Joining actions through effort sounds: Mothers and infants in routine activities

Iris Nomikou

University of Portsmouth, UK

ABSTRACT

This paper analyses the effort sounds made by caregivers in routine interactions with very young infants. Video recordings were made of 15 mother-infant dyads in Germany during nappy changing. The multimodal analysis of the interactions revealed that effort sounds were used when handling the infant’s body, such as dressing them or lifting them up, but also made to link to the sensations of the infant. The sounds achieved their meaning within sequences of actions and contextualised via temporality, phonetic variation and multimodality. With these vocalisations, I propose, parents can give infants’ sensations a voice, make them public, and thus achieve a co-ordination of experience.

1. Introduction

In the everyday lives of families, caregivers and children come together multiple times each day in various routines of care. Within such routines, such as getting dressed, feeding, bathing and getting ready to go out, children and caregivers need to coordinate their perspectives, their actions, and their bodies to get things done. Infants are thus born and immersed into a world of goal-directed actions which requires their participation and collaboration from day one. In the case of month-old infants, who are not yet capable of independently tending to their own basic bodily needs, routines involve being acted upon, being moved, and guided by others. For the caregivers it means performing tasks with an infant who despite limited motor skills, understanding of language, or knowledge of the task, might have other needs, and is capable of assisting or resisting being acted upon. How caregivers and infants manage these routines is what this paper focuses on by looking at nappy changing. More specifically, I focus on the ways visual, aural and haptic senses are coordinated during these interactions, and I describe a practice whereby the caregiver is sounding out movements performed on the infant’s body.

1.1. Family routines and child development

Family routines have featured in research as rich contexts for observing human social conduct (Goffman, 1959). They are means for the analysis of culture (Weisner, 2014), and allow us to see how children acquire cultural competence (Cole, 1985; Rogoff, 1990). Furthermore, it has been suggested that they are the “engines of development” (Bronfenbrenner, 1993;
Developmental psychology has emphasised the role of routines and early games in infancy. With their fixed and predetermined structure, routines allow infants to form expectations about activities and other people. For example, they can form predictions about being fed (Kochukhova and Gredebäck, 2010), being picked up (Reddy et al., 2013), and being involved in games (Bruner and Sherwood, 1976; Ratner and Bruner, 1978; Fantasia et al., 2014; Nomikou et al., 2017). Routines have been shown to facilitate coordination with the adult (Hubley, 1983; Trevarthen, 1977). Within the routines themselves, infants learn to take turns, participate in meaning making and the sharing of emotions, and calibrate attention. Yet, a large number of existing studies have focused on whether these activities enable developmental outcomes beyond the interaction, and have not investigated the way the routine itself is organised. Thus, routines contribute to the discovery of language (Bruner, 1977; Gratier, 2019; Raczaszek-Leonardi et al., 2018; Rohlfing et al., 2016) and the development of intersubjectivity (Trevarthen, 1979). Therefore, it is critical to understand the organisation of routines in detail, in order to see the practices that enable infant participation; that is, we should investigate the actual actions of managing the activity.

Everyday family routines of care have been the focus of research showing how multiple sensory modalities and materials are used by parents and older children to coordinate their actions (Tulbert and Goodwin, 2011; Goodwin and Cekaite, 2018; Wootton, 2005). A lot less is known about routines of care in infancy (but see Raczaszek-Leonardi et al., 2013; Wiggins, 2019; Wiggins and Keevallik, 2021a, 2021b; Takada, 2011, 2020; de León, 2021). Concerning nappy changing as a routine activity, research has been very scarce. Research within caregiving contexts points out that the social nature of nappy changing is overlooked because of the precedence of the physical needs of the infant (Degotardi and Davis, 2008). Indeed, the little research in the field seems to suggest that nappy changing is not very interactive because of its rigid structure, and that caregivers show less responsiveness to and, in turn, elicit fewer positive behaviours from infants (Maas et al., 2013). Also, nappy changing interactions have been found to be less attuned and harmonious than face-to-face and play interactions (Leyendecker et al., 1997). Yet these studies, as is very common in developmental psychology, have mostly used a top-down, behavioural coding approach based on pre-defined scales and coding schemes that do not emerge from looking at the details of the interactions themselves. Extant research in nappy changing has suggested instead that caregivers do not merely focus on the functional aspect of the task nor act on infants and objects instrumentally (Nomikou, 2018; Nomikou and Rohlfing, 2011; Rossmannith, 2017). Rather, infants are treated as active participants and their contributions are acted upon by caregivers in multiple modalities, such as vocal practices, intonation, facial expressions, gesture, movement and touch. The few detailed descriptions of this interactional setting suggest that nappy changing in Western societies reflects and enacts cultural values that are crucial for human interactions, such as care, collaboration, respect, attention and mutual agency (Raczaszek-Leonardi et al., 2022; Raczaszek-Leonardi and Nomikou, 2015; Lerner et al., 2011). This research furthermore conceptualises them even as a cognitive and affective context that may support musical and vocal development (Addessi, 2020).

