EMBODIED COMMUNICATION: USING A NOVEL MACHINE LEARNING METHOD TO IDENTIFY INFANT-CAREGIVER INTERACTIONAL EVENTS FROM NATURALISTIC AUDIO RECORDINGS

AUTHORS

Anna Mercier*[1 2], Sydney Segal*[1 2], Silver Liftin[1 2], Nathan Simon[1 2], Alexandra Kahn[1 2], Sylvie Shrum[1 2], Sam Melin[3], Eva Piras[1 2], Ruthe Foushee[1 2 4 5]

AFFILIATIONS

[1]The Human Language and Development Laboratory, [2]The New School for Social Research, [3]Ferkauf Graduate School of Psychology, [4]University of Chicago

RESEARCH QUESTIONS

RQ1: Can we use machine learning to identify meaningful interactional events from noisy naturalistic audio streams?

RQ2: Can this approach reliably capture the daily choreography of breastfeeding, infant vocalizations, and sleep-wake patterns from continuous daylong audio recordings?

BACKGROUND

- Healthy infant development depends on responsive, attuned caregiving.[1]
- Universal caregiver-infant interactional behaviors include gaze, vocalization, and touch.[2 3]
- Current conceptualization of caregiver-infant contingency largely draws from Western caregiving contexts emphasizing face-to-face interaction.[4 5]
- Western research links breastfeeding with caregiver–infant responsiveness and bonding, but beyond its biological function, its role in attachment formation across cultures remains unclear.[6 7]
- Existing measures assessing skin-to-skin contact and synchrony are not suited to baby-wearing contexts, where contact is hidden, gaze is minimal, and communication is embodied.[8 9]
- We use a human-in-the-loop machine learning pipeline developed for long-form audio recordings: humans code validated event categories in video-accompanied audio segments, which serve as training data for a classifier detecting events in daylong audiorecordings.[10]

We use audio-based machine learning to capture meaningful caregiver-infant interactions in a Tseltal Maya context, where infants spend most of the day worn on their mothers' backs.

METHODOLOGY

Participants: 68 dyads from Chiapas, Mexico
- **Mothers**: 15-39yrs, M=25.27, SD=7.19
- **Infants**: 0-14mos, M=7.44, SD=3.66

Figure 1. *Tseltal mother carrying infant on her back.*

Figure 2. *Naturalistic audio used in analyses came from high-quality audio-recorders worn by infants in the pockets of custom vests.*

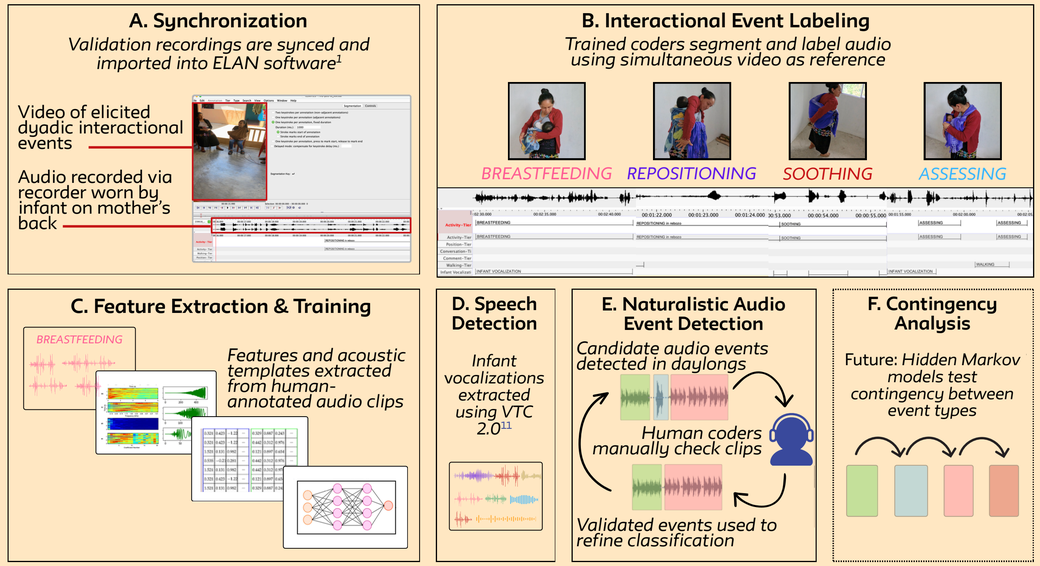

Figure 3. *Method training classifiers on hand-labeled audio clips of caregiver-child interaction to segment and label dyadic interactional events in noisy, naturalistic longform audiorecordings, enabling analyses of otherwise-invisible nonverbal communication between infants and caregivers.*

PRELIMINARY RESULTS

RQ1: From 26 daylong recordings with no visual or motion data, the model detected *some* meaningful interactional events.

EVENT CLASS	*n*	dur (s)	Top-1 Accuracy	Top-2 Accuracy
SOOTHING	75	816.17	85.33%	97.33%
BREASTFEEDING	53	1452.49	50.94%	90.57%
VOCALIZATIONS[11]	172	809.00	85.5%	N/A

Table 1. *Model Evaluation: Sample Interactional Event Classes and Initial Classification Accuracy*

Of the coded-for meaningful interactional events, **breastfeeding** had the clearest acoustic signature.

