Journal of Neurolinguistics 49 (2019) 248–251

Contents lists available at ScienceDirect

Journal of Neurolinguistics

journal homepage: www.elsevier.com/locate/jneuroling

Commentary on *Neuroemergentism: A framework for studying cognition and the brain.* The neurocomputations of neuroemergentism: Long-term memory + reinforcement learning = language?

Andrea Stocco[a],[b],[c],[d],[*]

[a] Department of Psychology, University of Washington, Campus Box 351525, Seattle, WA 98195, United States
[b] Institute for Learning and Brain Sciences, University of Washington, United States
[c] NSF Center for Sensorimotor Neural Engineering, University of Washington, United States
[d] UW Institute for Neuroengineering, University of Washington, United States

ABSTRACT

Hernandez and colleagues proposed a new framework, *neuroemergentism*, according to which complex functions (such as language) may originate from the interactions of existing neural mechanisms that have evolved for potentially different reasons, and whose recycling within a new network results in apparent functional specialization. Here, I will consider neuroemergentism from a computational perspective. In particular, I will examine how this approach could provide a way to make sense of linguistic abilities within two accepted, general-purpose neurocomputational frameworks: rational (Bayesian) retrieval of information from long-term memory, and reinforcement learning. These two frameworks might provide a natural way to account for the two most important components of language, lexicon and grammar, and they are consistent with the established neural substrates of linguistic processes. In agreement with the requirements of Hernandez and colleagues, this hypothesis is also compatible, in principle, with language genetics and language development.

1. Introduction

In their recent paper, Hernandez, Greene, Vaughn, Francis and Grigorenko (2015/8) proposed a new framework, *neuroemergentism*, that organizes apparently incompatible points of view on the relationship between neural specialization and cognitive function into a single, integrated model. The authors' central point is that complex functions (such as language) might originate through the interactions of existing mechanisms, and that this interaction provides the best lens through which we can analyze any complex cognitive phenomena. This idea merges several key intuitions from previous work on language, including Elizabeth Bates' view of language as a “new machine made of old parts” and Stanislas Dehaene's hypothesis on the recycling of existing neural circuits based on their functional connectivity.

In addition, Hernandez and colleagues consider two other important windows into understanding the mechanisms by which complex cognitive functions are supported by evolutionarily older brain regions. The first is *development*, which provides an insight into the different stages in which a function develops, and provides a way to track the corresponding changes in the underlying neural

---

[*] Department of Psychology, University of Washington, Campus Box 351525, Seattle, WA 98195, United States.
E-mail address: stocco@uw.edu.

https://doi.org/10.1016/j.jneuroling.2018.04.008
Received 7 March 2018; Received in revised form 21 April 2018; Accepted 22 April 2018
0911-6044/

hardware (Yeatman et al., 2011; Kuhl, 2004). The second is *genetics*, which provides powerful insights into the biological processes that support a function. As Hernandez himself has pointed out in a recent paper (Vaughn and Hernandez, 2018), genetic mechanisms and development interact–for instance, second language proficiency is associated with the expression of dopamine receptors in the striatum at early ages, but in the prefrontal cortex at later ones.

What I found most compelling about this hypothesis is how well it matches with some of the key insights of computational neuroscience. Emergentism is, of course, a well known concept in the field of computational neuroscience: the composition of two functions might not resemble any of the two functions; thus, two well-known brain circuits (described by separate functions) might interact to produce behavior that seems only apparently irreducible to their original functions (O'Reilly and Munakata, 2000).

Implicitly, the neuroemergentist approach calls into question some of the most common paradigms to investigate language, which are based on the top-down decomposition of linguistic operations from specific computational frameworks. As an example, consider the recent paper by Goucha and Friederici (2015) in which the authors identify a subregion of Broca's area responsible for a specific morphosyntactic operation. In this case, functional specialization of brain circuitry and functional decomposition of language go hand-in-hand. According to a neuroemergentist framework, however, this top-down decomposition of language function (and the corresponding fractionation of its neural substrate) might be a misleading approach.

