

Daylong Mobile Audio Recordings Reveal Multiscale Dynamics in Infants' Vocal Productions and Auditory Experiences

Anne S. Warlaumont¹ , Kunmi Sobowale², and Caitlin M. Fausey³

¹Department of Communication, University of California, Los Angeles; ²Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles; and ³Department of Psychology, University of Oregon

Current Directions in Psychological Science
2022, Vol. 31(1) 12–19

© The Author(s) 2021



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/09637214211058166
www.psychologicalscience.org/CDPS



Abstract

The sounds of human infancy—baby babbling, adult talking, lullaby singing, and more—fluctuate over time. Infant-friendly wearable audio recorders can now capture very large quantities of these sounds throughout infants' everyday lives at home. Here, we review recent discoveries about how infants' soundscapes are organized over the course of a day. Analyses designed to detect patterns in infants' daylong audio at multiple timescales have revealed that everyday vocalizations are clustered hierarchically in time, that vocal explorations are consistent with foraging dynamics, and that some musical tunes occur for much longer cumulative durations than others. This approach focusing on the multiscale distributions of sounds heard and produced by infants is providing new, fundamental insights on human communication development from a complex-systems perspective.

Keywords

infancy, vocalization, music, timescales, daylong audio

Over the course of a day, babies may babble playfully, coo socially, scream manipulatively, attempt to produce spoken words and phrases, laugh, cry, observe quietly, and sleep silently. And they may hear adult speech, siblings' screaming, soothing lullabies, recorded voices, running water, dogs' barking, rustling of clothes, and many, many other sounds. The sounds depend on the infant's age, physical environment, culture, family structure, personality, and other factors, some of which may be relatively stable and others of which may change within or across days, weeks, and months.

All theories intended to explain human communication development (atypical or typical) make assumptions (implicit or explicit) about the statistics of the inputs infants receive. And all must account for the statistics of the sounds children produce and how they change over time. It is therefore crucial that ecological data on children's input and productions be recorded in naturalistic settings and with durations long enough to capture the range of contexts and fluctuations infants actually experience and exhibit.

Thanks to innovation in infant-friendly wearable audio recorders and related tools for quantifying patterns in everyday soundscapes (Casillas & Cristia, 2019; Gilkerson et al., 2017; VanDam et al., 2016), researchers are now able to characterize infants' sound experiences over the course of an entire day. Foundational discoveries about how these everyday soundscapes matter for young children used machine estimates of overall quantities of specific event types (e.g., number of adult words heard over the day). Human listeners' annotations of short sections of audio sampled from daylong recordings have led to further insights. Now, an additional suite of discoveries is emerging through analyses that focus on how sounds are distributed over the course of a day.

Corresponding Author:

Anne S. Warlaumont, Department of Communication, University of California, Los Angeles
Email: warlaumont@ucla.edu

One overarching finding emerging from these studies is that structure in infants' auditory and vocal experiences is nested across seconds, minutes, and hours. There is a general tendency for acoustic events to be distributed nonuniformly. Sounds occur in nested clusters such that there are many short gaps between sounds along with relatively fewer large gaps. A few sound types occur a lot and cumulate to long total durations of experience with those sound types, and many other sound types are experienced less often. We suggest that this nonuniform organization has important implications for understanding and studying human communication development.

Hierarchical Clustering of Infants' and Adults' Vocalizations in Time

A *complex system* can be defined as a system comprising many interacting components organized at multiple levels. The human brain-body-environment system is one of many naturally occurring complex systems. Researchers have analyzed the behaviors of many natural and simulated complex systems in search of commonalities across domains. One result is an understanding that complex systems tend to generate behavior that fluctuates at multiple nested scales (Kello, 2013; Kello et al., 2010; Viswanathan et al., 2011). This leads to similarity in how a pattern looks when viewed zooming in or out, or *fractality*. There are many reasons scientists have found fractality in behavior intriguing. One reason is that the degree to which there is such nesting in animal behavior often correlates with environmental features. In one study, fractality of human spatial search on a computer screen was higher when resources were clustered than when they were uniformly randomly distributed (Kerster et al., 2016). In another study, fractality of albatross foraging was greater when food resources were scarce compared with when they were plentiful (Viswanathan et al., 2011). It is possible that changes in fractality of search patterns are adaptive to the organism's environment.

Another reason for interest in nested fluctuations is that changes in fractality can be predictive of important state transitions. For example, Stephen et al. (2009) found that there was a predictable peak (an increase followed by a decrease) in the amount of nested structure in adult participants' eye movements immediately before they exhibited instances of mathematical insight. Stephen et al. noted that this pattern—increase in nested structure during times of reorganization—is a common feature of complex systems. Often this reorganization is purely self-organized; that is, it results from the internal evolution of the system's state as its components interact with each other. Reorganization can also be initiated or influenced by external inputs to the system.

