Biological Psychiatry

Connecting Environmental and Genetic Explanations for Nonlinear Language Development in Infancy and Toddlerhood

Federica Bulgarelli

Language development is not linear, regardless of the lens through which it is examined. Across both comprehension (i.e., receptive vocabulary) and production (i.e., expressive vocabulary, grammar), development is characterized by slow and gradual progress followed by sudden spurts. For example, comprehension is thought to improve drastically around 14 months of age; while infants exhibit knowledge of a few words prior to this time, their word recognition quickly improves after this point (1). Expressive vocabulary knowledge is also slow at first; typically developing infants go from knowing just a few words to knowing at least 50, but typically more than 100, between 18 and 24 months of age (2). Grammatical development follows a similar trajectory. Infants begin producing words in isolation and begin combining them into simple sentences as they start learning verbs and prepositions. Exactly what leads to these seemingly sudden and measurable changes in language development is of interest to researchers across disciplines.

In the current issue of *Biological Psychiatry*, Verhoef *et al.* (3) investigated the genetic basis of receptive and productive vocabulary at 2 time points spanning these periods of nonlinear growth: the early phase (infancy, 15–18 months) and the later phase (toddlerhood, 24–38 months). They found that while all vocabulary measures spanning 15 to 38 months of age were moderately heritable, genetic correlations between early-phase expressive and later-phase receptive vocabulary were negligible. Such data suggest the existence of distinct genetic influences operating at these different time periods. In addition, while expressive vocabulary in both the early phase and the later phase was genetically linked to literacy-related measures, only later-phase expressive and receptive vocabulary were genetically correlated with intelligence and educational attainment across the lifespan. Thus, while both early and later language development milestones are important for lifelong cognitive and educational functioning, Verhoef *et al.* suggest that they may stem from partially distinct genetic bases.

If partially distinguishable sets of common genetic variants influence infant and toddler vocabulary skills, we would predict that spurts in language knowledge around toddlerhood might be related to mechanisms that may not be available to, or used by, learners until then. This possibility is consistent with the literature on constraints used for language learning. Before 18 months of age, children are thought to acquire words through a brute-force approach, learning word–object pairs through slow associative learning mechanisms. Around 18 months of age, infants are thought to begin making use of learning constraints, such as mutual exclusivity, to launch their word learning. Mutual exclusivity, a learner's assumption that objects should have one label and labels should refer to one object, can help infants quickly learn new labels for objects for which they lack a label, without needing to have them named directly. While it remains an open question whether mutual exclusivity does not develop or is just not used until 18 months of age, the fact that learners use this mechanism to boost word learning only after 18 months could be consistent with the idea that heritable biological mechanisms emerge during this period that are partially distinct from those that govern associative learning mechanisms supporting earlier word learning (4,5).

Another knowledge gap from the language development literature that is potentially relevant to the idea of distinct genetic influences on language development in infancy versus toddlerhood relates to late talkers. Late talkers are children 18 to 30 months of age who exhibit typical receptive language skills but delayed expressive milestones (e.g., do not produce at least 50 words by 24 months of age) yet go on to catch up and not receive later language disorder diagnoses. While some demographic factors such as preterm birth and socioeconomic status increase the chances of late language emergence at the population level, research to date has not been able to determine exactly why some children exhibit these short-lived expressive language challenges (6). Given the results of Verhoef *et al.* (3), late talkers could show delays in expressive language skills due to individual differences in when and how genes underlying these skills come online. Incorporating the examination of genetic variation into studies of late talkers could provide further evidence for this possibility.

Given evidence that genetic contributions do not explain the entirety of language development, it is also useful to consider the role of environmental experiences. Studies of environmental influences have taken different approaches to quantifying language input and relating language exposure to word learning. A recent meta-analysis found that both language quantity (e.g., number of words) and quality (e.g., vocabulary diversity and syntactic complexity) predicted children's language outcomes (7), suggesting that 20% to 30% of the variance in language outcomes can be explained by properties of linguistic input. Looking at more specific factors, in a recent analysis predicting the age at which infants will say specific words, we found that how often individual words were heard accounted for 7% of the variance, and how variable those words sounded acoustically accounted for an additional 5% of the variance (8). Thus, experiences have a sizable effect on language outcomes, when measured both at the vocabulary level and at the individual word level.

A potential challenge to the effects of environment on language development is that it is often impossible to separate

SEE CORRESPONDING ARTICLE ON PAGE 859

https://doi.org/10.1016/j.biopsych.2024.03.001
ISSN: 0006-3223

genetic from environmental influences. That is, biological parents are typically contributing both genes and language input to their children. Since genes relate to linguistic behaviors [as shown by Verhoef *et al.* (3)], parents could be more talkative because of their own genes, which they have passed on to their children who will therefore also be more talkative. Coffey *et al.* (9) tried to tease apart the effects of genes and the environment by studying the effects of parental language input for a group of adoptee children, who therefore do not share parental genes. They found that parental language input is still correlated with vocabulary growth, even in the absence of genetic confounds.

