

Language and other complex behaviors: unifying characteristics, computational models, neural mechanisms*

Shimon Edelman
Department of Psychology
Cornell University
se37@cornell.edu

Revised — April 4, 2017

Abstract

Similar to other complex behaviors, language is dynamic, social, multimodal, patterned, and purposive, its purpose being to promote desirable actions or thoughts in others and self (Edelman, 2017b). An analysis of the functional characteristics shared by complex sequential behaviors suggests that they all present a common overarching computational problem: dynamically controlled constrained navigation in concrete or abstract situation spaces. With this conceptual framework in mind, I compare and contrast computational models of language and evaluate their potential for explaining linguistic behavior and for elucidating the brain mechanisms that support it.

**Language Sciences*, to appear. DOI: <http://doi.org/10.1016/j.langsci.2017.04.003> © 2017. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

1 Introduction

How do brains compute language? The current consensus view in cognitive sciences, forty years after Marr and Poggio (1977) (cf. Marr, 1982; Poggio, 2012; Edelman, 2012), is that questions of cognitive computations and mechanisms cannot be settled without also addressing the other, complementary levels of understanding. Notably, there is the abstract or problem level: what is it, in terms of computation, that needs to be done, and why? Even this, methodologically more appropriate approach is, however, liable to lead nowhere if the problem-level hypotheses are mistaken, without being recognized as such — as seems to be the case both in Marr’s original field, vision, which is still widely and erroneously believed to hinge on “object recognition” (Edelman, 2017a), and in language, where problem-level thinking is dominated by the conception of communication as packaging meanings into messages and by the tripartite dogma of grammar, sentence, and well-formedness (Edelman, 2017b; more about this in a moment).

A remedy for this methodological impasse is to augment Marr’s three levels (problem, algorithm, and implementation) with two additional and related perspectives on the phenomenon in question, which are obligatory in biology: Mayr’s (1961) concerns about explaining causation (including the distinction between proximate and ultimate causes), and Tinbergen’s (1963) four questions — survival value, ontogeny (development), evolution, and behavioral causation. In language science, in particular, it is critically important to make questions of evolution, development, and behavior an integral part of the inquiry, as suggested next.

1.1 Language: postulates and reality

The reigning conceptual framework in linguistics, which underlies both the formalist and the functionalist theoretical outlooks,¹ rests on two postulates. First, with regard to the form or structure of language, it is held that the theoretical focus should be on sentences, whose well-formedness is underwritten by a set of formal rules, or grammar. Second, the function of language is believed to be communication, construed as an exchange of “meanings” — information packets that are fully formed by the speaker, to be shipped to and decoded by the listener.

As I argued at some length elsewhere (Edelman, 2017b), it is time for these two postulates to be traded off for a biologically better motivated, and perhaps more promising, approach. To that end, we should consider both the form and the function of language — and perforce the computational processes that comprise it and the brain mechanisms that implement those processes — in a broader evolutionary and behavioral context. Specifically, the extensive literature on animal signals (e.g. Smith, 1965; Green and Marler, 1979; Macedonia and Evans, 1993; Lachmann et al., 2001; Seyfarth and Cheney, 2003; Ouattara et al., 2010; Kershenbaum *et al.*, 2015) suggests that meaning in animal behavior in general and in language in particular should be conceived of not as a completed product to be “communicated” but as an open-ended interactive process to participate in — meaning-making — which is dynamically constrained by a range of factors that include the participants’ shared social background, personal history, and immediate environment (Neuman, 2006; Stolk et al., 2016). ***The goal of this process, and the overarching use of language, is to influence behavior*** (Edelman, 2017b, section 6.2).² Likewise, the structures — composite actions or strands of behavior — that mediate this influence need not be computed completely ahead of time or syntactically well-formed;

¹For a book-length treatment of the distinction between formalist and functionalist linguistics, see (Newmeyer, 1998).

²Similarly, Anderson (2016) writes that “Language works by presenting and manipulating cultural affordances that will cause one’s dialog partner(s) to see and do what the speaker intends to be seen and done.” Importantly, although this claim revolves around *behavior*, which virtually necessitates mentioning (Skinner, 1957), it is anything but “behaviorist,” as explained in (Edelman, 2017b).

rather, linguistic behavior, just like any other complex sequential behavior, must be controlled flexibly and dynamically, subject to just enough constraints to make a difference in the desired direction (LaPolla, 2015; Edelman, 2017b).

1.2 A plan for this paper

In light of the above considerations, and in accordance with the research program outlined in (Edelman, 2017b), in this paper I adopt a pluralist stance that views language not as a peculiar faculty with a unique and evolutionarily unprecedented core function and brain substrate, but rather as a cluster of adaptations that supervene on neurocomputational systems which have analogs in other species and which jointly facilitate a range of behavioral tasks.

The notion that language as it exists in humans co-evolved with major systems of traits and functions, many of them shared with other taxa, has implications for understanding both its computational nature and its brain basis. On the task and computational levels, it motivates a search for commonalities between language and other behaviors (Kolodny and Edelman, 2015), such as navigation and foraging, which would imply that the problem of language is closely related computationally to the dynamical control of behavior (Bullock, 2004; Cisek, 2012). On the level of brain mechanisms, this approach shuns the standard corticocentric dogma in favor of a broad consideration of neural mechanisms that underlie action selection (Bullock et al., 2009) and sociality (Syal and Finlay, 2010).

My main goals in this paper are (i) to establish functional parallels between language and other complex behaviors; (ii) to evaluate and compare existing computational models of language, considered as a complex behavior; and (iii) to use this comparison to generate some insights into the brain basis of language. Accordingly, the paper is structured as follows. In section 2, I briefly review the types of behavioral tasks that exemplify what might be termed the generalized problem of behavior; section 3 then casts language as an instance of this problem. In preparation for assessing possible modeling solutions to it, in section 4 I list a set of criteria for neurocomputational plausibility, and state the explanatory goals that a good model must meet. These are then used to group computational models of language by algorithmic approach and to rank them by plausibility and explanatory power. Section 4 ends with a synthesis intended to serve as a basis for developing a new computational approach to language. In section 5, I briefly review the literature on the brain mechanisms of language and other sequential behaviors, focusing on the signatures of the proposed computational approach. Finally, section 6 offers a summary and a concise prognosis.

2 Language and other behaviors

Because understanding the structure or form of any evolved system depends on understanding of its function — a standard notion in biology in general (Bock and von Wahlert, 1965), as well as in language (Lenneberg, 1967; LaPolla, 2015) — I begin by briefly discussing some of the key categories of language use by humans.

2.1 The uses of language

A widely held view in linguistics is that the primary use of language is communication (e.g., Pinker and Jackendoff, 2005).³ This view is, however, both misleading, insofar as it promotes the construal of linguistic

³See (Edelman, 2017b) for more references and for a comparison of this essentially functionalist view with the one held by the formalists (e.g., Everaert et al., 2015), according to which language is primarily a tool for structured thinking, communication being merely its “ancillary” function.

<i>verbalization of experience</i>	Narrative structure is a foundational property of language, which has been hypothesized to drive grammaticalization (Croft, 2010). Stories one tells about oneself mediate the construction of the self and of one’s life history (Hermans et al., 1992; Fivush and Haden, 2003; Locke and Bogin, 2006).
<hr/>	
<i>teaching by story-telling</i>	Telling stories in instructional settings has been shown effective in college-level science (Cooper et al., 1983), early math education (Casey et al., 2004), history (Farmer, 1990), and ethics (Tappan and Brown, 1989; Upright, 2002). Cues to the causal structure of the world can be found in stories (including fairy tales; Bettelheim, 1976; cf. Gardner, 1993) and are often the focus of literary fiction (Cámpora, 2011).
<hr/>	
<i>practical skill transmission</i>	Language has arguably been critical to the teaching of practical skills in ancestral settings (Kolodny and Edelman, 2016; Laland, 2017), as suggested by evidence from contemporary hunter-gatherer societies (Garfield et al., 2016).
<hr/>	
<i>social bonding</i>	The social roles of language include facilitation of group cohesion, possibly through the use of gossip as a form of “grooming” (Dunbar, 1993), as well as the use of narratives for shared memory building and acculturation (Sutton, 2002); social complexity itself may be a driver of communication complexity (Freeberg et al., 2012).
<hr/>	
<i>coordinating joint action</i>	On a more immediate level, purposive discourse, which is always situated and embodied Sebanz et al. (2006), is a key use of language (Clark and Schaefer, 1989), in which all humans are adept (Garrod and Pickering, 2004).
<hr/>	
<i>courtship</i>	Locke and Bogin (2006) make a case for the central role of courtship in language evolution. The structure of conversation that characterizes courtship situations is being revealed by computational analysis of discourse (McFarland et al., 2013). There seem also to be specific hormonal effects on attention to courtship language (Rosen and López, 2009).

Table 1: Some of the categories into which the many uses of language can be classified. Note that all of these are subsumed under the general rubric of influencing the behavior of others (and self), as stated in section 1.1 (see Edelman, 2017b for arguments in support of this view). A particular instance of linguistic behavior (e.g., the dialogue sample offered in Table 2) may belong to more than one of these categories.

communication as an exchange of coded “meanings” (as noted in section 1.1), and superficial, as it glosses over the many categories of use that humans make of language. Some of the categories of language use are listed in Table 1, in no particular order, each accompanied by a few select references. In contrast to the customary appeal to communication and meaning, studying these diverse uses of language allows us to get help from a number of allied disciplines that are all concerned with the behavior of humans and other animals: psychology, sociology, anthropology, ethology, ecology, biology, and evolutionary science — a wealth of data and theories, which I can barely afford to touch upon in this paper.

A closer look at the various uses of language reveals deep functional parallels with other behaviors, in the form of shared structural characteristics. Indeed, all but the first two categories listed in Table 1 are not specific to language. Linguistic and other behaviors exhibit (i) *serial order and concurrency*, as in sequential narrative that is accompanied by prosody and gestures and is playing out against the backdrop of external events; (ii) *complexity and hierarchy*, as in the build-up of narrative structure, or in the goal/task hierarchy in instructional discourse; (iii) *embodiment and situatedness*, as in the dependence of discourse on physical settings and social context; and (iv) *agency and contingency*, as in personal motives steering linguistic behavior, which moreover is modified on the fly in response to the dynamically unfolding situation. In the next section, I argue that these four categories of structure are found in all complex behaviors.

2.2 The generalized problem of behavior: key characteristics and a reduction

Animal behavior in the wild is driven by certain overarching goals, such as gaining access to food, avoiding predation, and gaining or maintaining social status. In the spirit of McCulloch (1950; Iberall and McCulloch, 1968) and Lashley (1951), one may call the problem of translating goals into actions *the generalized problem of behavior* (Kolodny and Edelman, 2015). The key characteristics of this problem are the ones that I have just highlighted:

- (i) *Serial order and concurrency*. Because a remote behavioral goal may not be attainable via a single action that would map the present state of affairs to the desired one, behavior is generally an extended sequential undertaking rather than one stimulus-response (S-R) association after another (Lashley, 1951; Edelman, 2015). This undertaking may be continuous rather than discrete, as when “stimuli” and “responses” cannot even be individuated,⁴ and it is always multidimensional (if only because any action typically requires synergy among many muscles), with the different dimensions processed concurrently and asynchronously (Kolodny and Edelman, 2015).
- (ii) *Complexity and hierarchy*. An animal’s goals may be not just remote in time and space but also too abstract to afford direct translation into actionable behavioral plans, and so must be approached hierarchically. For instance, the detour-taking behavior of salticid spiders stalking their prey (Hill, 1978; cf. Jackson and Cross, 2011, fig.2) suggests an ability to plan ahead, which, moreover, improves with practice (Edwards and Jackson, 1994). This behavior is hierarchical in that it consists of several distinct phases of potentially variable duration, properly sequenced and concatenated to form a single overarching whole.
- (iii) *Embodiment and situatedness*. One of the many consequences of the embodied brain’s evolution-enforced need to care about its state and the outcomes of its actions is that the cognitive processes that engage with the self and the environment are, at a sufficiently high level, always emotionally charged (Carver and Scheier, 1990; Pessoa, 2008; Lindquist and Feldman Barrett, 2012; Damasio and Carvalho, 2013). Furthermore, the environment, in addition to influencing goals and behavior through outcomes, interacts with the physical body to constrain (and, in well-adapted systems, to facilitate) the solutions to problems that arise (Anderson et al., 2012). The interaction between the self and the environment may also give rise to entire new classes of computational problems (as in the case of social behavior; Galef, Jr and Giraldeau, 2001; Lehmann et al., 2008; Weitekamp and Hofmann, 2014).

⁴Cf. Epstein (1991, p.362): “The click of a microswitch suggests, falsely, that a discrete ‘response’ has occurred, but the rat is active continuously, and what occurs is multidimensional and complex.”

- (iv) *Agency and contingency*. Behavior does not reduce to S-R associations not only for reasons just mentioned, but also because it is often endogenous, that is, initiated by the animal itself (e.g., Thurstone, 1923; Brembs, 2011). Furthermore, because the animal's goals may change with time or as a result of environmental contingencies, many types of behavior cannot be fully planned ahead, requiring instead dynamical control by the brain (Cisek and Kalaska, 2010; Cisek, 2012). Foraging herbivores, for example, typically balance search and consumption of different types of plants, access to water, predator avoidance, shelter-seeking, social drives, etc. (e.g., Kamil and Sargent, 1981; Sallabanks, 1993).

Critically, the general computational problem implied by the above characteristics is that of *control* (Llinás and Iberall, 1977; Cisek and Kalaska, 2010; Heisenberg, 2014), which I turn to next.

2.3 The ubiquity of control

The need for control is universal in the animal kingdom. A common assumption is that only multicellular animals that are equipped with a nervous system are capable of complex behavior and therefore are in need of control. This assumption is implicit in the following quote from Savage-Rumbaugh and Lewin (1994): “All organisms with complex nervous systems are faced with the moment-by-moment question that is posed by life: what shall I do next?” The evolutionary roots of behavioral control do, however, run deeper: as Llinás and Iberall (1977, p.233) remark, “All brains are basically a progressive elaboration of ancient (primitive) cellular themes [. . .] The command-control properties of multicellular organisms are viewed as derived from the elaboration of themes already present in single cell forms.”

Indeed, any animal with more than one task to perform has an implicit control problem: when to feed, when to reproduce, etc. These choices must be arbitrated by the animal's internal physiology, which is charged with carrying out the requisite computations. A type of activity that is particularly prone to giving rise to decision and control problems is locomotion, which, according to Heisenberg (2014), “came about early in evolution, already with the prokaryotes.”^[A]⁵ In metazoans that are capable of active locomotion, the need for control becomes pervasive particularly if locomotion can be used in the service of more than one behavioral goal, as noted above. The behavioral control problem increases in urgency, requiring fast solutions and a way of hierarchically organizing goals and relating them to contingencies — that is, a neural control system.

In language too, contingent control is critically important. The two postulates that underlie the standard view of language (section 1.1) imply that linguistic behavior consists of emitting (or parsing) a premeditated and preplanned sequence of elementary actions (vocal or other gestures), in which a disfluency cannot but signify a failure, from which the speaker may or may not recover. This, however, is not what language looks like in the wild, where utterances are worked on jointly by the interlocutors (see Table 2 for an example and section 3.3.6 for a discussion), where speakers take liberties with “sentence” completion and structural “well-formedness,” and where listeners not only do not find “disfluencies” disruptive, but are actually aided by them (Arnold et al., 2003; Ferreira and Bailey, 2004). Thus, normal spoken language shows all the signs of being generated and processed on the fly, incrementally (Jackendoff, 2011, p.600), which implies that it must be subject to real-time, dynamic control.^[B]

Over and above basic decision-making and behavioral control, the computational problem of real-time, multidimensional, dynamic control that arises in language is common to all complex behaviors, whether individual (monadic) or social (dyadic, etc.). A typical example of the former variety is solo downhill skiing in unfamiliar complex terrain that includes ice patches to be avoided, pistes to be skirted, makeshift

⁵Capital letters in brackets refer to longer notes, which have been relegated to the end of the main text.

1 KATHY; I don't know this one so,
 2 NATHAN; You don't know how to ζ^do this one?
 3 (1.1)
 4 So ^we in ^trouble.
 5 (1.4)
 6 KATHY; Well ^you apparently knew how to do it.
 7 (0.2)
 8 NATHAN; Did I get it ζ^right?
 9 (0.3)
 10 KATHY; (H) Well you didn't (0.3) get the ^whole thing right.
 11 (0.2)
 12 NATHAN; @@[@ (H) ##]
 13 KATHY; [(H)
 14 But you—] —
 15 Well you just ^missed one ^part of it.
 16 (0.4)
 17 NATHAN; So what's that ^problem.

Table 2: A snippet of dialogue from the Santa Barbara Corpus of Spoken American English, reproduced from (Du Bois, 2014, p.373). Numbers in parentheses are timed pauses, in seconds; (H) is audible inhalation; brackets [and] mark beginning and end of simultaneous speech; @ is laughter; for the other symbols, see (Du Bois, 2014, p.361).

trampolines to be used for “catching air,” and glades to be navigated. A skier cannot plan the entire descent ahead of time, except in abstract terms (as in deciding to follow blue-square rather than black-diamond trails), and must thus deal with various contingencies as they arise. If the trails are crowded, these include the need to take into account individual and social actions on part of other agents.

Sustained social behaviors that require control may be adversarial, as in dominance fights between bighorn rams (Shackleton and Shafak, 1984) or non-adversarial, as in the courtship behavior in the blue-capped cordon-bleu (Ota et al., 2015), in which both the male and the female vocalize, bob up and down on the perch, and tap-dance. In humans too, dance is a common social behavior; a vivid and memorable example of the kind of dynamic control and coordination that it requires is the tango *Recuerdo* from Carlos Saura's 1998 film *Tango, no me dejes nunca* (see Figure 1 for the formal notation that captures the first 30 seconds of this dance).⁶

Of course, in humans the primary medium of social interaction is *verbal* behavior or “*linguaging*” (Linell, 2016; see Table 2 for an illustrative sample of a dialogue). The ultimate problem of control in language is, however, the same as in any other behavior: whereas its general formulation is “What shall I do next?”, in language it becomes “What shall I say [*and* do] next?”. Thus, insofar as it requires choosing a course of action, behavioral control is always an exercise in *navigation*: “Where, in the space of possible choices, should I go?”

⁶The *Recuerdo* can be viewed on Youtube at https://www.youtube.com/watch?v=_JIZjpywxPk (music by Osvaldo Pugliese, choreography by Juan Carlos Copes).

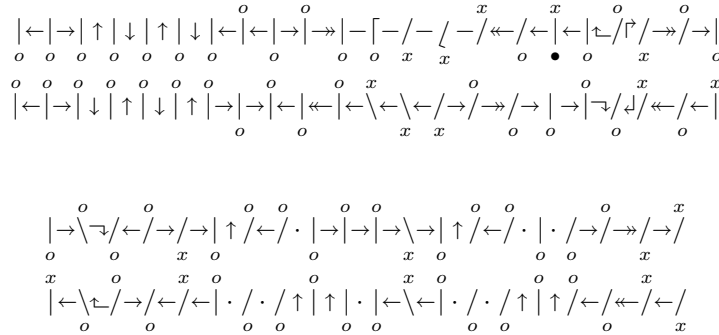


Figure 1: Dance notation for the first 30 seconds of the tango *Recuerdo* (Bodirsky, 2006). The two tracks describing the dancers’ moves include a line break, so as to make the entire passage fit on the page.

2.4 Navigation: the prototypical problem in behavioral control

That all problems in behavioral control, when considered at a proper level of abstraction, reduce to navigation is trivially true for actual locomotion and way-finding in physical space, as in the skiing example offered earlier. Less obviously, this is also true for the cases in which the “space” that is being navigated is abstract. For instance, vertebrate *posture* is naturally defined in the space of possible skeleto-muscular configurations, whose dimensions are the joint angles; any *movement*, including, say, a step in courtship dance, amounts to controlling the muscles so as to bring about an appropriate change in the configuration. Performing a vocalization while dancing amounts to the addition of some more dimensions, corresponding in the case of birdsong to the syrinx muscles, to the motor control space. Finally, the space can be entirely abstract, as in *problem solving*: for instance, *tool use* (e.g., in food retrieval in crows; Chappell and Kacelnik, 2002; Taylor et al., 2007; or stone knapping in humans; Rein et al., 2013) requires way-finding in the space of states and relative configurations of elements that define the problem.

The navigation metaphor applies also to socially coordinated/shared behaviors, such as the song-and-dance courtship routine of a cordon-bleu dyad or a tango performed by a pair of humans. The chief characteristic of multi-actor behaviors is the constraints on each individual’s actions that are imposed by the others’ actions and that must typically be accommodated in real time. I shall discuss socially constrained problem-space navigation in some detail in section 2.6 below.

In natural way-finding behaviors, some outcomes may be associated with reward. In physical-space navigation, one straightforward example is *prey-stalking*, where the hunter’s goal is to approach the location of the prey (which may change over time) unobserved. Another example is foraging, where the goal is to obtain resources that are distributed across a terrain. Interestingly, occasionally rewarded foraging (over and above navigation as such; Kazakov and Bartlett, 2005) can serve as a basis for the evolution of complex behaviors in abstract spaces, including language (Arsenijević, 2008; Kolodny et al., 2015). While the nature of reward may vary from problem to problem, it is central to the exploration of any behavior-related space, concrete or abstract. In particular, because outcomes always shape behavior, sequential *decision making* reduces to finding a path through the abstract space of actions (which may or may not include physical movement) that maximizes cumulative reward, as in reinforcement learning (Pack Kaelbling et al., 1996; Sutton and Barto, 1998).

In this connection, one may ask why not elevate *search*, rather than navigation, to the status of the Ur-problem in behavior. In addition to being applicable in any situation that involves a state space (physical or

abstract), just like navigation is, search is by definition goal-directed, the goal being the subset of states that need to be attained for the search to terminate successfully. The idea of search is central to the discipline of artificial intelligence (e.g., Dechter and Pearl, 1985; Silver et al., 2016), where a classical example is the reduction of a game of chess or Go to a search of the abstract space of all legal board configurations, subject to constraints over moves and the pursuit of a winning state (as prescribed by the rules). In psychological theory too, the generality of the concept of search has recently received some recognition (Hills et al., 2015).

Board games are, however, a far from perfect metaphor for natural behavior, in part because their rule sets are both inviolable and immutable, their reward schedules are fixed, and their state spaces are closed (Edelman, 2015). Likewise, the concept of search connotes the existence of a definitive goal, whose attainment terminates the process. In comparison, navigation — particularly of the exploratory kind⁷ — is a more open-ended, ongoing behavior that unfolds dynamically, usually under simultaneous control of multiple factors, such that neither these factors nor indeed the space itself are fixed ahead of time. In what follows, I shall, therefore, refer to my “working metaphor” for general complex behavior as navigation.^[D]

2.5 Constrained navigation

While many types of behavior reduce to navigating an “open” space that may be changing even as it is being traversed, the agent’s moves are typically still constrained by a variety of factors, just as in language (more about which in section 3) not all combination of elementary gestures are likely to be equally effective. Indeed, were it not for the constraints on the moves, behaviors such as language would not even be learnable (Edelman, 2008b, section 4.3).

