

2014

Neurocognitive Mechanisms of Statistical-Sequential Learning: what do Event-Related Potentials Tell Us?

Jerome Daltrozzo

Georgia State University, jdaltrozzo@gsu.edu

Christopher M. Conway

Georgia State University, cconway@gsu.edu

Follow this and additional works at: https://scholarworks.gsu.edu/psych_facpub



Part of the [Psychology Commons](#)

Recommended Citation

Daltrozzo, J., & Conway, C. M. (2014). Neurocognitive mechanisms of statistical-sequential learning: what do event-related potentials tell us? *Frontiers in Human Neuroscience*, 8, 437. <http://doi.org/10.3389/fnhum.2014.00437>

This Article is brought to you for free and open access by the Department of Psychology at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Psychology Faculty Publications by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.



Neurocognitive mechanisms of statistical-sequential learning: what do event-related potentials tell us?

Jerome Daltrozzo* and Christopher M. Conway

Department of Psychology, Georgia State University, Atlanta, GA, USA

Edited by:

Rachael D. Seidler, University of Michigan, USA

Reviewed by:

Ben Godde, Jacobs University Bremen, Germany

Rebecca Spencer, University of Massachusetts, USA

*Correspondence:

Jerome Daltrozzo, Department of Psychology, Georgia State University, P.O. Box 5010, Atlanta, GA 30302-5010, USA
e-mail: jdaltrozzo@gsu.edu

Statistical-sequential learning (SL) is the ability to process patterns of environmental stimuli, such as spoken language, music, or one's motor actions, that unfold in time. The underlying neurocognitive mechanisms of SL and the associated cognitive representations are still not well understood as reflected by the heterogeneity of the reviewed cognitive models. The purpose of this review is: (1) to provide a general overview of the primary models and theories of SL, (2) to describe the empirical research – with a focus on the event-related potential (ERP) literature – in support of these models while also highlighting the current limitations of this research, and (3) to present a set of new lines of ERP research to overcome these limitations. The review is articulated around three descriptive dimensions in relation to SL: the level of abstractness of the representations learned through SL, the effect of the level of attention and consciousness on SL, and the developmental trajectory of SL across the life-span. We conclude with a new tentative model that takes into account these three dimensions and also point to several promising new lines of SL research.

Keywords: sequential learning, statistical learning, implicit learning, procedural learning, artificial grammar, ERP, P300, P600

INTRODUCTION

From an ecological point of view, learning about temporal patterns in our environment, and using this information to make predictions about upcoming events and actions, is arguably of primary importance to humans and other higher-order organisms (Lashley, 1951; Conway et al., 2010; Goldstein et al., 2010). In the past 15 years, an increasingly established body of research has demonstrated that humans have a remarkable ability to learn statistical patterns – i.e., commonalities and underlying regularities – from among a set of stimuli, a phenomenon now referred to simply as “statistical learning” (Saffran et al., 1996, 1997). A related phenomenon, known as “implicit learning,” likewise reveals people's ability to learn predictive patterns without conscious intent or awareness (Cleeremans and McClelland, 1991; Berry and Dienes, 1993). Both statistical learning and implicit learning have been observed with many different types of input materials in sensory (e.g., music, speech, and visual patterns) and motor domains. In fact, due to the apparent commonalities between statistical learning and implicit learning, there is growing consensus that these two phenomena may actually tap into the same process (Perruchet and Pacton, 2006).

In the current review, we focus in particular on the learning of temporal or sequential patterns of stimuli and therefore use the term “statistical-sequential learning” or simply “sequential learning” (SL) for short. Because it is still an open question as to whether these learning abilities are also governed at least in part by explicit processes (e.g., Baddeley and Wilson, 1994; Cleeremans, 2006; Haider and Frensch, 2009; Jamieson and Mewhort, 2009; Dale et al., 2012), we avoid the use of the term “implicit” (and in subsequent sections we directly address the different contributions of

implicit and explicit processes). Under this definition, SL is the ability to learn underlying structured patterns that exist among a set of non-random, sequentially presented stimuli (Conway and Christiansen, 2001; Conway, 2012). Yet another term recently used that also captures this crucial aspect of statistical-sequential learning is “structured sequence processing” (Uddén and Bahlmann, 2012).

To date, the underlying cognitive and neural mechanisms of SL and the associated cognitive representations are still not well understood. SL has been explored through a combination of cognitive modeling and empirical studies using behavioral and neurophysiological measurements. The current outcome of these heterogeneous approaches is that the proposed theories of SL still need to be confirmed by empirical evidence. The purpose of this review is to provide an initial assessment of the current theories of SL and to identify the areas of empirical research that need further development. Due to the extensive behavioral and neural SL literature, the scope of this review will focus on the exploration of SL with a specific neural approach, the event-related potential (ERP) technique (for other neuroimaging techniques, see for example Seger et al., 2000; Bischoff-Grethe et al., 2001; Huettel et al., 2002; Skosnik et al., 2002; Lieberman et al., 2004; Petersson et al., 2004; Thomas et al., 2004; Forkstam et al., 2006; Turk-Browne et al., 2009; Uddén and Bahlmann, 2012).

Since SL has been observed in multiple modalities and domains, we draw upon a wide range of empirical studies, reviewing for instance studies on motor learning, visual-motor learning, visual-perceptual learning, auditory learning of different types of stimuli, language learning, and social learning. Recognizing the differences across these studies when relevant, we also focus on the

commonalities among them in order to bring to light what we believe is the cognitive process at the core of all of them.

We first summarize the primary theoretical views of SL. We then review the main approaches used by ERP research to study SL. This will point to a discrepancy between the theoretical and the empirical approaches, highlighting a series of fundamental unanswered questions. Finally, we provide suggestions for moving forward to address the most challenging aspects of SL research and provide a tentative new model of SL that incorporates much of the existing empirical and theoretical advances.

MODELS AND THEORIES OF SEQUENTIAL LEARNING

Three primary questions about the nature of SL have intrigued researchers over the decades, organized around a limited set of non-orthogonal (i.e., partly overlapping) dimensions: (1) the extent to which SL encodes and manipulates concrete versus abstract representations, (2) whether SL depends on the level of conscious awareness or attention, and (3) how SL changes across the life-span. We consider each of these issues in turn.

CONCRETE VERSUS ABSTRACT REPRESENTATIONS

Sequential learning could in principal encode either: (1) concrete features of the sequence, such as the frequencies of individual items (or exemplars) of the sequence, (2) or abstract features, e.g., abstract rule(s) that organize the to-be-learned sequence (Franco and Destrebecqz, 2012). This section refers primarily to the types of representations that are manipulated by the SL mechanism(s).

Reber (1967) – using an artificial grammar paradigm – was the first to propose that SL is the result of the implicit learning of abstract rules. This proposal was later endorsed by several others (e.g., McAndrews and Moscovitch, 1985; Mathews et al., 1989; Dienes et al., 1991; Knowlton et al., 1992; Knowlton and Squire, 1993, 1994, 1996; Manza and Reber, 1997; Marcus et al., 1999; Rosnagel, 2001; Kuhn and Dienes, 2005). The idea that the cognitive system was able to unconsciously process abstract information, the so-called “smart unconscious” hypothesis (Cleeremans et al., 1998), was for many researchers somewhat provocative and was challenged by connectionist computational modeling (Christiansen et al., 1998). Connectionist models showed that rather than the learning of abstract rules, several results of the SL literature could be successfully modeled using only concrete feature processing, such as the processing of chunks or transitional probabilities (Perruchet and Pacton, 2006).

Perhaps the best-known empirical demonstration of SL comes from Saffran et al. (1997), who used a word segmentation task in which a continuous sequence of syllables was presented (e.g., “bupadapatubitutibu”). The syllable sequence covertly consisted of artificial “words” (e.g., “bupada” and “patubi”) spliced together. Participants demonstrated above-chance performance in a subsequent recognition test, discriminating words from non-word syllable groupings. Saffran et al. (1997) proposed that such performance was achieved by exploiting the statistical regularities present in the sequence of syllables, such as transitional probabilities between successive

syllables (e.g., the probability that a given syllable A is immediately followed by another given syllable B) that are higher within words than between words. These statistical regularities are one type of concrete feature that could be learned in a sequence.

The acquisition of these concrete features is often referred to as “surface learning” or “fragmentary learning” (Perruchet and Pacteau, 1990; Servan-Schreiber and Anderson, 1990; Perruchet and Amorim, 1992; Meulemans and Van der Linden, 1997). Surface learning may be based on the encoding of item frequencies and item variability across the sequence (Maye et al., 2002; Perruchet et al., 2004; Clayards et al., 2008). Cleeremans et al. (1998) reviewed at least three types of concrete features that once learned could account for many results of the SL literature: fragment-based or chunk information, exemplars, and distributional information (Figure 1). In the same vein, several models have been proposed to account for surface learning based on the to-be-learned type of concrete information. Some models focused on conditional statistics between items of the sequence (Thiessen and Pavlik, 2013) and others on the use of temporal contingencies (Montague and Sejnowski, 1994) that may covary in a cause-effect relationship with the physical world (Gopnik et al., 2004).

These concrete feature-based models are computational and have been criticized as such. For instance, the simple-recurrent-network model (Elman, 1990; Cleeremans and McClelland, 1991), has been argued to suffer major weaknesses (McCloskey and Cohen, 1989; Goldstein et al., 2010) with (1) long range dependencies, as in “embedded sequences” (e.g., Uddén and Bahlmann, 2012); (2) sequences made of large sets of rules and items of the scale found in natural language, notably because they are designed to consider the entire corpus of input simultaneously, rather than in the proper temporal order (Goldstein et al., 2010); and (3) multimodal data (Goldstein et al., 2010).

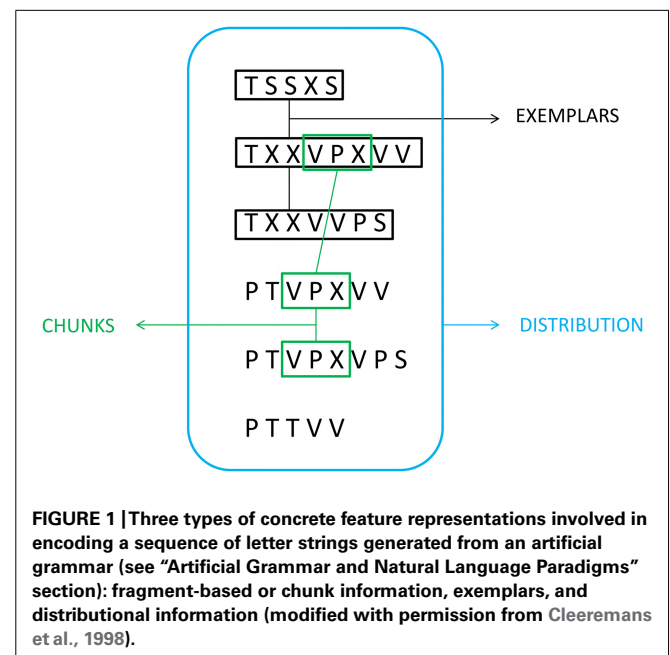


FIGURE 1 | Three types of concrete feature representations involved in encoding a sequence of letter strings generated from an artificial grammar (see “Artificial Grammar and Natural Language Paradigms” section): fragment-based or chunk information, exemplars, and distributional information (modified with permission from Cleeremans et al., 1998).

Related to the issue of abstractness, SL could result in modality-specific (more concrete) or amodal (more abstract) representations. For Reber, SL was a mainly amodal process (Reber, 1989); however, some research has suggested that both domain-general (Clegg et al., 1998; Kirkham et al., 2002; Bapi et al., 2005) and modality-specific SL might coexist (Keele et al., 2003; Conway and Christiansen, 2005; Conway and Pisoni, 2008; Turk-Browne et al., 2009; Shafto et al., 2012). For example, Keele et al. (2003) proposed two independent SL systems based on the available behavioral and neuroimaging findings at the time. One system integrates all sequential information regardless of the input modality (presumably relying on more “abstract” representations that are not tied to a particular input modality), while a second system captures only the patterns of a sequence within a single modality (more reliant on “concrete” or modality-specific representations), without suffering interference from intervening sequential information from other modalities. Keele et al.’s (2003) two-system model of SL is therefore consistent with the notion that SL might encode both concrete (stimulus-specific) and more abstract (domain-general) patterns.

Some models of SL in fact explicitly incorporate a multilayer structure. Clegg et al. (1998) suggested three levels of processing: (1) an abstract level storing higher-level goals that are neither stimulus- nor response-related; (2) an intermediate level encoding the type of action required (independently of the effector) or the stimulus specificity (independently of its exact identity); and (3) a low level acquiring highly specific information related to the exact stimulus and the associated final motor execution. Possibly a parallel could be drawn between the representations processed by these three layers and the concrete-abstract continuum. Multilayer models like Clegg et al. (1998) have the advantage of providing an account of both concrete feature learning and more abstract situations, such as the “transfer of learning” paradigm, which indicates that the representation of a sequence may not be tied to a particular effector or stimulus domain (Clegg et al., 1998).

Related to the issue of modality-specificity, it should be noted that the more concrete-based aspects of SL appear to show similarities to perceptual learning (PL), which allows for the development of spatio-temporal representations of the environment through learning along various levels of cortical processing (Sagi and Tanne, 1994; Skrandies and Fahle, 1994; Goldstone, 1998; Conway et al., 2007). Interestingly, PL and perceptual-based SL seem to activate similar neural networks (Turk-Browne et al., 2009). Like SL, PL can occur with rather short exposure to patterns, can have long lasting effects, and can occur without attention to or awareness of the patterns; however, PL can also be modulated by levels of attention and awareness (Goldstone, 1998; Alain et al., 2007; Sasaki et al., 2010; Lu et al., 2011; Aberg and Herzog, 2012; Byers and Serences, 2012; Kumano and Uka, 2013). Furthermore, PL is, like SL, often described as being at the root of language learning, particularly for the development of phonological and lexical representations (Goldstone, 1998; Cutler, 2008; Samuel and Kraljic, 2009; Werker, 2012) and is also proposed as a process required for motor preparation and execution (Hommel et al., 2001). According to a standard definition of SL – the ability to learn patterns of stimuli unfolding

in time – SL can be seen as the “temporal” subcategory of a the more general “spatio-temporal” PL, in which items frequently co-occurring in time (but not spatially) can form new perceptual “units” (Goldstone, 1998). If SL is viewed from this perspective, the development of concrete representations during SL could be explained in terms of properties of PL. Indeed, (temporal) statistical contingencies between items/percepts (e.g., transitional probabilities or perceptual units of co-occurring informations such as chunks, Czerwinski et al., 1992; Seriès and Seitz, 2013) could be captured and stored in cortical spatio-temporal representations.

One final way that, together with abstractness and modality, SL representations might be differentiated is by the types of input structures (Conway and Christiansen, 2001; Conway, 2012). Three types have been proposed: fixed patterns (i.e., invariant or repeating sequences); statistical patterns (sequences containing statistical regularities or distributional information across exemplars); and hierarchical patterns (i.e., embedded sequences with non-adjacent or self-recursive structures). Different neurocognitive mechanisms may be used in the service of each type of input structure (Bahmann et al., 2006; Uddén and Bahmann, 2012). These three types of input structures appear related to the concrete-abstract continuum: learning an invariant fixed pattern or statistical regularity is likely represented in a concrete fashion, whereas learning a self-recursive structure is likely represented more abstractly, allowing for generalization of the recursive rule to new exemplars.

It appears likely then that SL involves multiple processes, some that could be characterized as being more domain-general and that manipulate rather abstract representations, and others that are more input-specific and that encode more concrete features. This perspective is similar to the “more-than-one-mechanism” (MOM) hypothesis of language acquisition, stating that language is acquired via the manipulation of both rule-based and statistical representations (Endress and Bonatti, 2007). Several recent models of SL now combine feature-based learning with more abstract forms of rule-learning mechanisms. For instance, Pierrehumbert (2003) provided a model of how abstract rules could be extracted from a speech signal through the interaction between different high and low-level cognitive systems, including bottom-up processing of low-level acoustic and articulatory features. In this model, a phonological system would refine internal categorizations in its different levels through: (1) internal feedback mechanisms from higher level internal systems to lower levels internal systems, and (2) external feedback due to the interaction with the speech community.

The issue of the abstractness of the representations manipulated during SL is complex. Perhaps the most promising accounts of SL involve the processing of both concrete and abstract information (e.g., Clegg et al., 1998; Keele et al., 2003; Pierrehumbert, 2003). The exact interplay among these postulated processes remains unknown and opened to multiple model implementations. In a rather simple model, the two hypothetical mechanisms would work in parallel. One would encode and store modality-specific concrete features in a given format and another mechanism would encode and store domain-general

abstract information in another format. A second and perhaps more neurally plausible possibility is a cascading account, whereby the two mechanisms interact in a hierarchical manner, with concrete information being first encoded in a modality-specific format, followed, upon further processing or exposure to the input, by the development and encoding of more abstract and domain-general representations. Accordingly SL across input modalities (e.g., learning that a particular tone predicts a visual stimulus) would present a greater processing challenge than SL within an input modality (e.g., learning that a particular visual stimulus predicts another visual stimulus). This is, in fact, what recent findings appear to indicate (Walk and Conway, 2011).

IMPLICIT AND EXPLICIT MECHANISMS

In addition to dissociating the mechanisms of SL by the level of abstractness of the learned features, the level of attention (and consciousness) has also been recognized as a critical dimension of SL. The SL literature often refers to this issue in terms of “implicit” and “explicit” processing. Traditionally, SL is generally thought to involve the activation of incidental/implicit, automatic, and even unconscious processes (e.g., Saffran et al., 1996, 1997; Fiser and Aslin, 2001, 2002; Shanks and Perruchet, 2002; Turk-Browne et al., 2005; Shanks et al., 2006; Hannula and Ranganath, 2009; Rosenthal et al., 2010). Several empirical strands of research on SL have suggested that the level of awareness is irrelevant to SL performance (Curran and Keele, 1993; Goschke, 1998; Song et al., 2007). Clegg et al. (1998) not only acknowledge this implicit component of SL but go further by suggesting that SL does not manipulate explicit knowledge representations. Rather, they suggest that explicit knowledge emerges through the interaction of SL with other cognitive systems that can access and modify explicit memories (Clegg et al., 1998).