In addition to serving as a warning for experimental investigations, neuroemergentism also offers an alternative approach. Namely, instead of decomposing linguistic functions based on top-down task analysis, we should instead look for general, “bottom-up” computations that support linguistic functions but also have other, evolutionarily justified, uses (the “old parts” in Bates' metaphor). Ideally, these computational functions should be (1) evolutionarily plausible and advantageous, (2) provide a natural explanation for neural data.

2. A neuroemergentist, neurocomputational view of language

As an example of neuroemergentist “neurocomputational” approach, I will outline here an alternative way in which linguistic functions can be decomposed. This example is described here for purely illustrative purposes; although I (personally) find it convincing, I will certainly not claim that it will hold up to experimental scrutiny. Nor is this approach entirely original, as similar ideas have been suggested in the past. Finally, this approach is far from complete, and does not consider language in its entirety, focussing instead on central cognitive components at the expense of sensorimotor processes (visual processes for reading, auditory processes for phonology, and motor processes for writing and vocalization). Specifically, the two components I will consider are the *lexicon* (which ties specific combinations of sounds and meaning) and the *grammar* (which contains combinatorial rules). I will discuss them separately and provide evidence for their roles.

2.1. Component 1: lexicon and rational memory retrieval

The first claim is that the workings of the mental lexicon can be understood within the rules that govern long-term memory. Multiple models of long-term memory (LTM) exist, with several degrees of mathematical (Diller, Nobel, & Shiffrin, 2001; Shiffrin & Steyvers, 1997) and biological sophistication (McClelland, McNaughton, & O'Reilly, 1995; Norman & O'Reilly, 2003). Neurally, LTM has been linked to the hippocampus and the temporal lobe, which are also associated with the mental lexicon. In addition, the specific characteristics of memory retrieval, such as the role of frequency, recency, and context, have been linked to evolutionary survival through rational analysis, which, in essence, shows that the dynamics of long-term memory are optimal for the specific environment we live in Anderson and Schooler (1991), Anderson (1990).

2.2. Component 2: grammar and reinforcement learning

The second, and most controversial, claim, is that the grammatical system can be understood in terms of Reinforcement Learning (RL: Sutton & Barto, 1998). RL is a machine-learning theory that describes how an agent can learn optimal behavior through unsupervised (i.e., environmental or self-generated, rather than explicitly instructed) feedback. At its simplest, “model-free” form, RL theory provides algorithms that associated expected *rewards* to basic *actions* or states, and unifies classic behavioral theories, such as Pavlovian and instrumental conditioning. In its most complex, “model-based” version, however, RL includes mechanisms to create internal models of the task, and provides ways to learn the optimal concatenation of actions to achieve incredibly complex goals, such as planning ahead the optimal sequence of moves in a game. For this reason, RL is the cornerstone of recent AI achievements (Mnih et al., 2015; Silver et al., 2016).

In the field of language, RL can be used to understand how specific sequences of words are used to produce meaning. RL methods share commonalities with statistical learning methods (Sutton and Barto, 1998), but are characterized by being driven by reward signals. In the domain of language, *reward* is associated with successful communication (Kraus and White-Schwoch, 2015; Kuhl, 2011), and the *actions* are the selection of single words from the lexicon (through memory retrieval) and their morphological declination to fit specific roles.

Like memory retrieval, RL checks both marks for being a candidate “old part” to be used for the “new machine” of language. The evolutionary importance of RL is obvious since the entire theory is about optimal mechanisms for adapting behavior to maximize rewards. In addition, there is overwhelming evidence that RL theory explains multiple facets of neural activity. For example, it provided an elegant computational explanation for the otherwise bizarre behavior of midbrain dopamine neurons (Schultz, Dayan, & Montague, 1997), and has recently provided an elegant explanation for complex cognitive activity such as planning (Daw, Gershman, Seymour, Dayan, & Dolan, 2011) and response conflict (Holroyd & Yeung, 2012; Yeung, Botvinick, & Cohen, 2004).

