



Coupling Cognitive Expectations and Anticipations in Dialogue (as Based on Russian Mother-Child Interaction)

Anastasia Vladimirovna Kolmogorova 

Siberian Federal University

Russian Federation

nastiakol@mail.ru

Received 11 July 2020; accepted 30 January 2021; published 16 September 2021.

Abstract

The paper deals with the problem of how mothers and their babies use multi-scalar temporalities to couple their cognitive activity while interacting. Assuming that everyday communication is based on the coordinated behavior of two or more agents, I argue that the crucial cognitive principle which enables such coordination is the overlapping of cognitive expectations and cognitive anticipations performed by interacting people. To address the problem of the genesis of human capacity to expect and to anticipate the actions of others, I analyze a large corpus of video recordings of Russian mothers speaking and playing with their babies from 0 to 4 years old. Being applied to the video data, the method of Cognitive Event Analysis, supported by the use of Elan software, has showed that the baby learns to adjust their vocalizations, gestures, gazes and movements to how a caregiver's activity is directed to the infant from very early on. The multi-scalar perspective in analysis helped to degage “cognitive event pivots” in such interactions—the moments which divide the interaction into two parts: that of before and after. Seeking to attune their behavior to the mother's, the baby tries to imitate it mimetically. If they succeed, they both feel the satisfaction of understanding, which anchors such a “valuable” cognitive result in child cognitive experience. The research led to the conclusion that, while communicating, a mother-child dyad forms two brain-body systems whose coupling amalgamates in the moment and results in distributed cognition achievements.

Keywords: dialogue; mother-child interaction; distributed cognition; cognitive dynamics; multi-scalarity; everyday communication

1. Introduction

This paper pursues the problem of how human agency enables very small and slow changes in agents' cognitive experience to yield conceptual "jumps" across its relatively stable domains. Formulated in these terms, the research question demands to be examined around issues of cognitive dynamics.

In a broad sense, scholars understand cognitive dynamics as any changes in an organism that directly affects an organism's cognitive processing or cognitive capacities (Dietric & Markman, 2000, p. 7). Within the framework of distributed cognition, the concept comes to focus on how human agency yields cognitive achievements that affect conceptual domains of both agents. Referring to the example of "an exchange of greeting situation", S. Cowley details a number of factors (settings, time of day, mood, and feelings) that influence the intersubjective orientations of participants. He notes:

Multi-scale dynamics move dyads to feel, think, and act. As one person controls and hears, others perceive the speaking, gesturing, play of expression, shifts in posture etc. that constitute the cognitive dynamics (Cowley, 2012, p.22).

As such, cognitive dynamics are to be seen as the "key to understanding human sociality, culture and agency" (Steffensen & Fill, 2014, p. 19).

According to a distributed perspective on cognition, cognitive processes are based on "shallow thinking" (Cowley, 2004)—a kind of human reasoning reliant on the context-bound mind-body system that is triply distributed (Hollan et al., 2000): across time, across the supporting material environment, and across human community. The multi-faceted theoretical grounding of Huchins' work already suggests the idea of a multidisciplinary approach to the study of distributed phenomena. Adopting these principles, theories of language, development and communication drastically modify their objects—they shift attention from strings (of verbal signs or communicative signals), to the dynamic processes that unfold. It is like turning to a 3D dimension that focuses on languaging and interacting whose dynamics place time at the heart of observation and description (Cowley, 2011). The multi-scalar temporality in interaction enables humans to bring the past to the present and to anticipate future events (Cowley & Madsen, 2014).

Using the method of Cognitive Event Analysis (Steffensen, 2013), the paper shows how the baby, like an excellent social "tabula rasa", mimetically attunes his own behavior to that of his mother (or caregiver). Weight falls on how the pico- and microscales of gazing, eye and head movements, vocalizations, and wording influence how mother and baby unite their activity. In doing so, they establish mutual understanding that arises with the overlapping of cognitive expectations and anticipations.

The aim of the article is to describe the trajectory of slow changes affecting conceptually stabilized domains within the “mother-baby” couple and, thus, to validate the hypothesis that dialogicality¹ is primordial to agent, in that it ensures the integration of two body-brain systems in a given cultural and social environment. This, however, is achieved in both the moment and as part of a cultural settings that grants multi-scalarity to what happens.

In the next section of the paper (§ 2), I examine the concept of dialogical communication in the light of distributed cognition and, above all, treat dialogue as a minimal perimeter of coordinated behavior. Then, in what follows (§ 2), I describe the data and research methodology. After that (in § 4), I present the results of Cognitive Event Analysis (CEA) of the video recordings while emphasizing the multi-scalar organization of mother-baby cognitive interaction. In the penultimate section (§ 5), I use the material to spell out the paper’s theoretical implications. Finally, in (§6) I summarize how the concept of cognitive coupling relates to the very basic statements of the theory of Distributed Language Approach, and, mainly, its key notion of languaging.

2. Expectations and anticipations in dialogical communication: a distributed view

In the last two decades, the humanities have seen drastic changes in the notion of communication. Having been influenced by a number of innovative approaches in cognitive science, linguistics, biology, and semiology, it has acquired many facets and diverse dimensions. It became very different from the original concept that was associated with Shannon’s communication channel. For instance, U. Maturana and F. Varela (Maturana & Varela, 1992) traced communication to the bio-cognitive principal of body-world coupling with the natural environment. This way, communication gained new value as it took on a cognitive status. It increasingly became integrated with concepts such as intersubjectivity (Trevarthen, 2005), embodiment (Gibbs, 2005), interactivity (Thibault & King, 2016), enaction (Stewart et al., 2010), and agency (Cowley & Fester, 2017).

The distributed view on language and communication focuses particularly on the process of intersubjective coordination organized as a “set of practices that use *co-feeling* to shape activity around the understanding” (Cowley, 2004, p. 584).