1.2. Language and action in interaction

In these early interactions, the role of caregivers and people who have a close relationship to the infant is considered very important. Within developmental linguistics, there is a long tradition of research into parents’ talk to children. It is well known, for example, that when interacting with infants, caregivers use Child Directed Speech (Ferguson, 2004; Soderstrom, 2007), a broad repertoire of prosodic, grammatical, lexical, phonological, and discourse strategies. Their function is said to be to gain and hold the infants’ attention, establish an affective bond between infant and caregiver, and allow the early communication between them (Ferguson, 2004). Moreover, beyond being appealing to infants, research suggests that Child Directed vocal practices mark interesting aspects of the world (Csibra and Gergely, 2009). They isolate events for the infant out of the flow of an activity (Nomikou and Rohlfing, 2011) and provide an “acoustic packaging” (Hirsh-Pasek and Golinkoff, 1996), assisting infants in disentangling complex actions into their individual components.

Focusing on the coming together of people in inter-action, interactional linguistics suggests that language not only emerges from multisensorial strands of action (Mondada, 2019b), but actually never ceases to be embedded in action (Goodwin, 2000) and situated in the world and the body. These studies have revealed how vocal contributions are carefully positioned at particular points of action trajectories, how they project next actions, and make action recognizable and predictable to participants (Mazeland, 2019). Interestingly, these studies (for German see e.g. Barske and Golato, 2010) also confirm that the ways in which one individual times their verbal behaviour creates affordances that can be picked up by others. It seems, therefore, relevant to explore the coming together of vocal practices with mundane routines of care, to understand whether and how these might contribute to the management of these activities.

More recently, there has been a growing interest in vocal practices that go beyond that which has traditionally been treated as linguistic. Discourse particles, which have been seen as mainly managing verbal discourse, have been studied as markers of boundaries and transitions in activities (Keevallik, 2010). More recent research has looked at “sound objects”, revealing the systematicity with which people use, grunts/moans (Hofstetter, 2020, 2021; Hofstetter et al., 2021), sighs (Hoey, 2014), clicks (Ogden, 2020), whistles (Reber and Couper-Kuhlen, 2020), and effort sounds in their natural interactions. In some of these, for example, the bodily sensations come to the fore and become the focus of the communication; this type of
sound objects produced within those moments are physical and social events at the same time (Goffman, 1978; Hofstetter, 2021; Keevallik, 2023). This body of studies have forced us to rethink what minimal vocalisations can do. “Liminal signs” help streamline social interaction by providing a running commentary on states of mind (Dingemanse, 2020). They also function as expressions of emotion (Pehkonen, 2020; Ben-Moshe, 2023) that are acted upon by co-participants, and are therefore expressions of how people interacting with each other are persons emotionally involved with each other (Reddy, 2018).

In caregiver-infant interactions such vocalisations produced by caregivers are abundant. One type of vocalisation that has received much attention in the literature is emotive and expressive interjections (e.g., Stange, 2009), which have been mostly treated as expressions of affect and less so as interactional (for a discussion, see Dingemanse, 2023). The research on the affective aspect of interjections has mostly suggested that they serve as attention-eliciting, soothing stimuli, or are expressions of an emotion (Morikawa et al., 1988). Yet in the literature on Child Directed speech, other non-lexical vocalisations beyond interjections have been contrasted to what is considered the “high quality speech input” (for a recent review, see Rowe and Snow, 2020) and have been characterised as “nonsense” (Toda et al., 1990, p. 287) and “paralinguistic information” (Henning et al., 2005, p. 524; but see Wiggins, 2019; Wiggins and Keevallik, 2021b). The reason for the omission of non-lexical vocalisations from the discourse on Child Directed Speech is that lot of what we know about early caregiver-infant interactions comes from lab studies, and artificially quiet environments (e.g., Moore, 2006). This might give the false impression that all dyadic interactions involve a parent and child looking and smiling at each other, observing each other and looking out for interactional opportunities. This raises the question whether what we think we know about language development, Child-Directed Speech and Child-Directed-Behaviour in general can generalise beyond lab settings and (typically) Western cultures. For each particular care activity as well, it is equally important to acknowledge that the way it is carried out might also in part be culturally relative – the very posture and configuration can vary and thus fluctuate different body-speech activities. To embrace and understand the variability and gain a better understanding of the vocal practises used by caregivers when interacting with infants and what their functions might be, we need to contextualise observations.