The prime type of constraint on navigation is determined by the *connectivity* of the space in which it takes place: if a shipwreck strands you on the proverbial tropical island and you cannot swim, it does not avail you that the mainland is within sight. If you can swim, but the channel is infested with sharks, you’re barely better off: now you have to decide whether to stay on your island and hope to be discovered and rescued, or swim, on the off chance that the sharks will look the other way. This exemplifies the other key constraint on navigation, namely, the relative *cost* of choosing one available path or course of action over another. Lastly, by building a raft out of flotsam, you can change the connectivity of the space you’re in: the mainland may now be within reach.

The mathematical discipline concerned with matters of connectivity is topology. Informally, the concept of topological space is built around the notion of neighborhoods: sets of points that are related (that is, belong together according to some criteria) and interconnected (that is, can be visited without breaking continuity). The most familiar example of a general topological space is the physical space in which we reside. An empty area (or volume of space) is continuously connected: within it, it is possible to get from any point to any other point, by any route. Place a rock in the middle of the lot, and the connectivity ceases to be continuous. To get from point A to point B, one must now pass on one or the other side of the obstacle; there are now at least two distinct categories of possible routes, which cannot be continuously morphed into one another. Because of the added constraint, the problem of navigating the area in question has assumed a discrete structure.

Spaces characterized by discrete connectivity are known in mathematics as *graphs*. The basic graph is a formal structure that consists of a set of vertices and a set of possibly directed (anisotropic) edges that connect some of the vertices to others. Vertices and edges may both be labeled or otherwise annotated; for example, each edge may be assigned a real-number weight representing its strength or some other graded

⁷Think of the great transcontinental or ocean-going voyages of discovery.

quantity associated with it (see (Chartrand, 1992) for an overview of basic graph theory; Harel (1988) describes higraphs, which is a more expressive formalism).

Revisiting some of the behavioral problems mentioned earlier, it is easy to see how detour-taking and foraging can be described in terms of graph traversal. In the case of detour-taking, the vertices may stand for regions of space that differ qualitatively in their visibility to the intended prey, and the edges — for accessibility relations given the stalker’s locomotor options. Similarly, a foraging problem can be defined over a graph whose vertices represent locations in a terrain and are labeled with the resources they offer, and whose edges carry weights that stand for the metabolic cost of their traversal.⁸ Under this representation, foraging may be reduced to finding a path through the graph that strikes a balance between being short (in terms of total edge weight) and yielding a high cumulative return at the vertices.⁹

2.6 Navigating socially constrained and constructed spaces

Applying the constrained navigation metaphor to social behaviors raises several questions. Consider, for instance, how in a social species foraging or hunting (or skiing, as in the earlier example) may require that individuals take into account the presence and the actions on part of their neighbors while planning their own behavior (Stander, 1992; Galef, Jr and Giraldeau, 2001; Melis et al., 2006). In such cases, in what sense is the relevant space shared among the participants and how are one participant’s options affected by the others’ moves?

Whereas in group skiing the space on the slope is literally shared (thereby making the skiers’ moves interdependent), in a tango dance (Figure 1) only some of the dimensions, those of the physical volume of the room, are fully shared; others, such as the postural configuration dimensions of the dancers, are merely mutually constrained. Thus, one dancer’s actual location and posture rule out a set of the other’s otherwise perfectly plausible locations and postures entirely; reduce the probability of another subset; and leave the rest unaffected. It is up to the second dancer to pick from among the moves that remain possible the ones that best serve his or her current goals. One of these goals can be bending the behavior of one’s partner to one’s own ends. In a board game such as backgammon or Go, this can be done through strategic placement of one’s pieces so as to restrict the opponent’s moves, or re-weight their relative appeal, and thereby gain positional or material advantage.

In a crucial divergence from board game players, however, real-life agents navigate spaces to which they, or their partners, can add, on the fly, new paths (as when a new shortcut is made while traversing an otherwise familiar terrain), new regions (as when the adoption of a new tool extends the reach of an established behavior), or even entire new dimensions (as when hitherto purely vocal singing behavior gains an accompaniment on a newly invented musical instrument).¹⁰

3 The nature of language

We are now ready to apply the conceptual tools mustered in the preceding section — control, navigation, constraints — to try to understand the nature of language and its use, with a particular focus on how users of language could, by navigating their space of actions, influence others’ (and their own) future representational states and behavior. The following formulation will serve as a working hypothesis:

⁸For instance, monkeys that forage in a rain-forest tend to keep to certain distinct routes that form a graph (Di Fiore and Suarez, 2007, fig.1), presumably because travel outside established paths in the forest is significantly more costly.

⁹This is a well-known optimization problem, which is encountered, for instance, in goods distribution and shipping.

¹⁰For a discussion of evolutionary novelty and “ontological emergence,” see (Kauffman and Clayton, 2006).

Language behaviors reside in a graph-like space of sequences of multimodal gestures that embody a game of influence between interlocutors and that are subject to multiple, dynamically changing influences and contextual constraints.

The first order of business is to see why the space of language behaviors is “graph-like,” as opposed to simply being a graph.

3.1 Language as constrained navigation in a graph-like space

Spoken language production amounts to navigating the (abstract) space of sequences of phonemes — a space that is, to a good approximation, discrete (as are the phonologies of sign languages (Sandler, 2006), to which all the observations and arguments offered here also apply).^[E] Because natural speech is always multimodal (Vigliocco et al., 2014), the sequence of phonemes that comprises a speech act is accompanied by other, concurrent and coordinated signals, which include prosody and facial and manual gestures (Havas and Matheson, 2013) that carry both emotional/diffuse and specific additional information, visual attention shifts, etc. (Kolodny and Edelman, 2015).

On the level of phonemes, speaking thus amounts to the traversal of a graph, whose vertices number in the dozens, the exact number depending on the particular language. Much more importantly for the actual use of language, the universal principle of the duality of patterning (Hockett, 1960) states that speaking also amounts to the traversal of another graph, one whose vertices correspond to lexemes — informally, words. I shall call it the word graph (Edelman et al., 2004; Solan et al., 2005; Edelman, 2008a).

Different individuals who share a language have at their disposal similar, but not identical, word graphs — surrounded by similar, but not identical, spaces of additional elementary behavioral choices, some of which are graded (continuous) rather than discrete. Discourse, then, consists not of “meanings” being packed into sequences of words and shipped back and forth, but of an interactive game in which the conversants navigate graph-like spaces using their individual, dynamically shifting degrees of freedom to steer and constrain the others’ moves in real time. Even in the cases where declarative knowledge is imparted by these exchanges, it is never fully contained within the “message,” but rather resides in the partly shared conceptual spaces and prior experience of the interlocutors (Ramscar and Baayen, 2013; cf. Stolk et al., 2016, fig.2).

3.2 Language learning and change

The multidimensional space in which language behavior unfolds, including the word graph that anchors it, is perpetually changing. Developmentally, the word graph that is harbored by an individual listener/speaker begins with nothing — no words and no paths. This is in contrast to the multimodal “halo” that surrounds it: disorganized as it may be to begin with, the space of paraverbal behaviors is present to some extent already in the neonates. Guided by developmental factors, which include vastly important emotional and social cues (Goldstein et al., 2010), and even some explicit guidance on part of the caregivers (e.g., Moerk, 1976), the entire space grows and changes with experience as newly learned words join the set of vertices and newly encountered utterances modify the set of edges — and thereby the set of immediately available paths through the graph.¹¹

In addition to taking any of the available paths at a given juncture, a speaker can also decide to break with precedent, jump tracks, and proceed from an entirely different location in the space of possibilities, a

¹¹Cf. Hockett (1968): “A language is a kind of system in which every actual utterance, whether spoken aloud or merely thought to oneself, at one and the same time by and large conforms to (or manifests) the system, and changes the system.”

destination that is not connected directly (that is, by an existing edge of the graph) to the point of departure.¹² Because of such possibilities, the best analogy for the graph-like aspects of language may be not a railroad yard or a mass transit system (Edelman, 2008a, p.274), but the one offered by Sampson (2007, p.10): “the grammatical possibilities of a language are like a network of paths in open grassland.” In any case, the never-ending lexical and structural innovation in language, which in the subway system analogy may be likened to the building of new stations, excavation of new tunnels, and abandonment of old ones, amounts to what Beckner et al. (2009, p.15) called the “perpetual dynamics” of language change.

Because change can affect virtually all the aspects of the language representation space and its use, both the initial acquisition of language and its subsequent modification through use likely involve all the learning mechanisms to which the human brain has access (Atallah et al., 2004). In particular, to the extent that language is an effective tool for seeking social and other rewarding outcomes, such as influencing others or gaining access to resources (and may even have evolved in part under the selection pressures from these factors; Herrmann et al., 2007; Pinker, 2010), language acquisition and change must rely on reinforcement learning (RL; Pack Kaelbling et al., 1996; Littman, 2015), both of the model-free and model-based varieties.¹³

However, even model-based RL, in which the learner gradually builds an internal representation of the environment that it navigates, does not quite capture the nature of the problem at hand, because like RL in general, model-based RL still assumes that representational states are to be associated with specific actions. In comparison, in real-time language behavior the actions are picked dynamically, in response to moment-by-moment changes in the situations, which may well have no precedent in the agent’s history. For a detailed discussion of the reasons because of which the standard formulations of RL are less than adequate for realistic behaviors, see (Edelman, 2015).

3.3 Dependency: the universal type of structural constraint in language

The many diverse factors that dynamically shape language behavior can all be expressed in terms of a common computational primitive: *dependencies*. The concept of dependency is a relatively theory-neutral structural primitive, which can express all the “syntactic” characterizations of language, and which is therefore widely used in both formalist and computational linguistics (Phillips, 2003b; de Marneffe et al., 2014). At the same time, it is sufficiently powerful to express the other, “extra-linguistic” influences on the dynamic production and comprehension of language, if these processes are viewed as ongoing constraint satisfaction (Levy et al., 2009; Bullock et al., 2009).

Dependencies can therefore serve as a bridge between the behavior-general theoretical framework based on the computational concept of navigation on the one hand and the existing body of data in linguistics on the other hand.¹⁴ To help along with this conceptual bridging, the following brief breakdown of dependencies by type includes examples both from language and from other behaviors.

¹²As in the subterranean mining cart chase scene in Spielberg’s film *Indiana Jones and the Temple of Doom*; compare (Edelman, 2008a, Figure 7.21).

¹³Note the computational universality of the RL formulation: “any task with a computable description can be formulated in the RL framework” (Schmidhuber, 2015, p.100).

¹⁴This requires, however, that the idea of generative grammar be abandoned (Edelman, 2017b), in favor of a descriptive approach, such as Postal’s (2008) model-theoretic one.

3.3.1 Dependency: basic sequential structure

A clear tell-tale of the presence of structure in natural language utterances is non-random word sequencing: the “departures from equiprobability” (Harris, 1991, p.32) in the serial order of the words that comprise them. Indeed, in the simplest cases, a linguistic dependency is merely the elevated probability of encountering a word following another word, so that the former can be said to depend on the latter. More generally, it has been observed that “the presence of a word [...] depends not on particular other words but on the dependence properties of the other words” (Harris, 1991, p.52) — a provision that quickly escalates the intricacy of linguistic dependencies (Mel’čuk, 2003). Nevertheless, dependencies in language are fundamentally similar in kind to those found in other graph-based problems such as navigation or foraging.

Now, some word sequences are sufficiently cohesive to warrant being made into units in their own right. Thus, comparing a number of utterances to one another after aligning them on shared words reveals *collocations* — multi-word sequences, whose status as units of representation (e.g., Arnon and Snider, 2010; Bannard and Lieven, 2012) is statistically motivated (i.e., they have to be more common than warranted by chance; cf. Barlow’s (1990) notion of suspicious coincidence). Outside of language, collocations correspond to structures such as “chains” in habitual motor tasks (Aldridge and Berridge, 1998) and beaten paths in familiar-terrain navigation (Di Fiore and Suarez, 2007).

3.3.2 Dependency: hierarchical structure

If the newly discovered collocations are added to the lexicon and the discovery process is applied recursively, the lexicon becomes a set of units some of which possess *tree-like* hierarchical structure, similar to the constituent structure postulated by most linguistic theories (see Solan et al., 2005 and the references therein). This process corresponds to a rewiring and recoding of the graph that represents the listener/speaker’s experience: while some of its vertices are still isolated words, others are hierarchically structured supra-word units.¹⁵

The language user can generate utterances by reading out the units while traversing the grammar graph (descending in each case of a hierarchical unit down to the level of words) and can analyze an utterance by assigning it a properly structured representation (a “parse”), as dictated by the traversal; the “meaning” of the utterance — that is, its effect on the listener — is then constrained by its recovered structure. As before, note that such representations have parallels in navigation and in more abstract problem solving, in the form of recursive path or task decomposition (Botvinick, 2008; cf. Jackendoff and Pinker, 2005). Indeed, the basic type of hierarchy in dependency structures may be so common as to be shared between music-, math-, and language-related tasks (Van de Cavey and Hartsuiker, 2016).

Tree-structured hierarchical dependencies are commonly assumed to mediate a range of linguistic phenomena such as binding in anaphora¹⁶ (Phillips, 2003a). Differences in the posited hierarchical dependencies are also thought to account for cases in which there is a big difference in interpretation or in acceptability¹⁷ between two utterances that seem almost identical.

The so-called tough movement adjectives such as “eager” and “easy” exemplify this phenomenon: they work equally well in some contexts (“Abe is eager to read”; “Abe is easy to read”), but cannot be interchanged in others (compare “to read is easy” and “to read is eager”). This type of structure dependence of

¹⁵Technically, this makes it a higraph (Kolodny et al., 2015).

¹⁶For instance, the binding of “Abe” and “himself” in “Abe told Bob that he hurt himself.” The patterns of acceptability in anaphor binding may arise from thematic constraints (Pollard and Sag, 1992).

¹⁷While the perceived acceptability of utterances is always graded rather than all-or-none (Schütze, 1996), it is sometimes sharply graded.

path acceptability is, however, not as formidable as it sounds: it appears to be learnable from distributional (corpus) data.^[F] Moreover, sharp differences in behavioral generalization between items that are found in very similar contexts are not exclusive to language. Consider the difference between regular and French windows: both admit light, but the underlying scene structure induces people to disprefer exiting a room via the former. And just as in language, dispreferred does not mean absolutely ruled out: a person may well decide to exit via a window if the room is on fire and the door is blocked.

3.3.3 Dependency: long-range sequential structure

Sequential dependencies in utterances, such as number or gender *agreement* between noun and verb, can leap over one or more — sometimes, many — intervening words. If the intervening words are all subsumed under the same root in a tree-structured unit, the dependency could be considered short-range, skipping just one unit (this is the case, e.g., in the tree-adjoining grammar or TAG formalism; Joshi and Schabes, 1997). Given that human languages do tend to minimize actual word-level dependency length in their utterances (Futrell et al., 2015), it seems likely that long-range dependencies are indeed psychologically real.

In terms of graph structure, non-local dependencies correspond to several paths that are initially distinct, then traverse the same sequence of items, then part ways again. Tasks with this structure are commonly used in studies of animal navigation (e.g., the 8-maze; Wood et al., 2000). Following a proposal by Levy (1996, fig.2), Eichenbaum et al. showed that rats can learn to perform an analogous navigation task set in an abstract space of odor sequences (Agster et al., 2002; Fortin et al., 2002).

3.3.4 Dependency: categorical abstraction in context

If a set of paths through the word graph can be partially aligned so that some of their vertices match, but others do not, the non-matching vertices are thereby shown to belong to a *category* — namely, the equivalence class of words that are interchangeable in the context defined by the matching parts of the aligned paths (Harris, 1946; Solan et al., 2005). In principle, this procedure motivates syntactic categories such as parts of speech; in practice, for reasons of semantics such categories are much narrower than that.¹⁸

In non-linguistic tasks, categorical grouping corresponds, for instance, to situations where several alternative paths, perhaps differing in some quality related to another task, lead to the goal. The latter may be a physical location, as in navigation, or an abstract state, as in problem solving. In particular, while navigating from work to home, one can travel by a direct route, or by making a detour via a grocery store to pick up the ingredients for dinner.

A distinction is commonly made in language between open-class and closed-class word categories. Words that belong to open-class categories such as various subtypes of nouns, verbs, and adjectives, connote content and are relatively easily learnable: people can usually pick up a novel open-class word from context and gradually master its use. The much less numerous, yet much more frequently encountered, closed-class or function words anchor the form and import (“meaning”) of utterances. A novel item encountered in a slot in which a function word is expected (e.g., on the basis of phonological cues; Dąbrowska, 2004, pp.164-165) is much more difficult to make sense of. There is evidence that infants gradually learn to recognize function words over the first year (Shi et al., 2006) and that they use for this purpose distributional cues, especially the frequency of occurrence (Hochmann et al., 2010).

¹⁸Consider the relative merits of filling the slot in the expression “dead ____” by any random member of the noun category (say, “chair”), as opposed to just the members of the set $AnimateObjects \cup \{calm, ringer\}$.

A possible counterpart for the closed-class or function words outside language may be found in motor act planning and sequencing. The constraints of the human skeletomuscular system (say) allow a person to easily learn actions that are drawn from a very broad — and open-ended — class; at the same time, there is a small and difficult to extend repertoire of “special” actions. In free-style lap swimming in a pool, for instance, technical inhaling/exhaling and flip-turning are closed-class elements, while the actual swimming is composed of open-class elements.

3.3.5 Dependency: multimodal concurrency and synchronizing events

As already mentioned earlier, the symbolic or “digital” aspect of language — the sequences of discrete words uttered by speakers and analyzed by listeners — is always accompanied by several concurrent streams of additional information such as “analog” or continuously variable prosody, gestures, etc., typically spanning several modalities (Vigliocco et al., 2014; Kolodny and Edelman, 2015; Roy et al., 2015). In terms of dependencies, this implies that the presence and the precise form of a discrete element in the “main” sequence may depend not only on other discrete elements, past and future, but also on certain graded and possibly quite subtle features of the accompanying analog streams.

These additional streams exert a multidimensional, dynamically unfolding influence that is traditionally deemed “extralinguistic” and thus rarely discussed in connection with language, let alone included in the corpus data typically used by computational linguists and psychologists. Among these analog streams are the (vector-valued) fluid emotional states of the speaker and the listener, as well as the constantly shifting motivational variables, which affect on the fly the communication process and therefore also the ongoing choice of words (and intonation, and gestures) and their interpretation.

In making a case for the multimodal, concurrent nature of all complex behaviors, including language, Kolodny and Edelman (2015) discuss some of the peculiar computational problems that characterize them. In particular, the normally asynchronous processes that comprise behavior are occasionally checked, because one or more conditions on which further progress depends has not yet been met. This type of stop-and-go dynamics, which effectively synchronizes the processes involved in it, gives rise to *events* — time windows during which some of the variables in question remain practically constant. Synchronization can also be carried out by a periodic signal. An example of a periodic factor that induces a cross-component dependency is breathing: in olfaction, it helps shape the neural dynamics of the olfactory bulb activity (Skarda and Freeman, 1987), while in language it affects prosody, utterance segmentation, and turn-taking in conversation.

3.3.6 Dependency: structure distributed across individuals

Social aspects of language are central both to its acquisition by novice users and to its continuous shaping through use. The act itself of using language is a socially grounded experience, arguably even when we think/speak to ourselves (Lewis, 2002).¹⁹ On the social- and discourse-centered view of language discussed earlier, linguistic behavior is coordinated across speakers and listeners. This coordination manifests itself on many levels in dependencies that span the inter-subject “gap.” For example, interlocutors tend to structure their utterances similarly (Bock, 1986; Pickering and Branigan, 1999); the timing of their actions exhibits dynamical rapport (Dale and Spivey, 2006; Dale et al., 2013), and so do their brain activation patterns (Bilek et al., 2015). Instances of such coordination can be found in animal social behaviors, from the

¹⁹Cf. the “community view” (Candlish and Wrisley, 2014) of Wittgenstein’s argument in *Philosophical Investigations* regarding the impossibility of a private language (Wittgenstein, 1958, §§244–271).

courtship dance in *Drosophila* (Anderson and Perona, 2014) to coordinated hunting strategies in prides of lions (Stander, 1992).

In language, these phenomena are so pervasive as to have inspired a dialogic theory of syntax (Du Bois, 2014), according to which the patterns of linguistic dependencies can only be fully described and understood if one studies complete corpora, which include the contributions of all the participants in each bout of discourse dialogic discourse (cf. Clark and Schaefer, 1989; Lewis, 2002).²⁰ This view fits well within the general approach advocated here, according to which language behavior is generated by a multidimensional set of concurrent processes that transcend the boundaries between individual speakers, with foci of activity proceeding in parallel down multiple paths in each speaker/listener’s graph, all the while interacting dynamically within and between participants.

4 Neurocomputational models of language

An adequate computational understanding of these processes, without which they cannot be modeled, is still lacking. Revisiting Lashley’s (1951) problem of serial order in behavior, Houghton and Hartley (1996) claimed that “cognitive science has no (neuro-) psychologically grounded theory of serial order”^[G] — a claim that remains difficult to argue with twenty years later (Kolodny and Edelman, 2015) and that certainly applies to language *qua* serial behavior. The best one can do here, then, is round up potential candidates for such a neurally informed theory and see which directions seem the most promising.

To that end, I first state some methodological criteria for modeling, then proceed to survey the major computational approaches to modeling specifically the serial order in language. Because no existing model even begins to address the multimodal nature of language, let alone account for its ongoing dynamics, the discussion in this section focuses on how the various models deal with the problems arising in graph-constrained navigation, it being understood that these are only a small part of the big picture.

4.1 Methodological criteria

Any viable cognitive-computational model of language must meet several theoretical and empirical validity criteria (see also Edelman and Waterfall, 2007; Waterfall et al., 2010):

- I **Dynamic coordination.** It must support the coordination of activities and adequately convey the conceptual and goal-oriented intentions of interlocutors, while dynamically balancing the multiple constraints that shape discourse.
- II **Multimodality and concurrency.** It must allow integration with the “extralinguistic” aspects of language behavior, as discussed earlier.
- III **Learnability.** It must include an account of language acquisition that is specific and practical enough to allow learning from experience: ideally, from interacting with fluent speakers, or, at a minimum, from a realistic corpus of behavioral data.
- IV **Production and comprehension.** It must support all aspects of language behavior, and in particular account for both comprehension and production.
- V **Neural plausibility.** It must rely on neurocomputational mechanisms that appear plausible in light of known brain neuroanatomy and neurophysiology (in the sense of appendix A).

²⁰For a rather more philosophical take on dialogism, see (Linell, 2013; Steffensen, 2015).

- VI **Scalability.** In learning and use, it must scale up to data sets that are realistic in their composition and size.
- VII **Empirical adequacy.** It must replicate the many phenomena that characterize the behavior and performance of humans acquiring and using language.
- VIII **Formal power.** [This point is a legacy of the formalist approach to language, which should by rights be set aside (Edelman, 2017b). I include it here because some linguists and psychologists focus their efforts exclusively on issues of formal power.²¹] Its formal computational power (in the sense of appendix B) must match the structural characteristics of natural languages.

