PNAS BRIEF REPORT | PSYCHOLOGICAL AND COGNITIVE SCIENCES NEUROSCIENCE OPEN ACCESS

Recognizability and timing of infant vocalizations relate to fluctuations in heart rate

Jeremy I. Borjon[a,b,c,1], Manash K. Sahoo[a], Katherine D. Rhodes[b], Rebecca Lipschutz[d], and Johanna R. Bick[a,b]

Edited by Megan Gunnar, University of Minnesota Institute of Child Development, Minneapolis, MN; received October 2, 2024; accepted November 15, 2024

For human infants, producing recognizable speech is more than a cognitive process. It is a motor skill that requires infants to learn to coordinate multiple muscles of varying functions across their body. This coordination is directly linked to ongoing fluctuations in heart rate; a physiological process that can scaffold behavior. We investigated whether ongoing fluctuations in heart rate coincide with vocal production and word formation in 24-mo-old infants. Infants were most likely to produce a vocalization when heart rate fluctuations reached a peak (local maximum) or trough (local minimum). Vocalizations produced at the peak were longer than expected by chance. Functionally, vocalizations produced just before the trough, while heart rate is decelerating, were more likely to be recognized as a word by naive listeners. Thus, for the developing infant, heart rate fluctuations align with the timing of vocal productions and are associated with their duration and the likelihood of producing recognizable speech. Our results have broad and immediate implications for understanding normative language development, the evolutionary basis and physiological process of vocal production, and potential early indicators of speech and communication disorders.

language development | vocal production | autonomic processes

Language is a hallmark cognitive achievement in humans and understanding its biological substrates is a core pursuit of psychology and neuroscience. Producing speech is a physiological process recruiting multiple cognitive, linguistic, and motor systems and the result of a protracted developmental process. Infants learn to speak iteratively by generating vocalizations which train the vocal apparatus and its articulators, culminating in the emergence of their first recognizable words (1). While much research has focused on the coordination of the orofacial articulators during the emergence of speech (2), little is known about the extended physiological systems supporting vocal production in humans despite the fact that the very act of vocalizing recruits multiple motor (3) and autonomic (4) processes.

From breathing and chewing to directing visual attention, the muscles involved in vocal production are varied in their function (1). Like non-human primates and other mammals, human infant vocalizations are produced by vibrating the vocal folds of the larynx, pushing air through the airways, and coordinating multiple articulatory muscles (5). These processes dynamically interact with the baroreceptor reflex control system to produce rhythmic oscillations in heart rate and blood pressure (6, 7). Such emergent, autonomic oscillations influence behavior; nonhuman primates time contact calls to fluctuations in heart rate (8). While changes in heart rate have been associated with crying in very young infants (9), it is unknown whether and how these fluctuations relate to other vocalizations. As early infancy is characterized by emerging cortical control of behavior and cognition (10), the influence of autonomic processes on infant vocalizations likely extends beyond cries.

We measured the behavior of 24-mo-old infants while they played with their caregiver in a small playroom. Given infants in this age range are still learning to produce recognizable words, momentary increases in heart rate may accompany the motor coordination to vocalize. We therefore predicted that the timing of some vocalizations would be related to the peaks of heart rate fluctuations. As every vocalization produced by the infant contributes to training the vocal apparatus and articulators, we considered any sound emitted by the infant's mouth regardless of intentionality or linguistic content to be a vocalization. To determine whether vocalizations could be recognized as a word, a group of 4 listeners naive to the aims of the study scored whether each vocalization was recognizable as a word. Mobile electrocardiogram recordings were simultaneously conducted during play and converted to heart rate. For each session, the heart rate was converted into percentiles to control for individual variability in resting heart rate.

Results

A total of 2,708 vocalizations [mean (std): 79.64 (36.35) per infant] were emitted by 34 unique infants (20 male) between 18 and 27 mo of age [23.05 (2.62)]. Play sessions lasted from 354.93

Author affiliations: [a]Department of Psychology, University of Houston, Houston, TX 77204; [b]Texas Institute for Measurement, Evaluation, and Statistics, University of Houston, Houston, TX 77204; [c]Texas Center for Learning Disorders, University of Houston, Houston, TX 77204; and [d]Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, GA 30322

Author contributions: J.I.B., R.L., and J.R.B. designed research; R.L. and J.R.B. performed research; J.I.B., M.K.S., and K.D.R. analyzed data; and J.I.B. and J.R.B. wrote the paper.

The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: jiborjon@uh.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2419650121/-/DCSupplemental.

Published December 16, 2024.

