ANNUAL REVIEWS

Annual Review of Psychology

The Neurocognition of Developmental Disorders of Language

Michael T. Ullman,[1] F. Sayako Earle,[2] Matthew Walenski,[3] and Karolina Janacsek[4,5]

[1]Brain and Language Lab, Department of Neuroscience, Georgetown University, Washington, DC 20057, USA; email: michael@georgetown.edu

[2]Department of Communication Sciences and Disorders, University of Delaware, Newark, Delaware 19713, USA

[3]Department of Communication Sciences and Disorders, Northwestern University, Evanston, Illinois 60208, USA

[4]Institute of Psychology, Eotvos Lorand University (ELTE), H-1071 Budapest, Hungary

[5]Brain, Memory, and Language Lab; Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, H-1117 Budapest, Hungary

Annu. Rev. Psychol. 2020. 71:389–417

First published as a Review in Advance on July 23, 2019

The *Annual Review of Psychology* is online at psych.annualreviews.org

https://doi.org/10.1146/annurev-psych-122216-011555

Keywords

developmental language disorder, dyslexia, articulation disorder, childhood apraxia of speech, stuttering, specific language impairment, procedural circuit deficit hypothesis, procedural memory, basal ganglia, declarative memory

Abstract

Developmental disorders of language include developmental language disorder, dyslexia, and motor-speech disorders such as articulation disorder and stuttering. These disorders have generally been explained by accounts that focus on their behavioral rather than neural characteristics; their processing rather than learning impairments; and each disorder separately rather than together, despite their commonalities and comorbidities. Here we update and review a unifying neurocognitive account—the Procedural circuit Deficit Hypothesis (PDH). The PDH posits that abnormalities of brain structures underlying procedural memory (learning and memory that rely on the basal ganglia and associated circuitry) can explain numerous brain and behavioral characteristics across learning and processing, in multiple disorders, including both commonalities and differences. We describe procedural memory, examine its role in various aspects of language,

and then present the PDH and relevant evidence across language-related disorders. The PDH has substantial explanatory power, and both basic research and translational implications.

Contents

INTRODUCTION

Developmental disorders can impact various aspects of language. Developmental language disorder (DLD) broadly refers to childhood language problems not explained by factors such as hearing deficits or environmental deprivation (Am. Psychiatr. Assoc. 2013, Bishop et al. 2017). Developmental dyslexia, by contrast, affects reading (Goswami 2015), whereas developmental motor-speech disorders such as articulation disorder, childhood apraxia of speech (verbal dyspraxia), and developmental stuttering impair the production of fluent speech (Alm 2004, Bernthal et al. 2009). These disorders are not uncommon (e.g., 7% prevalence for DLD) and can persist into adulthood, with important social and economic consequences (Elbro et al. 2011, Leonard 2014).

Various accounts have been proposed to explain these disorders. DLD (formerly often referred to as specific language impairment) has generally been attributed either to a processing deficit (capacity limitations on processing, or problems with working memory, phonological processing, or temporal processing) or a specific linguistic deficit, especially of grammar, which tends to be strongly affected in DLD (Leonard 2014, Ullman & Pierpont 2005). Dyslexia has often been accounted for by deficits related to phonology or magnocellular circuitry (Ramus et al. 2003). Motor-speech disorders have generally been explained by impairments in the motor production of speech, with neural accounts for stuttering (Alm 2004, Bernthal et al. 2009).

Though these accounts have some explanatory power (Bernthal et al. 2009, Leonard 2014, Ramus et al. 2003), important gaps remain. First, many accounts are essentially functional, in that

they ascribe the problem to a particular function (e.g., working memory). Yet, these disorders must ultimately be explained by neurobiological disruptions. Thus, neurocognitive accounts that unify neurobiological and cognitive aspects should have more explanatory power than do purely functional (or purely neurobiological) accounts. Second, because these are developmental disorders, impairments in learning may play important roles, perhaps in addition to processing deficits—yet most of the accounts listed above focus on processing rather than learning. Third, most accounts focus on a particular disorder. However, many of the disorders' symptoms overlap, and the disorders are often comorbid with one another (Bernthal et al. 2009, Leonard 2014). Thus, accounts that unify these disorders may be better positioned to advance the field.

The procedural deficit hypothesis (PDH) attempts to address these gaps (Ullman 2004, Ullman & Pierpont 2005). The PDH was originally presented in depth for DLD (Ullman & Pierpont 2005). According to the PDH, DLD may be largely explained by abnormalities of brain structures underlying procedural memory, which is defined as the learning and memory that relies on the basal ganglia (BG) and its associated circuitry. Although the cerebellum may interact with this circuitry, we do not focus on the cerebellum here (see Bostan & Strick 2018, Ullman & Pierpont 2005). Reflecting the anatomical basis of the PDH, we henceforth refer to it as the Procedural circuit Deficit Hypothesis. The PDH of DLD was originally motivated by the grammar impairments in the disorder, because grammar has been linked to procedural memory as part of the declarative/procedural model of language (Ullman 2004, 2016). According to the PDH, individuals with procedural circuit abnormalities should show deficits not only of grammar and procedural memory, but also of apparently nonprocedural functions such as working memory that also depend on this circuitry. However, the PDH posits that declarative memory, defined as the learning and memory that relies on the medial temporal lobe (MTL) and its associated circuitry, remains largely normal in DLD and can compensate for procedural circuit deficits. Thus, although the PDH was motivated by functional problems (of grammar), it is a unifying neurocognitive account in that it posits neuroanatomical abnormalities that can explain patterns of spared and impaired learning and processing across both language and nonlanguage domains.

It is also a unifying account in that it can explain multiple disorders. Although the PDH has been laid out in the greatest depth for DLD, it has also been proposed for dyslexia (Ullman 2004). Nicolson & Fawcett (2007) posited a related account for dyslexia, though they specifically implicated cerebellar circuits. The PDH has not previously been applied to motor-speech disorders, except in the case of the KE family, who are commonly diagnosed with verbal dyspraxia (Ullman & Pierpont 2005).

Here we provide an up-to-date exposition of the PDH, unifying neural and functional characteristics for both learning and processing across DLD, motor-speech disorders, and dyslexia. First, we describe procedural memory with greater specificity than in the past, integrating findings from several fields. Building on this understanding of procedural memory, we then examine the predicted roles for this system in typical development across multiple aspects of language learning and processing. These sections in turn provide the foundation for the PDH and its predictions, which we examine across different developmental disorders. Finally, we conclude with a brief discussion of implications and avenues for future investigation.

PROCEDURAL MEMORY: BASAL GANGLIA–BASED LEARNING

What Are the Basal Ganglia?

The BG in humans and other primates include the striatum (nucleus accumbens, caudate nucleus, and putamen), the globus pallidus (external and internal segments), the subthalamic nucleus, and both the substantia nigra (pars compacta and pars reticulata) and ventral tegmental area in the

midbrain (Hélie et al. 2015). The BG receive inputs from much of cortex (which mainly projects to the striatum) and provide outputs to frontal and other cortical regions, via the thalamus, as well as to subcortical structures. BG circuitry is organized into direct and indirect pathways that together (dis)inhibit frontal and other cortical activity, thereby underlying the selection of motor and other cortically based representations (Frank 2005). Dopamine plays a key role, with the midbrain structures projecting dopaminergic neurons to the striatum.

The circuitry running through the BG is organized into parallel circuits (partially topographically and functionally segregated loops), each of which receives somewhat different cortical inputs and projects to largely distinct frontal and other (e.g., temporal) cortical regions, which in turn feed back into their respective circuits (Middleton & Strick 1996). For example, somatosensory–(posterior) putamen–frontal motor circuitry underlies motor functions; posterior parietal–anterior dorsal striatum (anterior caudate/putamen)–prefrontal circuitry subserves cognition, such as working memory and executive functions; and hippocampal/amygdala–ventral striatum (nucleus accumbens and ventral caudate/putamen)–orbitofrontal/anterior cingulate circuitry supports motivational functions (Draganski et al. 2008, Seger et al. 2010).

How Do the Basal Ganglia Learn?

BG-based learning involves generating predictions about associations and then evaluating these predictions on the basis of the outcome (i.e., information about the correctness of the prediction) (Balleine & O'Doherty 2010). In particular, learning in the BG occurs when such (dis)confirmatory information is promptly available after the prediction is generated. For example, the BG are involved when a rat learns to predict that when it sees a lever (stimulus) it should press the lever (response), with the goal of receiving an immediately subsequent food reward (signaling the prediction was correct). Similarly, in categorization tasks such as the weather prediction task, in which human participants select a (probabilistically) predictable output stimulus (e.g., rain) based on an input stimulus (particular cards), rapid (dis)confirmation (feedback) of the predicted association leads to BG-based learning; by contrast, slow feedback, or none at all, leads to learning in the MTL (Foerde & Shohamy 2011a, Poldrack et al. 2001). Knowledge learned in the BG appears to be implicit (not available to conscious awareness) (Ashby et al. 2007, Janacsek & Nemeth 2012).

Both the input and output elements of an association learned in the BG can be either a stimulus or a response. Thus, one can learn not only stimulus-response and stimulus-stimulus pairs, as in the examples above, but also response-response pairs (e.g., after sneezing, one should apologize) and response-stimulus pairs (e.g., if I smile, you will smile) (Balleine & O'Doherty 2010, Foerde & Shohamy 2011b). In both animal and human experimental paradigms, the feedback is generally separate from the association (as above), but it does not need to be: When predicting that a given stimulus or response should lead to a subsequent stimulus, the mere occurrence of the latter can serve as feedback (e.g., the next item in a sequence). More generally, the basic selection function of the BG (see above) may be thought of as selecting the predicted output, with learning occurring when information indicates that this selection was incorrect (Frank 2005).

Indeed, BG-based learning ensues from generating incorrect predictions (if a prediction is correct, no learning is needed). Unpredicted outcomes (e.g., the rat failed to predict that pressing the lever will lead to a food reward) create/strengthen the input-output associations that should have been predicted, whereas incorrect predictions weaken them, via increases or decreases, respectively, of dopamine in the striatum (Frank 2005). Crucially, such learning occurs if the prediction is incorrect not only about the outcome itself but also when it will occur (e.g., if we incorrectly predict the next item of a sequence will occur immediately). Thus, the BG learn to predict not only what should occur, but also when that happens.

Because prediction errors during learning are larger earlier in the process, more learning occurs earlier. Additionally, the more predictable the output is, given the input, the faster the learning. Thus, more predictable (higher probability) pairs are typically learned faster than less predictable (lower probability) pairs, and completely deterministic pairs are learned the fastest (Delgado et al. 2005). BG-based learning is therefore gradual but varies (in part) according to the predictability (probability) of the association. Eventually, after extensive training, learned associations become not only very fast and accurate, but also inflexible, that is, less sensitive to prediction errors from feedback—characteristics of habitual/automatic behaviors (Ashby & Crossley 2012, Graybiel & Grafton 2015).

The functional neuroanatomy of BG-based learning may reflect distinctions between earlier and later phases of learning. Anterior portions of the striatum (ventral striatum and anterior caudate/putamen) seem to be more important for earlier phases, whereas posterior portions (posterior caudate/putamen) play a larger role in later phases. The reasons for this anterior/posterior striatal distinction for earlier/later learning remain unclear, but may be related to the different parallel circuits. Anterior striatal circuits may support aspects of motivation (linked to ventral striatal circuitry) (see above) as well as working memory and executive functions (linked to the anterior caudate/putamen), which may underlie early-stage prediction-feedback learning of associations. In contrast, posterior portions may underlie aspects of motor and/or visual learning (motor and visual circuits rely on more posterior putamen/caudate) that may take place during the fine-tuning of performance in later stages of acquisition, and they may even underlie the processing of automatized associations (Doyon et al. 2009, Seger et al. 2010).

Learning in the BG also critically involves cortex. In particular, BG-based learning underlies the creation/strengthening of corticocortical connections between cortical neuronal populations that project to the BG (striatum) and those to which the BG project (via the thalamus), while these input and output neuronal populations (representing the input and output elements of the association) are both active (Hélie et al. 2015). This simultaneous activation allows the formation/strengthening of these corticocortical connections via Hebbian learning (they fire together and thus wire together). An increased reliance on these connections (systems consolidation) is associated with increased automatization, though the analogous input-output circuitry passing through the BG may continue to be relied upon (Ashby & Crossley 2012, Doyon et al. 2009). Systems consolidation may promote generalization, for example, to new members of a category in category learning (see below). Consistent with BG inputs originating importantly in parietal regions, and the BG projecting critically to frontal regions, the corticocortical circuits formed by the BG may play a key role in dorsal stream pathways (Ullman 2004, 2016).

What Do the Basal Ganglia Learn?

Associations learned by the BG provide the building blocks for learning various simple and complex functions and behaviors, such as habits, skills, sequences, and categories (Ashby & Crossley 2012, Graybiel & Grafton 2015)—though these can also rely on declarative memory (see below). Here we focus on skills, sequences, and categories.