To the best of my knowledge, no existing model meets all of these criteria. Indeed, most are concerned either with criterion VII (empirical power), or with criterion VIII (formal power). The former focus on replicating a selection of empirical findings (which is often extremely narrow, e.g., the formation of the past tense in English verbs, and nothing else; or the auxiliary fronting in forming polar interrogatives in English, and nothing else). The latter may aim to meet the formal requirements and so focus exclusively on grammar (a theoretically barren notion, as noted earlier), with no regard to any of the other constraints — not even neural plausibility.

While the plausibility of a particular mechanism is often a matter of debate, relative rather than absolute plausibility estimates should be less controversial and can still be useful. As an example, consider two operations on strings, *rewriting* and *adjoining*, each of which can be used to define a grammar (for whatever it is worth; see appendix B). Context-free (CF) and context-sensitive (CS) families of languages are generated by grammars consisting of rewriting rules such as $NP \rightarrow the N$ that involve both terminal strings (here, the determiner *the*) and non-terminals or variables (here, NP and N), which can in principle be bound to any terminal or other variable. In contrast, contextual languages (CL) are generated by adjoining (prepending and/or appending) terminal strings, drawn from specific categories, to existing predefined “seed” strings (Marcus et al., 1998). Arguably, the operation of adjoining is more neurally plausible than rewriting:

It is hard to imagine our brain using auxiliary intermediate sentences of a nonterminal type. Instead, it looks more *natural*, in the proper sense of the word, to start with a collection of well-formed sentences, maybe acquired from experience, and to produce new well-formed ones by adding further words, in pairs that can observe dependencies and agreements, and in accordance with specified selectors, which can ensure the preservation of grammaticality.

— Marcus et al. (1998, p.270)

The difficulty with rewriting-based approaches lies with their reliance on variables and symbolic binding. Although binding through synchrony or phase-locking between the responses of units that may be far apart in the brain (Singer and Gray, 1995; Engel and Singer, 2001) is often assumed to be available as a building block in neural modeling (e.g., Hummel and Biederman, 1992; Hummel and Holyoak, 1998; Hummel, 2001), this assumption sidesteps some key questions: by what means synchrony is achieved in the first place, whether or not it is as flexible as abstract variable binding needs to be.

Settling for a static (circuit-based) instead of dynamic (binding-based) approach to structure (Edelman and Intrator, 2003) resolves these issues, but necessarily results in less than fully flexible variable-value “binding.” Even such limited binding often exacts a price in terms of neural plausibility. For instance, the

²¹ As in the to-do surrounding on context-free languages of the form $a^n b^n$ (Fitch and Hauser, 2004; Gentner, Fenn, Margoliash, and Nusbaum, 2006; see Pinker and Jackendoff, 2005, p.216 for a critique).

model of Borensztajn et al. (2009) owes its ability to process syntactic structures to the built-in operation of substitution, which, despite being less general than global variable binding, requires “a capacity for the nodes in [the] network to transmit and store richer information than just activation levels.”

The trade-off between neural plausibility and model power is exemplified also by the recently developed model of “symbol-like processing” in the prefrontal cortex and basal ganglia, which is, in the words of its creators, limited in that “It can only assign pointers to memory locations with which it has had some previous experience, and those locations can only represent information that has been learned to be represented. Also, neural pointers cannot be nested at arbitrary levels of complexity or depth. In this respect, these neural representations fall short of qualifying as symbols in the most general sense” (Kriete et al., 2013, p.16390). Kriete et al. (2013) proceed to note that human symbolic computation is similarly constrained (cf. Johnson, 2004) — a point that resonates with my own past arguments (Edelman and Intrator, 2003; Edelman, 2007) and that in turn supports my contention that neural plausibility must serve as a useful and constructive criterion for choosing between competing language models.

4.2 Types of models

With the above-mentioned criteria in mind, I now offer a rough blueprint for classifying and evaluating cognitive-computational models of language. Traditionally, solutions to Lashley’s (1951) problem of serial order have been classified as associative (based on response chaining between successive elements), positional (based on coding the absolute order of elements in the sequence), or hierarchical (based on multiple levels of abstraction in representing sequential order), Lashley himself favoring the latter.

A rather different approach to serial order, competitive queuing (CQ), abandons sequentially structured representations altogether and shows instead that serial order can arise out of a parallel representation through a series of winner-take-all steps, in each of which the next element in the would-be sequence is selected via competition (Houghton and Hartley, 1996; Bullock, 2004). I shall return to this intriguing concept in section 4.6.

Because my prime concern in this paper is with language (rather than other, simpler serial behaviors) and with how it is implemented in the brain, the classification scheme that I offer below is based on how different models engage with language as a graph-structured problem (section 3.1) and how they account for dependencies (section 3.3), as well as how they could be mapped onto brain circuitry (section 5). On the basis of these characteristics, I discern four categories of models:

- *Symbolic models* (section 4.3). These models directly implement fully general, abstract structures and operations. An attempt to model language with a Probabilistic Context-Free Grammar (PCFG) based on a putative neural implementation of rewriting rules would belong under this rubric.
- *Recurrent (“folded”) distributed networks* (section 4.4). The popular Simple Recurrent Network (SRN) architecture belongs here. Its successive layers are fully interconnected and the representations that it maintains are uniform (no units dedicated to specific functions) and distributed. This architecture is folded in the sense that the full sequential structure of a typical utterance is implicit in the weights of the network; generating it requires multiple passes around the recurrent loop.
- *Extended (“unfolded”) localist networks* (section 4.5). These spell out the possible utterances explicitly, via the sequential patterns of connectivity in non-uniformly wired circuits of neurons, which are therefore functionally not necessarily all alike. The label “extended” alludes also to the distinction between extensional and intentional definitions of a set: by listing its members vs. by stating the common properties that they share.

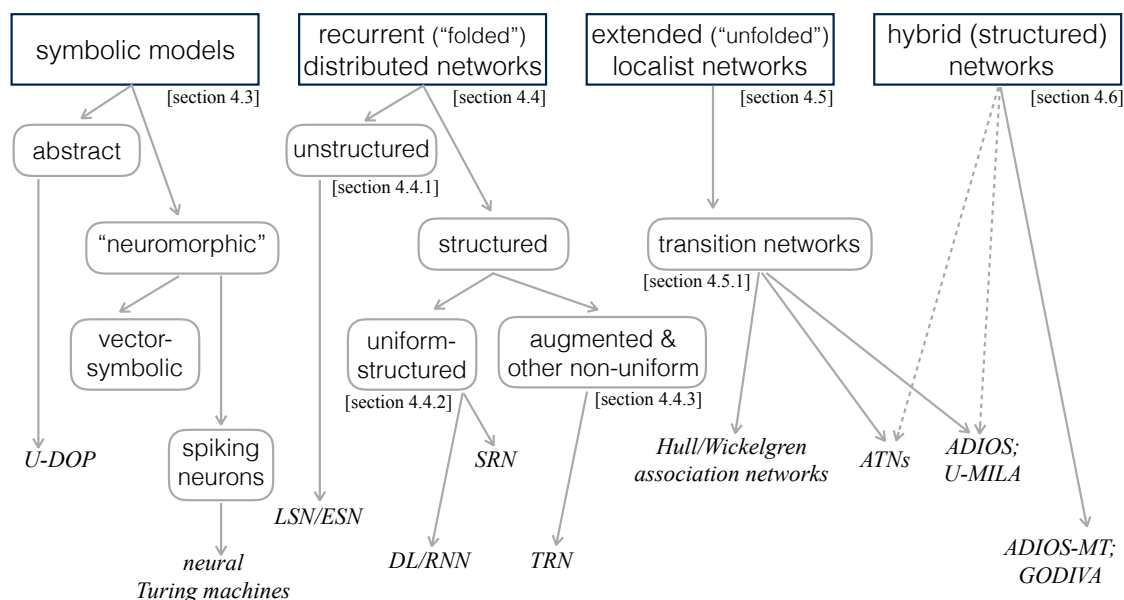


Figure 2: A sketch of the relationships among the language models discussed in section 4 (subsection numbers as indicated). The four classes of models appear in the boxes along the top of the figure. Some specific models, each of which may inherit features from more than one class (or sub-class, shown in cartouche) are named at the bottom, in italics (left to right): U-DOP, data-oriented parsing; (Bod, 2009); neural Turing machines, (LeCun et al., 2015); LSN/ESN, liquid state / echo state networks, (Maass, 2007; Tong et al., 2007); DL/RNN, deep learning / recursive neural networks, (LeCun et al., 2015); SRN, simple recurrent networks (Elman, 1990); association networks, (Hull, 1934; Wickelgren, 1979); ATNs, augmented transition networks, (Woods et al., 1972); ADIOS, automatic distillation of structure, (Solan et al., 2005); U-MILA, unsupervised memory-based incremental language acquisition, (Kolodny et al., 2015); ADIOS-MT, an application of ADIOS to machine translation (Edelman and Solan, 2009); GODIVA, gradient order directions into velocities of articulators, (Bohland et al., 2010).

- *Hybrid structured networks* (section 4.6). Approaches that combine elements of the above three strategies are possible and some have indeed been tried in the past. A hybrid architecture that I consider promising combines a localist representation of the elements of a sequence and an extended representation of element collocations with a recurrent loop that implements a flexible, yet constrained sequencing of the elements.

A sketch of this taxonomy appears in Figure 2. In the rest of this section, I discuss the four categories of models in some detail, including occasional remarks on their neural plausibility — the critically important issue to which section 5 will be devoted.

4.3 Symbolic models

In developing a symbolic model of brain function, the first issue is that of representation, where the many possibilities are often glossed over by highlighting the choice between distributed and localist schemes (Bowers, 2009). Choosing a distributed representation scheme leads to a particular formal framework —

the geometry of multidimensional vector spaces and manifolds — that is considered neurally plausible (Mumford, 1994; Edelman, 1999). To use it in modeling symbolic computation, one needs to be able to express the relevant formal operations such as variable-value binding in terms of operations on vectors, as it was done for instance by Plate (1995). Vector symbolic architectures have been used in modeling the representations of objects, relations, and sequences (Gallant and Okaywe, 2013), as well as syntactic dependencies (Basile et al., 2011). In language, however, it seems that such models do not easily scale up, having only been tested on the same semantic disambiguation and thematic mapping tasks that had been explored already by Plate (1995). In addition, as one may expect given that they implement abstract binding, vector symbolic models do not rank high on neural plausibility.

Models of symbolic computation that employ localist representations may or may not score better in that regard. A prominent example of a model that was motivated explicitly by what is known about the circuits connecting the prefrontal cortex with the basal ganglia is the already mentioned implementation of indirection and pointer-based processing by Kriete et al. (2013). These authors did not, however, aim to model all of language from the top down (indeed, the type of language tasks on which their model has been tested is the same as in Plate, 1995). In contrast, the spiking-neuron model of Fernando (2011) was designed explicitly at a direct implementation of rewrite rules, for which, however, it uses circuits tailored specifically for this purpose (rather than inspired by the circuitry of the brain). No tests of this model on actual language were reported.

The focus of symbolic neurocomputational models of language on rewrite rules — under the assumption that such rules capture the computational essence of language — illustrates a common feature of symbolic modeling: top-down flow of design choices. Alongside the choice of representation, it requires choosing the formal framework for describing language, such as PCFG (van Schijndel et al., 2012), which indeed involves rewrite rules. One notable exception is the work of Rodriguez and Granger (2016) on “the grammar of mammalian brain capacity.” Taking as a starting point the gross anatomy (including comparative data) and physiology of the cortex and thalamocortical and hippocampal-cortical circuits, they conclude that brain function is best seen in terms of “two constituent algorithms [. . .]: i) categorization of objects by similarity, and ii) chaining objects into sequences,” which together amount to the recursive application of rewrite rules. While this approach is not aimed specifically at modeling language (nor has it been tested on natural-language tasks), it is in a sense more “linguistic” than all the others mentioned here, in that it attempts to reduce everything that the brain does to following the rules of a grammar — an explanatory move that is diametrically opposite to the one that underlies the present paper.

Standing back from these specific efforts to model brain function in terms of symbolic computation, we should realize that the symbolic modeling framework may be fundamentally at odds with natural language. The dynamics of discourse and the fluid and distributed nature of meaning contribute to what Marcus et al. (1998, p.247) call “the ‘open’ character of natural languages, making it impossible to formulate a necessary and sufficient condition for a sentence to be well-formed.” The upshot of this insight is that one must abandon structural well-formedness — and perhaps the very notion of a complete “sentence” to which the concept of well-formedness would apply — as the centerpiece of linguistic theory.

This realization reveals critical shortcomings in the formal symbolic approaches that equate the knowledge of language with the possession of a grammar, conceived of as a complete set of rules for parsing and generating well-formed sentences (cf. Postal, 2004). The explanatory inadequacy of the formalist conception of grammar implies in turn that purely “syntactic” models, be they rule-based or statistical, are inherently incapable of explaining either production or comprehension. In production, the utterance that gets generated is never merely well-formed or highly probable (two characterizations that are difficult to distinguish empirically; see, e.g., Lau et al., 2016): it is the one that the utterer intends to produce, given the

effects that he or she or it intends the utterance to exert. Likewise, in comprehension, the “meaning” of an utterance is never fully contained within it (Ramscar and Baayen, 2013; Edelman, 2017b).

4.4 Recurrent (folded) networks

A useful representation of the knowledge of language (whether or not it takes the form of a set of symbolic rules or a grammar) must be more compact than the body of language that it can process; in other words, it must attain *compression* (Wolff, 1988; Grünwald, 1994; Clark, 2001; Goldsmith, 2001). Compression of training data, as stipulated by the Minimum Description Length principle (Rissanen, 1987), is also the only way of attaining some ability to generalize the learned regularities to unseen data (Solomonoff, 1964).²²

In the recurrent network-based approach to modeling language, compression is enforced, implicitly, by constraining the model to reuse the same limited-capacity representational medium in processing a potentially unlimited range of inputs. In *unstructured recurrent* models, such as liquid state machines (defined below), this constraint corresponds to the limitation on the size of the network. In *uniform structured recurrent* architectures, such as the SRN (Elman, 1990), it is imposed by the size of the “hidden” layer. I discuss the unstructured recurrent networks first.

4.4.1 Unstructured recurrent networks

On an abstract-computational level, some theoretical tools for the analysis of the behavior of recurrent networks are provided by the related concepts of chaotic itinerancy (Tsuda, 1996; Rabinovich et al., 2008; Tsuda, 2015) and structured flows on manifolds (Huys et al., 2014). While these ideas are beginning to find empirical support (e.g., in the study of hippocampal place cell sequence replay; Pfeiffer and Foster, 2015) and application in the modeling of sequential behaviors (such as handwriting; Huys et al., 2014), full-fledged models that employ unstructured recurrent networks tend to be more limited in their scope.

Consider, for instance, architectures known variously as Liquid State Networks (Maass, 2007) or Echo State Networks (Tong et al., 2007). In these networks, a set of randomly interconnected spiking neurons that jointly define the machine’s “liquid” state gives rise to a potentially very complex dynamics that can be modulated by an external stimulus (such as an image that needs to be classified) and read out by a neuron that computes a function (which may be linear) of the activities of the state neurons. Many different readout neurons can be trained to compute a variety of outputs from the same internal dynamics of the Liquid State Machine (LSM). This architecture can provably generate a broad family of input-output mappings (Maass and Markram, 2004).²³ It also proved effective in modeling the behavior of anatomically realistic circuits of spiking neurons, involved in supporting many different cognitive tasks in a variety of species (Buonomano and Maass, 2009; Harvey et al., 2012).

Given the computational power of LSM guaranteed by the theorems in (Maass and Markram, 2004), it is not surprising that a model consisting of an unstructured recurrent network (with noise input and a winner-take-all readout) could replicate many statistical properties of birdsong in Bengalese finches (Yamashita et al., 2011). The implications of this finding for understanding the neural basis of language are, however, unclear. In bird courtship the sheer complexity of the song is likely the entire point of the male’s performance (as an honest signal of his motivation or quality; Lachmann et al., 2001). In contrast, in language the main function is conveying/gleaning structured information that could be useful in shaping complex behavior. To

²²Interestingly, the ADIOS model of language acquisition (Solan et al., 2005), whose learning process does not involve explicit optimization of compression or of any other global performance criterion, ends up compressing the training corpus (Solan et al., 2003).

²³For a discussion of the limitations of the input-output or stimulus-response (S/R) conception of cognition, see (Edelman, 2015).

serve as a model of language, LSM or any other architecture must be shown capable of enforcing nontrivial structural regularities — “dependencies among dependencies” (Harris, 1991, p.52; recall section 3.3) — that transcend basic statistics such as symbol frequencies.²⁴

4.4.2 Uniform structured recurrent networks

The Simple Recurrent Network (SRN) architecture, introduced by Elman (1990), differs from unstructured networks such as LSM in that its units are separated into several functional categories arranged by layer: input, output, and hidden, as in the popular three-layer perceptron or “backpropagation network” (Rumelhart et al., 1986) from which the SRN was derived. SRN models have been used to replicate, one at a time, a variety of findings in psycholinguistics, generating an extensive literature (reviewed, for instance, in Elman et al., 1996; Christiansen and Chater, 2001). While most such models typically focus each on a limited set of phenomena, some efforts (e.g., Rohde, 2002; Chang, 2002) stand out for their broad coverage. However, the SRN architecture scores low on neural plausibility and even the most ambitious SRN-based models have not been shown to scale up to learning from, generating, and processing unconstrained natural language and to dealing with large corpora.

A renewed motivation for the attempts to scale up uniform structured recurrent network solutions for various cognitive tasks, including language, came with the development of deep network architectures and of new algorithms capable of training such networks (Hinton, 2007). Deep learning (DL) methods have resulted in some spectacular successes, achieving human-level performance in certain tasks, such as visual object recognition (see (LeCun et al., 2015) and the references therein). DL has also been applied to sequence learning (Sutskever and Hinton, 2007) and, more specifically, to language tasks such as parsing and translation (Collobert and Weston, 2008; Socher et al., 2011; Hermann and Blunsom, 2013; Sutskever et al., 2014). While the practical appeal of DL is clear, its contribution to the understanding of the human language faculty seems limited. In particular, as discussed in detail in (Edelman, 2015), (i) because they are essentially devices for supervised learning of function approximation, DL systems perform well on tasks of the stimulus/response variety (mapping inputs to outputs); (ii) their training requires massive amounts of labeled data; and (iii) the similarities between the DL architectures and brain circuits are partial and superficial.

In their critique of recurrent networks as models of serial order in behavior, Cooper and Shallice (2006) suggest that “explicit hierarchically organized and causally efficacious schema and goal representations are required to provide an adequate account of the flexibility of sequential behavior.” From the computational standpoint, framing the problem at hand as function approximation imposes few constraints on the architecture of a model that implements the solution, which is why uniform networks can do well in carrying it out and also why they falter on tasks that involve endogenous (agent-generated) goals and require flexible, open-ended reasoning about structured situations (Edelman, 2015).

Deep network models are not equipped to deal with this challenge. Consider, for instance, a model that integrates DL with reinforcement learning and that was recently shown capable of human-level performance in a variety of computer games (Mnih et al., 2015). While this system did better than humans in dozens of simpler games, its creators note that performance remained far below that of human players on games that involve complex, hierarchically structured environments.

For DL models, even just holding a goal (let alone a hierarchy of goals) in mind while working on it is problematic. As LeCun et al. (2015) observe with regard to recursive neural networks (RNNs), of which SRN is an instance,

²⁴Grammar-like structure in birdsong (Katahira et al., 2011; Menyhart et al., 2015) may evolve because the learner’s ability to pick up patterns from his tutor can signal the learner’s quality.

RNNs, once unfolded in time [...], can be seen as very deep feedforward networks in which all the layers share the same weights. Although their main purpose is to learn long-term dependencies, theoretical and empirical evidence shows that it is difficult to learn to store information for very long.

4.4.3 Augmented and other non-uniform recurrent networks

DL's problem with long-range dependencies stems from its reuse of the same units over and over at each processing cycle, as output is fed back through through the recurrent connections to become part of the next input. Two solutions to this problem suggest themselves. First, one may give up on the recurrent architecture in favor of “unfolded” circuits, in which reuse-induced interference among representations is easier to control (more about this in section 4.5). Second, one may counter the interference by breaking with the uniformity of the network design, for instance by augmenting it with a special module that implements persistent working memory, as in the LSTM (long short term memory) network architecture (Hochreiter and Schmidhuber, 1997b; Schmidhuber, 2015).

Continuing down that path leads to computationally extremely powerful models: memory networks and neural Turing Machines (see (LeCun et al., 2015) for references and discussion). Of course, the possibility of a neural implementation of the Turing Machine (TM) has been investigated already by McCulloch and Pitts (1943). This approach to modeling, which includes “connectionist” production systems of Lebiere and Anderson (1993) and the “human Turing machine” of Zylberberg et al. (2011), is similar in spirit to the abstract symbolic models mentioned in section 4.3. Like those models, it has not been applied specifically to language. In any case, in modeling language, the computational power of a TM is more of a handicap than an advantage: as pointed out by Chomsky (2004, p.92) and others, to capture, crucially, both no less and no more than the regularities of natural language, the power of the model must be constrained in sophisticated ways, for which an abstract, closed-form characterization is lacking, and which in fact may be unattainable in principle, given the open-ended nature of the constraints that shape language behavior, as discussed earlier.

It seems therefore, that a better design choice here is not to strive for computational power as such (let alone for TM-level power), but rather to augment the recurrent architecture as needed for modeling specific aspects of language acquisition and use, preferably in a brain-inspired manner. This approach leads to models such as the Temporal Recurrent Network of (Dominey, 2005; Dominey and Hoen, 2006), which is augmented with working memory for open-class words.²⁵ To date, such models (e.g., Dominey, 2005; Takac et al., 2012) aimed for neural plausibility at the expense of coverage and did not attempt to deal with realistic language or corpus data.

4.5 Extended (unfolded) networks

The extended-network model architectures discussed in this section represent a convergence of ideas from several sources. As mentioned earlier, an extended network can be obtained by unfolding a recurrent network, so that the successive time steps in its operation are represented by serially connected elements in the circuit (as in LeCun et al., 2015, fig.5). In neural computation, the use of space to represent time is quite common (see appendix A). It is found, for instance, in the circuit that implements binaural sound localization in birds and mammals, where axonal propagation delay serves to compensate for interaural time

²⁵Devolving into an abstract binding-based “model” (Dominey, 2005, fig.4). Sequence learning in a model of insect mushroom bodies (Arena et al., 2015); cf. vertebrate hippocampus.

differences, allowing the use of a bank of coincidence detector neurons to signal direction to the sound source (Joris et al., 1998; Shamma, 2001).

The extended-circuit idea can be seen as motivated by three mutually complementary considerations, which together neatly span the levels of the Marr-Poggio hierarchy. On the *computational* level, in machine learning, obtaining an extended circuit by unfolding a recurrent one aims to alleviate the difficulty of building and preserving sequential and hierarchical structure in a distributed recurrent architecture. On the *algorithmic* level, in psycholinguistic modeling, constructing an extended circuit incrementally from the bottom up constitutes a response to the posited constraint on elementary learning operations, namely, alignment, comparison, and chunking. On the *implementational* level, in birdsong, choosing an extended circuit as a model is motivated by basic neuroanatomical and neurophysiological constraints, such as the space-for-time representational trade-off.