Alternatively, other theories have argued for a more direct role of explicit processing in SL. For instance, Cleeremans (2006) suggested that a representation obtained from exposure to a sequence may become explicit when the strength of activation of this representation reaches a critical level. Similarly, explicit knowledge may emerge as the result of a search process that is triggered by unexpected events occurring during task processing and requiring an explanation (the unexpected-event hypothesis; Haider and Frensch, 2009). Some authors go even further by drawing a link between “general” consciousness/awareness (i.e., not only of sequence representations) and SL. Dale et al. (2012) proposed that predictive mechanisms such as those that are thought to account for SL may be at the root of the formation of conscious percepts or awareness (Morsella, 2005).

Between these two extreme views there exist proposals that acknowledge the development of both conscious and unconscious representations resulting from SL as well as the contribution of explicit and implicit mechanisms to SL. For instance, Baddeley and Wilson (1994), who analyzed the effect of explicit versus implicit learning in amnesic patients, suggested that implicit learning is strongly dependent on the efficiency of explicit learning, as the later would monitor errors while the former would be heavily impaired by errors during learning. Jamieson and Mewhort

(2009) reached a similar conclusion. In their model, they suggested that even though SL can occur without the participant's explicit knowledge of an underlying rule, SL would nevertheless require memory retrieval of association traces between the current stimulus, the response associated with it, and the context provided by the immediately preceding response. Importantly, they underline that this account of SL does not require implicit learning but instead memory retrieval, that may or may not be fully conscious.

Clearly, there is far from a consensus on the question of whether SL is subserved by implicit or explicit mechanisms, or a combination of both. Nevertheless, perhaps the most influential view to date is that both types of mechanisms contribute to SL (e.g., Curran and Keele, 1993). Importantly, this view finds support from neuroimaging data. Physically distinct brain networks, including dorsolateral prefrontal, medial frontal, and more dorsal posterior regions, appear to be activated when subjects become consciously aware of a sequence. These networks are not activated when subjects are unaware of the sequence rules (Grafton et al., 1995). Such results would be consistent with explicit knowledge leading to the use of working memory to process conscious representations of the sequence (Smith and Jonides, 1995), while areas commonly associated with motor control and/or perceptual processing, including motor cortex, primary sensory areas, and subcortical structures in the basal ganglia, would be activated under conditions of implicit learning (for a more complete discussion see Curran, 1998).

From a methodological point of view, one way to explore the extent of explicit and implicit learning in SL paradigms is to use rapid serial visual presentations (RSVP). Kim et al. (2009), for instance, used such a design together with a matching questionnaire to assess explicit learning and concluded that SL was performed through implicit mechanisms. But several critiques can be raised on the ability to assess purely explicit learning through questionnaire assessments. Thus, novel methods have been developed to better dissociate implicit from explicit learning, such as comparisons between direct and indirect tasks or the process-dissociation procedure (Jacoby, 1991). In direct tasks, such as questionnaire assessments or recognition judgments, subjects are explicitly instructed to respond based on their conscious knowledge. In indirect tasks, performance is measured in a manner that does not require conscious choice by the participants. If participants show greater SL as measured by an indirect task compared to a direct task, it is likely that SL occurred without accompanying conscious awareness (Cleeremans et al., 1998). Taking this logic one step further, Jacoby (1991) proposed the process-dissociation procedure as a method for dissociating implicit from explicit learning. This procedure allows one to separate memories acquired intentionally (i.e., consciously) from memories acquired automatically. Franco et al. (2011) applied this method to explore the cognitive mechanism(s) of SL. They found that statistical information acquired through two SL paradigms containing two different artificial grammars of syllables where only transition probabilities differed, can be consciously manipulated to differentiate these artificial languages. That is, the transitional probabilities became to some extent available to consciousness.

Even though these new methods have improved our ability to assess the contribution of the level of consciousness to SL mechanisms, the issue is far from settled. Some researchers still believe that the assessment of consciousness needs further improvements (Dale et al., 2012). Importantly, the debate about the interaction between consciousness and SL performance essentially distinguishes between two aspects of consciousness: the consciousness of the acquired knowledge (e.g., transitional probabilities) resulting from SL (see for instance Franco et al., 2011) and the level of consciousness available or required during the SL process itself, that is, whether learning was intentional or incidental. One recent empirical study incorporated this distinction by using a dual-task paradigm that induced a cognitive load either during an (incidental) encoding phase or during an (explicit) test phase, or both (Hendricks et al., 2013). Interestingly, the results demonstrated differential effects of the dual-task manipulation, impairing performance only during the explicit test phase, that is, during the manipulation of explicit knowledge, but not during the encoding phase. Furthermore, in a transfer condition in which the elements of each sequence were mapped onto a new subset of items, the dual-task condition eliminated SL regardless of whether it occurred during the encoding phase or during the test phase. This finding suggests that SL is largely an implicit process; however, the expression of previously learned knowledge gained through SL during an explicit test as well as the learning of abstract rules appears to require conscious awareness (Hendricks et al., 2013).

In summary, the literature remains highly heterogeneous in terms of the impact of the level of consciousness on SL performance. However, perhaps the most conservative view, similar to that discussed earlier, is that SL might not be governed by a single cognitive mechanism and might not store representations in a single – e.g., unconscious – format. Instead, SL is likely subserved by at least two mechanisms, one that is rather independent of the level of consciousness/attention and results in unconscious representations and one that depends more on attentional resources and leads to more conscious representations. We will see that ERPs can be helpful in testing this assumption.

DEVELOPMENTAL CONSIDERATIONS

Whether described in terms of the abstractness of the representations or on the consciousness/attentional dimension, SL can hardly be fully investigated without taking into account its developmental trajectory. Although most SL experiments have been performed with young adults, several studies have focused on SL in children (Saffran et al., 1997; Meulemans et al., 1998; Thomas and Nelson, 2001; Vicari et al., 2003; Thomas et al., 2004; Arciuli and Simpson, 2011; Arciuli and von Koss Torkildsen, 2012) and infants (Haith et al., 1988; Haith and McCarty, 1990; Saffran et al., 1996, 2001; Smith et al., 1997; Aslin et al., 1998; Clohessy et al., 2001; Fiser and Aslin, 2002; Shafto et al., 2012). There are also a handful of studies investigating SL in the elderly population (Prull et al., 2000; Dennis et al., 2003; Howard et al., 2004; Aizenstein et al., 2005; Humes and Floyd, 2005; Shea et al., 2006).

Despite the growing body of research that focuses on SL across the life-span, the developmental progression of SL is still largely unknown. The early literature on implicit learning

assumed that this cognitive ability was rather independent of age (Reber, 1993), while explicit learning would improve with aging (Schneider and Pressley, 1997; Parkin and Streete, 1988). Later on, this claim of developmental invariance was contradicted in several instances (Mecklenbräuker et al., 2003; Thomas et al., 2004; Barry, 2007; McNealy et al., 2010). In most cases where developmental differences in implicit learning have been found, young adults out-performed children. However, it appears that in at least some instances, the SL mechanisms of juvenile organisms may be more efficient than those of older ones (McNealy et al., 2010; Johnson and Wilbrecht, 2011); in natural language, this is evidenced by the difficulty with which adults acquire a second language (Gordon, 2000) compared to infants who can display efficient bilingual learning skills (Werker, 2012). Some proposals take the somewhat paradoxical stance that cognitive limitations may confer a computational advantage for learning, which may provide an alternative explanation for the presence of sensitive periods in language development (Newport, 1990; Elman, 1993; Conway et al., 2003). Additional research is needed to explore these ideas further.

In terms of how SL abilities develop later in life, the literature from the elderly population points either to no change in old age in the case of deterministic sequences (Howard and Howard, 1989, 1992; Frensch and Miner, 1994; Cherry and Stadler, 1995; Salthouse et al., 1999) but age-related deficits when sequences are probabilistic or have rather complex structures such as long range dependencies (Curran, 1997; Howard and Howard, 1997; Feeney et al., 2002; Howard et al., 2004). According to “the frontal lobe hypothesis of cognitive aging” (Hess, 2005), this deficit could stem from atypical activation of the dorsolateral prefrontal system, resulting in failures to properly represent and maintain context information (Braver et al., 2001), which in turn might be due to reduced working memory performance.

The model of Pierrehumbert (2003) takes clearly into account the developmental aspect. The author proposes that bottom-up mechanisms, including SL mechanisms – that encode concrete features of sequences – would be the main component of speech processing strategy in infants. Later on, with increased exposure to linguistic materials, this strategy would allow the development of categorizations at higher levels of the phonetic system, which in turn, would trigger top-down feedback mechanisms. Consistent with this model, children show evidence of categorization of the speech stream rather early, by age three (Nittrouer, 1996) and Hazan and Barrett (2000) showed that categorization of consonants in minimal pairs such as boat/goat continues to develop between 6 and 12 years. At age 12, such categorizations have still not reached young adult levels. According to Pierrehumbert, these later developments would result from top-down feedback mechanisms within the phonological system requiring a long process of elaboration and refinement. These top-down mechanisms would explain how initial preconscious levels of representation are progressively refined from childhood to adulthood. Such top-down accounts of SL mechanisms imply that low-level mechanisms of SL do not provide a full picture of the SL in adults and require one to take into account interactions between a more “basic” SL mechanism and information received from higher-level systems of the phonological system. Along this line,

one may hypothesize the existence of two types of SL mechanisms: a “basic” and an “expert” mechanism. Infants would benefit almost exclusively from the former, while children, adolescents, and young adults would benefit from the latter becoming increasingly developed as age increases into young adulthood. In older adults, however, the “expert” mechanism, presumably drawing upon working memory resources, might show signs of deficiency.

Thus, similarly to the dissociation of mechanisms of SL into explicit and implicit components, and into mechanisms encoding concrete and abstract representations, the Pierrehumbert (2003) developmental account of SL incorporates two systems that develop differentially. Such a multiple mechanism view of SL is consistent with Gervain and Mehler’s (2010) suggestion that a combination of language-specific, perceptual, and statistical learning mechanisms are all necessary for learning language (Gervain and Mehler, 2010). In their ACCESS model, these elements are combined together with social cues to explain language acquisition performance across the early life-span. Some learning mechanisms would work only on short time-scales while others would require the link of information at longer time-scales (Goldstein et al., 2010). Over short time-scales, infants would use surface structure such as transitional probabilities to extract co-located sequences of phonemes from a continuous input (Saffran et al., 1996; Pelucchi et al., 2009). Over longer time-scales, infants may benefit from social cues, such as parents’ use of common grammatical constructions and incorporate them in their own speech (Cameron-Faulkner et al., 2003). Importantly, such developmental models (i.e., involving multiple mechanisms) have received support from neuroimaging data. For instance, Thomas et al. (2004) provided evidence of a maturation of two distinct mechanisms of SL between childhood and adulthood: a process acting on unconscious representations and another that manipulates explicit knowledge.

In summary, SL may consist of at least two different systems. The first relies upon bottom-up implicit/perceptual mechanisms that result in unconscious representations, develop early in life, and are likely to exploit surface structure of input and hence can explain some of the impressive language-related abilities present in infants and children. The second system develops later in life, consisting of expert SL mechanisms that rely more on top-down information, are more dependent on the level of attention, and result in explicit knowledge of abstract rules that further improves language processing abilities (but see Marcus et al., 1999, suggesting that abstract information may already be processed by 7-months old as well). Thus, rather than a simple explanation of how a single SL ability progresses over time, it may be necessary to consider at least two different sub-systems and associated mechanisms to draw a complete picture of the developmental trajectory of SL. Understanding how each of these processes develops and interacts dynamically across the life-span remains a formidable research challenge. Based on the preceding discussion, we propose an initial and albeit simplified model showing the developmental progression of these two SL systems (Figure 2). In order to provide extra empirical validation of this model, we now turn to how ERPs have

contributed to a better understanding of the mechanisms of SL.

EXPLORING SEQUENTIAL LEARNING WITH EVENT-RELATED POTENTIALS

We will first summarize the main ERP paradigms that have been used to date in SL research (the main ERP components are described in Figure 3). We will then focus on how ERPs have been used to explore the three above-mentioned dimensions of SL mechanisms: the abstractness of the manipulated representation, the level of attention/consciousness of the mechanisms and the level of consciousness of the representations, and the development of SL across the life-span. After considering these three dimensions of SL, we then consider new avenues of research and then conclude with a re-evaluation of the two-system model of SL described in Figure 2.

MAIN ERP PARADIGMS OF SL RESEARCH

Oddball and SRT paradigms

A rather basic paradigm for testing a simple form of SL, referred to as the “Oddball” paradigm, contains a rare (or “deviant”) target stimulus presented along with more frequent (or “standard”) non-target stimuli in a serial input stream (Figure 4). This paradigm elicits a P300 ERP component, one of the most studied components of ERP research (for a review, see Polich, 2007). The P300 is thought to reflect a decision based on an evaluation or categorization of the stimulus. The amplitude of the P300 is highly sensitive to the stimulus probability and to the level of attention. In the oddball paradigm, the number of repetition of standards between two occurrences of a (target) deviant is randomized, such that the length of the sequence of interest is not fixed, but random. The perceiver is thought to “compute online” a conditional probability of the target occurrence. Stadler et al. (2006) were able to show how decision and preparatory mechanisms are affected by this conditional probability, by measuring the P300 and the contingent negative variation (CNV, Walter et al., 1964), respectively. In this paradigm, the target cannot be predicted by the occurrence of a given stimulus. However, as the number of consecutive standards increases, the probability of occurrence of the target increases too, which increases the likelihood of a motor response requirement, hence affecting: (1) the level of attention and/or motor decision mechanisms (as reflected by the P300), and (2) the amount of motor preparation (as reflected by the CNV). Stadler et al. interpreted their results as an indication that the level of activation of decision mechanisms indexed by the P300 were continuously increasing as the target conditional probability increased while the activation of preparatory motor mechanisms according to the CNV was much like an all-or-none phenomena.

Another well studied ERP component elicited by the oddball paradigm is the mismatch negativity (MMN), which typically is thought to reflect an automatic discrimination or echoic memory updating between the standard and the deviant stimulus (for a review, see Näätänen et al., 2012). Capitalizing on the fact that the MMN is less dependent on the level of attention than the P300, van Zuijen et al. (2006) recorded these two components simultaneously with an oddball paradigm to explore how the

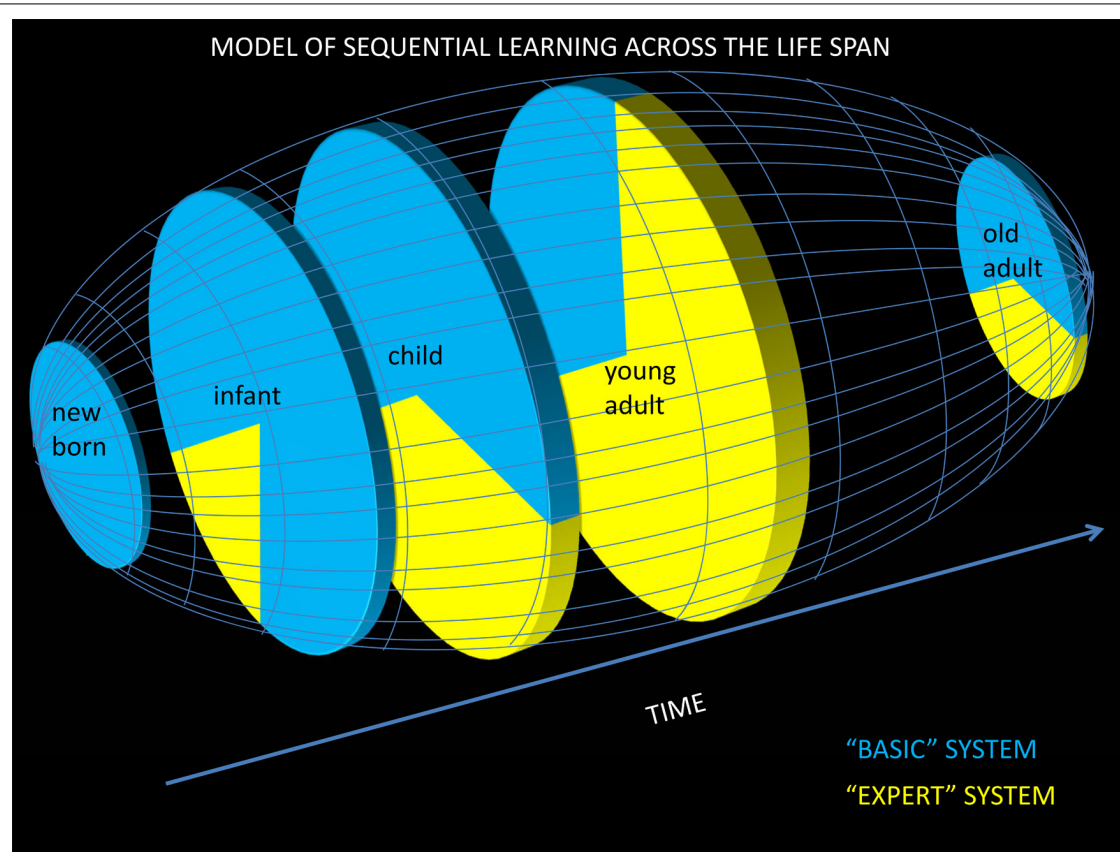


FIGURE 2 | Model of SL across the life span. We propose that SL is governed by two systems: a “basic” and an “expert” system. The “basic” system incorporates modality-specific predictive mechanisms that are mostly automatic and implicit and that capture concrete structures of sequences such as chunks and transition probabilities through a bottom-up process. The basic system, which is possibly a sub-system (in the temporal domain) of the (spatio-temporal) PL system, can be modeled by simple recurrent networks. The “basic” system is already available very early in life, allowing for the development of explicit long-term associative memories that become available to the expert SL system. The “expert” system, which relies on

