Cognitive Neuroscience

Current Debates, Research & Reports

ISSN: 1758-8928 (Print) 1758-8936 (Online) Journal homepage: www.tandfonline.com/journals/pcns20

Visual cortex through the lens of language

Haojie Wen & Yanchao Bi

COGNITIVE NEUROSCIENCE
https://doi.org/10.1080/17588928.2025.2590661

Routledge
Taylor & Francis Group

COMMENTARY

Visual cortex through the lens of language

Haojie Wen[a] and Yanchao Bi[b,c,d,e]

[a]School of Systems Science, Beijing Normal University, Beijing, China; [b]School of Psychological and Cognitive Sciences, Peking University, Beijing, China; [c]IDG/McGovern Institute for Brain Research, Peking University, Beijing, China; [d]Institute for Artificial Intelligence, Peking University, Beijing, China; [e]Beijing Key Laboratory of Behavior and Mental Health & Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing, China

ABSTRACT

How is the visual cortex organized? Ritchie et al. (this issue) argue for moving beyond category-selective accounts toward an emphasis on complex, behaviorally relevant functions—a perspective we fully endorse. Extending this view, we emphasize that human behaviors are diverse and differentially prioritized. Among these, language emerges as a uniquely critical domain. Converging evidence from developmental and cognitive neuroscience demonstrates that language exerts a powerful influence on visual processing. These findings underscore the need to incorporate language—alongside other high-priority behaviors—into frameworks seeking to elucidate the organizational principles of the human visual cortex.

ARTICLE HISTORY
Received 2 October 2025
Accepted 11 November 2025

KEYWORDS
Visual cortex; category selective; behavior; language

What are the organizational principles of the visual cortex? In addressing this question, we fully endorse the perspective articulated by Ritchie et al. (this issue), which emphasizes the behavioral relevance of visual properties in real-world environments. Building on this framework, they argue that the visual cortex is fundamentally organized to optimize and coordinate natural behavior. They further note that the visual cortex must integrate complex visual inputs with behavioral goals, relying on coding that is globally distributed yet locally sparse. We also concur that category selectivity is inherently limited, as what forms a relevant category here is not always clear, and the currently identified categorical representations cannot capture the richness of real-world environments or the diversity of behavioral goals. Supporting this view, we have also shown, by combining computational vision models, parametric-modulation fMRI, and natural image statistics, the organization of visual features in the ventral occipitotemporal cortex (VOTC) reflects their relevance to distinct behavioral computations—including fight-or-flight, navigation, and object manipulation—rather than to object categories (Fan et al., 2021). Together, these findings highlight the productivity in shifting from categorical profiling to emphasis on behavioral relevance to advance our understanding of the organizational principles of the visual cortex.

Then the key question becomes how behaviors or goals should be considered and prioritized with respect to visual cortex computations, which goes back to a general challenge to behaviorism. Humans engage in a wide spectrum of activities, ranging from basic biological functions, such as eating and sleeping, to complex, culturally embedded behaviors, such as festival celebrations or the invention of sophisticated instruments. Even within a single scenario—for example, the 'walking a dog' case discussed in the review—an individual may either continue jogging past the dog or pause to interact with it. Presumably, the brain does not weigh all behaviors equally; rather, it allocates resources preferentially to behaviors essential for survival and adaptive functioning. Among various contexts of object use, for instance, using an object as a tool is a behavioral profile that is relatively human-distinct. Brain networks supporting tool processing have been identified involving the human-specific recruitment of the inferior parietal lobule (IPL), along with intrinsic connectivity patterns (Bi et al., 2015, 2016; Buxbaum et al., 2014; Chao & Martin, 2000; Kastner et al., 2017; Lewis, 2006; Peelen et al., 2013; Peeters et al., 2009; Wen et al., 2022). Incorporating such behaviorally prioritized actions into theoretical frameworks is therefore essential for developing a more comprehensive account of the organizational principles of the visual cortex.