2. Data and methods

2.1. The interactional context

The care activity I focus on, nappy changing, takes place at close range, with the interaction partners in direct contact with each other as well as with objects. Towels, wipes, clothes, creams, nappies, but also plush toys, dummies are all available for both infant and caregiver. This introduces multiple “strands of engagement” (Rossmanith, 2017). Furthermore, the task of dressing/undressing and cleaning the infant poses a number of additional situational constraints, i.e., the need to prevent a wiggly slippery infant from rolling off the changing table, and make sure the soiled nappy does not contact clothes in the surroundings or the infant. Finally, to understand the vocal practices in this setting we need to take into account the infants’ behaviours and opportunities for action.

My study focusses on three-month-old infants. At three months of age, infants have passed the “two-month transition” (Lavelli and Fogel, 2005), which has been suggested as a period of developmental reorganization. Infants increasingly start to
engage with their interaction partners through gaze, facial expressions and phonation (Stern, 2000), appear interested in people’s faces and focus their gaze on them (Nomikou et al., 2013; Bakeman and Adamson, 1984). They are not capable of lifting their body weight independently (for a review, see Adolph et al., 2010), however, they are capable of stiffening their body (or not) and this can both assist or hinder other people during everyday care activities (Reddy et al., 2013). Overall, they are dependent on other people to move their body.

Looking at the interactional setting from such a contextualised perspective, it becomes clear that if we want to understand the vocal practices used by caregivers, we need to take into consideration the interplay of multiple modalities as well as strands of action and their different dynamics. The objects, the moving of bodies, the dangers, and also the mood of the infant – all these aspects will modulate how the routine takes place and how it is managed by caregiver and infant, providing crucial insight into the use of vocal practices in the service of joint actions.

2.2. The data

Data for this analysis were drawn from a 12-hour video corpus of mother-infant interactions during nappy changing in Germany (Nomikou and Rohlfing, 2011; Nomikou, 2018). The corpus included 6 monthly recordings of 15 German-speaking families with infants between three and eight months old.

Recordings ranged from 2 to 12 min per mother-infant dyad and were made using two HD video cameras (Canon HF10; Sony 3CMOS HDV 1080i) mounted on camera stands (see Fig. 1) An external microphone was mounted on one of the cameras in order to guarantee high-quality audio recording. The second camera recorded sound via its built-in microphone.

For the analysis presented in this study, I have used the videos from the first visit—when the infants were three months old—which amounts to approximately 2 h. From all the vocalisations of the mother in the recordings of the 15 families, effort sounds were collected and explored. These were vocalisations that featured either glottal onset and/or laryngeal constriction which was either retained throughout the vocalisation or was released and trailed into a breathy, audible outbreath (see also, Keevallik (2023)). The collection includes 39 vocalisations. Analysis of the phonetic details of the mothers’ vocalisations was done through repeated listening (Walker, 2012). Analysis of infant vocalisations was carried out using the Acoustic Phonetic Catalog (Buder et al., 2013) which distinguishes infant protophones in categories instead of using phonetic transcription. The vocalisations were annotated in PRAAT and then imported into ELAN for further multimodal annotations. The final files included annotations for mothers’ and infants’ vocalisations, gaze, smiles and some selective observations of other facial expressions and body movements. This step proved necessary as the temporal coordination of micromovements was particularly important for the analysis. The ELAN annotations were then used as a basis for the generation of transcripts using the conventions for multimodal transcription put forward by Mondada (2019a).

Multimodal interaction analysis (Goodwin, 2000; Mondada, 2019a) was used to analyse the use of mothers’ vocalisations during nappy changing. Each vocalisation was analysed in its sequential environment. Special attention was paid to its relationship with the ongoing activity and the ongoing talk, its sequential position, its precise timing and its relationship to other modalities.

3. Analysis

All effort sounds in the collection occurred when there was a tactile connection between caregiver and infant. Mutual gaze was not always present. In what follows I present four extracts which exemplify that effort sounds are a way in which caregivers “sound out” their infants’ movements.

3.1. Uniting actions and sensations across participants

In this extract, the mother is undressing the infant. In the sequence below we see how she pushes up the infant’s vest to free the nappy from contact with any clothing before she can open it. This involves pushing up the top part of the vest (on the infant’s belly) and the bottom part of the vest, which the infant is currently lying on. To do so the mother has to lift the infant’s lower body to clear the space for the movement. It is during this action that she utters the effort sound ‘ha’\(^1\). In particular, we will see that the vocalisation is timed perfectly with the moving of the clothing. It sounds out the mother’s own movement as well as the movement of the clothing the infant is sensing; this ties the actions of the mother (the person who is performing the action) and sensations of the infant (the person on whom the action is being performed) into one event.

