CellPress

Review

Local Patterns to Global Architectures: Influences of Network Topology on Human Learning

Elisabeth A. Karuza,[1,*] Sharon L. Thompson-Schill,[1,2] and
Danielle S. Bassett[3,4]

A core question in cognitive science concerns how humans acquire and represent knowledge about their environments. To this end, quantitative theories of learning processes have been formalized in an attempt to explain and predict changes in brain and behavior. We connect here statistical learning approaches in cognitive science, which are rooted in the sensitivity of learners to local distributional regularities, and network science approaches to characterizing global patterns and their emergent properties. We focus on innovative work that describes how learning is influenced by the topological properties underlying sensory input. The confluence of these theoretical approaches and this recent empirical evidence motivate the importance of scaling-up quantitative approaches to learning at both the behavioral and neural levels.

Relating Two Approaches

From the earliest stages of development, the human brain is tasked with the monumental feat of building and efficiently accessing an enormously complex constellation of knowledge. Even the most mundane interactions with our environment require a rich understanding of its component parts as well as of the scales at which they relate to form a larger system. Thus, knowledge can be represented at multiple levels, ranging from local associations between elements to complex networks built from those local associations. Until recently a dominant approach to human learning has focused on micro-level patterns, often the pairwise relationships between the constituents of sensory input. In the present review we turn our attention to exciting advances in the application of network science to the study of broader architectural patterns to which human learners are sensitive.

One source of compelling support for locally driven learning derives from demonstrations that infants can extract words from continuous speech based on the conditional probabilities between syllables [1]. Ongoing work continues to elucidate the power of statistical relationships exploited by both infants and adults, making 'statistical learning' one of the most robust and deeply explored phenomena in the field of cognitive science [2–5]. An underlying rationale has been that local associations, such as co-occurrence frequency or the conditional probabilities that facilitate word segmentation, assist in directing the learner to component parts of a cognitive system. Knowledge of these component parts not only opens up other informative cues to structure ([6] for review) but also spurs the development of sophisticated representations of dependencies between higher-order units (e.g., [7]). While evidence has thus supported a key role for local computations in complex learning environments, intriguing counter-evidence

Trends

Descriptive analytical approaches indicate that diverse facets of the environment adhere to a complex network structure.

Recent advances offer insight into how learners might acquire and access network representations. Specifically, higher-order topological properties of networks have been shown to facilitate learning.

Emerging neuroimaging techniques construe the brain itself as complex system, revealing how network dynamics support learning.

We suggest that network science approaches are compatible with statistical learning approaches to knowledge acquisition. That is, local statistical regularities extracted from sensory input form the building blocks of complex network structures. Broader architectural properties of network structures might then explain learning effects beyond sensitivity to local statistical information.

[1]Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104, USA
[2]Department of Neurology, University of Pennsylvania, Philadelphia, PA 19104, USA
[3]Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104, USA
[4]Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104, USA

*Correspondence:
ekaruza@sas.upenn.edu (E.A. Karuza).

suggests that statistical bootstrapping mechanisms may be overwhelmed by real-world cognitive systems (e.g., natural language [8]; but see [9,10]). Because the issues of scalability in statistical learning are as yet unresolved, we stress here the value of also considering the global network structure that emerges from pairwise relationships between constituent elements in the environment.

Under a complex systems approach, the network structure of a system is studied by determining its component elements (**nodes**, see Glossary) as well as the relational links between them (**edges**). Once this scaffolding is constructed, it is possible to probe large-scale topological and dynamical properties over and above those present in the pairwise relations between elements. In fact, one defining characteristic of complex systems is that the explanatory power of their global architecture exceeds that of their local architecture [11]. Network science is increasingly applied to answer questions about the structure of immensely complex information: how might we represent or navigate spatial maps [12,13], object features [14], semantic concepts [15–17], and grammatical relationships [18]? It has also been effectively harnessed by cognitive neuroscientists to examine how structural and functional connections in the brain give rise to various cognitive capacities [19–25]. Despite these many advances, the integration of network science and cognitive science has tended to focus either on (i) the description of networks derived from the sensory world, or (ii) the mechanisms by which the human brain engages with the sensory world, with little communication between these two areas. We focus here on a related but distinct line of questioning that may begin to bridge these branches of cognitive science. Namely, how can topological properties of sensory input drive the process of human learning?

In the subsequent sections we offer examples of complex networks present in our everyday environment, focusing particularly on descriptive analyses of language networks. Next, we detail a growing body of experimental work that links topological properties of networks to knowledge acquisition. We then discuss the intersection between distributional approaches to learning, which offer insight into the acquisition of local statistical patterns, and network-based approaches to learning, which offer complementary insight into the acquisition of higher-order patterns. Finally, we describe cutting-edge neuroimaging work that construes the brain itself as a dynamical complex system, highlighting the importance of bridging internal network models of brain function with higher-order patterns in external networks.

Complex Networks Are Pervasive

Complex systems approaches rest on the premise, not tied to any particular domain, that the world can be decomposed into parts, and that those parts interact with one another in meaningful ways. Therefore, diverse facets of human knowledge can and have been studied under the lens of network science. Cognitive systems are generally thought to adhere to a complex network structure, a type of graph structure that is neither truly random nor truly regular [26]. Random graphs are collections of nodes that are linked by edges selected at random from a uniform distribution of all possible connections. Regular graphs are collections of nodes that share connections to the same number of neighbors, thus having equivalent **degree**. Falling between these two extremes (Figure 1), complex networks display their own set of unique properties including, but not limited to: **community structure** (nodes may pattern in densely connected groupings), skewed degree distribution (a few nodes may be densely connected, forming 'hubs'), and distinctive mixing patterns (nodes may be more likely to share a link with other nodes that have either similar or dissimilar properties). As we will explore in detail in the following section, human learners are adept at exploiting topological properties such as these as they extract structure from sensory input.