Table 1. For each recognizability score, the percentage of vocalizations, their mean duration and SD in seconds, and percentage of each score for each bin of heart rate phase

			Heart rate phase bin					
Recognizability score	Percentage (%)	Duration (s)	Around peak 330° to 30° (%)	Beginning of fall 30° to 90° (%)	End of fall 90° to 150° (%)	Around trough 150° to 210° (%)	Beginning of rise 210° to 270° (%)	End of rise 270° to 330° (%)
0	15.29	0.96 (0.70)	18.71	15.3	16.09	15.5	13.81	13.83
1	41.21	1 (0.75)	40.74	44.85	35.34	41.7	38.62	42.96
2	33.20	0.93 (0.67)	31.38	31.4	35.06	32.1	36.57	34.22
3	9.38	0.87 (0.67)	7.99	8.18	12.93	9.83	9.21	8.25
4	0.92	0.93 (0.76)	1.17	0.26	0.57	0.87	1.79	0.73

to 903.64 seconds in duration [621.51 (76.89)]. Vocalizations varied in their duration from 0.23 to 14.07 seconds [1.17 (0.79)] with a minority (1.91%) determined to be cries. A group of 4 naive listeners independently scored each vocalization as an unrecognizable word (0) or a recognizable word (1), the sum of these scores across the individual listeners was considered the vocalization's recognizability score. This resulted in a range between 0 (completely unrecognizable as a word) to 4 (completely recognizable as a word). Completely recognizable words comprised 0.92% of the dataset while completely unrecognizable vocalizations comprised 15.29% (Table 1).

Unsupervised learning techniques revealed higher-order relationships between a vocalization's recognizability score, duration, and the percentile of heart rate at onset. Three clusters emerged: Cluster 1 (40.06% of the data) was found to be highly recognizable and of short duration, Cluster 2 (8.36% of the data) was moderately recognizable and of longer duration, and Cluster 3 (51.58% of the data) was not recognizable and comparable in duration to Cluster 1 (11). There was a significant effect of cluster membership on recognizability score [$F(2, 2498) = 2744.7$, $P < 0.0001$]. Post hoc analyses revealed the recognizability scores for each cluster was significantly different from each other (minimum $Z = 13.96$, maximum $Z = 32.48$, all $P < 0.0001$). There was a significant effect of cluster membership on the duration of vocalizations [$F(2, 2498) = 1457.29$, $P < 0.0001$]. Cluster 2 vocalizations were significantly longer than Cluster 1 ($Z = 23.18$, $P < 0.0001$) and Cluster 3 ($Z = 22.72$, $P < 0.0001$) vocalizations. There was no statistically significant effect of cluster membership on heart rate percentile at vocalization onset [$F(2, 2498) = 2.21$, $P = 0.11$].

Changes in heart rate are dynamic and the point in phase along increasing and decreasing heart rate may be more relevant to scaffolding vocal production than its percentile value. Ongoing heart rate percentiles were therefore converted into phase angles with 0° corresponding to the local peak and 180° corresponding to the local trough. Phase angles were divided into 6 bins: around the peak (330° to 30°), beginning of fall (30° to 90°), end of fall (90° to 150°), around the trough (150° to 210°), beginning of rise (210° to 270°), and end of rise (270° to 330°). A randomly permuted and bootstrapped significance test (*SI Appendix*) was conducted to determine whether there was significant clustering beyond what would be expected by chance (Table 2). Vocalization onset significantly clustered (Fig. 1*A*) around the peak ($P < 0.001$) and the trough ($P < 0.001$). We then determined whether changes in heart rate were related to the duration of vocalizations. The average duration of vocalizations occurring around the peak (Fig. 1*B*) was significantly longer than expected by chance ($P < 0.004$). We next asked whether recognizable words were more likely to be produced at specific periods of heart rate phase. Words were considered to be recognizable if their recognizability score was a 3 or 4. The probability of a vocalization being recognizable was calculated for each bin of phase angles. Vocalizations occurring at the end of falling heart rate (Fig. 1*C*) were significantly more likely to be recognized as a word by naive listeners ($P < 0.001$). While there was no significant relationship between period of phase and an increase in unrecognizable vocalizations, unrecognizable vocalizations were significantly less likely to occur at the end of falling heart rate ($P < 0.031$), when vocalizations are more likely to be recognized as a word.

Discussion

In human infants, ongoing fluctuations in heart rate and vocal production interact. Heart rate phase aligns with infant vocalizations and is related to their duration and perceived recognizability. These findings implicate the activity of the autonomic nervous system in early language production with clear consequences for language development as well as communication and speech disorders.

