ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Current Directions in Psychological Science
1–9

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/09637214211037665
www.psychologicalscience.org/CDPS
SAGE

The Development of Communication Across Timescales

Elise A. Piazza[1,2,3], Mira L. Nencheva[1], and Casey Lew-Williams[1]

[1]Department of Psychology, Princeton University; [2]Princeton Neuroscience Institute, Princeton University; and [3]Department of Brain & Cognitive Sciences, University of Rochester

Abstract

How do young children learn to organize the statistics of communicative input across milliseconds and months? Developmental science has made progress in elucidating how infants learn patterns in language and how infant-directed speech is engineered to ease short-timescale processing, but less is known about how children link perceptual experiences across multiple levels of processing within an interaction (from syllables to stories) and across development. In this article, we propose that three domains of research—statistical summary, neural processing hierarchies, and neural coupling—will be fruitful in uncovering the dynamic exchange of information between children and adults, both in the moment and in aggregate. In particular, we discuss how the study of brain-to-brain and brain-to-behavior coupling between children and adults will advance the field's understanding of how children's neural representations become aligned with the increasingly complex statistics of communication across timescales.

Keywords

communication, development, hierarchies, neural coupling, statistics, timescales

When adults and young children communicate, they exchange information across milliseconds, seconds, and minutes. Statistics of these exchanges accumulate through diverse interactions across hours, days, and months and have long-lasting consequences for children's cognition. Children are tasked not only with integrating communicative input across the set of shorter timescales from milliseconds to minutes (e.g., connecting related words into meaningful sentences and narratives), but also with aggregating experiences across many interactions (Altmann, 2017; Gogate & Hollich, 2010; McMurray, 2016). However, developmental experiments have often focused on relatively short-timescale processing and learning (e.g., sounds, words, and sentences), rather than on the integration of information into a larger narrative context and across multiple naturalistic interactions.

In this article, we call for the merging of three complementary frameworks—statistical summary, neural processing hierarchies, and neural coupling—that will enable exciting progress in understanding how young children organize the statistics of their communicative input across timescales. We primarily focus on how recent theoretical and methodological advances in neuroscience and psychology can provide insight into children's integration of complex, naturalistic input within single interactions (i.e., across timescales of milliseconds, seconds, and minutes). This confluence of ideas will facilitate novel lines of scientific inquiry into how children integrate lower-level information into gistlike summaries (e.g., how they aggregate syllables, words, and sentences into concepts and narratives), how adults package information when speaking to children in a way that facilitates this integration, and how patterns of adult-child coupling support the development of communication, including language processing, improvisatory play, and collaborative problem solving. In the final section, we propose an expansion of these approaches to the study of long-term development, which unfolds over timescales of days to years.

Corresponding Author:
Elise A. Piazza, Department of Brain & Cognitive Sciences, University of Rochester
Email: elise.piazza@gmail.com

Integrating the Statistics of Communicative Input Across Levels of Complexity and Across Time

To learn language from complex, multisensory input, infants must extract statistical representations over time. This process has been described as a form of invariance detection (Gogate & Hollich, 2010) in which infants pick up on relatively stable patterns in their caregivers' input. Models of this process (Altmann, 2017) suggest that infants accumulate knowledge by gradually transforming the sensorimotor details of individual episodes (e.g., the word "kitty" spoken with different prosodic contours and referring to various real, stuffed, and cartoon cats) into higher-order statistical representations. Such models help explain widespread evidence that infants infer novel word boundaries via transitional probabilities between syllables (Gómez, 2002; Saffran, 2020), aggregate across multiple individually ambiguous trials to learn word-referent mappings (Smith & Yu, 2008), and use the distributional statistics of speech sounds to guide their perception (Maye et al., 2002).

These characterizations of infants' processing and learning dovetail nicely with the distinct but highly related literature on statistical-summary processing in adult perception (Whitney & Yamanashi Leib, 2018; Zhao et al., 2011), which has primarily been studied at very short timescales (milliseconds to seconds). For example, adults can precisely estimate the average pitch of a sequence of auditory tones, even when they struggle to report information about individual tones (Piazza et al., 2013). This compression of local details of input into a more abstract, compact representation, or gist, is thought to contribute to the efficient recognition and retention of sounds (McDermott et al., 2013).

