function as controls in interaction.

The structured layer of vocal behavior, which emerges over the course of development under the influence of adults' speech, appears to be as systematic as evoking such formal-descriptive attempts or attempts at generative explanations on only one timescale (generative mechanisms "in the mind"). This possibility is attractive for the information-processing paradigm and is perhaps supported by folk intuitions concerning language as a system of vehicles for bounded meanings. However, focusing on the structured layer itself and disregarding the controlling role of language indeed leads to the necessity of introducing a number of abstract notions, such as reference or arbitrariness. There is no clear correspondence between those terms, as traditionally understood, and the terms native to the ecological approach that we endorse.

From the ecological perspective adopted by Rączaszek-Leonardi et al. (2018), formal relations among utterances or such traditionally defined concepts (e.g., reference, arbitrariness) cannot be reasonably treated as facts about language that require explanation, nor can they be used as elements of explanations of other phenomena. Nevertheless, these notions are employed by researchers of some orientations with respect to language qualities as they are grasped by them, and they have been found useful in multiple discourses on language. Rather than claiming that these researchers are plainly wrong, we ask: how is it possible for language to be viewed in this manner? How is it possible that the dynamic strands of coaction within a variety of time-orders can allow for the emergence of stabilities that are so potent as to be seen as a "system" and that can serve as the foundation for structuralist ideas, formal descriptions, generative machines, high-dimensional vector semantics and large-

---

[4] While it is tempting to draw parallels between this kind of systematicity in the realm of constraints on dynamics and the traditional notion of "compositionality", we carefully refrain from doing so here to avoid suggesting false analogies. The issue seems very complex and requires at least a paper on its own.

scale language models? Asking such questions could be very fruitful for the purpose of reconciling the ecological perspective with that of more traditional linguistic approaches, which focus on describing the structure of such a system.

Given the arguments made in this paper, one thing is certain: the processual, dynamic and open nature of the vocal (or, in the case of sign languages, the sign) layer and the copresence of pressures from different timescales manifested by its structure ensure that language constantly eludes, and will elude, both structuralist and neo-structuralist explanations. On the other hand, acknowledging the presence and replicability of complex forms, which are identifiable by approximate structural descriptions, is crucial to properly understanding the rich, structured environment of the child and the complex coordinating role that such forms might play.

The problem of how symbols emerge from dynamic interactions was the core of the "ungrounding problem" that we discussed in Rączaszek-Leonardi et al. (2018) and contrasted with the grounding problem, which assumes that language is a system of arbitrary symbols in need of interpretation (Searle, 1980; Harnad, 1990). In the present paper, we aimed to show that studying the timing and participation in the cocreations of nested structures in terms of the skills required to engage in and control such timing may help to shed additional light on this process and allow us to reformulate some of the questions that have been posed as problems of individual cognition in terms of problems of interactive coupling and coordination. We based this process on participation in one strand of behavior, supplemented by rich relational constraints, which may loosen (but never entirely lose) its connection with other strands. On the level of language development, the questions highlighted above take a particular form: how can something that is describable as a "symbolic system", with its own norms of "being proper" or "correct", emerge in development? This question is especially salient given the causal connections between speech and the flow of interaction, which are much tighter in the context of infant interaction than we observe among adults.

Importantly, we certainly do not suggest that the growing systematicity of utterances causes the language to fully liberate itself from the dynamics of ongoing interaction. Ever-present immersion in the two orders, the "double grounding", ensures that "[…], all linguistic action, of any kind, necessarily impacts what is immediately present and coupled to the speaker" (Steffensen and Harvey, 2018, p. 12). However, we argue that the rich structuring of the vocal layer, which is brought to the interaction by the parent and further elaborated and specified during the interactions, is as vital as its underlying dynamics. This layer of behavior exhibits its own rhythms, even though it remains partly dependent on the rhythms of other coactions in which children must also learn to engage.

In this context, the independence of the vocal layer and coaction is not understood in a binary fashion; rather, it is a matter of degree. However, studying these systematicities in the vocal layer may require methods that are different from the microanalysis of language-infused interactions, which are indeed more informative regarding concrete controlling processes. These methods must be able to capture various ranges of dependencies that are present in the vocal layer and that become stabilized to the point that they can be rendered in text. Such a focus does not entail abandoning the notion of language as "primarily an activity" but calls for a broad integration of methods. Certainly, much work remains to be done to study how the structure of speech arises from and controls coaction on multiple levels. Attention to timing, which we tried to highlight in this paper, introduces an additional dimension and calls for novel methods and measures, which can render these processes clearer.

