Title Page (with Author Details)

The Role of Mother-Infant Emotional Synchrony in Speech Processing in 9-month-old Infants

Monica Vanoncini*[1,2,3], Natalie Boll-Avetisyan[2], Birgit Elsner[1], Stefanie Hoehl[3], and Ezgi Kayhan[1,4]

Author Affiliations:

[1]Department of Developmental Psychology, University of Potsdam, Germany

[2]Department of Linguistics, University of Potsdam, Germany

[3]Department of Developmental and Educational Psychology, University of Vienna, Austria

[4]Max Planck Institute for Human Cognitive and Brain Sciences. Leipzig, Germany

*Corresponding author: Monica Vanoncini

Universität Potsdam

Campus Golm, Haus 14, 1.10

Karl-Liebknecht-Str. 24-25

14476 Potsdam, Germany

vanoncini@uni-potsdam.de

Policy and ethics: The research was conducted in accordance with APA ethical standards in the treatment of the study sample

Declarations of Interest

None

Acknowledgments

This work was made possible by the Research Focus Cognitive Sciences (RFCS) of the University of Potsdam and by the DFG (Grant number: 402789467). The authors would like to thank mothers and infants who participated in the study, and the Department of Obstetrics and Gynecology of the Vienna General Hospital for supporting our participant recruitment. We acknowledge Barbara Höhle and colleagues for the kind permission to use the stimuli of their study (Bartels et al., 2009). We thank Jessica Gärtner for sharing the R-script on CRQA with us. The authors are grateful to Markus Tünte, Liesbeth Forsthuber and all the research assistants (namely Asena Boyadzhieva, Johannes Bullinger, Sandra Gaisbacher, Lara Iven, Laura Neumann, Julia Otter, Lisa Triebenbacher, Jakob Weickmann, Felicia Wittmann, Moritz Wunderwald) for their help in setting up the experiment, data collection, and coding.

Published version:

This article was accepted for publication on 29th August 2022 in Infant Behavior and Development as part of special issue "The Interplay between Cognition and Emotion in Infancy and Toddlerhood" (available at https://doi.org/10.1016/j.infbeh.2022.101772).

Highlights

- Entropy is associated with infants’ word segmentation performance.
- We used cross-recurrence quantification analysis to quantify emotional synchrony.
- Entropy is a promising measure of mother-infant emotional dynamics.

Abstract

Rhythmicity characterizes both interpersonal synchrony and spoken language. Emotions and language are forms of interpersonal communication, which interact with each other throughout development. We investigated whether and how emotional synchrony between mothers and their 9-month-old infants relates to infants' word segmentation as an early marker of language development. Twenty-six 9-month-old infants and their German-speaking mothers took part in the study. To measure emotional synchrony, we coded positive, neutral and negative emotional expressions of the mothers and their infants during a free play session. We then calculated the degree to which the mothers' and their infants' matching emotional expressions followed a predictable pattern. To measure word segmentation, we familiarized infants with auditory text passages and tested how long they looked at the screen while listening to familiar versus novel words. We found that higher levels of predictability (i.e. low entropy) during mother-infant interaction is associated with infants' word segmentation performance. These findings suggest that individual differences in word segmentation relate to the complexity and predictability of emotional expressions during mother-infant interactions.

Keywords: mother-infant dyads, entropy, emotional synchrony, cross-recurrence quantification analysis, word segmentation, rhythmicity

1. Introduction

Before infants start saying their first words, they communicate their emotions and needs through facial expressions, voice, posture and gestures. These signals help their caregivers to understand whether their children are hungry or tired, or whether they are bored or excited. When their language skills advance, children progressively learn to communicate their feelings and internal states through words and sentences. Language and emotion, as communicative systems, have been argued to be acquired through similar learning processes (Ruba et al., 2022). Adults typically do not provide instructions about how language or emotion are used to communicate. When dealing with other humans, infants mainly learn to understand language and emotion by tracking regularities in sequences of syllables or phonemes in the case of language (Pelucchi et al., 2009; Romberg & Saffran, 2010), and in sequences of facial expressions of emotions (Mermier et al., 2022), respectively.

Previous studies have shown that language abilities play a role in children's emotional competence by making emotions explicit and communicable (Cole et al., 2009; Ornaghi et al., 2017, 2019). Moreover, delay in language development is associated with lower emotion recognition skills (Nelson et al., 2011). Less is known about whether certain attributes of infants' emotional experiences positively relate to language development. It has been proposed that sharing congruent emotions might provide an intersubjective foundation or a first medium of communication, which enhances infants' motivation to use language to communicate (Nicely et al., 1999). Notably, meaning is built through co-regulated coupling of behavior, including emotional expressions and vocalizations between two individuals (Hoehl & Bertenthal, 2021). To understand the role of emotions in language development, the present study investigated whether early emotional interaction dynamics between mothers and their infants predict infants' language processing abilities at 9 months of age.

1.1 Rhythm as the foundation of early interactions

Human beings establish their first interactions with their caregivers. Caregiver-infant exchanges follow a clear rhythmic structure in which sequences of behaviors are repeatedly carried out (Cohn & Tronick, 1987). This rhythmic structure, characterized by temporal regularity that caregivers and infants experience in the interactions across their daily routine, leads to interpersonal synchrony (Markova et al., 2019). Interpersonal synchrony has been defined as the temporal concordance of behavioral, neural and/or physiological states (Feldman et al., 2011; Markova et al., 2019). Synchrony allows the interaction partners to anticipate each other's internal states and actions enhancing bonding, prosocial behavior and positive affect (Feldman, 2012; Feldman et al., 1999; Mogan et al., 2017). Nonetheless, typical interactions of mothers and their infants are characterized by periods of mismatching and dyssynchrony, with intermittent phases of interactional "repair", and neutral affect (Tronick, 2017; Tronick & Gianino, 1986). Mismatching states can serve as a catalyst for interaction partners to adjust their behaviors, and establish a synchronous coordination. Finally, the temporal coordination between action and language as a form of maternal responsiveness is linked to children's later language development (Rohlfing & Nomikou, 2014).

Humans are presumed to be the only species where early social play relies on the matching of socio-affective expressions (Feldman et al., 2011). Mutual regulation of affect greatly contributes to self-regulation of children (Feldman et al., 1999). Emotions affect the alignment between individuals at behavioral, physiological and neural levels (Hoehl et al., 2021). The reasons for this may be bidirectional: On the one hand, facial expressions help to predict whether the interaction partner might act in a certain way. On the other hand, nonverbal synchrony (i.e., body movements) is associated with both higher positive affect and lower negative affect (Tschacher et al., 2014). In line with this, being in a happy mood has been found to enhance behavioral alignment, whereas a sad mood impedes alignment (Likowski et al., 2011). Importantly regarding infants' development, research has shown that emotions influence the alignment between mothers and infants: During moments of positive mutual affect and positive vocal synchrony, the synchrony between the mothers' and their infants' heart rhythms increases (Feldman et al., 2011). Moreover, maternal matching responses to 9-month-olds' affective expressions predict the timing of children's first word comprehension and production (Nicely et al., 1999). Building up on these studies, we examined whether the temporal dynamics of emotional expressions between mothers and their 9-month-old infants are associated with infants' language development at the same age.

1.2 Quantification of early interaction dynamics

Cross-recurrence quantification analysis (CRQA) is an ever growing technique to examine the degree to which two or more time series exhibit a similar pattern (Coco & Dale, 2014). While aggregative methods assume that behaviors are stationary across time, CRQA, as a nonlinear time-series method, captures the changes in behavioral time series (Xu et al., 2020). CRQA provides several outcome measures of dynamic organization. The simplest one is called Recurrence Rate (RR). RR quantifies the percentage of time points in synchrony (e.g., matching emotional states) in a two-dimensional matrix of time. A more complex metric is Entropy (ENTR). The concept of entropy, originally derived from thermodynamics, has been applied to both linguistics and social interactions. Entropy is a measure of disorder, and as such, it quantifies the degree of uncertainty of a distribution as a whole. Here is an example. We have three small boxes. Box A contains four apples; box B contains three apples and a banana; box C contains two apples and two bananas. In this case the degree of entropy will depend on the probability to grab the fruit you guessed. We know for certain that box A will give you an apple (i.e., low entropy). It is highly likely that you can pick an apple from box B (i.e., middle entropy). However, it is equally likely to grab an apple or a banana from box C (i.e., high entropy).

In language, the degree of entropy can, for example, vary depending on the number of words and frequency of words. That is, higher levels of entropy (i.e., high uncertainty) are characterized by a higher number of words in the input (e.g., dog, ball, tree, etc.), appearing at equal number of times. Lower levels of entropy (i.e., high predictability) are characterized by a lower number of words (e.g., tree, elderflower), having different levels of frequency (in this case "tree" is a high-frequency word, "elderflower" is a low-frequency word) (Radulescu et al., 2019). Remarkably, entropy has an impact on language learning: Children learn words better (i.e., word segmentation, object-label pairing) in low entropy conditions (Lavi-Rotbain & Arnon, 2019). Parents naturally simplify their speech to their young children. Entropy, as a measure of the degree of uncertainty in the language input, has been found to be lower when parents talk with younger infants (Tal et al., 2021). Only after having learned some words, infants can learn their native language's syntax (Saffran & Wilson, 2003). Spontaneously, parents increase the entropy rate of their speech when infants get more proficient in their language development (Tal et al., 2021).