Unlike typical experiments on infants' and adults' statistical processing, everyday interactions are not neatly divided into learning and test phases. Rich, real-life learning requires a constant formation of summary representations in real time and across parallel levels of processing. To navigate the real-time dynamics of speech comprehension and production during early interactions with caregivers, infants and toddlers must perform statistical computations for each of several features (e.g., prosody, semantic meaning) and timescales in parallel, while integrating information across these levels (Fig. 1a). Whereas the pitch height of an individual word might provide a cue to its momentary salience, pitch variability across a sentence might indicate a parent's emotional tone. Similarly, the average semantic-feature vector across a set of descriptions of a character contributes to the character's overall identity by the end of a conversation or story.

A developmental understanding of this computation and integration of statistics across multiple levels of processing would benefit from neuroscientific investigations of the representations of summary statistics during real-life communicative exchanges. For instance, whereas the auditory cortex represents the local acoustic features of speech with relatively fine temporal detail (e.g., pitches of individual syllables, differences between the sounds "cat" and "kitty"), regions further downstream must integrate statistics over longer timescales (e.g., the difference in overall contour between questions and sentences), and even higher-order regions likely suppress these acoustic details in the service of higher-level semantic representations that unfold over longer timescales (e.g., a story arc about a lost cat summarized across many local cat-related words and phrases; Lerner et al., 2011). However, the neural encoding of these representations at increasingly long timescales—and how they collapse relatively fine details in long-term memory—is underexplored. Experimental paradigms that track children's integration of such gist-like summaries over multiple timescales will provide a useful model of language processing that spans low-level perceptual averaging and higher-level summary representations of communicative information. Recent models of how the adult brain hierarchically processes the structure of communicative input across multiple timescales, discussed in the next section, could provide a key framework for tracking this integration of statistics at different developmental stages.

Hierarchical Processing of Communicative Statistics in the Brain

The auditory systems of many communicative species are structured hierarchically, reflecting organization of natural sounds from simple to relatively complex (Margoliash & Fortune, 1992). In humans, these levels likely result from parallel computations of incoming speech at shorter timescales (e.g., sounds and syllables) and longer timescales (sentences and entire narrative arcs). Recent neuroscience research has echoed hierarchical models of language processing (e.g., McClelland & Elman, 1986) in illuminating how the adult brain hierarchically organizes the successive building blocks of language and communication.

To understand how the brain processes information at different timescales, researchers have used experimental designs that disrupt meaning at one or more timescales while preserving meaning in others. One magnetoencephalography (MEG) study using this approach (Ding et al., 2016) found that distinct frequencies of neural activity entrain to different timescales of