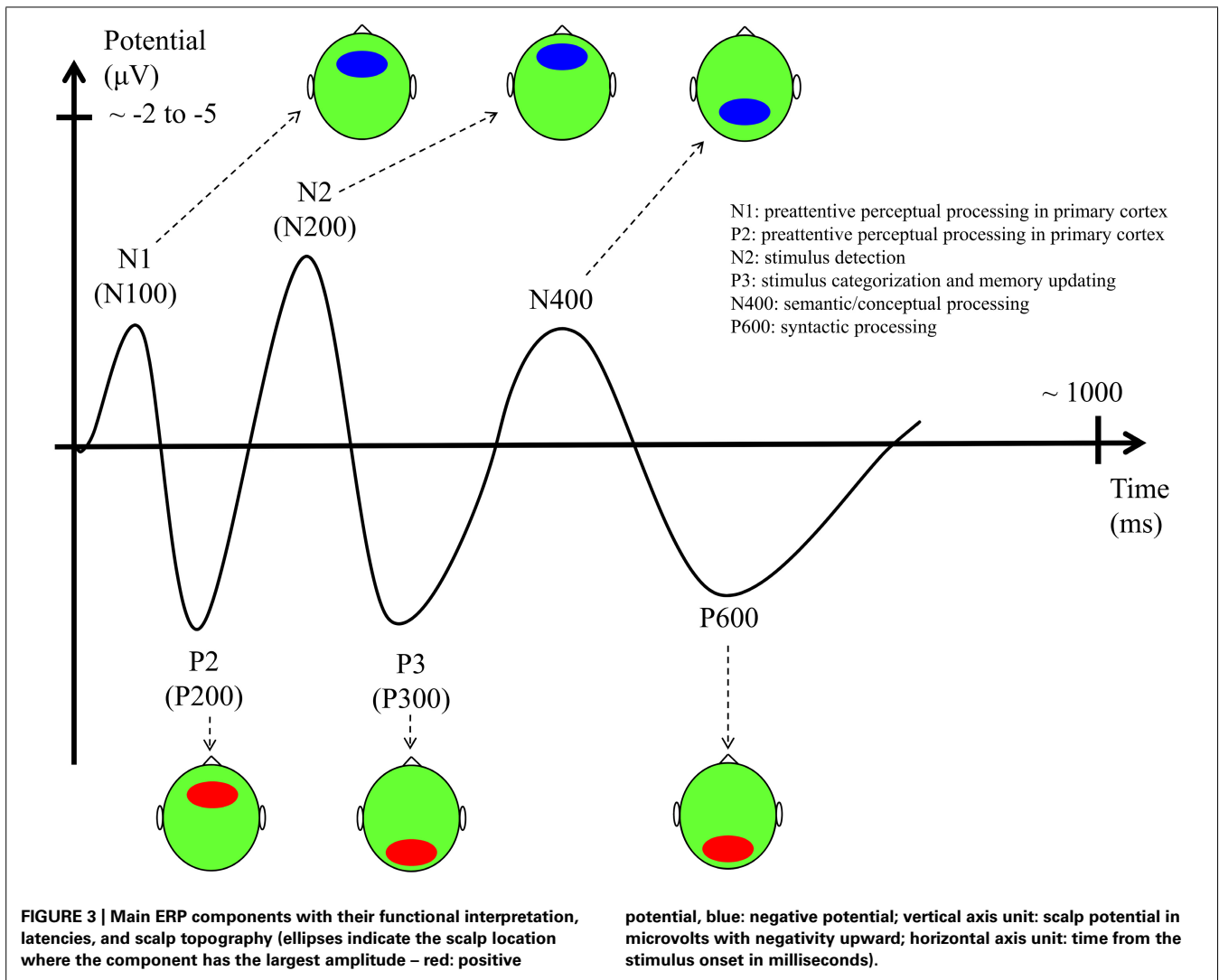
top-down explicit multimodal and retrospective mechanisms, depends on the level of intention (to learn) and attention (including selective attention through social cues). The “expert” system, which captures more abstract patterns, increasingly develops from childhood into adulthood and then declines in old age because of impaired working and sensory memories. Blue represents the proportion of SL governed by the basic system and yellow represents the proportion of SL governed by the expert system. Clearly, this model is tentative and highly speculative. In particular, the exact degree of contribution of the basic and expert systems at different ages of life remain currently unknown.

level of attention affects SL (more on this study in a subsequent section).

Some researchers have taken the standard oddball paradigm and used it to study SL processes that occur during the serial reaction time task (SRT; Nissen and Bullemer, 1987). The typical SRT task is a visuo-motor SL task where visual stimuli appear at different locations on a screen, as described by a particular rule or pattern (Figure 5). Response buttons correspond spatially to each location. SL is behaviorally demonstrated by a reduced response time to repeating/familiar sequences compared to novel or random sequences. The SRT has been subsequently adopted and modified by many others for various purposes (Cleeremans and McClelland, 1991; Perruchet and Amorim, 1992; Willingham et al., 1993; Reed and Johnson, 1994; Stadler, 1995; Jiménez et al., 1996; Perruchet et al., 1997; Frensch et al., 1998; Honda et al., 1998; Reber and Squire, 1998; Shanks and Johnstone, 1999; Destrebecqz and Cleeremans, 2001). Most relevant to the present purposes, the SRT has also been used with ERP

recordings, revealing ERP correlates of SL (Eimer et al., 1996; Baldwin and Kutas, 1997; Rüsseler and Rösler, 2000; Rüsseler et al., 2003a; Ferdinand et al., 2010; Meiri, 2011). Specifically, under an oddball-type version of the SRT that involves the presentation of deviant stimuli occurring in a sequence of standards, an enhancement of the N200 to deviants compared to standards has been reported (e.g., Eimer et al., 1996; Rüsseler and Rösler, 2000; Schlaghecken et al., 2000). Note that an important question is whether this modulation stems from SL *per se* or from a secondary effect of SL, for instance, an effect of attention. We will also come back to this issue in a subsequent section of this review.

One final variation of the oddball design comes from Jost et al. (2011). This paradigm included sequences of visual stimuli (colored circles) containing a frequent stimulus and a set of “deviant” stimuli. These deviants belonged to two different categories: “predictors” and “targets” (Figure 6). The participant is asked to respond to target stimuli without being told that certain

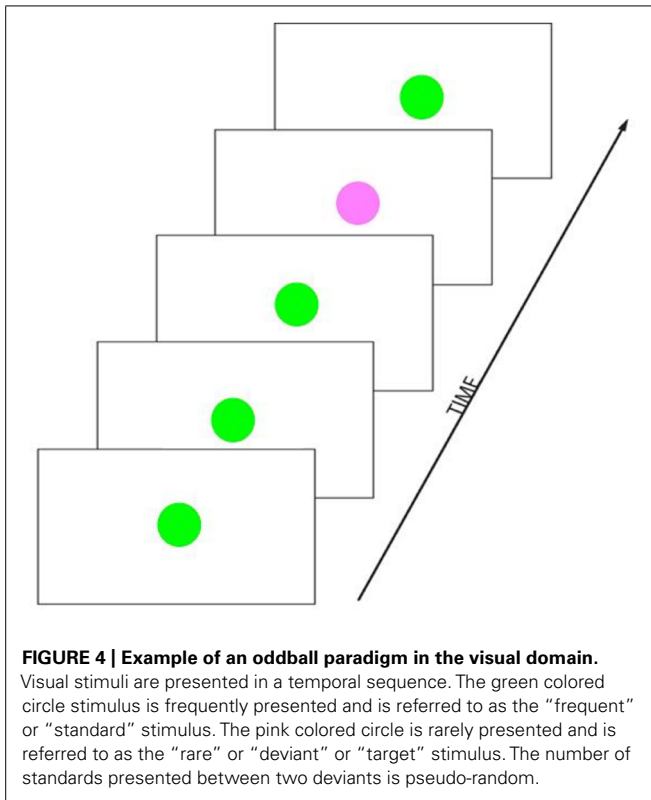


predictor stimuli predict the occurrence of the target with fixed contingent probabilities. That is, the occurrence of the predictor allows the participant to predict the target with varying probabilities. The assumption is that this design requires a kind of basic statistical learning of the contingent probabilities that links the predictors to the targets. Jost et al. (2011) reported a late positivity in response to the predictors between 300 and 600 ms post-predictor onset that increased as the contingent probability increased. This ERP effect was referred to as a P300-like component and interpreted as reflecting an index of SL. Similarly, Rose et al. (2001) reported an SL effect as reflected by an increased P300 to the first stimulus of a two-item sequence. According to these authors, since the task required a motor response to the second item, the ERP to the first item was also modulated by: (1) an increased lateralized readiness potential component (LRP, e.g., Hackley and Valle-Inclán, 2003), reflecting an increased motor preparation to the predictable second item (see also Eimer et al., 1996; Rüsseler et al., 2001), and (2) a decreased CNV, reflecting a reduced motor preparation to other alternative, non-predictable second items.

Unlike these oddball paradigms where the sequences embody rather simple contingent statistics, other ERP paradigms have been used to explore SL using more complex sequences, such as the “artificial grammar” paradigm.

Artificial grammar and natural language paradigms

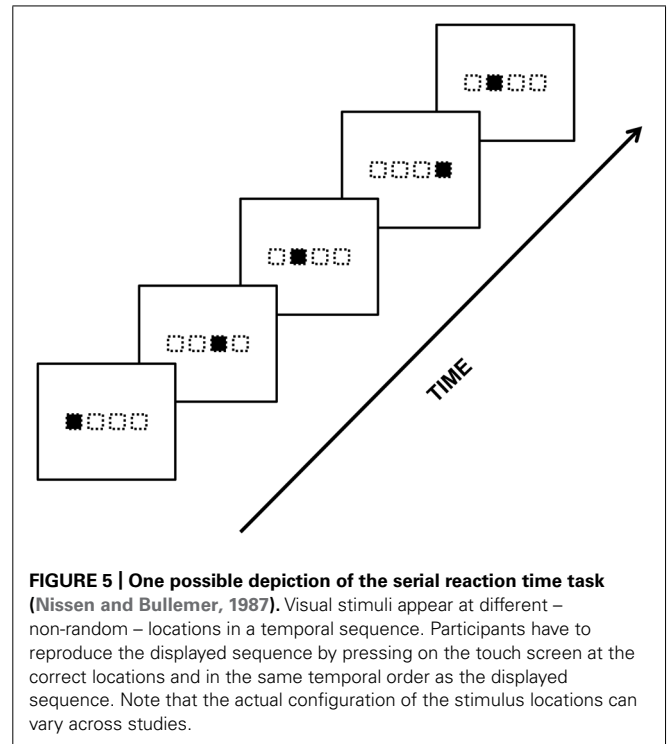
Artificial grammar learning (AGL) paradigms, which incorporate a set of rules that govern the structure of sequences (Figure 7), have been designed to mimic the complex structure of natural language while simultaneously removing other potentially confounding parameters such as semantic information. Converging evidence has suggested that this experimental design is a good model for testing the grammatical and structural processing of natural language (for a review see Christiansen et al., 2002). It should be noted that the AGL paradigms used in ERP research often incorporate aspects of the SRT paradigm, described above (Nissen and Bullemer, 1987). In such a combined SRT-AGL task, the structure of the sequence of stimuli follows the rules defined by an artificial grammar to determine what stimulus occurs next in the sequence.



The ERP research using AGL has shown that several ERP components known to index grammar/syntactic violation in natural language (e.g., Steinhauer et al., 2001) and in music perception (e.g., Patel et al., 1998) are also elicited by artificial grammar violations (Osterhout and Holcomb, 1992; Christiansen et al., 2012; Tabullo et al., 2013). The most commonly reported ERP indices of syntactic violation are an “early” negativity and a “late” positivity. The early negativity is usually found at left anterior cortical sites and between 200 and 400ms poststimulus-onset (but see for instance Hoen and Dominey, 2000), and hence is often referred to as the early left anterior negativity (ELAN) (e.g., De Diego Balaguer et al., 2007; Mueller et al., 2008). The late positivity, being often maximal around 600ms is usually referred to as the P600 (e.g., Steinhauer et al., 2001).

Using such AGL paradigms, it is possible for instance to test whether SL is processed by different mechanisms for different sequence structures. For instance, Bahlmann et al. (2006) reported two ERP components to grammar violation of CV syllables sequences, an early negativity within a 300–400 ms window that was evoked only by local violation [in (AB)_n sequences] and a late positivity within 400–750 ms that was evoked by both local and longer range violation (in center-embedded AnB_n sequences). These ERP results confirm earlier predictions of the existence of different cognitive mechanisms engaged for the processing of different types of input structures (e.g., Conway and Christiansen, 2001).

Other ERP components have also occasionally been reported as indices of SL during exposure to artificial grammars: the error-related negativity (ERN, Gehring et al., 1993), the N200, the slow



negative wave (SNW), and the N400. Rüsseler et al. (2003b) used an Erikson-like flanker task wherein a central imperative letter followed a sequence or was randomly chosen and reported sequence error monitoring as reflected by the ERN. This finding suggests that the detection of (artificial) syntactic violations is cognitively processed as a specific instance of a more general set of errors, as reflected by the ERN. Lang and Kotchoubey (2000) using an AGL paradigm based on sequences of vowels within a passive task not requiring a motor response reported two frontally distributed ERP effects to rule violations: one at a latency of 250 ms – a N200 – and another around 500 ms – a SNV. Lang and Kotchoubey (2000) suggest that this SNV may in fact be an instance of the “family” of N400 components. This would be in line with other studies that also propose the N400 (Kutas and Federmeier, 2011) as an index of SL processes (Sanders et al., 2002; Cunillera et al., 2006, 2009; Carrión and Bly, 2007; De Diego Balaguer et al., 2007; Abla et al., 2008; Buiatti et al., 2009).

The AGL paradigms allow one to test SL mechanisms independently of the effects of other language processes. However, natural language paradigms remain useful even with these potential confounds, as they allow one to better understand how SL might directly contribute/interfere with language processing. The research using natural language paradigms has mainly reported an ELAN and a left anterior negativity (LAN; for an overview see Friederici, 2002) as well as a P600 (e.g., Osterhout and Holcomb, 1992) as markers of syntactic violations. For instance, Friederici et al. (2002) reported a similar P600 and ELAN to artificial and natural language grammar violations in native-speakers. This result suggests that adult who are learning a new (artificial) language use the same learning mechanisms as

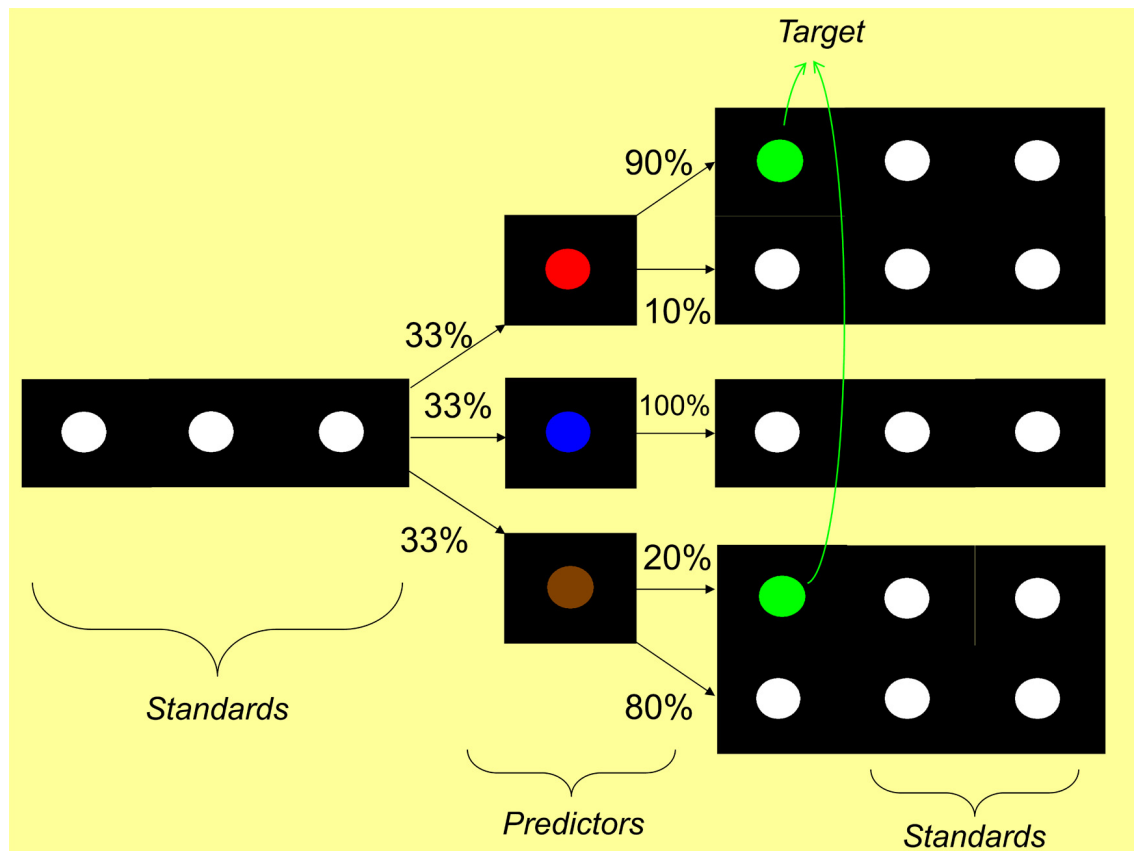


FIGURE 6 | Modified oddball paradigm of Jost et al. (2011). The standard stimulus is a white circle on a dark background. The paradigm comprises several deviant stimuli belonging to two different categories: “predictor” and “target.” Participants are asked to press a button when the target is presented. There are three types of predictors (corresponding to the three experimental conditions): a “high probability” predictor which is followed 90% of the trials by the target, a “low probability” predictor, followed 20% of

the trials by the target, and a “zero probability” predictor, which is never followed by the target. Participants are not told about these predictor-target variable statistical contingencies. SL is observed behaviorally when performance improves with higher statistical contingency. SL is observed neurophysiologically when the ERP to the predictors differ between the experimental conditions (e.g., a larger amplitude for the high probability predictor compared to the other two predictor types).

are used in natural language. A similar conclusion comes from Mueller et al. (2005), who reported similar ERP patterns from non-native Japanese speakers trained to learn a “Mini-Japanese” compared to native Japanese speakers. Similarly, Christiansen et al. (2012) found a P600 to syntactic violations in artificial grammars and natural language paradigms in the same set of participants. The amplitude of the P600 was correlated between the two tasks, suggesting that identical or similar underlying mechanisms were engaged in both non-linguistic SL and natural language processing. These studies suggest that a successful methodological approach is to combine the AGL and natural language paradigms in order to more fully understand SL and natural language processing.