In social interactions, the higher the complexity and uncertainty of the dyadic exchanges, the higher will be the entropy (Guevara et al., 2017). For example, entropy will be lower if a dynamic system revisits similar interval lengths of emotional states in a stable manner; whereas dynamic systems following fluctuating patterns of emotional states across time will have higher entropy (see § 2.4, Figures 2 and 3).

Effects of entropy in emotional synchrony have rarely been studied. However, a recent study has found that maternal mood regulation, as measured by mood entropy, is linked to the child's neurodevelopmental outcome. More specifically, higher maternal mood entropy was associated with lower cognitive development scores at 2 years of age and lower expressive language scores at 6–9 years of age (Howland et al., 2021). In the present study, we will assess two different aspects of emotional synchrony between mothers and their infants: 1) by examining the percentage of time points in emotional synchrony in a two-dimensional matrix of time (i.e., RR) and 2) by determining the degree of uncertainty of the interaction or predictability of emotional synchrony (i.e., ENTR).

1.3 Rhythm as the foundation of language development

A fundamental challenge in language development is that infants need to discover what the words are in the language they learn, which are embedded in a continuous speech signal that does not contain any clear-cut cues to word boundaries. Interestingly, in order to find words, infants make use of rhythmic information (Jusczyk, 2000; Jusczyk et al., 1999). Rhythm, which characterizes spoken language (Allen, 1975), is essential in perceiving, comprehending and producing speech. Language learners need to perceive the temporal organization of phonemes, syllables, words and phrases from an ongoing speech stream (Fujii & Wan, 2014). Therefore, word segmentation represents a milestone in language acquisition. If sensitivity to rhythmicity in language supports word segmentation, we reasoned that word segmentation may be linked to the dynamics of mother-infant interaction, which are organized through a rhythmic structure.

Infants exposed to a language with stress-timed rhythm characterized by many words beginning with a strong syllable (e.g., kingdom vs. guitar) can segment words on the basis of strong syllables (e.g., Dutch-learning infants at 9 months of age, see Houston et al., 2000; American English-learning infants at 7.5 months of age, see Jusczyk et al., 1999, German-learning infants at 9 months of age, see Bartels et al., 2009; Zahner et al., 2016). To sum up, by the second half of the 1st year, infants use the rhythmicity features of their native language to segment words from continuous speech, and to start building vocabulary. In the case of German, which is a stress-time language, infants segment words on the basis of strong syllables. We, therefore, assess the word segmentation performance of infants exposed to German using trochaic words as targets.

In order to test infants' speech segmentation ability, studies classically employ paradigms that use looking time as measurement such as the Head-Turn Preference Procedure (HPP; Fernald, 1985), or the Central Fixation Procedure (CF; Cooper & Aslin, 1990). Infants are familiarized with text passages containing one or more target words (i.e., which will become the familiar words). They are then tested on isolated novel and familiar words. Infants' looking time preferences for one type of stimulus (i.e., novel vs. familiar) are assumed to reflect infants' ability to segment target words during the familiarization phase. Whereas some studies indicate that infants show a familiarity preference (e.g., Bartels et al., 2009; Newman et al., 2006; Ota & Skarabela, 2018), others have shown that infants have a novelty preference (e.g., Zahner et al., 2016). In general, the differences in terms of novelty and familiarity findings might depend on three primary factors: Stimulus complexity, familiarization time (Hunter et al., 1983), and developmental stage (DePaolis et al., 2016). First, overly simple stimuli (i.e., highly predictable) might lead to a novelty preference, and overly complex stimuli (i.e., lowly predictable) might lead to a familiarity preference. Second, infants tend to show a familiarity preference if the familiarization phase is short, and vice versa. Third, lexically advanced infants might exhibit a novelty preference. However, an absence of looking preference does not mean an absence of discrimination (Aslin, 2007).

1.4 The current study: Research question and hypothesis

Previous studies have shown that rhythm is an important feature of language acquisition and social interaction. This raises the question of whether synchronizing emotions with caregivers (i.e., emotional synchrony) is related to infant's language development. We aimed to examine whether and how two aspects mother-infant emotional synchrony during a free play interaction predicts word segmentation ability of 9-month-old infants. We hypothesized that higher levels of emotional synchrony (i.e., higher RR) and higher levels of predictability of the length of the intervals of emotional synchrony (i.e., lower Entropy) are linked to better word segmentation at 9 months of age.

2. Methods

2.1 Participants

A total of 49 infants and their mothers participated in the study. Participants were recruited by contacting parents who previously expressed their interest in participating in developmental research. Families received travel cost reimbursement and either a 15€ voucher or a children's book for their participation. This study was approved by the Ethics Committee of the University of Vienna.

The final sample consisted of 26 9-month-old infants (*M*=9 months and 24 days, *SD*=11.57 days, 12 females) and their German-speaking mothers (*M*=32.65 years, *SD*=5.18). Data from 23 additional dyads were excluded due to not performing one or both tasks (*n*=7), emotional expressions not being codable due to sucking a pacifier (*n*=1), technical issues (*n*=6), or fussiness (*n*=3). Finally, we excluded infants (*n*=6) who did not provide a minimum of six test trials (i.e., three in each condition), and discarded trials with looking times shorter than 1 second in the word segmentation task (cf. Junge et al., 2020). Finally, this study was part of a bigger study. Each testing consisted of six paradigms with an overall duration of about three hours including breaks. Although the order of tasks was counterbalanced across dyads, this might have also contributed to the relatively higher data loss here in comparison to other studies.

Among the included dyads, regarding the mothers' education, *n*=6 had at least one university degree, *n*=5 completed technical or high school, *n*=4 obtained vocational training, and *n*=3 did not fill in the questionnaire. All children were predominantly exposed to German. Infants were born full term (min. 36 weeks of gestation) healthy (10-min APGAR Score > 9). Four mothers did not report the APGAR scores for their children.

2.2 Materials and design

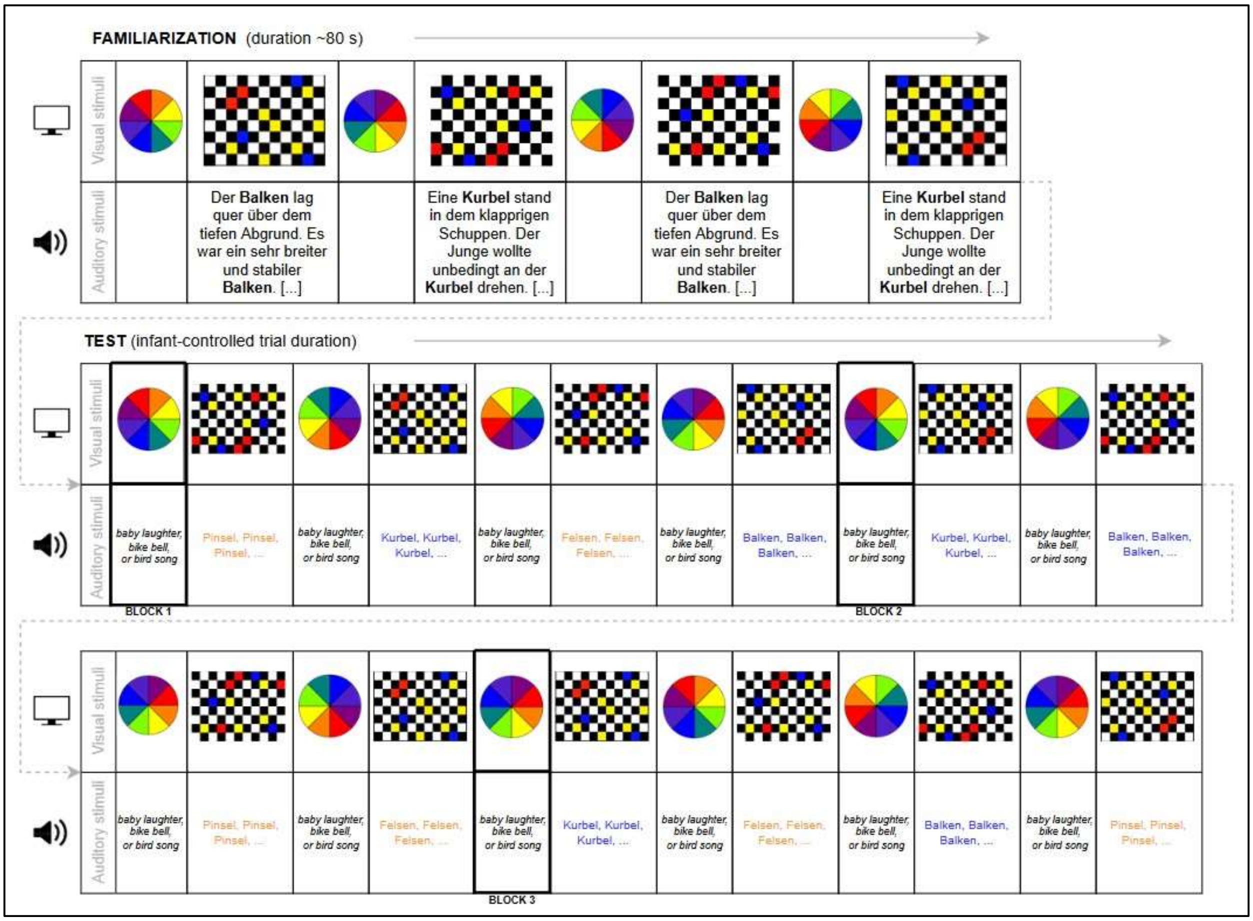
As illustrated in Figure 1, infants' word segmentation ability was examined with a central fixation paradigm (Cooper & Aslin, 1990). Looking times were measured with eye-tracking (EyeLink 1000 Plus). Building on Bartels and colleagues' original study (2009), we introduced a number of significant methodological changes. First, we adopted a *passage-to-word* order rather than word-to-passage order. Therefore, infants were familiarized with auditory text passages in German containing two target words. They were then tested with auditory novel and familiar words presented in isolation. Second, although we kept the same target words (i.e., *Balken*, *Felsen*, *Pinsel*, *Kurbel*) we divided the infants into two groups: Half of the infants (group A) were familiarized with passages containing the words *Pinsel* and *Felsen* (in English: paintbrush, rock), and half of the infants (group B) were familiarized with the passages containing *Balken* and *Kurbel* (in English: beam, crank). In Bartels and colleagues' study (2009), group A was familiarized with passages containing *Balken* and *Pinsel* whereas group B was familiarized with the words *Felsen* and *Kurbel*. These methodological changes were motivated by the following considerations. First, familiarization with text passages, as opposed to word lists, is relatively closer to infants' natural language input. Second, on average, infants start producing /p/ and /b/ earlier than /k/ and /f/ (Sander, 1972), and there is a well-known link between perception and production abilities in language development (Tsao et al., 2004). Hence, we balanced the stimuli by familiarizing each group of infants with one target word starting with an early-pronounced consonant (i.e., *Pinsel* or *Balken*), and one target word starting with a later-pronounced consonant (i.e., *Felsen* or *Kurbel*).