A key word mentioned in the above definition is *co-feeling* which yields understanding. The concept has been largely discussed over recent decades in child development studies. According to researchers, emotions not only underlie a child’s growing sociality (Hepach & Tomasello, 2020), but, furthermore, act as “the core capacities necessary for cognitive and language development” (Greenspan & Shanker, 2007, p.128). Early patterns of affect signaling prompt simultaneous evolution in language interaction of first order activities and second order structures (Fowler, 2010; Spurrett & Cowley, 2004). The former combine vocalizations, sights, gestures, and movements of cognitive agents that perform concerted intersubjective

¹ The concept is introduced in (Linell, 2007) to denote “the dynamic abilities to take part in interactions with others and with sociocultural contexts as well as physical environments” (p.618).

activities, while the latter represent the more commonly described language structures (combinations of signs of a given language) to which members of a given society attach the most significant values, meanings, and behavioral patterns.

During the first order activity or *linguaging*, people draw on wording to create a notional connection between structures of a common social world and language users in attaching meaning to things and events (Cowley, 2016). This perspective suggests that every day communication offers an excellent environment for linguaging, which inevitably implies the ability of agents to co-feel.

Nina Arutyunova, a famous Russian linguist, assumes that co-feeling can be the aim of everyday communication:

“In a situation of everyday talk the speaker expects that his partner seeks to adjust his behavior to fit the speaker’s feelings and ideas. The addressee is expected to accept this invitation to feel similarly to the way the speaker does. This is what the speaker wants” (Arutyunova, 1999, p. 655).

Linguistic tradition tends to underestimate the importance of co-feeling enabling first order activity—it seems to be something additional to “wording”. However, it is the most important way of mutual adjustment and intersubjective attunement not only in infants, but in adults too: “although the entire systemic behavior of human society depends on the cognitive properties of the components themselves (infants—A.K.), these cognitive properties emerge in the domain of linguaging (the relations between agents) as systemic behavior of the human society” (Kravchenko, 2016, p. 110).

As a living system, a newborn infant still has a limited number of emerging affordances that can be used to form relationships with the environment. However, at 3 months they are already able to fit activity to some primordial culturally based expectations of his caregivers (Cowley et al., 2004). It is worthy of note that this can be done long before the infant can say a word. By the age of entering primary school, the child can easily meet the norms of intracommunity communication and to “shape understanding” in joint activity with others. When and how does the infant manage to come up with the cultural settings shared by community members?

According to Jens Madsen (Madsen, 2017, p.170) therefore, it is necessary to think of human cognition as an emergent property that stems from the neural, the body, the interactions, *and* the temporal. The individual brain-body system meets distributed norms that frame communication against the cultural settings of a community whose patterns of coordinated behavior yield multi-scalar temporality action. As in everyday life, we tend to wonder “how” (to join our activities, to be a part of community, to feel oneself comfortably with the others etc.) and leave aside the “what” and, for this reason, find ourselves included in spatio-temporal aspects of living in a given environment. This confers everyday life with a particular *modus vivendi* (Syrov, 2000) whereby an individual brain-body system is able to attune its functions to the community’s cultural setting. It prompts a biological infant to follow the path of socially mediated individuation interpreted in terms of M. Bakhtin (Bakhtin, 1990, p.16) who demonstrated that as human beings “we are constantly and intently on the watch for reflections of our own life on the plane of other people’s consciousness, and, moreover, not just reflections of particular moments of our life, but even reflections of the whole of it.” Such a “path” in-

vitably builds on dialogue—and languaging—as the most natural form of intersubjective interaction. When dialoguing in our everyday life we constantly learn to anticipate our partner’s expectations and to fit them by attuning our behavior patterns. The model we usually run while dialoguing may look like the following:

I, communicant Z, expect that my partner X

Shares with me the same cultural settings of dialoguing, will use communicative patterns which fit such cultural settings and only if I could anticipate him saying Y and, finally, understand him.

And if I am communicant X, I anticipate expectations of Z

Using shared cultural settings and expected communicative patterns to be easily anticipated saying Y and, finally, understood.

This way, a dialogue is not only the initial form of speech (Vygotsky, 1987, p. 272), but also, in humans, the most prototypical perimeter of coordinated behavior. Genuine dialogue, which leads to mutual understanding, is the maximally full overlapping of expectations and anticipations.

Even being close to each other, two last terms have an important distinction deeply rooted in the ecological conception of human perception introduced by J. J. Gibson (Gibson, 1979) and developed in A. Berthoz’s theory of neural symplexity (Berthoz, 2009). Gibson’s major assumption was that perception and action are interrelated—the act of perceiving something is intrinsically coordinated with the plan of action. Berthoz echoes Gibson’s idea by introducing the concept of “projected brain”, which presumes that brains “project properties on the world” and “they predict and anticipate” (Berthoz, 2009, p.18). In other words, while interacting with the environment, an organism, backed by his previous experience and his current activities, has already a schema or a pre-image of what he will find in the world around him; he projects these predictions on the world and confers significance only to what fits to them. In such context, expectation is a basic cognitive schema that precedes the contact with the environment and anticipation, in its turn, is a kind of prediction generated on the basis of expectation and projected on the environment. In our research, we treat the “expectation—anticipation” dichotomy to denote and then distinguish the well-established behavior schemas or patterns prompting to generate predictions (expectations) and predictions themselves (anticipations).

As C. Trevarthen’s works (Trevarthen, 2005, 2011) show, human beings learn to anticipate expectations from at least the first weeks of life—even neonates appear to be disposed to attune to their mother’s behavior by coupling mother-child brain-body systems. Using the temporally limited pico-scales of gazing, voicing, moving, and vocalizing, both mother and baby amalgamate their activities. Such amalgam gives them the impression of co-feeling and entangles them into a sense-making activity whose result will influence the future baby’s behavior within the community. Thus, temporality enters the baby’s cognitive experience to link past cognitive events and the way of pursuing future ones. The ability to anticipate expectations within a community emerges here.

In the next section, I illustrate mother-child coordination in pico-scales that yield infant cognitive experience in two cultural settings. In the future, I suggest, these will ground a mechanism of anticipating community members' expectations in two situations. This occurs: 1) when someone wants to speak to you and 2) when you are dialoguing with someone. In so doing, the mother interacts with (and guides) her baby while, correspondingly, using practices of mother-child interaction: (1) the practice of «making the child sensitive to the invitation to communicate»; and (2) the practice of «making the child sensitive to the communicative cadency».

