HHS Public Access

Reading and the Neurocognitive Bases of Statistical Learning[1]

Oliver M. Sawi and **Jay G. Rueckl**

University of Connecticut & Haskins Laboratories

Abstract

The processes underlying word reading are shaped by statistical properties of the writing system. According to some theoretical perspectives (e.g. Harm & Seidenberg, 2004) reading acquisition should be understood as an exercise in statistical learning. *Statistical learning* (SL) involves the extraction of organizing principles from a set of inputs. Several lines of research provide convergent evidence supporting the connection between SL and reading acquisition (e.g., Arciuli & Simpson, 2012; Frost et al., 2014; Bogaerts et al., 2015). An obstacle to fully appreciating the theoretical and educational implications of these findings is that SL is itself not well understood. In this paper, we review the current literature on SL with a particular focus on organizing this literature by grounding it in theories of learning and memory more generally. This approach can clarify the nature of SL and provide a framework for understanding its role in reading, reading acquisition, and reading disorders.

Introduction

Over the last several decades, reading research has become increasingly grounded in the idea that a writing system can be characterized by the statistical regularities embodied in the mappings between the ortho- graphic, phonological, and semantic properties of printed words. For example, in English the mapping between written and spoken words is quasiregular—orthographic units such as letters and word bodies tend to represent specific phonological units (e.g., phonemes, rimes), but these regularities are rarely completely reliable. One way to account for how readers cope with challenges imposed by quasiregularity is to posit that reading is driven by knowledge of the statistical properties of the writing system, and hence of reading acquisition as an exercise in statistical learning (SL; Harm & Seidenberg, 2004); Seidenberg & McClelland, 1989). Although the statistical approach to reading was initially rooted in accounts of phonological decoding, over time it has been applied to a broad range of topics, including the representation of orthographic (Lerner, Armstrong, & Frost, 2014) and morphological (Rueckl, 2010; Seidenberg & Gonnerman, 2000) structure, cross-language differences (Frost, 2012; Seidenberg, 2011), reading acquisition (Treiman, Kessler, & Bick, 2003), and developmental dyslexia (Harm & Seidenberg, 1999).

In parallel to these developments in the science of reading, broader interest in the mechanisms underlying sensitivity to environmental regularities gave rise to a distinct and largely independent body of research. The primary goal of this research is to elucidate the neurocognitive processes responsible for SL, the extraction of the organizing principles or regularities from a set of inputs. For example, in studies that are both representative of and

seminal to this line of research, Saffran, Aslin, and Newport (1996a), Saffran, Newport, and Aslin (1996b) presented participants with sequences of spoken syllables structured such that any given syllable was more likely to be followed by some syllables than by others. Their results revealed that both adults (Saffran et al., 1996a) and infants (Saffran et al., 1996b) acquired knowledge about the statistical structure of these sequences. As reviewed next, the SL literature is now rather expansive, covering learning in a variety of circumstances and in a number of populations (see Siegelman, Bogaerts, Christiansen, & Frost, 2017; Thiessen, Kronstein, & Hufnagle, 2015, for reviews).

One focus of this expansive literature has been the relationship between SL and language processing. Researchers have investigated the role of SL in a variety of linguistic domains, including phonological learning (Maye, Werker, & Gerken, 2002), word segmentation (e.g., Saffran et al., 1996a, 1996b), early vocabulary development (Shafto, Conway, Field, & Houston, 2012), and lexical access (Mainela-Arnold & Evans, 2014). Individual differences in SL have been shown to predict performance in psycholinguistic tasks involving speech perception in noise (Conway, Bauernschmidt, Huang, & Pisoni, 2010), the processing of complex sentences (Misyak, 2010; Misyak & Christiansen, 2012), and syntactic comprehension (Kidd & Arciuli, 2016). Neuroimaging evidence suggests that language processing and SL activate overlapping cortical regions (Conway & Pisoni, 2008; Folia et al., 2008; Petersson, Folia, & Hagoort, 2012).

Of particular relevance to this special issue is the fact that SL has also been linked to reading and reading acquisition. For example, in preliterate children, SL ability is associated with skills that are predictive of early literacy achievement (e.g., oral language, vocabulary, and phonological processing) in children (Spencer, Kaschak, Jones, & Lonigan, 2015) and reading ability in typically developing children and adults (Arciuli & Simpson, 2012). In addition, in a study of English-speaking young adults learning Hebrew, performance on an SL task was found to predict the changes on several measures of Hebrew reading (Frost, Siegelman, Narkiss, & Afek, 2013). Important to note, in addition to demonstrating an association between SL ability and reading acquisition generally, some findings suggest a link between SL and developmental dyslexia. In particular, a number of studies have reported that individuals with dyslexia have significantly lower SL scores, relative to typically developing individuals, across a variety of tasks measuring sensitivity to statistical structure (e.g., Bogaerts, Szmalec, Hachmann, Page, & Duyck, 2015; Hachmann et al., 2014; Jiménez-Fernández, Vaquero, Jiménez, & Defior, 2010; Lum, Ullman, & Conti-Ramsden, 2013); Menghini, Hagberg, Caltagirone, Petrosini, & Vicari, 2006; Menghini et al., 2008; Stoodley, Harrison, & Stein, 2006; Stoodley, Ray, Jack, & Stein, 2008; Vicari, 2005; Vicari, Marotta, Menghini, Molinari, & Petrosini, 2003).

It is important to note, however, that not all findings support a link between SL and reading. For example, Nigro, Jiménez-Fernández, Simpson, and Defior (2015) failed to find a correlation between SL and several reading measures in a sample of young children learning to read their native Spanish. In addition, a number of studies have failed to find group-level differences between typically developing individuals and individuals with dyslexia (e.g., Bussy et al., 2011; Deroost et al., 2010; Gabay, Schiff, & Vakil, 2012; Kelly, Griffiths, & Frith, 2002; Menghini et al., 2010; Yang & Hong-Yan, 2011). Other studies have reported

mixed effects, with the presence or absence of a group difference contingent on methodological factors such as the sequence structure or the characteristics of the stimuli (Henderson & Warmington, 2017; Howard, Howard, Japikse, & Eden, 2006; Jimenez-Fernandez et al., 2010). This pattern of mixed findings is reflected in the conclusions of two recent meta-analyses. Lum et al. (2013) found a statistically significant group difference (between typically developing readers and individuals with dyslexia) that was modulated by factors such as age and test condition. In contrast, although Schmalz, Altoè, and Mulatti (2016, p. 1) also found a statistically significant (albeit small) effect in a meta-analysis of a different set of findings (including additional forms of SL), they concluded that "there is insufficient high-quality data to draw conclusions about the presence or absence of an effect."

It is likely that this inconsistent pattern of findings is in part due to methodological factors. For example, as Schmalz et al. (2016) noted, the criteria for classifying participants as dyslexic often differ from study to study. Moreover, Siegelman and Frost (2015) have observed that the test–retest reliability of the SL tasks used in these studies can vary widely and is often quite low.

This being said, it is also possible that the inconsistent pattern of results is theoretically informative. The question addressed by the aforementioned studies can be (and often is) simplified to "Does statistical learning ability predict reading ability?" However, neither reading nor SL are monolithic skills. Rather, both are fundamentally componential. Tasks of both sorts require the coordinated engagement of an ensemble of distinct neurocognitive processes. Therefore, one might expect that whether a relationship between SL and reading is found depends on which tasks are used to assess each skill and whether there is overlap in the component processes engaged by each task. The primary focus of the present review is the componential nature of SL. In the next sections, we review theoretical advances that are beginning to illuminate the componential structure of SL and discuss the relationship between the component processes of SL and key theoretical distinctions drawn in the broader literature on learning and memory. In the final section, we discuss the implications of this componential approach to SL for research on reading and reading acquisition.

Statistical learning

SL involves the extraction of statistical regularities from the environment. In principle, this definition subsumes the learning of many types of regularities under a variety of circumstances. In practice, however, the definition of SL has often been (implicitly) linked to the tasks used to measure it. As just noted, in their landmark studies Saffran et al. (1996a, 1996b) had participants listen to a series of syllables that varied in transitional probability and then tested whether the participants' behavior was shaped by these probabilities. Given the seminal importance of these studies in drawing the field's attention to SL as a phenomenon, it is not surprising that their experimental method has come to be understood as the canonical SL paradigm (we refer to this task and its variants as the canonical SL task hereafter). Important to note, however, although this paradigm and its variants (cf. Endress & Mehler, 2009; Newport & Aslin,2004; Siegelman & Frost 2015), are representative of how SL might be measured, SL should be understood as a neurocognitive process (or set of

processes) responsible for the extraction of environ- mental regularities. Indeed, because the publication of the Saffran et al. (1996a, 1996b) studies, a number of experimental tasks have come to be understood as "statistical learning" tasks. The characteristics of these tasks both help to clarify the range of relevant phenomena and suggest possible avenues for investigating the nature of the processes that underlie SL.