Among the diverse repertoire of human behaviors, language is worth highlighting in particular. It functions not only as a primary tool for communication but might also act as a foundational mechanism that brings other cognitive processes to a common abstract relational

CONTACT Yanchao Bi ybi@pku.edu.cn School of Psychological and Cognitive Sciences, Peking University, Beijing 100871, China

space (Morgan et al., 2015; Thibault et al., 2021). Human language is highly complex, supported by a frontal-parietal network, shaped by genetic influences, and exhibits a degree of human specificity (Fedorenko et al., 2010; Kong et al., 2020; Le Guen et al., 2018). Compelling developmental evidence indicates that language plays a prioritized role, compared to other cues, in shaping infants' categorization abilities (Balaban & Waxman, 1997; Ferguson & Waxman, 2017; Waxman & Markow, 1995). In a typical paradigm, infants are first familiarized with multiple exemplars of a category (e.g., animals). In the word condition, each exemplar is accompanied by a novel noun (e.g., 'look at the toma'); in the no-word condition, phrases direct attention to the objects but introduce no novel words (e.g., 'look at this'); and in the nonverbal cue conditions (e.g., tones), infants hear sine-wave tones (or bird vocal in Woodruff Carr et al., 2021; or speech with very different profiles from their native language in Perszyk & Waxman, 2019). During the test phase, infants are presented with two novel objects: one from the familiar category (e.g., a new animal) and one from an unfamiliar category (e.g., a fruit). Results show that 9-month-old infants successfully form categories only in the word condition, whereas those in the no-word or nonverbal conditions do not—demonstrating that language exerts a unique effect on categorization.

Consistent with these findings, recent computational studies have shown that models integrating language and vision (e.g., CLIP) show stronger alignment with human visual cortical activity than models relying exclusively on visual input (Chen et al., 2025; Wang et al., 2023). Patient studies indicate that neural representations in the sensory-derived ventral occipitotemporal cortex (VOTC) are modulated by disruption of its connections with the left dorsolateral anterior temporal lobe (LdlATL) within the language network (Liu et al., 2025). Together, these results suggest that language, as a key cognitive capacity, shapes specific functions of the visual cortex. Yet, it remains unclear how the visual-spatial encoding principles of the visual cortex are adjusted to support language processing, and how such adjustments are constrained by underlying connectivity.

In summary, indeed elucidating the organizational principles of the visual cortex requires moving beyond category-selective frameworks toward behaviorally grounded objectives. Human behavioral goals are inherently diverse and differentially prioritized, necessitating careful consideration of which behaviors shape cortical organization. In particular, language-related behaviors deserve focused investigation, both to clarify how the visual system supports linguistic functions and to uncover the underlying organizational principles that enable this integration.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the STI2030-Major Project (2021ZD0204100), the National Natural Science Foundation of China (31925020, 82021004), and the China Postdoctoral Science Foundation (2024M760231). We thank Yuxi Chu for her helpful discussions.

References

Balaban, M. T., & Waxman, S. R. (1997). Do words facilitate object categorization in 9-month-old infants? *Journal of Experimental Child Psychology*, *64*(1), 3–26. https://doi.org/10.1006/jecp.1996.2332

Bi, Y., Han, Z., Zhong, S., Ma, Y., Gong, G., Huang, R., Song, L., Fang, Y., He, Y., & Caramazza, A. (2015). The white matter structural network underlying human tool use and tool understanding. *The Journal of Neuroscience*, *35*(17), 6822–6835. https://doi.org/10.1523/JNEUROSCI.3709-14.2015

Bi, Y., Wang, X., & Caramazza, A. (2016). Object domain and modality in the ventral visual pathway. *Trends in Cognitive Sciences*, *20*(4), 282–290. https://doi.org/10.1016/j.tics.2016.02.002

Buxbaum, L. J., Shapiro, A. D., & Coslett, H. B. (2014). Critical brain regions for tool-related and imitative actions: A componential analysis. *Brain*, *137*(7), 1971–1985. https://doi.org/10.1093/brain/awu111

Chao, L. L., & Martin, A. (2000). Representation of manipulable man-made objects in the Dorsal Stream. *Neuroimage*, *12*(4), 478–484. https://doi.org/10.1006/nimg.2000.0635

Chen, H., Liu, B., Wang, S., Wang, X., Han, W., Zhu, Y., Wang, X., & Bi, Y. (2025). Language modulates vision: Evidence from neural networks and human brain-lesion models. *arXiv preprint arXiv:2501.13628*.