In principle, network analysis of cognitive systems requires only that a given dataset be parsed into discrete elements (nodes) and that some relationships between those elements be specified

Figure 1. Visualization of Network Types. Regular networks, also known as lattices, are collections of nodes with equivalent degree (left panel). Random networks are collections of nodes that are linked by edges selected at random from a uniform distribution of all possible connections. We show here a random network generated from an Erdős–Rényi model with an edge probability of 0.3 (right panel). In the center panel we display a complex network with community structure, much like a network that could be derived from a learner's language environment.

(edges). In practice, this process presents a number of challenges, not least among them is determining the appropriate level of granularity of the elements or what exactly constitutes a relationship (for a thoughtful assessment of these topics, see [27]). Nevertheless, network-based approaches remain extremely powerful and have been successfully applied to several pressing questions. In the visual domain, complex network modeling techniques have been used to understand processes essential to scene perception such as texture and shape discrimination [28–30]. Moreover, the burgeoning field of social network analysis has offered unprecedented insight into how humans transmit information and interact with one another [31,32]. Perhaps more than any other branch of cognitive science, quantitative linguistics has adopted network-based approaches as a cornerstone methodology [33–36]. Across levels of the language hierarchy, graph theoretical methods have been applied to the study of phonological [37,38], semantic [39–41], and syntactic dependency systems [42,43]. In a phonological network, for example, the nodes of the graph correspond to the phonetic transcription of a word (e.g., as drawn from a dictionary), and edges are placed between words if they differ by no more than one phoneme (Figure 2). In this way, network approaches to cognitive systems tend to be built on existing distributional measures (in this case, phonological neighborhood density [44]). However, as discussed below, the higher-order architectural properties of these cognitive networks likely have their own set of consequences for learnability.

Network Topology Influences Learning and Memory

Historically, complex network analyses of cognitive structures have had a descriptive focus. Naturally, the first step in understanding how networks structures develop [45–47] is to characterize existing topological properties (e.g., based on text or production-based corpora). Only recently has network topology been linked to empirical evidence of human learning. This relationship is typically examined in one of two ways: (i) by exposing adult learners to a set of tightly controlled stimuli and asking how they retain knowledge based on its underlying architecture, and (ii) by extracting the network properties of adult speech and charting its influence on the trajectory of word learning in young children. Most of this empirical work has taken place in the language domain, likely because the topologies of language systems are richly characterized and also because mastery of a complex language system represents an unequivocal learning challenge [35].

Central to the production and comprehension of language lies the construct of a mental lexicon. To interact successfully with our language environment it is essential to develop, adapt, and

High clustering coefficient

Low clustering coefficient
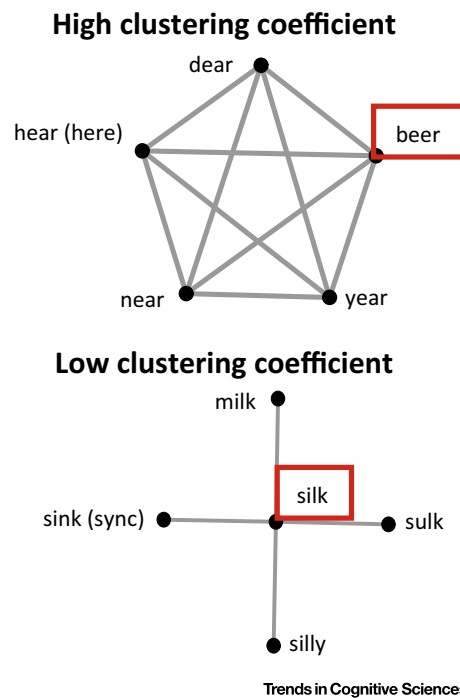

Trends in Cognitive Sciences

Figure 2. Sample (Not Exhaustive) Phonological Networks of Two English Words, 'Beer' and 'Silk', That Differ in Their Clustering Coefficients. Note how the phonological neighbors of 'beer' tend also to be phonological neighbors of each other, resulting in a high clustering coefficient. By contrast, the phonological neighbors of silk are not phonological neighbors, resulting in a low clustering coefficient.

efficiently access the conceptual, grammatical, and sound-based properties of words. Accumulating evidence suggests that phonological (sound-based) network properties play a key role in mediating various aspects of this process. Interestingly, the **clustering coefficient** of a word (construed in this case as a node in a phonological network) has been shown to predict how robustly that word is acquired [48]. The clustering coefficient indicates whether the neighboring nodes of a target word are also connected to each other. In Figure 2 we offer examples of words with high and low clustering coefficients. Words in the phonological neighborhood of 'beer' (i.e., those that overlap but for a single phoneme) also tend to overlap with each other (e.g., hear–year–near–dear), resulting in a high clustering coefficient. By contrast, words in the neighborhood of 'silk' (e.g., milk–sulk–silly–sink) do not share a strong phonological relationship with each other, resulting in a low clustering coefficient. One week after training on pseudoword–novel object pairings, adult participants were better able to match a pseudoword to its corresponding object if that pseudoword had a high clustering coefficient. Complementary results indicate that, after training on the phonological neighbors of a target word (but not the word itself), learners tended to erroneously endorse target words with low clustering coefficient. By contrast, results from a different long-term memory task (in which target words were not withheld) indicated that low clustering coefficient conferred a recognition advantage; correct endorsement was higher when compared to words with high clustering coefficient [49]. These findings are complemented by word recognition studies that demonstrate a processing [50,51] and production [52] advantage for words with low clustering coefficient, suggesting that words with these properties are more rapidly accessed from the mental lexicon. Taken together, these results suggest that dense interconnectedness between nodes in a network can confer an advantage for learning, but a disadvantage for retrieval. The effect of network topology on a given cognitive process may therefore differ depending on the precise nature of the task at hand.