Like the canonical SL task, several of the other tasks often used to study SL also involve the learning of sequential structure. One such task involves artificial grammar learning (e.g., Dienes, Broadbent, & Berry, 1991; Gomez & Gerken, 1999). This paradigm is similar to the canonical SL task except that the sequences presented to the participants are determined by a finite-state grammar. Of interest, although the artificial grammar and canonical SL paradigms are closely related methodologically, they stem from different research traditions and are often used to address different theoretical issues (see Perruchet & Pacton, 2006, for discussion). Another task that targets the learning of sequential statistics is the serial reaction time task (SRTT). The SRTT is a choice reaction-time task in which participants repeatedly respond to a small set of visual cues, typically by pressing a button paired with each cue. The sequence of cues is structured such that a particular cue is at least somewhat predictable on the basis of the previous cue or series of cues (Nissen & Bullemer, 1987; Robertson, 2007, Siegelman & Frost 2015). Finally, one other sequential learning task sometimes used to study SL is the Hebb repetition task (e.g., Bogaerts et al., 2015; Page, Cumming, Norris, Hitch, & Mcneil, 2006), a serial recall task in which participants hear or see a series of stimuli (e.g., syllables, digits) and attempt to recall the stimuli in order immediately after the presentation of the sequence. Recall typically improves with repeated presentation of the same sequence, indicating that participants have acquired knowledge of sequential order.

Although each of the tasks just described tracks the learning of sequential structure, it is important to note that SL encompasses learning of other kinds of regularities as well. For example, SL can involve the detection of regularities in the spatial relationships among stimuli rather than in their sequential order. One method that has been employed to investigate this aspect of SL is contextual cueing paradigm (e.g., Chun & Jiang, 1998; Goujon, Didierjean, & Thorpe, 2015). In this paradigm, partici- pants search for the presence of a visual target within a configuration of distractors. Learning is revealed by a reduction in search time for targets in repeated configurations relative to targets in novel configurations.

Not all SL tasks involve the extraction of regularities involving the sequential or spatial relationships among stimuli. Another form of SL involves the extraction of distributional statistics about the frequency and variability of exemplars in the input (cf. Maye et al., 2002; Smith & Yu, 2008). For example, Maye et al. (2002) exposed infants to two continua of stimuli (/da/-/ta/) with either a unimodal (single peak at the intermediate point between/da/and/ta/) or bimodal distribution (peaks at prototypical/da/and/ta/). Infants exposed to a continuum with a bimodal distribution were able to successfully discriminate between /da/ and /ta/ stimuli, suggesting they were able to generate two separate categories, whereas infants exposed to a unimodal distribution were not.

It is worth noting that given the differences among the tasks just described and the methodological variants afforded by each task, studies of SL vary on a number of important dimensions, including not only the type of statistical information to be learned (e.g., conditional or distributional statistics, spatial or sequential regularities, adjacent or nonadjacent dependences), but also the modality of the stimuli, the tasks used to measure learning, and the degree to which participants are explicitly directed toward the to- be-learned regularities during encoding or at test. Understanding the nuances in the methodological similarities and differences among these studies is particularly important for understanding SL as a construct.

Towards a Componential View of Statistical Learning

Many early descriptions of SL typically assumed that SL was a unitary, domain-general learning mechanism or capacity (Kirkham, Slemmer, & Johnson, 2002; Saffran, 2003). Siegelman et al. (2017) noted that most examinations do not mention specific underlying computations or mechanisms but rather a more abstracted system in which a unified capacity is controlled by a single learning system across all domains. However, recent evidence suggests that SL is in fact componential (see Frost, Armstrong, Siegelman, & Christiansen, 2015, for discussion). Therefore, a more nuanced understanding of the nature of SL is needed to help appreciate the potential differences underlying computations supporting various aspects of the tasks described.

There is particularly strong evidence regarding modality-specific components in SL. For example, in artificial grammar learning there is no cross-modality interference but strong intermodality interference (Conway & Christiansen, 2006) and learning does not transfer across modalities (Redington & Chater, 1996), suggesting that learning produces representations that are specific to the stimulus properties present in auditory, visual, and "tactile" (i.e., motor involvement) modalities (Conway & Christiansen, 2005, 2006). In addition, individual differences in learning on one SL task sometimes fail to predict learning on another SL task (e.g., SRTT and Hebb Repetition; Henderson & Warmington, 2017), and even the correlation across variants of the canonical SL task can be quite low (Siegelman & Frost, 2015).

In light of these findings, Frost et al. (2015) proposed a theoretical framework that holds that the mechanisms underlying SL are a set of interrelated, modality-specific processes. Thus, which brain regions are activated during a particular SL task is contingent on the modality of the stimuli presented during that task as well as other task demands. For example, SL tasks involving sequences of spoken syllables engage the inferior frontal gyrus and left temporal gyrus (Alba & Okanoya, 2008; Karuza et al., 2013), regions associated with speech perception more broadly (Hickok & Poeppel, 2007). Similarly, visual networks are activated by SL tasks involving sequences of visual stimuli (Bishoff-Grethe, Proper, Mao, Daniels, & Berns, 2000; Turk-Browne, Scholl, Chun, & Johnson, 2009) and motor regions (e.g., motor cortex, the cerebellum) are activated during the SRTT (Packard & Knowlton, 2002). Therefore, at least one aspect of the componentiality of SL involves the role of early, modality-specific processes (Frost et al., 2015).

However, the componential nature of SL is not simply driven by modality-specific constraints. Arciuli (2017) argued that SL draws on component processes related to the encoding, retention, and abstraction of statistical regularities. For example, older children performed better than younger children on an SL task, separate from differences in attention (Arciuli & Simpson, 2012). They posited that an implicit form of working memory is an underlying component of SL that is late developing, contributing to these age-related differences.

Turning to components related to the abstraction of statistical regularities, Arciuli (2017) asserted that inconsistent findings across studies, comparing SL performance in individuals with autism spectrum disorder (ASD) and typically developing individuals, may be due to the nature of the statistical regularities investigated. For example, in two studies investigating differences in task performance in SL with sequential regularities, there were no differences between groups (Brown, Aczel, Jiménez, Kaufman, & Grant, 2010; Mayo & Eigsti, 2012). However, a separate study found that individuals with ASD had superior task performance (relative to typically developing individuals) in SL with spatial regularities. The inconsistent findings in the connection between ASD and SL suggest that sensitivity to sequential and spatial regularities may be supported by different component processes (Arciuli, 2017). In addition, recent evidence suggests that several measures of SL do not correlate, even within modality. Siegelman and Frost (2015) examined individual differences in four versions of the canonical SL task that differed with regard to whether the stimuli were verbal or nonverbal stimuli and whether the sequences embodied adjacent or nonadjacent regularities. The weak and generally nonsignificant correlations between these measures suggest that they are supported by separate underlying component processes.

Further, Thiessen and Erickson (2013, 2015) modeled the underlying processes supporting sensitivity to multiple forms of statistical information across modality. In their model, conditional statistics and distributional statistics are modeled by different underlying computational, memory-based systems. However, these systems are also linked, as the output of computations related to conditional statistics (extraction) provide the input for processes involved in the computation of distributional statistics (integration).

Although there is strong evidence for modality-specific components of SL, there is also evidence that domain-general processes contribute to the learning of statistical regularities. These domain-general principles emerge in two ways (Frost et al., 2015). First, across modality similar computations are engaged to pull out statistical regularities in the input stream (as modeled by Thiessen & Erickson, 2013; Thiessen et al., 2015). Second, modality-specific information generated during initial encoding is further processed in multimodal regions. Information across all domains is therefore processed in the same brain networks and may be subject to similar processing demands. Specifically, these multimodal processing regions include aspects of the frontal (Alba & Okanoya, 2008; Karuza et al., 2013), striatal (Turk-Browne et al., 2009) and Medial Temporal Lobe (MTL) memory systems (Schapiro & Turk- Browne,2015; Turk-Browne et al., 2009).