Its **rhymicity resembles sleeping sounds**, and the classifier picked up these bouts as breastfeeding.

Human coders **further trained the classifier** to correctly identify sleeping vs. breastfeeding events.

Vocalizations were salient around events. VTC 2.0[11] was used to identify infant vocalizations.

Preliminary analyses were thus conducted with **breastfeeding**, **infant vocalizations**, and **sleeping** events.

EVENT CLASS	SEGMENTS
BREASTFEEDING	43
SLEEPING	89
INFANT VOCALIZATION	62,914[11]

Table 2. *Segments identified by the trained classifier and VTC 2.0.*[11]

RQ2: The data revealed structured daily choreography linking nighttime infant vocalization, breastfeeding, and sleep patterns.

INFANT VOCALIZATION → ~ 5 mins → BREASTFEEDING → < 10 mins → SLEEPING

Infant Vocalization x Breastfeeding

Infant vocalization higher in 5 mins before feeding compared to after in 60% of bouts.

Effect driven by nighttime feedings: 69% showed decreases (median 1.0 → 0.4 sec/min, p = .064).

Daytime effects were near chance, suggesting this is nighttime specific.

Sleeping x Breastfeeding

39.5% of breastfeeding bouts occurred within 10 mins of a sleep period.

Repeated feed-sleep-feed cycles were especially evident at night.

Near-sleep bouts were longer than other bouts (median 8.4 min vs. 3.8 min).

DISCUSSION

RQ1: Preliminary support for feasibility of identifying meaningful interactional events using audio-only data.

- Detection was intentionally conservative, prioritizing accuracy over capturing every occurrence.
- Lack of detected breastfeeding bouts for eight participants likely reflects shorter recordings or insufficient template matches rather than true absence of feeding.
- This provides a **novel method to extend understanding of developmental conditions and interactive behavior beyond face-to-face contexts.**

RQ2: Pilot data suggests breastfeeding functions as a transitional regulatory mechanism around sleep.

- Temporal proximity of breastfeeding and sleep, evidenced by repeated **nighttime feed-sleep-feed cycles** and **longer near-sleep feedings.**
- Nighttime decrease in infant vocalization following breastfeeding is consistent with the following pattern: **hunger cue → feeding → satiation**; suggesting coordinated regulation of infant state.
- Absence of a strong daytime effect suggests this regulatory choreography is most structured at night.

FUTURE DIRECTIONS

1. Infant vocalization as precipitant for breastfeeding: motivation and implications for language development in contexts without infant-directed speech.

2. More focus on embodied practices: insights into infant-caregiver attunement beyond psychological aspects centered in Attachment Theory.[6 12]

ACKNOWLEDGEMENTS

Thanks to Humbertina Gómez Pérez, Antonieta López Santiz, Petrona Girón López, Marisa Casillas, Dan Byrne, & Sarah Goodman.

REFERENCES

1. Ainsworth, M. S. (1979). Infant–mother attachment. *American Psychologist*.
2. Cameron-Faulkner, T., Malik, N., Steele, C., Coretta, S., Serratrice, L. and Lieven, E. (2021). A cross-cultural analysis of early prelinguistic gesture development and its relationship to language development. *Child Development*.
3. Hernik, M., & Broesch, T. (2019). Infant gaze following depends on communicative signals: An eye-tracking study of 5- to 7-month-olds in Vanuatu. *Developmental Science*.
4. Murray, L., De Pascalis, L., Bozicevic, L., Hawkins, L., Sclafani, V., & Ferrari, P. F. (2016). The functional architecture of mother-infant communication, and the development of infant social expressiveness in the first two months. *Scientific Reports*.
5. Beebe, B., & Steele, M. (2013). How does microanalysis of mother-infant communication inform maternal sensitivity and infant attachment? *Attachment & Human Development*.
6. Funke, L., Scheidecker, G., Chopin, B., Schmidt, W., El Ouardani, C., Chaudhary, N. (2023). Feeding, bonding, and the formation of social relationships. *Ethnographic Challenges to Attachment Theory and Early Childhood Interventions*.
7. Pearson, R. M., Lightman, S. L., & Evans, J. (2011). The impact of breastfeeding on mothers' attentional sensitivity towards infant distress. *Infant Behavior and Development*.
8. Feldman, R. (2012). Physiological measures of emotion from a developmental perspective: State of the science: Parent–infant synchrony: A biobehavioral model of mutual influences in the formation of affiliative bonds. *Monographs of the Society for Research in Child Development*.
9. Hertenstein, M. J. (2002). Touch: Its communicative functions in infancy. *Human Development*.
10. Burger, T., Marxer, R., & Barker, J. (2012). Manual annotation of environmental noise in audio streams. *Proceedings of Interspeech*.
11. Charlot, T., Kunze, T., Poli, M., Cristia, A., Dupoux, E., & Lavechin, M. (2025). BabyHuBERT: Multilingual self-supervised learning for segmenting speakers in child-centered long-form recordings. *arXiv:2509.15001*.
12. Reddy, V., Markova, G., & Wallot, S. (2013). Anticipatory adjustments to being picked up in infancy. *Public Library of Science*.