Acknowledgments

References

Abney, D.H., Paxton, A., Dale, R., Kello, C.T., 2014. Complexity matching in dyadic conversation. J. Exp. Psychol. Gen. 143 (6), 2304–2315. https://doi.org/10.1037/xge0000021.

Barwise, J., Perry, J.R., 1983. Situations and Attitudes. Bradford Books/The MIT Press, Cambridge, MA.

Bickhard, M.H., 2009. The interactivist model. Synthese 166 (3), 547–591. https://doi.org/10.1007/s11229-008-9375-x.

Bickhard, M.H., 2015. The social-interactive ontology of language. Ecol. Psychol. 27 (3), 265–277. https://doi.org/10.1080/10407413.2015.1068656.

Chemero, A., 2003. An outline of a theory of affordances. Ecol. Psychol. 15, 181–195. https://doi.org/10.4324/9780203726655-5.

Cox, R.F.A., van Dijk, M., 2013. Microdevelopment in parent-child conversations: from global changes to flexibility. Ecol. Psychol. 25 (3), 304–315.

Colombetti, G., 2018. Enacting affectivity. In: De Bruin, L., Newen, A., Gallagher, S. (Eds.), The Oxford Handbook of 4E Cognition. Oxford University Press, pp. 571–588.

Costall, A., 1995. Socializing affordances. Theor. Psychol. 5 (4), 467–481. https://doi.org/10.1177/0959354395054001.

De Barbaro, K., Johnson, C.M., Deák, G.O., 2013. Twelve-month 'social revolution' emerges from mother–infant sensorimotor coordination: a longitudinal investigation. Hum. Dev. 56, 223–248.

De Jaegher, H., Di Paolo, E., 2007. Participatory sense-making: an enactive approach to social cognition. Phenomenol. Cognitive Sci. 6, 485–507.

Deacon, T.W., 1997. The Symbolic Species: The Co-evolution of Language and the Brain. W.W. Norton, New York, NY.

Delafield-Butt, J.T., Trevarthen, C., 2015. The ontogenesis of narrative: from moving to meaning. Front. Psychol. 6. https://doi.org/10.3389/fpsyg.2015.01157.

Di Paolo, E., 2016. Participatory object perception. J. Conscious. Stud. 23, 228–258.

Di Paolo, E.A., Cuffari, E.C., De Jaegher, H., 2018. Linguistic Bodies: The Continuity between Life and Language. MIT Press, Cambridge MA.

Di Paolo, E.A., Rohde, M., Iizuka, H., 2008. Sensitivity to social contingency or stability of interaction? Modelling the dynamics of perceptual crossing. New Ideas Psychol. 26, 278–294.

Donald, M., 2007. The slow process: a hypothetical cognitive adaptation for distributed cognitive networks. J. Physiol. Paris 111, 214–222.

Enfield, N.J., 2013. Causal frames for understanding language. In: Enfield, N.J., Kockelman, P., Sidnell, J. (Eds.), The Cambridge Handbook of Linguistic Anthropology. Cambridge University Press, Cambridge.

Fowler, C.A., 1980. Coarticulation and theories of intrinsic timing. J. Phonetics 8, 113–133.

Gibson, J.J., 1960. The concept of the stimulus in psychology. Am. Psychol. 16, 694–703.

Gibson, James J., 1979. The Ecological Approach to Visual Perception. Houghton Mifflin.

Goldin-Meadow, S., Brentari, D., 2017. Gesture, sign, and language: The coming of age of sign language and gesture studies. Behav. Brain Sci. 40, E46. https://doi.org/10.1017/S0140525X15001247.

Gratier, M., Devouche, E., Guellai, B., Infanti, R., Yilmaz, E., Parlato-Oliveira, E., 2015. Early development of turn-taking in vocal interaction between mothers and infants. Front. Psychol. 6, 1167.

Gregoromichelaki, E., Kempson, R., 2013. Grammars as processes for interactive language use: incrementality and the emergence of joint intentionality. In: Perspectives on Linguistic Pragmatics. Springer, Cham, pp. 185–216.

Harnad, S., 1990. The symbol grounding problem. Phys. D 42, 335–346. https://doi.org/10.1016/0167-2789(90)90087-6.