A skill can be defined as a serial combination of (perceptuo-)motor and/or cognitive computations, which together constitute a new capacity and with sufficient practice can become automatized (e.g., driving a car, playing an instrument) (Graybiel & Grafton 2015). Skill learning is often modeled with perceptuomotor sequence learning. In humans, this is generally tested with the implicit serial reaction time (SRT) task, in which participants respond to a sequence of stimuli using corresponding buttons, without being informed that a sequence exists (and generally with little or no explicit knowledge of the sequence after learning). Each stimulus typically occurs rapidly after

the previous stimulus or response. The sequence in this task can be deterministic and/or probabilistic, in that any given item(s) may be always followed by the same item(s) (deterministic) or by different item(s) with varying probabilities (probabilistic) (Janacsek & Nemeth 2012). Learning involves predicting a subsequent stimulus, which also serves as feedback; as learning proceeds, anticipation responses become more common (Vakil et al. 2017). Associations (dependencies) of adjacent elements seem to be learned faster than of nonadjacent (long-distance) elements, which in turn are learned faster if the dependency (number of nonpredictive items between the elements) is smaller (Remillard 2008). Moreover, a dependency that requires more preceding (predicting) elements is learned slower than one that requires fewer (Remillard 2008). Converging evidence (e.g., from lesion and neuroimaging studies) suggests the following functional neuroanatomy for perceptuomotor sequence learning in implicit SRT and related tasks (Doyon et al. 2009, Penhune & Steele 2012). Early stages of learning depend importantly on the anterior striatum, mainly the anterior caudate/putamen (but not the cerebellum, at least in implicit SRT; Hardwick et al. 2013). In contrast, in later stages, during and after automatization, posterior portions of the caudate and putamen are engaged, as well as cortical [mainly (pre)motor as well as parietal] regions, though again apparently not the cerebellum.

Sequence learning is not limited to perceptuomotor sequences. We also encounter purely perceptual (e.g., auditory, visual) sequences with no overt responses involved, such as when listening to music. Such sequences are often investigated in statistical learning paradigms that involve the sequential presentation of stimuli (e.g., tones, shapes) without responses, with knowledge typically tested in judgment tasks after learning (Durrant et al. 2012, Turk-Browne et al. 2009). As with perceptuomotor sequences, items in these sequences are normally presented rapidly, and sequences can be deterministic or probabilistic. Here we only discuss tasks with nonlinguistic stimuli (for those with language-related stimuli, including artificial grammars, see below). These paradigms appear to rely importantly on BG-based learning. Learning, which occurs gradually, appears to involve prediction (e.g., of the subsequent item), resulting in the acquisition of stimulus-stimulus associations; moreover, learning and knowledge of the sequences are generally implicit (Abla et al. 2008, Turk-Browne et al. 2009). Functional imaging studies, which have focused on early learning stages, have implicated the anterior caudate (head/body) as well as frontal (mainly premotor) and other cortical regions, though apparently not the cerebellum (e.g., Turk-Browne et al. 2009). Literatures other than statistical learning also link learning or processing of perceptual sequences to the BG. For example, evidence suggests that beat (rhythm) perception (temporal processing) and aspects of music processing rely on the BG, particularly the striatum (as well as premotor cortex and the cerebellum for music processing) (Grahn & Rowe 2009).

Category learning involves learning to group elements into categories. It is modeled with various paradigms, though most involve learning to categorize perceptual (usually visual) stimuli (Seger & Miller 2010). For example, in the weather prediction task, participants learn which visually presented cards are probabilistically associated with which visually presented weather outcome (rain or sun). The rules of categorization in such tasks can be deterministic or probabilistic. Category learning typically exhibits characteristics of BG-based learning. Learning is gradual, and is generally implicit in that the categorization rules are not told to the participants, who, moreover, do not usually acquire this knowledge explicitly (at least for more complex rules) (Seger & Miller 2010). BG-based category learning involves predicting a category (rain) given an input (the cards), with rapid feedback regarding the correctness of the prediction (Foerde & Shohamy 2011b). During learning, input element–category associations are acquired and then generalized according to relevant features (Seger & Miller 2010). Generalization, that is, the ability to successfully categorize new elements or combinations of elements, appears to take place gradually, and may depend on systems consolidation, in particular the formation of feature-based corticocortical

connections (between input element and/or category neuronal populations) (Ashby et al. 2007). After substantial practice, categorization measures (selecting the appropriate category based on the input) become automatized: Speed and accuracy reach asymptote, secondary tasks do not impair performance, and the categorization behavior becomes inflexible (difficult to modify) (Ashby & Crossley 2012, Hélie et al. 2010b). Converging evidence suggests that early phases of category learning depend especially on the caudate head and nucleus accumbens (and in some studies the caudate body and tail), whereas later phases mainly involve the (posterior) putamen; the cerebellum does not seem to be heavily involved (Seger & Miller 2010, Wilkinson et al. 2014). Cortex, particularly frontal regions, is also involved: prefrontal cortex (especially ventral lateral prefrontal cortex) during earlier phases and premotor cortex during later phases and after automatization (Hélie et al. 2010a, Waldschmidt & Ashby 2011).

Basal Ganglia– Versus Medial Temporal Lobe–Based Learning

Skills and knowledge that can be learned in the BG (i.e., in procedural memory) may (also) be learned in the MTL (i.e., in declarative memory). Here we summarize MTL learning and then discuss the factors that modulate BG- versus MTL-based learning. We do not discuss other types of learning, such as purely cortical learning (Seger & Miller 2010).

As mentioned above, we define declarative memory as MTL-based learning and memory (Davachi 2006; Eichenbaum 2012; Ullman 2004, 2016). Learning (encoding) initially relies on the hippocampus and other MTL structures, which link (neo)cortical representations via Hebbian learning. Unlike BG-based learning, MTL-based learning (e.g., observational learning) does not depend on feedback or prediction errors. However, analogous to BG-based learning, the knowledge learned in the MTL increasingly depends on cortex, via the formation of corticocortical connections through systems consolidation. This systems consolidation leads to generalization, as corticocortical connections represent features common to individual associations learned in the MTL. Sleep promotes this consolidation (though forgetting still occurs), apparently more reliably than for consolidation in BG-based learning (King et al. 2017). Longer-term retention seems to be worse for declarative than for procedural memory (Foerde et al. 2006). Declarative memory has traditionally been linked to the acquisition of events (episodic memory) and facts (semantic memory), though it is actually much more flexible: It appears to be able to learn many types of information, including both arbitrary information (e.g., items and idiosyncratic associations) and regularities, and both implicit and explicit knowledge. Indeed, declarative memory seems to be the only learning and memory system to underlie explicit knowledge, though it is one of many (including procedural memory) to subserve implicit knowledge (Ullman 2016). Learning can be very rapid in the MTL, occurring even after one exposure (one-shot learning), though repeated exposures strengthen the memory traces. Learned information can be retrieved via recollection or familiarity. MTL learning is flexible not only in that it can learn most information, but also (in contrast to automatized BG-based knowledge) in that this information can be used flexibly for new purposes. The MTL (especially the hippocampus) may be involved in prediction, possibly by using the current context to predict items likely to occur in that context (Henson & Gagnepain 2010). MTL-based learning is closely linked to working memory, which plays important roles in encoding and recall in declarative memory, possibly including the maintenance of different elements (e.g., those occurring at different times) to be associated (Ullman 2016). Not surprisingly, brain structures underlying working memory and related functions (e.g., recall and executive functions), such as portions of the BG and frontal cortex that underlie these functions (various parts of the striatum as well as ventrolateral and other prefrontal regions) also play (related) roles in declarative memory (Scimeca & Badre 2012, Van der Linden et al. 2000).

Various material-, learning-, and subject-related factors modulate the relative dependence of learning and/or processing in declarative versus procedural memory. Here we summarize these factors, though we emphasize that not all factors have been investigated for all types of knowledge learned in the BG (i.e., for perceptuomotor sequences, perceptual sequences, categories). Various aspects of BG/MTL modulation have been referred to as redundancy, competition, and compensation; we discuss all of these here (Poldrack & Packard 2003; Ullman 2015, 2016).

First, material-level factors can affect whether information is learned in declarative or procedural memory. Individual items (e.g., specific objects)—to the extent that an item can exist without associations—can be learned in the MTL, in particular in perirhinal cortex (Davachi 2006), but might not be supported by procedural memory. Associations, in contrast, can be learned through either circuit. However, it appears that unique associations (e.g., Clemi's dog just smelled that rotten squirrel) can be learned only in the MTL, in particular the hippocampus, consistent with the role for the hippocampus in episodic memory (Davachi 2006). The BG can also learn associations, but apparently only under particular circumstances: gradually, with many presentations, in contexts involving prediction (e.g., given a squirrel, will Clemi's dog smell it?), and with rapid feedback. Importantly, the regularities shared by many presented associations are learned, whereas any unique information in each association is lost. Thus, BG-based learning critically involves learning regularities (Hélie et al. 2015). Nevertheless, regularities (commonalities) across associations can also emerge from MTL-based learning, which underlies the formation of long-term semantic knowledge from multiple individual episodes containing that knowledge (e.g., learning that the capital of Burkina Faso is Ouagadougou from many episodes of learning that information). It remains unclear whether there is any kind of information or learning context that actually requires the BG for learning (i.e., that the MTL or other structures cannot learn), though it may be that learning (automatized) associations involving motor responses (whether in the input or output element of a learned pair) (Rose et al. 2011) or trial-and-error learning (Hélie et al. 2015) might necessitate the BG. Thus, it appears that MTL- and BG-based learning are somewhat but not completely redundant, in that both can learn associations, though not necessarily of the same sort in the same learning conditions.

Second, various learning-related factors modulate the relative reliance on the two circuits for material that could be learned in either. Providing explicit information (e.g., of a sequence) can push learning toward the MTL (Schendan et al. 2003). Similarly, encouraging attention to underlying patterns (e.g., telling participants there is an underlying category) can result in explicit awareness of these patterns, and working memory/executive function–dependent hypothesis-driven learning in the MTL (Ashby & Maddox 2011). In contrast, BG-based learning is promoted in implicit versions of the same tasks (i.e., without providing explicit information or encouraging attention)—even more so for learning under dual-task conditions, which reduce attentional resources that could be applied to the search for underlying patterns (Ashby & Maddox 2011, Hazeltine et al. 1997). Additionally, during learning, the MTL is often involved in and relied on during early stages (consistent with its more rapid learning abilities), whereas the BG appear to be involved throughout acquisition and are increasingly relied on over the course of practice (though BG lesions can revert dependence to knowledge learned in the MTL) (Packard 2008, Schendan et al. 2003). Thus, MTL-based representations can apparently inhibit (block) BG-based representations, and vice versa, depending on which is predominant. The two systems can therefore be thought of as being in competition. Additionally, during later stages of learning, after more attention has been paid to the task and stimuli, explicit awareness of underlying patterns can emerge, potentially (again) increasing reliance on declarative memory (Hazeltine et al. 1997). Finally, as mentioned above, whereas rapid feedback in tasks involving prediction seems to result

in BG-based learning (even if initially learning might also involve the MTL), slow feedback or no feedback seems to lead primarily to MTL-based learning.

Third, within- and between-subject-level factors can moderate the relative dependence on declarative versus procedural memory. Any subject-level factor that enhances or depresses learning or processing in one system relative to the other may shift reliance toward the more functionally available system. One such within-subject factor is development. Whereas during infancy, MTL learning, working memory, and executive functions do not appear to be well-developed, these improve during childhood and may plateau during young adulthood (Lavenex & Lavenex 2013, Welsh et al. 1991). In contrast, BG learning may be relatively well established early in childhood, but then attenuates around adolescence (Janacsek et al. 2012). Indeed, although sequence learning relies mainly on procedural memory in both children and adults, it appears to depend more on procedural memory in children than adults and more on declarative memory in adults than children (Thomas et al. 2004).

Between subjects, MTL and/or BG learning can be modulated by various factors, such as genotype (e.g., different alleles of the gene for BDNF differentially affect MTL learning, while those for DARPP-32 modulate BG learning) and sex (females show MTL learning advantages as compared to males, likely due in part to estrogen) (Ullman 2016). Despite possible predictions for the relative dependence of MTL versus BG learning for these factors, we are not aware of any studies clearly testing these or similar predictions for (nonlinguistic) sequence or category learning. However, more extreme cases of between-subject differences in MTL or BG functioning, leading to the modulation of BG versus MTL learning, can be found in individuals with disorders. For example, amnesic patients with MTL damage likely rely on the BG for sequence learning (though it takes them longer than it does healthy controls, consistent with early MTL learning in controls) (Curran 1997). In contrast, patients with BG dysfunction, such as those with Parkinson's disease, appear to rely on declarative memory to compensate for category and sequence learning deficits (Ullman & Pullman 2015). Finally, there seems to be a seesaw effect, such that BG dysfunction may lead to enhanced MTL functioning, and vice versa, perhaps due to the absence of inhibition from the other system (Ullman 2004).