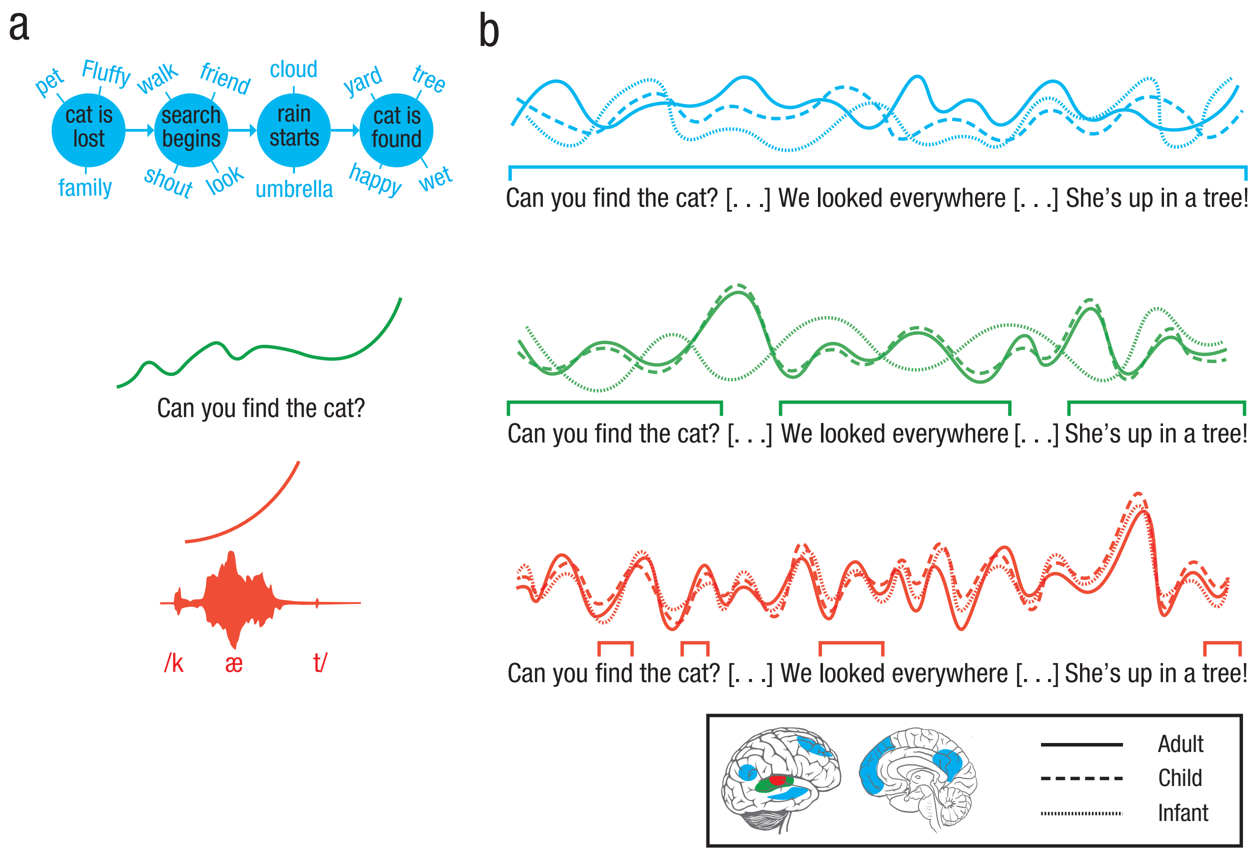

Fig. 1. Schematic diagrams of three levels of communicative processing, described in terms of (a) statistical summary and (b) neural hierarchies. The diagrams in (a) show statistical-summary representations of the word "cat" in the context of a story about a lost cat. At the single-word level (bottom row), "cat" is represented in terms of its acoustic features (e.g., pitch contour) and component phonemes. At the sentence level (middle row), the word is integrated into its nearby context, including the surrounding words and their pitch contour (which indicates a question in this example). At the narrative level (top row), the word is processed in terms of the full story arc comprising four events; the gist of each event is summarized in a circle, surrounded by the local details of that event. The diagrams in (b) illustrate neural entrainment to each of the three levels of processing in three age groups: adults, children, and infants. Auditory regions, such as primary auditory cortex (bottom row; red) and superior temporal gyrus (middle row; green), process fast dynamics of language (syllables, words) and might be fairly well synchronized across these age groups. However, the higher-order default-mode network (top row; blue) integrates over longer timescales, and the regions in this network might be synchronized across multiple adults but not across the three age groups.

meaningful linguistic information (words, phrases, sentences), and coupling between these frequencies is thought to coordinate information flow between brain regions that process speech at different levels (Giraud & Poeppel, 2012). Related MEG research has extended this hierarchy to characterize the transition from acoustic to lexical representations (Brodbeck et al., 2018).

Experiments using functional MRI have incorporated longer, more naturalistic stimuli and determined the contributions of different brain regions to each timescale of processing. In one study (Lerner et al., 2011), adults listened to a spoken story that had been scrambled at the word, sentence, or paragraph level. The between-subjects reliability (measured using intersubject temporal correlation; Hasson et al., 2004) of the responses of each brain region revealed its processing of specific timescales. The results showed a nested neural hierarchy for processing these three levels of complexity. Primary auditory cortex responded reliably across subjects even when words were scrambled, which indicates that it processes the local, moment-to-moment details of speech. In contrast, regions of the higher-order default-mode network (e.g., precuneus, frontal cortex) responded reliably only when paragraphs (and not words) were scrambled or when the story was intact. This hierarchy extends to musical structure as well (Farbood et al., 2015), which suggests that it supports the extraction of information over windows of various durations in domains beyond language. In adults, one important feature of brain regions that process input at long timescales is that they represent holistic, gistlike information, and thus respond similarly even when the low-level details of a stimulus are changed (e.g., when a story is translated into a bilingual's other language; Honey et al., 2012; Yeshurun et al., 2021).

Little is known about how children accumulate complex details into a narrative or concept. Expanding the range of timescales in studies of children's language processing will position researchers to learn how the richness and structure of children's representations evolve over time (see Fig. 1b for a schematic example of infants', children's, and adults' neural entrainment to information at the word, sentence, and narrative levels). For example, it may be the case that the default-mode network processes only shorter-timescale speech input early in infancy and then gradually converges onto mature, longer-timescale narrative-level representations.