Heft, H., 2001. Ecological Psychology in Context: James Gibson, Roger Barker, and the Legacy of William James's Radical Empiricism. Psychology Press, New York.

Hodges, B.H., 2007. Values define fields: the intentional dynamics of driving, carrying, leading, negotiating, and conversing. Ecol. Psychol. 19, 153–178.

Hodges, B.H., Baron, R.M., 1992. Values as constraints on affordances: Perceiving and acting properly. J. Theor. Soc. Behav. 22, 263–294.

Hodges, B.H., Rączaszek-Leonardi, J., 2022. Ecological values theory: beyond conformity, goal-seeking, and rule-following in action and interaction. Rev. Gen. Psychol. 26 (1), 86–103.

Jaffe, J., Beebe, B., Feldstein, S., Crown, C.L., Jasnow, M.D., 2001. Rhythms of dialogue in infancy: coordinated timing in development. Monogr. Soc. Res. Child Dev. 66 (2).

Jennings, R.E., Thompson, J.J., 2012. The biology of language and the epigenesis of recursive embedding. Interact. Stud. 13 (1), 80–102.

Kaiser, N., Butler, E., 2021. Introducing social breathing: a Model of Engaging in relational systems. Front. Psychol. 12, 1134.

Kelso, J.A.S., 1995. Dynamic Patterns: The Self-Organization of Brain and Behavior. MIT Press, Cambridge, MA.

Kelso, J.A.S., Fuchs, A., 2016. The coordination dynamics of mobile conjugate reinforcement. Biol. Cybern. 110 (1), 41–53.

Kelty-Stephen, D.G., 2017. Threading a multifractal social psychology through within-organism coordination to within-group interactions: a tale of coordination in three acts. Chaos, Solit. Fractals 104, 363–370.

Kempson, R., Cann, R., Gregoromichelaki, E., Chatzikyriakidis, S., 2016. Language as mechanisms for interaction. Theor. Ling. 42 (3–4), 203–276. https://doi.org/10.1515/tl-2016-0011.

Korbak, T., Zubek, J., Kuciński, Ł., Miłoś, P., Rączaszek-Leonardi, J., 2021. Interaction history as a source of compositionality in emergent communication. Interact. Stud. 22 (2), 212–243.

Korbak, T., Zubek, J., Rączaszek-Leonardi, J., 2020. Measuring non-Trivial Compositionality in Emergent Communication arXiv preprint arXiv:2010.15058.

Laland, K.N., Odling-Smee, F.J., Feldman, M.W., 1999. Evolutionary consequences of niche construction and their implications for ecology. Proc. Natl. Acad. Sci. USA 96 (18), 10242–10247.

Lee, D.N., 1976. A theory of visual control of braking based on information about time-to-collision. Perception 5, 437–459.

Leonardi, G., Nomikou, I., Rohlfing, K.J., Rączaszek-Leonardi, J., 2016. Vocal interactions at the dawn of communication: the emergence of mutuality and complementarity in mother-infant interaction. Proc. IEEE ICDL-EpiRob, 288–293. Cergy-Pontoise.

Levinson, S.C., 2006. On the human "interaction engine". In: Enfield, N.J., Levinson, S.C. (Eds.), Roots of Human Sociality: Culture, Cognition and Interaction. Berg, Oxford, pp. 39–69.

Malloch, S., Trevarthen, C., 2009. Musicality: Communicating the vitality and interests of life. Commun. Musical. Explor. Basis Human Compan. 1, 1–10.

Merleau-Ponty, M., 1962/2005. Phenomenology of Perception. Routledge, London and New York. Trans. Colin Smith.

Marsh, K.L., Johnston, L., Richardson, M.J., Schmidt, R.C., 2009. Toward a radically embodied, embedded social psychology. Eur. J. Soc. Psychol. 39, 1217–1225.

Mordatch, I., Abbeel, P., 2017. Emergence of Grounded Compositional Language in Multi-Agent Populations. AAAI.

Moreno-Núñez, A., Alessandroni, N., 2021. What's an early triadic interaction made of? A methodological proposal to study the musical dynamics of interaction. Infant Behav. Dev. 63, 101572.

Murray, L., Trevarthen, C., 1985. Emotional regulation of interactions between 2-month-olds and their mothers. In: Field, T.M., Fox, N.A. (Eds.), Social Perception in Infants. Ablex, Norwood, NJ, pp. 177–197.

Nomikou I. (under review). Joining Actions through Sounds: Mothers and Infants in Routine Activities. Language and Communication.

