

# HIDDEN COMMUNICATION: A MACHINE LEARNING APPROACH TO ANALYZING CAREGIVER-INFANT CONTINGENCY IN BABYWEARING CONTEXTS

## AUTHORS

Sydney Segal\*<sup>1</sup>, Anna Mercier\*<sup>1</sup>, Silver Liftin<sup>1</sup>, Nathan Simon<sup>1</sup>, Alexandra Kahn<sup>1</sup>, Sylvie Shrum<sup>1</sup>, Sam Melin<sup>2</sup>, Eva Piras<sup>1</sup>, Ruthe Foushee<sup>1,3</sup>

## AFFILIATIONS

<sup>1</sup>The New School for Social Research, <sup>2</sup>Ferkauf Graduate School of Psychology, <sup>3</sup>University of Chicago

## RESEARCH QUESTION

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

## BACKGROUND

- Healthy infant development depends on responsive, attuned caregiving.<sup>1</sup>
- Universal caregiver-infant interactional behaviors include gaze, vocalization, and touch.<sup>2,3</sup>
- Current conceptualization of caregiver-infant contingency largely draws from Western caregiving contexts emphasizing face-to-face interaction.<sup>4,5</sup>
- 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.<sup>6,7</sup>
- 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.<sup>9</sup>

## PARTICIPANTS

From Tzeltal-speaking households in Chiapas, Mexico, where infants are carried in rebozos until they can walk and talk.



Table 1. Participants.

	Validation Recordings	Daylong Recordings
N	33	68
Mothers' Age	16-40yrs M=25.16, SD=7.05	15-39yrs M=25.27, SD=7.19
Infants' Age	0-14mos M=7.09, SD=4.29	0-14mos M=7.44, SD=3.66

Figure 1. Tzeltal mother carrying infant on her back.

## METHODOLOGY

### 1. DAYLONG RECORDINGS & INTERVIEWS

#### 1A. Recording the Day

Recording devices are worn by infants in the pockets of custom vests inside the rebozo (sling in which the infant is carried); recording a naturalistic audio stream over a 24-hour period then used in analyses.



Figure 2. Recording devices inside the rebozos.

#### 1B. Interviews with Mothers

Interviews conducted with mothers are analyzed to iteratively determine meaningful interactional events (MIE) occurring within the dyad throughout the day.

### 2. VALIDATION RECORDINGS & CODING PROCESS

#### 2A. Validation Recordings

Collecting validation recordings of elicited dyadic Meaningful Interactional Events (MIEs) provides a gold standard against which model predictions are evaluated.



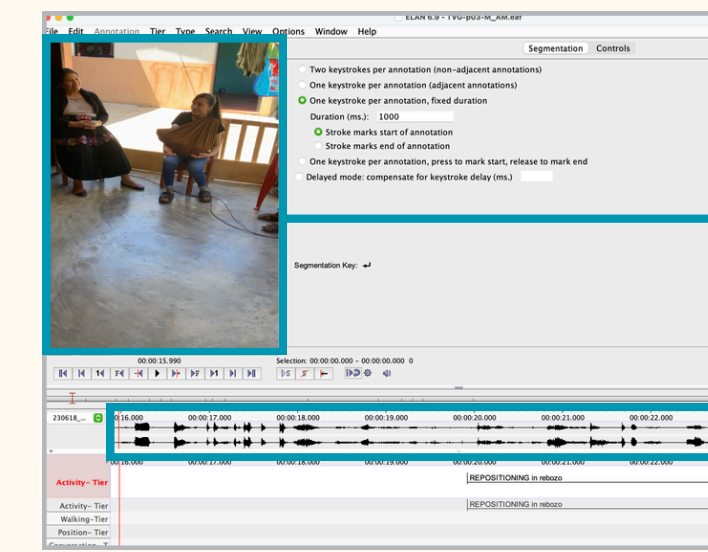
Figure 3. Entering the rebozo: Process of the caregiver placing the infant in the rebozo.

Table 2. Meaningful interactional event (MIE) categories.

IDENTIFIED MIEs
BREASTFEEDING
SOOTHING
ASSESSING
EATING
REMOVING FROM REBOZO
ENTERING THE REBOZO
REPOSITIONING IN REBOZO
WALKING
GREETING
LAYING DOWN

#### 2B. Interactional Event Labeling

Validation recordings are imported into ELAN software.<sup>10</sup> Trained coders segment and label audio clips using simultaneous video as reference.