Based on this case-study, I would like to illustrate the idea that development is the slow movement of the infant into the culturally embedded mutual coordination. The start point of developing movement is joint and mutual attention in interaction with others that forms mimetical schemas enabling, in their turn, the agreement of judgments existing between community members and leading to world conceptualization and languaging ability.

3. Research data and methods

The data corpus is built up with 81 hours of video recordings focusing on Russian mother-child interactions collected in the Krasnoyarsk and Kemerovo regions of Russia. The mothers are from 19 to 44 years old and the children from 0 to 7 years old. Each was asked to record themselves while interacting with their child in a very familiar environment—at home or on a walk. To ensure the objectivity of data, I provided mothers with a rather vague account of my purpose in describing the study aim as identifying the main steps in child speech development. The minimum duration of recording was 30 minutes and the maximum was 1 hour.

The research methodology is built on Cognitive Event Analysis (CEA) elaborated by S. von Steffensen (2013). The researcher starts from the assumption that interactivity could hardly be analyzed in a linear way, in terms of beginnings and ends. However, the coalition of external and internal sources used by agents to shape the understanding through co-action emerges in the alternation of a number of fine-grained action-perception cycles. Since they are not perceptible by simple observation, a special analytic procedure is to be applied to video records as data material. There are, specially, three heuristic elements being crucial for CEA: cognitive trajectory, cognitive event, and cognitive event pivot. Via the multiscale and microscopical focus on interbodily dynamics of gesture and movement on prosody and eye contact, the analysis seeks to elicit the emerging ecosystem cognitive trajectory, that is, “the dynamical and nonlinear path that the system creates as it achieves a given cognitive result” (Steffensen et al., 2016, p.83). The latter consists in “changes in the layout of affordances” (Chemero, 2000). Since the mentioned changes impact the ecosystem of interacting agents, it is considered a cognitive event. A transition point between two different cycles on cognitive trajectory, which makes the cognitive event happen, is perceived as a cognitive event pivot. In comparison with the well-established concept of the “aha-effect,” mainly used in research on problem-solving, the cognitive event pivot isn't supposed to be a moment of insight felt as such by an agent, but a moment of undergone achievement whose consequences on the agent-environment relations aren't directly realized by interacting people, although they are of particular importance for their future being. In analyzing video records, we follow the same five steps

which were proposed by S. V. Steffensen (Steffensen et al., 2016), but in a different order. Thus, applying CEA we 1) do the annotation of the video records by using Elan software; 2) identify a cognitive event; 3) do a cognitive trajectory segmentation in cycles; 4) identify a cognitive event pivot and 5) finally, analyze the whole dynamics of a cognitive trajectory.

Using time scaling and a detailed observation of the mother and baby's wordings, vocalizations, gazes, distances, and movements I sought transition points that divides the event trajectory into a phase 'before' and 'after' the event. Partly anticipated in (Cowley, 2003), the concept of mother-child practice as part of interaction (Kolmogorova, 2013) was also used in interpreting and systematizing the results. Such mother-child interaction practice arises in the specific ways of coupling of verbal, non-verbal, intentional, and affective behavior used by the mother in everyday communication with the child, as she sensitizes him/her to the demands of their community's interactional and cultural settings (p. 69). The concept is intended to be independent of any didactic strategy and has nothing to do with training the infant or developing a repertoire of skills. Mother-child interaction practice emerges from the principles of companionship and attachment that a mother uses "subconsciously" drawing on cultural instinct and implicit learning, which may have occurred in her own infancy.

4. Results of Cognitive Event Analysis

The first practice is that of «making the child sensitive to the invitation to communicate». As shown (Kolmogorova, 2013), Russian mothers use the practice to orient babies of 0–6 months for communicative contact with community members. In other words, the mother "pushes" her child to react vocally (verbally) to contact requests from other community members.

Cognitive Event Analysis especially focuses on a discrepancy of what the agents are doing and what happens on their cognitive level. In presenting the results of analysis, I shall often proceed with a redundant dichotomy "doings/happening".

The first example of this kind is a video recording of the mother, a 24 year old Russian woman, interacting with her baby at the end of the 8th week.

Their doings and wording (*in italic*):

01 mama	stays in the middle position from baby's head
02 mama	takes girl's hand
03 mama	looks from the side
04 mama	<i>Masha</i>
05 mama	<i>Masha my daughter let's talk</i>
06 mama	moves away from the baby
07 mama	raises baby's right hand
08 baby	raises her left leg
09 baby	raises her right hand
10 mama	releases girl's hand and takes her left leg
11 baby	raises her right leg
12 mama	<i>My little talker</i>
13 mama	pulls her baby closer to her and turns her around

14	mama	takes girl's legs
15	mama	pulls the baby to her
16	mama	releases girl's legs and takes both of the girl's hands
17	mama	moves closer to the baby
18	mama	<i>Now say AGU</i>
19	mama	looks straight into girl's eyes
20	mama	pulls girl's left hand
21	mama	moves baby's hand away
22	mama	<i>AGUUUU</i>
23	baby	raises her left leg
24	mama	remains holding baby's hands and moves them slightly
25	baby	moves her legs
26	mama	<i>AGUUU</i>
27	baby	lowers her legs
28	mama	stays in the closest position
29	mama	<i>Tsok-tsok-tsok</i>
30	baby	raises her legs and rests her feet on mama's breast
31	mama	<i>Daughter</i>
32	baby	shakes her head
33	baby	puts feet together
34	baby	is constantly looking at her mother
35	mama	constantly looks at her baby
36	mama	<i>Let's go</i>
37	baby	moves legs looking to the mother
38	baby	looks constantly at her mother and remains motionless
39	mama	<i>Haaah</i> (with admiration)
40	mama	<i>What what what</i> (with admiration and surprise)
41	mama	moves baby's hands away from girl's head

What is happening: mother begins by denoting the “Let’s talk” situation. She shows the **cultural settings** necessary for it: we should be **close to each other** (from lines 13-18 the mother, step by step, becomes more and more close to the baby), we should **look straight into the eyes of each other** (lines 19, 34, 35, and 38), and we need some **tactile contact** (to take the hands of the communicative partner).