Nomikou, I., Leonardi, G., Rohlfing, K., Rączaszek-Leonardi, J., 2016. Constructing interaction: the development of gaze dynamics. Infant Child Dev. 25 (3), 277–295. https://doi.org/10.1002/icd.1975 (IF=1,14), MNiSW 25.

Nomikou, I., Rohlfing, K.J., 2011. Language does something: body action and language in maternal input to three-month-olds. IEEE Trans. Auton. Mental Dev. 3 (2), 113–128. https://doi.org/10.1109/TAMD.2011.2140113.

Papoušek, M., Papoušek, H., 1981. Musical elements in the infant's vocalization: Their significance for communication, cognition, and creativity. Adv. Infancy Res. 1, 163–224.

Popova, Y.B., 2015. Stories, Meaning and Experience: Narrativity and Enaction. Routledge, New York.

Popova, Y.B., Cuffari, E., 2018. Temporality of sense-making in narrative interactions. Cognit. Semiot. 11 (1).

Popova, Y.B., Rączaszek-Leonardi, J., 2020. Enactivism and ecological psychology: the role of bodily experience in agency. Front. Psychol. 11, 539841. https://doi.org/10.3389/fpsyg.2020.539841.

Port, R.F., van Gelder, T. (Eds.), 1995. Mind as Motion: Explorations in the Dynamics of Cognition. The MIT Press.

Rączaszek-Leonardi, J., 2012. Language as a system of replicable constraints. In: Laws, Language and Life. Springer, pp. 295–333.

Rączaszek-Leonardi, J., 2016. How does a word become a message? An illustration on a developmental time-scale. New Ideas Psychol. 42, 46–55.

Rączaszek-Leonardi, J., Nomikou, I., 2015. Beyond mechanistic interaction: value-based constraints on meaning in language. Front. Psychol. 6, 1579. https://doi.org/10.3389/fpsyg.2015.01579.

Rączaszek-Leonardi, J., Nomikou, I., Rohlfing, K.J., 2013. Young children's dialogical actions: The beginnings of purposeful intersubjectivity. IEEE Trans. Autonom. Mental Dev. 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. Ecol. Psychol. 30 (1), 39–73.

Recanati, F., 2012. Compositionality, flexibility, and context-dependence. In: Hinzen, W., Werning, M., Machery, E. (Eds.), The Oxford Handbook of Compositionality. Oxford University Press, Oxford, pp. 175–191.

Reed, E.S., 1995. The ecological approach to language development: a radical solution to Chomsky's and Quine's problems. Lang. Commun. 15 (1), 1–29.

Reddy, V., 2018. Why engagement? a second-person Take on social cognition. In: De Bruin, L., Newen, A., Gallagher, S. (Eds.), The Oxford Handbook of 4E Cognition. Oxford University Press, pp. 433–452.

Reddy, V., 2019. Meeting infant affect. Dev. Psychol. 55 (9), 2020–2024. https://doi.org/10.1037/dev0000773.

Reddy, V., Morris, P., 2004. Participants don't need theories: knowing Minds in engagement. Theor. Psychol. 14 (5), 647–665. https://doi.org/10.1177/0959354304046177.

Rogoff, B., 2003. The Cultural Nature of Human Development. Oxford University Press.

Rohlfing, K.J., Leonardi, G., Nomikou, I., Rączaszek-Leonardi, J., Hüllermeier, E., 2019. Multimodal turn-taking: motivations, methodological challenges, and novel approaches. IEEE Trans. Cognit. Dev. Syst. 12 (2), 260–271.

Rossmanith, 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. Front. Psychol. 5, 1390. https://doi.org/10.3389/fpsyg.2014.01390.

Schegloff, E.A., Ochs, E., Thompson, S., 1996. Introduction. In: Ochs, E., Schegloff, E.A., Thompson, S. (Eds.), Interaction and Grammar. Cambridge University Press, Cambridge.

Searle, J.R., 1980. Minds, brains and programs. Behav. Brain Sci. 3, 417–424. https://doi.org/10.1017/S0140525X00005756.

Sheets-Johnstone, M., 1999. The Primacy of Movement. John Benjamins, Amsterdam.

Stawarska, B., 2015. Saussure's Philosophy of Language as Phenomenology: Undoing the Doctrine of the Course in General Linguistics. Oxford University Press, Oxford.