---

\(^1\) The hyphen here is used to denote the glottal stop.
3.1.1. Highlighting a part of an action

Extract 1 VP10_1T_2 (00:56 - 01:02)

In line 1, we see a preface to the upcoming action: the mother looks at the infant’s face with a smile (Wiggins and Keevallik, 2021a, 2021b; Nomikou et al., 2013) and announces the next step of the activity, i.e., lifting the infant’s body (line 1). With an inbreath “.hh” she pushes the infant’s legs slightly to the centre and continues “we make the bottom up,” while sliding her right arm under the infant’s leg and grabbing his thigh (Fig. 2.1). During this preparation phase, there is eye-contact between mother and infant. Following the announcement, in line 2 the mother moves her gaze down the infant’s body and pulls up the thigh, raising the infant’s bottom and creating space under the body; with her left hand she starts to pull the bottom part of the vest under the infant’s nappy. This whole movement takes about 0.7 s. At the peak of the lift the mother produces the vocalisation of interest “/C14-aha/C14” (line 3) which is synchronised with the swift movement of pushing the vest up his torso. Having cleared the space under the infant’s nappy she lowers her right hand and lets the infant’s bottom back down on the mat while saying “/C14bottom/C14 up” (line 4). The moment his body touches the mat the infant produces a soft vocalisation (line 5).

The effort sound here is whispered, with a glottal onset and a breathy back open vowel. It is a short burst accompanying a small, swift movement. Sequentially, it is not produced in isolation but rather forms part of a three-step sequence. The sequence is initiated by a prefacing turn which announces the next action (Fig. 2.4a and b) while grasping the infant’s body and preparing for the lift; then in a second step the actual lifting of the body occurs, and the production of the effort vocalisation while moving the clothing (Fig. 2.4c and d), followed by the lowering of the infant’s body back on the mat and returning to the starting position marked through the repetition of the same phrase (Fig. 2.4e). The vocalisation is part of the...
temporal unfolding of the action and marks one specific part of it, i.e., the moment the infant’s body has reached the maximum lift and the mother is moving the clothing the infant is lying on. The actual effort the mother has to put here is likely not big, as indicated by the smooth flow of the three-step action described above. This is also reflected in the form of the vocalisation, which is rather soft, and so one could argue that an effort sound is unnecessary here. Yet her vocalisation displays an orientation to her action and is timed so that it coincides with the point in time during which the infant’s body is in the air and he can feel clothing moved under him. This provides structure to what the infant is experiencing and at the same time links the mother’s action and the child’s sensation into a joint action. It is also interesting here that the mutual gaze between mother and infant is interrupted during the lifting, so it is possible that what the vocalisation here might also be doing is sustaining the connection between them.

In what follows, we have an example of a sequence whose smooth flow is interrupted by a slight complication, because the mother needs to readjust the position of the infant’s head after she has caused it to be pulled back. The effort vocalisation \( u::\text{wakha} \) (in line 3) is more complex here, as is the movement she has to perform – to realign the infant’s head and lift it. Like in the previous extract, we see how the effort in the mother’s voice matches the trajectory of action, but this time we can also observe how the effect of the weight of the infant’s head ostensibly regulates the quality of the mother’s vocalisation (the creakiness of the voice). For the infant, the vocalisation encapsulates the feeling of being moved and lifted from the mat like an acoustic envelope. It points both to the sensation of the infant as well as of the mother. This is an example of “mutual incorporation” in which the body of one person becomes incorporated into the voice of the other (Fuchs and De Jaegher, 2009) and it is the activity context and bodily configuration that enables this.

### 3.1.2. Voicing a mishap

Extract 2 VP9_1_1 (00:47 - 00:58)

In this sequence, the infant is still lying on his vest, so the mother needs to move his body to remove it. The mother does this by turning the infant to the side. This reduces the actual contact surface of the body with the mat and facilitates the removal of the clothing. In line 1, the mother looks at the infant’s body and grabs his right arm with her right hand (Fig. 3.1). She then looks at the infant and aligns the turning of the body with the whispered “come here” (Fig. 3.2). Having turned the infant to the side, she attempts to pull the clothing from under his body. Yet because the head of the infant was on the clothing as well, the infant’s head is pulled backwards (line 2 Fig. 3.3 and 3.4), bringing it into what might be an uncomfortable position.