Environment matters, and marked changes in language skills around 18 months of age could be due to children's environments changing to support new language learning strategies. However, longitudinal analyses of infants' linguistic input suggest remarkable consistency in environmental input over time. Bergelson (1) showed that from 6 to 17 months of age, noun input to infants remained highly stable (though there was high individual variance), regardless of whether the analysis looked at the total number of nouns, how easy it was to identify what object was being referred to, and the syntactic frames that noun occurred in. And yet, infant's in-lab comprehension improved drastically around 14 months. Thus, it is likely something about the learner, rather than the environment, that changes. In line with Verhoef *et al.* (3), one thing that might change is the underlying biology that supports infants' ability to make better use of the information at their disposal.

One limitation of the Verhoef *et al.* (3) study is that the authors only included children of European descent. This limitation is common in genome-wide studies, and the field is actively seeking to address this issue. Nonetheless, the sample composition limits the generalizability of their conclusions. The issue of representation has been an increasing focus in the behavioral language development literature as well, as the majority of the field's findings to date have also relied on Western, educated, industrialized, rich and democratic samples. Recent efforts to combat this issue have involved extending particularly naturalistic recording methods to less represented communities. As a concrete example, Casillas *et al.* (10) measured linguistic input to children in a Tzeltal village in Southern Mexico and found that children's linguistic input varied considerably from that of North American infants often studied, in that children rarely received speech directed to them (termed child-directed speech). Despite these differences, many aspects of language development remain consistent; for example, children in the Casillas *et al.* sample met milestones such as production of their first word at approximately the same time as North American infants. Thus, while child-directed speech is thought to be an important factor for language learning in North American samples, it is certainly not required for typical language development. Many different experiences thus support language learning, and these may vary cross-culturally and across racial and ethnic groups. Given that experience and genetics interact, it will be important to explore whether differences in the genes that underlie language development at different ages also vary as a function of ancestry.

In summary, Verhoef *et al.* (3) provide novel insights into how different genes may shape language development at earlier and later phases of development. These findings are consistent with the literature on language development showing nonlinear improvements around toddlerhood across a variety of measurements. Broadly surveying the literature on the effects of the environment on language development highlights the importance of considering both genetic and environmental influences in understanding individual language learning trajectories. While it may be challenging, future research should consider investigating the genetic and environmental bases of language development together, allowing, for example, elucidation of whether variable effects of language input on language outcomes are mediated or boosted by genetics. Combined with more representative samples, these results set the stage for better understanding rapid changes in language development in early childhood.

Acknowledgments and Disclosures

Early Career Investigator Commentaries are solicited in partnership with the Education Committee of the Society of Biological Psychiatry. As part of the educational mission of the Society, all authors of such commentaries are mentored by a senior investigator. This work was mentored by Alysa Doyle, Ph.D.

The author reports no biomedical financial interests or potential conflicts of interest.

Article Information

From the Department of Psychology, Department of Learning and Instruction, University at Buffalo, Buffalo, New York.

Address correspondence to Federica Bulgarelli, Ph.D., at fbulgare@buffalo.edu.

Received Feb 26, 2024; revised Mar 3, 2024; accepted Mar 4, 2024.

References

1. Bergelson E (2020): The comprehension boost in early word learning: Older infants are better learners. Child Dev Perspect 14:142–149.
2. Bloom P (2001): Roots of word learning. In: Bowerman M, Levinson S, editors. Language Acquisition and Conceptual Development. Cambridge: Cambridge University Press, 159–181.
3. Verhoef E, Allegrini AG, Jansen PR, Lange K, Wang CA, Morgan AT, *et al.* (2024): Genome-wide analyses of vocabulary size in infancy and toddlerhood: Associations with attention-deficit/hyperactivity disorder, literacy, and cognition-related traits. Biol Psychiatry 95:859–869.
4. Markman EM (2014): Constraints on word learning: Speculations about their nature, origins, and domain specificity. In: Gunnar MR, Maratsos M, editors. (2014), Modularity and Constraints in Language and Cognition. Washington, DC: American Psychological Association Press, 59–102.
5. Bohn M, Tessler MH, Merrick M, Frank MC (2021): How young children integrate information sources to infer the meaning of words. Nat Hum Behav 5:1046–1054.
6. Rescorla L (2011): Late talkers: Do good predictors of outcome exist? Dev Disabil Res Rev 17:141–150.
7. Anderson NJ, Graham SA, Prime H, Jenkins JM, Madigan S (2021): Linking quality and quantity of parental linguistic input to child language skills: A meta-analysis. Child Dev 92:484–501.
8. Bulgarelli F, Bergelson E (2023): Linking acoustic variability in the infants' input to their early word production. PsyArXiv. https://doi.org/10.31234/osf.io/su768.
9. Coffey JR, Shafto CL, Geren JC, Snedeker J (2022): The effects of maternal input on language in the absence of genetic confounds: Vocabulary development in internationally adopted children. Child Dev 93:237–253.
10. Casillas M, Brown P, Levinson SC (2020): Early language experience in a Tseltal Mayan village. Child Dev 91:1819–1835.