The baby tries to attune her movements to the mother’s behavior: while the mother raises baby’s right hand up (line 7), the baby raises her left leg (line 8); when the mother takes the baby’s left leg (line 10), the baby raises her right hand (line 11). Seemingly, the infant is searching for the sympathetic 'intersynchrony' (Trevarthen et al., 2009).

The mother persists repeating the same cycle of actions ending with *AGUU*² (lines 22-26) and expects a reaction from her daughter. The baby tries to attune her movements to her mother’s and, just when the mother expects her to say something (lines 31–36), the girl looks at her mom as if she was seeking out appraisal, and, finally, shows her concentration (line 37). Then she stays motionless looking at the mother for 3 seconds (line 38). At the moment of maximum

² A non-word in mother tongue that Russians use as an onomatopoeia to denote fist baby’s words.

concentration (line 39), the mother imagines that she has heard her daughter saying something and exhibits affection with an exclamation of delight—*haaah*. After, both mother and baby restart their usual bodily movement (mother goes to take baby’s dummy, baby moves her legs and hands looking around) breaking their coupling.

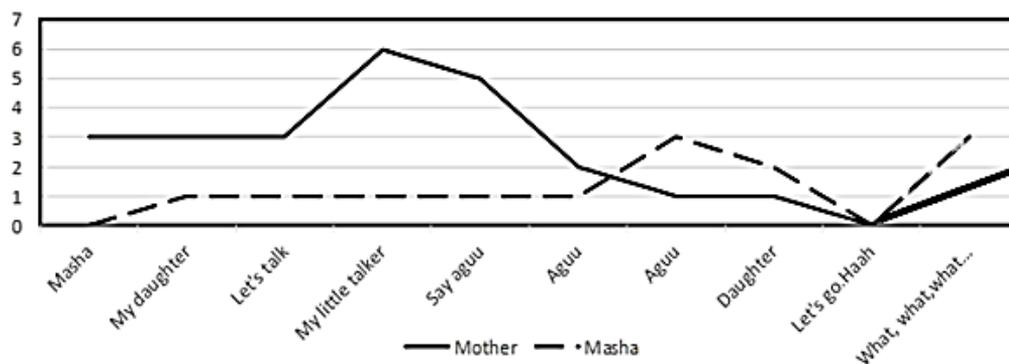


Figure 1. The intensity of mother’s and baby’s activities in their correspondence to mother’s wording

If we analyze the annotations using Elan to pick out the mother’s wordings, we can plot the fluctuation of Masha (the daughter) and the mother’s bodily and mimicking movements against wordings during the interaction (Fig. 1). The vertical axis shows the number of movements in relation to the wordings shown on the horizontal axis. The mother’s movements reach maximum intensity at the moment when she induces her baby to imitate mimetically (Zlatev, 2007) what she is now doing: *My little talker* and *Say aguu!* In response, the baby, who has been passive up to her mother’s utterance of “Aguu”, then increases her movements in ways that coincide with mother’s encouragement: *Aguu... Daughter*. At the moment denoted on the X axis as *Let’s go ...Haaah*, both decrease their activity drastically in response to the incredible concentration that they have shared. After that, the mother continues her routine movements.

Cognitive Event Analysis seeks to explore how modalities draw on separate timescales and, as shown in (Fig. 2), finds five cycles in the cognitive event. Four of them precede the event pivot or most important point of interaction (marked up on the timeline [Fig.2] as the “Haaah point”) and the fifth follows it. During the first 367 milliseconds, the mother attracts the baby’s attention in order to invite her to integrate their joint activity. In the second cycle, she spent 3 seconds denoting the situation as a “let’s talk situation.” By saying repetitively “let’s talk” she contributes to its conceptualization in the infant’s cognitive experience as a kind of very important social event and anchors the pattern of such intersubjective behavior to a second order pattern (Thibault, 2011)—to the key “let’s talk” utterance. In the third cycle, the longest one, the mother shows the sample of expected behavior conceptualized as “talk behavior”: taking the baby by the legs and then by the hands, she turns the little girl in order to ensure eye contact; she shortens the distance between herself and the baby and takes the baby’s hands to ensure body contact. Finally, she addresses the baby with four short utterances. The fourth

cycle of a 2 second duration has an encouragement function. The mother will correspond to the invitation “to become part of a dramatic narration of being in companionship with another person” (Trevarthen, 2011, p. 126). During the next 2 seconds, they both show extreme concentration which yields mutual understanding in a moment of sense-making that is followed by the mother’s affective appraisal. I consider this point as a moment of cognitive event pivot dividing the timeline of the interaction into “before” and “after.” Since this moment of co-feeling, of veritable mother-child coupling, the baby has acquired a sort of primordial grounding for future successful anticipation of community members’ communicative expectations in the “let’s talk situations.” After having lived this moment, they both feel a kind of relaxation and they continue with their routine occupations.