Bringing a nested-structure focus to study human infants' communication, Abney et al. (2016) used the LENA (Language Environment Analysis) system (Gilkerson et al., 2017) to assess the degree of hierarchical clustering in infants' and caregivers' vocalizations during daylong recordings. LENA enables recording up to 16 hr of infant-centered audio and provides automatic tagging of when infant and adult vocalizations occurred, categorizing infant vocalizations into prespeech sounds (cooing, babbling, squealing, talking, etc.) versus reflexive or vegetative sounds (cries, laughs, coughs, etc.). Abney et al. found that the difference between the number of vocalizations within one time interval and the number in the next consecutive interval is positively correlated with the size of the time intervals. In other words, there is more difference in vocalization quantity from one hour to the next hour than from one 5-min interval to the next 5-min interval, which is consistent with the nesting of vocalization clusters apparent in Figure 1.

In addition to demonstrating hierarchical clustering of both infants' and adults' vocalizations, Abney et al. (2016) found that the degree of nesting tended to match between infants and adults. This matching effect held even after controlling for matching in overall rates of vocalization and for temporal proximity between the vocalizations. Matching was also found to increase with infants' age because adults' scaling pattern became more similar to infants'.

Vocalization-to-Vocalization Changes: A Foraging Perspective

Analyses of foraging by humans and other animals have also yielded many examples of multitime-scale, nonuniform patterns in behavior. Foraging can be considered broadly to pertain to a wide range of resource types and realms being searched (Todd & Hills, 2020). For example, when animals forage in space for prey and when human adults forage in cognitive semantic networks for items of a particular type, resources tend to be found in nested clusters over time and space (Kerster et al., 2016; Montez et al., 2015; Viswanathan et al., 2011). One way to characterize foraging behavior is to quantify the transitions between consecutive resource-gathering events. From one event to the next (this transition is sometimes called a "step"), one can measure the distance individuals "travel" within a physical or feature space. One can also measure the time between the two events.

Using LENA recordings and their associated automatically identified vocalization onsets and offsets, Ritwika et al. (2020) analyzed the acoustic differences and time elapsed between one vocalization and the next (Fig. 2, left panel). They asked whether infants'

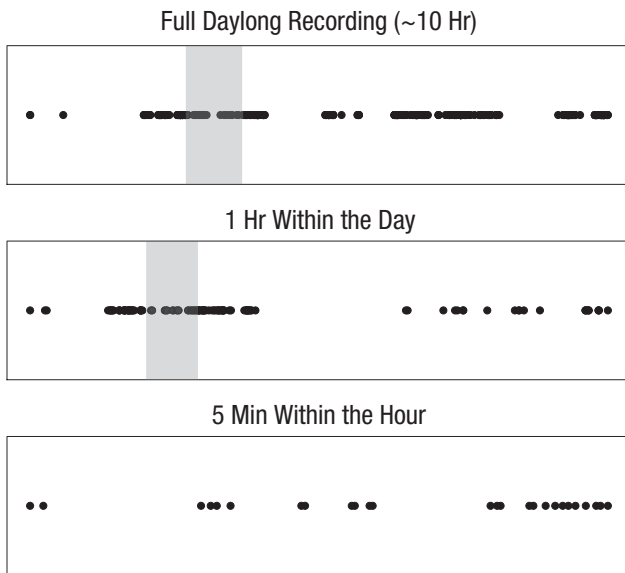


Fig. 1. Illustration of the nested structure of infants' vocalizations. The circles in the top panel show the onsets of automatically identified vocalizations in a 9-month-old infant's daylong audio recording, from earlier in the day (left) to later in the day (right). It is apparent that the infant vocalized in clusters over the day, and some of the clusters were denser and/or longer lasting than others. The area with the gray background is an hour-long period and forms the basis of the middle panel. The middle panel thus presents a zoomed-in version of a portion of the top panel. Within that hour, the infant vocalized in clusters, and the pattern of clustering is similar to the clustering at the day level even though the timescale is much smaller. The area with the gray background is a 5-min-long period and forms the basis of the bottom panel. Even within the 5-min period, the infant vocalized in clusters. Again, the clustering shows similarity in its patterning to the clustering at the hour-long and daylong scales. This figure thus illustrates fractality (i.e., self-similarity across scales) in the patterning of infants' vocalizations over time.

and adults' vocalizations could be construed as foraging through pitch and amplitude space. They also looked for evidence that vocal responses from other individuals serve as "resources" for the foraging individual. Inspired by prior foraging research, Ritwika et al. fitted mathematical distributions—normal, exponential, log-normal, or Pareto (power law)—to the observed acoustic step sizes (Fig. 2, right panel) and intervocalization intervals. These step sizes and time intervals spanned large ranges. Note that observing and measuring the longer step sizes and intervocalization intervals was possible only because of the length of the recordings.