To be clear, the clustering coefficient of a target word represents only one well-studied example of topological influences on learning and memory. When analyzing phonological properties of caregiver speech, it was found that children <30 months of age were most likely to produce words with a high degree and low **coreness** (a measure of how deeply a node is embedded in a network) [53]. Larger-scale architectural properties may also play a role in accessing the mental lexicon. One example of such a macro-level property is **assortative mixing** by degree. This network property encapsulates the phenomenon that words with many phonological neighbors (high degree) are linked to words that also have many phonological neighbors, while words with few phonological neighbors (low degree) are linked to words that share this property. Although assortative mixing has not been examined in a learning context, its effects are observable in psycholinguistic studies of lexical access. When the sound quality of a word with high degree is degraded, listeners are most likely to misperceive that item as a different word with equally high degree [54]. Other higher-level patterns in the phonological network have been implicated in memory processes. For example, words drawn from so-called 'giant components' exhibiting **small-world network** properties such as a high clustering coefficient and **short characteristic path length** are recalled less reliably than words drawn from 'lexical islands' that share no connection to the giant component [55].

Clearly, the network structure of the lexicon has been most commonly framed in terms of phonological patterns. Nonetheless, its semantic organization (i.e., how words are related to each other via their meaning or co-occurrence in a corpus) has also been shown to influence word learning. For example, the distinctness of the features of an object, such as shape and surface properties, has been shown to predict age of acquisition of the word corresponding to that object. Phrased another way, objects that are topologically disconnected from other objects in a semantic network are labeled at the youngest age [14]. Graph analytic techniques have also been successfully applied to model child output [56,57]. In one recent study, a semantic network was constructed such that a node in the network represented a word known by children aged 15–36 months, and each edge represented co-occurrence statistics as derived from a database of caregiver speech [57]. Word knowledge of typically-developing children exhibited characteristic small world properties, but the networks of late-talking children exhibited this quality to a lesser degree. Thus, language acquisition is influenced not only by the network architecture of input but also by differences in how that input is transformed to an individual's own distinct network topology. More broadly, these developmental studies highlight the potential power of harnessing network science to better understand individual differences in core cognitive capacities (see also [58]). In the following section we turn to the essential question of how the learner begins to build complex network representations.

Local Statistics Underpin Network Architecture

The Impact of Local Statistics

As introduced previously, statistical learning persists as an influential and well-supported theory of how learners extract structure from our external world. While we mainly focus on the effect of pair-based conditional probabilities, statistical learning fits into a broader distributional learning literature. In fact, longstanding interest in how we compute local contingencies can be traced to even earlier studies of associative learning mechanisms in animals (e.g., [59]). As used to describe complex knowledge acquisition in humans, the term 'distributional information' encompasses the context in which elements in the environment appear together and with what regularity (e.g., frequency or probability). Sensitivity to distributional information is thus evident in charting the influence of patterns of co-occurrence frequency (e.g., how often do groupings of elements appear together? [60], see also a related literature on 'chunking' in finite-state grammar learning [61]) or the conditional probabilities between elements (e.g., what is the probability that a given element will follow another?). Manipulation of distributional statistics has also used to examine the acquisition of phonetic categories [62,63], phrase structure [64,65], and syntactic

dependencies [66]. Outside such learning contexts, the influence of distributional statistics is also observed in the processing of already familiar input. During natural language comprehension, for example, knowledge about grammatical dependencies result in a processing benefit for expected (probable) structures [67,68]. Likewise, processing costs associated with improbable structures can be overcome by sharply shifting their frequency in the immediate sensory environment [69]. Thus, the computation of statistical regularities is a continuous process that permeates both the acquisition of novel representations and the processing of familiar structures (for further discussion of the intersection between these capacities see [65,70,71]).

While distributional sensitivity is often demonstrated via linguistic stimuli, we note that, outside the language domain, pairwise statistical patterns have been shown to drive the parsing of tonal groupings [72], visual events [73,74], and spatial scenes [75]. Evidence therefore points to statistical learning as a powerful domain-general mechanism, or one that can perform equivalent computations regardless of both input modality (e.g., auditory or visual; but see [76,77]). Despite its wide applicability, an important and growing area of research probes constraints on our distributional learning capacities (i.e., how might this learning mechanism break down in the context of multiple patterns [78,79]? At what levels of abstraction does it operate [80]? And how does learning interact with attentional processes [81,82]?).

### Local Statistics Give Rise to Complex Networks

While the evidence suggests that associations can be computed hierarchically (e.g., to order words into phrases and phrases into sentences [66]; more generally [83]), many learning studies solely involve the manipulation of non-hierarchical, pairwise statistical information such as co-occurrence frequencies or conditional probabilities. From a complex network perspective, this amounts to constructing **dyads** in the network. However, a notable recent study suggests that learners also capitalize on regularities that emerge from global network architecture [84]. Participants in this study viewed a continuous stream of images generated by a random walk through a graph with community structure (i.e., three distinct groupings of interconnected nodes). Each node of the graph corresponded to a unique image, and the random walk through the graph ensured that transition (conditional) probabilities between images in the sequence were uniform (i.e., not a cue to an event boundary, as in canonical segmentation studies; e.g., [1]). Despite the absence of pairwise probabilistic information, learners were able to segment groupings of images based solely on the community structure of the underlying graph, accurately detecting when the stream shifted from one cluster to the next. While this finding represents a tremendous step forward in our understanding of macro-level topological influences on the learning process, we stress that this account is by no means incompatible with existing statistical learning accounts. Exactly as early segmentation tasks demonstrated that conditional probabilities were sufficient to drive word extraction (for further discussion of this point see [4]), the present task demonstrates that community structure is also sufficient to drive event segmentation. In real-world situations, learners likely exploit a combination of the local and global-level cues that are available to them.