Fan, S., Wang, X., Wang, X., Wei, T., & Bi, Y. (2021). Topography of visual features in the human ventral visual pathway. *Neuroscience Bulletin*, *37*(10), 1454–1468. https://doi.org/10.1007/s12264-021-00734-4

Fedorenko, E., Hsieh, P.-J., Nieto-Castañón, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New method for fMRI investigations of language: Defining ROIs functionally in individual subjects. *Journal of Neurophysiology*, *104*(2), 1177–1194. https://doi.org/10.1152/jn.00032.2010

Ferguson, B., & Waxman, S. (2017). Linking language and categorization in infancy. *Journal of Child Language*, *44*(3), 527–552. https://doi.org/10.1017/S0305000916000568

Kastner, S., Chen, Q., Jeong, S. K., & Mruczek, R. E. B. (2017). A brief comparative review of primate posterior parietal cortex: A novel hypothesis on the human toolmaker. *Neuropsychologia*, *105*, 123–134. https://doi.org/10.1016/j.neuropsychologia.2017.01.034

Kong, X.-Z., Tzourio-Mazoyer, N., Joliot, M., Fedorenko, E., Liu, J., Fisher, S. E., & Francks, C. (2020). Gene expression correlates of the cortical network underlying sentence processing. *Neurobiology of Language*, *1*(1), 77–103. https://doi.org/10.1162/nol_a_00004

Le Guen, Y., Amalric, M., Pinel, P., Pallier, C., & Frouin, V. (2018). Shared genetic aetiology between cognitive performance and brain activations in language and math tasks. *Scientific Reports*, *8*(1), 17624. https://doi.org/10.1038/s41598-018-35665-0

Lewis, J. W. (2006). Cortical networks related to human use of tools. *The Neuroscientist*, *12*(3), 211–231. https://doi.org/10.1177/1073858406288327

Liu, B., Wang, X., Wang, X., Li, Y., Han, Y., Lu, J., Zhang, H., Wang, X., & Bi, Y. (2025). Object knowledge representation in the human visual cortex requires a connection with the language system. *PLOS Biology*, *23*(5), e3003161. https://doi.org/10.1371/journal.pbio.3003161

Morgan, T. J. H., Uomini, N. T., Rendell, L. E., Chouinard-Thuly, L., Street, S. E., Lewis, H. M., Cross, C. P., Evans, C., Kearney, R., De La Torre, I., Whiten, A., & Laland, K. N. (2015). Experimental evidence for the co-evolution of hominin tool-making teaching and language. *Nature Communications*, *6*(1), 6029. https://doi.org/10.1038/ncomms7029

Peelen, M. V., Bracci, S., Lu, X., He, C., Caramazza, A., & Bi, Y. (2013). Tool selectivity in left occipitotemporal cortex develops without vision. *Journal of Cognitive Neuroscience*, *25*(8), 1225–1234. https://doi.org/10.1162/jocn_a_00411

Peeters, R., Simone, L., Nelissen, K., Fabbri-Destro, M., Vanduffel, W., Rizzolatti, G., & Orban, G. A. (2009). The representation of tool use in humans and monkeys: Common and uniquely human features. *The Journal of Neuroscience*, *29*(37), 11523–11539. https://doi.org/10.1523/JNEUROSCI.2040-09.2009

Perszyk, D. R., & Waxman, S. R. (2019). Infants' advances in speech perception shape their earliest links between language and cognition. *Scientific Reports*, *9*(1), 3293. https://doi.org/10.1038/s41598-019-39511-9

Ritchie, J. B., Wardle, S. G., Vaziri-Pashkam, M., Kravitz, D. J., & Baker, C. I. (this issue). Rethinking category-selectivity in human visual cortex. *Cognitive Neuroscience*, 1–28. https://doi.org/10.1080/17588928.2025.2543890

Thibault, S., Py, R., Gervasi, A. M., Salemme, R., Koun, E., Lövden, M., Boulenger, V., Roy, A. C., & Brozzoli, C. (2021). Tool use and language share syntactic processes and neural patterns in the basal ganglia. *Science*, *374*(6569), eabe0874. https://doi.org/10.1126/science.abe0874

Wang, A. Y., Kay, K., Naselaris, T., Tarr, M. J., & Wehbe, L. (2023). Better models of human high-level visual cortex emerge from natural language supervision with a large and diverse dataset. *Nature Machine Intelligence*, *5*(12), 1415–1426. https://doi.org/10.1038/s42256-023-00753-y

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology*, *29*(3), 257–302. https://doi.org/10.1006/cogp.1995.1016

Wen, H., Xu, T., Wang, X., Yu, X., & Bi, Y. (2022). Brain intrinsic connection patterns underlying tool processing in human adults are present in neonates and not in macaques. *Neuroimage*, *258*, 119339. https://doi.org/10.1016/j.neuroimage.2022.119339

Woodruff Carr, K., Perszyk, D. R., & Waxman, S. R. (2021). Birdsong fails to support object categorization in human infants. *PLOS ONE*, *16*(3), e0247430. https://doi.org/10.1371/journal.pone.0247430