Ritwika et al. (2020) found that for both infants and adults, the more time that elapsed between consecutive vocalizations, the bigger the change in pitch and amplitude. This aligns with foraging in other domains (Montez et al., 2015; see also Hills et al., 2012). They also found that less time elapsed from one infant vocalization to the next infant vocalization when the first was followed within 1 s by an adult vocalization than when it was not. Similarly, less time elapsed from one

adult vocalization to the next adult vocalization when the first was followed within 1 s by an infant vocalization than when it was not. This fits the hypothesis that vocalization is a type of foraging for social responses. It also corresponds with prior research on infant-adult turn taking. Ritwika et al. also observed that infants' vocalization-to-vocalization pitch movements increased with age (which suggests increasing pitch exploration), whereas amplitude movements shrank. Adults' vocalization steps in both acoustic dimensions grew larger as the infants became older. These results connect existing research on infant-adult turn taking with interdisciplinary work on foraging dynamics. They indicate that vocalization can be construed as an exploratory foraging process and that infants' and adults' patterns of vocal exploration change over developmental time.

Multiple Timescales in the Musical Sound Types Infants Experience

Focusing now on what daylong audio can reveal about the distributions of specific types of sounds infants encounter, we turn to recent discoveries about musical sounds in the environment throughout the day. Mendoza and Fausey (2021) manually annotated daylong audio recordings second by second to identify when infants' soundscapes were musical as well as the specific voices and tunes in the soundscapes. Because full waking days were annotated, it was possible to observe relatively rare musical voices and tunes and to observe the proportional differences between more and less prevalent musical identities. As shown in Figure 3, instances of a given musical tune did not cumulate to the same proportion of daily music as instances of other musical tunes; rather, infants encountered certain tunes much more than others.

How does this distributional nonuniformity matter for infants' learning? One possibility is that highly familiar tunes ground musical recognition, providing a base of deep expertise from which infants can learn to generalize to novel tunes. Experiences of numerous less available tunes may help infants establish this generalization capability (see Smith et al., 2018, for related hypotheses about early learning in other domains). Indeed, it has long been known that word frequencies in natural language follow highly skewed (Zipfian) distributions and that this nonuniformity can help adults learn words (Hendrickson & Perfors, 2019). Skewed distributions can also improve adults' category generalization (Carvalho et al., 2021).

What could account for the nonuniformity of musical-tune distributions? As with many aspects of infant-caregiver interactions, possible factors include infant and caregiver preferences, the availability of a particular option in a caregiver's memory, the appeal of novelty,