The mother discontinues the pulling action, slides her left hand under the infant’s head (Fig. 3.5a) and at the moment of lifting his head she marks her effort and the infant’s potential awkward sensation by producing the vocalisation “\( u::\text{wakha} \)” (line 3). The second part “\( \text{wu} \)” is laryngealised, hearable as creakiness. This vocalisation accompanies the turning of the infant back on her left palm so that she can stabilize the head. The final part of the vocalisation “\( \text{kha} \)” is produced with a breathiness and corresponds to the lifting of the head (Fig. 3.5c). Having lifted the infant’s head, the complication has been resolved and the mother can remove the clothing.
In the vocalisation in line 3, we see a matching in time and form between the dynamics of the movement, both as effected by the mother and as experienced by the infant, and the vocalisation itself. Similarly to Extract 1, there is a segmentation of the action steps (positioning the hand under the head, stabilising it and lifting it up), with the laryngeal constriction marking one particular section of the action; the moment of stabilising the head, followed by an audible outbreath when the head is securely held and lifted. The difference here is that these action steps are incorporated in the form of a single longer vocalisation, rather than three separate units. This vocalisation is in the first instance reminiscent of Goffman’s (1978) spill-cries ("u:" and how they can accompany a loss of control, either of the person losing the control or the person experiencing the loss of control. Here, the physical connection of the participants enables the vocalisation to display both parties’ loss of control simultaneously. As already illustrated above, the vocalisation then morphs into an effort vocalisation "~wa" at the moment the mother is trying to stabilise the infant’s head in her hand. The completion of the action is marked with the release of tension in the outbreath "kha". The type of vocalisations I am describing lend themselves to modulation during the unfolding of actions and can be designed creatively (Dingemanse, 2014) to capture the trajectories and mishaps of joint action.

The extracts analysed so far illustrate how the mother in the data is using effort sounds to selectively foreground her action. I have argued that this practice not only provides some indication about the structure and management of the activities but also provides a connection between these actions and the sensory experience of the child. Vocalising through the movements to which the infant is subjected acknowledges the infant as a participant rather than a body that is handled and moved, and potentially connects movements and sensations across participants into one unified joint action. We will now turn to further examples of distributed agency, namely instances in which the effort sounds produced by the caregiver sound out actions initiated by the infant.

3.2. Sounding another’s sensation

The following example is of the infant stretching his body. The mother is orienting to the changes of the infant’s body tension and sounds out the experience of stretching providing the voice for the infant’s stretch. The non-lexical sound here, “hra:h”, is a grunt of release, marking the stretch and relaxation of the infant body. Using touch, the mother senses the tension and release of the infant body and is displaying his sensation and co-performing the experience as it is happening. Compared to the previous examples the vocalisation here is produced in a marked manner and the articulation is exaggerated.

Fig. 3.5. Mother puts her hand under the infant’s neck and lifts head up.
3.2.1. Co-performing a stretch

Extract 3 VP01_1_1 (00:40 - 01:04)

In this example we see the mother and infant in the beginning of the nappy change activity. The mother has previously been trying to wake the infant up. After some attempts of blowing softly on his face and calling his name the infant has opened his eyes and the mother starts to undress him. In line 1 we see the mother unbuttoning the infant’s onesie saying “show me here” (Fig. 4.1). As she reaches for the infant’s leg (Fig. 4.2), the infant stretches his arms over his head. The mother responds to this with a gasp (line 2). She orients to the infant’s action as that of being a stretch “stretch yourself” and acknowledges it as being an appropriate action “exactly”, followed by a number of quickly uttered self-repetitions of “stretch” which become faster as the stretch continues and culminate with a stretching of the final syllable of the last repeat of “stret::ch”. This turn extends to accompany the entire duration of the infant’s stretch (Okada, 2021; Mondada, 2017; Keevallik, 2020). In line 4, the infant stretches and stiffens their legs towards the mother’s stomach at the same time as the mother reaches over to continue undoing the buttons of the onesie (Fig. 4.3). The mother retracts her hand and grabs both of the infant’s stretched thighs, feeling the tension in the legs (Fig. 4.4) and taking a loud inbreath (line 4); then, with a turn beginning with “and” (Keevallik, 2020), she resumes the verbal accompaniment of the infant’s stretch with two more repetitions of “stretch”. At this point the infant first bends their legs slightly towards the stomach (Fig. 4.5). The mother treats this as a signal that the infant is finishing the stretch and produces the acknowledgment “yeah” (line 5) and then, in synchrony with the infant’s release of tension, produces the vocalisation “hra::h” (line 5). This sound is uttered in the far back of the mouth, it is laryngealized, has a glottal onset, and is followed by a breathy voiced back open vowel and a voiceless epiglottal fricative. Orienting to the experience of the stretch, she continues with an assessment in the form of an exclamation “is that nice!” (line 6).
The vocalisation in the above extract mirrors the sensation of the infant at the moment in which it is happening. The form of the utterance entails a tension (via the glottal onset) and then a subsequent release (through the aspirated vowel and the fricative) which is reminiscent of the release we observed in the second part of the vocalisation in Extract 2, only in Extract 3 the voice modification is exclusively used to join in on the sensation of another person. The second part of the vocalisation is also a good example of how utterances are co-constructed moment-by-moment by the participants involved, as the mother cuts off her turn (line 5) and incorporates the ongoing action of the infant in her vocalisation. We see here an enactment of the entire feeling of the stretch: first the build-up of tension through the repetition of the verb “stretch” and then by the positioning of a vocalisation to depict the tension release of the body of the infant followed by a ratification of this through an assessment. Additionally, in this sequence the muscle tension and release of the infant’s body are also likely experienced through the tactile connection. During the stretch, the mother squeezes the infant’s legs (Fig. 4.4) and releases the squeeze after the infant stretch, embracing the infant’s lower body loosely with her hands (Fig. 4.6). It looks as if the tension is travelling through the infant’s body into hers, allowing a shared experience.