In summary, recent advances in SL suggest that it is a componential construct. SL involves both modality-specific and domain-general processes subserved by a number of brain

regions differentially involved in the encoding, retention, and abstraction of statistical regularities. Although several recent accounts have addressed the componential character of SL (e.g. Arciuli, 2017; Frost et al., 2015), questions remain regarding, for example, the similarities and differences in SL across and within modality, the role of multimodal processing systems such as the MTL and striatum, and the developmental trajectories of the components of SL. One strategy for addressing these questions is to turn to the broader literature on learning and memory, where similar questions have been addressed in relation to phenomena beyond the domain of SL. As we discuss in the next section, grounding theories of SL in well-established memory theories, particularly those that draw contrasts between distinct underlying memory subsystems, can provide valuable insights into the componential nature of SL.

Statistical Learning & Multiple Memory Systems Frameworks

Memory comprises several distinguishable component processes evolved to support different types of information (Schacter, 1987). For example, one may have memories associated with specific events (e.g., one's birthday this year) and facts (e.g., the capitol of one's state) to which one has conscious awareness. On the other hand, one may have memories associated with skills or habits to which one does not have conscious awareness (e.g., riding a bike).

There are several frameworks used to dichotomize memory of these types. For example, declarative and procedural memory (e.g., Squire, 1992, 2004; Ullman, 2004) are typically characterized by dependence on specific anatomical regions such as the MTL or the striatum and neocortex, respectively (e.g., Squire, 1992, 2004). In addition, declarative memory is typically associated with memories to which individuals have conscious access such as memories of facts or events. Procedural memory refers to memories to which individuals do not have conscious access such as skills and habits (e.g., Squire, 1992, 2004). Learning in declarative memory occurs with conscious intention, whereas learning in procedural memory occurs over time without direct conscious awareness or intention. Declarative memory is also responsible for the learning of arbitrary relationships (associative binding) over short periods and is domain general (Cohen, Poldrack, & Eichenbaum, 1997; Eichenbaum & Cohen, 2001; Squire & Knowlton, 2000). Learning in procedural memory is related to understanding of relationships between complex sequences (sensorimotor or cognitive) over extended periods and is modality specific (Squire & Knowlton, 2000; Ullman, 2004).

A related and overlapping dichotomy is the distinction between implicit and explicit memory (e.g., Schacter, 1987). This dichotomy differs from the declarative/procedural distinction in relative emphasis on whether memory retrieval occurs intentionally or incidentally. Explicit memory is marked by the inten- tional and conscious retrieval of information and is typically measured by "direct" means such as recall or recognition. Implicit memory, on the other hand, involves the incidental and unconscious retrieval of information and is indexed by "indirect" measures such as repetition priming or skill acquisition. Explicit and implicit memory systems map onto similar neural correlates as declarative and procedural memory systems. The explicit memory system has been characterized to rely on the MTL system, whereas the implicit memory system seems to rely

on a circuit including frontal-striatal, as well as cortico-cortical, connections (Dew & Cabeza, 2011; Voss & Paller, 2008).

There is a related dichotomy examining explicit and implicit learning (Reber, 1992; Reber, Gitelman, Parrish, & Mesulam, 2003). Whereas the distinction between implicit and explicit memory is primarily a matter of whether the processes that occur at the time of retrieval involve the conscious intention to recollect (Schacter, 1987), the distinction between implicit and explicit learning is primarily a matter of whether the processes involved in the initial encoding and storage of information occur are intentionally engaged and whether the resulting knowledge is available to conscious awareness (e.g., Perruchet & Pacton, 2006; Reber, 2013). Thus, implicit learning is incidental and typically occurs with extended practice, whereas explicit learning is deliberate and often occurs on the basis of a single event.

Although each of the preceding contrasts is conceptually distinct, there is clearly much overlap in these theories. Therefore, although the nuanced differences between these contrasts are of importance in some contexts, for the purposes of the current review we use the terms "Implicit/Procedural Memory" (IPM) to refer to Procedural Memory and Implicit Memory and Learning and "Explicit/Declarative Memory" (EDM) to refer to Declarative Memory and Explicit Memory and Learning.

Although the conceptual distinctions made by the various multiple-memory-systems theories are clear, isolating the influence of each system has proven to be rather challenging. One issue is that although tasks are frequently described as indices of a particular type of memory (e.g., a "procedural task" or an "explicit memory task"), these tasks rarely index the operation of a single memory system in isolation (Dew & Cabeza, 2011; Voss & Paller, 2008). Although a number of experimental strategies have been devised to address this issue (e.g., Jacoby, 1991; Schacter, 1987), the lack of a straightforward one-to-one mapping between memory systems and experimental tasks complicates the interpretation of any experimental finding. A second issue is that when EDM and IPM are both engaged, they may interact in ways that depend on factors such as task demands, maturation, and the time-course of learning (Ullman, 2004; Wagner, Maril, & Schacter, 2000). For example, Poldrack et al. (2001) found that the activation of MTL and striatal systems in several memory tasks is negatively correlated across participants, suggesting that that EDM and IPM networks actually compete during the learning process. Relatedly, in a speech-category learning experiment, Yi, Maddox, Mumford, and Chandrasekaran (2014) found that participants used strategies associated with EDM early in training, with a gradual shift towards strategies associated with IPM.

EDM and IPM also have a degree of interactivity, particularly in their neural correlates. For example, IPM also engages MTL (Rose, Haider, Weiller, & Büchel, 2002; Schendan, Searl, Melrose, & Stern, 2003). Schendan et al. (2003) suggested that the mid-MTL (hippocampus) is involved in sequence learning in both EDM and IPM but engage anterior and posterior regions, respectively. Several theoretical frameworks explain MTL involvement in IPM. For example, Shohamy and Turk-Browne (2013) suggested that the hippocampus is highly connected to most cortices, including temporal, DLPFC, and the striatum (Goldman-Rakic,

Selemon, & Schwartz, 1984; Shohamy & Adcock, 2010; Suzuki & Amaral, 1994) and is involved in most behavioral functions. In light of the widespread connectivity of the MTL, Shohamy and Turk-Browne (2013) suggested that the hippocampus may exert direct control over the nature of the cognitive representations or modulate cognitive function. In this way MTL is involved in various processing streams.

The memory theories just discussed provide a useful framework for understanding SL. For example, component processes underlying memory (e.g., EDM, IPM) may also contribute to aspects of SL. Further, exploring the interactivity of these previously dichotomized systems can help us better understand the function of SL. For example, one can look into how the underlying neurobiology supporting SL may in fact shift during the learning process to rely on different memory systems or have differential activation patterns.

Is Statistical Learning Supported by a Specific Memory System?

Recent studies have focused on specific connections between SL and multiple memory systems. In fact, artificial grammar learning and SRTT, although also SL tasks, are seen as canonical measures of implicit and procedural learning respectively. Many conceptions of SL (see Newport & Aslin, 2004; Conway & Christiansen, 2006; Perruchet & Pacton, 2006; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997) assume it is a strictly IPM process. For example, some theories suggest that SL and IPM measures tap into the same mechanism (Perruchet & Pacton, 2006; Thiessen, 2017; Thiessen & Erickson, 2013; Thiessen et al., 2015) or SL is simply a subset of IPM (Conway & Christiansen, 2006). Various lines of research support this assertion. For example, most SL tasks do not give explicit instructions (e.g., Newport & Aslin, 2004; Saffran et al., 1997), and many individuals are not consciously aware of the patterns in the input (e.g., Turk-Browne et al., 2009). In addition, Saffran et al. (1997) found that SL occurred in infants, in the auditory domain, even when they were distracted by a concurrent drawing task, suggesting that SL can occur without direct attention or intention. Last, SL activates aspects of the frontal (Alba & Okanoya, 2008; Karuza et al., 2013) and striatal (Turk-Browne et al., 2009) learning systems (IPM).

Although there is strong evidence linking SL and IPM, recent findings suggest SL is also supported by EDM. One line of evidence involves the neural correlates of SL. For example, the MTL network has been implicated in SL across modalities (Schapiro & Turk-Browne, 2015; Turk-Browne et al., 2009). This suggests that SL is supported by multiple memory systems. Gómez (2017) examined this possibility from a developmental perspective. For example, adults retain statistical patterns learned after a single exposure, even after a 24-hr period (Durrant, Cairney, & Lewis, 2012; Durrant, Taylor, Cairney, & Lewis, 2011; Kim, Seitz, Feenstra, & Shams, 2009). However, infants display "fragile" overnight retention up to 15 months (Gómez, 2017; Simon et al., 2016). In addition, infant retention of statistical regularities takes repeated exposure and, crucially, seems to be related to learning processes in the neocortex and striatal networks (Gomez & Edgin, 2016). However, once the hippocampal learning system is online, individuals are able to quickly consolidate the statistical information. Further, in adults, hippocampal activity is important for the consolidation of memories overnight (Marshall & Born, 2007). However, before 2 years of age, the necessary connections (both within hippocampus and from hippocampus to

prefrontal cortex) have not matured and cannot support consolidation (Gómez & Edgin, 2016). Thus, evidence suggests that before the hippocampal-prefrontal cortex circuit is developed, SL mainly involves the neocortex and striatal networks, with the MTL (hippocampal) network developing more slowly.