The unifying characteristic of models driven by these considerations is that their formal structure — a graph (Chartrand, 1992) or a generalization thereof (Harel, 1988; Edelman, 2015) — has a close and explicit correspondence to the repertoire of sequential behaviors that they account for: a corpus of linguistic utterances or bird songs. Specifically, the behavioral sequences that such models generate and accept (parse) correspond to the serial ordering of graph nodes created by a walk through it. In other words, a sequence is generated (or accepted) as the “token” of activity is passed down the network’s links. Depending on the actual dynamics of this process (e.g., on the interplay of excitation and inhibition), there may be one focus of activation, in which case the network’s instantaneous state is represented by the identity of the one active node at a time; alternatively, activation may be distributed, in which case the state is composite — a possibility that is outside the scope of the present discussion (Kolodny and Edelman, 2015).

Note that, in contrast, in recurrent models (section 4.4) this correspondence is implicit, while symbolic models (section 4.3) do not even possess structure in the relevant sense of the word. Further, there is a large and diverse class of network models of grammar that are in this sense symbolic (e.g., Lamb, 1998; Mel’čuk, 2003; Hudson, 2007). A radical example is Hudson’s (2007) Word Grammar, according to which “language is nothing but a network – there are no rules, principles or parameters to complement the network” and, further, “syntax consists of nothing but pair-wise links among words.” Crucially, such network approaches posit multiple types of links. For instance, figure 2.63 on p.259 of (Hudson, 2007) illustrates the hierarchical relationships among no fewer than 11 subtypes of dependency between nodes. In a biological neural network, in comparison, all axons are functionally the same; synapses come in very few varieties, each of which is defined in terms of its action on the postsynaptic cell.²⁶ Whether or not all dependency grammars with typed links can be implemented in neural circuits under these constraints — perhaps by shifting the burden of type maintenance from links to nodes, through the use of distributed instead of localist representations — is an open question. In the spirit of the neural plausibility constraints of appendix A, in the following discussion I consider only those network models that do not use typed links. I refer to these as transition networks.

4.5.1 Transition network models

Transition networks have been used in several models of birdsong generation, where the sequential structure of song is represented explicitly by an extended circuit, in which synaptic transitions correspond to transitions between successive notes or “syllables” (e.g., Fiete and Seung, 2009; Jin, 2009; Katahira et al., 2011). Such models consist of a serially connected “backbone” that encodes the syllable sequence and a readout

²⁶The contrast between networks with abstract typed links and real neural networks brings to mind Johannes Müller’s “law of specific nerve energies” (1835).

circuit that maps each presently active representation to the corresponding motor output. If the song of the species being modeled includes more than one main motif, the single linear backbone becomes a branching and reconnecting graph (Jin, 2009), which corresponds, in formal computational terms, to the finite-state automaton that generates the (regular) language (Cane, 1978) of the song in question (Yamashita et al., 2011; Katahira et al., 2011; Menyhart et al., 2015).

The transition network approach has deep roots in psychology. In particular, it resembles Clark Hull’s diagrams of association networks offered as an account of sequential behavior in maze learning and traversal by rats (Hull, 1934, fig.10). A basic Hullian transition network is simply a chain of stimulus-response associations, in which each action on part of the animal is a response to an immediately preceding “stimulus” — an architecture criticized by Lashley (1951) as incapable of modeling behavioral sequences that contain optional elements or, more generally, that possess hierarchical structure. Lashley’s criticism was subsequently addressed by Wickelgren (1969), whose elaboration on the basic association network included a shallow hierarchy of control units (cf. appendix B.2).

Hull’s network is among the direct predecessors of the “structured connectionist networks” (Feldman and Ballard, 1982; Feldman et al., 1988), whose development ushered in the “connectionist” era in cognitive modeling.²⁷ As long as they do not use typed links and other abstract computational tools, the structured connectionist models are precisely what I call here transition networks.²⁸ Unlike the uniform “backprop” networks mentioned in section 4.4.2, structured connectionist models have all but disappeared from the cognitive modeler’s toolbox. When they are applied to language, it is in a form that is typically much more abstract than the straightforward transition network architecture (Chang et al., 2004; Chang and Mok, 2006; Chang, 2008).

Exceptions to this pattern are some data-driven models of language acquisition that build, from the bottom up, graph-structured grammars that act as transition networks (Solan et al., 2005; Kolodny et al., 2015). In these networks (which have been applied also to the modeling of birdsong and of other sequential behaviors and their evolution; Kolodny et al., 2014; Menyhart et al., 2015; Kolodny et al., 2015), each newly encountered utterance is adduced to the growing graph after some abstraction, which includes the chunking and reuse of commonly occurring subsequences (collocations) and the detection and coding of contextually similar and hence substitutable elements. Edelman (2015, esp. section 5) discusses at some length the reasons why the structured connectionist approach to modeling language may be particularly promising (see also section 4.6 below).

4.5.2 Path switching in transition networks

In transition network models of this type, each utterance that can be generated or parsed corresponds to a unique path through the graph. With lifelong learning enabled, the possible behaviors of such a model can become probabilistically weighted, in a manner that is driven by experience (as in the path selection model of Fernando et al., 2011; cf. Doursat and Bienenstock, 2006). It may seem that because of this one-to-one correspondence between action sequences and paths, the model’s behavior must be “hard-wired” — indeed, interrupting a sequence, switching paths dynamically, and even choosing the desired path to begin with (as opposed to the path that is the most probable according to prior experience²⁹) would all seem impossible.

The resolution of this fundamental issue lies in allowing the traversal of each link in the network to be

²⁷Walker’s (1992) survey of the early history of connectionism lists even earlier sources than Hull, such as the work of Spencer and James.

²⁸Connectionist PCFGs (e.g., van Schijndel et al., 2012) do use typed links.

²⁹Standard probabilistic language models (Goodman, 2001) are geared to rank utterances by their probability.

determined by a conditioning variable. This type of control corresponds to the abstract computational notion of a guarded command (see the discussion and references in Kolodny and Edelman, 2015). On the level of mechanism, it requires *gating* — either on/off switching, as in temporarily enabling or disabling a link, or routing, as in steering a signal from a source node to one of several possible destination nodes.

Perhaps unexpectedly for those who think of neural networks as hard-wired, it is rather easy to implement the gating functionality in a neural circuit. A locus classicus for this idea is the celebrated paper by McCulloch and Pitts (1943), in which they showed that gated networks of highly simplified formal “neurons” have the computational power of a Turing Machine. More recent examples include a probabilistic routing mechanism that uses distributed representations (Jin, 2009, fig.3), a neural router concept controlled by a steering signal (Miller and Cohen, 2001; Botvinick et al., 2009), a detailed interpretation of cortical circuitry implementing graphical models (Litvak and Ullman, 2009), and a (still very corticocentric) proposal of how the brain, operating in what a computer engineer would call a virtual machine mode (Sloman and Chrisley, 2003; Edelman, 2008a), may implement a TM (Zylberberg et al., 2010).

4.5.3 On the computational power of transition networks

As already mentioned in section 4.4.3, in modeling language Turing equivalence is more of a problem than a solution, because of the difficulty of constraining the TM’s power in just the right ways. An augmented transition network (ATN) architecture with the power of a TM has been used in the first large-scale models of natural language processing (Woods et al., 1972; Wanner and Maratsos, 1978). In discussing the implications of ATN’s excess power, Fodor and Frazier (1980) offered an interesting comment:

[...] the class of ATNs is more or less co-extensive with the class of conceivable parsing mechanisms and thus that ATN theory does not substantially contribute to the search for the human sentence parsing mechanism (HSPM)? We believe that ATN theory can be defended against this charge of emptiness, for there is an interpretation of it under which it does make an empirical claim (though we don’t know whether this is a claim to which the proponents of the theory take themselves to be committed). The idea is that the behavior of the HSPM can be modelled by an ATN because the relevant part of the human brain has exactly the structure and capacities (computational resources, storage capacity, control structures, etc.) of an ATN.

Because some of the computational tricks to which ATNs resort to achieve their power (notably, the stack-like push/pop operations) are highly implausible in the context of a biological brain (unless implemented slowly and effortfully on the level of a virtual machine; see appendix A), the above interpretation seems as unlikely to me as it did to Fodor and Frazier (1980).

4.6 Hybrid networks

Different types of models discussed so far may be particularly good at addressing different aspects of the control problem that underlies language behavior. For instance, the simple two-layer competitive queuing (CQ) network (Bohland et al., 2010, fig.1), mentioned in section 4.2, is effective in integrating multiple constraints that join forces in generating a sequence of actions from a parallel representation, while recurrent (folded) networks offer representational savings, and transition (unfolded) networks are easier to learn. These considerations motivate a hybrid, division-of-labor approach to modeling language. Two such hybrid approaches are described in the remainder of this section.

The first one has been developed as a model of machine translation (Edelman and Solan, 2009). This model uses the ADIOS algorithm (Solan et al., 2005) to learn structured transition networks (recall section 4.5.1) for the source and target languages, then uses the activation pattern induced by the input utterance in the source network to selectively activate elements in the target network. The resulting spread of activation in the target network is subject to the learned statistical-structural patterns of usage (“grammar”) in the target language. This process generates a probability-ranked list of competing output sentences, which is then reranked so as to take into account any additional thematic or contextual constraints. The ADIOS-MT model can therefore be seen to rely both on structured transition networks and on competitive constraint satisfaction, which contribute to its functioning in distinct stages of processing.

The second hybrid approach is found in the GODIVA model of speech production (Bohland et al., 2010), which integrates CQ with certain aspects of both (unfolded) transition networks and (folded) recurrent networks. This model is designed to produce hierarchically and contextually constrained sequences of categorical elements of speech (phonemes), reaching up to the phonology of isolated words, but stopping short of sentence construction. This exclusion is significant: as noted in section 3.1, the graph of the possible paths over phonemes is orders of magnitude smaller than the graph over words; the computational problem of controlling navigation over the word graph is much more challenging than the one addressed by GODIVA.

Even so, the GODIVA model is not learned; instead, the network weights that determine its dynamics are “hand-wired” — manually set to values that guarantee the desired outcomes when the differential equations describing the network are integrated (Bohland et al., 2010, p.1516). With regard to the model’s architecture, Bohland et al. (2010, p.1522) note that it “combines elements of both CQ and positional models,” as well as “‘serial chain’ representations.” Importantly, they offer an explicit mapping of the functional components of the GODIVA model onto certain brain regions and circuits implicated in language (more about this in section 5).

Taking into account the lessons from the entire preceding survey of various classes of models, it appears that our best bet is a hybrid approach that would combine elements of existing approaches (such as ADIOS-MT and GODIVA) and extend them so as to meet the functional-computational demands of language. The proposed approach would:

- (A) implement the graph-like space of elementary language actions — by an ensemble of transition networks, TN;
- (B) implement the dynamic CQ sequencing of these actions — by a recurrent network RN, controlling the transitions of TN;
- (C) implement the various constraints and dependencies — by steering the dynamics (activity-space trajectory) of RN, through the gating of control units by contingent inputs;
- (D) ground language behavior in general-purpose understanding of how the world works — by making the control of RN susceptible to influence from hierarchically structured representations of the world.

A sketch of the proposed architecture (which is as yet very far from being a model), showing only the main functional components, appears in Figure 3, with the above items (A)-(D) marked. In the next section, I discuss some of the neurobiological considerations that should be brought to bear on the design of a proper — computationally and structurally explicit and neurally plausible — model based on these ideas.

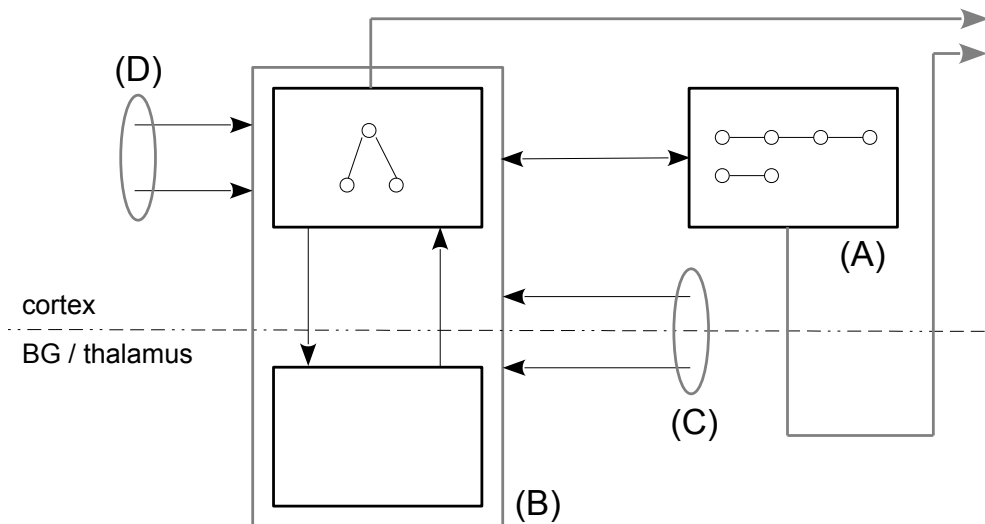


Figure 3: A bird’s eye view of the functional architecture of the proposed hybrid approach to modeling language behavior, combining ideas from (Edelman and Solan, 2009, fig.1) and (Bohland et al., 2010, fig.2). The two arrows exiting on the right are motor outputs. BG, basal ganglia. (A), ensemble of transition networks. (B), dynamic CQ control loop. (C), contingent real-time influences on control. (D), general-cognitive influences on control. For a rough mapping between these functional components and brain areas and circuits, see Table 3.

5 On the possible brain basis of language

Where in the brain should we look for language, and in particular for the neural counterparts to the computational building blocks identified and discussed above? Given that virtually every brain area for which functional imaging data are available has been shown to contribute to every task ever considered in those studies (Anderson, 2010, pp.257-258), one should expect to find language, along with other complex behaviors, to be computed and controlled cooperatively by a variety of circuits. The real questions with regard to the brain basis of language have therefore to do with the area- and circuit-level commonalities between different complex sequential behaviors and with the roles that the relevant circuits play in computing various aspects of language.

In considering the findings that can be brought to bear on these questions, one must be cognizant of several potential pitfalls. First, most of the studies to date had subjects perform quite unnatural tasks, including, notably, sentence acceptability judgments, which have very little, if anything, to do with real language use. The few exceptions, to be discussed later, include (Silbert et al., 2014) and (Brennan et al., 2016), whose subjects listened to narrated texts — still not natural dialogue of the kind illustrated in Table 2, but much better than, say, the task of reporting the number of nouns changed between two consecutive sentences in a trial (Santi and Grodzinsky, 2010).

The second issue, which has affected the gathering and interpretation of most of the available functional imaging data, is extreme corticocentrism — a bias that is not specific to language sciences (Parvizi, 2009), and that is far from having been dispelled by critique such as that of (Anderson, 2010). As a theoretical stance that dominates not only the reports of particular studies, but also integrative reviews and program-

matic papers (e.g., Poeppel et al., 2012; Bornkessel-Schlesewsky and Schlesewsky, 2013; Poeppel, 2014), corticocentrism remains a serious obstacle to the development of a broad and naturalistic understanding of the brain basis of language.

The third, methodological issue has to do with level of detail of models of language in the brain that are on the offer. Many such models, including the best-known and most influential ones, make no attempt to engage with either the circuitry or the computations underlying language. Rather, they are mere verbal descriptions that map the classical conceptual components of language such as “syntax” and “semantics” onto mostly cortical areas. Examples include the daring for its time “declarative-procedural model” of (Ullman, 2001, 2004; Walenski and Ullman, 2005; Ullman, 2006), as well as more recent attempts at theoretical synthesis such as (Hagoort, 2014; Fedorenko and Thompson Schill, 2014), which deal in networks of areas, not circuits or computations.

That said, it seems that the behavioral and computational take on language sketched earlier in this paper is hardly detailed enough to suggest a more explicit model of how the brain implements language. Rather than pretending that building such a model is already feasible, in the remainder of this section I offer some remarks on the classical view of the brain basis of language, followed by a series of hypotheses that link the various components of the alternative view of language — agency and social functions, embodiment, dynamic control, constrained navigation, multiple types of dependencies and constraints, multimodality and concurrency — to brain circuits that support them.

5.1 The classical view of language in the brain

There are two aspects to what I call the classical view of how language is implemented in the brain: it is, as already noted, corticocentric in its neuroscientific outlook; and it is syntactocentric³⁰ in its linguistic orientation. In mapping language onto the brain, theories entertained by the corticocentric mainstream differ little from the Broca/Wernicke caricature familiar to many of us from psychology textbooks. For instance, van der Lely and Pinker (2014) include the caudate nucleus of the basal ganglia as the only non-cortical region in their synthesis. In contrast, the imaging study of Silbert et al. (2014), whose subjects listened to an actor-narrated story, reports extensive spread of language production and comprehension over various brain areas.^[H]

Theories that focus exclusively on the cortex tend also to be beholden to textbook linguistic notions of syntactic well-formedness, the autonomy of the syntax, and communication by passing around meanings (see section 1.1 and Edelman, 2017b for a critique). For instance, the modeling effort of (Bornkessel-Schlesewsky and Schlesewsky, 2013) aims at uncovering “the functional neuroanatomy of the form-to-meaning mapping” culminating in an interface with the “cognitive control” module in the frontal lobe.

Among these theoretical commitments, the one to generative grammar and the reality of syntax (as, for example, in the unquestioning adoption by Friederici and Singer (2015) of the Minimalist concept of Merge) is probably the most counterproductive. The continued prominence of this stance is striking, given the extreme scarcity of independent empirical evidence for the behavioral and neural reality of syntax. While claims of such evidence are made regularly, the studies they are based on tend to be methodologically lacking and the findings weak at best.

In one typical example, the study of Ben-Shachar et al. (2003), which used planned contrasts in an imaging experiment to localize syntax in the cortex, relied on dubious assumptions regarding the argument structure of the verbs it employed.^[I] In another example, the study of Santi and Grodzinsky (2010) was constructed around an extremely artificial task: the subjects had to tell whether one or two nouns changed

³⁰Jackendoff’s (2011) term.

<i>section(s)</i> ← <i>function/computation</i>	<i>brain basis</i> →	<i>refs. in section(s)</i>	
2.2	serial order and concurrency	basal ganglia	5.2.2, 5.2.7
	complexity and hierarchy	prefrontal cortex; medial frontal cortex (pre-SMA and SMA)	5.2.1, 5.2.6
	embodiment and situatedness	hippocampus, insula, amygdala	5.2.5, 5.2.4
	agency and contingency	orbitofrontal cortex; thalamocortical loops	5.2.1, 5.2.7
2.3, 2.4	dynamic control (CQ)	prefrontal cortex; pre-SMA and SMA	5.2.1, 5.2.6
2.5, 2.6	constrained navigation of action space	BG / thalamus / frontal cortex loops	5.2.2
3.2	acquisition and change	hippocampus; pre-SMA and SMA	5.2.5, 5.2.6
3.3	dependencies	prefrontal cortex; hippocampus	5.2.1, 5.2.5
2++	THE USES OF LANGUAGE	THE ENTIRE BRAIN	5++

Table 3: Language in the brain: functional aspects and brain areas and circuits (see text for discussion). SMA, supplementary motor area. BG, basal ganglia. CQ, competitive queuing.

from the first to the second sentence of every trial. In addition, as the “syntactically identical” yet distinct sentences necessarily differed in some of the words, Santi and Grodzinsky (2010, p.1292) admit that it was not possible to attribute their findings exclusively to “syntax” (as distinct from semantics).

A more sound approach has been taken by Brennan et al. (2016), who used multiple regression (instead of planned contrasts) to compare the effects of syntactic complexity on the fMRI signal in various cortical areas, in subjects who listened to a story being read. While their analysis indicated that in some of the areas in question hierarchical estimates of syntactic complexity did make a unique contribution to the variance, the size of the effect was small: smaller, indeed, than that of bigram- and trigram-based measures of complexity. As I argued elsewhere (Edelman, 2017b), given how extraordinary the claims that such studies purport to substantiate are,³¹ this is not the kind of extraordinary evidence that is called for.

5.2 Bringing the rest of the brain and behavior into the picture

To distance the present approach to language from the corticocentrism and syntacticocentrism of the mainstream neurolinguistics, Table 3 lists, for the main components of the framework for language proposed in this paper, some of the candidate brain structures and circuits (subcortical and cortical) that may support them, according to published findings. Sections 5.2.1 through 5.2.7 contain the same entries, sorted by brain area/circuit, each accompanied by a few comments and select references. The list is both very partial and very terse; a proper review of the relevant neuroanatomy and neurophysiology is outside the scope of this paper and will not be attempted here.

5.2.1 Prefrontal cortex

Broca’s area in the prefrontal cortex needs no introduction as a language hub; I mention it here first, both to acknowledge this fact and to point out that modern reviews hold it to be but one hub in a much larger

³¹For example, the formalist framework used by the study in question posits extreme proliferation of nodes without lexical labels in the syntactic trees, which typically outnumber lexical nodes by a factor of 4 or 5 to 1.

“language network” that includes, notably, the basal ganglia (BG; Ullman, 2006), as well as many other areas (Hagoort, 2014).³² Among its internal subdivisions, Brodmann areas 44 and 45 are particularly closely associated with enforcing dependencies (“syntax”; Hagoort and Indefrey, 2014).

Another prefrontal area that I single out here is the dorsolateral prefrontal cortex (DLPFC), which participates in action selection — what Haggard and Chambon (2012) call prospective agency.³³ Endogenous selection of action, preceded by prospection and accompanied by a sense of agency is, of course, a critical functional component in any goal-oriented behavior, including language.

5.2.2 Basal ganglia

The basal ganglia, comprising the striatum, globus pallidus, substantia nigra, and subthalamic nucleus, serve as a hub in the set of cortical-basal-thalamic-cortical loops that play an essential role in a number of behavioral control tasks. Anatomically, the cortico-BG loops belong to five major types (Alexander et al., 1986, fig.3). A further subdivision into many parallel “split circuits” (Joel and Weiner, 1994, fig.5) or “stripes,” which are a prominent functional-anatomical feature of the prefrontal cortex, has been interpreted as supporting hierarchically structured tasks (O’Reilly and Frank, 2006; Badre, 2008; Botvinick, 2008).^[J]

The tasks supported by the cortico-BG circuit include action *sequencing* (Aldridge and Berridge, 1998; Nakahara et al., 2001; Aldridge and Berridge, 2003), habit formation and habitual sequencing (Graybiel, 2008), sequence parsing and concatenation (Jin et al., 2014), start-stop control (Jin and Costa, 2010), and, importantly, the *selection* of an action to be executed, while inhibiting its alternatives³⁴ (Bullock et al., 2009) — all under the influence of reward processing (Humphries and Prescott, 2010; Redgrave et al., 2011), as in reinforcement learning (Niv, 2009). It is widely acknowledged that all these tasks arise also in language; for instance, Longworth et al. (2005) describe the role of the striatum in language as “late integrational processes requiring inhibition of competing alternatives.”