PROCEDURAL MEMORY IN LANGUAGE

Grammar

Across different levels of language, linguistic elements are distributed sequentially and hierarchically into larger structures, following predictable patterns. Phonotactic rules constrain the combination of phonological segments (phonemes) into syllables, whereas grammatical rules constrain the combination of simple words and bound morphemes into complex words (e.g., meow + -ed; morphology) and phrases and sentences (e.g., the cat meowed; syntax). The distributional patterns of linguistic elements can be captured by (various kinds of) structural representations of language, including syntactic trees generated by phrase-structure rules (e.g., Chomsky 1995, Pollard & Sag 1994) and constructions in construction grammar (Goldberg 1995). These representations refer to abstract categories (e.g., noun, tense) in addition to individual words and morphemes. They capture both local (adjacent) dependencies between successive elements [e.g., between determiner (the) and noun (cat)] and long-distance dependencies, which can be arbitrarily far apart. Within the hierarchically structured sequence, the probability of particular subsequent (downstream) elements ranges from very low (e.g., noun phrase following noun phrase: The boy the girl kissed) to very high as well as deterministic (e.g., -ed following the verb stem for regular past tense).

Grammar appears to rely heavily on procedural memory. This is suggested, first of all, by the observation that grammar learning shares key characteristics with learning in procedural memory.

Both involve learning sequential and categorical knowledge, which is largely implicit. Compared with local dependencies, long-distance dependencies are learned more slowly in both procedural memory (see above) and grammar (Guasti 2017). Like learning in procedural memory, grammar is learned gradually and becomes largely automatic, that is, rapid, inflexible (grammatical relations become more rigidly delineated), and with decreased variability in performance (Love et al. 2009). Because linguistic input is generally fast paced, successive elements (categories, words) can provide rapid feedback for internally generated predictions.

Direct evidence also links grammar learning to procedural memory. A meta-analysis of correlational studies found that individuals with better procedural learning abilities also show better learning or knowledge of grammar (but not of words) (Hamrick et al. 2018). This pattern was observed in children in their first language as well as in adults learning a second language—though only at later stages of second language learning when a greater dependence on knowledge learned via procedural memory is expected (at earlier stages, second language correlated with learning abilities in declarative memory). The correlation between grammar and procedural learning held across different languages and linguistic structures (including syntax and inflectional morphology) as well as for both sequence and category learning tasks that depend on procedural memory. Grammar learning has also been tied specifically to the BG. A functional imaging neuroanatomical meta-analysis of early phases of adult language learning revealed that grammar (but not lexical) learning was associated with activation in the anterior caudate/putamen (and not the cerebellum) (Tagarelli et al. 2019), which underlies early phases of learning sequences and categories in procedural memory (see above). Moreover, grammar learning that is predicted to rely particularly on procedural memory (e.g., with implicit/uninstructed training) showed anterior caudate/putamen involvement, whereas grammar learning predicted to rely especially on declarative memory (e.g., with explicit/instructed training) showed hippocampal involvement. Thus, consistent with MTL/BG learning of sequences and categories, evidence from correlational and neuroanatomical meta-analyses suggests that (*a*) early stages of grammar learning involve both systems but rely on declarative memory especially under conditions favoring learning in that system and (*b*) when grammatical knowledge in procedural memory has become predominant in later stages, grammar depends mainly on the procedural system. Finally, evidence from the lesion method also suggests that grammar learning relies on procedural memory. For example, grammar learning is impaired in patients with early-stage Huntington's disease (characterized by striatal degeneration), while word learning remains relatively unimpaired (De Diego-Balaguer et al. 2008).

Evidence suggests that procedural memory underlies not only grammar learning but also grammatical processing. In expressive language, the meaning of a message is transformed into a hierarchically structured sequence of words, morphemes, and syllables. In receptive language, linear sequences of syllables, morphemes, and words are translated into structured grammatical representations from which meaning is recovered. Here we focus on receptive grammatical processing, which not only is automatized (rapid, inflexible, and less variable) (see above) in native language but also involves prediction: In first (more than second) language, automatized comprehension of rapidly unfolding sentences seems to depend on predicting upcoming information, including word category or form (Kaan 2014). Given that automatization is linked to an increased reliance on corticocortical connections formed through BG-based learning, automatized grammatical processing should depend largely on cortical structures. Indeed, whereas grammar learning depends importantly on the BG (see above), grammatical processing in native speakers depends instead largely on cortex, as shown by converging evidence from neuroimaging, lesion (e.g., in patients with Broca's aphasia), and other methods (Ullman 2004, 2016). Indeed, a recent neuroanatomical meta-analysis of functional neuroimaging studies of syntactic processing in native speakers implicated cortical regions but not the BG (or the cerebellum) (Walenski et al. 2019). Frontal

regions, particularly premotor areas and Broca's region (inferior frontal gyrus), may play especially important roles (Ullman 2004, 2016; Walenski et al. 2019).

Just as sequences and categories learned in procedural memory can also be learned in declarative memory, evidence suggests that aspects of grammar can be learned in either system. However, the computational basis of learning and representing grammar seems to differ between the two systems. Procedural memory appears to subserve the learning of grammatical knowledge that underlies (real-time) combination (Ullman 2004, 2016), perhaps through the acquisition of (probabilistic) relations between categories or specific units, which allows for the real-time prediction of downstream elements. In contrast, declarative memory can memorize (structured) chunks, such as meowed or the cat, or even abstract structures (e.g., Noun Phrase Verb Phrase) as in construction grammar. Additionally, declarative memory is involved in learning explicit (or possibly implicit) rules (e.g., nouns come after determiners) and appears to support generalization across similar stored associations (e.g., walk-walked and talk-talked may generalize to balk-balked) (Ullman 2004, 2016).

Similar factors modulate the degree to which grammar and nonlinguistic sequences/categories rely on each of the two systems. At least for adult language learning, both procedural and declarative memory support grammar during early stages, though in different ways: Whereas explicit knowledge, chunks, and similarity-based grammatical knowledge rely more on declarative memory, implicit knowledge of grammatical regularities relies especially on procedural memory (Tagarelli et al. 2019). In contrast, at later stages grammar appears to depend only on procedural memory (Hamrick et al. 2018). Local dependencies are particularly likely to be chunked, especially for highly frequent combinations; even surface syntactic structures may be stored (Ullman 2016). Compared with boys and men, girls and women may depend more on declarative memory (e.g., by chunking and associative generalization) for aspects of grammar, particularly for local dependencies (Ullman 2016). This sex difference may be modulated by estrogen, higher levels of which are associated with a greater reliance on chunked forms (Ullman 2016). In disorders affecting procedural memory structures (e.g., Parkinson's disease, Broca's aphasia), there may be a compensatory reliance on chunking of local dependencies, which may moreover interact with sex (with females relying more on chunking) (Johari et al. 2019, Ullman & Pullman 2015).

Lexicon

Language also depends on the lexicon, which contains (at least) arbitrary information that cannot be generated by the grammar and thus must be stored. At minimum, this includes word forms (e.g., cat), their meanings, and form-meaning associations, as well as information about irregular morphological forms (e.g., the form dug and its link to dig), the number and form of a verb's complements (e.g., devour takes a single direct object), and larger arbitrary form-meaning mappings (e.g., idioms).

Although the declarative/procedural model has primarily linked the lexicon to declarative memory, some aspects of lexical learning and knowledge appear to rely on procedural memory. First, learning the phoneme sequences of word forms seems to depend partly on gradually and implicitly segmenting these sequences out of the speech stream (Singh et al. 2012). Such word segmentation may depend on the anterior caudate/putamen during early phases of learning, underscoring its reliance on procedural memory (Karuza et al. 2013). Second, words whose meanings involve knowledge of motor skills that were likely learned in procedural memory (e.g., words for tools or actions) are also linked to the BG (Walenski et al. 2007). Third, closed-class words and morphemes (e.g., auxiliaries, inflectional affixes such as -ed), which are not tightly bound to conceptual meanings but depend strongly on grammatical structure, are

more closely linked to the BG than open-class words such as nouns (Ullman 2004, 2016). Despite the apparent suitability of procedural memory for learning aspects of conceptual/semantic categories (e.g., allowing one to group individual cat exemplars into a semantic category for cat) or for learning form-meaning associations, we are not aware of any evidence along these lines.

Most evidence in fact ties lexical learning and knowledge to declarative memory (Ullman 2004, 2016). First, correlational evidence shows that children with better declarative (but not procedural) learning abilities also show better lexical abilities (Hamrick et al. 2018). Second, consolidation patterns of newly learned word forms and semantic features are consistent with systems consolidation following hippocampal learning (Schreiner & Rasch 2017). Third, amnesiacs with MTL lesions show word-learning impairments (Davis & Gaskell 2009). Fourth, neuroimaging studies have linked word learning to the hippocampus (Davis & Gaskell 2009). Moreover, a neuroanatomical meta-analysis of neuroimaging studies of adult language learning has tied word learning to ventral stream structures, which are closely linked to the MTL (Tagarelli et al. 2019). Finally, the N400 event-related potential component, which is found in response to various lexical/semantic manipulations, relies importantly on MTL structures, especially perirhinal cortex (Fernández & Tendolkar 2006).

Speech-Sound Representations

Speech-sound representations are distributions of acoustic-phonetic features that can function as linguistic units. They include contrastive (phonemic) and noncontrastive (phonetic) categories as well as talker-specific information. The characteristics of speech-sound learning in infancy resonate with BG-based category learning. Young children appear to acquire speech-sound categories implicitly, via gradual exposure to distributions of phonetic features. Although infants are perceptually flexible, they then become less sensitive to features that are noncontrastive in their native language (Zhang et al. 2005).

Speech-sound category learning has generally been investigated in adults. This research suggests that such learning initially relies on a reflective learning system tied to declarative memory structures, whereas during later stages it relies on a reflexive learning system tied to procedural memory structures (Yi et al. 2014). Converging evidence suggests that automatization and other optimal outcomes of speech-sound category learning are contingent on this switch from declarative to procedural memory (Yi et al. 2014). Factors that modulate the relative dependence of nonlinguistic categories on the two systems play analogous roles in speech-sound category learning. For example, explicit knowledge of the speech-sound category label or of the criteria of category membership increases reliance on declarative memory (Chandrasekaran et al. 2016). In contrast, learning relies more on procedural memory when rapid feedback is provided (Tricomi et al. 2006), which dramatically improves the rate of category learning (Chandrasekaran et al. 2014). In sum, while both memory systems appear to be engaged in speech-sound category learning, efficient and ultimately successful learning of speech-sound categories seems to depend on the extent of procedural memory involvement.

This work in adults leads to the prediction that native-language speech-sound category learning in infants relies even more on procedural memory. Unlike in many adult studies, infants rarely encounter speech sounds in isolation. Instead, they encounter them as elements of a sequence wherein successive elements provide rapid feedback. Moreover, early-life weaknesses of declarative memory relative to procedural memory suggest that infants rely particularly on the latter. Thus, procedural memory likely plays a predominant role in native-language speech-sound category learning.

Articulation and Speech Production

Articulatory sequences seem to be learned in procedural memory much like other perceptuomotor skills. Learning to produce articulatory sequences occurs largely implicitly, and evidence suggests a role for rapid feedback in babbling behavior (Goldstein & Schwade 2008), a precursor to articulation. Infants learn to articulate in the context of a rapidly changing vocal tract, necessitating relative flexibility in the motor program, whereas the motor program stabilizes once the articulatory apparatus has matured. Over the lifespan, the production of articulatory sequences (i.e., speech production) becomes automatized, as variability is reduced and production rate increases (Smith & Zelaznik 2004).

Direct evidence from noninfant populations also ties articulatory learning to the BG. Studies of patients with Parkinson's disease implicate the BG in adapting preexisting speech-motor sequences to auditory feedback, which by its nature occurs rapidly (Defne Abur et al. 2018). Furthermore, the role of the BG in adaptive vocal learning has been widely explored in songbirds, which show clear behavioral and neural parallels with human vocal learning, an aspect of learning to articulate (Doupe & Kuhl 1999). For example, *FoxP2* expression in an avian correlate of the BG (area X) is positively associated with imitative vocal learning in songbirds (Haesler et al. 2007).