Table 2. For each bin of heart rate phases, the number of vocalizations, the observed and simulated probability of a vocalization occurring, average vocalization duration, probability of a recognizable vocalization, and probability of an unrecognizable vocalization

	Around peak 330° to 30°	Beginning of fall 30° to 90°	End of fall 90° to 150°	Around trough 150° to 210°	Beginning of rise 210° to 270°	End of rise 270° to 330°
Number of vocalizations	513	379	348	458	391	412
Observed probability of vocalization occurring	**0.19**	0.14	0.13	**0.17**	0.14	0.15
Simulated probability of vocalization occurring	0.16 (0.16, 0.16)	0.16 (0.16, 0.16)	0.16 (0.16, 0.16)	0.16 (0.16, 0.16)	0.16 (0.16, 0.16)	0.16 (0.16, 0.16)
Observed vocalization duration	**1.229**	1.19	1.21	1.15	1.09	1.12
Simulated vocalization duration	1.224 (1.223, 1.226)	1.23 (1.23 1.24)	1.238 (1.24, 1.24)	1.228 (1.23, 1.23)	1.233 (1.23, 1.24)	1.23 (1.23, 1.23)
Observed probability of recognizable vocalization	0.09	0.08	**0.14**	0.11	0.11	0.09
Simulated probability of recognizable vocalization	0.12 (0.11, 0.12)	0.12 (0.12 0.12)	0.12 (0.12, 0.12)	0.12 (0.12, 0.12)	0.12 (0.12, 0.12)	0.12 (0.12, 0.12)
Observed probability of unrecognizable vocalization	0.91	0.92	**0.86**	0.89	0.89	0.91
Simulated probability of unrecognizable vocalization	0.87 (0.87, 0.88)	0.87 (0.87, 0.87)	0.87 (0.87, 0.87)	0.87 (0.87, 0.88)	0.87 (0.87, 0.87)	0.87 (0.87, 0.87)

For simulated values, parentheses indicate the 2.5th and 97.5th percentiles of bootstrapped CI. Observed values in bold indicate statistically significant differences from simulated values.

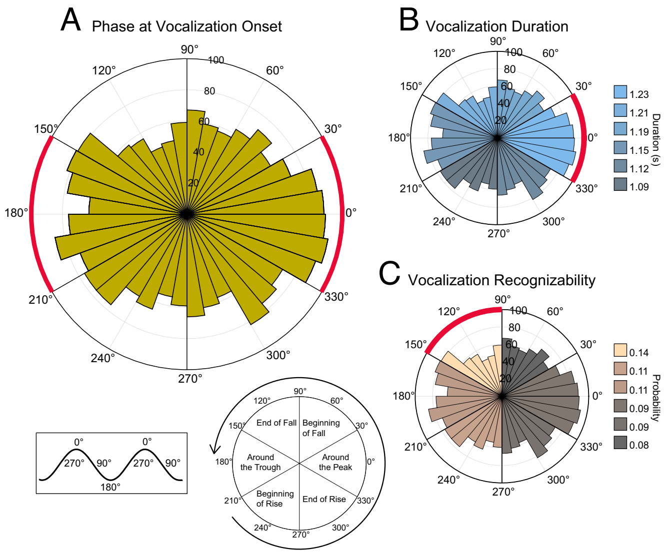

Fig. 1. Phase of heart rate relates to the timing and acoustic properties of infant vocalizations. *Bottom Left*, schematics of degrees against their location on a sine wave and phase bins. (*A*–*C*) Polar histograms of the phase of heart rate at vocalization onset. Bin height corresponds to the number of vocalizations. Bins with a perimeter highlighted in red exceeded a bootstrapped significance test for (*A*) the probability of a vocalization occurring, (*B*) vocalization durations, and (*C*) likelihood of being recognized as a word with color scales representing average per bin.

Control of the speech articulators (1) and motor coordination supporting vocal production emerges over the second year of life (3) and continues well into childhood. Still an immature motor skill at 2 y of age, infants are more likely to produce vocalizations when their heart rate has reached its relative extremes: either a local peak or trough. The coordination to produce a vocalization may be facilitated by the momentary quiescence of decelerative heart rate or the energetic support of increasing heart rate. Indeed, vocalizations produced at the local peak, when heart rate has completed its rise, are longer than expected by chance and vocalizations at the end of their deceleration are more likely to be recognized as a word. Our analyses show that the percentile value of heart rate does not appear to relate to a vocalization's duration or recognizability. Instead, heart rate's dynamic activity, whether it is increasing or decreasing, influences these properties.

The autonomic nervous system develops alongside the infant; the first years of life are marked by drastic changes in cardiorespiratory activity extending well throughout the lifespan (12). The emergence of recognizable vocalizations during decelerating heart rate implies the successful development of speech may depend on the experience of predictable ranges of autonomic activity through development. Infants who experience disruptions in their autonomic state due to stressful environments, chronic poverty, or trauma can exhibit general cognitive and language delays (13, 14).

