Assessing gain modulation as a cortical principle of natural language computation

Elliot Murphy[1,2]

1. Vivian L. Smith Department of Neurosurgery, McGovern Medical School, University of Texas Health Science Center
2. Texas Institute for Restorative Neurotechnologies, University of Texas Health Science Center

Recent work in language and computation in neural systems has begun to develop a model of linguistic computation relying on gain modulation, or the process whereby neurons combine multiple sources of information. In this brief article, I will review these proposals and provide two conclusions: First, gain modulation models of language face a number of internal conceptual and empirical difficulties. Second, our current understanding of gain modulation does not support its use as a core component of linguistic computation.

Language and gain modulation: The proposal

With ambitious interdisciplinary scope, Martin (2020) defends a multidimensional coordinate system for language based on neurophysiological models of sensory processing. Martin maintains that neural trajectories (of these coordinates) encode *sensory*, *motor* and *abstract linguistic states*. Crucial to Martin's model is the assumption that gain modulation tunes the path of these trajectories in accordance with behavior, and that this is "how latent structure is inferred" (2020). Similar to how Chomsky famously points to the paradox behind language use being appropriate to circumstance but not caused by circumstance (*Descartes' problem*), Martin points to a more recent, apparent paradox between the "marvelous expressive capacity" of language (i.e. its generative nature) and the "general remit" of the human brain to be "driven by statistical regularities in its environment". Martin adds that, lacking a neurophysiological account of core linguistic properties such as hierarchical structure, function application and scope, and compositionality, theories of the human brain "seem startingly incomplete".

Attempting to move beyond this situation, Martin (2020) proposes that linguistic structure building is "a form of perceptual inference". She adds that increasingly abstract linguistic structures (from syllables to morphemes to words) are inferred via gain modulation which she defines as "the way neurons combine information from two or more sources", with, for instance, inhibition and signal amplification being notable ways that the brain seems to combine and separate information.

Martin (2020) reports how gain modulation underlies "coordinate transform between sensory modalities and between sensory and motor systems". She cites classic work on cellular dynamics to support this (Salinas & Abbott 2001), but does not cite other important work by the same co-author (most importantly, Dayan & Abbott 2001) where the limitations of gain modulation are stressed, e.g. gaze-dependent gain modulation of retinotopic visual receptive fields is well-reported, but the relevance of gain modulation to higher cognition is not established; further, "[t]he mechanism by which sensory and modulatory inputs combine in a multiplicative way in gain-modulated neurons is not known" (Dayan & Abbott 2001: 17; although see discussion of Ferguson and Kardin 2020 below). Dayan and Abbott (2001: 48) note the potential that linear and non-linear feedforward and recurrent networks have to act as content-addressable memories *alongside* the exhibition of gain modulation, but it is not clear that gain modulation itself carries any explanatory scope certainly to the level of sentential representations. Martin (2020) does provide an interesting speculation that "[c]oordinate transform may be a computational requirement of any system with multiple data types or formats from multiple perceptors", yet the risk of centering gain modulation in these processes is that the genericity of this system-wide process leads to a situation in which one cannot easily test or falsify Martin's claims. Martin (2020) notes that "how [linguistic] units combine together via gain modulation over time is hypothetical and must be tested", but no empirical suggestions for testing are discussed. Martin (2020) later claims that "the role of gain modulation in the brain is likely to be much more broad than as a neuronal instantiation of attention" (Martin 2020). Unfortunately, Martin provides no empirical support for this claim.

Martin (2020) makes an additional claim that "gain modulation also offers a built-in system for predictive coding", citing Friston (2005); more recent, updated accounts (e.g. Friston 2018) rely on an interplay of a range of other processes, not gain modulation. Moreover, Friston's ventures into language (most prominently in Friston et al. 2017, exploring epistemic foraging and Bayesian belief updating) do not seem compatible with Martin's (2020) extensive use of gain modulation (see Shipp 2016 for discussion of some possible, motivated neural implementations of predictive coding, and also Heilbron & Chait 2018 for mechanisms more in line with the spirit of Martin's proposal).