Statistical learning and complex network approaches are therefore compatible, but the marriage of these two disciplines has the potential to offer additional insight into the acquisition of large-scale structural knowledge. As they are currently implemented, statistical learning approaches provide the building blocks of graph structure, while complex systems approaches can reveal higher-order relational patterns over and above those captured by co-occurrence statistics. We suggest here that the types of local relationships typically manipulated in statistical learning paradigms (e.g., syllable, word, or phrase co-occurrence) are not only correlated with particular micro-level topological properties of complex networks (i.e., the degree of a target node) but themselves underpin macro-level properties such as community structure (Figure 3). For example, in quantitative linguistics, co-occurrence networks generated from temporal adjacency between words in a

Input

…TU–DA–RO–BI–KU–TI–PI–GO–LA–
BU–DO–PA–BI–KU–TI–TU–DA–RO–
PI–GO–LA–TU–DA–RO–BU–DO–PA–
PI–GO–LA–BI–KU–TI–BU–DO–PA…

Pairwise statistical information
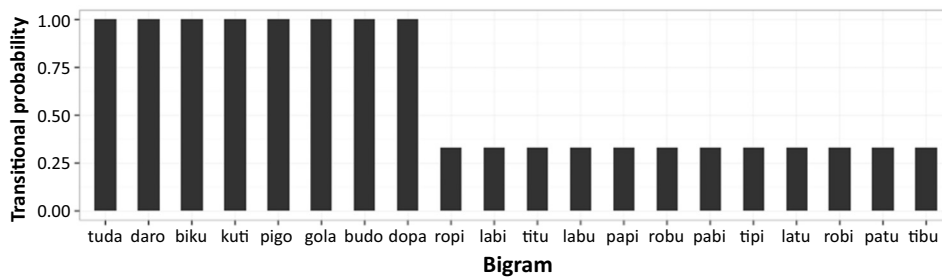
Community structure
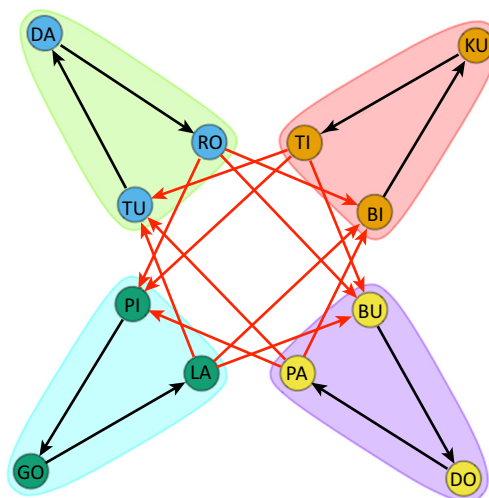
Trends in Cognitive Sciences

Figure 3. Pairwise Statistics Underpin Network Topology. When four pseudowords (*tudaro*, *bikuti*, *pigola*, *budopa*) are concatenated together to form a continuous stream of syllables, evidence from Saffran *et al.* (1996) indicates that these words can be segmented via the dip in transitional probabilities at word boundaries [1]. We show here that the co-occurrence between syllables can also be used to construct a weighted graph (black lines indicate a high bigram frequency and red lines indicate a low bigram frequency). A community detection algorithm consisting of a series of short random walks through this graph [108] will then reveal robust cluster structure corresponding to each word in the stream (shown in green, pink, purple, and blue).

given context have been used to probe broader topological properties of semantic systems (e.g., [85]). Moreover, because the edges of a network are interpreted as any relationship between two nodes, it is possible to build probabilistic information into the weights of an edge when modeling the graph structure of a system. While the densely clustered graph structure discussed previously [84] involved edges of equal weights (uniform transitional probabilities), a random walk on a weighted graph would result in the traversal of some edges more than others , potentially leading to novel segmentation patterns. Thus, probability of co-occurrence can be used to determine the regularity

with which graph edges are traversed (e.g., as a temporal sequence unfolds). Indeed, an important line of future work might be to explore how learning is influenced by topological properties of a complex network in addition to the specific sequence in which its edges are traversed. Preliminary evidence suggests that the impact of general topological properties outweighs that of algorithms dictating the order in which its constituent elements are revealed or retrieved [86].