Steffensen, S.V., 2011. Beyond mind: an extended ecology of languaging. In: Cowley, S.J. (Ed.), Distributed Language. John Benjamins, Amsterdam, pp. 185–210.

Steffensen, S.V., Fill, A., 2014. Ecolinguistics: The state of the art and future horizons. Lang. Sci. 41, 6–25.

Steffensen, S.V., Harvey, M.I., 2018. Ecological meaning, linguistic meaning, and interactivity. Cognit. Semiotic. 11 (1).

Steinert-Threlkeld, S., 2020. Toward the Emergence of nontrivial compositionality. Philos. Sci. 87 (5), 897–909.

Stern, D.N., 1999. Vitality contours: the temporal contour of feelings as a basic unit for constructing the infant's social experience. In: Rochat, P. (Ed.), Early Social Cognition: Understanding others in the First Months of Life. Lawrence Erlbaum Associates Publishers, pp. 67–80.

Stern, D.N., 2000. Putting time back into our considerations of infant experience: a microdiachronic view. Infant Ment. Health J. Off. Publ. World Associ. Infant Mental Health 21 (1–2), 21–28.

Stevanovic, M., Peräkylä, S., 2015. Experience sharing, emotional reciprocity, and turn-taking. Front. Psychol. 6, 450.

Stivers, T., Enfield, N.J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., et al., 2009. Universals and cultural variation in turn-taking in conversation. Proc. Natl. Acad. Sci. USA 106 (26), 10587–10592. https://doi.org/10.1073/pnas.0903616106.

Thelen, E., Smith, L.B., 1994. A Dynamical Systems Approach to the Development of Cognition and Action. Bradford-MIT Press, Cambridge, MA.

Trevarthen, C., Gratier, M., Osborne, N., 2014. The human nature of culture and education. Wiley Interdiscipl. Rev. Cognit. Sci. 5 (2), 173–192.

Tronick, E., Als, H., Adamson, L., Wise, S., Brazelton, T.B., 1978. The infant's response to entrapment between contradictory messages in face-to-face interaction. J. Am. Acad. Child Psychiatr. 17 (1), 1–13.

Valsiner, J., 1987. Culture and the Development of Children's Action. Wiley, Chichester, England.

Van Dijk, L., Rietveld, E., 2017. Foregrounding sociomaterial Practice in our Understanding of affordances: the skilled intentionality framework. Front. Psychol. 7, 1969. https://doi.org/10.3389/fpsyg.2016.01969.

Van Dijk, L., Withagen, R., 2016. Temporalizing agency: moving beyond on-and offline cognition. Theor. Psychol. 26, 5–26. https://doi.org/10.1177/0959354315596080.

Van Geert, P., van Dijk, M., 2021. Thirty years of focus on individual variability and the dynamics of processes. Theor. Psychol. 31 (3), 405–410.

Van Orden, G.C., Holden, J.G., Turvey, M.T., 2003. Self-organization of cognitive performance. J. Exp. Psychol. Gen. 132 (3), 331.

Von Bertalanffy, L., 1968. General System Theory: Foundations, Development, Applications. George Braziller, New York.

Vygotsky, L.S., 1978. Mind in Society: The Development of Higher Psychological Processes. Harvard University Press, Cambridge, MA.

Warlaumont, A.S., Richards, J.A., Gilkerson, J., Oller, D.K., 2014. A social feedback loop for speech development and its reduction in autism. Psychol. Sci. 25 (7), 1314–1324.

Webber Jr., C.L., Zbilut, J.P., 2005. Recurrence quantification analysis of nonlinear dynamical systems. Tutor. Contemp. Nonlinear Methods Behav. Sci., 26–94.

Westermann, G., Jones, S., 2021. Origins of dissociations in the English past tense: a synthetic brain imaging model. Front. Psychol. 12, 2452.

Wijnants, M., Cox, R., Hasselman, F., Bosman, A., Van Orden, G., 2012. A tradeoff study revealing nested timescales of constraint. Front. Psychol. 3, 116. https://doi.org/10.3389/fphys.2012.00116.

Wilson, M., Wilson, T.P., 2005. An oscillator model of the timing of turn-taking. Psychonomic Bull. Rev. 12 (6), 957–968.

Zapata-Fonseca, L., Dotov, D., Fossion, R., Froese, T., 2016. Time-series Analysis of embodied interaction: movement variability and complexity matching as dyadic properties. Front. Psychol. 7, 1940. https://doi.org/10.3389/fpsyg.2016.01940.