3. Plausibility of the approach

At least in its broader strokes, this approach is consistent with our current understanding of language. The role of the temporal lobe in supporting both LTM and mental lexicon is fairly uncontroversial; lexical access, for instance, shows the same effects of context, frequency, and recency of LTM retrieval. The RL component might appear a bit more problematic since the core computation of RL (the so-called *reward prediction error*) is linked to striatal dopamine, which is not typically considered part of the language network. However, both the role of dopamine (Wong, Morgan-Short, Ettlinger, & Zheng, 2012) and subcortical contributions to syntax (Wallesch et al., 1983) have been shown, and might have been overlooked in favor of cortical ones (Stocco, Lebiere, & Anderson, 2010).

Finally, it is worth noting that this approach is also analogous to Ullman's declarative/procedural model (Ullman et al., 1997) if one assumes that procedural knowledge can be equated to RL actions. This equivalence is commonly made since both RL computations and procedural knowledge share the same substrate, namely, the basal ganglia. This assumption has also been used in models of language development (Taatgen & Anderson, 2002), is widely assumed in cognitive architectures, such as ACT-R (Anderson, 2007) and Soar (Laird, 2012), and is common among neurocomputational models of the basal ganglia (Stewart, Choo, & Eliasmith, 2010; Stocco et al., 2010).

3.1. Development and neurogenetics

According to Hernandez and colleagues, a test of this model through a neuroemergentist lens should include an examination of this approach in terms of development and genetics. Again, the case is easier to make for LTM and lexicon, which follow similar developmental trajectories (Gathercole, 1998).

The case of RL as the framework for understanding grammar needs to be made. In many ways, because RL is a theory of *learning*, it already provides a natural way to think about language acquisition and development. The connection needs to be mediated by reward, which is the foundation of RL algorithms. Such a connection exists and has been brought into light in recent years, showing the role of reward in auditory learning (Kraus and White-Schwoch, 2015). In children, reward is likely taking the form of successful social interaction, which is essential for the survival of developing humans and seems to be correspondingly essential for language acquisition (Kuhl, 2011).

In terms of genetics, Hernandez himself has broken new ground showing the role played by striatal dopamine in language acquisition (Hernandez et al., 2015/8; Vaughn and Hernandez, 2018). Genetic mutations linked to selective language impairments have been linked to abnormally low levels of midbrain dopamine, and are also co-morbid with other psychiatric conditions linked to dopamine disorders (Newbury et al., 2010) Even FOXP2, the gene most commonly associated with language function and human evolution, is in fact related to striatal dopamine (Enard, 2011; Murugan, Harward, Scharff, & Mooney, 2013).

4. Conclusions

Hernandez and colleagues’ neuroemergentist approach provides a systematic way to reason about the neural constraints that underpin complex functions and, specifically, language. Their approach organizes and harmonizes previous theories that were subtly incongruent, and provides new directions for research. In this commentary, I have argued that the neuroemergentist approach is also consistent with the approach of computational neuroscience, and I have outlined a possible way in which language could be understood in terms of two well-established and powerful computational frameworks: Rational analysis of memory and Reinforcement Learning theory for action selection and composition. This approach is broadly consistent with the current literature and even with popular models of language function, and might provide a useful bridge to connect language research to widespread and important neurocomputational frameworks.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jneuroling.2018.04.008.

References

Anderson, J. R. (1990). *The adaptive character of thought.* L. Erlbaum Associates.

Anderson, J. R. (2007). *How can the mind occur in the physical universe?* Oxford University Press.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science, 2*(6), 396–408.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron, 69*(6), 1204–1215.

Diller, D. E., Nobel, P. A., & Shiffrin, R. M. (2001). An ARC–REM model for accuracy and response time in recognition and recall. Journal of Experimental Psychology. *Learning, Memory, and Cognition, 27*(2), 414.

Enard, W. (2011). FOXP2 and the role of cortico-basal ganglia circuits in speech and language evolution. *Current Opinion in Neurobiology, 21*(3), 415–424.

Gathercole, S. E. (1998). The development of memory. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 39*(1), 3–27.