Statistical learning researchers are already tackling new questions with concepts closely tied to network science: how does the sparseness or density of input influence generalization [87]? What latent structures support acquisition [88]? A more formal integration of network science methods with current statistical learning approaches will allow current experimental methods to be 'scaled up' to unprecedented levels (see criticisms in [8]). While mounting evidence indicates that statistical learning mechanisms act on larger lexicons (and that these learning effects can persist over years [89]), increased communication between statistical learning and complex network science will likely offer additional insight into the influence of macro-level topological properties on this process. Network topology might also have far-reaching impact on cognitive capacities that support or influence learning (e.g., attention, working memory, general intelligence; see [76]). For a more detailed treatment of how network-based approaches might be applied to the study of learning, refer to the Outstanding Questions.
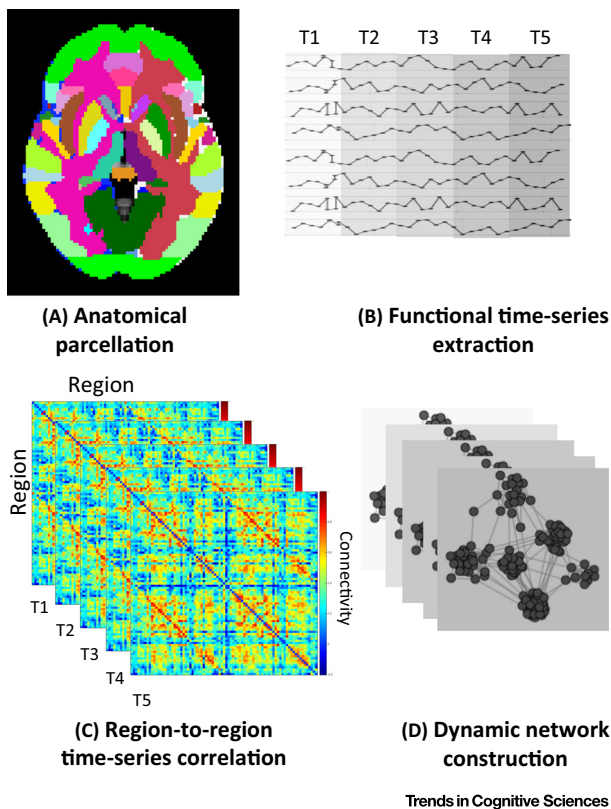
The Human Brain Is a Dynamical Complex System

As with all efforts to understand a cognitive process, probing how that process is implemented in the human brain must be considered in addition to patterns of human behavior. Particularly as related to linguistic functions such as processing and production, this integrative approach between brain and behavior has been applied with success [90]. With the advent of functional magnetic resonance imaging (fMRI), similar strides have been made in increasing our understanding of the neural regions recruited during statistical learning (e.g., in segmentation tasks [91–95]). However, this univariate approach is relatively coarse, revealing only areas that, on average, show increased neural activity above a significance threshold or related to a behavioral measure. To address this potential concern, we now survey one additional avenue of promising research: the study of the human brain itself as a dynamical complex system [19,21,23].

In functional brain networks, relationships arise from correlation or coherence in the neural response of pairs of neural areas (Figure 4). These relationships are taken to index how regions communicate with one another dynamically during a given task. Most relevant to the current review, advances have been made in the application of network-based approaches to uncovering functional neural mechanisms of the learning process [96,97]. For example, graph theoretical techniques have been applied to track changes in modular (community) organization as participants practiced a simple motor sequence [98]. In this case it was observed that modularity shifted dynamically over time, and also that it was possible to predict future behavior based on characteristics of network configuration in a given scanning session. Specifically, flexibility, defined by the number of changes in the connectivity of nodes to higher-level communities, was a significant predictor of reaction time speed-up. Especially as indexed in frontal areas, flexibility has also been linked to performance on various memory tasks, such as those involving short-term contextual recollection [99] and working memory components [100]. Related work has demonstrated that the separability of a richly connected 'core' of primary sensorimotor regions and a sparsely connected, flexible 'periphery' of association regions increases throughout learning, a measure also found to predict individual differences in performance on a motor learning task [101]. Ongoing work examines other dynamical properties that may drive learning (e.g., network centrality [102] and the autonomy of sensorimotor associations [103]).

As reviewed throughout this manuscript, a key line of research in cognitive science centers on the learnability of network structures that emerge from external input (e.g., via the topological

(A) Anatomical parcellation

(B) Functional time-series extraction

(C) Region-to-region time-series correlation

(D) Dynamic network construction

Trends in Cognitive Sciences

Figure 4. Evolution of Complex Network Structure in the Brain. To investigate topological properties of task-based functional connectivity, the brain is first parcellated into anatomically defined nodes, in this example using the Harvard-Oxford structural atlas (A). Next, the moment-to-moment activity within each of these regions is extracted at different points during a scanning session (B). To construct the edges between brain regions (nodes of the network), the correlation or coherence between any two time-series is computed, forming a pair-wise adjacency matrix at each timepoint (C). These matrices can then be used to probe the temporal dynamics of a functional brain network over the course of a given task (D).

features in language). We propose here that an equally important line of work is to examine how internal complex system dynamics gives rise to learning, and indeed how acquired knowledge might be reflected in observable topological patterns in the human brain. While most complex network approaches to brain connectivity begin with brain parcellation according to cytoarchitectonic or other anatomical division (Figure 4), some approaches have examined network structure via voxel-to-voxel associations (i.e., where individual voxels correspond to a network node [104–107]). In contrast to whole-brain connectivity approaches, which focus on the coordination of large-scale cognitive systems, voxel-based analysis techniques introduce the possibility of examining network-based representations as they develop. We submit that how learners translate network architectures in sensory input to network-based neural representations remains an essential but open question. More generally, we propose that the field of cognitive science stands to benefit enormously from increased incorporation of network-based approaches, specifically because the adoption of these methods offers a hitherto absent framework upon which to unify behavioral, computational, and neuroscientific studies of learning.

Concluding Remarks

Learners attain a complex and highly structured representation of the world. Currently, many quantitative approaches to learning hinge on sensitivity to local statistics such as co-occurrence frequencies and transitional probabilities between adjacent elements. While local statistics are clearly one salient source of structural information, evidence reviewed here suggests that learners also perceive global organizational patterns. In fact, exciting new results suggest that learners can acquire knowledge of these large-scale patterns even when local statistics are relatively uninformative. Thus, as we deepen our understanding of the natural world as a complex system, we can harness descriptions of its large-scale organizational properties to

then scale-up our approach to the study of learning. Moreover, exactly as we seek to understand how humans extract both local statistics and global architectural information, it is also useful to examine how micro- and macro-level brain dynamics support learning. In particular, an exciting avenue of future research will center on understanding how differences in network architecture of naturalistic stimuli lead to differences in learnability via unique neurophysiological mechanisms.