Extract 3 is different from the previous ones in that the mother’s vocalisation is a display of the tension release of the infant, rather than highlighting and foregrounding an aspect of her own action. Her body is not in tension although it should be mentioned that the tactile connection between her and the infant - through the grasp of his legs - is potentially allowing her some access to the sensation of his body. Therefore, in this case one could argue the vocalisation must be pursuing a different action. An idea put forward by Keevallik et al. (2023) is that sounding other’s sensations achieves intersubjectivity and empathy. Here the infant’s action is being noticed and oriented to. It is interesting to note here that the production of the vocalisation is exaggerated and dramatic; much more than in the previous cases, in which the mother was actually exerting physical effort. It is a performance, with the mother being an actor enacting the infant stretching.

3.3. Co-ordinating joint effort

We now turn to an example which exemplifies how the design of the mother’s effort vocalisation is modulated by the physical effort required by the infant to perform a particular movement. The effort vocalisation here accompanies the entire trajectory of a strenuous movement, both indicating that effort is needed to accomplish the action but also encouraging and supporting the infant to persevere and succeed. The example involves the mother pulling the arms of the infant so that his upper body is lifted to a seated position. This activity was quite common across many dyads at three months of age. The mothers reported doing this as a form of gymnastic training for the infant (Zukow-Goldring, 1997). In the videos it was observed that most infants at this age struggled to hold their head when transitioning from a supine to a seated position, as they need to build up strength in their neck and torso to be able to lift their head or keep their head steady. Also, it requires significant co-regulation for the dyad to achieve the pull. The pull enacted by the caregiver has to be of a specific intensity so that the infant will not be pulled by the force, and instead engage their neck and stiffen their entire body to hold steady (Fogel, 1993).

Extract 4 is an example in which the mother accompanies the infant’s bodily effort with an effort sound. As in Extract 2, the trajectory of the action is incorporated in the vocalisation. Only in this case, it is the infant’s effort, needed to keep their head steady while being lifted, that is voiced in the mother’s strained voice. When the infant is no longer able to hold the position, the vocalisation changes to a release outbreath that envelopes the action of returning back to supine position.

3.3.1. Pulling up together

Extract 4 - VP14_1 (14:33 - 17:54)

01 M: ++u::nd wieder, +(0.6s)
  ~a::nd again,
  +grabs inf hands; + slightly pulls arms up

02 I: @%(0.6s)
  @looks at mum--->
  %infant engages his body and arches neck slightly

03 M: -ha:h +(0.5s)
  m +releases pull and adjusts grip
  i @------------->

03 M: +(0.4s) bist du aber stark!% (. ) &+jha
  are you strong  (. ) yheah
  m +pulls arms  +relaxes pull
  i %some tension around the neck
  i @------------------------------------->
  i &smiles