Examining intersubject correlations between multiple children's brain responses to stimuli whose meaning is disrupted at different timescales will help show how children progress from representing smaller to larger units of language. Neural-decoding approaches have the power to reveal the richness of children's representations of input (e.g., patterns of voxel activation reflecting the shape of a toy, its category identity, or its overall role in a story), thereby providing insights into the processing of information at different levels of complexity that cannot be gained by observing children's behaviors. It would also be useful to examine how neural hierarchies vary across children and across age groups, and whether this variability is meaningfully related to the deployment of different processing mechanisms (e.g., prediction) in different contexts. Behavioral paradigms may be informative regarding how infants' attention and memory systems become increasingly capable of tracking hierarchically nested information (e.g., Bauer & Mandler, 1989; Rosenberg & Feigenson, 2013) and how parents package their language to help children build representational complexity (Schwab & Lew-Williams, 2016). Given that adults must ultimately communicate information across several representational levels to children, new measures of the development of neural processing hierarchies will elucidate how adults and children *jointly* coordinate the exchange of information across timescales.

Coupling Provides Insights Into the Real-Time Transfer of Representations Across Timescales

As infants and toddlers exchange information with their caregivers in real time and across time, they somehow progress toward mature hierarchical representation of words, sentences, and narratives. What dynamic adjustments do the two parties make in order to align their representations of communicative content? In particular, how do caregivers accommodate the limitations of infant cognition? For example, when two adults see a dark cloud, their shared understanding of weather allows them to predict that rain is likely. Although an adult and infant may have shared perceptual representations of a dark cloud, the infant's brain may lack the knowledge to predict upcoming rain, which relies on rich semantic associations stored in long-term memory. The adult will tend to provide scaffolds for such predictive leaps for the infant (e.g., "Uh-oh! I think we need an umbrella!"). Many related instances of such experiences over time may help the infant build longer-timescale representations (see Fig. 2).

Coupled interactions between infants and caregivers have been investigated behaviorally for a long time, and this work has yielded many insights about how caregivers tailor their speech and how infants actively contribute to multimodal communicative exchanges (Fernald et al., 1989; Piazza et al., 2017; Schwab & Lew-Williams, 2016; Soderstrom, 2007). Caregivers tailor their communication in ways that are initially optimized for shorter-timescale processing (McMurray, 2016), but over time they increase the complexity of their words and utterances (Schwab et al., 2018). They also change their behavior in response to infants' attentional focus; for example, they provide appropriately timed labels (Suanda et al., 2019) and try to align attention onto the same object (Suarez-Rivera et al., 2019). Over time, caregivers' and infants' behaviors become increasingly contingent on each other both within language (Abney et al., 2017; Hirsh-Pasek et al., 2015) and across domains, such that their gestures, gaze directions, and speech acts influence one another in a back-and-forth manner (Goldstein & Schwade, 2008; Gros-Luis et al., 2006). Thus, successful communication relies on mirrored behaviors and representations, as well as on nonmirrored, contingent responses.

Whereas behavioral coupling highlights the interplay between adults' and children's outward actions, neural (brain-to-brain) coupling provides unique access to the alignment of inner mental representations that are not always behaviorally measurable, especially in prelinguistic infants. In adults, neural coupling between a speaker and listener during storytelling predicts the listener's comprehension of the story, thus providing a measure of communication success and information transfer (Stephens et al., 2010). As in research on neural hierarchies, neural coupling in regions that support long-timescale processing reflects shared high-level understanding of language and cannot be explained