Video of elicited dyadic interactional events  
 Audio recorded via recorder worn by infant on mother's back

Figure 4. ELAN coding process.

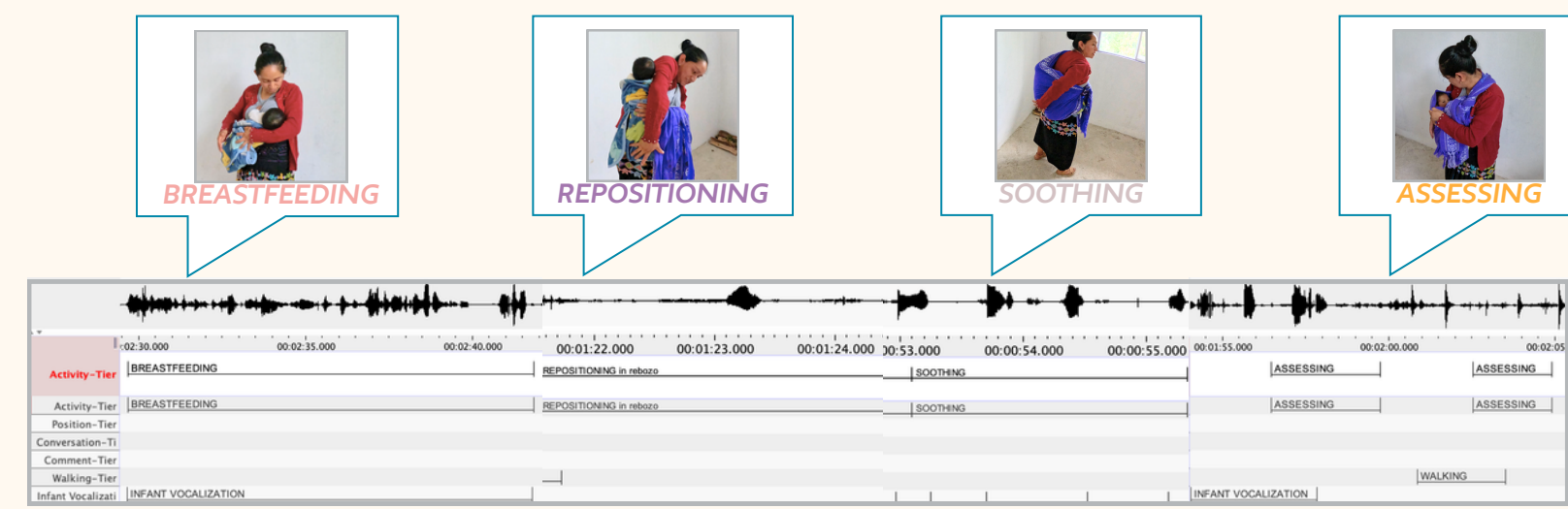
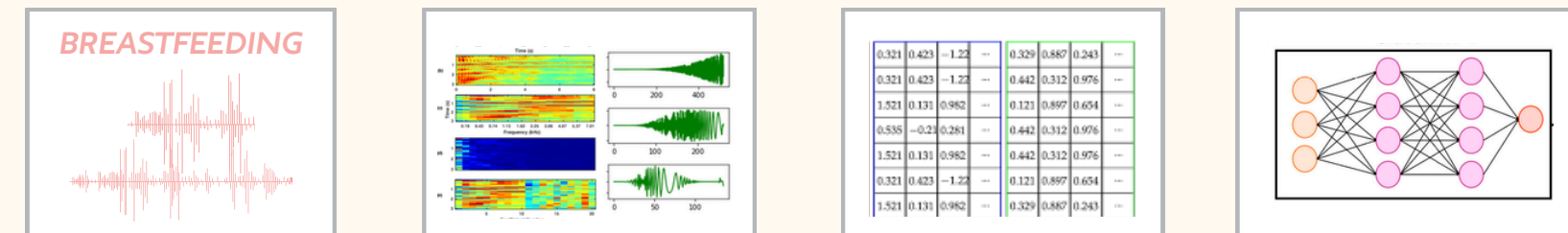


Figure 5. Labeling process.

### 3. MACHINE LEARNING PROCESS

#### 3A. Feature Extraction & Training

Features and acoustic templates extracted from 70% of human-annotated audio clips.