Speech production also relies on procedural memory circuitry (as well as the cerebellum). First, the BG are implicated in the timing of articulatory gestures, including in predicting that they will lead to appropriately timed auditory and tactile feedback (Kotz & Schwartze 2010). To illustrate, temporal processing impairments in patients with Parkinson's disease appear to be linked to deficits in speech production, which improve when external timing cues are provided (Thaut et al. 2001). Second, the BG act in concert with frontal regions, both in selecting the speech-motor programs apparently stored in the vicinity of BA 44 and in initiating these programs via the supplementary motor area (SMA) (Tourville & Guenther 2011). Indeed, patients with cortical lesions to BA 44/6 often present with acquired apraxia of speech, a disorder of speech-motor programming (Duffy 2006). Thus, procedural circuitry underlies not only the learning of speech-motor programs, but also the timing, prediction, selection, and initiation of these programs during speech production.

Speech Perception

Speech perception (the process of mapping an auditory speech signal to meaning) involves constraining various sources of information, including cross-modal (visual-auditory-motor) information about the signal, various types of linguistic knowledge, and perceptual experience (McClelland et al. 2006). To accomplish this at the rapid rate of spoken language, speech perception relies on predicting the upcoming signal on the basis of current and prior information (Kutas et al. 2011). Thus, language knowledge, including procedural memory–based knowledge (e.g., grammar, speech-sound categories, articulatory skills), will impact speech perception (Holt & Lotto 2010). Furthermore, converging evidence links such prediction to BG-based temporal processing (Kotz & Schwartze 2010). Thus, both language knowledge and temporally sensitive prediction that depend on procedural memory are critical for efficient speech perception.

THE PROCEDURAL CIRCUIT DEFICIT HYPOTHESIS: PRINCIPLES AND PREDICTIONS

The PDH of any given disorder (e.g., DLD, dyslexia) posits that abnormalities of procedural memory brain structures, i.e., of cortico-BG-thalamocortical circuitry, can partly if not largely explain the disorder (Ullman 2004, Ullman & Pierpont 2005). That is, such abnormalities can account

for—and predict additional—core as well as secondary characteristics of the disorder in a substantial number of afflicted individuals. Such individuals may be described as having Developmental Procedural circuit Disorder (DPD). To be clear, we do not claim that the PDH can explain all symptoms in all individuals diagnosed with any such disorder. Nevertheless, we suggest that the PDH has substantial explanatory power for multiple language (and other) disorders, though the extent of this power likely varies across disorders. Given the large number of possible predictions that follow from our understanding of procedural memory as well as its role in language, only key predictions are discussed here and examined below for each disorder.

The location(s), severity, and extent of the neuroanatomical abnormalities may vary across disorders, as well as across individuals within a disorder, in several ways (Ullman & Pierpont 2005). Different gray and white matter levels of the cortico-BG-thalamocortical circuitry can be affected, including within the BG itself. Moreover, abnormalities could involve different (sub)circuits (portions of the parallel circuitry) at any level (e.g., striatum, thalamus, cortex). In principle, either the direct or indirect pathway could be preferentially affected. Together, such anatomical variability could lead to substantial phenotypic variability within and between disorders, including in the type and severity of the deficits. For example, striatal dysfunction should lead to learning deficits, whereas cortical dysfunction may primarily affect the processing of automatized skills. Moreover, anterior caudate/putamen abnormalities should mainly affect earlier phases of learning, whereas posterior striatal impairments should affect later phases, including the process of automatization.

The etiology and pathology of the neuroanatomical abnormalities may also vary across and within disorders, because a variety of genetic and environmental factors can lead to the dysfunction of gray or white matter structures/pathways within the procedural memory circuitry (Ullman & Pierpont 2005). Therefore, like the acquired aphasias, disorders accounted for by the PDH may be best explained by their pattern of neuroanatomical abnormalities, largely independent of etiology or pathology (though different etiologies and pathologies could also lead to functional variability). The striatum may be especially vulnerable to dysfunction (Mitchell et al. 1999), perhaps (like the hippocampus) due in part to its relatively high level of learning-related plasticity, which is associated with high neuronal energy demands, and thus susceptibility (Bartsch & Wulff 2015, Harris et al. 2012, Kreitzer & Malenka 2008, Mitchell et al. 1999). Therefore, abnormalities of the striatum may be particularly common. Nevertheless, any given etiology may affect brain structures more broadly, both within and in addition to procedural memory circuitry. Thus, a range of brain structures (e.g., including cortex and the cerebellum) may be affected in disorders explained by the PDH, if not directly by the underlying etiology, then indirectly through diaschesis or downstream functional effects. For example, striatal abnormalities impairing learning could affect frontal or temporal cortical processing of skills that were not normally learned.

Individuals with disorders explained by the PDH are not unable to learn. Indeed, their learning abilities may still derive in part from procedural memory, because procedural memory dysfunction does not imply afunctionality, so some learning may take place within procedural memory itself. Additionally, the PDH posits that declarative memory remains relatively intact in individuals with DPD (Ullman & Pierpont 2005, Ullman & Pullman 2015). Given the flexibility of declarative memory, this system should support the learning of a vast range of information and capabilities in individuals with DPD, including by compensating for procedural memory and other dysfunctions. This in turn leads to further predictions. For example, impairments of functions for which declarative memory can more easily compensate (e.g., grammar more than motor functions and local- more than long-distance dependencies, particularly higher-frequency local relations) should be less evident and more likely to improve over time. Moreover, individuals with better declarative memory abilities should compensate more successfully, and may even compensate their way out of diagnosis. Therefore, individuals or groups with better declarative memory may show a lower rate

of diagnosis of these disorders (e.g., females versus males, those with higher estrogen, Val versus Met carriers of the Val66Met polymorphism of the *BDNF* gene) (Ullman & Pullman 2015).

These patterns have various consequences (Ullman & Pierpont 2005). First, all the expected sources of variability described above may contribute (though not exhaustively) to the heterogeneity across individuals within a disorder. Second, and in contrast, commonalities of a disorder across individuals (e.g., grammatical difficulties in DLD) may be explained by shared dysfunction of the procedural memory circuitry, even if the severity and etiology varies. Third, the expected variability, in particular differences in the location(s) of anatomical abnormalities, should help explain differences between disorders. For example, motor-speech disorders likely result from dysfunction of motor circuits, whereas DLD or dyslexia may involve other circuits. The dysfunction of additional structures (whether through direct or indirect effects) may also contribute to differences between disorders. Thus, disorders may profitably be distinguished, at least in part, on the basis of such expected anatomical differences. Finally, the commonalities and comorbidities among disorders (e.g., articulation and other motor problems found in both DLD and motor-speech disorders) can be at least partly explained by overlapping abnormalities of procedural memory circuitry.

DEVELOPMENTAL DISORDERS

A number of developmental disorders may be at least partly explained by the PDH. Here we focus on three (sets of) language-related disorders: DLD, developmental motor-speech disorders, and dyslexia. All are good candidates for the PDH given their core and additional symptoms, which involve deficits of functions that depend on cortico-BG-thalamocortical circuits (also see Krishnan et al. 2016). Indeed, as discussed below, despite their differences, these disorders share multiple brain and behavioral characteristics linked to procedural memory circuitry. Moreover, within at least certain disorders, the PDH unifies various impairments that have previously been targeted by different accounts (e.g., grammatical, working memory, phonological, and temporal processing deficits in DLD).

Developmental Language Disorder

Consistent with the PDH, structural anomalies in DLD are found reliably (a high proportion of studies show anomalies) in the BG, particularly the caudate nucleus, and in frontal cortex (see Ullman & Pierpont 2005, table 1; Mayes et al. 2015). Similarly, in functional imaging studies both frontal cortex (especially Broca's region) and the BG show reliable anomalies. The head of the caudate appears to be particularly implicated in DLD. No other brain structures, including superior temporal cortex, the cerebellum, and the MTL, seem to show as reliable structural or functional abnormalities. Moreover, as expected by the PDH, various genetic and environmental factors associated with the BG (and in some cases, with dopaminergic processes) seem to contribute to the symptoms of DLD, including the *FOXP2*, *CNTNAP2*, *SEMA6D*, *ANKK1*, and *D2D2* genes (Eicher et al. 2013, Ercan-Sencicek et al. 2012, Newbury et al. 2010), prenatal nicotine exposure (Eicher et al. 2013), and thiamine deficiency (Fattal et al. 2011).

Procedural memory has been shown to be impaired in DLD (indeed, after the PDH was proposed in Ullman & Pierpont 2005). Meta-analysis has demonstrated deficits in the implicit SRT task in DLD, particularly during early stages of learning (Clark & Lum 2017, Lum et al. 2014). Impairments in DLD have also been found in graphomotor learning (Adi-Japha et al. 2011), rotary pursuit (Lee & Tomblin 2015; but see Hsu & Bishop 2014), probabilistic category learning with rapid feedback (Kemeny & Lukacs 2009; Lee & Tomblin 2012, 2015; but see Mayor-Dubois et al. 2014), and perceptual sequence learning (statistical learning) tasks such as word segmentation (Evans et al. 2009, Mayor-Dubois et al. 2014)—all tasks that depend on procedural memory,

generally with impairments found during early stages of learning. Thus, (early-stage) learning of perceptuomotor sequences, perceptual sequences, and categories is affected, specifically consistent with abnormalities of the anterior caudate nucleus (see the section titled What Do the Basal Ganglia Learn?). The absence of learning deficits in some studies may be explained in part by insufficient power or by declarative memory compensation. However, implicit learning tasks that do not depend on procedural memory circuits are not predicted to be impaired (Conway et al. 2019). Some studies of procedural memory in DLD have found impairments primarily or only after post-training delays (Adi-Japha et al. 2011, Hedenius et al. 2011), also implicating consolidation deficits in this system. Indeed, motor skills, particularly those involving sequences, across oral (speech and nonspeech) and nonoral (e.g., hand) movements, are impaired in DLD (Hill 2001, Ullman & Pierpont 2005). Moreover, consistent with procedural memory impairments, temporal processing (e.g., in rhythm and meter perception, and of acoustic stimuli of short duration or presented in rapid succession) is impaired in DLD (Leonard 2014, Przybylski et al. 2013). Finally, other functions that depend on the BG, whether or not they are directly linked to procedural memory, are also impaired in DLD, including working memory and executive functions such as inhibition (Henry et al. 2012, Ullman & Pierpont 2005).

Grammar is clearly impaired in DLD (Leonard 2014, Ullman & Pierpont 2005). Deficits are commonly found across syntax, morphology, and phonology, both in expressive and receptive language. Evidence links grammar problems to procedural memory in the form of correlations between measures of grammar and procedural learning, including for both perceptuomotor sequence and category learning (Hedenius et al. 2011, Mayor-Dubois et al. 2014). Particular problems seem to be found in those aspects of grammar that (*a*) rely on procedural memory (especially if they are difficult to learn in this system) and (*b*) cannot easily be compensated for by declarative memory (Ullman & Pierpont 2005). For example, long-distance dependencies appear to be more problematic than are local dependencies (Purdy et al. 2014), which are easier to learn in procedural memory (see above) and easier to chunk in declarative memory (Ullman & Pullman 2015). Previously encountered local dependencies that are higher frequency (e.g., walked) are less impaired than lower-frequency ones, which also suggests chunking (Ullman & Pierpont 2005). Similarly, canonical syntactic structures are relatively spared in DLD (Leonard & Kueser 2019), consistent with learning structured chunks of abstract categories in declarative memory (e.g., in constructions). Aspects of grammar that likely depend importantly on declarative memory in typically developing individuals should remain relatively intact in DLD. Indeed, this is the case with argument structure (Ullman & Pierpont 2005). Grammar learning (e.g., in artificial grammar paradigms) is also impaired in DLD, especially for long(er)-distance dependencies that cannot easily be chunked (Hsu & Bishop 2010). Moreover, automatization of grammar is problematic in DLD (Hsu & Bishop 2010). Other procedural memory–related functions that are impaired in DLD may aggravate the grammatical problems. For example, syntactic processing difficulties in DLD may be linked to the temporal processing (timing) deficits in the disorder (Przybylski et al. 2013). In contrast, even though working memory can play roles in grammar and is impaired in DLD, it does not seem to contribute to grammatical problems found in the disorder (Lum et al. 2012).

Learning in declarative memory remains largely intact in DLD, especially for nonverbal material, but also in the verbal domain after controlling for working memory and language impairments (Lum et al. 2012, Ullman & Pullman 2015). Additionally, some evidence suggests that children with DLD may consolidate information in declarative memory better than typically developing children (Lukacs et al. 2017). Consistent with intact declarative memory, various lines of evidence suggest that declarative memory compensates for grammatical deficits, e.g., by chunking (see above), learning explicit rules, and associative generalization (Ullman and Pullman 2015). In

one study, grammatical abilities correlated with procedural learning in typically developing children but with declarative learning in children with DLD, suggesting DLD compensation with declarative memory (Lum et al. 2012). Some electrophysiological evidence suggests that whereas typical individuals show early left anterior negativities (these event-related components are fast, automatic, and have been linked to procedural memory) in response to syntactic anomalies, those with DLD instead show N400s, which have been linked to declarative memory (Ullman & Pullman 2015). Moreover, the prevalence of DLD is higher in males than in females, consistent with female advantages in declarative memory and thus a particularly high level of declarative memory compensation by girls (Ullman & Pullman 2015).