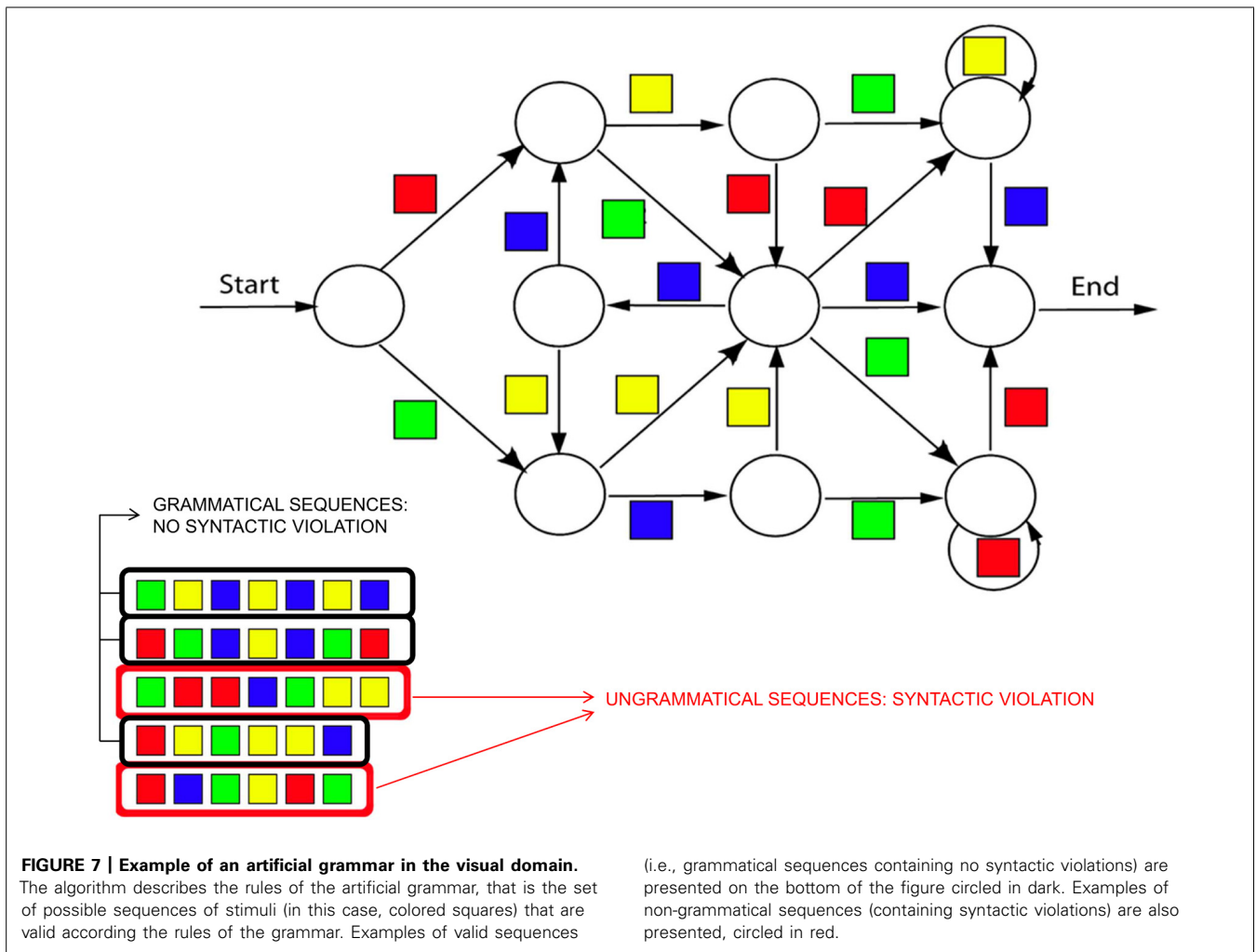
In summary, several ERPs components, such as the N200, the MMN, the N400, the ERN, the ELAN, the LAN, the P300, and the P600 seem to be modulated by SL in various experimental paradigms and hence may be used to better understand the cognitive mechanisms underlying SL. The variety of relevant paradigms ranges from simple sequence designs such as oddballs to more complex sequential stimuli involving natural

or artificial grammars. We now consider to what extent the ERP research helps elucidate questions about the underlying mechanisms of SL and the associated representations from the perspective of the three dimensions previously discussed: the level of abstractness, the level of attention or consciousness (i.e., implicit versus explicit mechanisms), and the developmental trajectory.

WHAT THE ERP FINDINGS TELL US ABOUT SL

ERP findings: level of abstractness

As previously discussed, SL is thought to stem from at least two different types of mechanisms, one that acts on rather concrete information and the other that acts on more abstract information. With concrete feature-learning mechanisms, SL is explained by the encoding of distributional properties of the sequence of items, such as item co-occurrences or the transitional probability between items. The alternative (or complementary) mechanism assumes that the perceiver encodes abstract rules (or discrete combinatorial systems).



One of the crucial results from ERP studies of SL is provided by Pulvermüller and Assadollahi (2007), who attempted to dissociate ERP correlates of SL mechanisms between those that process concrete versus abstract information. To this aim, these authors manipulated separately concrete features (item co-occurrences or transitional probability) and abstract features (syntactic rules or grammaticality) of sequences using ungrammatical word strings, very rare grammatical word strings (i.e., with low co-occurrence and low transitional probabilities), and common grammatical word strings (i.e., with high co-occurrence and high transitional probabilities). Pulvermüller and Assadollahi reported a magnetic MMN that differed between grammatical and non-grammatical word strings but was unaffected by the co-occurrence (or transitional probability) manipulation. These authors concluded that natural language grammar learning would stem from the encoding of discrete combinatorial systems (i.e., abstract rules) rather than the learning of co-occurrence and/or transitional probability (i.e., concrete features). However, an alternative interpretation could be drawn from their data: both mechanisms processing concrete and abstract features might occur during syntactic processing, but the magnetic MMN could be more sensitive to abstract compared to concrete features encoding. Put another

way, just because an ERP correlate was not observed for concrete feature learning does not mean that such a correlate does not exist; null effects in ERP research are notoriously difficult to interpret.

Lelekov et al. (2000) were also able to explore the issue of the level of abstractness of the information encoded during SL. Using an AGL paradigm, they presented instances of sequences of type ABCBAC and DEFEDF with different surface structure (i.e., different concrete distributional properties) but identical abstract structure. These authors reported a late positivity at 500ms, similar to the typical P600 to syntactic violation, in response to abstract structure violation, but no ERP effect to surface (concrete) structure violation. As with the Pulvermüller and Assadollahi's (2007) study, at least two conclusions could be drawn: either only SL mechanisms of abstract structures occur or both concrete and abstract structures are processed by the mechanisms of SL but in their paradigm the ERP are mainly sensitive to those mechanisms that act on abstract information and less sensitive to those related to concrete feature encoding.

Conversely, other studies have found ERP correlates – specifically, the MMN – related to concrete feature encoding (Deouell et al., 1998; Marco-Pallarés et al., 2005; Schröger et al., 2007). For

instance, Schröger et al. (2007) used standard and deviant tone pairs of different frequencies, either ascending or descending. The first tone of the pair had either a fixed frequency of 900 Hz or a random frequency within 600–1200 Hz using 10 Hz-steps. The second tone of the pair had a short or a long duration (200 or 400 ms). Schröger et al. (2007) referred to the condition with a fixed-frequency first tone as sequences with a “concrete rule” and to the condition with a random-frequency first tone as sequences with “abstract rules.” ERPs were time-locked to the second tone of the pairs. Schröger et al. (2007) reported a MMN to deviant pairs with both concrete and abstract sequences. These authors also used source localization analyses and concluded that the MMN sources elicited by abstract and concrete rule violations involved a similar neural network.

In summary, some of the few ERP studies that explored the level of abstractness of the encoded information during SL have been interpreted as evidence that SL is governed only by abstract rule-learning mechanisms. On the other end, other ERP research were taken as evidence that concrete-rule encoding can also be indexed by ERPs. Overall, it appears that the ERP research supports the assumption that both concrete and abstract feature encoding occurs in SL. The apparent inconsistency between these studies may be due simply to variation in experimental designs and a lack of sensitivity of ERP to adequately index particular mechanisms of SL.

ERP findings: level of attention and conscious awareness

Across various paradigms, not just those specifically looking at SL, almost all ERP components have been reported to be modulated by the level of attention (Kok, 2000; Barry et al., 2003; Correa et al., 2006). Thus, one might consider the possibility that several studies that interpreted ERP components as markers of SL were in fact pointing to a (top-down) attentional effect that may or may not be specific to SL itself. For instance, Sanders et al. (2002) reported an increased N100 to learned/segmented pseudowords compared to new/unfamiliar pseudowords with exposure to a speech-like stream of unfamiliar pseudo-words. These authors concluded that the N100 is an index of SL (or segmentation). However, an alternative top-down account of this result could be that, as pseudo-words become more and more familiar due to SL (or segmentation), the pseudo-words are better recognized, and hence are more likely to capture attention. The increased N100 across exposure to a speech-like stream would thus reflect a top-down attentional effect to items of this stream. If this attentional effect is indeed occurring, an important question is whether it contributes or not to the actual process of SL (or segmentation) itself.

The literature contains several other ERP studies that attribute to ERP components the property of indexing SL while often ignoring the alternative top-down attentional explanation (Rose et al., 2001; Sanders et al., 2002; Cunillera et al., 2006, 2009; De Diego Balaguer et al., 2007; Abia et al., 2008). For instance, Abia et al. (2008) and Sanders et al. (2002) interpreted an increased N100 and N400 to segmented/learned sequences of three items [tones in Abia et al. (2008) and syllables in Sanders et al. (2002)] as reflecting the indexing of SL mechanisms. Similarly, Rose et al. (2001) reported an increased P300 with SL to the first item of

a sequence of two items and several other SL studies concluded that the P200 is a marker of SL (Cunillera et al., 2006, 2009; De Diego Balaguer et al., 2007). As with the case of the Sanders et al. (2002) study, all of these ERP effects could instead be due to modulations of the level of attention, rather than SL *per se*. However, even if this top-down account is true, the ERP components still reflect an outcome of the SL process, that is, a learning-related change of attention to stimuli based on whether or not the stimuli are consistent with the previously learned patterns.

Whether SL requires conscious awareness is a hotly debated topic. The relation between implicit SL, explicit SL, and ERPs has been mostly explored through two approaches: by dissociating implicit from explicit learning according to whether participants acquired explicit knowledge of the patterns (e.g., Eimer et al., 1996; Baldwin and Kutas, 1997; Rüsseler and Rösler, 2000; Schlaghecken et al., 2000; Rüsseler et al., 2001) or by dissociating these two types of learning according to whether participants had or had not an intention to learn the rules (e.g., Rüsseler et al., 2003a,b).

In line with early behavioral studies of SL (Reber, 1967; Saffran et al., 1996, 1997), several ERP studies, using different experimental approaches, provide strong evidence that there is at least an implicit component of SL (Saarinen et al., 1992; Dell'Acqua et al., 2003; Carral et al., 2005; Kessler et al., 2005; Zachau et al., 2005; Kranczioch et al., 2006; van Zuijen et al., 2006; Trippe et al., 2007; Acqualagna et al., 2010; Yu et al., 2011; Batterink and Neville, 2013). Indeed, several ERP studies concluded that SL had occurred under conditions of minimal attention (Saarinen et al., 1992; Carral et al., 2005; Zachau et al., 2005; van Zuijen et al., 2006). For instance, van Zuijen et al. (2006) investigated the attentional issue by recording the MMN (assumed to reflect attention-independent discrimination, but see Arnott and Allan, 2002; Müller et al., 2002) and the P300 (assumed to be more dependent on the level of attention, but see Bennington and Polich, 1999). They used an oddball paradigm wherein standards are tone pairs with an ascending frequency and deviants are tone pairs with a descending frequency. Participants who after the ERP session did not report the presence of deviants, i.e., were subjectively unaware of them, showed only a MMN, while participants who were aware of the deviants showed also a P300. These findings suggest that both implicit and explicit SL can occur, each recruiting different neural mechanisms. In a study similar to van Zuijen et al. (2006), Gottselig et al. (2004) tested SL using an oddball paradigm containing eight-tone sequences [instead of tone pairs in van Zuijen et al. (2006)]. Deviant sequences differed from standard sequences only by the frequency of one tone. Similar to van Zuijen et al. (2006), Gottselig et al. (2004) were also able to record a MMN to deviants while participants' attention was focused on silent films, thus suggesting again that implicit SL of very basic input sequences is possible.

Still using the oddball paradigm and measuring the P300, but under rapid stimulus presentation - the so-called RSVP paradigm - other studies tested the perception of a deviant within an attentional blink (Dell'Acqua et al., 2003; Kessler et al., 2005; Kranczioch et al., 2006; Trippe et al., 2007; Acqualagna et al., 2010; Yu et al., 2011). These studies concluded that there was an implicit component of SL. A similar conclusion was also reported using

AGL paradigms (Baldwin and Kutas, 1997; Schröger et al., 2007) and syntactic violations within natural language (Batterink and Neville, 2013). For instance, Batterink and Neville (2013) reported early ERP deviations to such syntactic violations while the participant's attention was focused on a distractive task. This result indicates that SL of more complex rules than those found in an oddball paradigms might also be processed implicitly.

Importantly, none of the above-mentioned ERP studies rule out the possibility that explicit mechanisms of SL also contribute to the reported ERP effects. Indeed, the ERP research on SL mechanisms has abundantly explored the explicit component(s) of SL (Tiitinen et al., 1994; Eimer et al., 1996; Baldwin and Kutas, 1997; Rüsseler and Rösler, 2000; Schlaghecken et al., 2000; Rüsseler et al., 2003a,b; Miyawaki et al., 2005; Schröger et al., 2007). For instance, Schröger et al. (2007) reported a combination of implicit and explicit SL using violations of abstract auditory rules. Standard and deviant tone pairs of different frequencies were used, in which deviant and standard pairs could have either ascending or descending frequency and the second tone of the pair had a short or a long duration (200 or 400 ms). They manipulated the effect of attention on the rules by using three conditions: (1) a passive (i.e., no task) “ignore” condition wherein participants are asked to watch a soundless video, (2) an active rules task-irrelevant “distraction” condition wherein participants were asked to perform a two alternative-forced choice discrimination decision on duration, judging whether the second tone of each pair was short or long, and (3) an active rules task-relevant “detection” condition wherein participants were asked to detect deviant pairs after having been informed of the rising/falling frequency rule. Schröger et al. (2007) not only confirmed the above-mentioned reports of an implicit component of SL showing ERP effects to deviants modulated by the participants' performance on a non-rule related task, they also provided findings regarding the effect of the participant's intention. Schröger et al. (2007) results suggest that intention to learn improved the ability to perform the non-rule related task. All together, these data suggest that SL can be both implicitly and explicitly learned, depending on the participants' intention. A similar effect of the intention to learn sequences was found by Miyawaki et al. (2005). These authors presented sequences of eight digits and found that, after training, the amplitude of the N200 component (and behavioral performances in sequence free and cued recall) were higher with intention to learn compared to non-intention.

In addition, larger effects of learning (as measured by behavior and ERP) appear to be found in explicit compared to implicit conditions. For instance, Baldwin and Kutas (1997) provided evidence that behavioral measures of SL were roughly twice as large for explicit compared to implicit SL (**Figure 8**). In addition, these authors reported P300 effects to sequence violations that were, when explicit SL occurs, more than two times larger than those observed when only implicit SL was permitted (**Figure 8**). A similar “effect size doubling” on behavioral performance was reported by Eimer et al. (1996, see **Figure 9**) using 10-letter sequences with standard and deviant sequences. The effect size increase was even larger when measuring the amplitude of the N200. In the same vein, Rüsseler and Rösler (2000) and Schlaghecken et al. (2000),

reported N200 and P300 modulations to sequence violation only in participants that learned explicitly the sequence [according to post-experimental free recall and recognition tests in the Rüsseler and Rösler's (2000) study, and according to the “process dissociation procedure” of Jacoby, 1991 in the study of Schlaghecken et al. (2000)].

However, robust effects of explicit SL are not systematically reported. For instance, Rüsseler et al. (2003b) found similar behavioral and neurophysiological effects in implicit and explicit conditions. Rüsseler et al. (2003b) measured the ERN while participants performed an Erikson-like flanker task wherein a central imperative letter followed a sequence or was randomly chosen. The lack of difference between these conditions is likely to stem from the use of a rather unusual SL paradigm. Indeed, using a more typical SL paradigm with 16-letter-long sequences irregularly disrupted by deviant stimuli, Rüsseler et al. (2003a) were able to show a strong effect of intention on ERP effects of SL. These authors reported ERP effects on the N2b- and P3b-components only in participants who were informed of the presence of sequences and no ERP effects in a group of participants who were not previously informed of these stimulus patterns.

In summary, the ERP literature seems to support the existence of both implicit and explicit mechanisms of SL. Furthermore, the effect size of the SL measured behaviorally or neurophysiologically appears to increase with the intention to learn the rules and with the explicit knowledge of these rules. Therefore, when attempting to understand the mechanisms of SL, a very critical aspect appears to be the attentional/consciousness dimension. Importantly, since the level of attention can affect almost all ERP components, the interpretations of ERP correlates of SL must be cautious as in some instances there may be an alternative top-down explanation.

ERP findings: developmental trajectory

In general, there is a paucity of ERP research examining SL in young children. However, neural signatures of infant and children's early language learning mechanisms – presumably dependent in part on SL – have been documented using ERPs. Indeed, ERP studies have provided some evidence that the ability to extract statistical dependencies between adjacent elements in the speech stream appears to be present from birth, and infants can learn non-adjacent dependencies in a natural, non-native language by 4 months of age (Teinonen et al., 2009; Friederici et al., 2011). From about 9 months of age, familiar words evoke responses that are different in amplitude as well as in scalp distribution measurements from responses to unfamiliar words (Molfese, 1990; Vihman et al., 2007). By 11 months of age, phonetic learning can already be observed; by 14 months, responses to known words are observed; and by 2.5 years, semantic and syntactic learning is elicited (Kuhl and Rivera-Gaxiola, 2008). For instance, a P600 to sentence-level syntactic violations has been found in 30, 36, and 48 months old children that looked rather similar to the P600 found in young adults (Silva-Pereyra et al., 2007).