04 M: .ha [he he he] .ha bist |du aber st?ark, .ha are you strong,
In the beginning of the extract (line 1) the mother initiates the pull-up by taking hold of the infant arms and pulling them slightly. The infant here responds to this pull by engaging his body, as seen in the stiffening of his neck. Reddy et al., (2013) observed this movement in infants of the same age in preparation of them being picked up by the parents. This could be a go-ahead for the pull-up. The mother orients to this movement in line 3 with a short vocalisation of glottal onset “-ha:h”, but this vocalisation comes after the infant movement and does not overlap with it. The form of this vocalisation here is short and similar to the vocalisation in Extract 1 which was an outbreath after straining. The pull-up does not go ahead here. This could potentially be due to a lack of pull or bodily engagement from the infant but this cannot be attested with certainty from the recording. After this attempt, the mother adjusts her grip, potentially in preparation for the next pull, and in her next turn provides an assessment orienting to how the engagement of the body, i.e., putting effort, is a sign of strength with an exclamative (line 4) “are you strong!”. This theme of being strong continues throughout lines 4 and 7. During this interval, we see a few slight initiating pulls from the mother, which do not progress to full pulls. This could again be explained by the lack of a display of effort by the child here, as he is engaging in vocal play, laughter and smiling (lines 3, 5, 6 and 7). The mother responds to these vocalisations with confirmations of the child’s vocal contributions with “exactly” while initiating the next pull (line 7). Interestingly it is while the mother is uttering the word “strong” in line 7 that the infant engages his arms and neck which becomes visible in that he manages to hold his head in one line with his body (avoiding a backwards tilt) which in effect allows him to lift his head from the mat (Fig. 5a and 5b).

The mother utters the vocalisation of interest here (line 9) while holding the lift. This starts as a vocalisation of glottal onset produced with a strained whispered voice. The production of the vocalisation here mirrors the trajectory of the infant’s effort,
as the first part of the vocalisation is uttered in a strained voice “Σ a:: Σ” (Hofstetter et al., 2021) during the infant’s effort in holding his head at its peak (Fig. 5c). At this point, the angle of flexion of the infant arms increases, the arms move into a stretched position as well and the head starts to tilt back, indicating that the infant can no longer sustain the effortful position; the vocalisation now becomes more aspirated “-ha” and the mother stops pulling and lowers the infant again on the mat (Fig. 5d). This pull-up is followed by the infant’s cheerful squeal (line 10) and the mother joins in with laughter (line 11). In line 12, the mother reformulates the vocalisation, namely as being about the infant showing strength through his effort “very strong are you” (Wiggins, 2013). This is followed by a last attempt by the mother to initiate another pull-up (lines 11 and 12), but the infant has disengaged by looking to the side and is no longer responding with body engagement to the mother’s pull.

This final extract exemplifies a further aspect of joint action and the role of sounding out with effort sounds; in joint effort everybody has to do their part. Compared to the previous extract in which the mother’s sounding out enacts or performs the infant’s experience of stretching in a rather dramatic manner, the vocalisation here is whispered. As explained above, for the pull-up to be successful both mother and infant need to co-regulate the effort they put in; the mother needs to be pulling while the child needs to be engaging their body to resist the pull. The timing of the vocalisation here indicates that the mother is acknowledging the infant’s active contribution to the pull-up and therefore their joint effort is successful. The exclamative turns preceding and following the lift are celebrating the infant’s efforts, achievements and strengths. The temporal trajectory of the vocalisation here is similar to examples of incitement in sports activities (Reynolds, 2017; Hofstetter and Keevallik, 2023); it encourages (via the strained voice) the infant to keep trying, embraces (via the aspirated vowels) the infant’s efforts and limits, but also celebrates the successful joint effort or achieving a pull-up together.

4. Discussion

This work looked at early caregiver-infant interactions and explored caregivers’– in particular mothers’– use of effort sounds in interactions with their three-month old infants during nappy changing. The analysis of multiple examples revealed the rich sociality of these overlooked routines and different ways in which effort sounds are used to structure the activity and bring caregiver and infant together.

The analysed vocalisations had a glottal onset and/or were produced with laryngeal constriction and used by the caregivers to sound out actions they were performing on the infant body and the sensations that the infants are presumably experiencing at the same time. It was often the case (Extracts 1, 3 and 4) that vocalisations would initially have a glottal onset, or some laryngeal constriction (such as a creaky or strained voice) but then would develop in aspirated back vowels (Keevallik and Ogden, 2020) capturing the bodily effort and tension release of the participants. This can be explained by the fact that they accompanied longer actions and changed along with the action trajectory, setting them apart from the strict definitions of grunts and other strain sounds which are defined as short constricted vocalisations (Bordenave and McCune, 2021). A unifying feature of these vocalisations was that they were synchronous with the ongoing physical action and even prolonged to match the duration of an action. Their duration and intensity matched the actions they were synchronous to and the sensations they were orienting to. This work contributes to a research tradition within interactional linguistics on the ways in which language organises social action and connects not only people’s intentions but is also necessary for coordinating physical activities (Hofstetter and Keevallik, 2023). The effort vocalisations were positioned alongside the trajectories of physical action, at their peak or release, foregrounding some aspects of the ostensible sensory experiences of both caregiver and infant. This is particularly important in early routines, as vocalisations may contribute to infants’ sense of what is happening to their bodies in real time. They may also enable them to be active participants from early on.