One interesting advantage of Martin's model is the multiple-realisability of gain modulation, which is a deceptively more abstract process than its textbook definition would imply. Martin's (2020) supplied pseudocode for her model (Table 2) is instructively clear, yet it also begs the question as to the neuronal substrates of the proferred representations; "lexico-syntactic" relations are assumed *ab initio*, and one cannot simply stipulate that they align with "gain-field trajectories" without cashing out these higher-level abstractions within some form of compatible lower-level mechanism. In addition, Martin's (2020) account of minimal pair interpretation relies on inhibitory implementations of gain modulation, whereby activation of one representation inhibits the other, an interesting feature which might also extend to broader linguistic processes like polysemy. Still, as mentioned, the broader architectural setting of these operations is unclear. Relatedly, Martin only addresses in a footnote a core question at hand concerning the underlying non-linear vs linear (and Euclidean vs non-Euclidean) structure of the brain dynamics, but this is an essential issue to sustain a model relying on gain modulation.[1]

Language and gain modulation: A reassessment

How do these assumptions about the role of gain modulation fare when we examine the apparent mechanisms underlying cortical gain modulation? Ferguson and Kardin (2020) provide a comprehensive review of gain modulation, pointing to a common set of mechanisms: GABAergic inhibition, synaptically driven fluctuations in membrane potential, and changes in cellular conductance. Ferguson and Kardin (2020) note that diverse cortical

[1] Some other issues serve to make the exposition less clear. In Table 1 of Martin's (2020) paper, 'like an arrow' is presented as both a PP and an AdvP (incorrectly). The 'Role-filler binding calculus' in Table 1 is not informative: Representing 'Time flies like an arrow' as {'flies(time)'; 'like(an arrow)'} does not approach the underlying richness in thematic relations. Turning to Figure 1, the cartoon representation of words and morphemes is incorrectly matched, such that morphemes are represented as being composed of words, rather than the other way round. It is also unclear how Figure 2 provides conceptual or psycholinguistic clarity beyond the level of stating that 'Time flies like an arrow' contains five words with separate phrases.

functions such as information integration across cognitive, sensory and motor systems seem to be performed through gain modulation. Further, regulation of neural gain can provide “an integration mechanism whereby information from multiple sources can be non-linearly combined via multiplicative modulation of the cell’s response to inputs” (2020: 81). Morphological features of dendrites also regulate the extent to which gain control is neurons is possible; e.g. moderate branching in pyramidal neurons potentially promotes the greatest possible range of gain modulation (2020: 84).

Martin (2020) also briefly defends an additive, rather than a tensor-based multiplicative, formulation of natural language compositionality. On this issue of additive versus multiplicative gain modulation, Ferguson and Kardin (2020) review how visual properties that must be *decoded separately* are combined multiplicatively, whereas parts that must be *integrated* are combined additively. If gain modulation is to play a central role in linguistic computation, then one would assume that Martin’s additive-only model will be able to capture integrative functions of natural language compositionality, but that processes involved in parsing discrete linguistic (and embedded) representations that require structural separation may require multiplicative gain modulation.

What is more in line with Martin’s specific proposal is Ferguson and Kardin’s (2020: 88) summary that gain modulation serves the “rapid adaptation to varying ranges of input” in addition to “enhancing the salience of relevant information”, with the notion of salience being relevant to a range of linguistic processes. In addition, Ferguson and Kardin (2020: 88) note – using terminology familiar to syntacticians – that gain modulation may be able to *increase information transmission* and provide *computational efficiency* within a given neural network.

Nevertheless, it remains currently unknown how gain modulation at the single-cell level contributes to population coding that can be read out by downstream targets.

The authors conclude that “gain is regulated by a wide range of influences, including attention, learning, locomotion, arousal and neuromodulatory activity, and that these may act through a common set of cellular and circuit mechanisms” (2020: 89). As such, while the role of gain modulation in sensory-to-motor conversions is well-established, gain modulation has not currently been implicated in any sub-component of language-relevant cognition, with *attention* very much constituting the periphery of the B in FLB (“Faculty of Language Broad”; Hauser et al. 2002). Further, “the precise relationship between gain modulation of single neurons and the encoding and transmission of information at the population level is not well understood”, while “the contribution of neural gain control to perceptual and cognitive performance remains to be fully explored” (Ferguson & Kardin 2020: 89). As such, it seems premature to use gain modulation as a core component of any model of higher-level linguistic computation.