The closed-loop anatomy of the cortico-BG circuit suggests that its dynamics may be very much like that of the recurrent (folded) network models discussed in section 4.4. Indeed, Redgrave et al. (2011) refer to this circuit as “re-entrant loop.” It is important to remember, however, that the intricate anatomy of this circuit (Humphries and Prescott, 2010; Haber and Behrens, 2014) necessitates treating it as a highly structured, rather than uniform, network (using the terminology of section 4). The GODIVA model of (Bohland et al., 2010), discussed earlier, explicitly acknowledges this architectural constraint, as does the hybrid approach illustrated in Figure 3.

5.2.3 Cerebellum

Although the cerebellum is not often included in discussions of the brain basis of language, its developmental and acute-lesion disorders interfere not just with motor fluency, but also with a wide range executive functions, such as attentional and emotional control. Linguistic difficulties that ensue include “anomia, agrammatic speech, and abnormal syntactic structure, with abnormal prosody” (Koziol et al., 2014, p.156).

³²In the interests of brevity, the entire discussion here is heavily tilted toward production; Wernicke’s area and other temporal and parietal circuits (Hagoort, 2014) are left out.

³³It is worth repeating that when I put the DLPFC, or any other brain area A, on the spot in the context of this or that function F, it does *not* imply that F is exclusively implemented by A, or that A is dedicated exclusively to implementing F. In the case of DLPFC and agency, an entire network of areas must be intact for the functional association to hold (Frith and Dolan, 1996).

³⁴If only a single action at a time can be chosen, a problem arises in connection with concurrent (e.g., multimodal) aspects of behavior: it would seem that these must be bundled together as units to allow themselves to be switched together as one.

Because language behavior is multimodal and concurrent, another aspect of cerebellar function may be required for its smooth execution: “synchrony, or fluidly³⁵ coordinating sequences of thought and action,” which occurs as a result of “very rapid on-line processing and feedback between the cerebral cortex and the cerebellum, mediated through the thalamus [...]” (Bullock, 2004, p.160).

5.2.4 Amygdala and insula

The insular cortex and the amygdala complex, two heavily mutually interconnected structures that are typically discussed in the context of interoception and social and emotional functions, are another example of brain regions that only seem unlikely to be involved in language if one clings to the syntacticocentric dogma. A recent meta-analytical survey concluded that “the insula represents a core area in language processing” (Ardila et al., 2014). Interaction between the insula and the amygdala contributes to the processing of prosody (Leitman et al., 2017) and ironical language (Akimoto et al., 2014), which are important in social contexts. Because social cues are also central to normal language acquisition, it is not surprising that an association exists between the size of the amygdala in infancy and language abilities during preschool, as revealed by a longitudinal developmental study (Ortiz-Mantilla et al., 2010). The roles of the amygdala and the insula in social language dynamics have been directly corroborated by inter-subject synchronization analysis of fMRI phase signals (Nummenmaa et al., 2014).

5.2.5 Hippocampus

The classical notions of the hippocampal involvement in memory and navigation, as well as the more recent insights into its role in sequence learning and processing (Levy, 1996; Fortin et al., 2002; Schendan et al., 2003; Jensen and Lisman, 2005; Pastalkova et al., 2008; Pfeiffer and Foster, 2013; Albouy et al., 2013), all suggest that it should be central to the computational approach to language developed in the present paper. Even a cursory survey of the literature shows that this is indeed the case. Beginning in infancy, hippocampal volume predicts later fluency (Can et al., 2013); its lesions in early childhood derail acquisition (DeLong and Heinz, 1997; Weber et al., 2006). Because low socio-economic status (SES) during development negatively affects hippocampal anatomy and physiology (Hackman et al., 2010), hippocampal function likely mediates in part the influence of SES on linguistic ability (e.g., Hart and Risley, 1995; Huttenlocher et al., 2010). In adults, hippocampal activity distinguishes good from poor word learners (Breitenstein et al., 2005). More generally, it supports implicit learning of motor sequences (Gheysen et al., 2010), multimodal binding and memory (Kurczek et al., 2013), and flexible use and processing of language (Duff and Brown-Schmidt, 2012) and rule learning (Seger and Cincotta, 2006).

According to the “behavior as navigation” metaphor (section 2.4) and the classical notion of the hippocampus as a cognitive map (O’Keefe and Dostrovsky, 1971), the representations supporting all these tasks should be map-like. Indeed, a case has been made for this map to encode the topology (connectivity) rather than the geometry of the represented space (Dabaghian et al., 2014), that is, for it being a graph. The navigation of this graph can be flexible in that it may involve shortcuts (Tolman, 1948), as well as combining and recombining snippets of trajectories (Davidson et al., 2009; Gupta et al., 2010).

The functional interdependence of navigation and language is hinted at by the interference between the two: subjects in the study of (Meilinger et al., 2008) performed worse at wayfinding when required also to shadow a verbal stimulus. A more direct source of evidence for such interdependence comes from an electrophysiological study that documented the existence of hippocampal theta oscillations (known to reflect

³⁵Cerebellar volume in infancy predicts language fluency (Can et al., 2013).

the activity of place cells and related mechanisms) during natural language processing and their sensitivity to specifically linguistic information (Piai et al., 2016; Covington and Duff, 2016).

To be of use in language, the space represented and navigated must be abstract, as in the space (graph) of phoneme or word sequences. It turns out that even rats are capable of navigating abstract spaces.³⁶ Following an idea first proposed by Levy (1996), Agster et al. (2002) trained rats to dynamically switch between sequences of olfactory choices, exhibiting a capacity not just for following, and choosing between, abstract sequences, but also for establishing and later using transient memories of choices — the kind of long-distance dependency discussed in section 3.3.3.^[K] An early statement of the relationship between hippocampus-supported navigation³⁷ and language is found in (Fried et al., 1997, p.760): “conjunctions of attributes were represented in the activity of individual cells in the medial temporal lobe [of which the hippocampus is part]. These cells may be considered ‘place cells’ in an abstract space defined by stimulus features.”

The dynamics of the circuits centered on the hippocampus may contribute to the (agentic) goal-directed and sequence-generating aspects of wayfinding, language, and other navigation-like behaviors (Pezzulo et al., 2014; Pfeiffer and Foster, 2015). Some computationally explicit approaches to modeling these behaviors use the tools of classical symbolic computation (in the model of (Rodriguez et al., 2004), for example, the hippocampus acts as a stack for sequence learning/replay), while others resort to the mathematical machinery of dynamical system theory (Huys et al., 2014).

5.2.6 Medial frontal cortex (pre-SMA and SMA)

An early indication that the part of the medial frontal cortex originally designated as the supplementary motor area (now seen as consisting of SMA and pre-SMA) is involved in sequential behavior came from Tanji and Shima (1994), who found there, in the monkey, cells “exclusively related to a sequence of multiple movements performed in a particular order.” In her review of the cortico-striatal circuits supporting the chunking of motor programs, Graybiel (1998, p.127) remarks, in the section on the SMA, that “In the prefrontal, supplementary motor, and premotor cortex, there is a step-by-step building-up of a code for movement sequences — the neural representation of elements of a motor plan.” She notes further that, because learning in the basal ganglia circuits is slow, chunks may be first formed quickly in the SMA, followed by gradual consolidation in the basal ganglia (Graybiel, 1998, p.131).

Considerations of evolution and comparative neuroanatomy suggest that the “high cognitive abilities” in primates, including language in humans, are correlated with the complexity and sophistication of frontal-lobe motor circuits (Mendoza and Merchant, 2014). The common mechanisms of language and action (Pulvermüller, 2005) may have to do in part with timing: “[. . .] the core timing circuit should precisely align the predictive sensory signals, the motor representation of chains of movement sequences, and the signals associated with the embodiment of the different effectors intervening in a complex behavior” (Mendoza and Merchant, 2014, p.87). Salinas (2009) describes a working computational model of a circuit for generating motor sequences, whose composition and behavior are compatible with this idea. The model employs a type of neuron that is common in pre-SMA, SMA, and the basal ganglia and is selective for timing (specifically, the rank order) of elementary action initiation.

³⁶For an explicit study of the parallels in episodic memory and sequence learning between rats and humans, see (Allen et al., 2014).

³⁷Not the hippocampus alone, of course. The circuit in question includes at least the hippocampus and the striatum (Albouy et al., 2008, 2013). Compare (Chersi and Pezzulo, 2012): a “computational model of the hippocampal – ventral striatum circuit that implements a goal-directed mechanism of choice, with the hippocampus primarily involved in the mental simulation of possible navigation paths and the ventral striatum involved in the evaluation of the associated reward expectancies.”

5.2.7 Thalamus

The notion that motor control is functionally closely related to higher cognition, including language, is strongly reinforced by the emerging picture of their anatomical relationships. First, no cortical function at all seems to be possible without the involvement of the thalamus (Sherman, 2016). Second, any two cortical areas that are interconnected directly (for instance, the successive areas in the ventral visual stream) are always also connected via the thalamus (Sherman, 2016, p.536). Third, all cortical areas (even “sensory” ones, such as the primary visual area V1) project directly to motor controls circuits (Sherman and Guillery, 2006, p.361). Fourth, forward projection axons throughout the thalamocortical system, including the inputs to the first-order thalamic nuclei, send off collaterals to the motor system, which carry efference copy information (Sherman and Guillery, 2006, ch.10; Sherman, 2016).

The anatomical and functional inseparability of sensory and motor information in the brain has profound implications for understanding how brains give rise to behavior, including language. In particular, Sherman (2016, p.538) points out that “Coordinated behavior of any reasonably complex animal without efference copies is improbable.” The influence exerted by efference copies on control is dynamic and ongoing: the brain does not wait for information about the present state of affairs in the body and the world to reach the top of the processing hierarchy.³⁸ Rather, higher-order areas must resign to having merely modulatory effects on the ongoing behavior and on whatever changes that have already been introduced to it by lower-order mechanisms. Any attempt to elucidate the specific roles of the thalamus in language (e.g., Bullock et al., 2009; Crosson, 2013; Hebb and Ojemann, 2013; Klostermann et al., 2013; Barbas et al., 2013) must take into account these basic facts.

6 Summary and prognosis

The approach to language developed in this paper can be summarized by viewing it through the lenses of each of the levels of the explanatory hierarchy of Mayr (1961), Tinbergen (1963), and Marr/Poggio (1977).

On the levels of the uses of behavior and its evolution (the latter being largely outside the scope of this paper; see Kolodny and Edelman, 2017), my working hypothesis is that language is, fundamentally, a type of sequential behavior that is intended to be interpreted by and to *influence* others (and self, insofar as it facilitates thinking).

On the computational problem level, language, as a complex behavior, gives rise to the same basic needs for *control* as the others: sequencing multiple concurrent elementary actions; managing complexity and hierarchical planning; being attuned to the internal and external contexts, notably social ones; and setting goals and pursuing them amid dynamically arising contingencies.

On the level of representations and algorithms, *space* is viewed here as the first and final frontier in behavioral control: the problems that arise from complex behaviors are all about physical or abstract topological spaces, for which the solutions can therefore be viewed as *dynamic, constrained navigation*. The most promising computational approach to modeling navigation in the space of linguistic behaviors seems to be one that integrates a dynamically controlled recurrent network with a graph-like repository of sequences of elementary actions.

On the level of neural implementation, it is obvious by now that many parts of the brain contribute to the computation of language. These include, in addition to the classically defined cortical “language” areas, several key subcortical systems and circuits, which participate also in controlling other behaviors.

³⁸One such apex, having to do with the integration of sensory and motor information is found in the hippocampus (Merker, 2004); another one, which may underlie conscious awareness, is the pulvinar nucleus of the thalamus (Merker, 2013).

Of course, the proof of this many-layered conceptual pudding will be in the eating. Because talk is cheap,³⁹ the claims that I have advanced on each of the four levels need to be translated into predictions for, and eventually explanations of, *new* behavioral and neurobiological studies — a not insignificant undertaking, which I leave to future work. The first priority, however, should be the development of a detailed and explicit computational model of dynamic control of complex sequential behavior, followed by an attempt to demonstrate its effectiveness in capturing everything that matters about language.

Acknowledgments

I am indebted to Barbara Finlay for general guidance and for invaluable comments on early drafts of this paper. Paul Cisek pointed me in the direction of competitive queuing and suggested some key references. Oren Kolodny was, as always, a source of insight and inspiration. I am grateful to Frederick Callaway, Arnon Lotem, Zach Solan, Ryutaro Uchiyama, and William Warren for conversations about matters of language, as well as to two anonymous reviewers for constructive comments.

Appendix A Some general neurocomputational constraints

I now briefly introduce some observations about computation in the brain that are intended to serve as a backdrop for the comparison among neurocomputational models of language in section 4.

A.1 Native vs virtual-machine implementation

The most important point to consider is the distinction between the computations that are *native* to the system in question and its components and those that are *virtual*. In an electrical circuit, a resistor across which a certain voltage is applied computes the current that flows through it (or, equivalently, vice versa). This computation is native to the resistor because, as per Ohm's Law, the voltage and the current are explicitly, directly, and immutably related by a physical parameter: the value of its resistance. At the same time, the resistor may be part of a general-purpose computer, which can be programmed to simulate a virtual circuit (one that obeys Ohm's law, Kirchhof's laws, etc.), or, indeed, to simulate a formal scenario that has no relationship to the physics of the real world (cf. Edelman, 2008a, pp.205,338, and the references there).

In formulating a computational model of a cognitive function, care must be taken to specify which of its computations are posited as native to its declared implementational substrate and which, if any, are virtual. On the face of it, models that depend critically on virtual computation are much more powerful. For instance, a model of sentence processing (say) that already has the formal computational power of a Turing machine (TM) can be reprogrammed to model also a host of other cognitive processes. However, such models are also very difficult to anchor in actual neural computation. To do so, one must demonstrate that the requisite computational building blocks — in the case of a TM-based model, variables that can be bound to any values, as in λ -calculus (a Turing-equivalent formalism) — can be effectively implemented by neural circuits.

³⁹A fact that any attempt to place the origins of language in the evolution of animal signaling must contend with (Lachmann et al., 2001).

A.2 What neurons and neural circuits compute natively

The possibility of implementing a TM⁴⁰ by a circuit of formal neurons has been demonstrated, famously, already by McCulloch and Pitts (1943). Their proof, however, assumed that neurons act as logic gates — a view that has been abandoned as being of little relevance to the computations performed natively by real neurons.

Here is what seems to be safe to assume that individual neurons⁴¹ and simple⁴² networks of neurons compute natively:

- *Linear algebra.* If the activities of a set of neurons are taken to represent a point in a vector space (Mumford, 1994, p.144), a neuron is seen to compute the *inner (scalar) product* between the vector of its inputs and the vector of its synaptic strength values (weights). This implies that the neuron's output (before any nonlinearity) corresponds to the *projection* of the input vector onto the weight vector.⁴³
- *Dimensionality reduction.* Projection can be used to reduce the dimensionality of a representation space (e.g., so as to facilitate learning from examples). The particularly desirable methods for dimensionality reduction aim to preserve the similarity structure of a set of inputs. This can be done by resorting to *random projections* (Edelman, 1999, Appendix B).
- *Tuning.* A common response profile in sensory and motor systems is *graded tuning*, in which a neuron responds maximally to some prototypical or “landmark” stimulus, and progressively less strongly to stimuli that are less similar to the prototype (Edelman, 1999, ch.6). A neuron can implement tuning by piping the inner product between its weights and the input through a squaring nonlinearity (Duch and Jankowski, 1999, p.169). Tuning is an extremely useful functional building block, as illustrated by the following applications:
 - *Universal function approximation.* The tuned-response profile is equivalent to subtracting the input vector from a reference one, estimating a weighted norm of the difference, and passing the result through a nonlinearity. If the latter is Gaussian, the result is a Radial Basis Function unit, which is capable of approximating any sufficiently smooth input-output mapping from a set of examples, to an arbitrary degree of precision (Hornik et al., 1989).
 - *Associative memory.* Locality-sensitive hashing (LSH) is a modification of the classical approach to content-addressable memory that combines fast lookup with preservation of similarity structure of the content space (Andoni and Indyk, 2008). A neural implementation of LSH that makes use of tuned units has been outlined by Edelman and Shahbazi (2012).
 - *Kernels.* The so-called kernel trick (Jäkel et al., 2007) and a long list of computational methods that are based on it are indispensable in machine learning (Schölkopf and Smola, 2002) and in cognitive modeling (Jäkel et al., 2009) as a means of avoiding expensive computations in high-dimensional representation spaces. An equivalent and complementary view of kernels as

⁴⁰Minus the infinite memory “tape,” which is problematic for any implementation, and which in any case is not strictly necessary for practical applications.

⁴¹Vertebrate brains, including the human brain, contain dozens of morphologically, physiologically, and functionally distinct types of neurons; here, I focus on the cortical principal cells, which constitute about 80% of cortical neurons (Harris and Mrsic-Flogel, 2013).

⁴²Unstructured or uniformly structured, in the sense of section 4.

⁴³Projection followed by a nonlinearity is the basic building block of what neurobiologists call neural “integration” (Fetsch et al., 2013).

a measure of similarity, akin to tuned units, highlights their relevance to modeling the brain (Shahbazi et al., 2016).

The neural computations listed above do not explicitly refer to time. Taking into account the temporal dynamics and sequential nature of signals and representations leads to the following additional insights into native neural computation:

- *Learning.* In general, neural computation is continually modified by temporal dynamics of experience. For instance, the Hebbian learning rule, which calls for changing a synaptic weight in proportion to the product of pre- and post-synaptic activities, depends on the precise relative timing of incoming and outgoing spikes (spike timing dependent plasticity, or STDP; Sjöström and Gerstner, 2010). This type of learning serves various functions, such as improved dimensionality reduction via projection pursuit (Cooper and Bear, 2012).
- *Bayesian learning and inference.* The predictive coding theory holds (i) that the brain relies on the Bayes formula to update its beliefs about the world by combining priors with data-derived likelihoods, and (ii) that the “bottom-up” direction in neural signaling carries information about prediction errors — those aspects of data that are unaccounted for by the “top-down” Bayesian predictions (Rao and Ballard, 1999; Bastos et al., 2012; Park and Friston, 2013).
- *Coincidence detection.* On the most basic level, the fast dynamics of postsynaptic events precipitated by an arriving spike and the small contribution of individual spikes to postsynaptic activity combine to give rise to a natural means for temporal coincidence detection. Specifically, two spikes converging onto a neuron can be much more effective if they arrive within the same short time window of a few milliseconds — an arrangement that, together with binaural sensing and with axon length serving as a natural delay line, implements sound localization (Joris et al., 1998; Peña, 2011).
- *Dynamics and readout.* Nonlinear dynamics in neural circuits (even ones that consist of only a handful of neurons) is a powerful computational mechanism, which can be put to various uses by using input to steer the dynamics and by reading out the results through dedicated circuits (Maass, 2007; Buonomano and Maass, 2009). Interestingly, the readout function (cf. Buzsáki, 2010) can itself be quite simple: a linear combination (that is, projection, or weighted sum) of the state variables suffices for the purpose.
- *Synfire chains and sequencing of states.* A natural means for implementing sequences of states, each represented by the activity of a neural ensemble, is a synfire chain — a wave of coordinated activity that propagates from one ensemble to the next (Abeles, 1982). Recent work has explored the emergence of such chains through learning, their use in neural computation, and the evidence for their existence in the brain (Ikegaya et al., 2004; Izhikevich, 2007; Hosaka et al., 2008).

Appendix B On the formal power of language models

[Note: the following material appeals to the concepts of generative grammar and syntactic well-formedness, whose relevance to natural language behavior is questionable (Edelman, 2017b). It is included here mainly to serve as background for discussing symbolic models of language.]

For symbolic formalist models of language that focus exclusively on “syntax,” the standard requirement is that the formal expressive power of the model be capable of capturing just the structural complexity of

natural language (Frank, 2004) — no more and no less. Too little power, and the model will fail to account for certain psycholinguistic phenomena; too much, and its claim to psychological relevance suffers. Two families of models that attempt to restrict the grammatical complexity to just the right extent and in just the right ways (to reproduce the range of relevant psycholinguistic findings) are briefly discussed below. These are intended not to provide an exhaustive list of available formal options, but rather to serve as examples of such.

B.1 Mildly Context Sensitive Grammars

The standard (but not the only possible) formal framework for comparing the power of linguistic formalisms is the Chomsky hierarchy. From the bottom up, the major rungs in this language hierarchy are the regular, context-free, context-sensitive, and recursively enumerable languages (each successive family properly including the preceding ones). These correspond, respectively, to regular automata, nondeterministic push-down automata, linear bounded automata, and Turing machines (for details, see any textbook on formal language theory, such as Hopcroft and Ullman, 1979).

It is a matter of wide agreement in formalist linguistics that the formal language family with just enough power to capture the structural complexity of human languages is the Mildly Context-Sensitive Grammar or MCSG (Joshi, 1985; Stabler, 2013a). This is a category that includes formalisms such as the Tree-Adjoining Grammar or TAG (Joshi and Schabes, 1997), as well as the Minimalist Grammar (MG) and Multiple Context-Free Grammar (MCFG) (Stabler, 2013b). Some of the advantages of models that belong to this family are the relative feasibility of learning (Stabler et al., 2003; Clark and Lappin, 2011) and parsing (Stabler, 2013b).

As with any other formalism, the question always arises with respect to MCSG's neurocomputational plausibility. For MCSG, this includes specifically concerns, which have not yet been properly discussed in the literature, about its complex representational primitives, such as multi-feature variables (Joshi, 2004), and processes, such as a queue in which pointers to structures are held until "checked out" (Stabler (2013b).

B.2 Contextual grammars

An alternative approach to developing a formal grammatical formalism for natural languages has its roots in the long-standing psychology-theoretic principles of sequencing by association, hierarchy, and context sensitivity, which date back to Hull and Lashley. Attempts to apply these principles to language include the work of Wickelgren (1969) on a "context-sensitive associative" theory of word pronunciation based on phone n-grams (specifically, 3-grams) to specify allophones. In a later work, Wickelgren proposed an extension of these ideas to syntax, remarking that "Context-sensitive coding is an alternative to production systems or augmented transition networks for the representation of procedural knowledge" (Wickelgren, 1979, p.63).

A convergence between this psychology-motivated approach and work in formal language theory is exemplified by the Contextual Grammar (CG) theory, initially proposed by Marcus (1969). Marcus et al. (1998, p.256) note that "The generative process in a contextual grammar is based on two dual linguistic operations, which are among the most important in both natural and artificial languages: insertion of a string in a given context and adding a context to a given string."

The CG formalism offers an adequate computational power, while cutting across the levels of the Chomsky hierarchy, and using primitives that seem more apt from the standpoint of psychological reality and implementation in the brain (Marcus, Martín-Vide, and Păun, 1998; cf. section 4.1 in the main text). This

latter feature prompted Crespi Reghizzi and Braitenberg (2003) to single out CG as a formalism that holds promise as a “brain compatible theory of syntax.” Of course, if in fact “syntax” reflects not the properties of an underlying formal generative grammar, but rather — as I have argued here — a mere description of the effects that multiple behavioral, computational, and neural constraints have on production, the notion of “the best formalism” becomes largely moot.