During the familiarization phase, all infants listened to two different text passages (see Appendix, Table A), each of which contained six sentences including a target word (i.e., *Pinsel, Felsen* or *Balken, Kurbel)*. The target word was presented once in each sentence, at different positions across sentences (i.e., at the beginning, in the middle, or at the end). In total, there were 4 familiarization trials, as each text passage for the two target words was repeated twice. Before each trial, an attention grabber (i.e., a colorful rotating wheel) was displayed until the infant focused on the screen. The wheel was accompanied by the sound of either a baby laughter, a bike bell or a bird song. Once the infant fixated on the screen, the auditory text passage was played and a colorful checkerboard with blinking squares was shown on the screen. The audio was played until the end regardless of whether the infant fixated the screen or not. This scheme was then repeated for each familiarization trial, resulting in a total duration of about 80 seconds.

Following the familiarization phase, all infants went through the same 12 test trials. Each trial contained a repeated presentation of a word, which could be familiar (i.e., target words presented in the familiarization phase) or novel (i.e., the familiarized words of the other group of infants). Four trials, with four different words, formed a block. For group A, familiar words were *Pinsel* and *Felsen*, whereas novel words were *Balken* and *Kurbel*. For group B, familiar words were *Balken* and *Kurbel*, while novel words were *Pinsel* and *Felsen*. Each word was repeated with varied prosody for a maximum of 32 seconds and the number of words was set accordingly (i.e., *Balken* repeated 31 times, *Felsen* 27, *Kurbel* 28, *Pinsel* 28). During the test phase, first an attention grabber was displayed (i.e., the colorful rotating wheel and interesting sound). Once the infant fixated on the screen, the repeated auditory presentation of the first word began, which was either a familiarized or a novel word, while a colorful checkerboard with blinking squares was shown on the screen. Differently from the familiarization phase, the duration of the trials in the test phase was infant-controlled. If the infant looked away from the screen for longer than two seconds, the presentation was automatically terminated and the attention grabber was presented until the infant looked back at the screen. At this, the next test trial (with repeated presentation of another word) started. The test trials appeared in a pseudo randomized sequence. Different looking times at the screen while listening to novel words versus familiar words were taken to indicate that infants had segmented the target words from the familiarization text passages.

Figure 1

Word segmentation task



Note. An illustration of the word segmentation task for group B (i.e., infants who were familiarized with *Balken* and *Kurbel*). Upper rows: Visual stimuli presented on the screen (i.e., a rotating wheel and a blinking checkerboard). Lower rows: Auditory stimuli. For the test trials, novel words are marked in orange, familiar words in blue. The test phase included three blocks (see boxes with bold lines), each containing the repeated presentations of the four words in a different order.

2.3 Procedure

The study was part of a larger study in which participants completed six paradigms. In order to test the link between word segmentation ability and emotional synchrony, we used two tasks: Word segmentation task and free play interaction. The order of tasks was counterbalanced.

The paradigm was conducted with an arm-mounted EyeLink 1000 eye tracker sampling at 500 Hz with a 16 mm lens. Infants' gaze position was approximately 55 cm from the eye tracker. Stimuli were displayed on a 21.5 inch BenQ GL2250 - LED monitor. Two Logitech Z200 speakers were located behind the monitor, which played the audio stimuli at an intensity of around 65 dBA. The experiment was programmed in MATLAB (R2018b). The infant sat in a Maxi-Cosi in front of the screen, and the mother sat on a chair right behind the infant. Mothers were instructed (1) not to interact with the infant once the experiment started, (2) not to point at the screen, (3) to keep their gaze straight toward the monitor. The experiment started after a 3-points calibration procedure.

In order to measure emotional synchrony, mothers were instructed to play with their child as they would do at home. Due to Covid-19 regulations, we provided them with a small range of toys (e.g., a stacking ring toy, a ball, a doll) and, if they preferred, they could use toys they brought from home. The free play interaction lasted 5 minutes. In case the infant was fussy, we gave breaks.

2.4 Data coding and analysis

The sessions were video recorded with 25 frames by second and coded offline for positive, negative, and neutral emotional expressions following the criteria defined in Feldman and colleagues' study (2011). Each expression needed to last for a minimum of 1 second in order to be coded. Positive facial expressions were defined as smiles with lips turned upward. Negative facial expressions were defined as a curled mouth or grimacing expressing discontentment, distress, anger. Neutral facial expressions were indicated by the mouth staying in a baseline position (i.e., neither upward nor downward). Finally, in case the mouth was not visible, or the person was eating, we coded as not codable. The micro-coded time-series analysis was performed using Mangold Interact (version 16). After an inter-rater reliability was achieved for the 20% (*N*=5) of the dyads (kappa = 0.87 for infant facial expressions; kappa = 0.91 for maternal facial expressions), two observers coded half of the dyads' videos each.

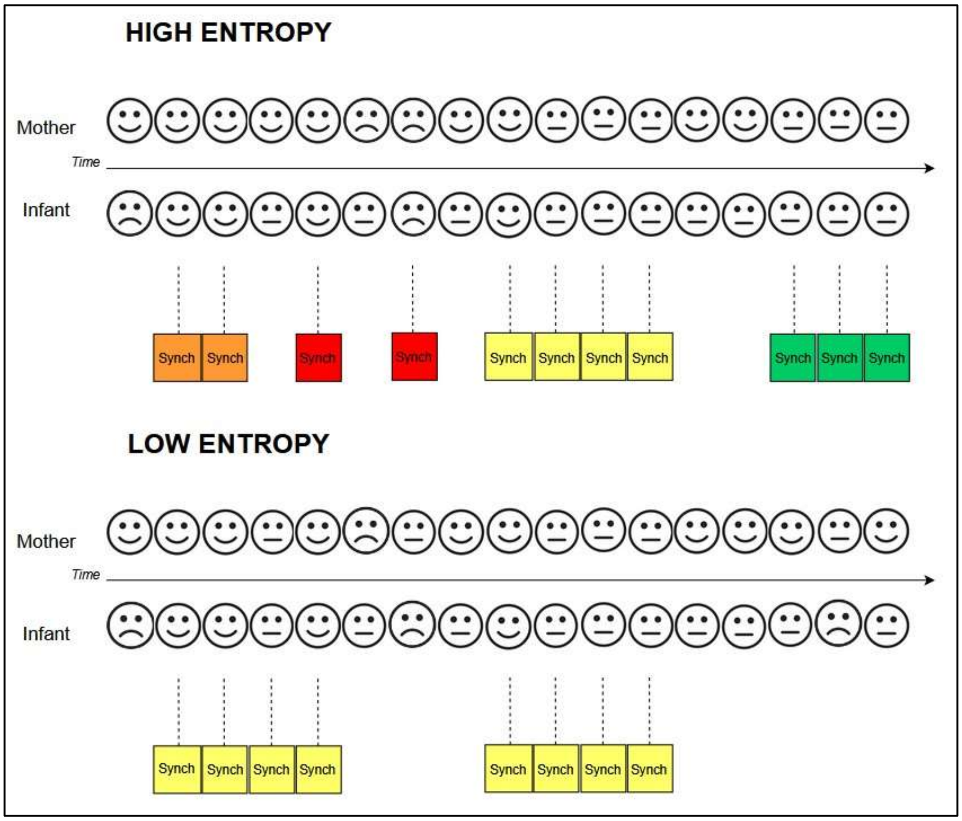
Emotional synchrony was defined as the co-occurrence of matching facial expressions: Time points (i.e., video frames) in which both the mother and infant showed the same expression (i.e., positive, neutral, or negative). To examine emotional synchrony, we used CRQA for categorical time series (Coco & Dale, 2014). CRQA examines how and the degree to which two or more time series exhibit a similar pattern. The graphical display of a recurrence analysis is the recurrence plot (RP). A RP is a two-dimensional graph which shows the times at which a state in one dynamical system occurs simultaneously in a second dynamical system. In the current study, on the x-axis there are mother's facial expressions across time; on the y-axis there are infant's facial expressions across time. For each time point, which is sampled every 40 msec, in case mother and infant share the same facial expression there will be a black dot, which is defined as "recurrence point"; if mother and infant have different facial expressions there will be a white dot, which is defined as "non-recurrence point." It is possible to extract different measures to quantify the structure of a RP, such as recurrence rate (RR) and entropy (ENTR).