Word learning and knowledge remain largely normal in DLD, unless they involve functions that depend on procedural memory or its neural substrates. Thus, children with DLD can learn new words (and in some cases show enhanced performance), especially in rich semantic contexts or with repetitive input, and appear to have normal lexical/semantic organization (Ullman & Pierpont 2005). They also show largely intact N400 responses for both content words and lexical/semantic processing (Ullman & Pierpont 2005, Ullman & Pullman 2015). In contrast, word learning that depends on procedural memory functions, such as word segmentation and learning in grammatical contexts, seems to be impaired in DLD (Evans et al. 2009, Mayor-Dubois et al. 2014, Ullman & Pierpont 2005). This may help explain why children with DLD have smaller vocabularies than do their age-matched peers (Lee 2011). Additionally, they appear to have shallower semantic representations for known words (McGregor 2009), consistent with procedural memory–based category learning difficulties. They also have problems inhibiting competing words in word recognition (McMurray et al. 2014), consistent with deficits in executive function. Word finding and production are especially impaired in DLD (particularly with little contextual support and during rapid naming), as compared to receptive lexical abilities, consistent with abnormalities of frontal/BG circuits (which underlie recall) and speech-motor (articulatory) difficulties in DLD (Ullman & Pierpont 2005). Finally, children with DLD have more problems with function (closed-class) words than with content (open-class) words and with verbs than nouns, consistent with the strong grammatical roles associated with function (versus content) words and verbs (versus nouns), and the action knowledge linked to verbs (Ullman & Pierpont 2005).

How about other language-related functions that rely on procedural memory, such as speech-sound category learning, articulation, and speech perception? Although we are not aware of any DLD research on speech-sound category learning particularly linked to procedural memory, articulation may be affected in DLD (Hill 2001, Ullman & Pierpont 2005). Moreover, speech perception (e.g., identifying words in noise) is impaired in the disorder (Ziegler et al. 2005). Although this impairment has generally been ascribed to various processing deficits (Leonard 2014, Ullman & Pierpont 2005), it can also be explained by processes or knowledge linked to procedural memory that are critical for speech perception and are impaired in DLD, namely speech-motor representations (Katz et al. 1992), perception of timing (Corriveau & Goswami 2009), and prediction (Hestvik et al. 2010).

Developmental Motor-Speech Disorders

Motor-speech disorders refer, as an umbrella term, to disorders in the production of speech. Idiopathic developmental motor-speech disorders include articulation disorder, childhood apraxia of speech (CAS) (also referred to as verbal dyspraxia), and developmental stuttering. Articulation disorder refers to problems producing discrete sounds, classes of sounds, or sequences of sounds. In rare cases, children with severe articulation problems are diagnosed with CAS, which involves difficulties planning and programming motor movements in speech production in the absence of

musculoskeletal atypicalities in the speech production apparatus (Bernthal et al. 2009). Thus, both articulation disorder and CAS describe problems with aspects of producing the motor sequences required for speech production. By contrast, developmental stuttering is a fluency disorder that affects the timing, patterning (e.g., prosody), and rhythm of speech (Alm 2004).

Neuroanatomical evidence suggests that procedural memory circuits are involved in developmental motor-speech disorders. Most such research has focused on affected members of the KE family, who can be characterized as having CAS (Alcock et al. 2000a). These individuals have a deleterious mutation of *FOXP2*, a gene that is strongly associated with the striatum, especially the caudate nucleus (but not with the hippocampus) (Ullman & Pierpont 2005). Broca's region and motor regions show the most consistent abnormalities across both structural and functional imaging in affected members, whereas structural anomalies alone are clustered in the caudate nucleus, along with superior temporal cortex and the cerebellum (see Ullman & Pierpont 2005, table 1). Although the neuroanatomical basis of other populations of CAS or articulation disorder has not been well studied, we are aware of one functional imaging study of nonword repetition in adolescents with articulation disorder. This study found activation anomalies in the putamen, cerebellum, and various cortical regions, including the inferior frontal gyrus and SMA (Tkach et al. 2011). In individuals with developmental stuttering, a robust literature points to structural abnormalities in cortico-BG-thalamocortical circuits, such as reduced gray matter volumes in the putamen and cortical regions, and altered connectivity among the putamen, thalamus, and the SMA and other cortical areas (Alm 2004, Lu et al. 2010). Moreover, activity in the caudate head correlates with the severity of disfluencies in affected individuals (Giraud et al. 2008). Overall, the data suggest that the fluency symptoms of developmental stuttering may be rooted in the striatum and its associated circuitry (Alm 2004).

There is some, although not extensive, evidence linking these motor-speech conditions to procedural memory dysfunction. Children with articulation disorder have difficulties adapting their articulatory motor programs to perturbed auditory feedback, suggesting problems in BG-based articulatory learning (Terband et al. 2014). Adults with developmental stuttering are impaired at both perceptuomotor and perceptual sequence learning (Smits-Bandstra & De Nil 2007). Moreover, similar to DLD, all three conditions are associated with concomitant deficits in motor functions (Busan et al. 2013, Hill 2001), auditory timing (e.g., rhythm perception) (Alcock et al. 2000b, Kenney et al. 2006, Wieland et al. 2015), and working memory (Bajaj 2007, Kenney et al. 2006), all of which are tied to procedural memory or its neural substrates.

The most compelling behavioral evidence regarding the PDH for motor-speech disorders may be in their core speech-motoric and associated linguistic deficits. It would follow from the predicted role of BG-based learning in articulation that procedural memory deficits should impede the learning and automatization of speech-motoric sequences. This in turn may lead to increased gestural variability between productions as well as a slower speech rate. Indeed, both children with articulation disorder and those with CAS show both such symptoms (Flipsen Jr. 2002, Shriberg et al. 2003, Terband et al. 2011). Beyond these core characteristics, articulation disorder and CAS are both associated with deficits in phonology, morphology, and syntax (Alcock et al. 2000a, Mortimer & Rvachew 2010) and with eventual problems in learning how to read (Gillon & Moriarty 2007, Peterson et al. 2009). This suggests the possibility that these motor-speech disorders are associated with procedural impairments in these domains as well, consistent with their comorbidities with DLD and dyslexia. The core deficits of stuttering are also linked to procedural memory and its neural substrates. Indeed, their deficits in the timing and initiation of speech-motor programs for the fluent production of speech are expected given a dysfunction in the BG-SMA circuit (Alm 2004). Moreover, developmental stuttering is associated with morphological

and syntactic problems and with particular difficulties with verbs (Bauman et al. 2012, Smith & Weber 2017), suggesting similarities with DLD in procedural language impairments.

We are unaware of any studies that have directly investigated declarative memory in the three disorders, though it is predicted to remain largely intact. Nevertheless, some evidence suggests compensatory roles for this system. Declarative memory may help compensate for impairments in CAS, given that explicit strategies are effective behavioral therapies for this disorder (Dale & Hayden 2013). Additionally, like DLD, syntactic anomalies yield N400s in developmental stuttering (Smith & Weber 2017). Intriguingly, all three disorders are more prevalent in boys than girls (Bernthal et al. 2009, Yairi & Ambrose 1999), and developmental stuttering is more likely to resolve by adulthood in women than in men (Porfert & Rosenfield 1978), consistent with a female advantage in compensating with declarative memory.

Developmental Dyslexia

Developmental dyslexia refers to a reading disability involving difficulties with fluent decoding (processing grapheme-phoneme mappings) that may cascade into reading comprehension problems—as opposed to reading comprehension deficits that may be secondary to language comprehension problems (Am. Psychiatr. Assoc. 2013, Gough & Tunmer 1986).

Structural anatomical abnormalities in dyslexia are found most consistently in superior temporal/temporoparietal and inferior/ventral temporal regions as well as the cerebellum (Eckert et al. 2016), though they are also observed in other regions, including the putamen (Eckert et al. 2005, Pernet et al. 2009) and caudate head (Brown et al. 2001). Although functional imaging abnormalities must be interpreted carefully due to their task dependence, they have been observed in anatomical meta-analyses in the anterior caudate, putamen/globus pallidus, motor and inferior frontal regions, and the temporal regions listed above (Paulesu et al. 2014, Richlan et al. 2011). In an SRT study of dyslexia, activation abnormalities were found in the putamen as well as in the SMA, cerebellum, and parietal regions (Menghini et al. 2006). Thus, procedural memory brain structures are implicated in dyslexia, including the BG (anterior caudate and putamen) and frontal (motor and inferior frontal) structures. It remains unclear whether the underlying etiologies affect the procedural circuit and/or other structures directly or whether some of the abnormalities are (also) due to indirect effects (diaschesis, downstream effects).

As with DLD, procedural memory is clearly impaired in dyslexia. Meta-analysis has demonstrated perceptuomotor sequence learning deficits in the implicit SRT task, particularly in early stages of learning (Clark & Lum 2017, Lum et al. 2013). Impairments have also been found in perceptual sequence (statistical) learning (Gabay et al. 2015a) and nonlinguistic visual category learning (in the weather prediction task) and auditory category learning (Gabay & Holt 2015, Gabay et al. 2015b). Thus, similar to DLD, (early phases of) perceptuomotor sequence, perceptual sequence, and category learning are all impaired, specifically consistent with anterior caudate abnormalities (but not with cerebellar anomalies) (see the section titled What Do the Basal Ganglia Learn?). In addition, dyslexia is linked to deficits of other functions associated with procedural memory or its underlying circuits, including impaired automaticity in speech production (Catts 1989), (visual) artificial grammar learning difficulties (van Witteloostuijn et al. 2017), atypical (morpho-)syntactic development and processing (Cantiani et al. 2013, Scarborough 1991), atypical temporal processing (Farmer & Klein 1995), impaired linguistic prediction (Huettig & Brouwer 2015), and poor working memory (Smith-Spark & Fisk 2007) and inhibition (Wang et al. 2012).

The reading problems in dyslexia may be at least partly explained by problems with procedural memory and its circuitry. First, the core decoding impairment in dyslexia is commonly attributed to the phonological deficits that clearly accompany the disorder, and often specifically to apparent

problems with speech-sound representations (Ramus et al. 2003). These speech-sound representation problems have in turn been explained by atypical perception, such as impairments of rapid auditory processing or temporal processing (Farmer & Klein 1995, Ramus et al. 2003). The PDH can account for at least some of these perceptual difficulties. Second, the PDH might expect impairments in learning speech-sound categories. Such impairments could result in a lack of cohesion in phonemic category representations, which is indeed consistent with observed weaknesses in dyslexia in phonemic anchoring (Ahissar 2007) and talker-specific adaptation (Perrachione et al. 2011) and with an overreliance on noncontrastive features of speech (Serniclaes 2011). Moreover, these characteristics appear to be consistent with learning (suboptimal) episodic (that is, instance specific) features of speech in declarative memory, rather than (optimal) categories in procedural memory (Shohamy et al. 2008). Third, magnocellular circuitry is closely linked with the dorsal stream (Goswami 2015), which in turn is closely linked to procedural memory (see above), suggesting that posited magnocellular deficits in dyslexia (see the section titled Introduction) may be related to the impairments posited by the PDH. Finally, given the role of procedural memory in learning predictive associations, it may underlie learning grapheme-phoneme (and phoneme-grapheme) mappings (though these mappings are likely also learned in declarative memory, especially at early stages of learning). Thus, procedural dysfunction could also contribute directly to dyslexia-associated impairments in learning such mappings.

Declarative memory seems to remain largely intact in dyslexia. Nonverbal learning in declarative memory, including in implicit tasks such as spatial contextual learning, does not appear to be impaired (Ullman & Pullman 2015). Verbal learning, including form-meaning mappings in word learning, is problematic, though the observed deficits may be explained by procedural memory–related dysfunctions such as with phonology or working memory (Litt & Nation 2014, Ullman & Pullman 2015). Indeed, though encoding of verbal material (which depends on such functions) is impaired, retention of learned material seems to remain intact. For example, children with dyslexia may learn fewer words during encoding in a list-learning task but have no difficulty subsequently remembering the words they learned (Kramer et al. 2000). Some evidence suggests that declarative memory may even be enhanced in dyslexia (Ullman & Pullman 2015).