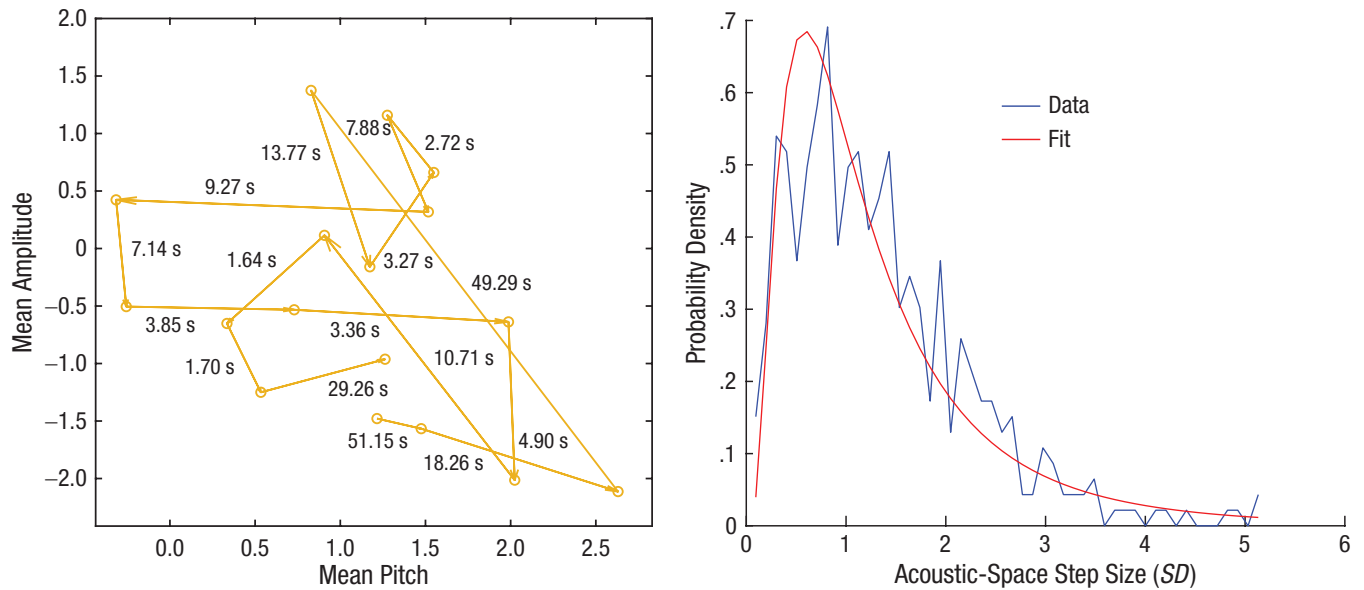


Fig. 2. Examples from Ritwika et al.'s (2020) investigation of foraging patterns in infants' vocalization. The diagram on the left shows a sample of some of the vocalization "movements," or "steps," of a 3-month-old infant's prespeech sounds. Each point represents the mean pitch (log-transformed and normalized with respect to the entire infant-vocalization data set) and the mean intensity (also normalized) of a single vocalization. Each arrow corresponds to one step. The number next to each arrow represents the time that elapsed between the two vocalizations (i.e., the intervalization interval). Acoustic-space step size was defined as the distance in the two plotted acoustic dimensions between the two vocalization points. The graph on the right shows the acoustic step-size distribution for prespeech sounds in a 2-month-old infant's recording, in cases when the first vocalization did not receive an adult's response. The graph shows that smaller step sizes were generally more frequent, but that larger step sizes (spanning more than 2 SD and up to 5 SD in the acoustic dimensions) did occur. The blue curve shows the histogram of step sizes from the raw data. In this case, a log-normal distribution was the best type of function to fit the histogram. The log-normal fit is shown by the red curve. Comparing the specific parameters of log-normal fits across recordings and interactive contexts can provide information about how infants' vocal dynamics change, such as with age or in relation to whether the infant is or is not engaged in vocal interaction with caregivers. The panel on the right was adapted from "Exploratory Dynamics of Vocal Foraging During Infant-Caregiver Communication," by V. P. S. Ritwika, G. M. Pretzer, S. Mendoza, C. Shedd, C. T. Kello, A. Gopinathan, and A. S. Warlaumont, 2020, *Scientific Reports*, 10, Article 10469, Fig. S3 in the Supplementary Information (<https://doi.org/10.1038/s41598-020-66778-0>). The original article is available under the Creative Commons CC-BY license.

and the comfort of familiarity. Given the many related factors involved, and that complex systems composed of many interacting components often self-organize to generate multiscale patterns of behavior, the answer is likely to be complicated.

Future work along these lines may enable researchers to compare how distributions of musical (or other audio) stimuli are affected by differing living situations, family structures, and early-childhood education experiences. Mendoza and Fausey's (2021) initial work, especially given that the data set is available for reuse by other researchers, may also enable machine-learning researchers to test whether the nonuniformities of input experienced by human infants yield improved capacities for machine learning (see also Bambach et al., 2018; Ossmy et al., 2018). Such research may, in turn, enhance understanding of how these distributional features affect human infants' perceptual learning.

Detecting Events Within Daylong Audio: Automatic Versus Manual Annotation

Automated algorithms available for annotating daylong child-centered audio include the LENA system's proprietary software as well as a handful of open-source alternatives (Le Franc et al., 2018; Räsänen et al., 2021; Schuller et al., 2017). LENA annotates recordings with a closed set of mutually exclusive sound-source labels and estimates counts of adults' words, the child's vocalizations, and back-and-forth conversational turns between the child and adults. One huge advantage of automated annotation is that annotation time does not scale prohibitively with recording length. Another advantage is that exactly the same algorithm can be used across projects, which eliminates variation that can occur when human annotators with different life and professional experiences interpret sounds differently.