RR is the percentage of time points in emotional synchrony in a two-dimensional matrix of time. Specifically, it is equal to the number of recurrence points (i.e., points in which mother and infant show the same facial expressions) divided by the total number of points (i.e., recurrence and non-recurrence points). In the current study the sampling rate equals to 40 msec. Overall, RR quantifies the tendency of the dyadic system to repeat itself. In the current study, RR reflects the following: The more points of emotional synchrony there are, the higher will be the value of RR; the less points of emotional synchrony there are, the lower will be the value of RR. To give an example: Across the 5-minute free-play period, dyad X gets into emotional synchrony in 70% of the time (i.e., for 3.5 minutes). Dyad Y gets into emotional synchrony in 40% of the time (i.e., for 2 minutes). Dyad X will have a higher RR compared to dyad Y.

In addition, ENTR is a measure based on Shannon's entropy and it represents the degree of uncertainty associated with the duration of intervals of recurrence points. ENTR is equal to Σ p(x) log (1/p(x)), where "p" stands for probability, "x" stands for the length of intervals of recurrence points, and "p(x)" equals the number of times we can observe an interval of length x divided by the total number of intervals. Namely, ENTR is the sum (Σ) of the probabilities to find an interval of length x (i.e., p(x)) multiplied by their degree of uncertainty (i.e., log (1/p(x)). In the current study, ENTR reflects the following: The more stable (i.e., certain) the length of emotional synchrony intervals is, the lower will be the value of ENTR; the more fluctuating (i.e., uncertain) the length of emotional synchrony intervals is, the higher will be the value of ENTR (see Figure 2;3). To give an example: Across the 5-minute free-play period, dyad A and dyad B get into emotional synchrony in, say, 50% of time (i.e., 2.5 minutes). In dyad A, the intervals of emotional synchrony last for about the same amount of time (e.g., often for 4 seconds). In contrast, in dyad B, some of the intervals last for 3 seconds, others for 10 seconds, then 2 seconds, and so on. Therefore, RR would be comparable/identical for the two dyads (see above), but they differ in ENTR: ENTR is low for dyad A, where the length of the intervals is stable (i.e., certain), and ENTR is high for dyad B, where the length of the intervals fluctuates (i.e., is uncertain).

Figure 2

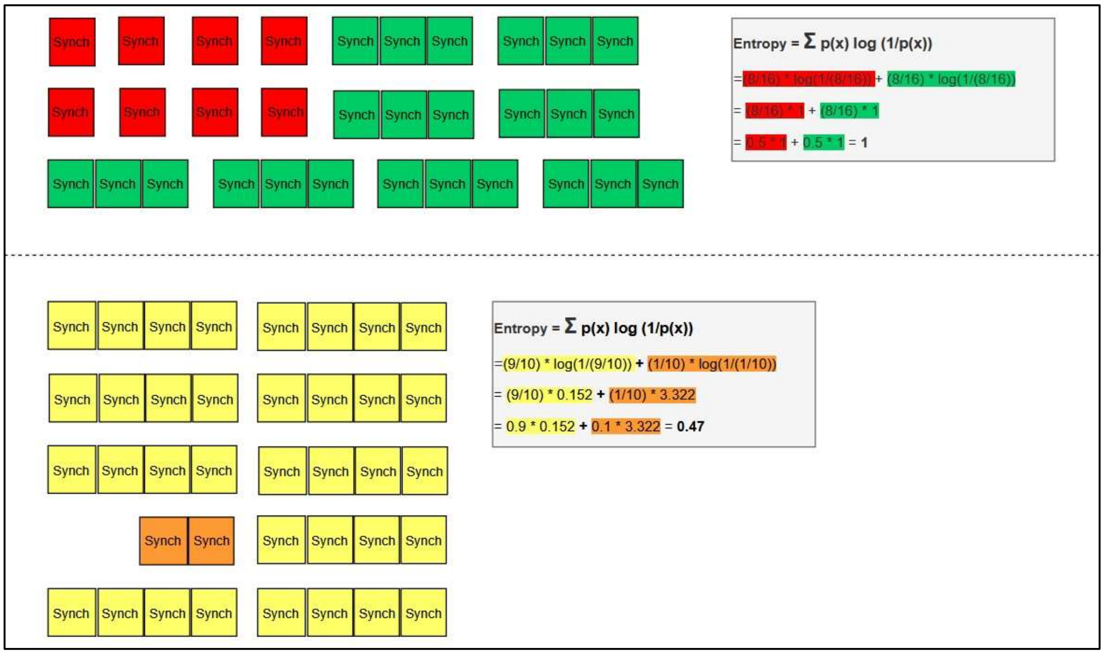
Entropy illustrated across time lines



Note. The figures represent simplified time lines to illustrate entropy in mother-infant interactions in the context of emotional expressions. For each time point, which is sampled every 40 msec, we coded a facial expression. In case facial expressions of mother and infant match for a time point (i.e., emotional synchrony), there is a colored box (i.e., "Synch"). A high level of entropy (above) results from a higher degree of variability or uncertainty of interval lengths of emotional synchrony. Namely, the dyad stays in emotional synchrony for 2 time points (i.e., orange), then for 1 time point (i.e., red), then for 1 time point, then for 4 time points (i.e., yellow), finally for 3 time points (i.e., green). It is thus difficult to predict for how long the dyad will synchronize the next time. A low level of entropy (below) is characterized by a lower degree of variability or uncertainty of interval lengths of emotional synchrony. In other words, the dyad stays in emotional synchrony repeatedly for the same amount of time (in this case 4 time points). This decreases uncertainty, and makes it easy to predict for how long a dyad will synchronize the next time.

Figure 3

Entropy calculation



Note. This figure shows examples of how ENTR is calculated for two hypothetical dyads. For each dyad, we gathered the interval lengths of emotional synchrony. During the interaction, the dyad at the top synchronizes half of the time for 1 time point (i.e., red boxes), and half of the time for 3 time points (i.e., green boxes). The dyad at the bottom synchronizes most of the time for 4 time points (i.e., yellow boxes), and once for 2 time points (i.e., orange boxes). The calculations of ENTR (see formulas on the right) is the sum of the probabilities to observe a specific interval length multiplied by the level of uncertainty related to that specific interval length. ENTR values show that the dyad at the top has a higher degree of uncertainty (i.e., higher ENTR) compared to the dyad at the bottom (i.e., lower ENTR).

In sum, both RR and ENTR inform about aspects of emotional synchrony in relation to time. However, RR, as a proportion, takes into account both recurrence and non-recurrence points; whereas ENTR takes into account recurrence points only. Moreover, RR does not capture intervals, but simply expresses how often the dyad emotionally synchronizes (i.e., proportion of recurrence points); whereas ENTR represents the uncertainty (i.e., fluctuation) of the interval length of emotional synchrony (i.e., how consistent in length they are). In other words, for ENTR it is relevant that mother and infant synchronize for recurring intervals of a certain duration rather than the frequency with which the dyad synchronizes in their emotional expression.

2.5 Data analysis

The data analysis was conducted with RStudio (RStudio Team, 2021). The primary dependent variable of interest was *Looking Times* (LTs) by trial during the test phase of the word segmentation task. Looking times were defined as times in milliseconds in which the infant looked at the screen during each test trial. Because the looking time data was not normally distributed, as commonly observed in infant data, we log-transformed it (i.e., log(LT); Csibra et al., 2016). The independent variables related to the word segmentation task was trial type (familiar vs. novel words). Independent variables concerning emotional synchrony were the ENTR and RR. The two measures of emotional synchrony were mean-centered and separately analyzed using the lme4 package (Bates et al., 2015). Participants were included as a random effect.

Two linear mixed models were constructed to measure the effect of trial type, ENTR or RR, and their interaction, on log (LT). We used a likelihood ratio test to compare two models with the same random effects specification. Below are the models including ENTR:

- E0: log(LT) ~ Trial type + (1|Participant)
- E1: log(LT) ~ Trial type * ENTR + (1|Participant)

Here are the models including RR:

- R0: log(LT) ~ Trial type + (1|Participant)
- R1: log(LT) ~ Trial type * RR + (1|Participant)

In both cases, the resulting model with a significant *p*-value as well as the lowest Akaike Information Criterion (AIC) was further analyzed.

3. Results

The free play interaction was mostly characterized by matching states: Mothers and infants spent on average 64% of the codable interaction time in emotional synchrony (i.e., same facial expression at the same time). Moreover, the most recurrent facial expression in infants, as well as mothers, was the neutral one. Descriptive results are reported in Table 1. Entropy and Recurrence Rate were extracted from the CRQA analysis. ENTR ranged from 3.80 to 5.35 (M=4.81; SD=0.38); whereas RR ranged from 7.12 to 78.75 (M= 43.30; SD=18.77). Note that it is not possible to interpret whether ENTR values are low, high, or average compared to ENTR values in other studies, as ENTR is a measure that goes from zero to infinite and its values are idiosyncratic to each study. ENTR and RR presented a strong positive relationship (r=0.64), $p < 0.001$): the higher RR, the higher ENTR (see Figure 4).