Although SL is assumed to be important for language acquisition, few studies have directly examined the relationship between SL and language outcomes. Recently, the link between SL and children's language performance has received new support. Rosas

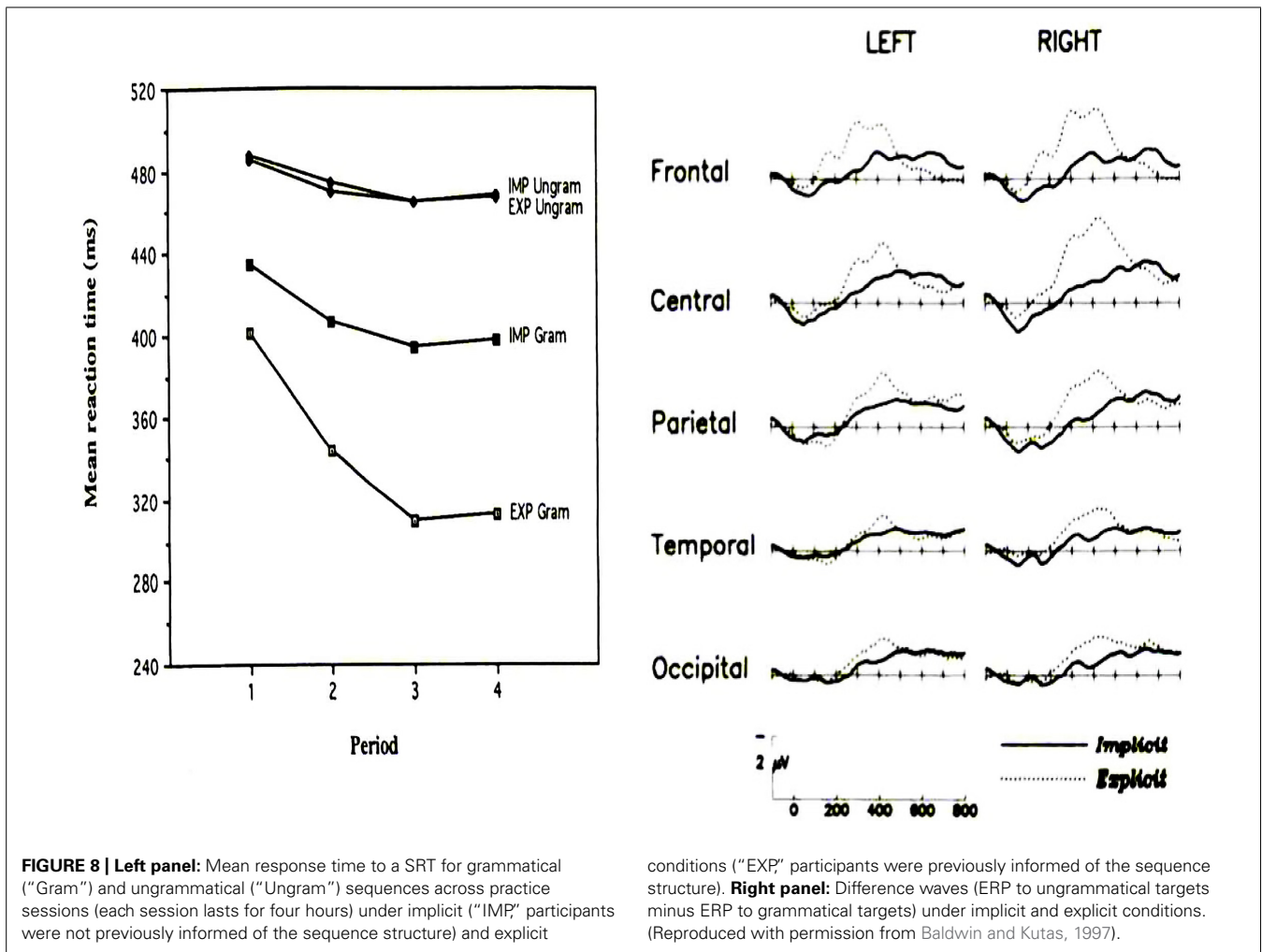


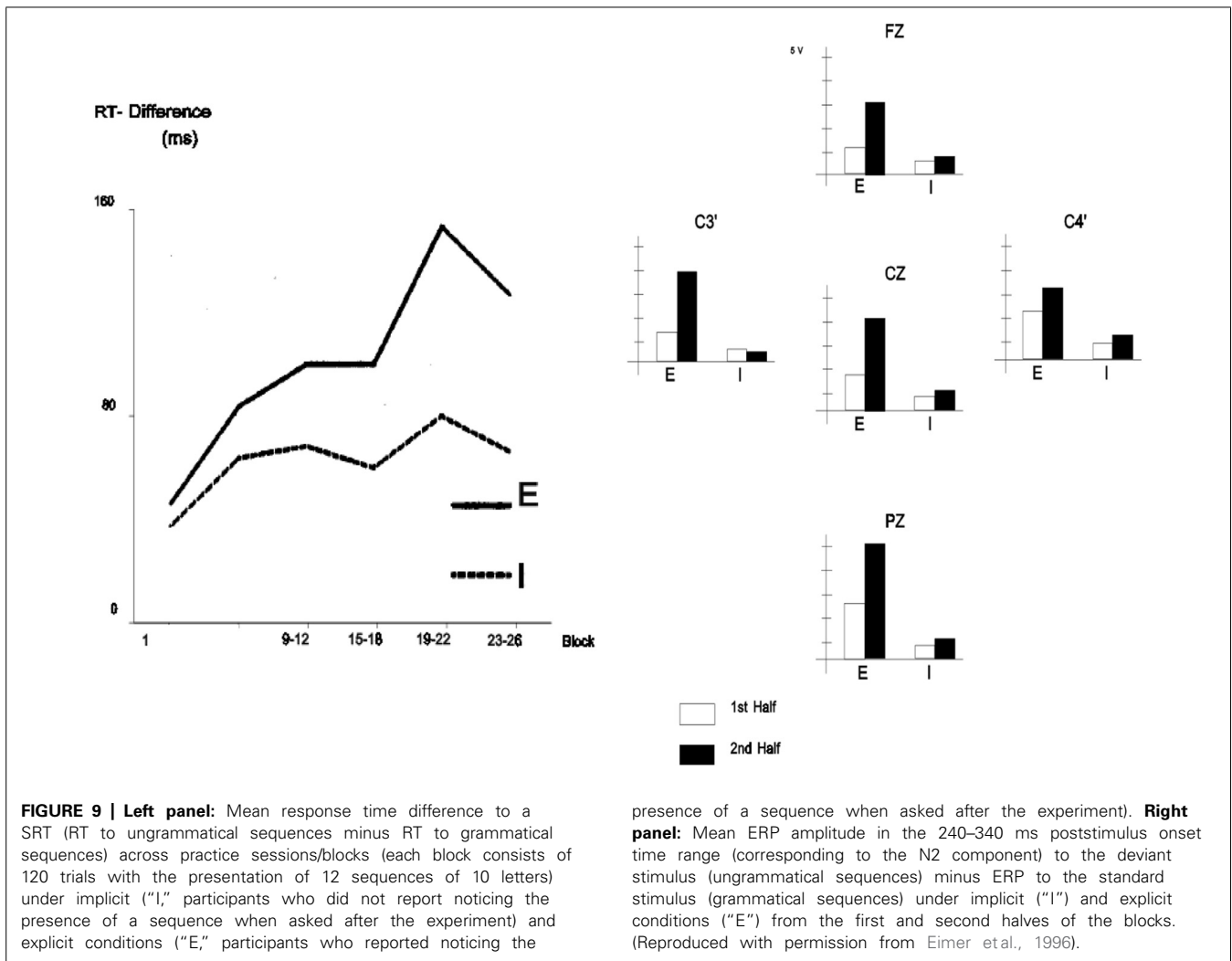
FIGURE 8 | Left panel: Mean response time to a SRT for grammatical ("Gram") and ungrammatical ("Ungram") sequences across practice sessions (each session lasts for four hours) under implicit ("IMP" participants were not previously informed of the sequence structure) and explicit

conditions ("EXP" participants were previously informed of the sequence structure). **Right panel:** Difference waves (ERP to ungrammatical targets minus ERP to grammatical targets) under implicit and explicit conditions. (Reproduced with permission from Baldwin and Kutas, 1997).

et al. (2010) reported an ERP study of SL in children (6–11 years) using visual sequences. The authors compared two groups of children: one with and one without attention deficit hyperactive disorder. Rosas et al. found that both behavioral and ERP findings pointed to the occurrence of SL in both experimental groups. However, their most striking ERP result seems to be a considerable difference in ERP amplitude between the two groups of children on a late positivity (between 400 and 800 ms post-stimulus onset) similar to the P600, suggesting that non-linguistic SL incorporates mechanisms also used for language learning (as reflected by the P600). The fact that the two groups differed on the magnitude of the P600 also suggests that differences in attention can modulate the P600 effects to SL in children. The relation between SL and natural language is predictive from a developmental perspective, as the early mastery of the sound patterns of one's native language provides a foundation for later language learning. Indeed, children who show enhanced ERP responses to phonemes at 7.5 months show faster advancement in language acquisition between 14 and 30 months of age (Kuhl et al., 2008).

As concerns older populations, the literature about ERP correlates of SL is scarce and mostly involves oddball paradigms that

elicit, for example, the MMN and the P300 (Fabiani and Friedman, 1995; Fabiani et al., 1998; Berti et al., 2013; Cheng et al., 2013). In line with behavioral data suggesting more age-related SL deficits for structures that include long range dependencies (Curran, 1997; Howard and Howard, 1997; Feeney et al., 2002; Howard et al., 2004), MMN studies show more age-related decline with interstimulus intervals larger than 2 s (Czigler et al., 1992; Pekkonen et al., 1993, 1996; Cooper et al., 2006; Ruzzoli et al., 2012) compared to shorter intervals (Cheng et al., 2013). This decline has been interpreted in terms of faster sensory memory trace decay in the older compared to the younger adults (Pekkonen, 2000; Näätänen et al., 2007). These results suggest that the behavioral studies showing age-related decline of SL due to impaired abilities to represent and maintain context information (Braver et al., 2001) might not only stem from working memory-related deficits but also from sensory memory impairments, as reflected by the MMN attenuation. Regardless as to whether or not working memory and sensory memory share underlying mechanisms (Jääskeläinen et al., 2011), these ERP studies of aging seem to point to an age-related impairment of memory systems that might in turn affect SL ability.



Clearly, the developmental trajectory of SL still has many unexplored fundamental questions. We believe the ERP technique has not been used to explore SL across the lifespan to its fullest potential. This research gap in the developmental dimension as well as opened questions left by the previously discussed models of SL lead us now to consider several new lines of ERP research that we believe could offer new insights into SL, some of which are amenable to developmental approaches.

NEW DIRECTIONS FOR RESEARCH

As mentioned earlier, SL mechanisms can be explored on the dimensions of the abstractness of the manipulated representations (i.e., whether it reflects abstract rule-learning or concrete/distributional learning) and attention (i.e., the question of implicit versus explicit SL). For these two approaches, ERPs, allowing the assessment of "online" cognition, could make a nice contribution if new paradigms are applied to control for the amount of concrete information available in the input and the level of attention (or consciousness) brought to bear. In this regard, the control of concrete information could be performed using the so-called "balanced chunk strength design" (e.g.,

Knowlton and Squire, 1996). This procedure allows one to control for the amount of potential chunks or fragments that can emerge from the stimuli, independent of whether or not the stimuli conform to grammatical rules. As concerns the level of attention, further insights about the underlying implicit and explicit mechanisms of SL could be explored with ERPs using, for instance, the process-dissociation procedure (Jacoby, 1991). This method seems particularly promising when combined with the balanced chunk strength design and ERP, as questions such as whether there exist chunks that participants are not aware of could be tested. Furthermore, it is important to attempt to tease apart the encoding of input (during a "training" phase) versus the expression of knowledge (during a "test" phase) as the level of attention may differentially impact each process (Hendricks et al., 2013). Such a line of research could be used to test the 2-step theory of Perruchet and Pacton (2006), who posited that chunks are unconsciously extracted via a bottom-up process, and then become consciously available, in a second step, for top-down processing.

We mentioned earlier that almost all ERP effects observed in SL paradigms can be interpreted either as indices of the SL process

itself or as a consequence of SL, which modulates the level of attention to the learned material (be it the full stimuli, or fragments of it, i.e., chunks). Future approaches could control the level of attention using for instance subliminal stimulation (Daltrozzo et al., 2011). Finding ERP effects of SL under subliminal stimulation would rule out the alternative attentional explanation, indicating that these ERP effects are indices of SL *per se* and not indirect effects of increased attention to newly learned materials.

Another area where ERPs can be fruitfully used is to explore the nature of multisensory SL and the ways in which different subsystems of SL interact and integrate information across domains. For instance, Walk and Conway (2011) have recently proposed that SL of multisensory patterns proceeds initially via modality-specific mechanisms, and then only at a later stage of information processing, are cross-modal contingencies learned. This type of two-stage theory, in which an earlier process is posited to be followed by a later one, is perfectly amenable for exploration by ERPs, which provide a precise temporal profile of information processing. For instance, ERPs could be used to measure within-modal versus cross-modal violations in an SL paradigm, with the prediction following from Walk and Conway (2011) that cross-modal processing will occur at a later latency than within-modal processing.

In addition, future ERP research could focus more on examining the developmental time-course of SL. This issue can be assessed either on a short time scale, with for instance the analysis of the development of SL across trials within a single experiment; or on a longer time scale, with groups of participants of various ages. Both approaches have been followed for instance by Jost et al. (2011). Indeed, the short time scale approach is particularly well-suited for ERP research because it provides an online assessment of cognitive processing. At a longer time scale, the ERP technique presents also some advantages as compared to other techniques, such as behavioral measures. Whereas behavioral data, which can be rather messy to collect from children, might show a particular developmental pattern, ERP data, which can be elicited even without a behavioral response, might show an entirely different pattern of results. For example, in Jost et al. (2011), two groups of children of different ages and one group of young adults participated in an SL task while ERPs were recorded. Despite the behavioral data showing SL only in the adults group, the ERPs indicated SL also in the children.

Importantly, developmental approaches should not be restricted to comparisons between age groups. What is also needed to better explore SL at a longer time-scale are longitudinal studies (as previously suggested by Conway et al., 2011 and Arciuli and von Koss Torkildsen, 2012). The use of longitudinal studies, for instance, would help provide evidence for a causal relationship between SL and language performance. The demonstration of causality, by showing that SL at a young age predicts language outcomes later in life, would in turn have important implications for clinical intervention. So far, recent research has found a strong link at the neural level between SL and language performance using correlational research strategies (Christiansen et al., 2012; Tabullo et al., 2013). But this type of research design only allows one to conclude that there exists an association between SL and language performance, not necessarily a causal relationship.

In this manner, one potentially important way that ERPs can be used is to assess to what extent SL is amenable to cognitive or behavioral intervention. Because it has been argued and empirically demonstrated that SL is related to language performance (Conway et al., 2011; Daltrozzo et al., 2013), incorporating novel training techniques in an attempt to improve SL could have a causal impact on (i.e., transfer to) language ability (Daltrozzo et al., 2013). In this vein, using ERPs to monitor changes in SL and language abilities after receiving SL training is an important next step. Such an intervention might be even more efficient if it is combined with a biofeedback procedure. Research indicates that the combination of ERP monitoring and biofeedback shows impressive results in terms of neuronal plasticity (e.g., Miltner et al., 1986; Rosenfeld, 1990; Kotchoubey et al., 2000; Birbaumer et al., 2006).

We also suggest that additional research ought to attempt to tackle more realistic learning situations. For example, some models of SL have incorporated the interaction with the speech community and other social cues. Goldstein et al. (2010) and Tomasello (2000) have proposed models that include a bottom-up analysis of statistical regularities reinforced by a top-down attentional mechanism driven by social context cues. The influence of the social environment on SL could be accounted for by an associative memory component, or a retrospective mechanism, which facilitates processing of the stimulus (McClelland, 1979; Dale et al., 2012). According to Dale et al., SL is explained by both a predictive mechanism, as modeled by simple recurrent networks (Cleeremans and McClelland, 1991; Misyak et al., 2009), and a retrospective mechanism, which facilitates subsequent processing in a top-down manner (see also Conway et al., 2010). More research is needed to tease apart the potential role of such top-down processing in more realistic social and linguistic situations, and how this impacts SL.

Finally, it is essential that future research also recognizes the need for exploring several dimensions of SL together, because by only assessing one dimension alone, we may suffer from an overly simplistic and perhaps inaccurate view of the underlying mechanisms of SL. For instance, it might be that different aspects of SL such as the level of abstractness of the encoded representations and the level of consciousness of the learned patterns may develop along different developmental trajectories (although our proposed model predicts that these two aspects develop in parallel, **Figure 2**). As indicated earlier, ERPs are particularly well-suited to explore each of these dimensions and could also be used to explore combinatory modulations of each of these dimensions.

CONCLUSION: AN INTEGRATIVE MODEL

SL mechanisms can be described along several partially-overlapping dimensions: the level of abstractness of the encoded sequential information, the level of attention/consciousness of these representations and the mechanisms that manipulate them, as well as the developmental trajectory. Based on these descriptors, several cognitive and computational models have been proposed. Although many disagreements and unanswered questions remain about these views, a general picture emerges. As an integrative model, we propose that SL is most likely governed by at least two

types of systems whose respective contributions vary across the life-span (**Figure 2**).

In many regards, the results of the ERP research are in line with a two-systems view of SL, as opposed to just one system. However, ERP findings appear to provide inconsistent evidence with regard to the relative involvement of concrete versus abstract rule-learning components. This could be merely due to the greater sensitivity of ERPs to one or the other process and therefore the extent that ERPs are a reliable index of different mechanisms of SL. On the other hand, this might not be an intrinsic weakness of ERP but instead may point to methodological weaknesses in the assessment of consciousness, attention, and intention. To overcome these limitations, several methodological improvements could be used in conjunction with ERP research, including the process-dissociation procedure (Jacoby, 1991) or dual-task methodology (Hendricks et al., 2013), with the aim to test the two-systems hypothesis along the dimensions of consciousness, attention, and intention. Furthermore, more nuanced ways of investigating the level of abstractness of the information encoded through SL could rely upon balanced-chunk strength designs (Knowlton and Squire, 1996).

In sum, this review has explored to what extent ERP findings can be used to better understand the neurocognitive mechanisms of SL. Rather than continuing to argue over a simple dichotomy of abstract versus concrete feature encoding or implicit versus explicit mechanisms, future research must be more aware of the potential complex relationships among multiple neurocognitive mechanisms that may differ along one or more of these dimensions based on the task at hand. Furthermore, ERPs can be used to shed light on the developmental progression of these various mechanisms. If the two-system view of SL (**Figure 2**) is correct, then this helps frame our understanding of the nature of many related aspects of cognition including motor skill development, perceptual processing, and language acquisition. One potential outcome of an improved understanding of the mechanisms of SL is the ability to design novel language rehabilitation interventions, capitalizing on the assumption that improving performance on SL could have a transfer effect and thereby improve the performance of other cognitive processes, such as language, that stem from it.