A second line of evidence suggesting that SL is supported by both IPM and EDM involves the effect of task instructions on SL performance. Learning is facilitated by instructions encouraging participants to attend to regularities in the input, although the benefit of explicit instructions occurs only under some conditions and not others (cf. Arciuli, Torkildsen, Stevens, & Simpson, 2014; Batterink, Reber, Neville, & Paller, 2015; Dienes et al., 1991; Frensch & Miner, 1994; Gómez, 2017; Hamrick & Rebuschat, 2012; Jiménez, Méndez, & Cleeremans, 1996). For example, Witt, Puspitawati, and Vinter (2013) found that the effect of instructions on an artificial grammar learning task changes as a function of age (children ages 5–8). In their study, older children had significantly higher task performance than younger children when given explicit instructions but not implicit instructions. This suggests that older children were better able to engage EDM processes to improve performance, whereas younger participants were not, and both older and younger children engaged IPM processes to a similar degree (Witt et al., 2013).

Of interest, instructions appear to impact which brain regions are activated during an SL task. In a recent study of artificial grammar learning, Yang and Li (2012) observed similar patterns of activation in MTL and frontal-striatal networks regardless of whether participants were given implicit or explicit instructions. However, there was greater activation of the caudate, a structure association with IPM, in the implicit condition than in the explicit condition. In contrast, explicit instructions led to greater activation in the precuneus, a structure associated with EDM.

Finally, the conclusion that SL is supported by both IPM and EDM is also supported by consideration of the measures used to index SL. For example, in a study employing the canonical SL task with auditory stimuli, Batterink et al. (2015) measured learning with an old/new recognition (direct) task and a syllable-detection (indirect) task. ERP data suggested that the recognition (direct) and syllable-detection (indirect) tasks elicited specific ERP components related to EDM (LPC) and IPM (P300), respectively. In an individual-differences analysis, these two measures of learning were not correlated even though most participants showed learning in both tasks. Batterink et al. (2015) asserted that these findings suggest both explicit and implicit representations may be developed concurrently.

In summary, SL is supported by aspects of both EDM and IPM. In light of these findings, under- standing SL in the context of multiple-memory-systems theories has interesting implications for elucidating the nuanced nature of the involvement of domain-general, multimodal brain regions in SL processing across development, for further expanding understanding of SL as a componential construct, and for theories of individual differences in SL. For example, SL is subject to both age-related (Gomez, 2017) and task-demand-related (Arciuli et al., 2014) shifts in the differential engagement of these memory systems.

Lastly, integration of SL into well-established theoretical frameworks allows for questions in SL to be guided by previously discovered phenomena. For example, SL can be affected by instructions (incidental/intentional divide in explicit and implicit memory literature), measurement (direct/indirect), and awareness of retrieval. This further expands investigation of SL to more explicitly examine the nature of the representations generated in SL and to look into online measures of learning (Siegelman et al., 2017). Understanding these additional dimensions and developmental shifts in engagement of multiple memory systems has implications for under- standing the componential nature of SL and the connection between SL and reading and reading acquisition.

Statistical Learning & Reading

From one prominent perspective, the organization of reading processes is shaped by the statistical structure of the writing system and learning to read is thus fundamentally a form of SL (Harm & Seidenberg, 2004; Rueckl, 2016). An open question is whether and how "statistical learning" in this context is related to learning in so-called statistical learning tasks such as the canonical SL task or SRTT. If a meaningful relationship does exist, reading scientists would be afforded new avenues of research for understanding the processes underlying reading acquisition, which could in turn inform the design of educational practices related to instruction, diagnosis, and intervention.

As noted earlier, although there is substantial evidence that reading acquisition is meaningfully related to the processes supporting learning in SL tasks, there are also a significant number of results supporting the opposite conclusion. In our view, this pattern of conflicting results is due in part to methodological issues such as the poor test–retest reliability of some SL tasks (see Siegelman, Bogaerts, & Frost, 2017), but it also reflects the componential nature of both SL and reading. At this point, further advances in the emerging understanding of SL as a componential process are needed before a comprehensive account of extant results can be provided. However, given the developments to date, it is possible to begin to lay out some speculative hypotheses.

First, becoming a skilled reader entails learning about a wide variety of statistical regularities. Some of these involve the relationships between different kinds of lexical properties (e.g., the mappings between orthographic, phonological, and semantic codes). Others involve statistical relationships within a given domain, including the clustering of features resulting in the formation orthographic units (e.g., letters), phonological segments (e.g., phonemes), and semantic concepts, as well as regularities in the sequential and spatial structure of spoken and written words (i.e., phonotactic and orthotactic regularities). These regularities vary in their reliability and are found at a variety of grains sizes (e.g., letters, bigrams, word bodies, phonemes, syllables, and so forth). Moreover, although they often involve adjacent elements, this is not always the case. (For example, some orthographic–phonological regularities in English involve letter clusters such as ph and ea, but others involve nonadjacent units such as a_e and i_e.). Finally, a large body of neuroimaging results has revealed that word reading engages a distributed network of cortical regions (Pugh et al., 2000; Rueckl et al., 2015) and that different kinds of statistical regularities are stored in

different parts of this network (Graves, Desai, Humphries, Seidenberg, & Binder, 2010; Taylor, Rastle, & Davis, 2013).

Given the variety of statistical regularities that must be learned, it is clear that the relationship between SL and reading could take a variety of forms. For example, if both SL and learning to read are primarily driven by the same domain-general mechanism, then individual differences on the performance of any SL task should be associated with almost any reading measure (assuming sufficiently reliable measures of both sorts). In contrast, if (as the emerging evidence suggests) variability on SL tasks reflects the operation of modality- or domain-specific learning mechanisms, then the relationship between SL and reading should depend on the characteristics of the tasks used to measure each skill. For example, SL and reading tasks that engage the same cortical regions should be associated. Similarly, task-specific relation- ships between SL and reading would be expected if different SL mechanisms underlie the learning of within- versus between-domain regularities, for example, if within-domain regularities are learned by processes specific to that domain, whereas the learning of between-domain regularities is mediated by domain-general processes involving MTL or frontal-striatal circuits.

A particularly interesting test case concerns the division of labor between phonological and lexical/semantic processes in word reading. The division of labor between these two reading "pathways" has been of long-standing theoretical interest (cf. Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Harm & Seidenberg, 2004), particularly with regard to individual differences (e.g., Baron & Strawson, 1976; Strain & Herdman, 1999; Woollams, Ralph, Madrid, & Patterson, 2016). In addition to mapping written forms to two different domains, the statistical regularities embodied in these mappings differ in several respects. In English, for example, the mapping from orthography to phonology embodies a more systematic set of regularities that generally involve smaller orthographic units. Thus, not only might we expect that the phonological and semantic pathways are linked to different domain-specific components of SL, but it is also plausible that they are differentially linked to putative domain-general SL mechanisms as well. For example, it has been proposed that the MTL plays a critical role in the learning of arbitrary associations (McClelland, McNaughton, & O'Reilly, 1995; Squire, 1992), suggesting that MTL mediation would be particularly important in learning along the semantic pathway. In contrast, given both the articulatory/ gestural grounding of phonological representations (Browman & Goldstein, 1989; Liberman & Mattingly, 1985) and the systematic structure of the orthographic-phonological mapping, it might by hypothesized that the procedural/frontal-striatal system is of particular importance in the learning of this mapping. Of interest, because the developmental trajectories of the MTL and frontal-striatal systems differ (see Gomez, 2017, for review), the contribution of these components of SL to reading acquisition might change over the course of the lifespan (cf. Ullman & Pierpont, 2005).