Table 1

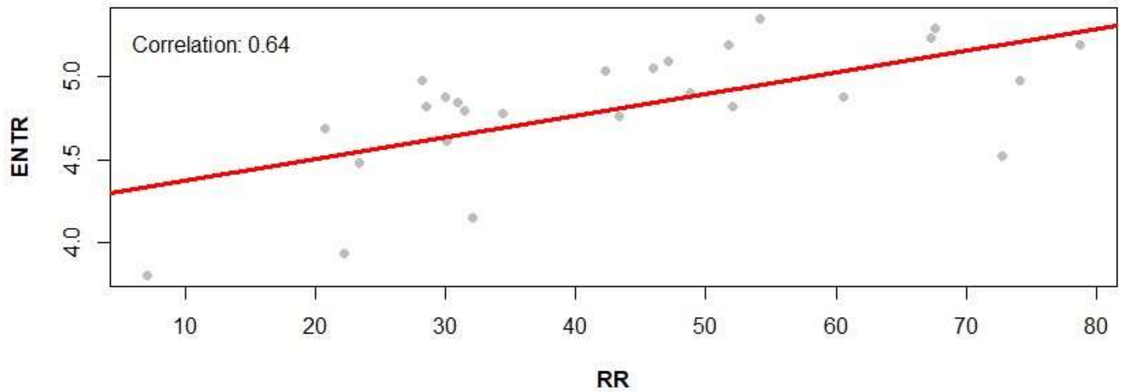
Descriptive summary of facial expressions

	Infant		**Mother**		**Mother-infant matching**	
	Duration *Mean (SD)*	Proportion	Duration *Mean (SD)*	Proportion	Duration *Mean (SD)*	Proportion
Neutral	170.77 s (73.96)	79.07%	176.35 s (64.89)	66.13%	117.59 s (68.65)	55.50%
Positive	29.32 s (30.09)	13.93%	85.80 s (47.25)	33.67%	14.82 s (13.67)	8.62%
Negative	14.21 s (20.07)	7.0%	0.59 s (2.60)	0.20%	0.08 s (0.42)	0.03%

Note. The table shows a descriptive summary of the facial expressions coded during the 5-minute mother-infant interaction. It presents the average amount of time in seconds, and the average percentage of time for each facial expression when codable.

Figure 4

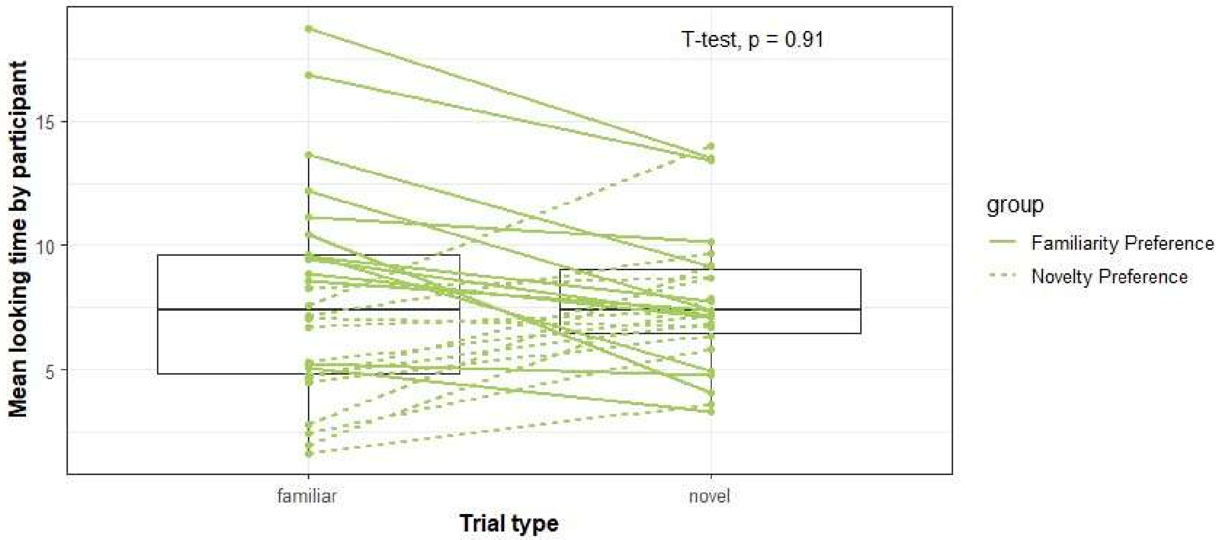
Entropy and recurrence rate by participant



For the word segmentation task, infants on average looked at the screen for 59.75 seconds (*SD*=13.38) during the 80-second familiarization phase. In test trials, they on average looked at the screen while listening to familiar words for 7.96 seconds (*SD*=7.25) and to novel words for 7.75 seconds (*SD*=6.10). The results show a lack of main effect of Trial Type in the model (β = 0.05; p = 0.52). However, it seems that half of the infants (n=13) tend to show a familiarity preference, and the other half (n=13) tend to show a novelty preference (see Figure 5). On the group level, these different tendencies likely canceled each other out.

Figure 5

Word segmentation looking times by trial type



Note. Boxplots representing mean looking times in seconds by participant while infants listened to familiar (left side) and novel words (right side) in the test trials. The solid lines represent the 13 participants who tended to look longer to the familiar than novel test trials. The dashed lines represent the 13 participants who tended to look longer to the novel than familiar test trials.

3.1 Word segmentation looking times and entropy

Following model comparisons using a maximum likelihood ratio test, the most representative model for ENTR, E1, revealed an AIC value of 635.23, with log-likelihoods represented in a chi-square statistic, $\chi 2$ (2) = 7.123, *p*= .02 (for a summary, see Table 2). Assumptions of linearity, normality of the residuals (Shapiro-Wilk test: W = 0.99, *p* = .10), homoscedasticity, absence of autocorrelation, and multicollinearity (average Variance Inflation Factor = 1.21) were checked and fulfilled before interpreting the model.

The E1 model showed that ENTR is a significant predictor of the looking times for test trials with an estimate of 0.563 (*p* = .01). Higher entropy (i.e., more uncertainty associated with the length of the intervals spent in emotional synchrony) during the 5-minute free-play period was associated with longer looking times in the test trials of the word segmentation task, regardless of Trial Type (novel or familiar). Moreover, there was a significant interaction between Trial Type and ENTR ($\beta = -0.449$; $p < .05$). The lower the ENTR during free-play, the longer infants looked at the screen while hearing the novel words as compared to the familiar words. However, at high levels of ENTR, infants looked for about the same amount of time while hearing the novel vs the familiar words (see Figure 6).

Table 2

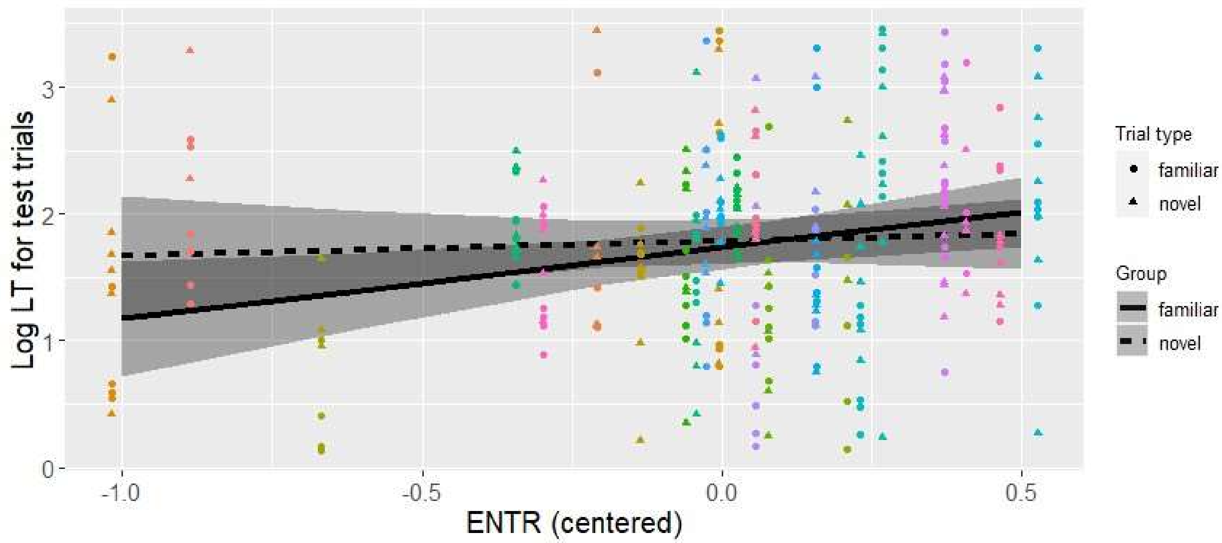
Effects of mother-infant emotional entropy on looking times during word segmentation test trials

Predictors	*Estimates*	*CI*	*p*
(intercept)	1.73	1.57 – 1.89	**<0.001**
Trial type	0.05	-0.11 – 0.21	n.s.
Entropy	0.56	0.13 – 0.99	**<0.05**
Trial type*Entropy	-0.45	-0.89 - -0.00	**<0.05**
Random effects			
σ2	0.48		
τ00 Participant	0.09		
ICC	0.15		
N Participant	26		

Note. Number of observations: 284. The reference level for Trial type is *familiar*.