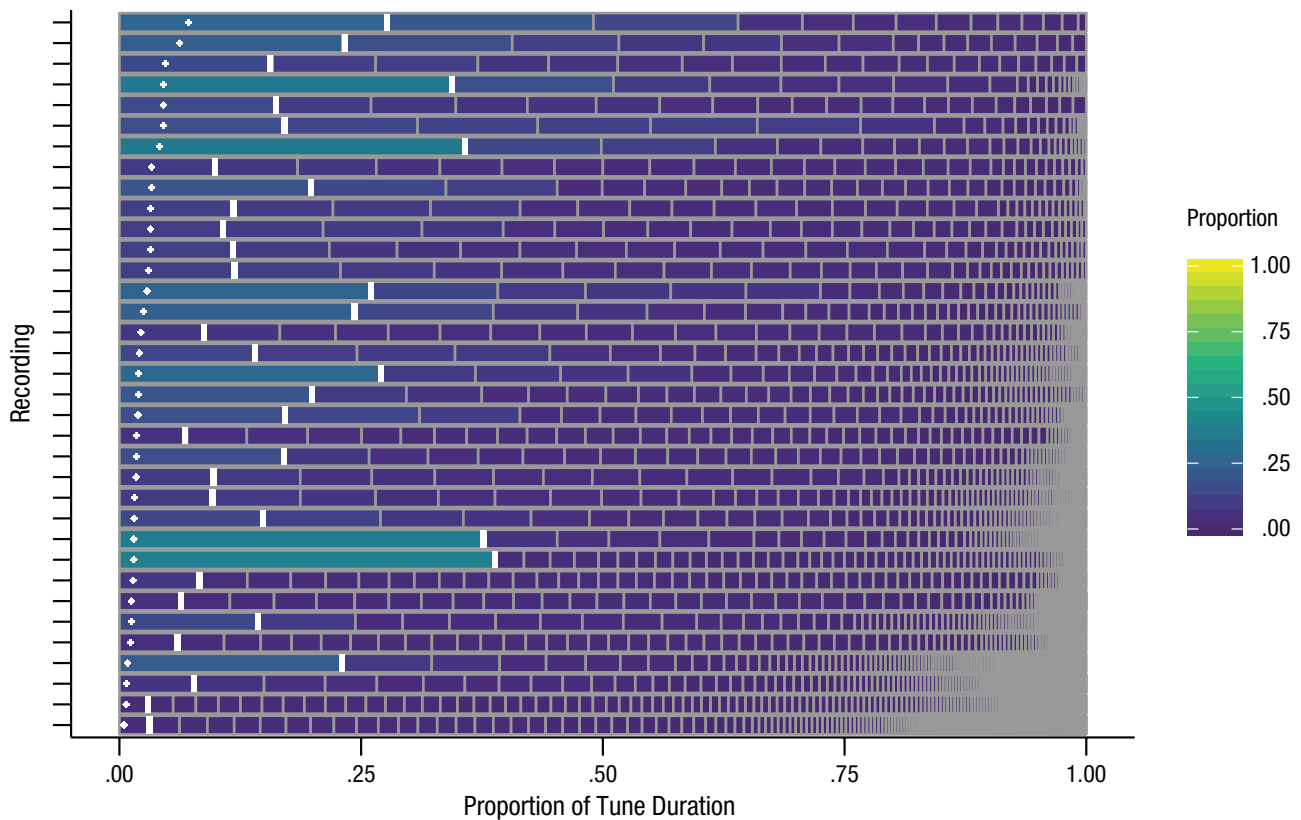


Fig. 3. The daily distribution of different tunes for each of 35 infants in Mendoza and Fausey’s (2021) study. Each row corresponds to one daylong audio recording, segmented into unique tune identities (e.g., “Twinkle, Twinkle, Little Star,” “Itsy Bitsy Spider,” “Shake It Off,” “Everybody Loves Potatoes,” “Short Whistle”). Within a row, each distinct tune’s relative duration (i.e., its proportion of the day’s musical time) is shown, in order from the tune most available to the infant (left) to the tune least available (right). The observed proportion of each recording’s most available tune is marked by the thick white vertical line. The small white plus signs (+) on the left show the proportion that would be expected per tune if each tune were equally available to the infant. Recordings are sorted with those containing the fewest total number of distinct tunes on the top. Adapted from “Everyday Music in Infancy,” by J. K. Mendoza and C. M. Fausey, 2021, *Developmental Science*, 24(6), Article e13122, Fig. 4A (<https://doi.org/10.1111/desc.13122>). The original article is available under the Creative Commons CC-BY-NC license.

However, the accuracy of automatic annotation is often lower than that of human listeners (e.g., Ferjan Ramírez et al., 2021). Further, algorithms originally trained with specific data sets for specific purposes may not generalize well. For example, LENA was developed for the purpose of obtaining word, vocalization, and turn counts at the 5-min, 1-hr, and daylong levels and in home settings (Gilkerson et al., 2017); using the detailed sound onset and offset times that the LENA algorithm provides might be unsuitable for research projects that demand very high accuracy and precision. Moreover, no automatic algorithms are currently up to the task of annotating many of the meaningful units within everyday recordings (e.g., Adolph, 2020). One issue with daylong child-centered audio recordings is that they are among the most difficult types of conversational speech data for automated systems to tag accurately (Casillas & Cristia, 2019).

An alternative is for human listeners to perform annotation. This can be an enormous undertaking—for example, 6,400 person hours were required to manually annotate the features, voices, and tunes in Mendoza and Fausey’s (2021) 35 daylong audio recordings. Infrastructure supporting sharing data and protocols (e.g., Gilmore et al., 2018; VanDam et al., 2016) helps to maximize the value of such investments. For example, shared manual annotations can provide training and evaluation for machine algorithms (e.g., Le Franc et al., 2018; Räsänen et al., 2021; Schuller et al., 2017), which in turn can provide new tools for annotating daylong recordings.

We expect that as algorithms for speech recognition and other automatic audio processing improve, and as data sets of human-annotated audio become increasingly available, it will become possible to automatically identify words, emotions, and more within

child-centered daylong audio recordings. Such advances will permit analyses of nested clustering in additional domains. They might also enable the detection of interactions across domains that partially contribute to the skewed distributions and nested clustering patterns within domains.

Capturing daylong real-world audio recordings also raises privacy concerns. Researchers must explain the issues and enable participants to make informed decisions about participation and use of their data. Some devices, such as the TILES recorder (Feng et al., 2018), provide investigators the flexibility to extract and collect data on only specific features (e.g., speech onset and offset time, pitch estimates) from the audio input. Collecting feature data alone may better preserve privacy but may not be suitable for every research question and limits reanalysis when improved automatic audio-processing tools become available.