Similar considerations suggest that the relationship between SL and reading might vary across languages. For example, in recent treatments the orthographic depth hypothesis (Frost, Katz, & Bentin, 1987), a key organizing principle for understanding cross-language differences, has been recast in terms of the statistical properties of the writing system such that the division of labor between the phonological and semantic pathways is determined by

the relative reliability of ortho-phonological and ortho-semantic correspondences (Frost, 2012; Seidenberg, 2011). Thus, the relative importance of various SL component processes might be hypothesized to vary across languages along the lines discussed in the previous paragraph. Moreover, it is worth noting that writing systems differ on a number of other dimensions as well, including the number of unique orthographic characters, the visual complexity of these characters (Chang, Plaut, & Perfetti, 2016), and the relevance of regularities involving nonadjacent letters or phonemes. For example, in Hebrew, most words are formed by interleaving triconsonant root morphemes with morphologically informative word patterns. Consequently, morphological regularities involve patterns of nonadjacent letters or phonemes (Frost, 2012; Lerner et al., 2014). Each of these differences could give rise to language-specific differences in the relative importance of different SL processes in learning to read.

Finally, with regard to dyslexia, the componential approach to SL may help us make sense of the conflicting pattern of results just discussed. For example, Jiménez-Fernández et al. (2010) tested participants on implicit and explicit forms of an SL task and observed that, relative to typically developing controls, individuals with dyslexia were impaired on the implicit task but not the explicit task. Similarly, Howard et al. (2006) found that individuals with dyslexia performed relatively poorly on a serial response time task and relatively better on an SL task in which regularities in the spatial configuration of a visual display signaled the location of a target. Because these tasks are differentially associated with the procedural/frontal-striatal and declarative/MTL systems, respectively, these results were taken as evidence that dyslexia is associated with differences in the operation of the former and not the latter. It is important to note, however, that dyslexia is a heterogeneous condition associated with a variety of risk factors, and thus it is rather unlikely that all individuals with dyslexia would exhibit the same pattern of SL deficits (Schmalz et al., 2016). A more likely possibility is that (in at least some cases) the factors that give rise to reading disability are associated with differences in specific SL processes. These factors likely give rise to measurable differences in reading behavior as well (see Harm & Seidenberg, 1999, for discussion), raising the possibility that associations between SL measures indexing specific SL processes and reading measures indexing specific reading processes could be particularly informative in revealing why some children struggle to learn to read.

Although much remains to be done to strengthen our theoretical understanding of the relationship between reading and SL, it is perhaps worth considering some of the potential implications of such an understanding for educational practice and treatment. Here we suggest three. First, considerable effort has been dedicated to identifying different "types" of development dyslexia (Castles & Coltheart, 1993; Morris et al., 1998). As the preceding paragraph suggests, the systematic investigation of the relation between components of SL and variation in the kinds of reading deficits exhibited by children with dyslexia could both advance our understanding of the etiology of dyslexia and lead to the development of diagnostic tools using SL measures. Second, developing a deeper understanding of the relationship between reading and the implicit and explicit aspects of SL could guide the development of instructional practice, revealing, for example, whether explicit instruction is especially beneficial (or costly) for certain kinds of regularities, at certain points in the acquisition process, or for certain individuals. Third, an SL perspective on reading

acquisition could provide guidance on how to structure instructional materials to best promote reading achievement. For example, in a recent study Apfelbaum, Hazeltine, and McMurray (2013) demonstrated that the learning of grapheme–phoneme correspondences were best learned when the target grapheme–phoneme correspondences were embedded in more variable word contexts, as suggested by findings in the SL literature.

To conclude, we note that pursuing the research agenda suggested by the foregoing remarks will require both theoretical and methodological advances. With regard to SL, the development of more reliable SL measures is crucial (Siegelman et al., 2017). Equally important, however, is the need for theoretical advances clarifying the componential nature of SL and identifying which tasks index each of these components. Similarly, with regard to reading, there is a need for tasks that clearly index specific component processes rather than providing a global measure of overall reading achievement. In addition, because the time scale of reading acquisition is so different from the time scale of learning in typical SL learning tasks, there would be substantial value in developing protocols for investigating the processes underlying learning to read at a time scale commensurate with that of SL paradigms.

Acknowledgments

References

Alba D, & Okanoya K (2008). Statistical segmentation of tone sequences activates the left inferior frontal cortex: A near-infrared spectroscopy study. Neuropsychologia, 46(11), 2787–2795. doi: 10.1016/j.neuropsychologia.2008.05.012 [PubMed: 18579166]

Apfelbaum KS, Hazeltine E, & McMurray B (2013). Statistical learning in reading: Variability in irrelevant letters help children learn phonics skills. Developmental Psychology, 49(7), 1348–1365. doi:10.1037/a0029839 [PubMed: 22924367]

Arciuli J (2017). The multi-component nature of statistical learning. Philosophical Transactions of the Royal Society B, 372(1711), 20160058. doi:10.1098/rstb.2016.0058

Arciuli J, & Simpson IC (2012). Statistical learning is related to reading ability in children and adults. Cognitive Science, 36(2), 286–304. doi:10.1111/j.1551-6709.2011.01200.x [PubMed: 21974775]

Arciuli J, Torkildsen JVK, Stevens DJ, & Simpson IC (2014). Statistical learning under incidental versus intentional conditions. Frontiers in Psychology, 5, 747. doi:10.3389/fpsyg.2014.00747 [PubMed: 25071692]

Baron J, & Strawson C (1976). Use of orthographic and word-specific knowledge in reading words aloud. Journal of Experimental Psychology: Human Perception and Performance, 2(3), 386–393. doi:10.1037//0096-1523.2.3.386

Batterink LJ, Reber PJ, Neville HJ, & Paller KA (2015). Implicit and explicit contributions to statistical learning. Journal of Memory and Language, 83, 62–78. doi:10.1016/j.jml.2015.04.004 [PubMed: 26034344]

Bischoff-Grethe A, Proper SM, Mao H, Daniels KA, & Berns GS (2000). Conscious and unconscious processing of nonverbal predictability in Wernicke's area. Journal of Neuroscience, 20(5), 1975–1981. [PubMed: 10684898]

Bogaerts L, Szmalec A, Hachmann WM, Page MP, & Duyck W (2015). Linking memory and language: Evidence for a serial-order learning impairment in dyslexia. Research in Developmental Disabilities, 43–44(106–122). doi:10.1016/j.ridd.2015.06.012

Browman CP, & Goldstein L (1989). Articulatory gestures as phonological units. Phonology, 6(02), 201–251. doi:10.1017/s0952675700001019

Brown J, Aczel B, Jiménez L, Kaufman SB,… Grant KP, (2010). Intact implicit learning in autism spectrum conditions. Quarterly Journal of Experimental Psychology, 63(9), 1789–1812.

Bussy G, Krifi-Papoz S, Vieville L, Frenay C, Curie A, Rousselle C,… Herbillon V (2011). Apprentissage procédural implicite dans la dyslexie de surface et la dyslexie phonologique. Revue De Neuropsychologie, 3(3), 141. doi:10.3917/rne.033.0141

Castles A, & Coltheart M (1993). Varieties of developmental dyslexia. Cognition, 47(2), 149–180. doi: 10.1016/0010-0277(93)90003-e [PubMed: 8324999]

Chang L, Plaut DC, & Perfetti CA (2016). Visual complexity in orthographic learning: Modeling learning across writing system variations. Scientific Studies of Reading, 20(1), 64–85. doi: 10.1080/10888438.2015.1104688

Chun MM, & Jiang Y (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. Cognitive Psychology, 36(1), 28–71. doi:10.1006/cogp.1998.0681 [PubMed: 9679076]

Cohen NJ, Poldrack RA, & Eichenbaum H (1997). Memory for items and memory for relations in the procedural/declarative memory framework. Memory, 5(1–2), 131–178. doi:10.1080/741941149 [PubMed: 9156097]

Coltheart M, Rastle K, Perry C, Langdon R, & Ziegler J (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. Psychological Review, 108(1), 204–256. doi: 10.1037//0033-295x.108.1.204 [PubMed: 11212628]

Conway CM, Bauernschmidt A, Huang SS, & Pisoni DB (2010). Implicit statistical learning in language processing: Word predictability is the key. Cognition, 114(3), 356–371. doi:10.1016/j.cognition.2009.10.009 [PubMed: 19922909]

Conway CM, & Christiansen MH (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. Journal of Experimental Psychology: Learning, Memory, and Cognition, 31(1), 24–39. doi:10.1037/0278-7393.31.1.24

Conway CM, & Christiansen MH (2006). Statistical learning within and between modalities. Psychological Science, 17(10), 905–912. doi:10.1111/j.1467-9280.2006.01801.x [PubMed: 17100792]