Turning to related concerns, Ferguson and Kardin (2020: 89) note that “the precise relationship between gain modulation of single neurons and the encoding and transmission of information at the population level is not well understood”, and that “the reliability and repeatability of gain modulation of single neurons and cortical networks is unknown”. Given this, it seems not well-motivated to stipulate (as Martin 2020 does) a close connection between gain modulation and the full range of oscillatory frequency interactions which seem to index the processing of linguistic representations. Since gain modulation does in fact appear relevant to enabling networks of neurons to produce distributed representations of stimulus features (as opposed to generating inferences to higher abstract structures from such features) , it may be that gain modulation goes the way of entrainment: Initially viewed as a crucial mechanism for all manner of linguistic processes (again, from syllables to sentences) before being relegated to a role in auditory cortex for *some* aspects of lower-level speech perception (Murphy 2020).

Keeping to these systems-level concerns, while Martin has previously sided against multiplexing as a viable mechanism for grounding linguistic computation, in Martin (2020) she now proposes that cortical oscillations structure speech input into linguistic representations via gain-modulated multiplexing. Multiplexing models of language have been proposed elsewhere (e.g. Benítez-Burraco & Murphy 2019, Murphy 2016a, 2016b, 2018, 2020 utilize multiplexing as operationalised via a range of distinct cross-frequency couplings; see also Gross et al. 2013) and, given the above-mentioned problems with relating our current understanding of gain modulation to higher-order language processing, it is unclear how Martin's (2020) *gain-modulated multiplexing* is a step forward. In addition, Martin puts forward one oscillation-related hypothesis, that "[t]he theta oscillation may be a likely carrier signal for linguistic sensory input, but more carrier signals and coherence between them must exist for the perceptual inference of linguistic structure, which itself is not recoverable from the sensory codes alone". Which sensations in particular are carried by theta, and how such a multiplexing mechanism plays a role in the broader architecture Martin alludes to but does not develop, are questions not explored but presumably need to be addressed in any given model of *language* as a system, rather than its sub-systems. In addition, while Martin's intuition of gain modulation being an elegant algorithmic-level description of compositionality is surely novel and appealing (putting aside the objections I have made here), and while Martin's ideas are also not incompatible with a range of oscillatory-models of language in the literature, Martin does not discuss any limitations of her model.

Finally, one potential successful application of gain modulation models is to learning, which appears to be more amenable to implementations via gain modulation (Ferguson & Kardin 2020). Nevertheless, testing Martin’s (2020) gain modulation model will require a must deeper understanding of the relationship between gain modulation and processes that we can more easily record in the human brain.

References

Benítez-Burraco, A., & Murphy, E. (2019). Why brain oscillations are improving our understanding of language. *Frontiers in Behavioral Neuroscience* 13: 190.

Dayan, P., & Abbott, L.F. (2001). *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems.* Cambridge, MA: MIT Press.

Ferguson, K.A., & Cardin, J.A. (2020). Mechanisms underlying gain modulation in the cortex. *Nature Neuroscience* 21: 80–92.

Friston, K. (2005). A theory of cortical responses. Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences 360: 815–836.

Friston, K.J., Rosch, R., Parr, T., Price, C., & Bowman, H. (2017). Deep temporal models and active inference. *Neuroscience and Biobehavioral Reviews* 77: 388–402.

Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., & Garrod, S. (2013). Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biology* 11: e1001752.

Hauser, M., Chomsky, N., & Fitch, W.T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science* 298(5598): 1569–1579.

Heilbron, M., & Chait, M. (2018). Great expectations: is there evidence for predictive coding in auditory cortex? *Neuroscience* 389: 54–73.

Martin, A. (2020). A compositional neural architecture for language. *Journal of Cognitive Neuroscience* 1–20. doi.org/10.1162/jocn_a_01552.

Murphy, E. (2016a). A theta-gamma neural code for feature set composition with phase-entrained delta nestings. *UCL Working Papers in Linguistics* 28: 1-23.

Murphy, E. (2016b). The human oscillome and its explanatory potential. *Biolinguistics* 10: 6-20.

Murphy, E. (2018). Interfaces (travelling oscillations) + recursion (delta-theta code) = language. Luef, E., & Manuela, M. (Eds). *The Talking Species: Perspectives on the Evolutionary, Neuronal and Cultural Foundations of Language.* Graz: Unipress Graz Verlag. 251-269.

Murphy, E. (2020). *The Oscillatory Nature of Language*. Cambridge: Cambridge University Press.

Salinas, E., & Abbott, L.F. (2001). Coordinate transformations in the visual system: how to generate gain fields and what to compute with them. *Progress in Brain Research* 130: 175–190.

Shipp, S. (2016). Neural elements for predictive coding. *Frontiers in Psychology* 7: 1792.