Notes

^[A]As an example, one may consider the behavioral needs of the *E. coli* bacterium, which can follow a chemical gradient to get to a place with a higher concentration of food, but is physically incapable of actively setting the direction of its cilia-propelled swim; this implies that it must exert a primitive kind of control by alternating swimming with random tumbling (Parkinson et al., 1983). A comparable degree of control underlies flagellar motility (Lertsethtakarn et al., 2011) and social behaviors (Strassmann et al., 2011) in various other species of bacteria.

^[B]In a sense, this is cybernetics (Wiener, 1948) redux.

^[C]The basic conventions of Bodirsky’s (2006) notation scheme are as follows: “For each partner, we annotate the movements by an alternating sequence of a symbol for the upper body, a symbol for the relative state of the legs, and a symbol for the leg movement. The symbol for the relative state is written below or above the symbol for the upper body, depending on whether the weight of the person is on the left or on the right leg. As we read from left to right, we imagine the leading person dancing forward from left to right.”

^[D]This choice of terms does not imply a rejection of search as a tool in behavior. Indeed, the game of chess itself, where the search paradigm reigns, illustrates the distinction between behavioral problems that can be given a tentative solution in advance and those whose solution must be constructed on the fly as required by the changing circumstances. Because a full search of the problem space in chess is computationally prohibitive, players must settle for an ongoing, open-ended approach — a limited-depth exploration that is undertaken anew before each move (Campitelli and Gobet, 2004).

^[E]A few linguists go as far as denying that discrete units such as phonemes or words have any psychological reality (Port and Leary, 2005; Ramscar and Port, 2016). On this account, language is fully and exclusively analog and continuous. A more balanced view, which I prefer, acknowledges the analog effects while holding that the basic sequential-discrete component of language is real — not the least because in the space of all possible acoustic signals of a given duration, any two phonemes (or lexemes) are disconnected in the sense that morphing them into one another would necessarily take you outside of the (discrete) space of categorically recognizable phonemes (or lexemes). Moreover, a language that is not discrete and combinatorial would be either severely limited in its expressivity or computationally impossible to learn (Edelman, 2008b,c).

^[F]Edelman et al. (2004) trained a model of language acquisition on phrases exemplifying tough movement (“is easy to read”; “is easy to please”; “is eager to read”; “is eager to please”; “to read is easy”; “to please is easy”) and discovered that the model learned tree-structured representations that stopped short of over-generalizing to incorrect phrases (“to read is eager” and “to please is eager”). Certain characteristics of child-directed speech may make such distinctions particularly accessible to learners: in one study, a significant proportion of variation sets (a corpus pattern known to facilitate learning; Onnis, Waterfall, and Edelman, 2008) in caregivers’ language pivoted on tough-movement adjectives (H. R. Waterfall, B. Sandbank, L. Onnis, and S. Edelman, unpublished observations). A variation set is a set of utterances that share some of the same words and that appear in close temporal proximity to each other; a shared word is a pivot if it is common to all the utterances that form a variation set.

^[G]The full quote from (Houghton and Hartley, 1996), which I reproduce here, is particularly poignant in light of the distinction between native and virtual computation, discussed in appendix A: “In artificial intelligence and computer science, analogous objects plus recursive serial processing are provided by computer programming languages. In such a context, serial order per se will not appear to be any kind of problem at all. Thus, although these devices have never been defended or tested on empirical grounds, their availability and computational power suffice to obscure the fact that cognitive science has no (neuro-) psychologically grounded theory of serial order. In neurophysiology and psychology, the problem has been largely ignored or workers have fallen back on the very position that Lashley attacked, associative chaining theory.”

^[H]Silbert et al. (2014) used *rehearsed* speech produced by professional actors (after discussing the near impossibility to have regular subjects reproduce a natural speech segment in full detail over several repetitions). Also: “the network reliably involved during speech production is tightly tied to the content of the produced speech” (no correlation across different stories told spontaneously by the same speaker).

^[I]For instance, Ben-Shachar et al. (2003) assumed that the verb “help” is “less complex, dyadic (two-argument)” than the verb “tell,” which they described as “triadic.” Of course, in reality “help” can be used with three arguments (“I helped them see”) and

“tell” — with two (“I told them”). Thus, when “help” was “included in the putatively more complex [...] sentences” it need not have necessarily worked as a “conservative” feature in their stimulus design.

^[J]The role of the cortico-BG circuit in goal-directed, dynamic behavioral control has been described by Haber and Calzavara (2009) as follows: “The development and modification of goal-directed behaviors require continual processing of complex chains of events, which is reflected in the feed-forward organization of both the striato-nigral connections and the thalamo-cortical connections. Information can thus, be channeled from limbic, to cognitive, to motor circuits, to produce decision-making processes that integrate different functional information, allowing the individual to respond appropriately to environmental cues.”

^[K]A propos the switching between sequences, Levy et al. (2005) writes that “temporal compression with a backward cascade produces a recoding that is suitable for generating predictions (forecasts) based on the hippocampal recordings themselves — so long as a decoder exists. Thus we conjecture that temporally compressed sequences, followed by their neocortical encoding, allow the neocortex to forecast without the hippocampus. This conjecture rests on the supposition that the sequence completion problem in the hippocampus becomes a pattern completion problem in neocortex. In particular, the highly compressed and overlapped encodings are suitable for rapid pattern completion by the more symmetrically connected recurrent networks of neocortex.”

References

- Abeles, M. (1982). Role of cortical neuron: integrator or coincidence detector? *Israel J. Med. Sci.* 18, 83–92.
- Agster, K. L., N. J. Fortin, and H. Eichenbaum (2002). The hippocampus and disambiguation of overlapping sequences. *The Journal of Neuroscience* 22, 5760–5768.
- Akimoto, Y., M. Sugiura, Y. Yomogida, C. M. Miyauchi, S. Miyazawa, and R. Kawashima (2014). Irony comprehension: social conceptual knowledge and emotional response. *Human Brain Mapping* 35, 1167–1178.
- Albouy, G., B. R. King, P. Maquet, and J. Doyon (2013). Hippocampus and striatum: dynamics and interaction during acquisition and sleep-related motor sequence memory consolidation. *Hippocampus* 23, 985–1004.
- Albouy, G., V. Sterpenich, E. Balteau, G. Vandewalle, M. Desseilles, T. Dang-Vu, A. Darsaud, P. Ruby, P.-H. Luppi, C. Degueldre, P. Peigneux, A. Luxen, and P. Maquet (2008). Both the hippocampus and striatum are involved in consolidation of motor sequence memory. *Neuron* 58, 261–272.
- Aldridge, J. W. and K. C. Berridge (1998). Coding of serial order by neostriatal neurons: a ‘natural action’ approach to movement sequence. *Journal of Neuroscience* 18, 2777–2787.
- Aldridge, J. W. and K. C. Berridge (2003). Basal ganglia neural coding of natural action sequences. In A. M. Graybiel, M. R. DeLong, and S. T. Kitai (Eds.), *The Basal Ganglia VI*, Volume 54 of *Advances in Behavioral Biology*. New York: Springer.
- Alexander, G. E., M. R. DeLong, and P. L. Strick (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience* 9, 357–381.
- Allen, T. A., A. M. Morris, A. T. Mattfeld, C. E. Stark, and N. J. Fortin (2014). A sequence of events model of episodic memory shows parallels in rats and humans. *Hippocampus* 24, 1178–1188.
- Anderson, D. J. and P. Perona (2014). Toward a science of computational ethology. *Neuron* 84, 18–31.
- Anderson, M. L. (2010). Neural re-use as a fundamental organizational principle of the brain. *Behavioral and Brain Sciences* 34, 245–266.

- Anderson, M. L. (2016). Précis of After Phrenology: Neural Reuse and the Interactive Brain. *Behavioral and Brain Sciences* 39, e120.
- Anderson, M. L., M. J. Richardson, and A. Chemero (2012). Eroding the boundaries of cognition: Implications of embodiment. *Topics in Cognitive Science* 4, 717–730.
- Andoni, A. and P. Indyk (2008). Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions. *Communications of the ACM* 51, 117–122.
- Ardila, A., B. Bernal, and M. Rosselli (2014). Participation of the insula in language revisited: A meta-analytic connectivity study. *Journal of Neurolinguistics* 29, 31–41.
- Arena, P., M. Calí, L. Patané, A. Portera, and R. Strauss (2015). Modelling the insect mushroom bodies: Application to sequence learning. *Neural Networks* 67, 37–53.
- Arnold, J. E., M. Fagnano, and M. K. Tanenhaus (2003). Disfluencies signal thee, um, new information. *Journal of Psycholinguistic Research* 32, 25–37.
- Arnon, I. and N. Snider (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language* 62, 67–82.
- Arsenijević, B. (2008). From spatial cognition to language. *Biolinguistics* 2, 3–23.
- Atallah, H. E., M. J. Frank, and R. C. O’Reilly (2004). Hippocampus, cortex, and basal ganglia: Insights from computational models of complementary learning systems. *Neurobiology of Learning and Memory* 82, 253–267.
- Badre, D. (2008). Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends in Cognitive Sciences* 12, 193–200.
- Bannard, C. and E. Lieven (2012). Formulaic language in L1 acquisition. *Annual Review of Applied Linguistics* 32, 3–16.
- Barbas, H., M. Á. García-Cabezas, and B. Zikopoulos (2013). Frontal-thalamic circuits associated with language. *Brain & Language* 126, 49–61.
- Barlow, H. B. (1990). Conditions for versatile learning, Helmholtz’s unconscious inference, and the task of perception. *Vision Research* 30, 1561–1571.
- Basile, P., A. Caputo, and G. Semeraro (2011). Encoding syntactic dependencies by vector permutation. In S. Padó and Y. Peirsman (Eds.), *Proceedings of the GEMS 2011 Workshop on Geometrical Models of Natural Language Semantics, EMNLP 2011*, Edinburgh, Scotland, pp. 43–51.
- Bastos, A. M., W. M. Usrey, R. A. Adams, G. R. Mangun, P. Fries, and K. J. Friston (2012). Canonical microcircuits for predictive coding. *Neuron* 76, 695–711.
- Beckner, C., N. C. Ellis, R. Blythe, J. Holland, J. Bybee, J. Ke, M. H. Christiansen, D. Larsen-Freeman, W. Croft, and T. Schoenemann (2009). Language is a complex adaptive system: position paper. *Language Learning* 59, 1–26. Suppl. 1.

- Ben-Shachar, M., T. Hendler, I. Kahn, D. Ben-Bashat, and Y. Grodzinsky (2003). The neural reality of syntactic transformations: evidence from functional magnetic resonance imaging. *Psychological Science* 14, 433–440.
- Bettelheim, B. (1976). *The uses of enchantment*. New York, NY: Knopf.
- Bilek, E., M. Ruf, A. Schäfer, C. Akdeniz, V. D. Calhoun, C. Schmahl, C. Demanuele, H. Tost, P. Kirsch, and A. Meyer-Lindenberg (2015). Information flow between interacting human brains: Identification, validation, and relationship to social expertise. *Proceedings of the National Academy of Science* 112, 5207–5212.
- Bock, J. K. (1986). Syntactic priming in language production. *Cognitive Psychology* 18, 355–387.
- Bock, W. J. and G. von Wahlert (1965). Adaptation and the form-function complex. *Evolution* 19, 269–299.
- Bod, R. (2009). From exemplar to grammar: A probabilistic analogy-based model of language learning. *Cognitive Science* 33, 752–793.
- Bodirsky, M. (2006). A notation system for tango. Unpublished ms.
- Bohland, J. W., D. Bullock, and F. H. Guenther (2010). Neural representations and mechanisms for the performance of simple speech sequences. *Journal of Cognitive Neuroscience* 22, 1504–1529.
- Borensztajn, G., W. Zuidema, and R. Bod (2009). The hierarchical prediction network: towards a neural theory of grammar acquisition. In *Proc. Annual Meeting of the Cognitive Science Society*, pp. 2974–2978.
- Bornkessel-Schlesewsky, I. and M. Schlewsky (2013). Reconciling time, space and function: A new dorsal-ventral stream model of sentence comprehension. *Brain & Language* 125, 60–76.
- Botvinick, M. M. (2008). Hierarchical models of behavior and prefrontal function. *Trends in Cognitive Sciences* 12, 201–208.
- Botvinick, M. M., Y. Niv, and A. C. Barto (2009). Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. *Cognition* 113, 262–280.
- Bowers, J. S. (2009). On the biological plausibility of grandmother cells: implications for neural network theories in psychology and neuroscience. *Psychological Review* 116, 220–251.
- Breitenstein, C., A. Jansen, M. Deppe, A.-F. Foerster, J. Sommer, T. Wolbers, and S. Knecht (2005). Hippocampus activity differentiates good from poor learners of a novel lexicon. *NeuroImage* 25, 958–968.
- Brembs, B. (2011). Towards a scientific concept of free will as a biological trait: Spontaneous actions and decision-making in invertebrates. *Proceedings of the Royal Society London (B)* 278, 930–939.
- Brennan, J. R., E. P. Stabler, S. E. Van Wagenen, W.-M. Luh, and J. Hale (2016). Abstract linguistic structure correlates with temporal activity during naturalistic comprehension. *Brain and Language* 157-158, 81–94.
- Bullock, D. (2004). Adaptive neural models of queuing and timing in fluent action. *Trends in Cognitive Sciences* 8, 426–433.

- Bullock, D., C. O. Tan, and Y. J. John (2009). Computational perspectives on forebrain microcircuits implicated in reinforcement learning, action selection, and cognitive control. *Neural Networks* 22, 757–765.
- Buonomano, D. V. and W. Maass (2009). State-dependent computations: spatiotemporal processing in cortical networks. *Nature Reviews Neuroscience* 10, 113–125.
- Buzsáki, G. (2010). Neural syntax: cell assemblies, synapsembles, and readers. *Neuron* 68, 362–385.
- Campitelli, G. and F. Gobet (2004). Adaptive expert decision making: Skilled chessplayers search more and deeper. *Journal of the International Computer Games Association* 27, 209–216.
- Cámpora, M. (2011). The providential apotheosis of his industry: Display of causal systems in Borges. *The New Centennial Review* 11, 125–141.
- Can, D. D., T. Richards, and P. K. Kuhl (2013). Early gray-matter and white-matter concentration in infancy predict later language skills: A whole brain voxel-based morphometry study. *Brain & Language* 124, 34–44.
- Candlish, S. and G. Wrisley (2014). Private language. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2014 ed.).
- Cane, V. R. (1978). On fitting low-order Markov chains to behaviour sequences. *Animal Behavior* 26, 332–338.
- Carver, C. S. and M. F. Scheier (1990). Origins and functions of positive and negative affect: a control-process view. *Psychological Review* 97, 19–35.
- Casey, B., J. E. Kersh, and J. M. Young (2004). Storytelling sagas: an effective medium for teaching early childhood mathematics. *Early Childhood Research Quarterly* 19, 167–172.
- Chang, F. (2002). Symbolically speaking: a connectionist model of sentence production. *Cognitive Science* 93, 1–43.
- Chang, N. and E. Mok (2006). A structured context model for grammar learning. In *International Joint Conference on Neural Networks*, Vancouver, BC.
- Chang, N. C. (2008). *Constructing grammar: A computational model of the emergence of early constructions*. Ph. D. thesis, University of California, Berkeley.
- Chang, N. C., J. A. Feldman, and S. Narayanan (2004). Structured connectionist models of language, cognition, and action. In *Proc. 9th Neural Computation and Psychology Workshop (NCPW9)*, Singapore, pp. 57–67.
- Chappell, J. and A. Kacelnik (2002). Tool selectivity in a non-primate, the New Caledonian crow (*Corvus moneduloides*). *Animal Cognition* 5, 71–78.
- Chartrand, G. (1992). *Applied and Algorithmic Graph Theory*. New York: McGraw-Hill.
- Chersi, F. and G. Pezzulo (2012). Using hippocampal-striatal loops for spatial navigation and goal-directed decision-making. *Cognitive Processes* 13 (Suppl 1), S125–S129.

- Chomsky, N. (2004). *The Generative Enterprise Revisited*. Berlin: Mouton de Gruyter. Discussions with Riny Huybregts, Henk van Riemsdijk, Naoki Fukui and Mihoko Zushi.
- Christiansen, M. H. and N. Chater (2001). Connectionist psycholinguistics: Capturing the empirical data. *Trends in Cognitive Sciences* 5, 82–88.
- Cisek, P. (2012). Making decisions through a distributed consensus. *Current Opinion in Neurobiology* 22, 927–936.
- Cisek, P. and J. F. Kalaska (2010). Neural mechanisms for interacting with a world full of action choices. *Annual Review of Neuroscience* 33, 269–298.
- Clark, A. (2001). *Unsupervised Language Acquisition: Theory and Practice*. Ph. D. thesis, School of Cognitive and Computing Sciences, University of Sussex.
- Clark, A. and S. Lappin (2011). *Linguistic Nativism and the Poverty of the Stimulus*. Oxford: Wiley-Blackwell.
- Clark, H. H. and E. F. Schaefer (1989). Contributing to discourse. *Cognitive Science* 13, 259–294.
- Collobert, R. and J. Weston (2008). A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th International Conference on Machine Learning*, Helsinki, Finland, pp. 1–8.
- Cooper, C., D. Orban, R. Henry, and J. Townsend (1983). Teaching and storytelling: an ethnographic study of the instructional process in the college classroom. *Instructional Science* 12, 171–190.
- Cooper, L. N. and M. F. Bear (2012). The BCM theory of synapse modification at 30: interaction of theory with experiment. *Nature Reviews Neuroscience* 13, 798–810.
- Cooper, R. P. and T. Shallice (2006). Hierarchical schemas and goals in the control of sequential behavior. *Psychological Review* 113, 887–916.
- Covington, N. V. and M. C. Duff (2016). Expanding the language network: direct contributions from the hippocampus. *Trends in Cognitive Sciences* 20, 869–870.
- Crespi Reghizzi, S. and V. Braitenberg (2003). Towards a brain compatible theory of syntax based on local testability. In C. Martín-Vide and V. Mitrana (Eds.), *Grammars and Automata for String Processing: From Mathematics and Computer Science to Biology, and Back*, pp. 17–32. CRC Press.
- Croft, W. (2010). The origins of grammaticalization in the verbalization of experience. *Linguistics* 48, 1–48.
- Crosson, B. (2013). Thalamic mechanisms in language: A reconsideration based on recent findings and concepts. *Brain & Language* 126, 73–88.
- Dabaghian, Y., V. L. Brandt, and L. M. Frank (2014). Reconceiving the hippocampal map as a topological template. *eLife* 3, e03476.
- Dąbrowska, E. (2004). *Language, mind, and brain: some psychological and neurological constraints on theories of grammar*. Georgetown University Press.

- Dale, R., R. Fusaroli, N. D. Duran, and D. Richardson (2013). The self-organization of human interaction. In B. Ross (Ed.), *Psychology of Learning and Motivation*, Volume 59, pp. 43–95. Elsevier.
- Dale, R. A. and M. J. Spivey (2006). Unraveling the dyad: Using recurrence analysis to explore patterns of syntactic coordination between children and caregivers in conversation. *Language Learning* 56, 391–430.
- Damasio, A. and G. B. Carvalho (2013). The nature of feelings: evolutionary and neurobiological origins. *Nature Reviews Neuroscience* 14, 143–152.
- Davidson, T. J., F. Kloosterman, and M. A. Wilson (2009). Hippocampal replay of extended experience. *Neuron* 63, 497–507.
- de Marneffe, M.-C., T. Dozat, N. Silveira, K. Haverinen, F. Ginter, J. Nivre, and C. D. Manning (2014). Universal Stanford Dependencies: A cross-linguistic typology. In N. Calzolari (Ed.), *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014)*, Reykjavik, Iceland.
- Dechter, R. and J. Pearl (1985). Generalized best-first search strategies and the optimality of A*. *Journal of the ACM* 32, 505–536.
- DeLong, G. R. and E. R. Heinz (1997). The clinical syndrome of early-life bilateral hippocampal sclerosis. *Annals of Neurology* 42, 11–17.
- Di Fiore, A. and S. A. Suarez (2007). Route-based travel and shared routes in sympatric spider and woolly monkeys: cognitive and evolutionary implications. *Animal Cognition* 10, 317–329.
- Dominey, P. F. (2005). From sensorimotor sequence to grammatical construction: evidence from simulation and neurophysiology. *Adaptive Behavior* 13, 347–361.
- Dominey, P. F. and M. Hoen (2006). Structure mapping and semantic integration in a construction-based neurolinguistic model of sentence processing. *Cortex* 42, 476–479.
- Doursat, R. and E. Bienenstock (2006). The self-organized growth of synfire patterns. In *10th International Conference on Cognitive and Neural Systems (ICCN 2006)*, Boston University, MA.
- Du Bois, J. W. (2014). Towards a dialogic syntax. *Cognitive Linguistics* 25, 359–410.
- Duch, W. and N. Jankowski (1999). Survey of neural transfer functions. *Neural Computing Surveys* 2, 163–213.
- Duff, M. C. and S. Brown-Schmidt (2012). The hippocampus and the flexible use and processing of language. *Frontiers in Human Neuroscience* 6, 69.
- Dunbar, R. I. M. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences* 16, 681–735.
- Edelman, S. (1999). *Representation and recognition in vision*. Cambridge, MA: MIT Press.
- Edelman, S. (2007). Bridging language with the rest of cognition: computational, algorithmic and neurobiological issues and methods. In M. Gonzalez-Marquez, I. Mittelberg, S. Coulson, and M. J. Spivey (Eds.), *Proc. of the Ithaca workshop on Empirical Methods in Cognitive Linguistics*, pp. 424–445. John Benjamins.

- Edelman, S. (2008a). *Computing the mind: how the mind really works*. New York, NY: Oxford University Press.
- Edelman, S. (2008b). On the nature of minds, or: Truth and consequences. *Journal of Experimental and Theoretical AI* 20, 181–196.
- Edelman, S. (2008c). A swan, and pike, and a crawfish walk into a bar. *Journal of Experimental and Theoretical AI* 20, 261–268.
- Edelman, S. (2012). Vision, reanimated and reimagined. *Perception* 41, 1116–1127. Special issue on Marr’s *Vision*.
- Edelman, S. (2015). The minority report: some common assumptions to reconsider in the modeling of the brain and behavior. *Journal of Experimental and Theoretical Artificial Intelligence* 27, 1–26.
- Edelman, S. (2017a). Perception of object shapes. In S. Gepshtein and L. Maloney (Eds.), *The Oxford Handbook of Computational Perceptual Organization*. New York, NY: Oxford University Press. Forthcoming.
- Edelman, S. (2017b). Verbal behavior without syntactic structures: beyond Skinner and Chomsky. In C. Behme (Ed.), *Chomsky’s Legacy*. Forthcoming.
- Edelman, S. and N. Intrator (2003). Towards structural systematicity in distributed, statically bound visual representations. *Cognitive Science* 27, 73–109.
- Edelman, S. and R. Shahbazi (2012). Renewing the respect for similarity. *Frontiers in Computational Neuroscience* 6, 45.
- Edelman, S. and Z. Solan (2009). Machine translation using automatically inferred construction-based correspondence and language models. In B. T’sou and C. Huang (Eds.), *Proc. 23rd Pacific Asia Conference on Language, Information, and Computation (PACLIC)*, Hong Kong.
- Edelman, S., Z. Solan, D. Horn, and E. Ruppin (2004). Bridging computational, formal and psycholinguistic approaches to language. In K. Forbus, D. Gentner, and T. Regier (Eds.), *Proc. of the 26th Conference of the Cognitive Science Society*, Chicago, IL, pp. 345–350.
- Edelman, S. and H. R. Waterfall (2007). Behavioral and computational aspects of language and its acquisition. *Physics of Life Reviews* 4, 253–277.
- Edwards, G. B. and R. R. Jackson (1994). The role of experience in the development of predatory behaviour in *Phidippus regius*, a jumping spider (Araneae, Salticidae) from Florida. *New Zealand Journal of Zoology* 21, 269–277.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science* 14, 179–211.
- Elman, J. L., E. A. Bates, M. H. Johnson, A. Karmiloff-Smith, D. Parisi, and K. Plunkett (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.
- Engel, A. K. and W. Singer (2001). Temporal binding and the neural correlates of sensory awareness. *Trends in Cognitive Sciences* 5, 16–25.