ACKNOWLEDGMENTS

Preparation of this manuscript was supported by the National Institutes of Health (NIH R01DC012037). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

- Aberg, K. C., and Herzog, M. H. (2012). About similar characteristics of visual perceptual learning and LTP. *Vision Res.* 61, 100–106. doi: 10.1016/j.visres.2011.12.013
- Abla, D., Katahira, K., and Okanoya, K. (2008). On-line assessment of statistical learning by event-related potentials. *J. Cogn. Neurosci.* 20, 952–964. Erratum in: *J. Cogn. Neurosci.* 21, 1 p preceding 1653. doi: 10.1162/jocn.2008.20058
- Acqualagna, L., Treder, M. S., Schreuder, M., and Blankertz, B. (2010). A novel brain-computer interface based on the rapid serial visual presentation paradigm. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2010, 2686–2689. doi: 10.1109/IEMBS.2010.5626548
- Aizenstein, H. J., Butters, M. A., Figurski, J. L., Stenger, V. A., Reynolds, C. F. III, and Carter, C. S. (2005). Prefrontal and striatal activation during sequence learning in geriatric depression. *Biol. Psychiatry* 58, 290–296. doi: 10.1016/j.biopsych.2005.04.023
- Alain, C., Snyder, J. S., He, Y., and Reinke, K. S. (2007). Changes in auditory cortex parallel rapid perceptual learning. *Cereb. Cortex* 17, 1074–1084. doi: 10.1093/cercor/bhl018
- Arciuli, J., and Simpson, I. C. (2011). Statistical learning in typically developing children: the role of age and speed of stimulus presentation. *Dev. Sci.* 14, 464–473. doi: 10.1111/j.1467-7687.2009.00937.x
- Arciuli, J., and von Koss Torkildsen, J. V. (2012). Advancing our understanding of the link between statistical learning and language acquisition: the need for longitudinal data. *Front Psychol.* 3:324. doi: 10.3389/fpsyg.2012.00324
- Arnott, S. R., and Allan, C. (2002). Stepping out of the spotlight: MMN attenuation as a function of distance from the attended location. *Neuroreport* 13, 2209–2212. doi: 10.1097/00001756-200212030-00009
- Aslin, R. N., Saffran, J. R., and Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychol. Sci.* 9, 321–324. doi: 10.1111/1467-9280.00063
- Baddeley, A., and Wilson, B. A. (1994). When implicit learning fails: amnesia and the problem of error elimination. *Neuropsychologia* 32, 53–68. doi: 10.1016/0028-3932(94)90068-X
- Bahlmann, J., Gunter, T. C., and Friederici, A. D. (2006). Hierarchical and linear sequence processing: an electrophysiological exploration of two different grammar types. *J. Cogn. Neurosci.* 18, 1829–1842. doi: 10.1162/jocn.2006.18.11.1829
- Baldwin, K. B., and Kutas, M. (1997). An ERP analysis of implicit structured sequence learning. *Psychophysiology* 34, 74–86. doi: 10.1111/j.1469-8986.1997.tb02418.x
- Bapi, R. S., Chandrasekhar Pammi, V. S., Miyapuram, K. P., and Ahmed, A. (2005). Investigation of sequence processing: a cognitive and computational neuroscience perspective. *Curr. Sci.* 89, 1690–1698.
- Barry, E. (2007). Does conceptual implicit memory develop? The role of processing demands. *J. Genet. Psychol.* 168, 19–36. doi: 10.3200/GNTP.168.1.19-36
- Barry, R. J., Johnstone, S. J., and Clarke, A. R. (2003). A review of electrophysiology in attention-deficit/hyperactivity disorder: II. Event-related potentials. *Clin. Neurophysiol.* 114, 184–198. doi: 10.1016/S1388-2457(02)00363-2
- Batterink, L., and Neville, H. J. (2013). The human brain processes syntax in the absence of conscious awareness. *J. Neurosci.* 33, 8528–8533. doi: 10.1523/JNEUROSCI.0618-13.2013
- Benington, J. Y., and Polich, J. (1999). Comparison of P300 from passive and active tasks for auditory and visual stimuli. *Int. J. Psychophysiol.* 34, 171–177. doi: 10.1016/S0167-8760(99)00070-7
- Berry, D. C., and Dienes, Z. (1993). *Implicit Learning: Theoretical and Empirical Issues*. Hillsdale, NJ: Erlbaum.
- Berti, S., Grunwald, M., and Schröger, E. (2013). Age dependent changes of distractibility and reorienting of attention revisited: an event-related potential study. *Brain Res.* 1491, 156–166. doi: 10.1016/j.brainres.2012.11.009
- Birbaumer, N., Weber, C., Neuper, C., Buch, E., Haapen, K., and Cohen, L. (2006). Physiological regulation of thinking: brain-computer interface (BCI) research. *Prog. Brain Res.* 159, 369–391. doi: 10.1016/S0079-6123(06)59024-7
- Bischoff-Grethe, A., Martin, M., Mao, H., and Berns, G. S. (2001). The context of uncertainty modulates the subcortical response to predictability. *J. Cogn. Neurosci.* 13, 986–993. doi: 10.1162/089892901753165881
- Braver, T. S., Barch, D. M., Keys, B. A., Carter, C. S., Cohen, J. D., Kaye, J. A., et al. (2001). Context processing in older adults: evidence for a theory relating cognitive control to neurobiology in healthy aging. *J. Exp. Psychol. Gen.* 130, 746–763. doi: 10.1037/0096-3445.130.4.746
- Buiatti, M., Peña, M., and Haene-Lambertz, G. (2009). Investigating the neural correlates of continuous speech computation with frequency-tagged neuroelectric responses. *Neuroimage* 44, 509–519. doi: 10.1016/j.neuroimage.2008.09.015
- Byers, A., and Serences, J. T. (2012). Exploring the relationship between perceptual learning and top-down attentional control. *Vision Res.* 74, 30–39. doi: 10.1016/j.visres.2012.07.008
- Cameron-Faulkner, T., Lieven, E., and Tomasello, M. (2003). A construction based analysis of child directed speech. *Cogn. Sci.* 27, 843–873. doi: 10.1207/s15516709cog2706_2

- Carral, V., Corral, M. J., and Escera, C. (2005). Auditory event-related potentials as a function of abstract change magnitude. *Neuroreport* 16, 301–305. doi: 10.1097/00001756-200502280-00020
- Carrión, R. E., and Bly, B. M. (2007). Event-related potential markers of expectation violation in an artificial grammar learning task. *Neuroreport* 18, 191–195. doi: 10.1097/WNR.0b013e328011b8ae
- Cheng, C. H., Hsu, W. Y., and Lin, Y. Y. (2013). Effects of physiological aging on mismatch negativity: a meta-analysis. *Int. J. Psychophysiol.* 90, 165–171. doi: 10.1016/j.ijpsycho.2013.06.026
- Cherry, K. E., and Stadler, M. A. (1995). Implicit learning of a non-verbal sequence in younger and older adults. *Psychol. Aging* 10, 379–394. doi: 10.1037/0882-7974.10.3.379
- Christiansen, M. H., Allen, J., and Seidenberg, M. S. (1998). Learning to segment speech using multiple cues: a connectionist model. *Lang. Cogn. Process.* 13, 221–268. doi: 10.1080/016909698386528
- Christiansen, M. H., Conway, C. M., and Onnis, L. (2012). Similar neural correlates for language and sequential learning: evidence from event-related brain potentials. *Lang. Cogn. Process.* 27, 231–256. doi: 10.1080/01690965.2011.606666
- Christiansen, M. H., Dale, R. A., Ellefson, M. R., and Conway, C. M. (2002). “The role of sequential learning in language evolution: computational and experimental studies,” in *Simulating the Evolution of Language*, eds A. Cangelosi and D. Parisi (London: Springer), 165–187.
- Clayards, M., Tanenhaus, M. K., Aslin, R. N., and Jacobs, R. A. (2008). Perception of speech reflects optimal use of probabilistic speech cues. *Cognition* 108, 804–809. doi: 10.1016/j.cognition.2008.04.004
- Cleeremans, A. (2006). “Conscious and unconscious cognition: a graded, dynamic perspective,” in *Progress in Psychological Science Around the World. I. Neural, Cognitive, and Developmental Issues*, eds Q. Jing, M. R. Rosenzweig, G. d’Ydewalle, H. Zhang, H.-C. Chen, and K. Zhang (Hove: Psychology Press), 401–418.
- Cleeremans, A., Destrebecqz, A., and Boyer, M. (1998). Implicit learning: news from the front. *Trends Cogn. Sci.* 2, 406–416. doi: 10.1016/S1364-6613(98)01232-7
- Cleeremans, A., and McClelland, J. L. (1991). Learning the structure of event sequences. *J. Exp. Psychol. Gen.* 120, 235–253. doi: 10.1037/0096-3445.120.3.235
- Clegg, B. A., DiGirolamo, G. J., and Keele, S. W. (1998). Sequence learning. *Trends Cogn. Sci.* 2, 275–281. doi: 10.1016/S1364-6613(98)01202-9
- Clohessy, A. B., Posner, M. I., and Rothbart, M. K. (2001). Development of the functional visual field. *Acta Psychol. (Amst)* 106, 51–68. doi: 10.1016/S0001-6918(00)00026-3
- Conway, C. M. (2012). “Sequential learning,” in *Encyclopedia of the Sciences of Learning*, ed. R. M. Seel (New York, NY: Springer Publications), 3047–3050.
- Conway, C. M., Bauernschmidt, A., Huang, S. S., and Pisoni, D. B. (2010). Implicit statistical learning in language processing: word predictability is the key. *Cognition* 114, 356–371. doi: 10.1016/j.cognition.2009.10.009
- Conway, C. M., and Christiansen, M. H. (2001). Sequential learning in non-human primates. *Trends Cogn. Sci.* 5, 539–546. doi: 10.1016/S1364-6613(00)01800-3
- Conway, C. M., and Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 24–39. doi: 10.1037/0278-7393.31.1.24
- Conway, C. M., Ellefson, M. R., and Christiansen, M. H. (2003). “When less is less and when less is more: starting small with staged input,” in *Proceedings of the 25th Annual Conference of the Cognitive Science Society* (Mahwah, NJ: Lawrence Erlbaum), 810–815.
- Conway, C. M., Goldstone, R. L., and Christiansen, M. H. (2007). “Spatial constraints on visual statistical learning of multi-element scenes,” in *Proceedings of the 29th Annual Meeting of the Cognitive Science Society* (Mahwah, NJ: Lawrence Erlbaum), 185–190.
- Conway, C. M., and Pisoni, D. B. (2008). Neurocognitive basis of implicit learning of sequential structure and its relation to language processing. *Ann. N. Y. Acad. Sci.* 1145, 113–131. doi: 10.1196/annals.1416.009
- Conway, C. M., Pisoni, D. B., Anaya, E. M., Karpicke, J., and Henning, S. C. (2011). Implicit sequence learning in deaf children with cochlear implants. *Dev. Sci.* 14, 69–82. doi: 10.1111/j.1467-7687.2010.00960.x
- Cooper, R. J., Todd, J., McGill, K., and Michie, P. T. (2006). Auditory sensory memory and the aging brain: a mismatch negativity study. *Neurobiol. Aging* 27, 752–762. doi: 10.1016/j.neurobiolaging.2005.03.012
- Correa, A., Lupiáñez, J., Madrid, E., and Tudela, P. (2006). Temporal attention enhances early visual processing: a review and new evidence from event-related potentials. *Brain Res.* 1076, 116–128. doi: 10.1016/j.brainres.2005.11.074
- Cunillera, T., Càmarà, E., Toro, J. M., Marco-Pallarès, J., Sebastián-Gallès, N., Ortiz, H., et al. (2009). Time course and functional neuroanatomy of speech segmentation in adults. *Neuroimage* 48, 541–553. doi: 10.1016/j.neuroimage.2009.06.069
- Cunillera, T., Toro, J. M., Sebastián-Gallès, N., and Rodríguez-Fornells, A. (2006). The effects of stress and statistical cues on continuous speech segmentation: an event-related brain potential study. *Brain Res.* 23, 168–178. doi: 10.1016/j.brainres.2006.09.046
- Curran, T. (1997). Effects of aging on implicit sequence learning: accounting for sequence structure and explicit knowledge. *Psychol. Res.* 60, 24–41. doi: 10.1007/BF00419678
- Curran, T. (1998). “Implicit sequence learning from a cognitive neuroscience perspective: what, how, and where?” in *Handbook of Implicit Learning*, eds M. A. Stadler and P. Frensch (Thousand Oaks, CA: Sage Publications, Inc.), 365–400.
- Curran, T., and Keele, S. W. (1993). Attentional and non-attentional forms of sequence learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 19, 189–202. doi: 10.1037/0278-7393.19.1.189
- Cutler, A. (2008). The abstract representations in speech processing. *Q. J. Exp. Psychol.* 61, 1601–1619. doi: 10.1080/13803390802218542
- Czerwinski, M., Lightfoot, N., and Shiffrin, R. M. (1992). Automatization and training in visual search. *Am. J. Psychol.* 105, 271–315. doi: 10.2307/1423030
- Czigler, I., Csibra, G., and Csontos, A. (1992). Age and inter-stimulus interval effects on event-related potentials to frequent and infrequent auditory stimuli. *Biol. Psychol.* 33, 195–206. doi: 10.1016/0301-0511(92)90031-O
- Dale, R., Duran, N. D., and Morehead, J. R. (2012). Prediction during statistical learning, and implications for the implicit/explicit divide. *Adv. Cogn. Psychol.* 8, 196–209. doi: 10.2478/v10053-008-0115-z
- Daltrozzo, J., Conway, C. M., and Smith, G. N. L. (2013). Rehabilitating language disorders by improving sequential processing: a review. *J. Macro Trends Health Med.* 1, 41–57.
- Daltrozzo, J., Signoret, C., Tillmann, B., and Perrin, F. (2011). Subliminal semantic priming in speech. *PLoS ONE* 6:e20273. doi: 10.1371/journal.pone.0020273
- De Diego Balaguer, R., Toro, J. M., Rodríguez-Fornells, A., and Bachoud-Lévi, A. C. (2007). Different neurophysiological mechanisms underlying word and rule extraction from speech. *PLoS ONE* 2:e1175. doi: 10.1371/journal.pone.0001175
- Dell’Acqua, R., Jolicoeur, P., Pesciarelli, F., Job, C. R., and Palomba, D. (2003). Electrophysiological evidence of visual encoding deficits in a cross-modal attentional blink paradigm. *Psychophysiology* 40, 629–639. doi: 10.1111/1469-8986.00064
- Dennis, N. A., Howard, J. H., and Howard, D. V. (2003). Age deficits in learning sequences of spoken words. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 58, P224–P227. doi: 10.1093/geronb/58.4.P224
- Deouell, L. Y., Bentin, S., and Giard, M. H. (1998). Mismatch negativity in dichotic listening: evidence for interhemispheric differences and multiple generators. *Psychophysiology* 35, 355–365. doi: 10.1111/1469-8986.3540355
- Destrebecqz, A., and Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychon. Bull. Rev.* 8, 343–350. doi: 10.3758/BF03196171
- Dienes, Z., Broadbent, D. E., and Berry, D. C. (1991). Implicit and explicit knowledge bases in artificial grammar learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 17, 875–887. doi: 10.1037/0278-7393.17.5.875
- Eimer, M., Goschke, T., Schlaghecken, F., and Stürmer, B. (1996). Explicit and implicit learning of event sequences: evidence from event-related brain potentials. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 970–987. Erratum in: *J. Exp. Psychol. Learn. Mem. Cogn.* 23, 279. doi: 10.1037/0278-7393.22.4.970
- Elman, J. L. (1990). Finding structure in time. *Cogn. Sci.* 14, 179–211. doi: 10.1207/s15516709cog1402_1
- Elman, J. L. (1993). Learning and development in neural networks: the importance of starting small. *Cognition* 48, 71–99. doi: 10.1016/0010-0277(93)90058-4
- Endress, A., and Bonatti, L. (2007). Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition* 105, 247–299. doi: 10.1016/j.cognition.2006.09.010
- Fabiani, M., and Friedman, D. (1995). Changes in brain activity patterns in aging: the novelty oddball. *Psychophysiology* 32, 579–594. doi: 10.1111/j.1469-8986.1995.tb01234.x
- Fabiani, M., Friedman, D., and Cheng, J. C. (1998). Individual differences in P3 scalp distribution in older adults, and their relationship to frontal lobe function. *Psychophysiology* 35, 698–708. doi: 10.1111/1469-8986.3560698