Figure 6

Looking times in the word segmentation test trials as a function of entropy



Note. Scatterplot of log transformed Looking Times by trial type as a function of the entropy of emotional synchrony during the free play interaction. Data points represent individual test trials. Each infant is illustrated with a different color. The continued regression line fits the familiar trials; the dashed regression line fits the novel trials.

3.2 Word segmentation looking times and recurrence rate

Following model comparisons, the large *p*-value of R1, $\chi 2$ (2) = 1.980, $p > .05$, indicated that including RR did not improve the model fit (see Appendix, Figure A). Hence we chose the more parsimonious R0 (AIC = 638.36), which included only Trial Type as fixed effect. The R0 model showed no significant main effect of Trial Type for the looking times during test trials (β = 0.053; $p > .05$).

4. Discussion

We investigated whether and how interpersonal emotional synchrony between mothers and their infants (i.e., mutual neutral expressions, mutual positive expressions, and mutual negative expressions) during a free play interaction, predict infants' word segmentation ability. Specifically, we measured 1) RR and ENTR through a cross-recurrence analysis to quantify emotional synchrony as well as predictability of emotional synchrony and 2) infants' looking times following a familiarization phase while listening to auditory presentations of familiarized versus novel words in an eye-tracking-based word segmentation task. The results support our hypothesis that experience with stable dynamics in the emotional configuration of the mother-infant dyad is linked to infants' word segmentation performance. Our main findings are that levels of entropy interacted with trial type to predict infants' word segmentation performance. As hypothesized, the lower the entropy during free play interaction, the longer infants looked during auditory presentation of novel as compared to familiar words at test, indicating successful word segmentation performance. Interestingly, the higher the entropy during the interaction, the longer infants' looked at the screen, regardless of whether novel or familiar words were presented. Moreover, mothers and infants showed neutral expressions more frequently than positive and negative expressions during free play. This is in line with previous research on typical mother-infant interactions (Tronick, 2017; Tronick & Gianino, 1986). Contrary to our expectations, the proportion of times the infants and their mothers showed the same emotional expression, as measured by RR, did not play a role in infants' word segmentation.

The indices generated from the CRQA provided information on the dynamics of emotional synchrony between infants and their caregivers. The lower the entropy during the interaction, the better was infants' word segmentation, as reflected by their tendency to look longer to novel compared to familiar test trials. Nonetheless, higher levels of entropy during interaction were linked to infants' longer looking times at the screen at test, regardless of the trial type. There are a couple of takeaways from these findings. First, finding an association between ENTR and word segmentation suggests that rhythmicity, as a common feature, may be the linking element in the often documented association of emotion exchanges and spoken language (e.g., Nicely et al., 1999). Second, the direction of the association is in line with previous studies on language development suggesting that early language acquisition benefits from low entropy conditions. More specifically, the negative relationship between entropy during mother-infant interaction and infants' word segmentation mirrors findings showing that 9-month-old infants learn words better in low as compared to high entropy conditions. Although this negative relationship seems to contradict literature showing that greater variety of words predicts larger vocabulary, we highlight that our participants were 9-month-old. At an early stage, parental input is characterized by low entropy (i.e., more redundancy) (Tal et al., 2021). This type of input might make the first linguistic achievements (i.e., word segmentation and object-label pairing) easier for infants (Tal & Arnon, 2022). In fact, redundant/frequent words are acquired earlier and may serve as an anchor for subsequent language development (Lavi-Rotbain & Arnon, 2019). Once infants have acquired some words, they can proceed with acquiring syntax, on which they take advantage of higher quality and quantity of parental input. Third, the entropy measure provides a better estimate of the complexity and variability of the dyadic interaction, in comparison to RR, which indicates the overall proportion of emotional synchrony because ENTR captures the uncertainty of emotional exchanges (Howland et al., 2021). To sum up, caregiver-infant interactions characterized by stable dynamics may help infants to detect rhythms in social exchanges as well as in the spoken language. Stable sequences of interaction between partners might make interpersonal exchanges tractable and less cognitively demanding.

Data further revealed that the percentage of times in which the mothers and their 9-month-old infants expressed the same emotional expression, as measured by RR, does not predict word segmentation. The lack of support for our hypothesis in terms of percentage of time points in emotional synchrony in a two-dimensional matrix of time can be explained as follows:

The rhythmicity that characterizes typical mother-infant interactions is made up of matching states as well as mismatching states[1], and the repair of the latter. Both are equally important in dyadic interactions. Being in matched states means that there is an engagement and coordination among partners. Being in mismatched states represents an occasion for conflict resolution, and the repair of mismatched states is evidence of adaptive behaviors and flexibility (Beebe et al., 2016; Biringen et al., 1997; Tronick, 2017). Simply calculating RR or the overall proportion of emotional synchrony may not capture information on the time spent in each matched emotional interval. In addition, ENTR, by measuring the uncertainty of interval lengths, informs on the predictability of emotional synchrony. Predicable aspects of the environment have been argued to be easier to pay attention to than unpredictable ones (Wass, 2022). Any self-organizing agent tends to reduce the entropy of its sensory states, therefore its "free energy", which is the difference between an organism's predictions and the sensations it faces (Friston, 2010). Minimizing free energy is cognitively demanding. Therefore, environments that are less uncertain and unpredictable minimize the "free energy" and lead brains to be more energetically efficient (Peters et al., 2017). This could explain why a coarse metric such as RR does not correlate with word segmentation performance. Accordingly, measures that quantify emotional synchrony should also capture the uncertainty of the sequences of states. Therefore, we argue that entropy, which takes into account the lengths of sequences of emotional states across time and their predictability, is a more informative measure compared to recurrence rate.

Interestingly, the lower the entropy was in the interaction, the clearer was infants' looking preference for novel words. This is in line with previous findings suggesting that children with a more advanced level of language processing show a novelty preference in word

---

[1] In the current study, mothers and infants spent on average 64.15% of the codable interaction time in matched states, and 35.85% of time in mismatched states.

segmentation tasks (DePaolis et al., 2016). The tendency to look longer at the novel words as compared to familiar words, which was linked to lower the entropy levels, suggests that infants were able to discriminate between familiar versus novel words. This would indicate that they segmented the target words from the text passages during the familiarization phase. Therefore, the overall shorter looking time is not due to general fatigue (Nordt et al., 2016).

The finding that infants' overall looking times differed to the extent to which they segmented the words or not is in line with the infant learning literature. Previous research has found that infants allocate their cognitive resources strategically in order to maximize their learning. For instance, they dedicate their attentional resources to stimuli that are neither too simple nor too complex (Kidd et al., 2012). In other words, infants drive their attention to environmental stimuli that offer information gain (Poli et al., 2020). It is conceivable that infants who already successfully segmented the words during familiarization recognized them quicker at test, resulting in shorter looking times. Relatedly, infants' longer looking times at the screen regardless of the trial type might suggest that those infants were still sampling information about the structure of the experiment or the stimuli.

We would like to add a final remark on the results of the word segmentation task. The best linear mixed-model (i.e., model with the lowest AIC and significant p-value) did not reveal a significant main effect of trial type. This could have been due the fact that some infants had segmented the words (resulting in a novelty preference), whereas other infants were still learning (resulting in a familiarity preference; Hunter & Ames, 1988). The non-significant main effect of trial type is in contrast to a previous study that reported a familiarity preference in the same task in 8- to 9-month-old German infants (Bartels et al., 2009). Among the potential interpretations, we suggest that the non-significant main effect of trial type might be due to the stimuli being presented in Standard German. This may have been more suitable for the participants in the previous study, who were from Potsdam and Berlin, than for the participants of the current study, who were from an Austrian town with a rather diverse German-speaking population. This seems in line with previous studies suggesting that, within the same language, dialectal variations could result in contrasting trajectories in classic language development paradigms (Floccia et al., 2016).

Limitations and future directions

Although it is common in infancy research, the relatively small sample size and short interaction time (i.e., 5 minutes) might be considered as limitations. In the current study, we focused on emotional synchrony. This methodological choice was made to investigate commonalities of emotion and language. In principle, CRQA can be used to analyze a variety of interaction dynamics. Future studies should investigate multi-dimensional properties of interpersonal synchrony that might relate to infant language development, such as mutual gaze, infants' vocalizations, or parents' language input. Moreover, caregivers and infants adjust their behaviors to each other. Future research may examine the directionality aspect of interpersonal synchrony. Finally, longitudinal designs could be used to understand the direction of effects between emotional synchrony, rhythmicity, and word segmentation. For example, it would be interesting to examine the question of whether infants' advanced language skills contribute to more predictable rhythmicity in their interactions with their mothers. This is our future work and it is beyond the scope of the current paper. Although this is only a first step, our study provides a clear foundation for understanding the link between social interactions and language development.

References

Allen, G. D. (1975). Speech rhythm: Its relation to performance universals and articulatory timing. *Journal of Phonetics*, *3*(2), 75–86. https://doi.org/10.1016/S0095-4470(19)31351-8

Aslin, R. N. (2007). What's in a look? *Developmental Science*, *10*(1), 48–53. https://doi.org/10.1111/j.1467-7687.2007.00563.x

Bartels, S., Darcy, I., & Höhle, B. (2009). Schwa syllables facilitate word segmentation for 9-month-old German-learning infants. In *BUCLD 33 Processing* (pp. 73–84). Cascadilla Press.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Beebe, B., Cohen, P., & Lachmann, F. (2016). Patterns of communication become expected and shape infant development. In *The mother-infant interaction picture book: Origins of attachment* (pp. 31–43). WW Norton & Co.