- Epstein, R. (1991). Skinner, creativity, and the problem of spontaneous behavior. *Psychological Science* 2, 362–370.
- Everaert, M. B. H., M. A. C. Huybregts, N. Chomsky, R. C. Berwick, and J. J. Bolhuis (2015). Structures, not strings: linguistics as part of the cognitives sciences. *Trends in Cognitive Sciences* 19, 729–743.
- Farmer, A. (1990). Story-telling in history. *Teaching History* 58, 17–23.
- Fedorenko, E. and S. L. Thompson Schill (2014). Reworking the language network. *Trends in Cognitive Sciences* 18, 120–126.
- Feldman, J. A. and D. H. Ballard (1982). Connectionist models and their properties. *Cognitive Science* 6, 205–254.
- Feldman, J. A., M. A. Fianty, N. H. Goddard, and K. J. Lynne (1988). Computing with structured connectionist networks. *Communications of the ACM* 31, 170–187.
- Fernando, C. (2011). Symbol manipulation and rule learning in spiking neuronal networks. *Journal of Theoretical Biology* 275, 29–41.
- Fernando, C., V. Vasas, E. Szathmáry, and P. Husbands (2011). Evolvable neuronal paths: A novel basis for information and search in the brain. *PLoS ONE* 6(8), e23534.
- Ferreira, F. and K. G. D. Bailey (2004). Disfluencies and human language comprehension. *Trends in Cognitive Sciences* 8, 231–237.
- Fetsch, C. R., G. C. DeAngelis, and D. E. Angelaki (2013). Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. *Nature Reviews Neuroscience* 14, 429–442.
- Fiete, I. R. and H. S. Seung (2009). Birdsong learning. In L. Squire (Ed.), *Encyclopedia of Neuroscience*, pp. 227–239. New York: Elsevier.
- Fitch, W. T. and M. D. Hauser (2004). Computational constraints on syntactic processing in nonhuman primates. *Science* 303, 377–380.
- Fivush, R. and C. A. Haden (2003). *Autobiographical Memory and the Construction of a Narrative Self: Developmental and Cultural Perspectives*. Mahwah, NJ: Erlbaum.
- Fodor, J. D. and L. Frazier (1980). Is the human sentence parsing mechanism an ATN? *Cognition* 8, 417–459.
- Fortin, N. J., K. L. Agster, and H. B. Eichenbaum (2002). Critical role of the hippocampus in memory for sequences of events. *Nature Neuroscience* 5, 458–462.
- Frank, R. (2004). Restricting grammatical complexity. *Cognitive Science* 28, 669–697.
- Freeberg, T. M., R. I. M. Dunbar, and T. J. Ord (2012). Social complexity as a proximate and ultimate factor in communicative complexity. *Phil. Trans. R. Soc. B* 367, 1785–1801.
- Fried, I., K. A. MacDonald, and C. L. Wilson (1997). Single neuron activity in human hippocampus and amygdala during recognition of faces and objects. *Neuron* 18, 753–765.

- Friederici, A. D. and W. Singer (2015). Grounding language processing on basic neurophysiological principles. *Trends in Cognitive Sciences* 19, 329–338.
- Frith, C. and R. Dolan (1996). The role of the prefrontal cortex in higher cognitive functions. *Cognitive Brain Research* 5, 175–181.
- Futrell, R., K. Mahowald, and E. Gibson (2015). Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Science* 112, 10336–10341.
- Galef, Jr, B. G. and L.-A. Giraldeau (2001). Social influences on foraging in vertebrates: causal mechanisms and adaptive functions. *Animal Behaviour* 61, 3–15.
- Gallant, S. I. and T. W. Okaywe (2013). Representing objects, relations, and sequences. *Neural Computation* 25, 2038–2078.
- Gardner, H. (1993). From conflict to clarification: a comment on Egan’s *Narrative and Learning: A Voyage of Implications*: *Linguistics and Education* 5, 181–185.
- Garfield, Z. H., M. J. Garfield, and B. S. Hewlett (2016). A cross-cultural analysis of hunter-gatherer social learning. In H. Terashima and B. S. Hewlett (Eds.), *Social Learning and Innovation in Contemporary Hunter-Gatherers*, Replacement of Neanderthals by Modern Humans, pp. 19–34. Springer Japan.
- Garrod, S. and M. J. Pickering (2004). Why is conversation so easy? *Trends in Cognitive Sciences* 8, 8–11.
- Gentner, T. Q., K. M. Fenn, D. Margoliash, and H. C. Nusbaum (2006). Recursive syntactic pattern learning by songbirds. *Nature* 440, 1204–1207.
- Gheysen, F., F. Van Opstal, C. Roggeman, H. Van Waelvelde, and W. Fias (2010). Hippocampal contribution to early and later stages of implicit motor sequence learning. *Exp. Brain Res.* 202, 795–807.
- Goldsmith, J. (2001). Unsupervised learning of the morphology of a natural language. *Computational Linguistics* 27, 153–198.
- Goldstein, M. H., H. R. Waterfall, A. Lotem, J. Halpern, J. Schwade, L. Onnis, and S. Edelman (2010). General cognitive principles for learning structure in time and space. *Trends in Cognitive Sciences* 14, 249–258.
- Goodman, J. T. (2001). A bit of progress in language modeling. *Computer Speech and Language* 15, 403–434.
- Graybiel, A. M. (1998). The basal ganglia and chunking of action repertoires. *Neurobiology of Learning and Memory* 70, 119–136.
- Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Annual Review of Neuroscience* 31, 359–387.
- Green, S. and P. Marler (1979). The analysis of animal communication. In P. Marler and J. G. Vandenberg (Eds.), *Handbook of behavioral neurobiology: Vol.3. Social behavior and communication*, pp. 73–158. New York: Plenum Press.

- Grünwald, P. (1994). A minimum description length approach to grammar inference. In G. Scheler, S. Wernter, and E. Riloff (Eds.), *Connectionist, statistical and symbolic approaches to learning for natural language*, Volume 1004 of *Lecture Notes in AI*, pp. 203–216. Berlin: Springer Verlag.
- Gupta, A. S., M. A. A. van der Meer, D. S. Touretzky, and A. D. Redish (2010). Hippocampal replay is not a simple function of experience. *Neuron* 65, 695–705.
- Haber, S. N. and T. E. J. Behrens (2014). The neural network underlying incentive-based learning: implications for interpreting circuit disruptions in psychiatric disorders. *Neuron* 83, 1019–1029.
- Haber, S. N. and R. Calzavara (2009). The cortico-basal ganglia integrative network: The role of the thalamus. *Brain Research Bulletin* 78, 69–74.
- Hackman, D. A., M. J. Farah, and M. J. Meaney (2010). Socioeconomic status and the brain: mechanistic insights from human and animal research. *Nature Reviews Neuroscience* 11, 651–659.
- Haggard, P. and V. Chambon (2012). Sense of agency. *Current Biology* 22, R390–R392.
- Hagoort, P. (2014). Nodes and networks in the neural architecture for language: Broca’s region and beyond. *Current Opinion in Neurobiology* 28, 136–141.
- Hagoort, P. and P. Indefrey (2014). The neurobiology of language beyond single words. *Annual Review of Neuroscience* 37, 347–362.
- Harel, D. (1988). On visual formalisms. *Commun. ACM* 31, 514–530.
- Harris, K. D. and T. D. Mrsic Flogel (2013). Cortical connectivity and sensory coding. *Nature* 503, 51–58.
- Harris, Z. S. (1946). From morpheme to utterance. *Language* 22, 161–183.
- Harris, Z. S. (1991). *A theory of language and information*. Oxford: Clarendon Press.
- Hart, B. and T. R. Risley (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brooks.
- Harvey, C. D., P. Coen, and D. W. Tank (2012). Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature* 484, 62–68.
- Havas, D. A. and J. Matheson (2013). The functional role of the periphery in emotional language comprehension. *Frontiers in Psychology* 4, 294.
- Hebb, A. O. and G. A. Ojemann (2013). The thalamus and language revisited. *Brain & Language* 126, 99–108.
- Heisenberg, M. (2014). The beauty of the network in the brain and the origin of the mind in the control of behavior. *J. Neurogenetics* 28, 389–399.
- Hermann, K. M. and P. Blunsom (2013). The role of syntax in vector space models of compositional semantics. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, Sofia, Bulgaria, pp. 894–904.

- Hermans, H. J. M., H. J. G. Kempen, and R. J. P. van Loon (1992). The dialogical self: beyond individualism and rationalism. *American Psychologist* 47, 23–33.
- Herrmann, E., J. Call, M. Lloreda, B. Hare, and M. Tomasello (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science* 317, 1360–1366.
- Hill, D. E. (1978). *Orientation by jumping spiders of the genus Phidippus (Araneae, Salticidae) during the pursuit of prey*. Ph. D. thesis, The University of Florida.
- Hills, T. T., P. M. Todd, D. Lazer, A. D. Redish, and I. D. Couzin (2015). Exploration versus exploitation in space, mind, and society. *Trends in Cognitive Sciences* 19, 46–54.
- Hinton, G. E. (2007). Learning multiple layers of representation. *Trends in Cognitive Sciences* 11, 428–434.
- Hochmann, J.-R., A. D. Endress, and J. Mehler (2010). Word frequency as a cue for identifying function words in infancy. *Cognition* 115, 444–457.
- Hochreiter, S. and J. Schmidhuber (1997b). Long Short-Term Memory. *Neural Computation* 9, 1735–1780.
- Hockett, C. (1968). *The State of the Art*. The Hague: Mouton.
- Hockett, C. F. (1960). The origin of speech. *Scientific American* 203, 88–96.
- Hopcroft, J. E. and J. D. Ullman (1979). *Introduction to Automata Theory, Languages, and Computation*. Reading, MA: Addison-Wesley.
- Hornik, K., M. Stinchcombe, and H. White (1989). Multilayer feedforward networks are universal approximators. *Neural Networks* 2, 359–366.
- Hosaka, R., O. Araki, and T. Ikeguchi (2008). STDP provides the substrate for igniting synfire chains by spatiotemporal input patterns. *Neural Computation* 20, 415–435.
- Houghton, G. and T. Hartley (1996). Parallels models of serial behaviour: Lashley revisited. *Psyche* 2(25). Symposium on Implicit Learning.
- Hudson, R. (2007). *Language networks: the new word grammar*. New York, NY: Oxford University Press.
- Hull, C. L. (1934). The concept of the habit-family hierarchy and maze learning: Part I. *Psychological Review* 41, 33–54.
- Hummel, J. E. (2001). Complementary solutions to the binding problem in vision: Implications for shape perception and object recognition. *Visual Cognition* 8, 489–517.
- Hummel, J. E. and I. Biederman (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review* 99, 480–517.
- Hummel, J. E. and K. J. Holyoak (1998). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review* 104, 427–466.
- Humphries, M. D. and T. J. Prescott (2010). The ventral basal ganglia, a selection mechanism at the crossroads of space, strategy, and reward. *Progress in Neurobiology* 90, 385–417.

- Huttenlocher, J., H. R. Waterfall, M. Vasilyeva, J. Vevea, and L. V. Hedges (2010). Sources of variability in children's language growth. *Cognitive Psychology* 61, 343–365.
- Huys, R., D. Perdikis, and V. K. Jirsa (2014). Functional architectures and structured flows on manifolds: a dynamical framework for motor behavior. *Psychological Review* 121, 302–336.
- Iberall, A. and W. S. McCulloch (1968). Behavioral model of man. His chains revealed. *Currents in Modern Biology* 1, 337–352.
- Ikegaya, Y., G. Aaron, R. Cossart, D. Aronov, I. Lampl, D. Ferster, and R. Yuste (2004). Synfire chains and cortical songs: Temporal modules of cortical activity. *Science* 304, 559–564.
- Izhikevich, E. M. (2007). Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cerebral Cortex* 17, 2443–2452.
- Jackendoff, R. (2011). What is the human language faculty: two views. *Language* 87, 586–624.
- Jackendoff, R. and S. Pinker (2005). The nature of the language faculty and its implications for evolution of language (reply to Fitch, Hauser, and Chomsky). *Cognition* 97, 211–225.
- Jackson, R. R. and F. R. Cross (2011). Spider cognition. *Advances in Insect Physiology* 41, 115–174.
- Jäkel, F., B. Schölkopf, and F. A. Wichmann (2007). A tutorial on kernel methods for categorization. *Journal of Mathematical Psychology* 51, 343–358.
- Jäkel, F., B. Schölkopf, and F. A. Wichmann (2009). Does cognitive science need kernels? *Trends in Cognitive Sciences* 13, 381–388.
- Jensen, O. and J. E. Lisman (2005). Hippocampal sequence-encoding driven by a cortical multi-item working memory buffer. *Trends in Neurosciences* 28, 67–72.
- Jin, D. Z. (2009). Generating variable birdsong syllable sequences with branching chain networks in avian premotor nucleus HVC. *Physical Review E* 80, 051902.
- Jin, X. and R. M. Costa (2010). Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature* 466, 457–462.
- Jin, X., F. Tecuapetla, and R. M. Costa (2014). Basal ganglia subcircuits distinctively encode the parsing and concatenation of action sequences. *Nature Neuroscience* 17, 423–433.
- Joel, D. and I. Weiner (1994). The organization of the basal ganglia-thalamocortical circuits: open interconnected rather than closed segregated. *Neuroscience* 63, 363–379.
- Johnson, K. E. (2004). On the systematicity of language and thought. *Journal of Philosophy* CI, 111–139.
- Joris, P. X., P. H. Smith, and T. C. T. Yin (1998). Coincidence detection in the auditory system: 50 years after Jeffress. *Neuron* 21, 1235–1238.
- Joshi, A. K. (1985). How much context-sensitivity is necessary for characterizing structural descriptions. In D. Dowty, L. Karttunen, and A. Zwicky (Eds.), *Natural Language Processing: Theoretical, Computational and Psychological Perspectives*, pp. 206–250. New York, NY: Cambridge University Press.

- Joshi, A. K. (2004). Starting with complex primitives pays off: complicate locally, simplify globally. *Cognitive Science* 28, 637–668.
- Joshi, A. K. and Y. Schabes (1997). Tree-Adjoining Grammars. In G. Rozenberg and A. Salomaa (Eds.), *Handbook of Formal Languages*, Volume 3, pp. 69–124. Berlin: Springer.
- Kamil, A. C. and T. D. Sargent (1981). *Foraging behavior: Ecological, ethological, and psychological approaches*. New York: Garland.
- Katahira, K., K. Suzuki, K. Okanoya, and M. Okada (2011). Complex sequencing rules of birdsong can be explained by simple Hidden Markov processes. *PLoS ONE* 6:9, e24516.
- Kauffman, S. A. and P. Clayton (2006). On emergence, agency, and organization. *Biology and Philosophy* 21, 501–521.
- Kazakov, D. and M. Bartlett (2005). Could navigation be the key to language? In *Proc. of the 2nd Symposium on the Emergence and Evolution of Linguistic Communication (EELC 2005)*, pp. 50–55.
- Kershenbaum *et al.*, A. (2015). Acoustic sequences in non-human animals: A tutorial review and prospectus. *Biological Reviews*.
- Klostermann, F., L. K. Krugel, and F. Ehlen (2013). Functional roles of the thalamus for language capacities. *Frontiers in Systems Neuroscience* 7:32, 1–8.
- Kolodny, O. and S. Edelman (2015). The problem of multimodal concurrent serial order in behavior. *Neuroscience and Biobehavioral Reviews* 56, 252–265.
- Kolodny, O. and S. Edelman (2016, November). The leap of language: the ecological context of language evolution. Poster presented at the Sackler colloquium on the Extension of Biology Through Culture, UC Irvine.
- Kolodny, O. and S. Edelman (2017). The evolution of the capacity for language: the ecological context of a process of cognitive hijacking. In preparation.
- Kolodny, O., S. Edelman, and A. Lotem (2014). The evolution of continuous learning of the structure of the environment. *Journal of the Royal Society Interface* 11, 20131091.
- Kolodny, O., S. Edelman, and A. Lotem (2015). Evolution of protolinguistic abilities as a byproduct of learning to forage in structured environments. *Proceedings of the Royal Society of London B* 282, 20150353.
- Kolodny, O., A. Lotem, and S. Edelman (2015). Learning a generative probabilistic grammar of experience: a process-level model of language acquisition. *Cognitive Science* 39, 227–267.
- Koziol, L. F., D. Budding, N. Andreasen, S. D’Arrigo, S. Bulgheroni, H. Imamizu, M. Ito, M. Manto, C. Marvel, K. Parker, G. Pezzulo, N. Ramnani, D. Riva, J. Schmammann, L. Vandervert, and T. Yamazaki (2014). Consensus paper: the cerebellum’s role in movement and cognition. *Cerebellum* 13, 151–177.
- Kriete, T., D. C. Noelle, J. D. Cohen, and R. C. O’Reilly (2013). Indirection and symbol-like processing in the prefrontal cortex and basal ganglia. *Proceedings of the National Academy of Science* 110, 16390–16395.

- Kurczek, J., S. Brown-Schmidt, and M. Duff (2013). Hippocampal contributions to language: Evidence of referential processing deficits in amnesia. *Journal of Experimental Psychology: General* 142, 1346–1354.
- Lachmann, M., S. Számadó, and C. T. Bergstrom (2001). Cost and conflict in animal signals and human language. *Proceedings of the National Academy of Science* 98, 13189–13194.
- Laland, K. N. (2017). The origins of language in teaching. *Psychonomic Bulletin and Review* 24, 225–231.
- Lamb, S. M. (1998). *Pathways of the brain: the neurocognitive basis of language*. Amsterdam: John Benjamins.
- LaPolla, R. J. (2015). On the logical necessity of a cultural and cognitive connection for the origin of all aspects of linguistic structure. In R. De Busser and R. J. LaPolla (Eds.), *Language Structure and Environment: Social, cultural and natural factors*, pp. 31–44. Amsterdam: John Benjamins.
- Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral Mechanisms in Behavior*, pp. 112–146. New York: Wiley.
- Lau, J. H., A. Clark, and S. Lappin (2016). Grammaticality, acceptability, and probability: a probabilistic view of linguistic knowledge. *Cognitive Science*. In press.
- Lebiere, C. and J. R. Anderson (1993). A connectionist implementation of the ACT-R production system. In *Proceedings of the 15th Annual Conference of the Cognitive Science Society*, pp. 635–640.
- LeCun, Y., Y. Bengio, and G. Hinton (2015). Deep learning. *Nature* 521, 436–444.
- Lehmann, L., K. R. Foster, E. Borenstein, and M. W. Feldman (2008). Social and individual learning of helping in humans and other species. *Trends in Ecology and Evolution* 23, 664–671.
- Leitman, D. I., J. C. Edgar, J. Berman, K. Gamez, S. Frühholz, and T. P. Robert (2017). Amygdala and insula contributions to dorsal-ventral pathway integration in the prosodic neural network. *PLoS Biology*. Submitted.
- Lenneberg, E. H. (1967). *Biological foundations of language*. New York, NY: Wiley.
- Lertsethtakarn, P., K. M. Ottemann, and D. R. Hendrixson (2011). Motility and chemotaxis in *Campylobacter* and *Helicobacter*. *Annual Review of Microbiology* 65, 389–410.
- Levy, R., K. Bicknell, T. Slattery, and K. Rayner (2009). Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. *Proceedings of the National Academy of Science* 106, 21086–21090.
- Levy, W. B. (1996). A sequence predicting CA3 is a flexible associator that learns and uses context to solve hippocampal-like tasks. *Hippocampus* 6, 579–590.
- Levy, W. B., A. B. Hocking, and X. Wu (2005). Interpreting hippocampal function as recoding and forecasting. *Neural Networks* 18, 1242–1264.
- Lewis, M. D. (2002). The dialogical brain: contributions of emotional neurobiology to understanding the dialogical self. *Theory & Psychology* 12, 175–190.

- Lindquist, K. A. and L. Feldman Barrett (2012). A functional architecture of the human brain: emerging insights from the science of emotion. *Trends in Cognitive Sciences* 16, 533–540.
- Linell, P. (2013). Distributed language theory, with or without dialogue. *Language Sciences* 40, 168–173.
- Linell, P. (2016). On agency in situated languaging: Participatory agency and competing approaches. *New Ideas in Psychology* 42, 39e45.
- Littman, M. L. (2015). Reinforcement learning improves behaviour from evaluative feedback. *Nature* 521, 445–451.
- Litvak, S. and S. Ullman (2009). Cortical circuitry implementing graphical models. *Neural Computation* 21, 3010–3056.
- Llinás, R. R. and A. Iberall (1977). A global model of neuronal command-control systems. *BioSystems* 8, 233–235.
- Locke, J. L. and B. Bogin (2006). Language and life history: A new perspective on the development and evolution of human language. *Behavioral and Brain Sciences* 29, 259–280.
- Longworth, C. E., S. E. Keenan, R. A. Barker, W. D. Marslen-Wilson, and L. K. Tyler (2005). The basal ganglia and rule-governed language use: evidence from vascular and degenerative conditions. *Brain* 128, 584–596.
- Maass, W. (2007). Liquid computing. In *Proceedings of the CiE'07 Conference: Computability in Europe 2007*, Lecture Notes in Computer Science, Berlin. Springer.
- Maass, W. and H. Markram (2004). On the computational power of recurrent circuits of spiking neurons. *Journal of Computer and System Sciences* 69, 593–616.
- Macedonia, J. M. and C. S. Evans (1993). Variation among mammalian alarm call systems and the problem of meaning in animal signals. *Ethology* 93, 177–197.
- Marcus, S. (1969). Contextual grammars. *Revue Roum. Math. Pures Appl.* 14, 1473–1482.
- Marcus, S., C. Martín-Vide, and G. Păun (1998). Contextual grammars as generative models of natural languages. *Computational Linguistics* 24, 245–274.
- Marr, D. (1982). *Vision*. San Francisco, CA: W. H. Freeman.
- Marr, D. and T. Poggio (1977). From understanding computation to understanding neural circuitry. *Neurosciences Res. Prog. Bull.* 15, 470–488.
- Mayr, E. (1961). Cause and effect in biology. *Science* 134, 1501–1506.
- McCulloch, W. S. (1950). Brain and behavior. In W. C. Halstead (Ed.), *Comparative Psychology Monograph*, Volume 20, pp. 39–50. Berkeley, CA: U. of Calif. Press.
- McCulloch, W. S. and W. Pitts (1943). A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 5, 115–133. Reprinted in *Embodiments of Mind*, pp.46-66, MIT Press, 1965.