References

1. Saffran, J.R. *et al.* (1996) Statistical learning by 8-month-old infants. *Science* 274, 1926–1928
2. Romberg, A.R. and Saffran, J.R. (2010) Statistical learning and language acquisition. *Wiley Interdiscip. Rev. Cogn. Sci.* 1, 906–914
3. Aslin, R.N. and Newport, E.L. (2012) Statistical learning: from acquiring specific items to forming general rules. *Curr. Dir. Psychol. Sci.* 21, 170–176
4. Aslin, R.N. and Newport, E.L. (2014) Distributional language learning: mechanisms and models of category formation. *Lang. Learn.* 64, 86–105
5. Schapiro, A. and Turk-Browne, N.B. (2015) Statistical learning. In *In Brain Mapping: An Encyclopedic Reference* (Toga, A.W., ed.), pp. 501–506, Elsevier
6. Saffran, J.R. *et al.* (2007) The infant's auditory world: hearing, speech, and the beginnings of language. In *In Handbook of Child Psychology (Vol. 6)* (Damon, W. and Lerner, R.M., eds), pp. 58–108, John Wiley & Sons
7. Saffran, J.R. (2001) The use of predictive dependencies in language learning. *J. Mem. Lang.* 44, 493–515
8. Johnson, E.K. and Tyler, M.D. (2010) Testing the limits of statistical learning for word segmentation. *Dev. Sci.* 13, 339–345
9. Pelucchi, B. *et al.* (2009) Statistical learning in a natural language by 8-month-old infants. *Child Dev.* 80, 674–685
10. Estes, K.G. and Lew-Williams, C. (2015) Listening through voices: infant statistical word segmentation across multiple speakers. *Dev. Psychol.* 51, 1517–1528
11. Nicolis, G. and Nicolis, C. (2012) *Foundations of Complex Systems,* World Scientific
12. Amancio, D.R. *et al.* (2011) On the concepts of complex networks to quantify the difficulty in finding the way out of labyrinths. *Phys. A Stat. Mech. its Appl.* 390, 4673–4683
13. Chrastil, E.R. and Warren, W.H. (2015) Active and passive spatial learning in human navigation: acquisition of graph knowledge. *J. Exp. Psychol. Learn. Mem. Cogn.* 41, 1162–1178
14. Engelthaler, T. and Hills, T.T. (2016) Feature biases in early word learning: network distinctiveness predicts age of acquisition. *Cogn. Sci.* Published online February 29, 2016. http://dx.doi.org/10.1111/cogs.12350
15. Steyvers, M. and Tenenbaum, J.B. (2005) The large-scale structure of semantic networks: statistical analyses and a model of semantic growth. *Cogn. Sci.* 29, 41–78
16. Borge-Holthoefer, J. and Arenas, A. (2010) Semantic networks: Structure and dynamics. *Entropy* 12, 1264–1302
17. Goñi, J. *et al.* (2010) Switcher-random-walks: a cognitive-inspired mechanism for network exploration. *Int. J. Bifurc. Chaos* 20, 913–922
18. Ferrer i Cancho, R. *et al.* (2004) Patterns in syntactic dependency networks. *Phys. Rev. E* 69, 051915
19. Bullmore, E. and Sporns, O. (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198
20. Siegelmann, H.T. (2010) Complex systems science and brain dynamics: a frontiers in computational neuroscience special topic. *Front. Comput. Neurosci.* 4, 7
21. Sporns, O. (2013) Structure and function of complex brain networks. *Dialogues Clin. Neurosci.* 15, 247–262
22. Papo, D. *et al.* (2014) Complex network theory and the brain. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 369, 47–97
23. Goldenberg, D. and Galván, A. (2015) The use of functional and effective connectivity techniques to understand the developing brain. *Dev. Cogn. Neurosci.* 12, 155–164
24. Medaglia, J.D. *et al.* (2015) Cognitive network neuroscience. *J. Cogn. Neurosci.* 27, 1471–1491
25. Sporns, O. and Betzel, R.F. (2016) Modular brain networks. *Annu. Rev. Psychol.* 67, 613–640
26. Strogatz, S.H. (2001) Exploring complex networks. *Nature* 410, 268–276
27. Butts, C.T. *et al.* (2009) Revisiting the foundations of network analysis. *Science* 325, 414–416
28. Backes, A.R. *et al.* (2009) A complex network-based approach for boundary shape analysis. *Pattern Recognit.* 42, 54–67
29. Backes, A.R. and Bruno, O.M. (2010) Shape classification using complex network and multi-scale fractal dimension. *Pattern Recognit.* 31, 44–51
30. Backes, A.R. *et al.* (2013) Texture analysis and classification: 1 complex network-based approach. *Inf. Sci.* 219, 168–180
31. Wang, Z. *et al.* (2013) Impact of social punishment on cooperative behavior in complex networks. *Sci. Rep.* 3, 3055
32. Zhang, F. *et al.* (2015) Community detection based on links and node features in social networks. In *MultiMedia Modeling: 21st International Conference, Part I* (He, T. *et al.*, eds), pp. 418–429, Springer International Publishing
33. Cong, J. and Liu, H. (2014) Approaching human language with complex networks. *Phys. Life Rev.* 11, 598–618
34. Vitevitch, M.S. and Castro, N. (2015) Using network science in the language sciences and clinic. *Int. J. Speech. Lang. Pathol.* 17, 13–25
35. Beckage, N.M. and Colunga, E. (2015) Language networks as models of cognition: understanding cognition through language. In *In Towards a Theoretical Framework for Analyzing Complex Linguistic Networks* (Mehler, A. *et al.*, eds), pp. 3–28, Springer
36. Nastase, V. *et al.* (2015) A survey of graphs in natural language processing. *Nat. Lang. Eng.* 21, 665–698
37. Arbesman, S. *et al.* (2010) The structure of phonological networks across multiple languages. *Int. J. Bifurc. Chaos* 20, 679–685
38. Siew, C.S.Q. (2013) Community structure in the phonological network. *Front. Psychol.* 4, 553
39. Gravino, p. *et al.* (2012) Complex structures and semantics in free word association. *Adv. Complex Syst.* 15, 1250054
40. Morais, A.S. *et al.* (2013) Mapping the structure of semantic memory. *Cogn. Sci.* 37, 125–145
41. Utsumi, A. (2015) A complex Network approach to distributional semantic models. *PLoS One* 10, e0136277
42. Corominas-Murtra, B. *et al.* (2009) The ontogeny of scale-free syntax networks: phase transitions in early language acquisition. *Adv. Complex Syst.* 12, 371–392