Conway CM, & Pisoni DB (2008). Neurocognitive basis of implicit learning of sequential structure and its relation to language processing. Annals of the New York Academy of Sciences, 1145(1), 113–131. doi:10.1196/annals.1416.009 [PubMed: 19076393]

Deroost N, Zeischka P, Coomans D, Bouazza S, Depessemier P, & Soetens E (2010). Intact first- and second-order implicit sequence learning in secondary-school-aged children with developmental dyslexia. Journal of Clinical and Experimental Neuropsychology, 32(6), 561–572. doi: 10.1080/13803390903313556 [PubMed: 19859852]

Dew IT, & Cabeza R (2011). The porous boundaries between explicit and implicit memory: Behavioral and neural evidence. Annals of the New York Academy of Sciences, 1224(1), 174–190. doi: 10.1111/j.1749-6632.2010.05946.x [PubMed: 21486300]

Dienes Z, Broadbent D, & Berry DC (1991). Implicit and explicit knowledge bases in artificial grammar learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 17(5), 875–887. doi:10.1037//0278-7393.17.5.875

Durrant SJ, Cairney SA, & Lewis PA (2012). Overnight consolidation aids the transfer of statistical knowledge from the medial temporal lobe to the striatum. Cerebral Cortex, 23(10), 2467–2478. doi:10.1093/cercor/bhs244 [PubMed: 22879350]

Durrant SJ, Taylor C, Cairney S, & Lewis PA (2011). Sleep-dependent consolidation of statistical learning. Neuropsychologia, 49(5), 1322–1331. doi:10.1016/j.neuropsychologia.2011.02.015 [PubMed: 21335017]

Eichenbaum H, & Cohen NJ (2001). From conditioning to conscious recollection: Memory systems of the brain. New York, NY: Oxford University Press.

Endress AD, & Mehler J (2009). The surprising power of statistical learning: When fragment knowledge leads to false memories of unheard words. Journal of Memory and Language, 60(3), 351–367.

Folia V, Uddén J, Forkstam C, Ingvar M, Hagoort P, & Petersson KM (2008). Implicit learning and dyslexia. Annals of the New York Academy of Sciences, 1145(1), 132–150. doi:10.1196/annals. 1416.012 [PubMed: 19076394]

Frensch PA, & Miner CS (1994). Effects of presentation rate and individual differences in short-term memory capacity on an indirect measure of serial learning. Memory & Cognition, 22(1), 95–110. doi:10.3758/bf03202765 [PubMed: 8035689]

Frost R (2012). Towards a universal model of reading. Behavioral and Brain Sciences, 35(05), 263– 279. doi:10.1017/s0140525×11001841 [PubMed: 22929057]

Frost R, Armstrong BC, Siegelman N, & Christiansen MH (2015). Domain generality versus modality specificity: The paradox of statistical learning. Trends in Cognitive Sciences, 19(3), 117–125. doi: 10.1016/j.tics.2014.12.010 [PubMed: 25631249]

Frost R, Katz L, & Bentin S (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison. Journal of Experimental Psychology: Human Perception and Performance, 13(1), 104–115. doi:10.1037//0096-1523.13.1.104 [PubMed: 2951484]

Frost R, Siegelman N, Narkiss A, & Afek L (2013). What predicts successful literacy acquisition in a second language? Psychological Science, 24(7), 1243–1252. doi:10.1177/0956797612472207 [PubMed: 23698615]

Gabay Y, Schiff R, & Vakil E (2012). Dissociation between online and offline learning in developmental dyslexia. Journal of Clinical and Experimental Neuropsychology, 34(3), 279–288. doi:10.1080/13803395.2011.633499 [PubMed: 22221291]

Goldman-Rakic P, Selemon L, & Schwartz M (1984). Dual pathways connecting the dorsolateral prefrontal cortex with the hippocampal formation and parahippocampal cortex in the rhesus monkey. Neuroscience, 12(3), 719–743. doi:10.1016/0306-4522(84)90166-0 [PubMed: 6472617]

Gómez RL (2017). Do infants retain the statistics of a statistical learning experience? Insights from a developmental cognitive neuroscience perspective. Philosophical Transactions of the Royal Society B: Biological Sciences, 372(1711), 20160054. doi:10.1098/rstb.2016.0054

Gómez RL, & Edgin JO (2016). The extended trajectory of hippocampal development: Implications for early memory development and disorder. Developmental Cognitive Neuroscience, 18, 57–69. doi:10.1016/j.dcn.2015.08.009 [PubMed: 26437910]

Gómez RL, & Gerken L (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. Cognition, 70(2), 109–135. doi:10.1016/s0010-0277(99)00003-7 [PubMed: 10349760]

Goujon A, Didierjean A, & Thorpe S (2015). Investigating implicit statistical learning mechanisms through contextual cueing. Trends in Cognitive Sciences, 19(9), 524–533. doi:10.1016/j.tics. 2015.07.009 [PubMed: 26255970]

Graves WW, Desai R, Humphries C, Seidenberg MS, & Binder JR (2010). Neural systems for reading aloud: A multiparametric approach. Cerebral Cortex, 20(8), 1799–1815. doi:10.1093/cercor/ bhp245 [PubMed: 19920057]

Hachmann WM, Bogaerts L, Szmalec A, Woumans E, Duyck W, & Job R (2014). Short-term memory for order but not for item information is impaired in developmental dyslexia. Annals of Dyslexia, 64(2), 121–136. doi:10.1007/s11881-013-0089-5 [PubMed: 24488229]

Hamrick P, & Rebuschat P (2012). How implicit is statistical learning? Statistical Learning and Language Acquisition. doi:10.1515/9781934078242.365

Harm MW, & Seidenberg MS (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. Psychological Review, 106(3), 491–528. doi:10.1037//0033-295x.106.3.491 [PubMed: 10467896]

Harm MW, & Seidenberg MS (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. Psychological Review, 111(3), 662– 720. doi:10.1037/0033-295x.111.3.662 [PubMed: 15250780]

Henderson LM, & Warmington M (2017). A sequence learning impairment in dyslexia? It depends on the task. Research in Developmental Disabilities, 60, 198–210. doi:10.1016/j.ridd.2016.11.002 [PubMed: 27856107]

Hickok G, & Poeppel D (2007). The cortical organization of speech processing. Nature Reviews Neuroscience, 8(5), 393–402. doi:10.1038/nrn2113 [PubMed: 17431404]

Howard JH, Howard DV, Japikse KC, & Eden GF (2006). Dyslexics are impaired on implicit higher-order sequence learning, but not on implicit spatial context learning. Neuropsychologia, 44(7), 1131–1144. doi:10.1016/j.neuropsychologia.2005.10.015 [PubMed: 16313930]

Jacoby LL (1991). A process dissociation framework: Separating automatic from intentional uses of memory. Journal of Memory and Language, 30(5), 513–541. doi:10.1016/0749-596X(91)90025-F

Jiménez L, Méndez C, & Cleeremans A (1996). Comparing direct and indirect measures of sequence learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22(4), 948–969. doi:10.1037//0278-7393.22.4.948

Jiménez-Fernández G, Vaquero JM, Jiménez L, & Defior S (2010). Dyslexic children show deficits in implicit sequence learning, but not in explicit sequence learning or contextual cueing. Annals of Dyslexia, 61(1), 85–110. doi:10.1007/s11881-010-0048-3 [PubMed: 21082295]

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 and Language, 127(1), 46–54. doi:10.1016/j.bandl.2012.11.007 [PubMed: 23312790]

Kelly SW, Griffiths S, & Frith U (2002). Evidence for implicit sequence learning in dyslexia. Dyslexia, 8(1), 43–52. [PubMed: 11990224]

Kidd E, & Arciuli J (2016). Individual differences in statistical learning predict childrens comprehension of syntax. Child Development, 87(1), 184–193. doi:10.1111/cdev.12461 [PubMed: 26510168]

Kim R, Seitz A, Feenstra H, & Shams L (2009). Testing assumptions of statistical learning: Is it long-term and implicit? Neuroscience Letters, 461(2), 145–149. doi:10.1016/j.neulet.2009.06.030 [PubMed: 19539701]

Kirkham NZ, Slemmer JA, & Johnson SP (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. Cognition, 83(2). doi:10.1016/s0010-0277(02)00004-5

Lerner I, Armstrong BC, & Frost R (2014). What can we learn from learning models about sensitivity to letter-order in visual word recognition? Journal of Memory and Language, 77, 40–58. doi:10.1016/j.jml.2014.09.002 [PubMed: 25431521]