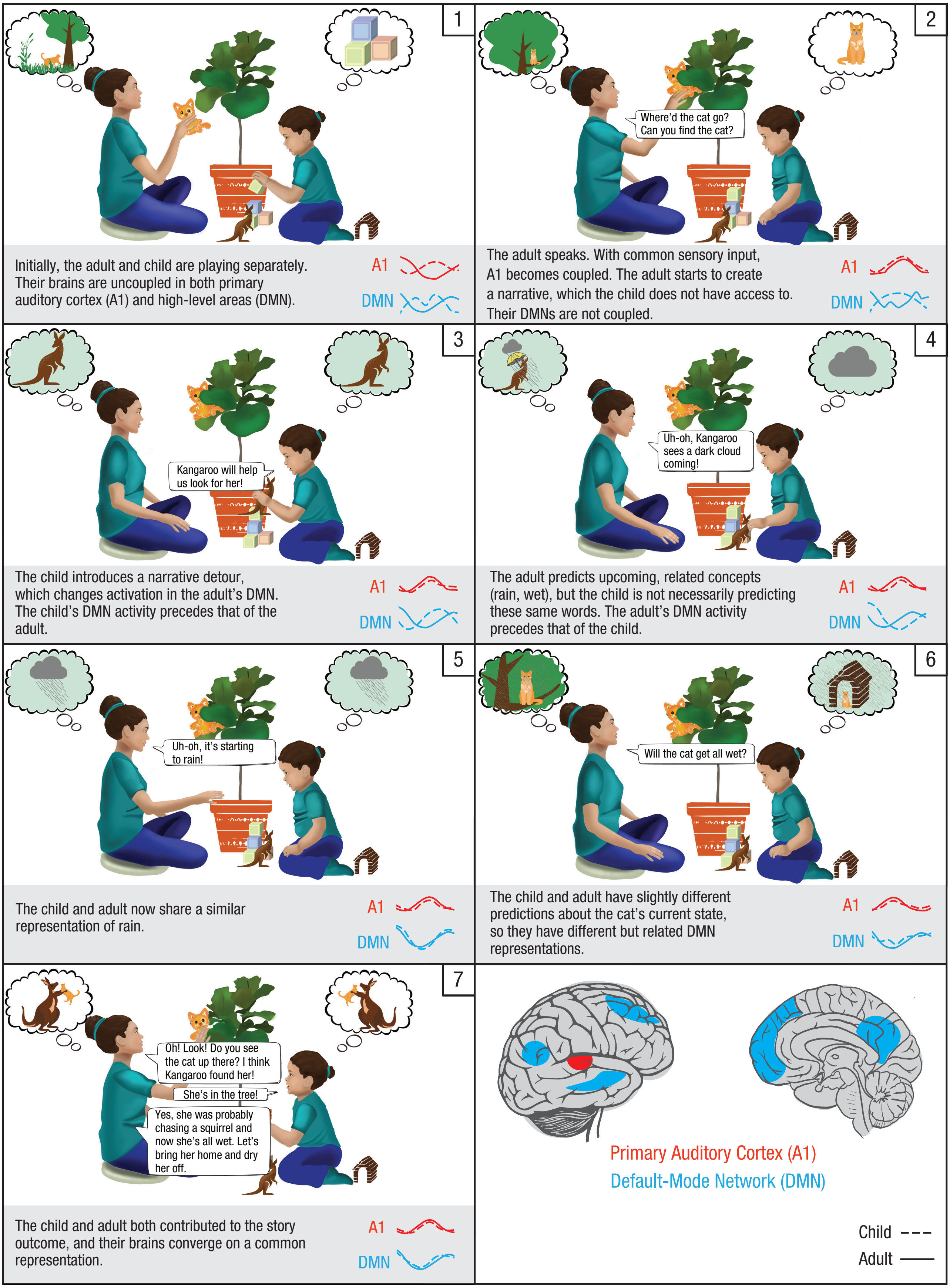

Fig. 2. *(continued on next page)*

Fig. 2. Cartoon example of a communicative interaction between a mother and child during real-life play. The red and blue curves at the bottom of each panel depict possible neural time series from one early sensory brain region (primary auditory cortex) and a network of higher-order regions (default-mode network). As the mother and child progress through the interaction, they flow through several states of neural coupling in each region. When they are not interacting (and there is no shared sensory input), their brains are uncoupled in both regions. When they are hearing the same speech, or viewing the same object, shared input drives sensory coupling. Frequently, the adult anticipates predictable content before the child does because the adult has access to richer semantic associations and narrative schemas than the child. This happens both at relatively short timescales (e.g., a dark cloud will be followed by rain) and at longer ones (e.g., a canonical ending to the "lost cat" schema). Sometimes the child reroutes the conversation in a surprising way (e.g., a kangaroo will help find the lost cat), and the adult adapts to this detour. These are all examples of leader-follower dynamics, facilitated by behaviors that guide the other person toward a joint state of understanding. Whenever the adult and child converge on that joint state (e.g., they represent rain in a related way), there is mirrored coupling between them. By the end of the interaction, each person has dynamically adapted to the other to create this story, so the interaction as a whole represents synergistic coupling.

simply as the processing of the same sensory input at the same time (Honey et al., 2012; Piazza et al., 2020; Yeshurun et al., 2017). Furthermore, it has been proposed that coupling plays a mechanistic role in learning by ensuring that the learner's brain enters a phase of high excitability during moments that are optimal for encoding information (Wass et al., 2020).