- McFarland, D. A., D. Jurafsky, and C. Rawlings (2013). Making the connection: social bonding in courtship situations. *American Journal of Sociology* 118, 1596–1649.
- Meilinger, T., M. Knauff, and H. H. Bühlhoff (2008). Working memory in wayfinding — a dual task experiment in a virtual city. *Cognitive Science* 32, 755–770.
- Melis, A. P., B. Hare, and M. Tomasello (2006). Chimpanzees recruit the best collaborators. *Science* 311, 1297–1300.
- Mel'čuk, I. (2003). Levels of dependency in linguistic description: Concepts and problems. In V. Agel, L. Eichinger, H.-W. Eroms, P. Hellwig, H. J. Herringer, and H. Lobin (Eds.), *Dependency and Valency. An International Handbook of Contemporary Research*, Volume 1, pp. 188–229. Berlin - New York: W. de Gruyter.
- Mendoza, G. and H. Merchant (2014). Motor system evolution and the emergence of high cognitive functions. *Progress in Neurobiology* 122, 73–93.
- Menyhart, O., O. Kolodny, M. H. Goldstein, T. Devoogd, and S. Edelman (2015). Juvenile zebra finches learn the underlying statistical regularities in their father's song. *Frontiers in Psychology* 6, 571.
- Merker, B. (2004). Cortex, countercurrent context, and dimensional integration of lifetime memory. *Cortex* 40, 559–576.
- Merker, B. (2013). The efference cascade, consciousness, and its self: naturalizing the first-person pivot of action control. *Frontiers in Psychology* 4(501), 1–20.
- Miller, E. K. and J. D. Cohen (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience* 24, 167–202.
- Mnih, V., K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis (2015). Human-level control through deep reinforcement learning. *Nature* 518, 529–533.
- Moerk, E. L. (1976). Processes of language teaching and training in the interactions of mother-child dyads. *Child Development* 47, 1064–1078.
- Mumford, D. (1994). Neuronal architectures for pattern-theoretic problems. In C. Koch and J. L. Davis (Eds.), *Large-scale neuronal theories of the brain*, Chapter 7, pp. 125–152. Cambridge, MA: MIT Press.
- Nakahara, H., K. Doya, and O. Hikosaka (2001). Parallel cortico-basal ganglia mechanisms for acquisition and execution of visuomotor sequences — a computational approach. *Journal of Cognitive Neuroscience* 13, 626–647.
- Neuman, Y. (2006). A theory of meaning. *Information Sciences* 176, 1435–1449.
- Newmeyer, F. (1998). *Language Form and Language Function*. Cambridge, MA: MIT Press.
- Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology* 53, 139–154.

- Nummenmaa, L., H. Saarimäki, E. Glerean, A. Gotsopoulos, I. P. Jääskeläinen, R. Hari, and M. Sams (2014). Emotional speech synchronizes brains across listeners and engages large-scale dynamic brain networks. *NeuroImage* 102, 498–509.
- O’Keefe, J. and J. Dostrovsky (1971). The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely moving rat. *Brain Research* 34, 171–175.
- Onnis, L., H. R. Waterfall, and S. Edelman (2008). Learn locally, act globally: Learning language from variation set cues. *Cognition* 109, 423–430.
- O’Reilly, R. C. and M. J. Frank (2006). Making working memory work: a computational model of learning in the frontal cortex and basal ganglia. *Neural Computation* 18, 283–328.
- Ortiz-Mantilla, S., M. Choe, J. Flax, P. E. Grant, and A. A. Benasich (2010). Associations between the size of the amygdala in infancy and language abilities during the preschool years in normally developing children. *NeuroImage* 49, 2791–2799.
- Ota, N., M. Gahr, and M. Soma (2015). Tap dancing birds: the multimodal mutual courtship display of males and females in a socially monogamous songbird. *Nature Scientific Reports* 5, 16614.
- Ouattara, K., A. Lemasson, and K. Zuberbühler (2010). Campbell’s monkeys concatenate vocalizations into context-specific call sequences. *Proceedings of the National Academy of Science* 106, 22026–22031.
- Pack Kaelbling, L., M. L. Littman, and A. W. Moore (1996). Reinforcement learning: A survey. *Journal of Artificial Intelligence Research* 4, 237–285.
- Park, H.-J. and K. J. Friston (2013). Structural and functional brain networks: from connections to cognition. *Science* 342, 1238411.
- Parkinson, J. S., S. R. Parker, P. B. Talbert, and S. E. Houts (1983). Interactions between chemotaxis genes and flagellar genes in *Escherichia coli*. *Journal of Bacteriology* 155, 265–274.
- Parvizi, J. (2009). Corticocentric myopia: old bias in new cognitive sciences. *Trends in Cognitive Sciences* 13, 354–359.
- Pastalkova, E., V. Itskov, A. Amarasingham, and G. Buzsáki (2008). Internally generated cell assembly sequences in the rat hippocampus. *Science* 321, 1322–1327.
- Peña, B. J. F. J. L. (2011). Owl’s behavior and neural representation predicted by Bayesian inference. *Nature Neuroscience* 14, 1061–1067.
- Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience* 9, 148–158.
- Pezzulo, G., M. A. A. van der Meer, C. S. Lansink, and C. M. A. Pennartz (2014). Internally generated sequences in learning and executing goal-directed behavior. *Trends in Cognitive Sciences* 18, 647–657.
- Pfeiffer, B. E. and D. J. Foster (2013). Hippocampal place-cell sequences depict future paths to remembered goals. *Nature* 497, 74–79.

- Pfeiffer, B. E. and D. J. Foster (2015). Autoassociative dynamics in the generation of sequences of hippocampal place cells. *Science* 349, 180–183.
- Phillips, C. (2003a). Linear order and constituency. *Linguistic Inquiry* 34, –. in press.
- Phillips, C. (2003b). Syntax. In L. Nadel (Ed.), *Encyclopedia of Cognitive Science*, Volume 4, pp. 319–329. London: Macmillan.
- Piai, V., K. L. Anderson, J. J. Lin, C. Dewar, J. Parvizi, N. F. Dronkers, and R. T. Knight (2016). Direct brain recordings reveal hippocampal rhythm underpinnings of language processing. *Proceedings of the National Academy of Science* 113, 11366–11371.
- Pickering, M. J. and H. P. Branigan (1999). Syntactic priming in language production. *Trends in Cognitive Sciences* 3, 136–141.
- Pinker, S. (2010). The cognitive niche: Coevolution of intelligence, sociality, and language. *Proceedings of the National Academy of Science* 107(Suppl. 2), 8993–8999.
- Pinker, S. and R. Jackendoff (2005). The faculty of language: What’s special about it? *Cognition* 95, 201–236.
- Plate, T. A. (1995). Holographic Reduced Representations. *IEEE Transactions on Neural Networks* 6, 623–641.
- Poeppl, D. (2014). The neuroanatomic and neurophysiological infrastructure for speech and language. *Current Opinion in Neurobiology* 28, 142–149.
- Poeppl, D., K. Emmorey, G. Hickok, and L. Pylkkänen (2012). Towards a new neurobiology of language. *The Journal of Neuroscience* 32, 14125–14131.
- Poggio, T. (2012). The levels of understanding framework, revised. *Perception* 41, 1017–1023.
- Pollard, C. and I. A. Sag (1992). Anaphors in English and the scope of binding theory. *Linguistic Inquiry* 23, 261–303.
- Port, R. F. and A. P. Leary (2005). Against formal phonology. *Language* 81, 927–964.
- Postal, P. (2008). Edge-based model-theoretic syntax. Unpublished ms.
- Postal, P. M. (2004). *Skeptical linguistic essays*. New York: Oxford University Press.
- Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience* 6, 576–582.
- Rabinovich, M. I., R. Huerta, P. Varona, and V. S. Afraimovich (2008). Transient cognitive dynamics, metastability, and decision making. *PLoS Comput. Biol.* 4(5), e1000072.
- Ramscar, M. and H. Baayen (2013). Production, comprehension, and synthesis: a communicative perspective on language. *Frontiers in Psychology* 4, 233.
- Ramscar, M. and R. F. Port (2016). How spoken languages work in the absence of an inventory of discrete units. *Language Sciences* 53, 58–74.

- Rao, R. P. and D. H. Ballard (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience* 2, 79–87.
- Redgrave, P., N. Vautrelle, and J. N. J. Reynolds (2011). Functional properties of the basal ganglia’s re-entrant loop architecture: selection and reinforcement. *Neuroscience* 198, 138–151.
- Rein, R., B. Bril, and T. Nonaka (2013). Coordination strategies used in stone knapping. *American Journal of Physical Anthropology* 150, 539–550.
- Rissanen, J. (1987). Minimum description length principle. In S. Kotz and N. L. Johnson (Eds.), *Encyclopedia of Statistic Sciences*, Volume 5, pp. 523–527. J. Wiley and Sons.
- Rodriguez, A. and R. Granger (2016). The grammar of mammalian brain capacity. *Theoretical Computer Science* 633, 100–111.
- Rodriguez, A., J. Whitson, and R. Granger (2004). Derivation and analysis of basic computational operations of thalamocortical circuits. *Journal of Cognitive Neuroscience* 16, 856–877.
- Rohde, D. L. T. (2002). *A Connectionist Model of Sentence Comprehension and Production*. Ph. D. thesis, Carnegie Mellon University. CMU-CS-02-105.
- Rosen, M. L. and H. H. López (2009). Menstrual cycle shifts in attentional bias for courtship language. *Evolution and Human Behavior* 30, 131–140.
- Roy, B. C., M. C. Frank, P. DeCamp, M. Miller, and D. Roy (2015). Predicting the birth of a spoken word. *Proceedings of the National Academy of Science*.
- Rumelhart, D. E., G. E. Hinton, and R. J. Williams (1986). Learning representations by back-propagating errors. *Nature* 323, 533–536.
- Salinas, E. (2009). Rank-order-selective neurons form a temporal basis set for the generation of motor sequences. *J. Neurosci.* 29, 4369–4380.
- Sallabanks, R. (1993). Hierarchical mechanisms of fruit selection by an avian frugivore. *Ecology* 74, 1326–1336.
- Sampson, G. R. (2007). Grammar without grammaticality. *Corpus Linguistics and Linguistic Theory* 3, 1–32.
- Sandler, W. (2006). An overview of sign language linguistics. In K. Brown (Ed.), *Encyclopedia of Language and Linguistics*, 2nd ed., Volume 11, pp. 328–338. Amsterdam: Elsevier.
- Santi, A. and Y. Grodzinsky (2010). fMRI adaptation dissociates syntactic complexity dimensions. *NeuroImage* 51, 1285–1293.
- Savage-Rumbaugh, S. and R. Lewin (1994). *Kanzi: The Ape at the Brink of the Human Mind*. New York: Wiley.
- Schendan, H. E., M. M. Searl, R. J. Melrose, and C. E. Stern (2003). An fMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. *Neuron* 37, 1013–1025.

- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks* 61, 85–117.
- Schölkopf, B. and A. J. Smola (2002). *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge, MA: MIT Press.
- Schütze, C. T. (1996). *The empirical base of linguistics: grammaticality judgments and linguistic methodology*. Chicago, IL: University of Chicago Press.
- Sebanz, N., H. Bekkering, and G. Knoblich (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences* 10, 70–76.
- Seger, C. A. and C. M. Cincotta (2006). Dynamics of frontal, striatal, and hippocampal systems in rule learning. *Cerebral Cortex* 16, 1546–1555.
- Seyfarth, R. M. and D. L. Cheney (2003). Meaning and emotion in animal vocalizations. *Annals of the New York Academy of Science* 1000, 32–55.
- Shackleton, D. M. and C. C. Shafak (1984). A review of the social behavior of feral and wild sheep and goats. *Journal of Animal Science* 58, 500–509.
- Shahbazi, R., R. Raizada, and S. Edelman (2016). Similarity, kernels, and the fundamental constraints on cognition. *Journal of Mathematical Psychology* 70, 21–34.
- Shamma, S. (2001). On the role of space and time in auditory processing. *Trends in Cognitive Sciences* 5, 340–348.
- Sherman, S. M. (2016). Thalamus plays a central role in ongoing cortical functioning. *Nature Neuroscience* 19, 533–541.
- Sherman, S. M. and R. W. Guillery (2006). *Exploring the Thalamus and Its Role in Cortical Function*. Cambridge, MA: MIT Press.
- Shi, R., J. F. Werker, and A. Cutler (2006). Recognition and representation of function words in English-learning infants. *Infancy* 10, 187–198.
- Silbert, L. J., C. J. Honey, E. Simony, D. Poeppel, and U. Hasson (2014). Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. *Proceedings of the National Academy of Science* 111, E4687–E4696.
- Silver, D., A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis (2016). Mastering the game of Go with deep neural networks and tree search. *Nature* 529, 484–503.
- Singer, W. and C. M. Gray (1995). Visual feature integration and the temporal correlation hypothesis. *Annual review of neuroscience* 18, 555–586.
- Sjöström, J. and W. Gerstner (2010). Spike-timing dependent plasticity. *Scholarpedia* 5(2), 1362.
- Skarda, C. and W. J. Freeman (1987). How brains make chaos in order to make sense of the world. *Behavioral and Brain Sciences* 10, 161–195.

- Skinner, B. F. (1957). *Verbal Behavior*. New York, NY: Appleton-Century-Crofts.
- Sloman, A. and R. Chrisley (2003). Virtual machines and consciousness. *Journal of Consciousness Studies* 10, 113–172.
- Smith, W. J. (1965). Message, meaning, and context in ethology. *The American Naturalist* 99, 405–409.
- Socher, R., C. C.-Y. Lin, A. Y. Ng, and C. D. Manning (2011). Parsing natural scenes and natural language with recursive neural networks. In L. Getoor and T. Scheffer (Eds.), *Proc. 28th Intl. Conf. on Machine Learning*, pp. 129–136.
- Solan, Z., D. Horn, E. Ruppin, and S. Edelman (2005). Unsupervised learning of natural languages. *Proceedings of the National Academy of Science* 102, 11629–11634.
- Solan, Z., E. Ruppin, D. Horn, and S. Edelman (2003). Automatic acquisition and efficient representation of syntactic structures. In S. Thrun (Ed.), *Advances in Neural Information Processing*, Volume 15, Cambridge, MA: MIT Press.
- Solomonoff, R. J. (1964). A formal theory of inductive inference, parts A and B. *Information and Control* 7, 1–22, 224–254.
- Stabler, E. (2013a). The epicenter of linguistic behavior. In M. Sanz, I. Laka, and M. K. Tanenhaus (Eds.), *Language Down the Garden Path: The Cognitive and Biological Basis of Linguistic Structures*, pp. 316–323. New York, NY: Oxford University Press.
- Stabler, E. P. (2013b). Two models of Minimalist, incremental syntactic analysis. *Topics in Cognitive Science* 5, 611–633.
- Stabler, E. P., T. C. Collier, G. M. Kobele, Y. Lee, Y. Lin, J. Riggle, Y. Yao, and C. E. Taylor (2003). The learning and emergence of mildly context sensitive languages. In W. Banzhaf, J. Ziegler, T. Christaller, P. Dittrich, and J. T. Kim (Eds.), *Proc. 7th European Conference, ECAL 2003*, Dortmund, Germany, pp. 525–534.
- Stander, P. E. (1992). Cooperative hunting in lions: the role of the individual. *Behavioral Ecology and Sociobiology* 29, 445–454.
- Steffensen, S. V. (2015). Distributed Language and Dialogism: notes on non-locality, sense-making and interactivity. *Language Sciences* 50, 105–119.
- Stolk, A., L. Verhagen, and I. Toni (2016). Conceptual alignment: how brains achieve mutual understanding. *Trends in Cognitive Sciences* 20, 180–191.
- Strassmann, J. E., O. M. Gilbert, and D. C. Queller (2011). Kin discrimination and cooperation in microbes. *Annual Review of Microbiology* 65, 349–367.
- Sutskever, I. and G. Hinton (2007). Learning multilevel distributed representations for high-dimensional sequences. In M. Meila and X. Shen (Eds.), *Proc. Eleventh International Conference on Artificial Intelligence and Statistics*, pp. 544–551.

- Sutskever, I., O. Vinyals, and Q. V. Le (2014). Sequence to sequence learning with neural networks. In C. Cortes and N. Lawrence (Eds.), *Proc. 27th Neural Information Processing Systems Conference (NIPS)*, Montreal, Canada.
- Sutton, J. (2002). Cognitive conceptions of language and the development of autobiographical memory. *Language & Communication* 22, 375–390.
- Sutton, R. S. and A. G. Barto (1998). *Reinforcement Learning*. Cambridge, MA: MIT Press.
- Syal, S. and B. L. Finlay (2010). Thinking outside the cortex: Social motivation in the evolution and development of language. *Developmental Science* 14, 417–430.
- Takac, M., L. Benuskova, and A. Knott (2012). Mapping sensorimotor sequences to word sequences: A connectionist model of language acquisition and sentence generation. *Cognition* 125, 288–308.
- Tanji, J. and K. Shima (1994). Role for supplementary motor area cells in planning several movements ahead. *Nature* 371, 413–416.
- Tappan, M. B. and L. M. Brown (1989). Stories told and lessons learned: toward a narrative approach to moral development and moral education. *Harvard Educational Review* 59, 182–205.
- Taylor, A., G. Hunt, J. Holzhaider, and R. Gray (2007). Spontaneous metatool use by New Caledonian crows. *Current Biology* 17, 1504–1507.
- Thurstone, L. L. (1923). The stimulus-response fallacy in psychology. *Psychological Review* 30, 354–369.
- Tinbergen, N. (1963). On aims and methods in ethology. *Zeitschrift für Tierpsychologie* 20, 410–433.
- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review* 55, 189–208.
- Tong, M. H., A. D. Bickett, E. M. Christiansen, and G. W. Cottrell (2007). Learning grammatical structure with Echo State Networks. *Neural Networks* 20, 424–432.
- Tsuda, I. (1996). A new type of self-organization associated with chaotic dynamics in neural networks. *International Journal of Neural Systems* 7, 451–459.
- Tsuda, I. (2015). Chaotic itinerancy and its roles in cognitive neurodynamics. *Current Opinion in Neurobiology* 31, 67–71.
- Ullman, M. T. (2001). A neurocognitive perspective on language: the declarative/procedural model. *Nature Reviews Neuroscience* 2, 717–727.
- Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative/procedural model. *Cognition* 92, 231–270.
- Ullman, M. T. (2006). Is Broca's area part of a basal ganglia thalamocortical circuit? *Cortex* 42, 480–485.
- Upright, R. L. (2002). To tell a tale: the use of moral dilemmas to increase empathy in the elementary school child. *Early Childhood Education Journal* 30, 15–20.

- Van de Cavey, J. and R. J. Hartsuiker (2016). Is there a domain-general cognitive structuring system? Evidence from structural priming across music, math, action descriptions, and language. *Cognition* 146, 172–184.
- van der Lely, H. K. J. and S. Pinker (2014). The biological basis of language: insight from developmental grammatical impairments. *Trends in Cognitive Sciences* 18, 586–595.
- van Schijndel, M., A. Exley, and W. Schuler (2012). Connectionist-inspired incremental PCFG parsing. In *Proceedings of the 3rd Workshop on Cognitive Modeling and Computational Linguistics*, Stroudsburg, PA, pp. 51–60. Association for Computational Linguistics.
- Vigliocco, G., P. Perniss, and D. Vinson (2014). Language as a multimodal phenomenon: implications for language learning, processing and evolution. *Phil. Trans. R. Soc. B* 369, 20130292.
- Walenski, M. and M. T. Ullman (2005). The science of language. *The Linguistic Review* 22, 327–346.
- Walker, S. F. (1992). A brief history of connectionism and its psychological implications. In A. Clark and R. Lutz (Eds.), *Connectionism in Context*, pp. 123–144. Berlin: Springer.
- Wanner, E. and M. Maratsos (1978). An ATN approach to comprehension. In M. Halle, J. Bresnan, and G. A. Miller (Eds.), *Linguistic theory and psychological reality*, pp. 119–161. Cambridge, MA: MIT Press.
- Waterfall, H. R., B. Sandbank, L. Onnis, and S. Edelman (2010). An empirical generative framework for computational modeling of language acquisition. *Journal of Child Language* 37(Special issue 03), 671–703.
- Weber, B., J. Wellmer, M. Reuber, F. Mormann, S. Weis, H. Urbach, J. Ruhlmann, C. E. Elger, and G. Fernandez (2006). Left hippocampal pathology is associated with atypical language lateralization in patients with focal epilepsy. *Brain* 129, 346–351.
- Weitekamp, C. A. and H. A. Hofmann (2014). Evolutionary themes in the neurobiology of social cognition. *Current Opinion in Neurobiology* 28, 22–27.
- Wickelgren, W. (1969). Context-sensitive coding, associative memory, and serial order in (speech) behavior. *Psychological Review* 76, 1–15.
- Wickelgren, W. A. (1979). I liked the postcard you sent Abe and I: Context-sensitive coding of syntax and other procedural knowledge. *Bulletin of the Psychonomic Society* 13, 61–63.
- Wiener, N. (1948). *Cybernetics: Control and Communication in the Animal and the Machine*. Cambridge, MA: MIT Press.
- Wittgenstein, L. (1958). *Philosophical Investigations* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall. Translated by G. E. M. Anscombe.
- Wolff, J. G. (1988). Learning syntax and meanings through optimization and distributional analysis. In Y. Levy, I. M. Schlesinger, and M. D. S. Braine (Eds.), *Categories and Processes in Language Acquisition*, pp. 179–215. Hillsdale, NJ: Lawrence Erlbaum.

- Wood, E., P. A. Dudchenko, R. J. Robitsek, and H. Eichenbaum (2000). Hippocampal neurons encode information about different types of memory episodes occurring in the same location. *Neuron* 27, 623–633.
- Woods, W. A., R. Kaplan, and B. Nash-Webber (1972). The LUNAR sciences natural language information system: Final report. BBN report 2378, Bolt Beranek and Newman.
- Yamashita, Y., T. Okumura, K. Okanoya, and J. Tani (2011). Cooperation of deterministic dynamics and random noise in production of complex syntactical avian song sequences: a neural network model. *Frontiers in Computational Neuroscience* 5:18, 1–12.
- Zylberberg, A., S. Dehaene, P. R. Roelfsema, and M. Sigman (2011). The human Turing machine: a neural framework for mental programs. *Trends in Cognitive Sciences* 15, 293–300.
- Zylberberg, A., D. F. Slezak, P. R. Roelfsema, S. Dehaene, and M. Sigman (2010). The brain’s router: a cortical network model of serial processing in the primate brain. *PLoS Computational Biology* 6:4, e1000765.