- Feeney, J. J., Howard, J. H. Jr., and Howard, D. V. (2002). Implicit learning of higher order sequences in middle age. *Psychol. Aging* 17, 351–355. doi: 10.1037/0882-7974.17.2.351
- Ferdinand, N. K., Runger, D., Frensch, P. A., and Mecklinger, A. (2010). Event-related potential correlates of declarative and non-declarative sequence knowledge. *Neuropsychologia* 48, 2665–2674. doi: 10.1016/j.neuropsychologia.2010.05.013
- Fiser, J., and Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychol. Sci.* 12, 499–504. doi: 10.1111/1467-9280.00392
- Fiser, J., and Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proc. Natl. Acad. Sci. U.S.A.* 99, 15822–15826. doi: 10.1073/pnas.232472899
- Forkstam, C., Hagoort, P., Fernandez, G., Ingvar, M., and Petersson, K. M. (2006). Neural correlates of artificial syntactic structure classification. *Neuroimage* 32, 956–967. doi: 10.1016/j.neuroimage.2006.03.057
- Franco, A., Cleeremans, A., and Destrebecqz, A. (2011). Statistical learning of two artificial languages presented successively: how conscious? *Front. Psychol.* 2:229. doi: 10.3389/fpsyg.2011.00229
- Franco, A., and Destrebecqz, A. (2012). Chunking or not chunking? How do we find words in artificial language learning? *Adv. Cogn. Psychol.* 8, 144–154. doi: 10.5709/acp-0111-3
- Frensch, P. A., Lin, J., and Buchner, A. (1998). Learning versus behavioral expression of the learned: the effects of a secondary tone-counting task on implicit learning in the serial reaction task. *Psychol. Res.* 61, 83–98. doi: 10.1007/s004260050015
- Frensch, P. A., and Miner, C. S. (1994). Effects of presentation rate and individual differences in short-term memory capacity on an indirect measure of serial learning. *Mem. Cogn.* 22, 95–110. doi: 10.3758/BF03202765
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends Cogn. Sci.* 6, 78–84. doi: 10.1016/S1364-6613(00)01839-8
- Friederici, A. D., Mueller, J., and Oberecker, R. (2011). Precursors to natural grammar learning: preliminary evidence from 4-month-old infants. *PLoS ONE* 6:e17920. doi: 10.1371/journal.pone.0017920
- Friederici, A. D., Steinhauer, K., and Pfeifer, E. (2002). Brain signatures of artificial language processing: evidence challenging the critical period hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 99, 529–534. doi: 10.1073/pnas.012611199
- Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., and Donchin, E. (1993). A neural system for error detection and compensation. *Psychol. Sci.* 4, 385–390. doi: 10.1111/j.1467-9280.1993.tb00586.x
- Gervain, J., and Mehler, J. (2010). Speech perception and language acquisition in the first year of life. *Annu Rev. Psychol.* 61, 191–218. doi: 10.1146/annurev.psych.093008.100408
- Goldstein, M. H., Waterfall, H. R., Lotem, A., Halpern, J. Y., Schwade, J. A., Onnis, L., et al. (2010). General cognitive principles for learning structure in time and space. *Trends Cogn. Sci.* 14, 249–258. doi: 10.1016/j.tics.2010.02.004
- Goldstone, R. L. (1998). Perceptual learning. *Annu. Rev. Psychol.* 49, 585–612. doi: 10.1146/annurev.psych.49.1.585
- Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., and Danks, D. (2004). A theory of causal learning in children: causal maps and Bayes nets. *Psychol. Rev.* 111, 3–32. doi: 10.1037/0033-295X.111.1.3
- Gordon, N. (2000). The acquisition of a second language. *Eur. J. Paediatr. Neurol.* 4, 3–7. doi: 10.1053/ejpn.1999.0253
- Goschke, T. (1998). “Implicit learning of perceptual and motor sequences: evidence for independent systems,” in *Handbook of Implicit Learning*, eds M. A. Stadler and P. Frensch (Thousand Oaks, CA: Sage Publications, Inc.), 401–444.
- Gottselig, J. M., Brandeis, D., Hofer-Tinguely, G., Borbely, A. A., and Achermann, P. (2004). Human central auditory plasticity associated with tone sequence learning. *Learn. Mem.* 11, 162–171. doi: 10.1101/lm.63304
- Grafton, S. T., Hazeltine, E., and Ivry, R. (1995). Functional mapping of sequence learning in normal humans. *J. Cogn. Neurosci.* 7, 497–510. doi: 10.1162/jocn.1995.7.4.497
- Hackley, S. A., and Valle-Inclan, F. (2003). Which stages of processing are speeded by a warning signal? *Biol. Psychol.* 64, 27–45. doi: 10.1016/S0301-0511(03)00101-7
- Haider, H., and Frensch, P. A. (2009). Conflicts between expected and actually performed behavior lead to verbal report of incidentally acquired sequential knowledge. *Psychol. Res.* 73, 817–834. doi: 10.1007/s00426-008-0199-6
- Haith, M. M., Hazan, C., and Goodman, G. S. (1988). Expectation and anticipation of dynamic visual events by 3.5-month-old babies. *Child Dev.* 59, 467–479. doi: 10.2307/1130325
- Haith, M. M., and McCarty, M. E. (1990). Stability of visual expectations at 3.0 months of age. *Dev. Psychol.* 26, 68–74. doi: 10.1037/0012-1649.26.1.68
- Hannula, D. E., and Ranganath, C. (2009). The eyes have it: hippocampal activity predicts expression of memory in eye movements. *Neuron* 63, 592–599. doi: 10.1016/j.neuron.2009.08.025
- Hazan, V., and Barrett, S. (2000). The development of phonemic categorization in children aged 6–12. *J. Phon.* 28, 377–396. doi: 10.1006/jpho.2000.0121
- Hendricks, M. A., Conway, C. M., and Kellogg, R. T. (2013). Using dual-task methodology to dissociate automatic from non-automatic processes involved in artificial grammar learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 39, 1491–1500. doi: 10.1037/a0032974
- Hess, T. M. (2005). Memory and aging in context. *Psychol. Bull.* 131, 383–406. doi: 10.1037/0033-2909.131.3.383
- Hoer, M., and Dominey, P. F. (2000). ERP analysis of cognitive sequencing: a left anterior negativity related to structural transformation processing. *Neuroreport* 11, 3187–3191. doi: 10.1097/00001756-200009280-00028
- Hommel, B., Musseler, J., Aschersleben, G., and Prinz, W. (2001). The theory of event coding (TEC): a framework for perception and action planning. *Behav. Brain Sci.* 24, 849–937. doi: 10.1017/S0140525X01000103
- Honda, M., Deiber, M. P., Ibanez, V., Pascual-Leone, A., Zhuang, P., and Hallett, M. (1998). Dynamic cortical involvement in implicit and explicit motor sequence learning: a PET study. *Brain* 121, 2159–2173. doi: 10.1093/brain/121.11.2159
- Howard, D. V., and Howard, J. H. Jr. (1989). Age differences in learning serial patterns: direct versus indirect measures. *Psychol. Aging* 4, 357–364. doi: 10.1037/0882-7974.4.3.357
- Howard, D. V., and Howard, J. H. Jr. (1992). Adult age differences in the rate of learning serial patterns: evidence from direct and indirect tests. *Psychol. Aging* 7, 232–241. doi: 10.1037/0882-7974.7.2.232
- Howard, J. H. Jr., and Howard, D. V. (1997). Age differences in implicit learning of higher order dependencies in serial patterns. *Psychol. Aging* 12, 634–656. doi: 10.1037/0882-7974.12.4.634
- Howard, D. V., Howard, J. H. Jr., Japikse, K., DiYanni, C., Thompson, A., and Somberg, R. (2004). Implicit sequence learning: effects of level of structure, adult age, and extended practice. *Psychol. Aging* 19, 79–92. doi: 10.1037/0882-7974.19.1.79
- Humes, L. E., and Floyd, S. S. (2005). Measures of working memory, sequence learning, and speech recognition in the elderly. *J. Speech Lang. Hear. Res.* 48, 224–235. doi: 10.1044/1092-4388(2005)016
- Huetzel, S. A., Mack, P. B., and McCarthy, G. (2002). Perceiving patterns in random series: dynamic processing of sequence in prefrontal cortex. *Nat. Neurosci.* 5, 485–490. doi: 10.1038/nn841
- Jaaskelainen, I. P., Ahveninen, J., Andermann, M. L., Belliveau, J. W., Raij, T., and Sams, M. (2011). Short-term plasticity as a neural mechanism supporting memory and attentional functions. *Brain Res.* 1422, 66–81. doi: 10.1016/j.brainres.2011.09.031
- Jacoby, L. L. (1991). A process dissociation framework: separating automatic from intentional use of memory. *J. Mem. Lang.* 30, 513–541. doi: 10.1016/0749-596X(91)90025-F
- Jamieson, R. K., and Mewhort, D. J. K. (2009). Applying an exemplar model to the serial reaction-time task: anticipating from experience. *Q. J. Exp. Psychol.* 62, 1757–1783. doi: 10.1080/17470210802557637
- Jimenez, L., Mendez, C., and Cleeremans, A. (1996). Comparing direct and indirect measures of sequence learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 948–969. doi: 10.1037/0278-7393.22.4.948
- Johnson, C., and Wilbrecht, L. (2011). Juvenile mice show greater flexibility in multiple choice reversal learning than adults. *Dev. Cogn. Neurosci.* 1, 540–551. doi: 10.1016/j.dcn.2011.05.008
- Jost, E., Conway, C. M., Purdy, J. D., and Hendricks, M. A. (2011). Neurophysiological correlates of visual statistical learning in adults and children. *Paper Presented at 33rd Annual meeting of the Cognitive Science Society*, Boston, MA.
- Keele, S. W., Ivry, R., Mayr, U., Hazeltine, E., and Heuer, H. (2003). The cognitive and neural architecture of sequence representation. *Psychol. Rev.* 110, 316–339. doi: 10.1037/0033-295X.110.2.316
- Kessler, K., Schmitz, F., Gross, J., Hommel, B., Shapiro, K., and Schnitzler, A. (2005). Target consolidation under high temporal processing demands as revealed by MEG. *Neuroimage* 26, 1030–1041. Erratum in: *Neuroimage* 35, 989–990. doi: 10.1016/j.neuroimage.2005.02.020

- Kim, R., Seitz, A., Feenstra, H., and Shams, L. (2009). Testing assumptions of statistical learning: is it long-term and implicit? *Neurosci. Lett.* 461, 145–149. doi: 10.1016/j.neulet.2009.06.030
- Kirkham, N. Z., Slemmer, J. A., and Johnson, S. P. (2002). Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition* 83, B35–B42. doi: 10.1016/S0010-0277(02)00004-5
- Knowlton, B. J., Ramus, S. J., and Squire, L. R. (1992). Intact artificial grammar learning in amnesia: dissociation of classification learning and explicit memory for specific instances. *Psychol. Sci.* 3, 172–179. doi: 10.1111/j.1467-9280.1992.tb00021.x
- Knowlton, B. J., and Squire, L. R. (1993). The learning of categories: parallel brain systems for item memory and category knowledge. *Science* 262, 1747–1749. doi: 10.1126/science.8259522
- Knowlton, B. J., and Squire, L. R. (1994). The information acquired during artificial grammar learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 79–91. doi: 10.1037/0278-7393.20.1.79
- Knowlton, B. J., and Squire, L. R. (1996). Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 169–181. doi: 10.1037/0278-7393.22.1.169
- Kok, A. (2000). Age-related changes in involuntary and voluntary attention as reflected in components of the event-related potential (ERP). *Biol. Psychol.* 54, 107–143. doi: 10.1016/S0301-0511(00)00054-5
- Kotchoubey, B., Haisst, S., Daum, I., Schugens, M., and Birbaumer, N. (2000). Learning and self-regulation of slow cortical potentials in older adults. *Exp. Aging Res.* 26, 15–35. doi: 10.1080/036107300243669
- Kranczioch, C., Debener, S., Herrmann, C. S., and Engel, A. K. (2006). EEG gamma-band activity in rapid serial visual presentation. *Exp. Brain Res.* 169, 246–254. doi: 10.1007/s00221-005-0139-2
- Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., and Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 979–1000. doi: 10.1098/rstb.2007.2154
- Kuhl, P., and Rivera-Gaxiola, M. (2008). Neural substrates of language acquisition. *Annu. Rev. Neurosci.* 31, 511–534. doi: 10.1146/annurev.neuro.30.051606.094321
- Kuhn, G., and Dienes, Z. (2005). Implicit learning of non-local musical rules: implicitly learning more than chunks. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 1417–1432. doi: 10.1037/0278-7393.31.6.1417
- Kumano, H., and Uka, T. (2013). Neuronal mechanisms of visual perceptual learning. *Behav. Brain Res.* 249, 75–80. doi: 10.1016/j.bbr.2013.04.034
- Kutas, M., and Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu. Rev. Psychol.* 62, 621–647. doi: 10.1146/annurev.psych.093008.131123
- Lang, S., and Kotchoubey, B. (2000). Learning effects on event-related brain potentials. *Neuroreport* 11, 3327–3331. doi: 10.1097/00001756-200010200-00013
- Lashley, K. S. (1951). “The problem of serial order in behavior,” in *Cerebral Mechanisms in Behavior*, ed. L. A. Jeffress (New York: John Wiley & Sons), 112–136.
- Lelekov, T., Dominey, P. F., and Garcia-Larrea, L. (2000). Dissociable ERP profiles for processing rules vs instances in a cognitive sequencing task. *Neuroreport* 11, 1129–1132. doi: 10.1097/00001756-200004070-00043
- Lieberman, M. D., Chang, G. Y., Chiao, J., Bookheimer, S. Y., and Knowlton, B. J. (2004). An event-related fMRI study of artificial grammar learning in a balanced chunk strength design. *J. Cogn. Neurosci.* 16, 427–438. doi: 10.1162/089892904322926764
- Lu, Z. L., Hua, T., Huang, C. B., Zhou, Y., and Doshier, B. A. (2011). Visual perceptual learning. *Neurobiol. Learn. Mem.* 95, 145–151. doi: 10.1016/j.nlm.2010.09.010
- Manza, L., and Reber, A. S. (1997). “Representing artificial grammars: transfer across stimulus forms and modalities,” in *How Implicit is Implicit Learning?* ed. C. Diane (Oxford: Oxford University Press), 73–106.
- Marco-Pallarés, J., Grau, C., and Ruffini, G. (2005). Combined ICA-LORETA analysis of mismatch negativity. *Neuroimage* 25, 471–477. doi: 10.1016/j.neuroimage.2004.11.028
- Marcus, G. F., Vijayan, S., Rao, S. B., and Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science* 283, 77–80. doi: 10.1126/science.283.5398.77
- Mathews, R. C., Buss, R. R., Stanley, W. B., Blanchard-Fields, F., Cho, J. R., and Druhan, B. (1989). Role of implicit and explicit processes in learning from examples: a synergistic effect. *J. Exp. Psychol. Learn. Mem. Cogn.* 15, 1083–1100. doi: 10.1037/0278-7393.15.6.1083
- Maye, J., Werker, J. F., and Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82, B101–B111. doi: 10.1016/S0010-0277(01)00157-3
- McAndrews, M. P., and Moscovitch, M. (1985). Rule-based and exemplar classification in artificial grammar learning. *Mem. Cogn.* 13, 469–475. doi: 10.3758/BF03198460
- McClelland, J. L. (1979). On the time relations of mental processes: an examination of systems of processes in cascade. *Psychol. Rev.* 86, 287–330. doi: 10.1037/0033-295X.86.4.287
- McCloskey, M., and Cohen, N. J. (1989). Catastrophic interference in connectionist networks: the sequential learning problem. *Psychol. Learn. Motiv.* 24, 109–164. doi: 10.1016/S0079-7421(08)60536-8
- McNealy, K., Mazziotta, J., and Dapretto, M. (2010). The neural basis of speech parsing in children and adults. *Dev. Sci.* 13, 385–406. doi: 10.1111/j.1467-7687.2009.00895.x
- Mecklenbräuer, S., Hupbach, A., and Wippich, W. (2003). Age-related improvements in a conceptual implicit memory test. *Mem. Cogn.* 31, 1208–1217. doi: 10.3758/BF03195804
- Meiri, H. (2011). Implicit learning processes of compensated dyslexic and skilled adult readers. *Dev. Neuropsychol.* 36, 939–943. doi: 10.1080/87565641.2011.606419
- Meulemans, T., and Van der Linden, M. (1997). Associative chunk strength in artificial grammar learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 23, 1007–1028. doi: 10.1037/0278-7393.23.4.1007
- Meulemans, T., Van der Linden, M., and Perruchet, P. (1998). Implicit sequence learning in children. *J. Exp. Child Psychol.* 69, 199–221. doi: 10.1006/jecp.1998.2442
- Miltner, W., Larbig, W., and Braun, C. (1986). Biofeedback of visual evoked potentials. *Int. J. Neurosci.* 29, 291–303. doi: 10.3109/00207458608986158
- Misyak, J. B., Christiansen, M. H., and Tomblin, J. B. (2009). “Statistical learning of non-adjacencies predicts on-line processing of long-distance dependencies in natural language,” in *Proceedings of the 31st Annual Meeting of the Cognitive Science Society*, eds N. T. Taatgen and H. van Rijn (Austin, TX: Cognitive Science Society), 177–182.
- Miyawaki, K., Sato, A., Yasuda, A., Kumano, H., and Kuboki, T. (2005). Explicit knowledge and intention to learn in sequence learning: an event-related potential study. *Neuroreport* 16, 705–708. doi: 10.1097/00001756-200505120-00010
- Molfese, D. L. (1990). Auditory evoked responses recorded from 16-month-old human infants to words they did and did not know. *Brain Lang.* 38, 345–363. doi: 10.1016/0093-934X(90)90120-6
- Montague, P. R., and Sejnowski, T. J. (1994). The predictive brain: temporal coincidence and temporal order in synaptic learning mechanisms. *Learn. Mem.* 1, 1–33.
- Morsella, E. (2005). The function of phenomenal states: supramodular interaction theory. *Psychol. Rev.* 112, 1000–1021. doi: 10.1037/0033-295X.112.4.1000
- Mueller, J. L., Bahlmann, J., and Friederici, A. D. (2008). The role of pause cues in language learning: the emergence of event-related potentials related to sequence processing. *J. Cogn. Neurosci.* 20, 892–905. doi: 10.1162/jocn.2008.20511
- Mueller, J. L., Hahne, A., Fujii, Y., and Friederici, A. D. (2005). Native and non-native speakers’ processing of a miniature version of Japanese as revealed by ERPs. *J. Cogn. Neurosci.* 17, 1229–1244. doi: 10.1162/089892905002463
- Müller, B. W., Achenbach, C., Oades, R. O., Bender, S., and Schall, U. (2002). Modulation of mismatch negativity by stimulus deviance and modality of attention. *Neuroreport* 13, 1317–1320. doi: 10.1097/00001756-200207190-00021
- Näätänen, R., Kujala, T., Escera, C., Baldeweg, T., Kreegipuu, K., Carlson, S., et al. (2012). The mismatch negativity (MMN)—a unique window to disturbed central auditory processing in ageing and different clinical conditions. *Clin. Neurophysiol.* 123, 424–458. doi: 10.1016/j.clinph.2011.09.020
- Näätänen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* 118, 2544–2590. doi: 10.1016/j.clinph.2007.04.026
- Newport, E. L. (1990). Maturational constraints on language learning. *Cogn. Sci.* 14, 11–28. doi: 10.1207/s15516709cog1401_2
- Nissen, M. J., and Bullemer, P. (1987). Attentional requirements of learning: evidence from performance measures. *Cogn. Psychol.* 19, 1–32. doi: 10.1016/0010-0285(87)90002-8