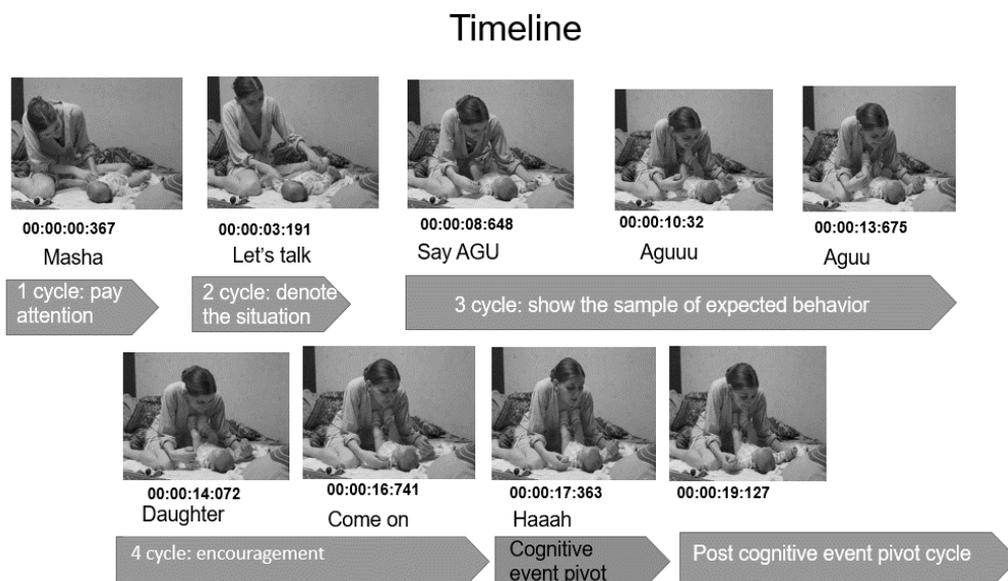


Figure 2. The Cognitive Event Trajectory of the mother-child interaction in “let’s talk situation”.

The next practice I would like to put under examination is the practice of «making a child sensitive to communicative cadency». In this case, I use observations to show that mothers prefer this practice in communicating with children of 12–34 months old. During this long period, the interactivity canvas doesn’t remain the same. At the beginning, the role of the mother is crucial—she sensitizes the baby, making them aware of what happens when two persons converse. As the child becomes more skilled, the mother’s monitoring weakens and, finally, by the end of the child’s third year, they both perceive it as a familiar interplay—it has already become a part of the child’s cultural experience of the world.

In a common sense, the cadence is a rhythmic sequence or flow of sounds in language that are bound up with the beat, time, or measure of rhythmical movement that accompanies activity. I use the term to denote a very particular, but necessary skill consisting in feeling and following

the rhythmus of dialogue so that one can find a good moment to participate vocally in exchange. To be successful in it, one should always be aware of conventional greetings, invitations, and requests. Such skills are not easy to acquire. However, towards school age, children often have the experience of communicating in families that makes it possible.

A mother exhibits a practice whereby a child learns to feel and to apply the principles of communicative cadence and attunement in dialogue. The example in question comes from a video recording of a 26 year old Russian mother interacting with her daughter of 31 months. The recording was made by the mother herself and she is not always visible, plus we can't always hear her voice.

What the mother and child are doing: they are using a well-known (by both child and mother) nursery rhyme as the mother invites her daughter to pronounce some of its parts between those which are articulated by the mother.

What happens: in doing so, they imitate the real interchange of utterances while dialoguing.

1	child	looks at the toy starts to undress the doll	
2		keeps a neutral face, is concentrated	
3			
4	mother		<i>my phone rings 00:00:00.030 - 00:00:03.300</i>
5	child	has a glance at her mother eye contact stops	
6		looking at the mother examines her toy says	
7			<i>rings 00:00:03.860 - 00:00:05.030</i>
8			
9	mother	speaks with fascination in her voice	<i>who speaks? 00:00:05.030 - 00:00:06.910</i>
10	child	holds the doll in her hands doesn't look at her	
11		mother keeps a neutral face	
12			<i>elephant 00:00:06.950 - 00:00:07.660</i>
13	mother	says modulating her voice to show the expres- sion	<i>from where? 00:00:08.390 - 00:00:09.700</i>
14	child	holds the doll in her hands bites her lip eye	
15		contact opens her mouth wide pushes her face	
16		forward and smiles	
17		looks down doesn't smile any more	<i>from the camel 00:00:16.100 - 00:00:18.820</i>
18			
19	mother	with admiration	<i>what do you want? 00:00:17.780 - 00:00:20.230</i>
20	child	tilts her head to the left and smiles	
21		looks up eye contact	<i>some chocolate 00:00:20.725 - 00:00:22.015</i>
22			

23	no eye contact examines her toy keeps a neutral face	
24		
25	mother with admiration	<i>for whom?</i> 00:00:22.230 - 00:00:24.630
26	child eye contact	
27	pushes her head forward and smiles	<i>for my son</i> 00:00:24.880 - 00:00:25.910
28	keeps a neutral face no eye contact looks down	
29		

The mother enables the child to use experience in developing the ability to anticipate the next step of the communicative partner.

Firstly (line 7), the child uses the citation principle, repeating the last word of what the mother has said. To meet the mother’s expectations in dialogue takes some time—the pause between mother’s first utterance and child’s response lasts 830 milliseconds (Table 1, line 1). Then (line 12), the child leaves the citation principal, succeeds to anticipate the dialogical expectation of her communicative partner, and completes the sequence by an adjacent turn: *who speaks?—elephant*. Furthermore, she does it rather quickly—the pause duration is 50 milliseconds. It is the shortest pause in the sample. However, in these two first sequences (Table 1.) the girl remains emotionless, seems to be very concentrated on trying to comprehend what kind of behavior she is expected to show: she looks at her doll, then starts to undress the doll, keeping a neutral face the whole time. She looked at her mother and stayed in eye contact with her only once—before she echoes “rings” (line 7).

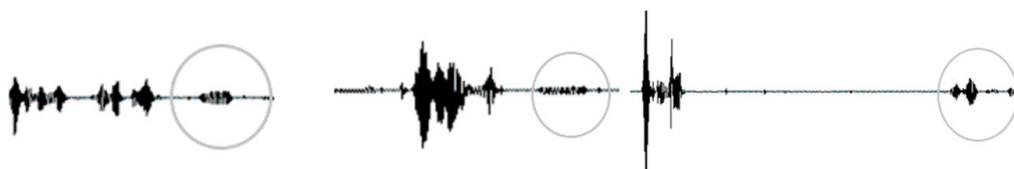
Table 1. The pausing duration in mother and child dialogue-recitation

№	Sequences in conversation	Pauses between turns
1	<i>my phone rings—rings</i>	830 milliseconds
2	<i>who speaks?—elephant</i>	50 milliseconds
3	<i>from where?—from the camel</i>	6400 milliseconds
4	<i>what do you want?—some chocolate</i>	495 milliseconds
5	<i>for whom?—for my son</i>	250 milliseconds

Seeing her mother persisting and offering her emotional support—by voicing, the mother always shows her fascination of what they are doing together—the girl not only completes the next sequence by an adjacent turn (line 17; Table 1, line 3), but also imitates her mother’s prosody. Although, it is a very difficult step for the child and she spends 6400 milliseconds to prepare but finally she succeeds. As shown (Fig. 3 A-B), while completing sequences 1 and 2

the child is not aware of the intonation: it is smooth and differs drastically from the mother's intonation. However, in the third sequence, her prosody mirrors the doubly repeated rise-fall-rise intonation of her mother's voicing (Fig. 3 C).