There is also converging evidence from adults and nonhuman primates that fluctuations in cardiac activity structure neural activity and behavior. In adults, periodic heart rate oscillations, which occupy multiple frequency bands, organize large-scale neural activity (15). Heart rate decelerations have been observed in preparation of motor coordination (16) with the frontal cortex implicated in the inhibition of heart rate at rest (17). Given the immaturity of the cortex and developing motor capacities, fluctuations in heart rate may serve as a temporal regularity for the infant brain to time neural signals to vocalization-related motor effectors.

While adult marmoset monkeys structure the timing of some types of vocalizations to fluctuations in heart rate (8), their vocal behavior emerges from an interaction between autonomic activity and distance to social partners (18). For both humans and nonhuman primates, the autonomic nervous system is embedded in and reactive to a social world. Understanding the coupling between autonomic activity and behavior, and whether they decouple over development, is a critical avenue of future research for understanding the physiological emergence of cognitive processes including language production and the physiological risk factors for atypical language development.

Materials and Methods

Thirty-four caregivers consented to research approved by the Institutional Review Board at the University of Houston. Caregiver-infant dyads participated in an approximately ten-minute session of naturalistic play while wearing lightweight sensors. For *Extended Methods*, *SI Appendix*.

Data, Materials, and Software Availability. Anonymized .csv data have been deposited in Open Science Framework (https://osf.io/p7fqz/) (11).

ACKNOWLEDGMENTS. We thank Lydia Borjon for her careful reading of this manuscript. This work was supported by NIH Grant R00HD105920 (J.I.B.).

1. R. D. Kent, The maturational gradient of infant vocalizations: Developmental stages and functional modules. *Infant. Behav. Dev.* **66**, 101682 (2022).
2. M. M. Vihman, *Phonological Development: The First Two Years* (Wiley-Blackwell, 2014).
3. J. I. Borjon, D. H. Abney, C. Yu, L. B. Smith, Infant vocal productions coincide with body movements. *Dev. Sci.*, **27**, e13491 (2024). 10.1111/desc.13491.
4. K. J. Reilly, C. A. Moore, Respiratory sinus arrhythmia during speech production. *J. Speech, Lang. Hear. Res.* **46**, 164–177 (2003).
5. A. A. Ghazanfar, D. Rendall, Evolution of human vocal production. *Curr. Biol.* **18**, R457–R460 (2008).
6. C. Julien, The enigma of Mayer waves: Facts and models. *Cardiovasc Res.* **70**, 12–21 (2006).
7. P. Grossman, Respiratory sinus arrhythmia (RSA), vagal tone and biobehavioral integration: Beyond parasympathetic function. *Biol. Psychol.* **186**, 108739 (2024).
8. J. I. Borjon, D. Y. Takahashi, D. C. Cervantes, A. A. Ghazanfar, Arousal dynamics drive vocal production in marmoset monkeys. *J. Neurophysiol.* **116**, 753–764 (2016).
9. P. S. Zeskind, "Infant crying and the synchrony of arousal" in *The Evolution of Emotional Communication: From Sounds in Nonhuman Mammals to Speech and Music in Man*, E. Altenmüller, S. Schmidt, E. Zimmermann, Eds. (OUP Oxford, 2013), pp. 155–174.
10. M. S. Blumberg, K. E. Adolph, Protracted development of motor cortex constrains rich interpretations of infant cognition. *Trends Cogn. Sci.* **27**, 233–245 (2023).
11. J. I. Borjon, M. K. Sahoo, K. D. Rhodes, R. Lipschutz, J. R. Bick, Data from "Recognizability and timing of infant vocalizations relate to fluctuations in heart rate". Open Science Framework. https://osf.io/p7fqz/. Deposited 11 November 2024.
12. S. Fleming *et al.*, Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: A systematic review of observational studies. *Lancet* **377**, 1011–1018 (2011).
13. D. A. Hackman, M. J. Farah, Socioeconomic status and the developing brain. *Trends Cogn. Sci.* **13**, 65–73 (2009).
14. M. L. Engel, M. R. Gunnar, "The development of stress reactivity and regulation during human development" in *International Review of Neurobiology*, A. Clow, N. Smyth, Eds. (Academic Press, 2020), pp. 41–76.
15. K. S. Sargent *et al.*, Oscillatory coupling between neural and cardiac rhythms. *Psychol. Sci.* **35**, 517–528 (2024).
16. R. Jennings, M. van der Molen, Cardiac timing and the central regulation of action. *Psychol. Res.* **66**, 337–349 (2002).
17. E. Patron, R. Mennella, S. Messerotti Benvenuti, J. F. Thayer, The frontal cortex is a heart-brake: Reduction in delta oscillations is associated with heart rate deceleration. *Neuroimage* **188**, 403–410 (2019).
18. D. A. Liao, Y. S. Zhang, L. X. Cai, A. A. Ghazanfar, Internal states and extrinsic factors both determine monkey vocal production. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 3978–3983 (2018). 10.1073/pnas.1722426115.