Evidence also suggests that declarative memory plays compensatory roles in dyslexia. Individuals with dyslexia appear to rely disproportionately on declarative memory for reading, using at least three strategies: chunking, a reliance on semantics, and the use of explicit knowledge (Ullman & Pullman 2015). First, whereas people with dyslexia have particular problems reading made-up words, they are less impaired at reading real words (whose grapheme and/or phonological strings and mappings could have been memorized), especially those that occur with higher frequency. Second, reading with supportive semantic contexts is particularly helpful in dyslexia, and improvements in reading during development appear to be due to an increased reliance on semantic knowledge. Third, explicit instruction of grapheme-phoneme correspondences (phonics) is an effective therapy in dyslexia (Ehri et al. 2001). Indeed, successful behavioral interventions in dyslexia lead to increased (compensatory) activation in the hippocampus and other MTL structures as well as to increased hippocampal gray matter (Ullman & Pullman 2015). As with DLD, syntactic violations elicit N400 responses in dyslexia (Cantiani et al. 2013), suggesting declarative memory compensation for syntactic impairments. Learning speech-sound categories may rely on declarative memory–based episodic features in individuals with dyslexia (see above) but mainly on procedural memory in typically developing individuals (see the section titled Speech-Sound Representations). Evidence also suggests that better declarative memory correlates with better reading in individuals with dyslexia but not in typically developing individuals (Ullman & Pullman 2015). Finally, as with the other disorders discussed above, dyslexia may be more prevalent in boys than in girls, though this pattern remains unclear (Hawke et al. 2009).

Other Developmental Disorders

The PDH may at least partially explain other disorders, whether or not they affect aspects of language. Indeed, abnormalities of procedural memory and its underlying circuitry, as well as compensation by declarative memory, have been found in a variety of other developmental disorders—which are often comorbid with each other and with the language disorders discussed above. These disorders include attention-deficit/hyperactivity disorder, autism spectrum disorder, Tourette syndrome, obsessive-compulsive disorder, developmental coordination disorder, mathematical disability (e.g., developmental dyscalculia), and schizophrenia (Bradshaw & Enticott 2014; Clark & Lum 2017; Evans & Ullman 2016; Ullman 2004; Ullman & Pierpont 2005; Ullman & Pullman 2015; Walenski et al. 2006, 2007, 2010). In many of these disorders, procedural memory seems to be impaired, yet in others (perhaps in subgroups) abnormalities manifest as procedural memory enhancements, including of grammar (e.g., in Tourette syndrome and possibly autism spectrum disorder) (Dye et al. 2016; Takács et al. 2018; Walenski et al. 2007, 2014). Finally, impairments of procedural memory and its neural substrates, as well as compensation by declarative memory, have been found in various adult-onset disorders, including Parkinson's disease and nonfluent/agrammatic aphasia (Ullman 2004, Ullman & Pullman 2015). Thus, the core tenets of the PDH may extend beyond developmental disorders.

CONCLUSION

The PDH has numerous implications. In terms of basic research, the pattern of findings underscores the importance of the BG in language—in particular the anterior caudate nucleus, which likely plays a critical role in early phases of learning grammar and other aspects of language. Indeed, the evidence presented above not only strengthens the basic principles of the declarative/procedural model regarding lexicon and grammar (Ullman 2004, 2015, 2016), but also extends them to other core aspects of language (speech sound representations, articulation, speech production, and speech perception) as well as other functions such as reading and math.

We have explored only a subset of the predictions that follow from the PDH. The scientific community's quite specific understanding of procedural and declarative memory and their underlying circuitry, as well as of the language and nonlanguage functions that depend on them, leads to a wide range of predictions and potentially promising lines of research—only some of which have begun to be explored. For example, disorders explained by the PDH may be associated with deficits of numerous functions and behaviors that may rely on the affected circuitry, as diverse as driving, social skills, habit formation, musical abilities, dorsal stream functions, BG-based selection, and even motivation and related affective functions. Our knowledge of declarative memory also leads to various intriguing possibilities. For example, are higher estrogen points of the menstrual cycle associated with fewer deficits? Do individuals with worse sleep show more impairments (because sleep supports declarative memory consolidation), and thus could sleep play a role in the disorders?

The PDH also makes translational predictions. It suggests the potential utility of pharmacological (e.g., dopaminergic), behavioral, and other interventions that have been successfully employed to enhance procedural memory, including in other disorders in which procedural memory circuitry is affected (de Vries et al. 2010, Ullman & Pierpont 2005). The compensatory role of declarative memory also suggests the likely value of pharmacological and other interventions aimed at enhancing such compensation (Ullman & Pullman 2015). Additionally, diagnoses of these disorders may be aided by the presence of particular neuroanatomical anomalies in procedural memory structures (and lack thereof elsewhere, such as in the MTL) as well as by the presence of deficits

(e.g., long-distance grammatical dependencies) that are difficult to compensate for in declarative memory (Ullman & Pullman 2015).

In sum, the PDH offers an explanatory framework for multiple aspects of the brain and behavioral correlates of various language (and other) developmental disorders, as well as their commonalities, comorbidities, and differences. It thus constitutes a powerful unifying neurocognitive account of these developmental disorders. Given the wide range of specific predictions that follow from the PDH, only the tip of the basic and translational research iceberg has been examined. We hope the present review serves to guide research that probes the depths of this iceberg.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank Gregory Ashby, Christopher Conway, Laurie Cutting, Jarrad Lum, Maximillian Riesenhuber, Lauren Russell, Carol Seger, Andrew Tiu, and Christine Weber for helpful input on this article. This work was supported in part by National Science Foundation BCS 1439290, National Institute of Health R21 HD 087088, and a research grant from the Tourette Association of America to M.T.U.; by National Institute of Health R21 DC 016391 to F.S.E.; and by Hungarian Scientific Research Fund OTKA PD 124148, National Brain Research Program project 2017-1.2.1-NKP-2017-00002, and a Janos Bolyai Research Fellowship of the Hungarian Academy of Sciences to K.J.

LITERATURE CITED

Abla D, Katahira K, Okanoya K. 2008. On-line assessment of statistical learning by event-related potentials. *J. Cogn. Neurosci.* 20:952–64

Adi-Japha E, Strulovich-Schwartz O, Julius M. 2011. Delayed motor skill acquisition in children with language impairment. *Res. Dev. Disabil.* 32:2963–71

Ahissar M. 2007. Dyslexia and the anchoring-deficit hypothesis. *Trends Cogn. Sci.* 11:458–65

Alcock KJ, Passingham RE, Watkins KE, Vargha-Khadem F. 2000a. Oral dyspraxia in inherited speech and language impairment and acquired dysphasia. *Brain Lang.* 75:17–33

Alcock KJ, Passingham RE, Watkins K, Vargha-Khadem F. 2000b. Pitch and timing abilities in inherited speech and language impairment. *Brain Lang.* 75:34–46

Alm PA. 2004. Stuttering and the basal ganglia circuits: a critical review of possible relations. *J. Commun. Disord.* 37:325–69

Am. Psychiatr. Assoc. 2013. *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: APA. 5th ed.

Ashby FG, Crossley MJ. 2012. Automaticity and multiple memory systems. *Wiley Interdiscip. Rev. Cogn. Sci.* 3:363–76

Ashby FG, Ennis JM, Spiering BJ. 2007. A neurobiological theory of automaticity in perceptual categorization. *Psychol. Rev.* 114:632

Ashby FG, Maddox WT. 2011. Human category learning 2.0. *Ann. N.Y. Acad. Sci.* 1224:147–61

Bajaj A. 2007. Working memory involvement in stuttering: exploring the evidence and research implications. *J. Fluen. Disord.* 32:218–38

Balleine BW, O'Doherty JP. 2010. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* 35:48

Bartsch T, Wulff P. 2015. The hippocampus in aging and disease: from plasticity to vulnerability. *Neuroscience* 309:1–16

Bauman J, Hall NE, Wagovich SA, Weber-Fox CM, Ratner NB. 2012. Past tense marking in the spontaneous speech of preschool children who do and do not stutter. *J. Fluen. Disord.* 37:314–24

Bernthal JE, Bankson NW, Flipsen P. 2009. *Articulation and Phonological Disorders: Speech Sound Disorders in Children*. Boston, MA: Pearson

Bishop DV, Snowling MJ, Thompson PA, Greenhalgh T, CATALISE-2 Consort. 2017. Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: terminology. *J. Child Psychol. Psychiatry* 58:1068–80

Bostan AC, Strick PL. 2018. The basal ganglia and the cerebellum: nodes in an integrated network. *Nat. Rev. Neurosci.* 19:338–50

Bradshaw JL, Enticott PG. 2014. *Developmental Disorders of the Frontostriatal System: Neuropsychological, Neuropsychiatric and Evolutionary Perspectives*. East Sussex, UK: Psychology Press

Brown WE, Eliez S, Menon V, Rumsey JM, White CD, Reiss AL. 2001. Preliminary evidence of widespread morphological variations of the brain in dyslexia. *Neurology* 56:781–83

Busan P, D'Ausilio A, Borelli M, Monti F, Pelamatti G, et al. 2013. Motor excitability evaluation in developmental stuttering: a transcranial magnetic stimulation study. *Cortex* 49:781–92

Catts HW. 1989. Speech production deficits in developmental dyslexia. *J. Speech Hear. Disord.* 54:422–28

Cantiani C, Lorusso ML, Perego P, Molteni M, Guasti MT. 2013. Event-related potentials reveal anomalous morphosyntactic processing in developmental dyslexia. *Appl. Psycholinguist.* 34:1135–62

Chandrasekaran B, Yi HG, Maddox WT. 2014. Dual-learning systems during speech category learning. *Psychon. Bull. Rev.* 21:488–95

Chandrasekaran B, Yi HG, Smayda KE, Maddox WT. 2016. Effect of explicit dimensional instruction on speech category learning. *Atten. Percept. Psychophys.* 78:566–82

Chomsky N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press

Clark GM, Lum JA. 2017. Procedural learning in Parkinson's disease, specific language impairment, dyslexia, schizophrenia, developmental coordination disorder, and autism spectrum disorders: a second-order meta-analysis. *Brain Cogn.* 117:41–48

Conway CM, Arciuli J, Lum JAG, Ullman MT. 2019. Seeing problems that may not exist: a reply to West et al.'s (2018) questioning of the procedural deficit hypothesis. *Dev. Sci.* 2019:e12814

Corriveau KH, Goswami U. 2009. Rhythmic motor entrainment in children with speech and language impairments: tapping to the beat. *Cortex* 45:119–30

Curran T. 1997. Higher-order associative learning in amnesia: evidence from the serial reaction time task. *J. Cogn. Neurosci.* 9:522

Dale PS, Hayden DA. 2013. Treating speech subsystems in childhood apraxia of speech with tactual input: the PROMPT approach. *Am. J. Speech-Lang. Pathol.* 22:644–61

Davachi L. 2006. Item, context and relational episodic encoding in humans. *Curr. Opin. Neurobiol.* 16:693–700

Davis MH, Gaskell MG. 2009. A complementary systems account of word learning: neural and behavioural evidence. *Philos. Trans. B* 364:3773–800

De Diego-Balaguer R, Couette M, Dolbeau G, Dürr A, Youssov K, Bachoud-Lévi AC. 2008. Striatal degeneration impairs language learning: evidence from Huntington's disease. *Brain* 131:2870–81

de Vries MH, Ulte C, Zwitserlood P, Szymanski B, Knecht S. 2010. Increasing dopamine levels in the brain improves feedback-based procedural learning in healthy participants: an artificial-grammar-learning experiment. *Neuropsychologia* 48:3193–97

Defne Abur R, Lester-Smith RA, Daliri A, Lupiani AA, Guenther FH, Stepp CE. 2018. Sensorimotor adaptation of voice fundamental frequency in Parkinson's disease. *PLOS ONE* 13:e0191839

Delgado MR, Miller MM, Inati S, Phelps EA. 2005. An fMRI study of reward-related probability learning. *NeuroImage* 24:862–73

Doupe AJ, Kuhl PK. 1999. Birdsong and human speech: common themes and mechanisms. *Annu. Rev. Neurosci.* 22:567–631

Doyon J, Bellec P, Amsel R, Penhune V, Monchi O, et al. 2009. Contributions of the basal ganglia and functionally related brain structures to motor learning. *Behav. Brain Res.* 199:61–75

Draganski B, Kherif F, Klöppel S, Cook PA, Alexander DC, et al. 2008. Evidence for segregated and integrative connectivity patterns in the human basal ganglia. *J. Neurosci.* 28:7143–52

Duffy JR. 2006. Apraxia of speech in degenerative neurologic disease. *Aphasiology* 20:511–27

Durrant SJ, Cairney SA, Lewis PA. 2012. Overnight consolidation aids the transfer of statistical knowledge from the medial temporal lobe to the striatum. *Cereb. Cortex* 23:2467–78

Dye CD, Walenski M, Mostofsky SH, Ullman MT. 2016. A verbal strength in children with Tourette syndrome? Evidence from a non-word repetition task. *Brain Lang.* 160:61–70

Eckert MA, Berninger VW, Vaden KI Jr., Gebregziabher M, Tsu L. 2016. Gray matter features of reading disability: a combined meta-analytic and direct analysis approach. *eNeuro* 3:ENEURO.0103-15. 2015

Eckert MA, Leonard CM, Wilke M, Eckert M, Richards T, et al. 2005. Anatomical signatures of dyslexia in children: unique information from manual and voxel based morphometry brain measures. *Cortex* 41:304–15

Ehri LC, Nunes SR, Stahl SA, Willows DM. 2001. Systematic phonics instruction helps students learn to read: evidence from the National Reading Panel's meta-analysis. *Rev. Educ. Res.* 71(3):393–447

Eichenbaum H. 2012. Memory systems. *Handbook of Psychology*, Vol. 3, ed. I Weiner, RJ Nelson, SJ Mizumori, pp. 551–73. Somerset, NJ: Wiley. 2nd ed.