Broader Implications and Future Directions

It is clear that, over the course of a day, infants' vocalizations and auditory experiences are organized in patterns that unfold at multiple timescales, from seconds to hours. The patterns likely extend to longer timescales (days, weeks, months), as well as to shorter timescales within utterances (Kello et al., 2017). Such multiscale behavior is characteristic of complex systems involving many interacting components, such as networks of neurons and networks of locally interacting social agents (Kello et al., 2010). It fits with the view that infant development emerges within a complex system of richly interacting components within and external to the infant (Frankenhuis et al., 2019; Oakes & Rakison, 2020; Wozniak et al., 2016).

Future research should explore how patterns of productions and input at shorter and longer timescales are related to other features of the physical and social environment (e.g., material resources, culture, family structure). Such work could help identify some of the mechanisms contributing to the multitimescale patterns we have described. It would also build bridges with other disciplines, such as anthropology (Cristia et al., 2017; Frankenhuis et al., 2019).

Future research should also explore the extent to which fractal analyses provide unique information that is not obtainable with other methods used to analyze time-series data that do not focus on the degree of self-similarity across timescales (e.g., Jebb et al., 2015). An explicit focus on dynamics across timescales and ecological contexts enables such comparisons.

Multiscale patterns in human infants' auditory and vocal experiences may also relate to brain plasticity, mental and physical health, and cognitive development. Research with


adult humans has documented individual differences in the balance between searching for resources in new places (exploration) and gathering resources in the current location (exploitation) across a range of spatial and cognitive foraging tasks. These differences are often consistent across domains and associated with performance (Todd & Hills, 2020). Regarding early-life development, experimental research using rodent models suggests that differences in the physical environment (e.g., a cage having or not having adequate nesting materials) can lead to differences in the predictability of maternal behavior, which in turn can lead to changes in offspring's brain development and variations in cognition, memory, and anhedonia of the offspring (Glynn & Baram, 2019). Predictable, repeated interactions between a caregiver and infant may signal safety and slow the maturation of the brain's corticolimbic circuitry, increasing plasticity and improving future emotion regulation (Gee & Cohodes, 2021). However, unpredictable, rare positive experiences, such as listening to a New Year's holiday song, also seem to prolong brain plasticity (Tooley et al., 2021). Most findings about how environmental experiences affect brain plasticity derive from animal models. Translating this research to humans will be facilitated by detailed data on the distributions of different event types at daylong timescales and in highly naturalistic contexts; such assays would enable researchers to measure predictability and environmental enrichment in human development.

Conclusion

Data on infants' vocal productions and auditory experiences acquired from daylong real-world recordings reveal multitimescale fluctuations and skewed distributions of event types across domains. Such patterns often arise through self-organization of complex systems of many interacting components. The findings thus support a complex-systems orientation to human development and underscore the richness and complexity of development as it unfolds in a diverse range of physical, social, and physiological contexts.

Recommended Reading

- Casillas, M., & Cristia, A. (2019). (See References). Provides a tutorial on how to start working with daylong audio recordings.
- Cychosz, M., Romeo, R., Soderstrom, M., Scaff, C., Ganek, H., Cristia, A., Casillas, M., de Barbaro, K., Bang, J. Y., & Weisleder, A. (2020). Longform recordings of everyday life: Ethics for best practices. *Behavior Research Methods*, 52(5), 1951–1969. <https://doi.org/10.3758/s13428-020-01365-9>. Provides an extensive discussion of ethical considerations when working with daylong audio recordings of children.
- Kello, C. T., Brown, G. D. A., Ferrer-i-Cancho, R., Holden, J. G., Linkenkaer-Hansen, K., Rhodes, T., & Van Orden,