Neural coupling has been measured using multiple modalities (functional MRI, functional near-infrared spectroscopy, electroencephalography) both between individuals listening to the same story in different sessions (e.g., Liu et al., 2017; Stephens et al., 2010) and between members of a dyad engaging in an interactive task during the same session (e.g., Piazza et al., 2020; see Wass et al., 2020, for a review of adult-child studies). Although most studies of adult-child neural coupling have focused on mirrored synchrony (one-to-one alignment of the neural dynamics between two individuals), it is not always optimal for an adult's brain and a child's brain to process the same information in the same way at the same time (e.g., the child's creative narrative detour in the third panel of Fig. 2). Many interactions also involve leading and following (Piazza et al., 2020) or synergistic, mutual adaptation between two individuals (Hasson & Frith, 2016); such patterns of nonmirrored coupling are likely to support improvisational aspects of creative play and problem solving.

The study of neural coupling will illuminate the ways in which adults and children align to each other across different timescales. For example, during communicative interactions, infants' relatively short-timescale processing may support the formation of sound categories and speech segmentation, whereas adults' broad ability to process information at longer timescales may facilitate sentence-level predictions and semantic understanding of narrative arcs. If so, there may be a pattern of progression from synchrony in early sensory regions to nonmirrored, leader-follower dynamics in higher-order regions (likely with moments of weak neural alignment). The convergence of children's neural representations with adults' could proceed linearly across development or could be linked to particular milestones, such as spikes in vocabulary acquisition, vocal production, or perspective taking. Another ripe area for future research is the contingency of infants' neural representations on adults' behaviors. At the neural level, how does high-quality, temporally contingent feedback from an adult improve the sophistication of children's communicative output in real time (Goldstein et al., 2003)? Does the joint pattern of coupling within a parent-child dyad (e.g., leader-follower, contingency, mutual adaptation) predict communicative success or learning more effectively than each individual's brain representations? Such investigations will require a widening of the temporal and spatial windows of analysis of coupling, to account for differences in the timing of adults' and children's neural processes, as well as in the brain regions performing complementary communicative functions at different stages.

Development as a Natural Model of Long-Timescale Integration

We have suggested that new, naturalistic experimental paradigms will provide insights into how—within a single interaction—adults and children share their representations of the world, which inherently span different timescales of processing. The longest timescale we have discussed thus far maps onto extraction of the overall narrative arc of a story or conversation. However, the process of development itself provides the ultimate model of truly long-timescale integration, as children must actively learn to communicate over the course of thousands of interactions. This development-level timescale is the most challenging to study and will benefit from the creative merging of approaches and fields. For example, how do different brain regions perform statistical-summary computations to integrate over diverse instances of a word, separated in time and space and articulated by multiple people? How do these computations enable a child to form a rich, unified, and usable representation of a concept? How do the types of invariance that adults emphasize—and children attend to—coevolve throughout childhood? Development-level investigations will ultimately expand the definition of the neural processing hierarchy beyond within-interaction timescales to include learning and communication across hours, days, and even years, and in doing so may reveal exciting discoveries about the neural processes that support long-term integration. This type of investigation will advance the field's understanding of a range of cognitive systems; the literature on the development of memory, for example, currently lacks explanations for how statistical learning is retained over timescales longer than an experimental session (Gómez, 2017; McMurray et al., 2012) and how the default-mode network communicates with the hippocampus during long-term learning.

Neuroimaging studies that capture fine-grained changes in patterns of brain activity over the course of development, rather than at a single time point in the lab, will serve as powerful tests of well-known models of learning. For example, dynamic systems models (Smith & Thelen, 2003) propose that subtle, short-timescale changes gradually move a system toward a destination, such as the first instance of walking or the production of a new word. The ability to "peer under the hood" throughout these processes will help to illuminate how changes in children's and adults' neural representations during coupled interactions contribute to advances in children's behavior. Finally, understanding the evolution of parent-child coupling during communicative interactions over developmental time could provide key insights about individual differences in learning, as well as in critical outcomes, such as school readiness.

Conclusion

In this article, we have called for the application and integration of three scientific frameworks to understand the development of children's ability to process information at multiple timescales. First, statistical summary may provide one mechanism through which a child learns to integrate experiences over time, such as by averaging word tokens across contexts into a semantic concept or forming the gist of a story across diverse words and events. Second, studying the development of neural processing hierarchies could help explain how children build up neural representations of communicative information that unfolds over time, from milliseconds to minutes. Third, interactive experiments measuring neural and behavioral coupling between adults and young children will illuminate how adults share their mental representations with children in real time, and how children contribute to their own learning by actively guiding adults' actions. By studying these phenomena across the longest timescale—human development itself—researchers will be positioned to build an integrative model of how the statistics of adult-infant communication give rise to children's learning outcomes.