Biringen, Z., Emde, R. N., & Pipp-Siegel, S. (1997). Dyssynchrony, conflict, and resolution: Positive contributions to infant development. *American Journal of Orthopsychiatry*, *67*(1), 4–19. https://doi.org/10.1037/h0080207

Coco, M. I., & Dale, R. (2014). Cross-recurrence quantification analysis of categorical and continuous time series: An R package. *Frontiers in Psychology*, *5*, 510. https://doi.org/10.3389/fpsyg.2014.00510

Cohn, J. F., & Tronick, E. Z. (1987). Mother–infant face-to-face interaction: The sequence of dyadic states at 3, 6, and 9 months. *Developmental Psychology*, *23*(1), 68–77. https://doi.org/10.1037/0012-1649.23.1.68

Cole, P. M., Dennis, T. A., Smith-Simon, K. E., & Cohen, L. H. (2009). Preschoolers' emotion regulation strategy understanding: Relations with emotion socialization and child self-regulation. *Social Development*, *18*(2), 324–352. https://doi.org/10.1111/j.1467-9507.2008.00503.x

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, *61*(5), 1584–1595. https://doi.org/10.2307/1130766

Csibra, G., Hernik, M., Mascaro, O., Tatone, D., & Lengyel, M. (2016). Statistical treatment of looking-time data. *Developmental Psychology*, *52*(4), 521–536. https://doi.org/10.1037/dev0000083

DePaolis, R. A., Keren-Portnoy, T., & Vihman, M. (2016). Making sense of infant familiarity and novelty responses to words at lexical onset. *Frontiers in Psychology*, *7*(715), 1–12. https://doi.org/10.3389/fpsyg.2016.00715

Feldman, R. (2012). Parent–infant synchrony: A biobehavioral model of mutual influences in the formation of affiliative bonds. *Monographs of the Society for Research in Child Development*, *77*(2), 42–51. https://doi.org/10.1111/j.1540-5834.2011.00660.x

Feldman, R., Greenbaum, C. W., & Yirmiya, N. (1999). Mother-Infant Affect Synchrony as an Antecedent of the Emergence of Self-Control. *Developmental Psychology*, *35*(1), 223–231. https://doi.org/10.1037/0012-1649.35.1.223

Feldman, R., Magori-Cohen, R., Galili, G., Singer, M., & Louzoun, Y. (2011). Mother and infant coordinate heart rhythms through episodes of interaction synchrony. *Infant Behavior and Development*, *34*(4), 569–577. https://doi.org/10.1016/j.infbeh.2011.06.008

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, *8*(2), 181–195. https://doi.org/10.1016/S0163-6383(85)80005-9

Floccia, C., Keren-Portnoy, T., DePaolis, R., Duffy, H., Delle Luche, C., Durrant, S., White, L., Goslin, J., & Vihman, M. (2016). British English infants segment words only with exaggerated infant-directed speech stimuli. *Cognition*, *148*, 1–9. https://doi.org/10.1016/j.cognition.2015.12.004

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Fujii, S., & Wan, C. Y. (2014). The role of rhythm in speech and language rehabilitation: The SEP hypothesis. *Frontiers in Human Neuroscience*, *8*(777), 1–15. https://doi.org/10.3389/fnhum.2014.00777

Guevara, M., Cox, R. F. A., van Dijk, M., & van Geert, P. (2017). Attractor dynamics of dyadic interaction: A recurrence based analysis. *Nonlinear Dynamics, Psychology, and Life Sciences*, *21*(3), 289–317.

Hoehl, S., & Bertenthal, B. I. (2021). An interactionist perspective on the development of coordinated social attention. *Advances in Child Development and Behavior*, *61*, 1–41. https://doi.org/10.1016/bs.acdb.2021.05.001

Hoehl, S., Fairhurst, M., & Schirmer, A. (2021). Interactional synchrony: Signals, mechanisms and benefits. *Social Cognitive and Affective Neuroscience*, *16*(1), 5–18. https://doi.org/10.1093/scan/nsaa024

Houston, D. M., Jusczyk, P. W., Kuijpers, C., Coolen, R., & Cutler, A. (2000). Cross-language word segmentation by 9-month-olds. *Psychonomic Bulletin & Review*, *7*(3), 504–509. https://doi.org/10.3758/BF03214363

Howland, M. A., Sandman, C. A., Davis, E. P., Stern, H. S., Phelan, M., Baram, T. Z., & Glyn, L. M. (2021). Prenatal maternal mood entropy is associated with child neurodevelopment. *Emotion*, *21*(3), 489–498. https://doi.org/10.1037/emo0000726

Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research*, *5*, 69–95.

Hunter, M. A., Ames, E. W., & Koopman, R. (1983). Effects of stimulus complexity and familiarization time on infant preferences for novel and familiar stimuli. *Developmental Psychology*, *19*(3), 338–352. https://doi.org/10.1037/0012-1649.19.3.338

Junge, C., Everaert, E., Porto, L., Fikkert, P., de Klerk, M., Keij, B., & Benders, T. (2020). Contrasting behavioral looking procedures: A case study on infant speech segmentation. *Infant Behavior and Development*, *60*, 101448. https://doi.org/10.1016/j.infbeh.2020.101448

Jusczyk, P. W. (2000). How speech perception develops during the first year. In *The discovery of spoken language* (pp. 73–109). MIT Press.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39*(3–4), 159–207. https://doi.org/10.1006/cogp.1999.0716

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS ONE*, *7*(5), e36399. https://doi.org/10.1371/journal.pone.0036399

Lavi-Rotbain, O., & Arnon, I. (2019). Children learn words better in low entropy. *CogSci*, 631–637.

Likowski, K. U., Weyers, P., Seibt, B., Stöhr, C., Pauli, P., & Mühlberger, A. (2011). Sad and lonely? Sad mood suppresses facial mimicry. *Journal of Nonverbal Behavior*, *35*(2), 101–117. https://doi.org/10.1007/s10919-011-0107-4

Markova, G., Nguyen, T., & Hoehl, S. (2019). Neurobehavioral interpersonal synchrony in early development: The role of interactional rhythms. *Frontiers in Psychology*, *10*(2078), 1–6. https://doi.org/10.3389/fpsyg.2019.02078

Mermier, J., Quadrelli, E., Turati, C., & Bulf, H. (2022). Sequential learning of emotional faces is statistical at 12 months of age. *Infancy*, *27*(3), 479–491. https://doi.org/10.1111/infa.12463

Mogan, R., Fischer, R., & Bulbulia, J. A. (2017). To be in synchrony or not? A meta-analysis of synchrony's effects on behavior, perception, cognition and affect. *Journal of Experimental Social Psychology*, *72*, 13–20. https://doi.org/10.1016/j.jesp.2017.03.009

Nelson, K. E., Welsh, J. A., Trup, E. M. V., & Greenberg, M. T. (2011). Language delays of impoverished preschool children in relation to early academic and emotion recognition skills. *First Language*, *31*(2), 164–194. https://doi.org/10.1177/0142723710391887

Newman, R., Ratner, N. B., Jusczyk, A. M., Jusczyk, P. W., & Dow, K. A. (2006). Infants' early ability to segment the conversational speech signal predicts later language development: A retrospective analysis. *Developmental Psychology*, *42*(4), 643–655. https://doi.org/10.1037/0012-1649.42.4.643

Nicely, P., Tamis-LeMonda, C. S., & Bornstein, M. H. (1999). Mothers' attuned responses to infant affect expressivity promote earlier achievement of language milestones. *Infant*

Behavior and Development, *22*(4), 557–568. https://doi.org/10.1016/S0163-6383(00)00023-0

Nordt, M., Hoehl, S., & Weigelt, S. (2016). The use of repetition suppression paradigms in developmental cognitive neuroscience. *CORTEX*, *80*, 61–75. https://doi.org/10.1016/j.cortex.2016.04.002

Ornaghi, V., Brazzelli, E., Grazzani, I., Agliati, A., & Lucarelli, M. (2017). Does training toddlers in emotion knowledge lead to changes in their prosocial and aggressive behavior toward peers at nursery? *Early Education and Development*, *28*(4), 396–414. https://doi.org/10.1080/10409289.2016.1238674

Ornaghi, V., Pepe, A., Agliati, A., & Grazzani, I. (2019). The contribution of emotion knowledge, language ability, and maternal emotion socialization style to explaining toddlers' emotion regulation. *Social Development*, *28*(3), 581–598. https://doi.org/10.1111/sode.12351

Ota, M., & Skarabela, B. (2018). Reduplication facilitates early word segmentation. *Journal of Child Language*, *45*(1), 204–218. https://doi.org/10.1017/S0305000916000660

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Statistical Learning in a Natural Language by 8-Month-Old Infants: **Statistical Learning in a Natural Language**. *Child Development*, *80*(3), 674–685. https://doi.org/10.1111/j.1467-8624.2009.01290.x

Peters, A., McEwen, B. S., & Friston, K. (2017). Uncertainty and stress: Why it causes diseases and how it is mastered by the brain. *Progress in Neurobiology*, *156*, 164–188. https://doi.org/10.1016/j.pneurobio.2017.05.004