- G. C. (2010). (See References). Discusses multiscale dynamics in cognitive science in relation to underlying system characteristics and in connection to other scientific domains.
- Rowe, M. L., & Snow, C. E. (2019). Analyzing input quality along three dimensions: Interactive, linguistic, and conceptual. *Journal of Child Language*, *47*(1), 5–21. <https://doi.org/10.1017/S0305000919000655>. Reviews research on language input and its role in language development.
- Warlaumont, A. S. (2020). Infant vocal learning and speech production. In J. J. Lockman & C. S. Tamis-LeMonda (Eds.), *The Cambridge handbook of infant development: Brain, behavior, and cultural context* (pp. 602–631). Cambridge University Press. <https://doi.org/10.1017/978108351959.022>. Reviews research on infants' vocal production and speech development.
- ### Transparency
- Action Editor:* Robert L. Goldstone
Editor: Robert L. Goldstone
Declaration of Conflicting Interests
 The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.
- Funding*
 This material is based partly on work supported by the National Science Foundation under Grants BCS-1529127 and SMA-1539129/1827744 and by a James S. McDonnell Foundation Scholar Award in Understanding Human Cognition, awarded to A. S. Warlaumont. K. Sobowale is supported by the U. S. National Institute of Mental Health (Grant T32 MH073517-14).
- ### ORCID iD
- Anne S. Warlaumont  <https://orcid.org/0000-0001-9450-1372>
- ### Acknowledgments
- We thank Ritwika VPS and Jennifer Mendoza for assisting with the figures.
- ### References
- Abney, D. H., Warlaumont, A. S., Oller, D. K., Wallot, S., & Kello, C. T. (2016). Multiple coordination patterns in infant and adult vocalizations. *Infancy*, *22*(4), 514–539. <https://doi.org/10.1111/infa.12165>
- Adolph, K. E. (2020). Oh, behave! *Infancy*, *25*(4), 374–392. <https://doi.org/10.1111/infa.12336>
- Bambach, S., Crandall, D. J., Smith, L. B., & Yu, C. (2018). Toddler-inspired visual object learning. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, & N. Cesa-Bianchi (Eds.), *NIPS '18: Proceedings of the 32nd International Conference on Neural Information Processing Systems* (pp. 1209–1218). Association for Computing Machinery. <https://dl.acm.org/doi/abs/10.5555/3326943.3327054>
- Carvalho, P. F., Chen, C., & Yu, C. (2021). The distributional properties of exemplars affect category learning and generalization. *Scientific Reports*, *11*, Article 11263. <https://doi.org/10.1038/s41598-021-90743-0>
- Casillas, M., & Cristia, A. (2019). A step-by-step guide to collecting and analyzing long-format speech environment (LFSE) recordings. *Collabra: Psychology*, *5*(1), Article 24. <https://doi.org/10.1525/collabra.209>
- Cristia, A., Dupoux, E., Gurven, M., & Stieglitz, J. (2017). Child-directed speech is infrequent in a forager-farmer population: A time allocation study. *Child Development*, *90*(3), 759–773. <https://doi.org/10.1111/cdev.12974>
- Feng, T., Nadarajan, A., Vaz, C., Booth, B., & Narayanan, S. (2018). TILES audio recorder: An unobtrusive wearable solution to track audio activity. In *WearSys '18: Proceedings of the 4th ACM Workshop on Wearable Systems and Applications* (pp. 33–38). Association for Computing Machinery. <https://doi.org/10.1145/3211960.3211975>
- Ferjan Ramirez, N., Hippe, D. S., & Kuhl, P. K. (2021). Comparing automatic and manual measures of parent-infant conversational turns: A word of caution. *Child Development*, *92*(2), 672–681. <https://doi.org/10.1111/cdev.13495>
- Frankenhuis, W. E., Nettle, D., & Dall, S. R. X. (2019). A case for environmental statistics of early-life effects. *Philosophical Transactions of the Royal Society B*, *374*, Article 20180110. <https://doi.org/10.1098/rstb.2018.0110>
- Gee, D. G., & Cohodes, E. M. (2021). Influences of caregiving on development: A sensitive period for biological embedding of predictability and safety cues. *Current Directions in Psychological Science*, *30*(5), 376–383. <https://doi.org/10.1177/09637214211015673>
- Gilkerson, J., Richards, J. A., Warren, S. F., Montgomery, J. K., Greenwood, C. R., Oller, D. K., Hansen, J. H. L., & Paul, T. D. (2017). Mapping the early language environment using all-day recordings and automated analysis. *American Journal of Speech-Language Pathology*, *26*(2), 248–265. https://doi.org/10.1044/2016_AJSLP-15-0169
- Gilmore, R. O., Kennedy, J. L., & Adolph, K. E. (2018). Practical solutions for sharing data and materials from psychological research. *Advances in Methods and Practices in Psychological Science*, *1*(1), 121–130. <https://doi.org/10.1177/2515245917746500>
- Glynn, L. M., & Baram, T. Z. (2019). The influence of unpredictable, fragmented parental signals on the developing brain. *Frontiers in Neuroendocrinology*, *53*, Article 100736. <https://doi.org/10.1016/j.yfrne.2019.01.002>
- Hendrickson, A. T., & Perfors, A. (2019). Cross-situational learning in a Zipfian environment. *Cognition*, *189*, 11–22. <https://doi.org/10.1016/j.cognition.2019.03.005>
- Hills, T. T., Jones, M. N., & Todd, P. M. (2012). Optimal foraging in semantic memory. *Psychological Review*, *119*(2), 431–440. <https://doi.org/10.1037/a0027373>
- Jebb, A. T., Tay, L., Wang, W., & Huang, Q. (2015). Time series analysis for psychological research: Examining and forecasting change. *Frontiers in Psychology*, *6*, Article 727. <https://doi.org/10.3389/fpsyg.2015.00727>
- Kello, C. T. (2013). Critical branching neural networks. *Psychological Review*, *120*(1), 230–254. <https://doi.org/10.1037/a0030970>