Figure 3. Oscillograms of the three first sequences in the mother-child dialoguing: A. *my phone rings—rings*; B. *who speaks?—elephant*; C. *from where?—from the camel*. The child's turns are given in circles.



In the next step (lines 20-22; Table 1, line 3), the girl goes further—she not only anticipates the utterance and imitates the intonation, but she begins gesturing in the same way her mother does: tilts her head to the left and smiles. The mother shows her admiration.

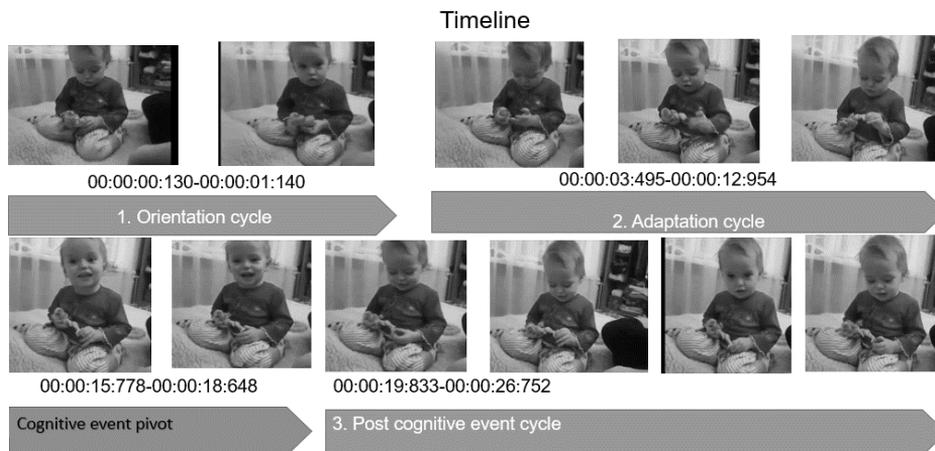
By the end of the interaction, we can see that the child's emotional excitement declines, but the principle of anticipation is already gained: the girl needs less time than in her previous attempts to complete the sequence initiated by her mother (250 milliseconds) and adjusts her prosody to her mother's voicing.

If we consider such an activity as a cognitive event, we can see 3 cycles (Figure 4). Two of them—the orientation cycle and adaptation cycle—precede the cognitive event pivot, and the third activity follows it. The orientation cycle occurs when the mother starts by showing her communicative expectations in a dialogue. It is the shortest (1 second): the girl only repeats her mother's words and she hasn't yet managed to link the mother orienting behavior to her own actions. The adaptation cycle during which the child tries to adjust her behavior to what the mother expects of her is the longest one. The cycle lasts 9 seconds, of which 6 seconds are a pause.

During this period, the child understands that the mother expects her not only to repeat words, but also to complete a sequence by a coupled turn articulated in tune to her mother's intonation and movements.

At the point of the cognitive event pivot, the mother and daughter experience a wonderful moment of co-feeling: they are about to shape the mutual understanding happening on multiple scales—they form dialogical sequences by using paired turns (wording) and, simultaneously, they maintain an “embodied conversation” (Cassell et al., 2000) by attuning their articulation and gesturing. They achieve a multi-scalar, sense saturated coordination. Even if their speech capacities are always different, the use of well-known nursery rhymes helps them overcome this difficulty. Since they are not any more constrained by words, they are allowed to pay attention to more complex principles of human agency: attention, embodiment, overlapping of communicative expectations and anticipations.

Figure 4. The Cognitive event trajectory of the mother-child interaction in a “reciting nursery rhymes” situation.



After the cognitive event pivot, the duration of the pause decreases—the child has successfully assimilated the principle of cadency and can repeat it in following communications.

5. Discussion

Our observations show a kind of discrepancy existing between the mother-child interaction and the things that really happen in their communication. After a first, very superficial look at their behavior, we will not see anything, apart from funny conversations, which have nothing to do with the sense-saturated communication.

But, upon a closer inspection, things become more complicated.

The fact that in the examined recordings the interaction does not rely mostly on linguistic behavior helps to highlight an important thing—cognitive dynamics emerges beyond linguistic forms and it does not necessarily need words. Although Masha, the little baby of about three months, is not yet able to pronounce any syllables suitable for linguistic analysis, she can act as a social agent, making her behavior predictable. At the moment when she is expected to “speak”, she keeps incredible concentration on maintaining body and eye contact with her caregiver. Although the two year old girl has not yet enough competences to maintain a conversation, prompted by her mother’s affective support, she grasps the crucial thing to do while dialoguing—she is sensitive to the interlocutor’s expectations.

Both examples evidence the dynamics of interindividual synergy within a coupled brain-body system.

In each case, it starts with several preparatory cycles which aim to orientate the infant, to provide them with a good model of expected behavior, to encourage them, and to create a domain for joined attention and activity. After this is done, it is time for the cognitive event pivot.

The important changes in behavior outline it: in the “let’s talk” situation the cognitive event pivot launches an increasing number of mother and daughter activities after a moment of zero action. In a “reciting nursery rhymes” situation, the silence between two adjacent turns starts to decrease drastically after an extensive pause. According to (Keevallik, 2018, p. 120), “long between-turn silences are by-products of the simultaneous maintaining of two activities.” Our example above shows that such an “extensive silence” is a marker of intensive cognitive processing work underlying current conversational routines in the child’s activity. This illustrates what a cognitive event pivot is. It is a moment of hard cognitive work in the coupled brain-body systems resulting in a significant jump in the cognitive experience of one or both agents.