Liberman AM, & Mattingly IG (1985). The motor theory of speech perception revised. Cognition, 21(1), 1–36. doi:10.1016/0010-0277(85)90021-6 [PubMed: 4075760]

Lum JA, Ullman MT, & Conti-Ramsden G (2013). Procedural learning is impaired in dyslexia: Evidence from a meta-analysis of serial reaction time studies. Research in Developmental Disabilities, 34(10), 3460–3476. doi:10.1016/j.ridd.2013.07.017 [PubMed: 23920029]

Mainela-Arnold E, & Evans JL (2014). Do statistical segmentation abilities predict lexical-phonological and lexical-semantic abilities in children with and without SLI? Journal of Child Language, 41(02), 327–351. doi:10.1017/s0305000912000736 [PubMed: 23425593]

Marshall L, & Born J (2007). The contribution of sleep to hippocampus-dependent memory consolidation. Trends in Cognitive Sciences, 11(10), 442–450. doi:10.1016/j.tics.2007.09.001 [PubMed: 17905642]

Maye J, Werker JF, & Gerken L (2002). Infant sensitivity to distributional information can affect phonetic discrimination. Cognition, 82(3). doi:10.1016/s0010-0277(01)00157-3

Mayo J, & Eigsti IM (2012). Brief report: A comparison of statistical learning in school-aged children with high functioning autism and typically developing peers. Journal of Autism and Developmental Disorders, 42(11), 2476–2485. [PubMed: 22382606]

McClelland JL, McNaughton BL, & O'Reilly RC (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. Psychological Review, 102(3), 419–457. doi:10.1037//0033-295x.102.3.419 [PubMed: 7624455]

Menghini D, Finzi A, Benassi M, Bolzani R, Facoetti A, Giovagnoli S,… Vicari S (2010). Different underlying neurocognitive deficits in developmental dyslexia: A comparative study. Neuropsychologia, 48(4), 863–872. doi:10.1016/j.neuropsychologia.2009.11.003 [PubMed: 19909762]

Menghini D, Hagberg GE, Caltagirone C, Petrosini L, & Vicari S (2006). Implicit learning deficits in dyslexic adults: An fMRI study. NeuroImage, 33(4), 1218–1226. doi:10.1016/j.neuroimage.2006.08.024 [PubMed: 17035046]

Menghini D, Hagberg GE, Petrosini L, Bozzali M, Macaluso E, Caltagirone C, & Vicari S (2008). Structural correlates of implicit learning deficits in subjects with developmental dyslexia. Annals of the New York Academy of Sciences, 1145(1), 212–221. doi:10.1196/annals.1416.010 [PubMed: 19076399]

Misyak J (2010). On-line individual differences in statistical learning predict language processing. Frontiers in Psychology, 1. doi:10.3389/fpsyg.2010.00031

Misyak JB, & Christiansen MH (2012). Statistical learning and language: An individual differences study. Language Learning, 62(1), 302–331. doi:10.1111/j.1467-9922.2010.00626.x

Morris RD, Stuebing KK, Fletcher JM, Shaywitz SE, Lyon GR, Shankweiler DP,… Shaywitz BA (1998). Subtypes of reading disability: Variability around a phonological core. Journal of Educational Psychology, 90 (3), 347–373. doi:10.1037//0022-0663.90.3.347

Newport EL, & Aslin RN (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. Cognitive Psychology, 48(2), 127–162. doi:10.1016/s0010-0285(03)00128-2 [PubMed: 14732409]

Nigro L, Jiménez-Fernández G, Simpson IC, & Defior S (2015). Implicit learning of written regularities and its relation to literacy acquisition in a shallow orthography. Journal of Psycholinguistic Research, 44(5), 571–585. doi:10.1007/s10936-014-9303-9 [PubMed: 24935792]

Nissen MJ, & Bullemer P (1987). Attentional requirements of learning: Evidence from performance measures. Cognitive Psychology, 19, 1. doi:10.1016/0010-0285(87)90002-8

Packard MG, & Knowlton BJ (2002). Learning and memory functions of the basal ganglia. Annual Review of Neuroscience, 25(1), 563–593. doi:10.1146/annurev.neuro.25.112701.142937

Page MP, Cumming N, Norris D, Hitch GJ, & Mcneil AM (2006). Repetition learning in the immediate serial recall of visual and auditory materials. Journal of Experimental Psychology: Learning, Memory, and Cognition, 32(4), 716–733. doi:10.1037/0278-7393.32.4.716

Perruchet P, & Pacton S (2006). Implicit learning and statistical learning: One phenomenon, two approaches. Trends in Cognitive Sciences, 10(5), 233–238. doi:10.1016/j.tics.2006.03.006 [PubMed: 16616590]

Petersson K, Folia V, & Hagoort P (2012). What artificial grammar learning reveals about the neurobiology of syntax. Brain and Language, 120(2), 83–95. doi:10.1016/j.bandl.2010.08.003 [PubMed: 20943261]

Poldrack RA, Clark J, Paré-Blagoev EJ, Shohamy D, Moyano JC, Myers C, & Gluck MA (2001). Interactive memory systems in the human brain. Nature, 414(6863), 546–550. doi:10.1038/35107080 [PubMed: 11734855]

Pugh KR, Mencl WE, Shaywitz BA, Shaywitz SE, Fulbright RK, Constable RT, … Gore JC (2000). The angular gyrus in developmental dyslexia: Task-specific differences in functional connectivity within posterior cortex. Psychological Science, 11(1), 51–56. doi:10.1111/1467-9280.00214 [PubMed: 11228843]

Reber AS (1992). The cognitive unconscious: An evolutionary perspective. Consciousness and Cognition, 1(2), 93–133.

Reber PJ (2013). The neural basis of implicit learning and memory: A review of neuropsychological and neuroima- ging research. Neuropsychologia, 51(10), 2026–2042. doi:10.1016/j.neuropsychologia.2013.06.019 [PubMed: 23806840]

Reber PJ, Gitelman DR, Parrish TB, & Mesulam MM (2003). Dissociating explicit and implicit category knowledge with fMRI. Journal of Cognitive Neuroscience, 15(4), 574–583. doi:10.1162/089892903321662958 [PubMed: 12803968]

Redington M & Chater N (1996). Transfer in artificial grammar learning: A reevaluation. Journal of Experimental Psychology: General, 125(2), 123.

Robertson EM (2007). The serial reaction time task: Implicit motor skill learning? Journal of Neuroscience, 27(38), 10073–10075. doi:10.1523/jneurosci.2747-07.2007 [PubMed: 17881512]

Rose M, Haider H, Weiller C, & Büchel C (2002). The role of medial temporal lobe structures in implicit learning. Neuron, 36(6), 1221–1231. doi:10.1016/s0896-6273(02)01105-4 [PubMed: 12495634]

Rueckl J (2010). Connectionism and the role of morphology in visual word recognition. The Mental Lexicon, 5(3), 371–400. [PubMed: 22934123]

Rueckl JG (2016). Toward a theory of variation in the organization of the word reading system. Scientific Studies of Reading, 20(1), 86–97. doi:10.1080/10888438.2015.1103741 [PubMed: 26997862]

Rueckl JG, Paz-Alonso PM, Molfese PJ, Kuo W, Bick A, Frost SJ,… Frost R (2015). Universal brain signature of proficient reading: Evidence from four contrasting languages. Proceedings of the National Academy of Sciences, 112(50), 15510–15515. doi:10.1073/pnas.1509321112

Saffran JR (2003). Statistical language learning. Current Directions in Psychological Science, 12(4), 110–114. doi:10.1111/1467-8721.01243

Saffran JR, Aslin RN, & Newport EL (1996a). Statistical learning by 8-month-old infants. Science, 274(5294), 1926–1928. doi:10.1126/science.274.5294.1926 [PubMed: 8943209]

Saffran JR, Newport EL, & Aslin RN (1996b). Word segmentation: The role of distributional cues. Journal of Memory and Language, 35(4), 606–621. doi:10.1006/jmla.1996.0032

Saffran JR, Newport EL, Aslin RN, Tunick RA, & Barrueco S (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. Psychological Science, 8(2), 101–105. doi:10.1111/j.1467-9280.1997.tb00690.x

Schacter DL (1987). Implicit memory: History and current status. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13(3), 501–518. doi:10.1037//0278-7393.13.3.501

Schapiro A, & Turk-Browne N (2015). Statistical learning. Brain Mapping, 501–506. doi:10.1016/b978-0-12-397025-1.00276-1

Schendan HE, Searl MM, Melrose RJ, & Stern CE (2003). An fMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. Neuron, 37(6), 1013–1025. doi:10.1016/s0896-6273(03)00123-5 [PubMed: 12670429]

Schmalz X, Altoè G, & Mulatti C (2016). Statistical learning and dyslexia: A systematic review. Annals of Dyslexia, 67(2), 147–162. doi:10.1007/s11881-016-0136-0 [PubMed: 27766563]

Seidenberg MS (2011). Reading in different writing systems: One architecture, multiple solutions In McCardle P, Miller B, Lee JR, & Tzeng OJL (Eds.), The extraordinary brain series. Dyslexia across languages: Orthography and the brain–Gene–Behavior link (pp. 146–168). Baltimore, MD: Brookes.