Eicher JD, et al. 2013. Associations of prenatal nicotine exposure and the dopamine related genes *ANKK1* and *DRD2* to verbal language. *PLOS ONE* 8:e63762

Elbro C, Dalby M, Maarbjerg S. 2011. Language-learning impairments: a 30-year follow-up of language-impaired children with and without psychiatric, neurological and cognitive difficulties. *Int. J. Lang. Commun. Disord.* 46:437–48

Ercan-Sencicek AG, et al. 2012. A balanced t(10;15) translocation in a male patient with developmental language disorder. *Eur. J. Med. Genet.* 55:128–31

Evans JL, Saffran JR, Robe-Torres K. 2009. Statistical learning in children with specific language impairment. *J. Speech Lang. Hear. Res.* 52:321–55

Evans TM, Ullman MT. 2016. An extension of the procedural deficit hypothesis from developmental language disorders to mathematical disability. *Front. Psychol.* 7:1318

Farmer ME, Klein RM. 1995. The evidence for a temporal processing deficit linked to dyslexia: a review. *Psychon. Bull. Rev.* 2:460–93

Fattal I, Friedmann N, Fattal-Valevski A. 2011. The crucial role of thiamine in the development of syntax and lexical retrieval: a study of infantile thiamine deficiency. *Brain* 134:1720–39

Fernández G, Tendolkar I. 2006. The rhinal cortex: gatekeeper of the declarative memory system. *Trends Cogn. Sci.* 10:358–62

Flipsen P Jr. 2002. Longitudinal changes in articulation rate and phonetic phrase length in children with speech delay. *J. Speech Lang. Hear. Res.* 45:100–10

Foerde KE, Knowlton BJ, Poldrack RA. 2006. *Retention of classification learning after training under single and dual task conditions.* Paper presented at the Society for Neuroscience, Atlanta, GA, Oct. 14–18

Foerde K, Shohamy D. 2011a. Feedback timing modulates brain systems for learning in humans. *J. Neurosci.* 31:13157–67

Foerde K, Shohamy D. 2011b. The role of the basal ganglia in learning and memory: insight from Parkinson's disease. *Neurobiol. Learn. Mem.* 96:624–36

Frank MJ. 2005. Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and non-medicated parkinsonism. *J. Cogn. Neurosci.* 17:51–72

Gabay Y, Holt LL. 2015. Incidental learning of sound categories is impaired in developmental dyslexia. *Cortex* 73:131–43

Gabay Y, Thiessen ED, Holt LL. 2015a. Impaired statistical learning in developmental dyslexia. *J. Speech Lang. Hear. Res.* 58:934–45

Gabay Y, Vakil E, Schiff R, Holt LL. 2015b. Probabilistic category learning in developmental dyslexia: evidence from feedback and paired-associate weather prediction tasks. *Neuropsychology* 29:844

Gillon GT, Moriarty BC. 2007. Childhood apraxia of speech: children at risk for persistent reading and spelling disorder. In *Seminars in Speech and Language*, 28:48–57. New York: Thieme Med.

Giraud AL, Neumann K, Bachoud-Levi AC, von Gudenberg AW, Euler HA, et al. 2008. Severity of dysfluency correlates with basal ganglia activity in persistent developmental stuttering. *Brain Lang*. 104:190–99

Goldberg A. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago, IL: Univ. Chicago Press

Goldstein MH, Schwade JA. 2008. Social feedback to infants' babbling facilitates rapid phonological learning. *Psychol. Sci.* 19:515–23

Goswami U. 2015. Sensory theories of developmental dyslexia: three challenges for research. *Nat. Rev. Neurosci.* 16:43

Gough PB, Tunmer WE. 1986. Decoding, reading, and reading disability. *Remedial Spec. Educ.* 7:6–10

Grahn JA, Rowe JB. 2009. Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. *J. Neurosci.* 29:7540–48

Graybiel AM, Grafton ST. 2015. The striatum: where skills and habits meet. *Cold Spring Harb. Perspect. Biol.* 7:a021691

Guasti MT. 2017. *Language Acquisition: The Growth of Grammar*. Cambridge, MA: MIT Press

Haesler S, Rochefort C, Georgi B, Licznerski P, Osten P, Scharff C. 2007. Incomplete and inaccurate vocal imitation after knockdown of *FoxP2* in songbird basal ganglia nucleus area X. *PLOS Biol.* 5:e321

Hamrick P, Lum JA, Ullman MT. 2018. Child first language and adult second language are both tied to general-purpose learning systems. *PNAS* 115:1487–92

Hardwick RM, Rottschy C, Miall RC, Eickhoff SB. 2013. A quantitative meta-analysis and review of motor learning in the human brain. *Neuroimage* 67:283–97

Harris JJ, Jolivet R, Attwell D. 2012. Synaptic energy use and supply. *Neuron* 75:762–77

Hawke JL, Olson RK, Willcut EG, Wadsworth SJ, DeFries JC. 2009. Gender ratios for reading difficulties. *Dyslexia* 15:239–42

Hazeltine E, Grafton ST, Ivry R. 1997. Attention and stimulus characteristics determine the locus of motor-sequence encoding. A PET study. *Brain* 120(Pt. 1):123–40

Hedenius M, Persson J, Tremblay A, Adi-Japha E, Veríssimo J, et al. 2011. Grammar predicts procedural learning and consolidation deficits in children with specific language impairment. *Res. Dev. Disabil.* 32:2362–75

Hélie S, Ell SW, Ashby FG. 2015. Learning robust cortico-cortical associations with the basal ganglia: an integrative review. *Cortex* 64:123–35

Hélie S, Roeder JL, Ashby FG. 2010a. Evidence for cortical automaticity in rule-based categorization. *J. Neurosci.* 30:14225–34

Hélie S, Waldschmidt JG, Ashby FG. 2010b. Automaticity in rule-based and information-integration categorization. *Atten. Percept. Psychophys.* 72:1013–31

Henry LA, Messer DJ, Nash G. 2012. Executive functioning in children with specific language impairment. *J. Child Psychol. Psychiatry* 53:37–45

Henson RN, Gagnepain P. 2010. Predictive, interactive multiple memory systems. *Hippocampus* 20:1315–26

Hestvik A, Schwartz RG, Tornyova L. 2010. Relative clause gap-filling in children with specific language impairment. *J. Psycholinguist. Res.* 39:443–56

Hill EL. 2001. Non-specific nature of specific language impairment: a review of the literature with regard to concomitant motor impairments. *Int. J. Lang. Commun. Disord.* 36:149–71

Holt LL, Lotto AJ. 2010. Speech perception as categorization. *Atten. Percept. Psychophys.* 72:1218–27

Hsu HJ, Bishop DV. 2010. Grammatical difficulties in children with specific language impairment: Is learning deficient? *Hum. Dev.* 53:264–77

Hsu HJ, Bishop DV. 2014. Sequence-specific procedural learning deficits in children with specific language impairment. *Dev. Sci.* 17:352–65

Huettig F, Brouwer S. 2015. Delayed anticipatory spoken language processing in adults with dyslexia—evidence from eye-tracking. *Dyslexia* 21:97–122

Janacsek K, Fiser J, Nemeth D. 2012. The best time to acquire new skills: age-related differences in implicit sequence learning across the human lifespan. *Dev. Sci.* 15:496–505

Janacsek K, Nemeth D. 2012. Predicting the future: from implicit learning to consolidation. *Int. J. Psychophysiol.* 83:213–21

Johari K, Walenski M, Reifegerste J, Ashrafi F, Ullman MT. 2019. Sex, dopamine, and hypokinesia: a study of inflectional morphology in Parkinson's disease. *Neuropsychology* 33:508–22

Kaan E. 2014. Predictive sentence processing in L2 and L1: What is different? *Linguist. Approaches Biling.* 4:257–82

Karuza EA, Newport EL, Aslin RN, Starling SJ, Tivarus ME, Bavelier D. 2013. The neural correlates of statistical learning in a word segmentation task: an fMRI study. *Brain Lang.* 127:46–54

Katz WF, Curtiss S, Tallal P. 1992. Rapid automatized naming and gesture by normal and language-impaired children. *Brain Lang.* 43:623–41

Kemeny F, Lukacs A. 2009. Impaired procedural learning in language impairment: results from probabilistic categorization. *J. Clin. Exp. Neuropsychol.* 32:249–58

Kenney MK, Barac-Cikoja D, Finnegan K, Jeffries N, Ludlow CL. 2006. Speech perception and short-term memory deficits in persistent developmental speech disorder. *Brain Lang.* 96:178–90

King BR, Hoedlmoser K, Hirschauer F, Dolfen N, Albouy G. 2017. Sleeping on the motor engram: the multifaceted nature of sleep-related motor memory consolidation. *Neurosci. Biobehav. Rev.* 80:1–22

Kotz SA, Schwartze M. 2010. Cortical speech processing unplugged: a timely subcortico-cortical framework. *Trends Cogn. Sci.* 14:392–99

Kramer JH, Knee K, Delis DC. 2000. Verbal memory impairments in dyslexia. *Arch. Clin. Neuropsychol.* 15:83–93

Kreitzer AC, Malenka RC. 2008. Striatal plasticity and basal ganglia circuit function. *Neuron* 60:543–54

Krishnan S, Watkins KE, Bishop DV. 2016. Neurobiological basis of language learning difficulties. *Trends Cogn. Sci.* 20:701–14

Kutas M, DeLong KA, Smith NJ. 2011. A look around at what lies ahead: prediction and predictability in language processing. In *Predictions in the Brain: Using Our Past to Generate a Future*, ed. M Bar, pp. 190–207. New York: Oxford Univ. Press

Lavenex P, Lavenex PB. 2013. Building hippocampal circuits to learn and remember: insights into the development of human memory. *Behav. Brain Res.* 254:8–21

Lee J. 2011. Size matters: early vocabulary as a predictor of language and literacy competence. *Appl. Psycholinguist.* 32:69–92

Lee JC, Tomblin JB. 2012. Reinforcement learning in young adults with developmental language impairment. *Brain Lang.* 123:154–63

Lee JC, Tomblin JB. 2015. Procedural learning and individual differences in language. *Lang. Learn. Dev.* 11:215–36

Leonard LB. 2014. *Children with Specific Language Impairment*. Cambridge, MA: MIT Press. 2nd ed.

Leonard LB, Kueser JB. 2019. Five overarching factors central to grammatical learning and treatment in children with developmental language disorder. *Int. J. Lang. Commun. Disord.* 54:347–61

Litt RA, Nation K. 2014. The nature and specificity of paired associate learning deficits in children with dyslexia. *J. Mem. Lang.* 71:71–88

Love T, Walenski M, Swinney D. 2009. Slowed speech input has a differential impact on on-line and off-line processing in children's comprehension of pronouns. *J. Psycholinguist. Res.* 38:285–304

Lu C, Peng D, Chen C, Ning N, Ding G, et al. 2010. Altered effective connectivity and anomalous anatomy in the basal ganglia-thalamocortical circuit of stuttering speakers. *Cortex* 46:49–67

Lukacs A, Kemeny F, Lum JAG, Ullman MT. 2017. Learning and overnight retention in declarative memory in specific language impairment. *PLOS ONE* 12:e0169474

Lum JA, Conti-Ramsden G, Page D, Ullman MT. 2012. Working, declarative and procedural memory in specific language impairment. *Cortex* 48:1138–54

Lum JA, Ullman MT, Conti-Ramsden G. 2013. Procedural learning is impaired in dyslexia: evidence from a meta-analysis of serial reaction time studies. *Res. Dev. Disabil.* 34:3460–76

Lum JAG, Conti-Ramsden GM, Morgan AT, Ullman MT. 2014. Procedural learning deficits in specific language impairment (SLI): a meta-analysis of serial reaction time task performance. *Cortex* 51:1–10

Mayes AK, Reilly S, Morgan AT. 2015. Neural correlates of childhood language disorder: a systematic review. *Dev. Med. Child Neurol.* 57:706–17

Mayor-Dubois C, Zesiger P, Van der Linden M, Roulet-Perez E. 2014. Nondeclarative learning in children with specific language impairment: predicting regularities in the visuomotor, phonological, and cognitive domains. *Child Neuro-Psychol.* 20:14–22

McClelland JL, Mirman D, Holt LL. 2006. Are there interactive processes in speech perception? *Trends Cogn. Sci.* 10:363–69

McGregor KK. 2009. Semantic deficits across populations. In *Handbook of Child Language Disorders*, ed. RG Schwartz, pp. 365–87. New York: Psychology Press

McMurray B, Munson C, Tomblin JB. 2014. Individual differences in language ability are related to variation in word recognition, not speech perception: evidence from eye-movements. *J. Speech Lang. Hear. Res.* 57:1344–62

Menghini D, Hagberg GE, Caltagirone C, Petrosini L, Vicari S. 2006. Implicit learning deficits in dyslexic adults: an fMRI study. *NeuroImage* 33:1218–26

Middleton FA, Strick PL. 1996. The temporal lobe is a target of output from the basal ganglia. *PNAS* 93:8683–87

Mitchell IJ, Cooper AJ, Griffiths MR. 1999. The selective vulnerability of striatopallidal neurons. *Prog. Neurobiol.* 59:691–719

Mortimer J, Rvachew S. 2010. A longitudinal investigation of morpho-syntax in children with speech sound disorders. *J. Commun. Disord.* 43:61–76

Newbury DF, Fisher SE, Monaco AP. 2010. Recent advances in the genetics of language impairment. *Genome Med.* 2:6

Nicolson RI, Fawcett AJ. 2007. Procedural learning difficulties: reuniting the developmental disorders? *Trends Neurosci.* 30:135–41

Packard MG. 2008. Neurobiology of procedural learning in animals. In *Concise Learning and Memory: The Editor's Selection*, ed. JH Byrne, pp. 341–56. London: Elsevier Sci.