Our analysis of two mother-child interactions shows that the cognitive event occurs in the moment of co-feeling when the dialogical interaction yields mutual understanding linked to the overlapping of agents’ anticipations and expectations. When anticipating, one integrates “both sensory and motor processes in anticipatory meaning generation based on the history of experience and cultural context [...]” (Järvilehto et al., 2009, p. 509). Furthermore, in the context of this paper, expectations mean previous experience obtained in the past to construe beliefs about future events.

In the “let’s talk” example, the baby learns to anticipate the contact-building situation in a dialogue. Russian cultural traditions mean that the baby is expected to be in the closest position to the interlocutor, maintaining eye and body contact. In the “reciting nursery rhymes” situation, the child achieves the principle of communicative cadency in the dialogue. This happens because she knows that while dialoguing, she is expected to take turns with another person, adjusting mimetically her voicing and gesturing to her partner’s vocalizations and movements.

However, such cognitive results are not easy to achieve, as they demand a lot of effort from both agents. The enormous silent work done within the two “mother-child” dyads becomes visible when analyzed on pico-scales. Small changes in vocalizations, movements, mimics, and distances mapped onto the time scales make the scaffolding of a coupled cognitive system formed by the mother and her child more evident.

Moreover, this is a dynamic and distributed system where small changes achieved in multi-agent and multi-scalar interactions slowly affect the conceptually stabilized domain of values, norms, socially relevant judgments, and cultural settings.

6. Conclusion

The mother-child dialogical communication gives us a nice example of how two brain-body systems are coupled and amalgamated at a specific moment in time. This interaction results in distributed cognition achievements.

The baby tries to attune his behavior to the behavior of his mother and to imitate it mimetically. Within a dyad, they develop co-understanding by drawing on consistent mother-infant practice where the pico-dynamics play a crucial role in forming relationships and enables both parties to draw on the ‘signs of culture’ (Cowley et al., 2004). The sensorimotor dynamics helps the baby assimilate cultural norms by accommodating to the mother’s practices. It is done through building the intrinsic motives that allow her to meet and shape the mother’s expectations. If

they succeed, they both feel the satisfaction of understanding, which anchors such “valuable” cognitive result in the child’s cognitive experience. Babies individuate such that they can draw on socioculture to understand and to come up with things in the moment—in their own ways.

The example of the babies’ growing communication skills in a dialogue demonstrates that languaging is embedded in the cognitive event structure. Our analysis suggests that the cognitive simplicity of this dialogue goes further than, for instance, Bruner’s idea of Language Acquisition Support System (Bruner, 1985), because it tends to broaden his major statement that learning to talk is not simply learning words but learning how to do things with them till the implication that learning to talk is simply a top of the iceberg of “learning to shape understanding within a culture”.

The event pivot represents a transition point when the caregiver shares with the baby her own cognitive niche. Due to the sympathetic mirroring, the baby, seeking a place in the community, gains new socially relevant and culturally meaningful patterns of verbal and non-verbal behavior for which not only words, but the whole interaction, including vocalizations, gazing, gestures, and eventually words are significant.

Finally, it should be pointed out that the phenomenon of cognitive coupling might be considered not only as a mechanism allowing the child to socially grow but also as the general foundation ensuring the drift of cognitive dynamic systems specific of human beings.

Acknowledgements

This article is developed from a contribution to the workshop on “Multi-scalar temporality: the case of languaging” organized as a part of the 4th Avant-Conference “Trends in Interdisciplinary Studies”, Porto, Portugal, October, 2019. I would like to thank all the participants for their valuable comments. I am particularly indebted to Stephen Cowley whose insightful comments helped to shape an earlier draft of the paper.

References

- Arutyunova, N.D. (1999). *Yazyk i mir cheloveka* [Language and the human world]. Moscow: Yaziki russkoy kultury. (In Russian)
- Bakhtin, M.M. (1990). Author and Hero in Aesthetic Activity. In M. Holquist, V. Liapunov (Eds), *Art and Answerability* (pp.4–257). Austin: Texas UP, 1990.
- Bruner, J. (1985). Child’s Talk: Learning to Use Language. *Child Language Teaching and Therapy*, 1(1):111–114. DOI:10.1177/026565908500100113
- Berthoz, A. (2009) The Human Brain “Projects” upon the World, Simplifying Principles and Rules for Perception. In: Berthoz A., Christen Y. (eds) *Neurobiology of “Umwelt”*. Research and Perspectives in Neurosciences. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-85897-3_3.