43. Cech, R. and Macutek, J. (2009) Word form and lemma syntactic dependency networks in Czech: a comparative study. *Glotto-metrics* 19, 85–98

44. Luce, P.A. and Pisoni, D.B. (1998) Recognizing spoken words: the neighborhood activation model. *Ear Hear.* 19, 1–36

45. Ke, J. and Yao, Y. (2008) Analysing language development from a Network approach. *J. Quant. Linguist.* 15, 70–99

46. Nematzadeh, A. *et al.* (2014) A cognitive model of semantic network learning. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 244–254

47. Hills, T.T. *et al.* (2009) Longitudinal analysis of early semantic networks: preferential attachment or preferential acquisition? *Psychol. Sci.* 20, 729–739

48. Goldstein, R. and Vitevitch, M.S. (2014) The influence of cluster-ing coefficient on word-learning: how groups of similar sounding words facilitate acquisition. *Front. Psychol.* 5, 1307

49. Vitevitch, M.S. *et al.* (2012) Complex network structure influences processing in long-term and short-term memory. *J. Mem. Lang.* 67, 30–44

50. Chan, K.Y. and Vitevitch, M.S. (2009) The influence of the phono-logical neighborhood clustering coefficient on spoken word recog-nition. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1934–1949

51. Yates, M. (2013) How the clustering of phonological neighbors affects visual word recognition. *J. Exp. Psychol. Learn. Mem. Cogn.* 39, 1649–1656

52. Chan, K.Y. and Vitevitch, M.S. (2010) Network structure influ-ences speech production. *Cogn. Sci.* 34, 685–697

53. Carlson, M.T. *et al.* (2014) How children explore the phonological network in child-directed speech: a survival analysis of children's first word productions. *J. Mem. Lang.* 75, 159–180

54. Vitevitch, M.S. *et al.* (2014) Insights into failed lexical retrieval from network science. *Cogn. Psychol.* 68, 1–32

55. Siew, C.S.Q. and Vitevitch, M.S. (2015) Spoken word recognition and serial recall of words from components in the phonological network. *J. Exp. Psychol. Learn. Mem. Cogn.* 42, 394–410

56. Kenett, Y.N. *et al.* (2013) Semantic organization in children with cochlear implants: computational analysis of verbal fluency. *Front. Psychol.* 4, 543

57. Beckage, N. *et al.* (2011) Small worlds and semantic network growth in typical and late talkers. *PLoS One* 6, e19348

58. Kenett, Y.N. *et al.* (2014) Investigating the structure of semantic networks in low and high creative persons. *Front. Hum. Neuro-sci.* 8, 407

59. Pavlov, I.P. (1927) *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*, Oxford University Press

60. Arnon, I. and Snider, N. (2010) More than words: frequency effects for multi-word phrases. *J. Mem. Lang.* 62, 67–82

61. Perruchet, P. and Pacton, S. (2006) Implicit learning and statisti-cal learning: one phenomenon, two approaches. *Trends Cogn. Sci.* 10, 233–238

62. Maye, J. *et al.* (2002) Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82, B101–B111

63. Liu, R. and Holt, L.L. (2015) Dimension-based statistical learning of vowels. *J. Exp. Psychol. Hum. Percept. Perform.* 41, 1783–1798

64. Gómez, R.L. (2002) Variability and detection of invariant struc-ture. *Psychol. Sci.* 13, 431–436

65. Misyak, J.B. *et al.* (2010) Sequential expectations: the role of prediction-based learning in language. *Top. Cogn. Sci.* 2, 138–153

66. Thompson, S.P. and Newport, E.L. (2007) Statistical learning of syntax: the role of transitional probability. *Lang. Learn. Dev.* 3, 1–42

67. Levy, R. (2008) Expectation-based syntactic comprehension. *Cognition* 106, 1126–1177

68. MacDonald, M.C. (2013) How language production shapes lan-guage form and comprehension. *Front. Psychol.* 4, 226

69. Fine, A.B. *et al.* (2013) Rapid expectation adaptation during syntactic comprehension. *PLoS One* 8, e77661

70. Conway, C.M. *et al.* (2011) Implicit statistical learning in language processing: word predictability is the key. *Cognition* 114, 356–371

71. Karuza, E.A. *et al.* (2013) On-line measures of prediction in a self-paced statistical learning task. In *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (Bello, P. *et al.*, eds), pp. 725–730, Cognitive Science Society

72. Creel, S.C. *et al.* (2004) Distant melodies: statistical learning of nonadjacent dependencies in tone sequences. *J. Exp. Psychol. Learn. Mem. Cogn.* 30, 1119–1130

73. Fiser, J. and Aslin, R.N. (2002) Statistical learning of higher-order temporal structure from visual shape sequences. *J. Exp. Psy-chol. Learn. Mem. Cogn.* 28, 458–467