In the data analysed, the vocalisations in question were observed as part of sequences; they were often preceded by turns, broadcasting the upcoming action steps, and followed by elaborations or explanations of what they were achieving (Wiggins, 2013). Furthermore, capitalising on the flexibility of their phonetic design they could also be prosodically modulated to incorporate trajectories of actions as well as the actions of both mother and infant in one vocalisation. On the one hand, this gave the impression of a repeated pattern, a multimodal gestalt (Mondada, 2014) that could become a recognisable pattern for the infant. On the other hand, while some vocalisations would be recognisable, others sounded as if they were improvised – fabricated in the here-and-now. It seems that they borrow their meaning from the interactional moments in which they occur and cannot exist beyond these moments.

As Dingemanse (2020, p. 188) writes, “For talk-in-interaction to work as smoothly as it does, perhaps some aspects of it must seem like not-talk.” The semantic underspecification (Keevallik and Ogden, 2020) of the vocalisations analysed in this paper might open up a broader space of potential meanings (Ogden, 2020; Reber, 2012) and therefore more possibilities for people to create meanings together. When uttered they are contextualised via temporality, phonetic variation and multisensoriality. This suggests that we have at our disposal flexible tools which we can use to creatively go beyond the linear transmission of information.

Furthermore, the analysis showed that vocalisations which would have traditionally been considered to be emotional spillages or side effects of physical strain and individual effort were part of structured activities (Pehkonen, 2020), and were finely timed and coordinated to capture the rhythms and dynamics of movements of both participants and foregrounded physical activity and its effects as something public, something that can be communicated about (Goffman, 1978; Keevallik, 2023). The moment-to-moment coordination of actions originating from multiple sensory modalities and held together by vocal practices created a togetherness uniting caregiver and infant into a joint experience.
For the developmental perspective, it is important that the vocalisations used involve phonetic patterns that the infants are already capable of producing. Laryngeal vocalisations belong developmentally to the earliest infant vocalisations (Buder et al., 2013). They are considered important in the process of development of vocal learning, and they have been suggested to also have social functions for young infants (McCune, 2021). Accordingly, they may be selected as communicative contributions that the infants are capable of reciprocating. This idea goes back to Papousek and Papousek (1989) and their study of parents’ vocal matching of infant vocalisations. They proposed that by providing vocalisations in a form in which the infants themselves are capable of producing, parents are offering infants the opportunity to provide adequate responses. Vocalisations that are produced as a result of body movement could be particularly accessible to infants (Gratier, 2019).

Capitalising on the proximal and intimate context of the activity dominated by constant touch, these vocalisations are also born out of intercorporeal engagements (Meyer et al., 2017) as caregiver and infant are actors and receivers of actions at the same time. The sounds produced by the caregivers in the data were expressions not only of their own bodily sensations but also those of the infant, crossing the boundaries of individual bodies and achieving “compresence” (Merleau-Ponty, 1964). In doing so, the complex task of managing an infant body and performing a complicated acrobatic task became something collaborative, a joint effort in which the infant was acknowledged as a participant, their sensations recognised and incorporated in meaningful sequences, their effort was encouraged and mishaps highlighted and resolved.

These carefully timed non-lexical vocalisations, I propose, create opportunities of co-experiencing and being attuned, making actions and sensations accessible to each other and enable mutually accessible meaning. The vocal behaviours enveloping those joint experiences become moments in which the infant is invited to engage with language (Rączaszek-Leonardi et al., 2018, 2022) and enter into patterns of participation that will develop into “linguistic modes of existence” (Di Paolo et al., 2018, p.260).

**Declaration of competing interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Acknowledgements**

The data were collected as part of my doctoral thesis in the project “Symbiosis of language and action” within the Volkswagen Foundation Dilthey Fellowship to Katharina Rohlfing. I would like to thank the families who invited me into their homes and agreed to being recorded. Special thanks to Vasu Reddy, Mark Dingemanse and Alessandra Fasulo for their encouraging and insightful feedback on different versions of this paper. I am indebted to the two anonymous reviewers and editors, Leelo Keevalik and Emily Hofstetter for understanding what I wanted to say and providing such detailed and constructive feedback.

**References**