- Cassell, J., Stone, M., Yan, H. (2000). Coordination and context-dependence in the generation of embodied conversation. *Proceeding of the First International Conference on Natural Language Generation*, 14, 171–178.
- Chemero, A., Klein, C. & Cordeiro, W. (2003). Events as Changes in the Layout of Affordances. *Ecological Psychology*, 15:1, 19-28. DOI: 10.1207/S15326969ECO1501_02
- Cowley, S. J. (2003). Distributed cognition at three months: caregiver-infant dyads in kwa Zulu. *Natal. Alternation*, 10(2), 229–257.
- Cowley, S.J. (2004). Contextualizing bodies: how human responsiveness constrains distributed cognition. In D. Spurrett (Ed.), Special issue on Integrational Linguistics and Distributed Cognition. *Language Sciences*, 26/6, 565–591.
- Cowley, S. J., Moodley, S., & Fiori-Cowley, A. (2004). Grounding signs of culture: Primary intersubjectivity in social semiosis. *Mind, Culture, and activity*, 11(2), 109–132.
- Cowley, S.J. (2011). Distributed language. In S.J. Cowley (Ed.), *Distributed language* (pp. 1–14). Amsterdam: Benjamins.
- Cowley, S.J. (2012). Cognitive dynamics: language as values realizing activity. In A. Kravchenko (Ed). *Cognitive Dynamics and Linguistic Interactions*, pp. 15-46. Newcastle upon Tyne: Cambridge Scholars Press.
- Cowley, S. & Madsen, J. (2014). Time and Temporality: Linguistic Distribution in Human Life-games. *Cybernetics & Human Knowing*, 21, 172–185.
- Cowley, S. (2016). Changing the idea of language: Nigel Love's perspective. *Language Sciences*, 61, 43–55.
- Cowley, S.J.& Fester, M.T. (2017). Voices in human agency. The power of presence. *Reti. saperi. linguaggi*. 4 (2), 303–320.
- Dietrich, E. & Markman, A.B. (2000). Cognitive Dynamics: Computation and representation regained. In E.Dietrich, A.B. Markman (Eds.), *Cognitive Dynamics* (pp.1–29). Mahwah, NJ: Lawrence Erlbaum Associates.
- Fowler, C. A. (2010). Embodied, embedded language use. *Ecological Psychology*, V, 4 (22), 286–303.
- Gibbs, R.W. (2005). *Embodiment and Cognitive Science*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511805844>
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. New York, NY: Psychology Press.
- Greenspan, S., & Shanker S. (2007). The developmental pathways leading to pattern recognition, joint attention, language and cognition. *New Ideas in Psychology*, 25 (2), 128–142. <https://doi.org/10.1016/j.newideapsych.2007.02.007>.
- Hepach, R. & Tomasello, M. (2020). Young children show positive emotions when seeing someone get the help they deserve. *Cognitive Development*, 56. DOI: 10.1016/j.cogdev.2020.100935

- Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 7(2), 174–196.
- Järvilehto, T., Nurkkala, V.M. & Koskela, K. (2009). The role of anticipation in reading. *Pragmat. Cognit.* 17 (3), 509–526.
- Keevallik, L. (2018). The temporal organization of conversation while mucking out a sheep stable. In A. Deppermann, J. Streeck (Eds.), *Time in Embodied Interaction: Synchronicity and Sequentiality of Multimodal Resources* (pp. 97–122), Amsterdam, John Benjamins.
- Kolmogorova, A.V. (2013). *Mamin yazyk: praktiki materinskogo obshcheniya v kontekste raspredelennoy modeli yazika i kognotsii* [The motherese: practices of mother-child communication in the light of distributed theory of language and cognition]. Moscow: Flinta. (In Russian).
- Kravchenko, A.V. (2016). Two views on language ecology and ecolinguistics. *Language Sciences*, 54 (2016), 102–113.
- Linell, P. (2007). Dialogicality in languages, minds and brains: is there a convergence between dialogism and neuro-biology? *Language Sciences*, 29, 605–620.
- Maturana, H.R.&Varela F.J. (1992). *The Tree of Knowledge: The biological roots of human understanding* (Revised ed.). Boston, MA: Shambala Publishing, Inc.
- Madsen, J. K. (2017). Time during time: Multi-scalar temporal cognition. In S.J. Cowley and F. Vallée-Tourangeau (Eds.), *Cognition Beyond the Brain: Computation, Interactivity and Human Artifice* (2nd ed.) (pp.155–174). Dordrecht: Springer, 155–174.
- Steffensen, S.V. & Fill, A. (2014). Ecolinguistics: the state of the art and future horizons. *Language Sciences*, 41, 6–25.
- Steffensen, S.V. (2013). Human interactivity: problem-solving, solution-probing, and verbal patterns in the wild. In: Cowley, S.J., Vallée-Tourangeau, F. (Eds.), *Cognition Beyond the Body: Interactivity and Human Thinking* (pp. 195–221). Dordrecht: Springer.
- Steffensen, S.V., Vallée-Tourangeau, F. & Vallée-Tourangeau, G. (2016). Cognitive events in a problem-solving task: a qualitative method for investigating interactivity in the 17 Animals problem. *Journal of Cognitive Psychology*, 28:1, 79-105. DOI: 10.1080/20445911.2015.1095193
- Stewart, J., Gapenne, O., & Di Paolo, E. (Eds.). (2010). *Enaction: Toward a new paradigm for cognitive science*. Cambridge, MA: The MIT Press.
- Syrov, V.N. (2000). O statuse i strukture povsednevnosti (metodologicheskie aspekty) [Towards the discussion on the status and the structure of everyday life (methodological aspects)]. *Personality. Culture. Society*, Vol.2, 147–159. (In Russian)
- Thibault, P. J. (2011). First order languaging dynamics and second order language: The distributed language view. *Educational Psychology*, V, 32, 210–245.
- Thibault, P.J.& King, M.E. (2016). Interactivity, values and the microgenesis of learning in a tertiary setting. In: C.C. Ng, R. Fox, M. Nakano (Eds.), *Reforming Learning and Teaching in Asia-Pacific Universities* (173–211). Springer, Singapore.

- Trevarthen, C. (2005). Stepping away from the mirror: Pride and shame in adventures of companionship: Reflections on the nature and emotional needs of infant intersubjectivity. In C.S. Carter, L. Ahnert, K. E. Grossman, S. B. Hardy, M. E. Lamb, S. W. Porges, & N. Sachser (Eds.), *Attachment and bonding: A new synthesis* (pp.55–84). Cambridge, MA: The MIT Press.
- Trevarthen, C. (2011). What is it like to be a person who knows nothing? Defining the active intersubjective mind of a newborn human being. *Inf. Child Develop.*, 20, 119–135. doi:10.1002/icd.689
- Trevarthen, C., Delafield-Butt, J. & Schögler, B. (2009). Psychobiology of musical gesture: Innate rhythm, harmony and melody in movements of narration. In A. Gritten & E. King (Eds.), *Music and Gesture – 2* (pp.11–43). Aldershot, UK: Ashgate.
- Vygotsky, L. S. (1987). *Cognition and language. The collected works of L. S. Vygotsky*, Vol. 1. Problems of general psychology (R. W. Rieber & A. S. Carton, Eds.). Plenum Press.
- Zlatev, J. (2007). Intersubjectivity, Mimetic Schemas and the Emergence of Language. *Intellectica*, Vol. 2–3 (46–48), 123–151.