74. Kirkham, N.Z. *et al.* (2002) Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition* 83, B35–B42

75. Fiser, J. and Aslin, R.N. (2001) Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychol. Sci.* 12, 499–504

76. Siegelman, N. and Frost, R. (2015) Statistical learning as an individual ability: theoretical perspectives and empirical evidence. *J. Mem. Lang.* 81, 105–120

77. Frost, R. *et al.* (2015) Domain generality versus modality speci-ficity: the paradox of statistical learning. *Trends Cogn. Sci.* 19, 117–125

78. Gebhart, A.L. *et al.* (2009) Changing structures in midstream: learning along the statistical garden path. *Cogn. Sci.* 33, 1087–1116

79. Franco, A. and Destrebecqz, A. (2012) Chunking or not chunk-ing? How do we find words in artificial language learning? *Adv. Cogn. Psychol.* 8, 144–154

80. Emberson, L.L. and Rubinstein, D.Y. (2016) Statistical learning is constrained to less abstract patterns in complex sensory input (but not the least). *Cognition* 153, 63–78

81. Turk-Browne, N.B. *et al.* (2005) The automaticity of visual statis-tical learning. *J. Exp. Psychol. Gen.* 134, 552–564

82. Musz, E. *et al.* (2015) Visual statistical learning is not reliably modulated by selective attention to isolated events. *Atten. Per-cept. Psychophys.* 77, 78–96

83. Friston, K. (2008) Hierarchical models in the brain. *PLoS Comput. Biol.* 4, e1000211

84. Schapiro, A.C. *et al.* (2013) Neural representations of events arise from temporal community structure. *Nat. Neurosci.* 16, 486–492

85. Howard, M.W. *et al.* (2011) Constructing semantic representa-tions from a gradually-changing representation of temporal con-text. *Top. Cogn. Sci.* 3, 48–73

86. Abbott, J. *et al.* (2015) Random walks on Semantic networks can resemble optimal foraging. *Psych. Review.* 122, 558–569

87. Reeder, P.A. *et al.* (2013) From shared contexts to syntactic categories: the role of distributional information in learning lin-guistic form-classes. *Cogn. Psychol.* 66, 30–54

88. Ellis, N.C. *et al.* (2013) Usage-based language: investigating the latent structures that underpin acquisition. *Lang. Learn.* 63, 25–51

89. Frank, M.C. *et al.* (2013) Learning and long-term retention of large-scale artificial languages. *PLoS One* 8, e52500

90. Friederici, A.D. and Singer, W. (2015) Grounding language proc-essing on basic neurophysiological principles. *Trends Cogn. Sci.* 19, 329–338

91. McNealy, K. *et al.* (2006) Cracking the language code: neural mechanisms underlying speech parsing. *J. Neurosci.* 26, 7629–7639

92. Turk-Browne, N.B. *et al.* (2009) Neural evidence of statistical learning: efficient detection of visual regularities without aware-ness. *J. Cogn. Neurosci.* 21, 1934–1945

93. Cunillera, T. *et al.* (2009) Time course and functional neuroanat-omy of speech segmentation in adults. *Neuroimage* 48, 541–553

94. Turk-Browne, N.B. *et al.* (2010) Implicit perceptual anticipation triggered by statistical learning. *J. Neurosci.* 30, 11177–11187

95. Karuza, E.A. *et al.* (2013) The neural correlates of statistical learning in a word segmentation task: an fMRI study. *Brain Lang.* 127, 46–54

96. Hampshire, A. *et al.* (2016) Network mechanisms of intentional learning. *Neuroimage* 127, 123–134

97. Parkin, B.L. *et al.* (2015) Dynamic network mechanisms of relational integration. *J. Neurosci.* 35, 7660–7673

98. Bassett, D.S. *et al.* (2011) Dynamic reconfiguration of human brain networks during learning. *Proc. Natl. Acad. Sci. U.S.A.* 108, 7641–7646

99. Fornito, A. *et al.* (2012) Competitive and cooperative dynamics of large-scale brain functional networks supporting recollection. *Proc. Natl. Acad. Sci.* 109, 12788–12793

100. Braun, U. *et al.* (2015) Dynamic reconfiguration of frontal brain networks during executive cognition in humans. *Proc. Natl. Acad. Sci. U.S.A.* 112, 11678–11683

101. Bassett, D.S. *et al.* (2013) Task-based core-periphery organization of human brain dynamics. *PLoS Comput. Biol.* 9, e1003171

102. Mantzaris, A.V. *et al.* (2013) Dynamic network centrality summarizes learning in the human brain. *J. Complex Networks* 1, 83–92

103. Bassett, D.S. *et al.* (2015) Learning-induced autonomy of sensorimotor systems. *Nat. Neurosci.* 18, 744–751

104. de Reus, M.A. and van den Heuvel, M.P. (2013) The parcellation-based connectome: limitations and extensions. *Neuroimage* 80, 397–404

105. Schwarz, A.J. *et al.* (2012) Voxel scale complex networks of functional connectivity in the rat brain: neurochemical state dependence of global and local topological properties. *Comput. Math. Methods Med.* 2012, 1–15

106. Hayasaka, S. and Laurienti, P.J. (2010) Comparison of characteristics between region-and voxel-based network analyses in resting-state fMRI data. *Neuroimage* 50, 499–508

107. Yu, Q. *et al.* (2011) Altered small-world brain networks in temporal lobe in patients with schizophrenia performing an auditory oddball task. *Front. Syst. Neurosci.* 5, 7

108. Pons, P. and Latapy, M. (2005) Computing communities in large networks using random walks. *Lect Notes Comput Sc* 3733, 284–293