Goucha, T., & Friederici, A. D. (2015). The language skeleton after dissecting meaning: A functional segregation within Broca’s Area. *NeuroImage, 114*, 294–302.

Hernandez, A. E., Greene, M. R., Vaughn, K. A., Francis, D. J., & Grigorenko, E. L. (2015/8). Beyond the bilingual advantage: The potential role of genes and environment on the development of cognitive control. *Journal of Neurolinguistics, 35*, 109–119.

Holroyd, C. B., & Yeung, N. (2012). Motivation of extended behaviors by anterior cingulate cortex. *Trends in Cognitive Sciences, 16*(2), 122–128.

Kraus, N., & White-Schwoch, T. (2015). Unraveling the biology of auditory learning: A cognitive-sensorimotor-reward framework. *Trends in Cognitive Sciences, 19*(11), 642–654.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews. Neuroscience, 5*(11), 831–843.

Kuhl, P. K. (2011). Social mechanisms in early language acquisition: Understanding integrated brain systems supporting language. In Jean Decety, & John (Eds.). *The oxford handbook of social neuroscience*. Oxford University Press.

Laird, J. E. (2012). *The soar cognitive architecture.* MIT Press.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*(3), 419–457.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., et al. (2015). Human-level control through deep reinforcement learning. *Nature, 518*(7540), 529–533.

Murugan, M., Harward, S., Scharff, C., & Mooney, R. (2013). Diminished FoxP2 levels affect dopaminergic modulation of corticostriatal signaling important to song variability. *Neuron, 80*(6), 1464–1476.

Newbury, D. F., Fisher, S. E., & Monaco, A. P. (2010). Recent advances in the genetics of language impairment. *Genome Medicine, 2*(1), 6.

Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychological Review, 110*(4), 611–646.

O’Reilly, R. C., & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain.* A Bradford Book.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275*(5306), 1593–1599.

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM—retrieving effectively from memory. *Psychonomic Bulletin & Review, 4*(2), 145–166.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., et al. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature, 529*(7587), 484–489.

Stewart, T. C., Choo, X., & Eliasmith, C. (2010). Dynamic behaviour of a spiking model of action selection in the basal ganglia. *Proceedings of the 10th international conference on cognitive modeling (pp. 235–240). Citeseer*.

Stocco, A., Lebiere, C., & Anderson, J. R. (2010). Conditional routing of information to the cortex: A model of the basal ganglia's role in cognitive coordination. *Psychological Review, 117*(2), 541–574.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*, *Vol. 1*. MIT press Cambridge.

Taatgen, N. A., & Anderson, J. R. (2002). Why do children learn to say “Broke”? A model of learning the past tense without feedback. *Cognition, 86*(2), 123–155.

Ullman, M. T., Corkin, S., Coppola, M., Hickok, G., Growdon, J. H., Koroshetz, W. J., et al. (1997). A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience, 9*(2), 266–276.

Vaughn, K. A., & Hernandez, A. E. (2018). Becoming a balanced, proficient bilingual: Predictions from age of acquisition & genetic background. *Journal of Neurolinguistics, 46*, 69–77.

Wallesch, C. W., Kornhuber, H. H., Brunner, R. J., Kunz, T., Hollerbach, B., & Suger, G. (1983). Lesions of the basal ganglia, thalamus, and deep white matter: Differential effects on language functions. *Brain and Language, 20*(2), 286–304.

Wong, P. C. M., Morgan-Short, K., Ettlinger, M., & Zheng, J. (2012). Linking neurogenetics and individual differences in language learning: The dopamine hypothesis. *Cortex; Journal Devoted to the Study of the Nervous System and Behavior, 48*(9), 1091–1102.

Yeatman, J. D., Dougherty, R. F., Rykhlevskaia, E., Sherbondy, A. J., Deutsch, G. K., Wandell, B. A., et al. (2011). Anatomical properties of the arcuate fasciculus predict phonological and reading skills in children. *Journal of Cognitive Neuroscience, 23*(11), 3304–3317.

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review, 111*(4), 931–959.