- Kello, C. T., Brown, G. D. A., Ferrer-i-Cancho, R., Holden, J. G., Linkenkaer-Hansen, K., Rhodes, T., & Van Orden, G. C. (2010). Scaling laws in cognitive sciences. *Trends in Cognitive Sciences*, *14*(6), 223–232. <https://doi.org/10.1016/j.tics.2010.02.005>
- Kello, C. T., Dalla Bella, S., Médé, B., & Balasubramaniam, R. (2017). Hierarchical temporal structure in music, speech and animal vocalizations: Jazz is like a conversation, humpbacks sing like hermit thrushes. *Journal of the Royal Society Interface*, *14*(135), Article 20170231. <https://doi.org/10.1098/rsif.2017.0231>
- Kerster, B. E., Rhodes, T., & Kello, C. T. (2016). Spatial memory in foraging games. *Cognition*, *148*, 85–96. <https://doi.org/10.1016/j.cognition.2015.12.015>
- Le Franc, A., Riebling, E., Karadayi, J., Wang, Y., Scaff, C., Metze, F., & Cristia, A. (2018). The ACLEW DiViMe: An easy-to-use diarization tool. In *INTERSPEECH 2018* (pp. 1383–1387). International Speech Communication Association. <https://doi.org/10.21437/Interspeech.2018-2324>
- Mendoza, J. K., & Fausey, C. M. (2021). Everyday music in infancy. *Developmental Science*, *24*(6), Article e13122. <https://doi.org/10.1111/desc.13122>
- Montez, P., Thompson, G., & Kello, C. T. (2015). The role of semantic clustering in optimal memory foraging. *Cognitive Science*, *39*(8), 1925–1939. <https://doi.org/10.1111/cogs.12249>
- Oakes, L. M., & Rakison, D. H. (2020). *Developmental cascades*. Oxford University Press.
- Ossmy, O., Hoch, J. E., MacAlpine, P., Hasan, S., Stone, P., & Adolph, K. E. (2018). Variety wins: Soccer-playing robots and infant walking. *Frontiers in Neurorobotics*, *12*, Article 19. <https://doi.org/10.3389/fnbot.2018.00019>
- Räsänen, O., Seshadri, S., Lavechin, M., Cristia, A., & Casillas, M. (2021). ALICE: An open-source tool for automatic measurement of phoneme, syllable, and word counts from child-centered daylong recordings. *Behavior Research Methods*, *53*(2), 818–835. <https://doi.org/10.3758/s13428-020-01460-x>
- Ritwika, V. P. S., Pretzer, G. M., Mendoza, S., Shedd, C., Kello, C. T., Gopinathan, A., & Warlaumont, A. S. (2020). Exploratory dynamics of vocal foraging during infant-caregiver communication. *Scientific Reports*, *10*, Article 10469. <https://doi.org/10.1038/s41598-020-66778-0>
- Schuller, B., Steidl, S., Batliner, A., Bergelson, E., Krajewski, J., Janott, C., Amatuni, A., Casillas, M., Seidl, A., Soderstrom, M., Warlaumont, A. S., Hidalgo, G., Schnieder, S., Heiser, C., Hohenhorst, W., Herzog, M., Schmitt, M., Qian, K., Zhang, Y., . . . Zafeiriou, S. (2017). The INTERSPEECH 2017 Computational Paralinguistics Challenge: Addressee, cold & snoring. In *INTERSPEECH 2017* (pp. 3442–3446). International Speech Communication Association. <https://doi.org/10.21437/Interspeech.2017-43>
- Smith, L. B., Jayaraman, S., Clerkin, E., & Yu, C. (2018). The developing infant creates a curriculum for statistical learning. *Trends in Cognitive Sciences*, *22*(4), 325–336. <https://doi.org/10.1016/j.tics.2018.02.004>
- Stephen, D. G., Boncoddio, R. A., Magnuson, J. S., & Dixon, J. A. (2009). The dynamics of insight: Mathematical discovery as a phase transition. *Memory & Cognition*, *37*(8), 1132–1149. <https://doi.org/10.3758/MC.37.8.1132>
- Todd, P. M., & Hills, T. T. (2020). Foraging in mind. *Current Directions in Psychological Science*, *29*(3), 309–315. <https://doi.org/10.1177/0963721420915861>
- Tooley, U. A., Bassett, D. S., & Mackey, A. P. (2021). Environmental influences on the pace of brain development. *Nature Reviews Neuroscience*, *22*(6), 372–384. <https://doi.org/10.1038/s41583-021-00457-5>
- VanDam, M., Warlaumont, A. S., Bergelson, E., Cristia, A., Soderstrom, M., De Palma, P., & MacWhinney, B. (2016). HomeBank: An online repository of daylong child-centered audio recordings. *Seminars in Speech and Language*, *37*(2), 128–142. <https://doi.org/10.1055/s-0036-1580745>
- Viswanathan, G. M., Da Luz, M. G. E., Raposo, E. P., & Stanley, H. E. (2011). *The physics of foraging*. Cambridge University Press.
- Wozniak, R. H., Leezenbaum, N. B., Northrup, J. B., West, K. L., & Iverson, J. M. (2016). The development of autism spectrum disorders: Variability and causal complexity. *WIREs Cognitive Science*, *8*(1–2), Article e1426. <https://doi.org/10.1002/wcs.1426>