Recommended Reading

Gogate, L. J., & Hollich, G. (2010). (See References). A discussion of infants' detection of statistical regularities in communicative input.

Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). (See References). An introduction to neural hierarchies for processing increasingly complex information.

McMurray, B. (2016). (See References). An overview of developmental processes that unfold across multiple timescales.

Piazza, E. A., Sweeny, T. D., Wessel, D., Silver, M. A., & Whitney, D. (2013). (See References). A report on experiments providing evidence of statistical-summary integration in adults' auditory perception.

Wass, S. V., Whitehorn, M., Haresign, I. M., Phillips, E., & Leong, V. (2020). (See References). A review of adult-child coupling during dyadic interactions.

Transparency

Action Editor: Amy Needham
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 work was supported by a Princeton University C. V. Starr Fellowship to E. A. Piazza and by a grant from the National Institute of Child Health and Human Development to C. Lew-Williams (R01HD095912).

ORCID iDs

Elise A. Piazza https://orcid.org/0000-0001-6729-8559
Mira L. Nencheva https://orcid.org/0000-0003-4854-4608

References

Abney, D. H., Warlaumont, A. S., Oller, D. K., Wallot, S., & Kello, C. T. (2017). Multiple coordination patterns in infant and adult vocalizations. *Infancy*, *22*(4), 514–539.

Altmann, G. T. M. (2017). Abstraction and generalization in statistical learning: Implications for the relationship between semantic types and episodic tokens. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), Article 20160060. https://doi.org/10.1098/rstb.2016.0060

Bauer, P. J., & Mandler, J. M. (1989). One thing follows another: Effects of temporal structure on 1- to 2-year-olds' recall of events. *Developmental Psychology*, *25*(2), 197–206. https://doi/org/10.1037/0012-1649.25.2.197

Brodbeck, C., Hong, L. E., & Simon, J. Z. (2018). Rapid transformation from auditory to linguistic representations of continuous speech. *Current Biology*, *28*(24), 3976–3983. https://doi.org/10.1016/j.cub.2018.10.042

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, *19*(1), 158–164. https://doi.org/10.1038/nn.4186

Farbood, M. M., Heeger, D. J., Marcus, G., Hasson, U., & Lerner, Y. (2015). The neural processing of hierarchical structure in music and speech at different timescales. *Frontiers in Neuroscience*, *9*, Article 157. https://doi.org/10.3389/fnins.2015.00157

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, *16*(3), 477–501. https://doi.org/10.1017/S0305000900010679

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511–517. https://doi.org/10.1038/nn.3063

Gogate, L. J., & Hollich, G. (2010). Invariance detection within an interactive system: A perceptual gateway to language development. *Psychological Review*, *117*(2), 496–516. https://doi.org/10.1037/a0019049

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences, USA*, *100*(13), 8030–8035.

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, *19*(5), 515–523. https://doi.org/10.1111/j.1467-9280.2008.02117.x

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*(5), 431–436. https://doi.org/10.1111/1467-9280.00476

Gómez, R. L. (2017). Do infants retain the statistics of a statistical learning experience? Insights from a developmental cognitive neuroscience perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), Article 20160054. https://doi.org/10.1098/rstb.2016.0054

Gros-Louis, J., West, M. J., Goldstein, M. H., & King, A. P. (2006). Mothers provide differential feedback to infants' prelinguistic sounds. *International Journal of Behavioral Development*, *30*(6), 509–516. https://doi.org/10.1177/0165025406071914

Hasson, U., & Frith, C. D. (2016). Mirroring and beyond: Coupled dynamics as a generalized framework for modelling social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1693), Article 20150366. https://doi.org/10.1098/rstb.2015.0366

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science*, *303*(5664), 1634–1640.

Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Tresch Owen, M., Golinkoff, R. M., Pace, A., Yust, P. K. S., & Suma, K. (2015). The contribution of early communication quality to low-income children's language success. *Psychological Science*, *26*(7), 1071–1083.

Honey, C. J., Thompson, C. R., Lerner, Y., & Hasson, U. (2012). Not lost in translation: Neural responses shared across languages. *Journal of Neuroscience*, *32*(44), 15277–15283.

Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *Journal of Neuroscience*, *31*(8), 2906–2915. https://doi.org/10.1523/JNEUROSCI.3684-10.2011

Liu, Y., Piazza, E. A., Simony, E., Shewokis, P. A., Onaral, B., Hasson, U., & Ayaz, H. (2017). Measuring speaker–listener neural coupling with functional near infrared spectroscopy. *Scientific Reports*, *7*, Article 43292. https://doi.org/10.1038/srep43293

Margoliash, D., & Fortune, E. S. (1992). Temporal and harmonic combination-sensitive neurons in the zebra finch's HVc. *Journal of Neuroscience*, *12*(11), 4309–4326.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), B101–B111. https://doi.org/10.1016/S0010-0277(01)00157-3

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86. https://doi.org/10.1016/0010-0285(86)90015-0

McDermott, J. H., Schemitsch, M., & Simoncelli, E. P. (2013). Summary statistics in auditory perception. *Nature Neuroscience*, *16*(4), 493–498. https://doi.org/10.1038/nn.3347

McMurray, B. (2016). Language at three timescales: The role of real-time processes in language development and evolution. *Topics in Cognitive Science*, *8*(2), 393–407.

McMurray, B., Horst, J. S., & Samuelson, L. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review*, *119*(4), 831–877.

Piazza, E. A., Hasenfratz, L., Hasson, U., & Lew-Williams, C. (2020). Infant and adult brains are coupled to the dynamics of natural communication. *Psychological Science*, *31*(1), 6–17.

Piazza, E. A., Iordan, M. C., & Lew-Williams, C. (2017). Mothers consistently alter their unique vocal fingerprints when communicating with infants. *Current Biology*, *27*(20), 3162–3167.

Piazza, E. A., Sweeny, T. D., Wessel, D., Silver, M. A., & Whitney, D. (2013). Humans use summary statistics to perceive auditory sequences. *Psychological Science*, *24*(8), 1389–1397.

Rosenberg, R. D., & Feigenson, L. (2013). Infants hierarchically organize memory representations. *Developmental Science*, *16*(4), 610–621.

Saffran, J. R. (2020). Statistical language learning in infancy. *Child Development Perspectives*, *14*(1), 49–54. https://doi.org/10.1111/cdep.12355

Schwab, J. F., & Lew-Williams, C. (2016). Language learning, socioeconomic status, and child-directed speech. *WIREs*

Cognitive Science, *7*(4), 264–275. https://doi.org/10.1002/wcs.1393

Schwab, J. F., Rowe, M. L., Cabrera, N., & Lew-Williams, C. (2018). Fathers' repetition of words is coupled with children's vocabularies. *Journal of Experimental Child Psychology*, *166*, 437–450. https://doi.org/10.1016/j.jecp.2017.09.012

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–1568. https://doi.org/10.1016/j.cognition.2007.06.010

Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive Sciences*, *7*(8), 343–348. https://doi.org/10.1016/S1364-6613(03)00156-6

Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review*, *27*(4), 501–532.

Stephens, G. J., Silbert, L. J., & Hasson, U. (2010). Speaker–listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences, USA*, *107*(32), 14425–14430. https://doi.org/10.1073/pnas.1008662107

Suanda, S. H., Barnhart, M., Smith, L. B., & Yu, C. (2019). The signal in the noise: The visual ecology of parents' object naming. *Infancy*, *24*(3), 455–476.

Suarez-Rivera, C., Smith, L. B., & Yu, C. (2019). Multimodal parent behaviors within joint attention support sustained attention in infants. *Developmental Psychology*, *55*(1), 96–109.

Wass, S. V., Whitehorn, M., Haresign, I. M., Phillips, E., & Leong, V. (2020). Interpersonal neural entrainment during early social interaction. *Trends in Cognitive Sciences*, *24*(4), 329–342.

Whitney, D., & Yamanashi Leib, A. (2018). Ensemble perception. *Annual Review of Psychology*, *69*, 105–129.

Yeshurun, Y., Nguyen, M., & Hasson, U. (2021). The default mode network: Where the idiosyncratic self meets the shared social world. *Nature Reviews Neuroscience*, *22*(3), 181–192. https://doi.org/10.1038/s41583-020-00420-w

Yeshurun, Y., Swanson, S., Simony, E., Chen, J., Lazaridi, C., Honey, C. J., & Hasson, U. (2017). Same story, different story: The neural representation of interpretive frameworks. *Psychological Science*, *28*(3), 307–319. https://doi.org/10.1177/0956797616682029