Seidenberg MS, & Gonnerman LM (2000). Explaining derivational morphology as the convergence of codes. Trends in Cognitive Sciences, 4(9), 353–361. doi:10.1016/s1364-6613(00)01515-1 [PubMed: 10962617]

Seidenberg MS, & Mcclelland JL (1989). A distributed, developmental model of word recognition and naming. Psychological Review, 96(4), 523–568. doi:10.1037//0033-295x.96.4.523 [PubMed: 2798649]

Shafto CL, Conway CM, Field SL, & Houston DM (2012). Visual sequence learning in infancy: Domain-general and domain-specific associations with language. Infancy, 17(3), 247–271. doi:10.1111/j.1532-7078.2011.00085.x [PubMed: 22523477]

Shohamy D, & Adcock RA (2010). Dopamine and adaptive memory. Trends in Cognitive Sciences, 14(10), 464–472. doi:10.1016/j.tics.2010.08.002 [PubMed: 20829095]

Shohamy D, & Turk-Browne NB (2013). Mechanisms for widespread hippocampal involvement in cognition. Journal of Experimental Psychology: General, 142(4), 1159–1170. doi:10.1037/a0034461 [PubMed: 24246058]

Siegelman N, Bogaerts L, Christiansen MH, & Frost R (2017). Towards a theory of individual differences in statistical learning. Philosophical Transactions of the Royal Society B: Biological Sciences, 372(1711), 20160059. doi:10.1098/rstb.2016.0059

Siegelman N, Bogaerts L, & Frost R (2017). Measuring individual differences in statistical learning: Current pitfalls and possible solutions. Behavior Research Methods, 49(2), 418–432. [PubMed: 26944577]

Siegelman N, & Frost R (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. Journal of Memory and Language, 81, 105–120. doi:10.1016/j.jml.2015.02.001 [PubMed: 25821343]

Simon KN, Werchan D, Goldstein MR, Sweeney L, Bootzin RR, Nadel L, & Gómez RL (2016). Sleep confers a benefit for retention of statistical language learning in 6.5 month old infants. Brain and Language, 167, 3–12. doi:10.1016/j.bandl.2016.05.002 [PubMed: 27291337]

Smith L, & Yu C (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. Cognition, 106 (3), 1558–1568. doi:10.1016/j.cognition.2007.06.010 [PubMed: 17692305]

Spencer M, Kaschak MP, Jones JL, & Lonigan CJ (2015). Statistical learning is related to early literacy-related skills. Reading and Writing, 28(4), 467–490. doi:10.1007/s11145-014-9533-0 [PubMed: 26478658]

Squire LR (1992). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. Journal of Cognitive Neuroscience, 4(3), 232–243. [PubMed: 23964880]

Squire LR (2004). Memory systems of the brain: A brief history and current perspective. Neurobiology of Learning and Memory, 82(3), 171–177. doi:10.1016/j.nlm.2004.06.005 [PubMed: 15464402]

Squire LR, & Knowlton BJ (2000). The medial temporal lobe, the hippocampus, and the memory systems of the brain. The New Cognitive Neurosciences, 2, 756–776.

Stoodley CJ, Harrison EP, & Stein JF (2006). Implicit motor learning deficits in dyslexic adults. Neuropsychologia, 44(5), 795–798. doi:10.1016/j.neuropsychologia.2005.07.009 [PubMed: 16125207]

Stoodley CJ, Ray NJ, Jack A, & Stein JF (2008). Implicit learning in control, dyslexic, and garden-variety poor readers. Annals of the New York Academy of Sciences, 1145(1), 173–183. doi:10.1196/annals.1416.003 [PubMed: 19076396]

Strain E, & Herdman CM (1999). Imageability effects in word naming: An individual differences analysis. Canadian Journal of Experimental Psychology/Revue Canadienne De Psychologie Expérimentale, 53(4), 347–359. doi:10.1037/h0087322 [PubMed: 10646206]

Suzuki WL, & Amaral DG (1994). Perirhinal and parahippocampal cortices of the macaque monkey: Cortical afferents. The Journal of Comparative Neurology, 350(4), 497–533. doi:10.1002/cne.903500402 [PubMed: 7890828]

Taylor JSH, Rastle K, & Davis MH (2013). Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. Psychological Bulletin, 139(4), 766–791. doi:10.1037/a0030266 [PubMed: 23046391]

Thiessen ED (2017). Whats statistical about learning? Insights from modelling statistical learning as a set of memory processes. Philosophical Transactions of the Royal Society B: Biological Sciences, 372(1711), 20160056. doi:10.1098/rstb.2016.0056

Thiessen ED, & Erickson LC (2013). Beyond word segmentation. Current Directions in Psychological Science, 22 (3), 239–243. doi:10.1177/0963721413476035

Thiessen ED, Kronstein AT, & Hufnagle DG (2015). The extraction and integration framework: A two-process account of statistical learning. Psychological Bulletin, 139(4), 792–814. doi:10.1037/a0030801

Treiman R, Kessler B, & Bick S (2003). Influence of consonantal context on the pronunciation of vowels: A comparison of human readers and computational models. Cognition, 88(1), 49–78. doi:10.1016/s0010-0277(03)00003-9 [PubMed: 12711153]

Turk-Browne NB, Scholl BJ, Chun MM, & Johnson MK (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. Journal of Cognitive Neuroscience, 21(10), 1934–1945. doi:10.1162/jocn.2009.21131 [PubMed: 18823241]

Ullman MT (2004). Contributions of memory circuits to language: The declarative/procedural model. Cognition, 92 (1–2), 231–270. doi:10.1016/j.cognition.2003.10.008 [PubMed: 15037131]

Ullman MT, & Pierpont EI (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. Cortex, 41(3), 399–433. doi:10.1016/s0010-9452(08)70276-4 [PubMed: 15871604]

Vicari S (2005). Do children with developmental dyslexia have an implicit learning deficit? Journal of Neurology, Neurosurgery & Psychiatry, 76(10), 1392–1397. doi:10.1136/jnnp.2004.061093

Vicari S, Marotta L, Menghini D, Molinari M, & Petrosini L (2003). Implicit learning deficit in children with developmental dyslexia. Neuropsychologia, 41(1), 108–114. doi:10.1016/s0028-3932(02)00082-9 [PubMed: 12427569]

Voss JL, & Paller KA (2008). Brain substrates of implicit and explicit memory: The importance of concurrently acquired neural signals of both memory types. Neuropsychologia, 46(13), 3021–3029. doi:10.1016/j.neuropsychologia.2008.07.010 [PubMed: 18691605]

Wagner AD, Maril A, & Schacter DL (2000). Interactions between forms of memory: When priming hinders new episodic learning. Journal of Cognitive Neuroscience, 12(Suppl. 2), 52–60. doi: 10.1162/089892900564064 [PubMed: 11506647]

Witt A, Puspitawati I, & Vinter A (2013). How explicit and implicit test instructions in an implicit learning task affect performance. PLoS ONE, 8(1). doi:10.1371/journal.pone.0053296

Woollams AM, Ralph MA, Madrid G, & Patterson KE (2016). Do you read how i read? Systematic individual differences in semantic reliance amongst normal readers. Frontiers in Psychology, 7. doi:10.3389/fpsyg.2016.01757

Yang J, & Li P (2012). Brain networks of explicit and implicit learning. PLoS ONE, 7(8). doi:10.1371/journal.pone.0042993

Yang Y, & Hong-Yan B (2011). Unilateral implicit motor learning deficit in developmental dyslexia. International Journal of Psychology, 46(1), 1–8. doi:10.1080/00207594.2010.509800 [PubMed: 22044127]

Yi H, Maddox WT, Mumford JA, & Chandrasekaran B (2014). The role of corticostriatal systems in speech category learning. Cerebral Cortex, 26(4), 1409–1420. doi:10.1093/cercor/bhu236 [PubMed: 25331600]