Paulesu E, Danelli L, Berlingeri M. 2014. Reading the dyslexic brain: multiple dysfunctional routes revealed by a new meta-analysis of PET and fMRI activation studies. *Front. Hum. Neurosci.* 8:830

Penhune VB, Steele CJ. 2012. Parallel contributions of cerebellar, striatal and M1 mechanisms to motor sequence learning. *Behav. Brain Res.* 226:579–91

Pernet CR, Poline JB, Demonet JF, Rousselet GA. 2009. Brain classification reveals the right cerebellum as the best biomarker of dyslexia. *BMC Neurosci.* 10:67

Perrachione TK, Del Tufo SN, Gabrieli JD. 2011. Human voice recognition depends on language ability. *Science* 333:595

Peterson RL, Pennington BF, Shriberg LD, Boada R. 2009. What influences literacy outcome in children with speech sound disorder? *J. Speech Lang. Hear. Res.* 52:1175–88

Poldrack RA, Clark J, Pare-Blagoev EJ, Shohamy D, Creso Moyano J, et al. 2001. Interactive memory systems in the human brain. *Nature* 414:546–50

Poldrack RA, Packard MG. 2003. Competition among multiple memory systems: converging evidence from animal and human brain studies. *Neuropsychologia* 41:245–51

Pollard C, Sag IA. 1994. *Head-Driven Phrase Structure Grammar*. Chicago, IL: Univ. Chicago Press

Porfert AR, Rosenfield DB. 1978. Prevalence of stuttering. *J. Neurol. Neurosurg. Psychiatry* 41:954–56

Przybylski L, Bedoin N, Krifi-Papoz S, Herbillon V, Roch D, et al. 2013. Rhythmic auditory stimulation influences syntactic processing in children with developmental language disorders. *Neuropsychology* 27:121

Purdy JD, Leonard LB, Weber-Fox C, Kaganovich N. 2014. Decreased sensitivity to long-distance dependencies in children with a history of specific language impairment: electrophysiological evidence. *J. Speech Lang. Hear. Res.* 57:1040–59

Ramus F, Rosen S, Dakin SC, Day BL, Castellote JM, et al. 2003. Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain* 126:841–65

Remillard G. 2008. Implicit learning of second-, third-, and fourth-order adjacent and nonadjacent sequential dependencies. *Q. J. Exp. Psychol.* 61:400–24

Richlan F, Kronbichler M, Wimmer H. 2011. Meta-analyzing brain dysfunctions in dyslexic children and adults. *Neuroimage* 56:1735–42

Rose M, Haider H, Salari N, Büchel C. 2011. Functional dissociation of hippocampal mechanism during implicit learning based on the domain of associations. *J. Neurosci.* 31:13739–45

Scarborough HS. 1991. Early syntactic development of dyslexic children. *Ann. Dyslexia* 41:207–20
Schendan HE, Searl M, Melrose R, Stern C. 2003. An fMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. *Neuron* 37:1013–25
Schreiner T, Rasch B. 2017. The beneficial role of memory reactivation for language learning during sleep: a review. *Brain Lang.* 167:94–105
Scimeca JM, Badre D. 2012. Striatal contributions to declarative memory retrieval. *Neuron* 75:380–92
Seger CA, Miller EK. 2010. Category learning in the brain. *Annu. Rev. Neurosci.* 33:203–19
Seger CA, Peterson EJ, Cincotta CM, Lopez-Paniagua D, Anderson CW. 2010. Dissociating the contributions of independent corticostriatal systems to visual categorization learning through the use of reinforcement learning modeling and Granger causality modeling. *NeuroImage* 50:644–56
Serniclaes W. 2011. Allophonic perception in dyslexia: an overview. *Escr. Psicol.* 4:25–34
Shohamy D, Myers CE, Kalanithi J, Gluck MA. 2008. Basal ganglia and dopamine contributions to probabilistic category learning. *Neurosci. Biobehav. Rev.* 32:219–36
Shriberg LD, Green JR, Campbell TF, McSweeny JL, Scheer AR. 2003. A diagnostic marker for childhood apraxia of speech: the coefficient of variation ratio. *Clin. Linguist. Phonet.* 17:575–95
Singh L, Reznick JS, Xuehua L. 2012. Infant word segmentation and childhood vocabulary development: a longitudinal analysis. *Dev. Sci.* 15:482–95
Smith A, Weber C. 2017. How stuttering develops: the multifactorial dynamic pathways theory. *J. Speech Lang. Hear. Res.* 60:2483–505
Smith A, Zelaznik HN. 2004. Development of functional synergies for speech motor coordination in childhood and adolescence. *Dev. Psychobiol.* 45:22–33
Smith-Spark JH, Fisk JE. 2007. Working memory functioning in developmental dyslexia. *Memory* 15:34–56
Smits-Bandstra S, De Nil LF. 2007. Sequence skill learning in persons who stutter: implications for cortico-striato-thalamo-cortical dysfunction. *J. Fluen. Disord.* 32:251–78
Tagarelli KM, Shattuck KF, Turkeltaub PE, Ullman MT. 2019. Language learning in the adult brain: a neuroanatomical meta-analysis of lexical and grammatical learning. *NeuroImage* 193:178–200
Takács Á, Kóbor A, Chezan J, Éltető N, Tárnok Z, et al. 2018. Is procedural memory enhanced in Tourette syndrome? Evidence from a sequence learning task. *Cortex* 100:84–94
Terband H, Maassen B, van Lieshout P, Nijland L. 2011. Stability and composition of functional synergies for speech movements in children with developmental speech disorders. *J. Commun. Disord.* 44:59–74
Terband H, Van Brenk F, van Doornik-van der Zee A. 2014. Auditory feedback perturbation in children with developmental speech sound disorders. *J. Commun. Disord.* 51:64–77
Thaut MH, McIntosh KW, McIntosh GC, Hoemberg V. 2001. Auditory rhythmicity enhances movement and speech motor control in patients with Parkinson's disease. *Funct. Neurol.* 16:163–72
Thomas KM, Hunt RH, Vizueta N, Sommer T, Durston S, et al. 2004. Evidence of developmental differences in implicit sequence learning: an fMRI study of children and adults. *J. Cogn. Neurosci.* 16:1339–51
Tkach JA, Chen X, Freebairn LA, Schmithorst VJ, Holland SK, Lewis BA. 2011. Neural correlates of phonological processing in speech sound disorder: a resonance imaging study. *Brain Lang.* 119:42–49
Tourville JA, Guenther FH. 2011. The DIVA model: a neural theory of speech acquisition and production. *Lang. Cogn. Process.* 26:952–81
Tricomi E, Delgado MR, McCandliss BD, McClelland JL, Fiez JA. 2006. Performance feedback drives caudate activation in a phonological learning task. *J. Cogn. Neurosci.* 18:1029–43
Turk-Browne NB, Scholl BJ, Chun MM, Johnson MK. 2009. Neural evidence of statistical learning: efficient detection of visual regularities without awareness. *J. Cogn. Neurosci.* 21:1934–45
Ullman MT. 2004. Contributions of memory circuits to language: the declarative/procedural model. *Cognition* 92:231–70
Ullman MT. 2015. The declarative/procedural model: a neurobiologically motivated theory of first and second language. In *Theories in Second Language Acquisition*, ed. B VanPatten, J Williams, pp. 135–58. New York: Routledge
Ullman MT. 2016. The declarative/procedural model: a neurobiological model of language learning, knowledge, and use. In *Neurobiology of Language*, ed. G Hickok, S Small, pp. 953–68. New York: Elsevier

Ullman MT, Pierpont EI. 2005. Specific language impairment is not specific to language: the procedural deficit hypothesis. *Cortex* 41:399–433

Ullman MT, Pullman MY. 2015. A compensatory role for declarative memory in neurodevelopmental disorders. *Neurosci. Biobehav. Rev.* 51:205–22

Vakil E, Bloch A, Cohen H. 2017. Anticipation measures of sequence learning: manual versus oculomotor versions of the serial reaction time task. *Q. J. Exp. Psychol.* 70:579–89

Van der Linden M, Meulemans T, Marczewski P, Collette F. 2000. The relationships between episodic memory, working memory, and executive functions: the contribution of the prefrontal cortex. *Psychol. Belg.* 40:275–97

van Witteloostuijn M, Boersma P, Wijnen F, Rispens J. 2017. Visual artificial grammar learning in dyslexia: a meta-analysis. *Res. Dev. Disabil.* 70:126–37

Waldschmidt JG, Ashby FG. 2011. Cortical and striatal contributions to automaticity in information-integration categorization. *NeuroImage* 56:1791–802

Walenski M, Europa E, Caplan D, Thompson CK. 2019. Neural networks for sentence comprehension and production: an ALE-based meta-analysis of neuroimaging studies. *Hum. Brain Mapp.* 40:2275–304

Walenski M, Mostofsky SH, Ullman MT. 2007. Speeded processing of grammar and tool knowledge in Tourette's syndrome. *Neuropsychologia* 45:2447–60

Walenski M, Mostofsky SH, Ullman MT. 2014. Inflectional morphology in high-functioning autism: evidence for speeded grammatical processing. *Res. Autism Spectr. Disord.* 8:1607–21

Walenski M, Tager-Flusberg H, Ullman MT. 2006. Language in autism. In *Understanding Autism: From Basic Neuroscience to Treatment*, ed. SO Moldin, JLR Rubenstein, pp. 175–204. Boca Raton, FL: Taylor & Francis

Walenski M, Weickert TW, Maloof CJ, Ullman MT. 2010. Grammatical processing in schizophrenia: evidence from morphology. *Neuropsychologia* 48:262–69

Welsh MC, Pennington BF, Groisser DB. 1991. A normative-developmental study of executive function: a window on prefrontal function in children. *Dev. Neuropsychol.* 7:131–49

Wieland EA, McAuley JD, Dilley LC, Chang SE. 2015. Evidence for a rhythm perception deficit in children who stutter. *Brain Lang.* 144:26–34

Wilkinson L, Tai YF, Lin CS, Lagnado DA, Brooks DJ, et al. 2014. Probabilistic classification learning with corrective feedback is associated with in vivo striatal dopamine release in the ventral striatum, while learning without feedback is not. *Hum. Brain Mapp.* 35:5106–15

Yairi E, Ambrose NG. 1999. Early childhood stuttering I: persistency and recovery rates. *J. Speech Lang. Hear. Res.* 42:1097–112

Yi HG, Maddox WT, Mumford JA, Chandrasekaran B. 2014. The role of corticostriatal systems in speech category learning. *Cereb. Cortex* 26:1409–20

Wang LC, Tasi HJ, Yang HM. 2012. Cognitive inhibition in students with and without dyslexia and dyscalculia. *Res. Dev. Disabil.* 33:1453–61

Zhang Y, Kuhl PK, Imada T, Kotani M, Tohkura YI. 2005. Effects of language experience: neural commitment to language-specific auditory patterns. *NeuroImage* 26:703–20

Ziegler JC, Pech-Georgel C, George F, Alario FX, Lorenzi C. 2005. Deficits in speech perception predict language learning impairment. *PNAS* 102:14110–15

Annual Review of Psychology

Volume 71, 2020

Contents

Indexes

Errata

An online log of corrections to *Annual Review of Psychology* articles may be found at http://www.annualreviews.org/errata/psych