Poli, F., Serino, G., Mars, R. B., & Hunnius, S. (2020). Infants tailor their attention to maximize learning. *Science Advances*, *6*(39), eabb5053. https://doi.org/10.1126/sciadv.abb5053

Radulescu, S., Wijnen, F., & Avrutin, S. (2019). Patterns bit by bit: An entropy model for rule induction. *Language Learning and Development*, *16*(2), 109–140. https://doi.org/10.1080/15475441.2019.1695620

Rohlfing, K. J., & Nomikou, I. (2014). Intermodal synchrony—As a form of maternal responsiveness—Is associated with language development. *Language, Interaction and Acquisition*, *5*(1), 117–136. https://doi.org/10.1075/lia.5.1.06roh

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *WIREs Cognitive Science*, *1*(6), 906–914. https://doi.org/10.1002/wcs.78

RStudio Team. (2021). *RStudio: Integrated Development Environment for R*. RStudio, PBC, Boston, MA URL. http://www.rstudio.com/

Ruba, A. L., Pollak, S. D., & Saffran, J. R. (2022). Acquiring complex communicative systems: Statistical learning of language and emotion. *Topics in Cognitive Science*, *0*, 1–19. https://doi.org/10.1111/tops.12612

Saffran, J. R., & Wilson, D. P. (2003). From Syllables to Syntax: Multilevel Statistical Learning by 12-Month-Old Infants. *Infancy*, *4*(2), 273–284. https://doi.org/10.1207/S15327078IN0402_07

Sander, E. K. (1972). When are speech sounds learned? *Journal of Speech and Hearing Disorders*, *37*(1), 55–63. https://doi.org/doi:10.1044/jshd.3701.55

Tal, S., & Arnon, I. (2022). Redundancy can benefit learning: Evidence from word order and case marking. *Cognition*, *224*, 105055. https://doi.org/10.1016/j.cognition.2022.105055

Tal, S., Grossman, E., & Arnon, I. (2021). Infant-directed speech becomes less redundant as infants grow: Implications for language learning. *PsyArXiv*. https://doi.org/10.31234/osf.io/bgtzd

Tronick, E. Z. (2017). The caregiver–infant dyad as a buffer or transducer of resource enhancing or depleting factors that shape psychobiological development. *Australian and New Zealand Journal of Family Therapy*, *38*(4), 561–572. https://doi.org/10.1002/anzf.1274

Tronick, E. Z., & Gianino, A. (1986). Interactive mismatch and repair: Challenges to the coping infant. *Zero to Three*, *6*(3), 1–6.

Tsao, F.-M., Liu, H.-M., & Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development*, *75*(4), 1067–1084. https://doi.org/10.1111/j.1467-8624.2004.00726.x

Tschacher, W., Rees, G. M., & Ramseyer, F. (2014). Nonverbal synchrony and affect in dyadic interactions. *Frontiers in Psychology*, *5*(1323), 1–13. https://doi.org/10.3389/fpsyg.2014.01323

Wass, S. (2022). Contingencies and periodicities: Two types of predictability in children's real-world environment that might influence the executive control of attention. *PsyArXiv*. https://doi.org/10.31234/osf.io/z6rjc

Xu, T. L., de Barbaro, K., Abney, D. H., & Cox, R. F. A. (2020). Finding structure in time: Visualizing and analyzing behavioral time series. *Frontiers in Psychology*, *11*(1457), 1–19. https://doi.org/10.3389/fpsyg.2020.01457

Zahner, K., Schönhuber, M., & Braun, B. (2016). The limits of metrical segmentation: Intonation modulates infants' extraction of embedded trochees. *Journal of Child Language*, *43*(6), 1338–1364. https://doi.org/10.1017/S0305000915000744

Appendix

Table A

Text passages presented during the Familiarization phase of the word segmentation task

Target word: BALKEN (in English: *beam*)

Der Balken lag quer über dem tiefen Abgrund. Es war ein sehr breiter und stabiler Balken. Über diesen Balken spazierten die Wanderer auf die andere Seite des Tales. Manchmal liefen auch Kinder den Balken entlang. Auf dem dicken Balken zu spielen war sehr gefährlich. Zum Glück ist noch nie ein Kind vom Balken gefallen.	*The beam lay across the deep abyss. It was a very wide and stable beam. Over this beam the hikers walked to the other side of the valley. Sometimes children also walked over the beam. Playing on the thick beam was very dangerous. Fortunately, no child ever fell from the beam.*

Target word: KURBEL (in English: *crank*)

Eine Kurbel stand in dem klapprigen Schuppen. Der Junge wollte unbedingt an der Kurbel drehen. Doch die Kurbel war völlig eingerostet und bewegte sich nicht. Das Kind holte seinen Vater zu der alten Kurbel. Er schmierte unsere Kurbel mit fettigem Öl aus einer schwarzen Kanne. Jetzt ist diese Kurbel wieder kinderleicht zu bedienen.	*A crank was standing in the rickety shed. The boy was eager to turn the crank. But the crank was completely rusty and would not move. The child fetched his father to the old crank. He lubricated our crank with greasy oil from a black can. Now this crank is child's play to operate again.*

Target word: PINSEL (in English: *paintbrush*)

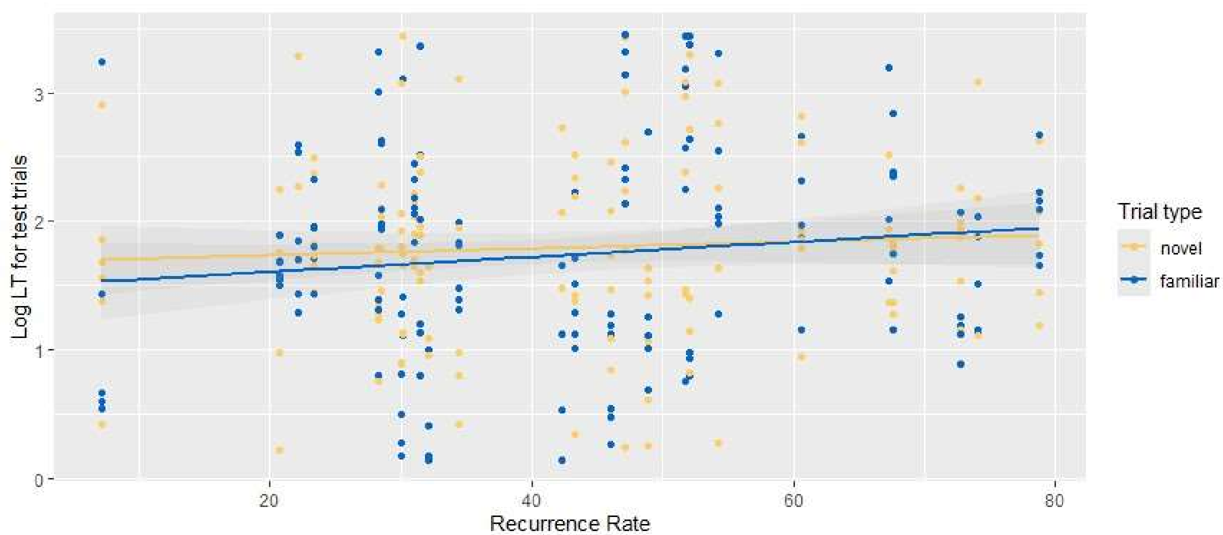
Der Pinsel lag in der hintersten Ecke der hölzernen Schublade. Das Mädchen wollte mit Pinsel und Farbe ein buntes Bild malen. Sie suchte ihren Pinsel im ganzen Kinderzimmer. Schließlich fand sie den Pinsel in ihrem Schreibtisch. Doch dieser alte Pinsel hatte kaum noch Borsten. So musste das Kind erst einen neuen Pinsel kaufen.	*The paintbrush was in the far corner of the wooden drawer. The girl wanted to paint a colorful picture with her brush and paint. She looked for her paintbrush all over her room. Finally, she found the paintbrush in her desk. But this old paintbrush had hardly any bristles left. So the child had to buy a new brush first.*

Target word: FELSEN (in English: *rock*)

Ein Felsen ragte hoch aus der Umgebung empor. Ein kleiner Bach floss unten am Felsen entlang. Auf dem Felsen stand ein verwunschenes Schloss. Viele Besucher stiegen auf den riesigen Felsen hinauf. Vom Felsen aus konnte man die ganze Landschaft überblicken. Am Abend war dieser Felsen wieder allein in der untergehenden Sonne.	*A rock rose high from the surroundings. A small stream flowed along the bottom of the rock. On the rock, there stood an enchanted castle. Many visitors climbed the huge rock. From the rock one could overlook the whole landscape. In the evening, this rock was again alone in the setting sun.*

Figure A

Looking times in the word segmentation test trials as a function of recurrence rate



Note. Scatterplot of log transformed Looking Times by trial type as a function of the Rate of Recurrence of emotional synchrony during the free play interaction. Data points represent individual test trials of each infant.

CRediT author statement

Monica Vanoncini: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, Visualization, Roles/Writing - original draft, Writing - review & editing. **Natalie Boll-Avetisyan:** Conceptualization, Funding acquisition, Methodology, Supervision, Writing - review & editing. **Birgit Elsner:** Funding acquisition, Writing - review & editing. **Stefanie Hoehl:** Conceptualization, Resources, Software, Writing - review & editing. **Ezgi Kayhan:** Conceptualization, Data curation, Funding acquisition, Methodology, Project administration, Supervision, Writing - review & editing.