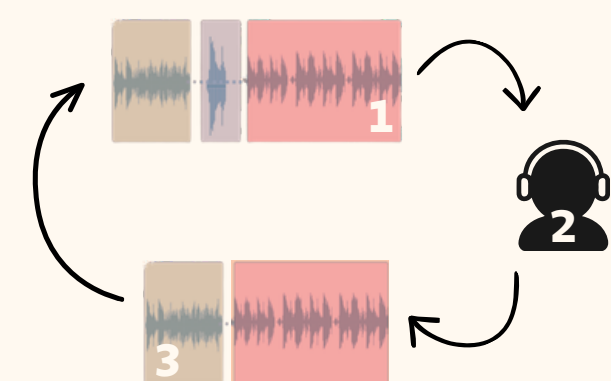


#### 3B. Model Evaluation

Classification accuracy calculated on test set (remaining 30%).



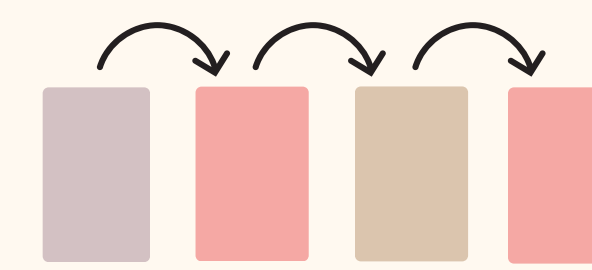
#### 3C. Naturalistic Audio Event Detection (Human in the Loop)



- Candidate audio events detected in daylongs.
- Human coders manually check clips.
- Validated events used to refine classification.

#### 3D. NEXT STEP: Contingency Analysis

Hidden Markov models test contingency between event types.



## INITIAL OBSERVATIONS

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

- Certain MIEs have more readily identifiable acoustic signatures (e.g., Soothing and Breastfeeding), resulting in better detection (Table 3).
- The classifier continues to be refined to better detect MIEs with rustling in the acoustic signature, and improve its detection of identifiable MIEs.
- This nevertheless provides a novel method to extend understanding of developmental conditions and interactive behavior beyond face-to-face contexts.

Table 3. Model evaluation: Sample interactional event classes and initial classification accuracy.

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%

## FUTURE DIRECTIONS

- Potential avenues from training the model**  
 Further developmentally meaningful structure may be derived from acoustic data alone (e.g., choreography of events and co-occurrences).
- More focus on embodied practices**  
 Insights into infant-caregiver attunement beyond psychological aspects centered in Attachment Theory.<sup>11,12</sup>
  - Co-occurrence of breastfeeding and soothing suggests feeding may serve an attachment and regulating function.<sup>12</sup>

## ACKNOWLEDGEMENTS

Thanks to Humbertina Gómez Pérez, Antonieta López Santiz, Petrona Girón López, Marisa Casillas, Dan Byrne, Sarah Goodman, & The Center for Gesture, Sign, and Language at The University of Chicago for funding.

## REFERENCES

- Ainsworth, M. S. (1979). Infant-mother attachment. *American Psychologist*.
- Cameron-Faulkner, T., Malik, N., Steele, C., Corsetto, S., Serratrice, L. and Lieven, E. (2021). A cross-cultural analysis of early prelinguistic gesture development and its relationship to language development. *Child Development*.
- Hermik, 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*.
- Murray, L., De Pasqualis, L., Borcic, L., Hawkins, L., Scalfani, 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*.
- Beebe, R., & Steele, M. (2013). How does microanalysis of mother-infant communication inform maternal sensitivity and infant attachment? *Attachment & Human Development*.
- 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*.
- Hertenstein, M. J. (2002). Touch: Its communicative functions in infancy. *Human Development*.
- Burger, T., Mauer, R., & Barker, J. (2012). Manual annotation of environmental noise in audio streams. *Proceedings of Interspeech*.
- ELAN (Version 6.9) Max Planck Institute for Psycholinguistics. <https://archive.mpi.nl/tilburg>
- Reddy, V., Markova, G., & Wallat, S. (2013). Anticipatory adjustments to being picked up in infancy. *Public Library of Science*.
- Funk, L., Scheidecker, G., Chapin, B., Schmidt, W., El Quasbi, C., Chaudhry, N. (2023). Feeding, bonding, and the formation of social relationships. *Ethnographic Challenges to Attachment Theory and Early Childhood Interventions*.
- Harlow, H. F. (1958). The nature of love. *American Psychologist*, 13(12), 673-685.