- Nittrouer, S. (1996). Discriminability and perceptual weighting of some acoustic cues to speech perception by 3-year-olds. *J. Speech Hear. Res.* 39, 278–297.
- Osterhout, L., and Holcomb, P. J. (1992). Event-related brain potentials by syntactic anomaly. *J. Mem. Lang.* 31, 785–806. doi: 10.1016/0749-596X(92)90039-Z
- Parkin, A. J., and Streete, S. (1988). Implicit and explicit memory in young children and adults. *Br. J. Psychol.* 79, 361–369. doi: 10.1111/j.2044-8295.1988.tb02295.x
- Patel, A. D., Gibson, E., Ratner, J., Besson, M., and Holcomb, P. J. (1998). Processing syntactic relations in language and music: an event-related potential study. *J. Cogn. Neurosci.* 10, 717–733. doi: 10.1162/089892998563121
- Pekkonen, E. (2000). Mismatch negativity in aging and in Alzheimer's and Parkinson's diseases. *Audiol. Neurootol.* 5, 216–224. doi: 10.1159/000013883
- Pekkonen, E., Jousmäki, V., Partanen, J., and Karhu, J. (1993). Mismatch negativity area and age-related auditory memory. *Electroencephalogr. Clin. Neurophysiol.* 87, 321–325. doi: 10.1016/0013-4694(93)90185-X
- Pekkonen, E., Rinne, T., Reinikainen, K., Kujala, T., Alho, K., and Näätänen, R. (1996). Aging effects on auditory processing: an event-related potential study. *Exp. Aging Res.* 22, 171–184. doi: 10.1080/03610739608254005
- Pelucchi, B., Hay, J. F., and Saffran, J. R. (2009). Statistical learning in a natural language by 8-month-old infants. *Child Dev.* 80, 674–685. doi: 10.1111/j.1467-8624.2009.01290.x
- Perruchet, P., and Amorim, M. A. (1992). Conscious knowledge and changes in performance in sequence learning: evidence against dissociation. *J. Exp. Psychol. Learn. Mem. Cogn.* 18, 785–800. doi: 10.1037/0278-7393.18.4.785
- Perruchet, P., Bigand, E., and Benoit-Gonin, F. (1997). The emergence of explicit knowledge during the early phase of learning in sequential reaction time tasks. *Psychol. Res.* 60, 4–13. doi: 10.1007/BF00419676
- Perruchet, P., and Pacteau, C. (1990). Synthetic grammar learning: implicit rule abstraction or explicit fragmentary knowledge. *J. Exp. Psychol. Gen.* 119, 264–275. doi: 10.1037/0096-3445.119.3.264
- Perruchet, P., and Pacton, S. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *Trends Cogn. Sci.* 10, 233–238. doi: 10.1016/j.tics.2006.03.006
- Perruchet, P., Tyler, M., Galland, N., and Peereboom, R. (2004). Learning non-adjacent dependencies: no need for algebraic-like computations. *J. Exp. Psychol. Gen.* 133, 573–583. doi: 10.1037/0096-3445.133.4.573
- Petersson, K. M., Forkstam, C., and Ingvar, M. (2004). Artificial syntactic violations activate Broca's region. *Cogn. Sci.* 28, 383–407. doi: 10.1207/s15516709cog2803_4
- Pierrehumbert, J. B. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Lang. Speech* 46(Pt 2–3), 115–154. doi: 10.1177/00238309030460020501
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148. doi: 10.1016/j.clinph.2007.04.019
- Prull, M. W., Gabrieli, J. D., and Bunge, S. A. (2000). “Age-related changes in memory: a cognitive neuroscience perspective,” in *The Handbook of Aging and Cognition*, 2nd Edn, eds F. I. M. Craik and T. A. Salthouse (Mahwah, NJ: Erlbaum), 91–153.
- Pulvermüller, F., and Assadollahi, R. (2007). Grammar or serial order?: Discrete combinatorial brain mechanisms reflected by the syntactic mismatch negativity. *J. Cogn. Neurosci.* 19, 971–980. doi: 10.1162/jocn.2007.19.6.971
- Reber, A. S. (1967). Implicit learning of artificial grammars. *J. Verbal Learning Verbal Behav.* 6, 855–863. doi: 10.1016/S0022-5371(67)80149-X
- Reber, A. S. (1989). Implicit learning and tacit knowledge. *J. Exp. Psychol. Gen.* 118, 219–235. doi: 10.1037/0096-3445.118.3.219
- Reber, A. S. (1993). *Implicit Learning and Tacit Knowledge: An Essay on the Cognitive Unconscious*. New York: Oxford University Press.
- Reber, P. J., and Squire, L. R. (1998). Encapsulation of implicit and explicit memory in sequence learning. *J. Cogn. Neurosci.* 10, 248–263. doi: 10.1162/0898929985626281
- Reed, J., and Johnson, P. (1994). Assessing implicit learning with indirect tests: determining what is learned about sequence structure. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 585–594. doi: 10.1037/0278-7393.20.3.585
- Rosas, R., Ceric, F., Tenorio, M., Mourgues, C., Thibaut, C., Hurtado, E., et al. (2010). ADHD children outperform normal children in an artificial grammar Implicit learning task: ERP and RT evidence. *Conscious. Cogn.* 19, 341–351. doi: 10.1016/j.concog.2009.09.006
- Rose, M., Verleger, R., and Wascher, E. (2001). ERP correlates of associative learning. *Psychophysiology* 38, 440–450. doi: 10.1111/1469-8986.3830440
- Rosenfeld, J. P. (1990). Applied psychophysiology and biofeedback of event-related potentials (brain waves): historical perspective, review, future directions. *Biofeedback Self Regul.* 15, 99–119. doi: 10.1007/BF00999142
- Rosenthal, C. R., Aimola Davies, A., Maller, J., Johnson, M. R., and Kennard, C. (2010). “Impairment of higher-order but not simple sequence learning in a case of bilateral hippocampal organic amnesia,” in *Poster Session Presented at the Cognitive Neuroscience Society Annual Meeting*, Montreal, QC.
- Rossnagel, C. S. (2001). Revealing hidden covariation detection: evidence for implicit abstraction at study. *J. Exp. Psychol. Learn. Mem. Cogn.* 27, 1276–1288. doi: 10.1037/0278-7393.27.5.1276
- Rüsseler, J., Hennighausen, E., Münte, T. F., and Rösler, F. (2003a). Differences in incidental and intentional learning of sensorimotor sequences as revealed by event-related brain potentials. *Brain Res. Cogn. Brain Res.* 15, 116–126. doi: 10.1016/S0926-6410(02)00145-3
- Rüsseler, J., Kuhlicke, D., and Münte, T. F. (2003b). Human error monitoring during implicit and explicit learning of a sensorimotor sequence. *Neurosci. Res.* 47, 233–240. doi: 10.1016/S0168-0102(03)00212-8
- Rüsseler, J., Hennighausen, E., and Rösler, F. (2001). Response anticipation processes in the learning of a sensorimotor sequence. *J. Psychophysiol.* 15, 95–105. doi: 10.1027//0269-8803.15.2.95
- Rüsseler, J., and Rösler, F. (2000). Implicit and explicit learning of event sequences: evidence for distinct coding of perceptual and motor representations. *Acta Psychol. (Amst)* 104, 45–67. doi: 10.1016/S0001-6918(99)00053-0
- Ruzzoli, M., Pirulli, C., Brignani, D., Maioli, C., and Miniussi, C. (2012). Sensory memory during physiological aging indexed by mismatch negativity (MMN). *Neurobiol. Aging* 33, e621–e630. doi: 10.1016/j.neurobiolaging.2011.03.021
- Saarienen, J., Paavilainen, P., Schöger, E., Tervaniemi, M., and Näätänen, R. (1992). Representation of abstract attributes of auditory stimuli in the human brain. *Neuroreport* 3, 1149–1151. doi: 10.1097/00001756-199212000-00030
- Saffran, J., Senghas, A., and Trueswell, J. C. (2001). The acquisition of language by children. *Proc. Natl. Acad. Sci. U.S.A.* 98, 12874–12875. doi: 10.1073/pnas.231498898
- Saffran, J. R., Aslin, R. N., and Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science* 274, 1926–1928. doi: 10.1126/science.274.5294.1926
- Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., and Barrueco, S. (1997). Incidental language learning: listening (and learning) out of the corner of your ear. *Psychol. Sci.* 8, 101–105. doi: 10.1111/j.1467-9280.1997.tb00690.x
- Sagi, D., and Tanne, D. (1994). Perceptual learning: learning to see. *Curr. Opin. Neurobiol.* 4, 195–199. doi: 10.1016/0959-4388(94)90072-8
- Salthouse, T. A., McGuthry, K. E., and Hambrick, D. Z. (1999). A framework for analyzing and interpreting differential aging patterns: application to three measures of implicit learning. *Aging Neuropsychol. Cogn.* 6, 1–18. doi: 10.1076/ane.6.1.1.789
- Samuel, A. G., and Kraljic, T. (2009). Perceptual learning for speech. *Atten. Percept. Psychophys.* 71, 1207–1218. doi: 10.3758/APP.71.6.1207
- Sanders, L. D., Newport, E. L., and Neville, H. J. (2002). Segmenting non-sense: an event-related potential index of perceived onsets in continuous speech. *Nat. Neurosci.* 5, 700–703. doi: 10.1038/nm873
- Sasaki, Y., Nanez, J. E., and Watanabe, T. (2010). Advances in visual perceptual learning and plasticity. *Nat. Rev. Neurosci.* 11, 53–60. doi: 10.1038/nrn2737
- Schlaghecken, F., Stürmer, B., and Eimer, M. (2000). Chunking processes in the learning of event sequences: electrophysiological indicators. *Mem. Cognit.* 28, 821–831. doi: 10.3758/BF03198417
- Schneider, W., and Pressley, M. (1997). *Memory Development between 2 and 20*, 2nd Edn. Mahwah, NJ: Erlbaum.
- Schröger, E., Bendixen, A., Trujillo-Barreto, N. J., and Roeber, U. (2007). Processing of abstract rule violations in audition. *PLoS ONE* 2:e1131. doi: 10.1371/journal.pone.0001131
- Seeger, C. A., Prabhakaran, V., Poldrack, A., and Gabrieli, J. D. E. (2000). Neural activity between explicit and implicit learning of artificial grammar strings: an fMRI study. *Psychobiology* 3, 283–292.
- Seriès, P., and Seitz, A. R. (2013). Learning what to expect (in visual perception). *Front. Hum. Neurosci.* 7:668. doi: 10.3389/fnhum.2013.00668
- Servan-Schreiber, E., and Anderson, J. R. (1990). Learning artificial grammars with competitive chunking. *J. Exp. Psychol. Learn. Mem. Cogn.* 16, 592–608. doi: 10.1037/0278-7393.16.4.592

- Shafto, C. L., Conway, C. M., Field, S. L., and Houston, D. M. (2012). Visual sequence learning in infancy: domain-general and domain-specific associations with language. *Infancy* 17, 247–271. doi: 10.1111/j.1532-7078.2011.00085.x
- Shanks, D. R., Channon, S., Wilkinson, L., and Curran, H. V. (2006). Disruption of sequential priming in organic and pharmacological amnesia: a role for the medial temporal lobes in implicit contextual learning. *Neuropsychopharmacology* 31, 1768–1776. doi: 10.1038/sj.npp.1300935
- Shanks, D. R., and Johnstone, T. (1999). Evaluating the relationship between explicit and implicit knowledge in a sequential reaction time task. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 1435–1451. doi: 10.1037/0278-7393.25.6.1435
- Shanks, D. R., and Perruchet, P. (2002). Dissociation between priming and recognition in the expression of sequential knowledge. *Psychon. Bull. Rev.* 9, 362–367. doi: 10.3758/BF03196294
- Shea, C. H., Park, J. H., and Braden, H. W. (2006). Age-related effects in sequential motor learning. *Phys. Ther.* 86, 478–488.
- Silva-Pereyra, J., Conboy, B. T., Klarman, L., and Kuhl, P. K. (2007). Grammatical processing without semantics? An event-related brain potential study of preschoolers using jabberwocky sentences. *J. Cogn. Neurosci.* 19, 1050–1065. doi: 10.1162/jocn.2007.19.6.1050
- Skosnik, P. D., Mirza, F., Gitelman, D. R., Parrish, T. B., Mesulam, M.-M., and Reber, P. J. (2002). Neural correlates of artificial grammar learning. *Neuroimage* 17, 1306–1314. doi: 10.1006/nimg.2002.1291
- Skrandies, W., and Fahle, M. (1994). Neurophysiological correlates of perceptual learning in the human brain. *Brain Topogr.* 7, 163–168. doi: 10.1007/BF01186774
- Smith, E. E., and Jonides, J. (1995). “Working memory in humans: neuropsychological evidence,” in *The Cognitive Neurosciences*, ed. M. S. Gazzaniga (Cambridge, MA: MIT Press), 1009–1020.
- Smith, P. H., Loboschewski, T. W., Davidson, B. K., and Dixon, W. E. Jr. (1997). Scripts and checkerboards: the influence of ordered visual information on remembering locations in infancy. *Infant Behav. Dev.* 20, 549–552. doi: 10.1016/S0163-6383(97)90044-8
- Song, S., Howard, J. H., and Howard, D. V. (2007). Implicit probabilistic sequence learning is independent of explicit awareness. *Learn. Mem.* 14, 167–176. doi: 10.1101/lm.437407
- Stadler, M. A. (1995). Role of attention in sequence learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 21, 674–685. doi: 10.1037/0278-7393.21.3.674
- Stadler, W., Klimesch, W., Pouthas, V., and Ragot, R. (2006). Differential effects of the stimulus sequence on CNV and P300. *Brain Res.* 1123, 157–167. doi: 10.1016/j.brainres.2006.09.040
- Steinhauer, K., Friederici, A. D., and Pfeifer, E. (2001). “ERP recordings while listening to syntax errors in an artificial language: evidence from trained and untrained subjects,” in *Poster Presented at the 14th Annual CUNY Conference on Human Sentence Processing*, Philadelphia, PA.
- Tabullo, A., Sevilla, Y., Segura, E., Zanutto, S., and Wainseboim, A. (2013). An ERP study of structural anomalies in native and semantic free artificial grammar: evidence for shared processing mechanisms. *Brain Res.* 1527, 149–160. doi: 10.1016/j.brainres.2013.05.022
- Teinonen, T., Fellmann, V., Näätänen, R., Alku, P., and Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neurosci.* 10:21. doi: 10.1186/1471-2202-10-21
- Thiessen, E. D., and Pavlik, P. I. (2013). iMinerva: a mathematical model of distributional statistical learning. *Cogn. Sci.* 37, 310–343. doi: 10.1111/cogs.12011
- Thomas, K. M., Hunt, R. H., Vizuetta, N., Sommer, T., Durston, S., Yang, Y., et al. (2004). Evidence of developmental differences in implicit sequence learning: an fMRI study of children and adults. *J. Cogn. Neurosci.* 16, 1339–1351. doi: 10.1162/0898929042304688
- Thomas, K. M., and Nelson, C. A. (2001). Serial reaction time learning in preschool- and school-age children. *J. Exp. Child Psychol.* 79, 364–387. doi: 10.1006/jecp.2000.2613
- Tiitinen, H., May, P., Reinikainen, K., and Näätänen, R. (1994). Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature* 372, 90–92. doi: 10.1038/372090a0
- Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition* 74, 209–253. doi: 10.1016/S0010-0277(99)00069-4
- Trippie, R. H., Hewig, J., Heydel, C., Hecht, H., and Miltner, W. H. (2007). Attentional Blink to emotional and threatening pictures in spider phobics: electrophysiology and behavior. *Brain Res.* 1148, 149–160. doi: 10.1016/j.brainres.2007.02.035
- Turk-Browne, N. B., Jungé, J. A., and Scholl, B. J. (2005). Attention and automaticity in visual statistical learning. *Talk Presented at Vision Sciences Society Conference*, Sarasota, FL.
- Turk-Browne, N. B., Scholl, B. J., Chun, M. M., and Johnson, M. K. (2009). Neural evidence of statistical learning: efficient detection of visual regularities without awareness. *J. Cogn. Neurosci.* 21, 1934–1945. doi: 10.1162/jocn.2009.21131
- Uddén, J., and Bahlmann, J. (2012). A rostro-caudal gradient of structured sequence processing in the left inferior frontal gyrus. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 2023–2032. doi: 10.1098/rstb.2012.0009
- van Zuijen, T. L., Simoens, V. L., Paavilainen, P., Näätänen, R., and Tervaniemi, M. (2006). Implicit, intuitive, and explicit knowledge of abstract regularities in a sound sequence: an event-related brain potential study. *J. Cogn. Neurosci.* 18, 1292–1303. doi: 10.1162/jocn.2006.18.8.1292
- Vicari, S., Marotta, L., Menghini, D., Molinari, M., and Petrosini, L. (2003). Implicit learning deficit in children with developmental dyslexia. *Neuropsychologia* 41, 108–114. doi: 10.1016/S0028-3932(02)00082-9
- Vihman, M. M., Thierry, G., Lum, J., Keren-Portnoy, T., and Martin, P. (2007). Onset of word form recognition in English, Welsh, and English-Welsh bilingual infants. *Appl. Psycholinguist.* 28, 475–493. doi: 10.1017/S0142716407070269
- Walk, A. M., and Conway, C. M. (2011). “Multisensory statistical learning: can cross-modal associations be acquired?” in *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, eds L. Carlson, C. Hoelscher, and T. F. Shipley (Austin, TX: Cognitive Science Society), 3337–3342.
- Walter, W. G., Cooper, R., Aldridge, V. J., Mccallum, W. C., and Winter, A. L. (1964). Contingent Negative Variation: an electric sign of sensorimotor association and expectancy in the human brain. *Nature* 203, 380–384. doi: 10.1038/203380a0
- Werker, J. (2012). Perceptual foundations of bilingual acquisition in infancy. *Ann. N.Y. Acad. Sci.* 1251, 50–61. doi: 10.1111/j.1749-6632.2012.06484.x
- Willingham, D. B., Greeley, T., and Bardone, A. M. (1993). Dissociation in a serial response time task using a recognition measure: Comment on Perruchet and Amorim (1992). *J. Exp. Psychol. Learn. Mem. Cogn.* 19, 1424–1430. doi: 10.1037/0278-7393.19.6.1424
- Yu, K., Shen, K., Shao, S., Ng, W. C., Kwok, K., and Li, X. (2011). Common spatio-temporal pattern for single-trial detection of event-related potential in rapid serial visual presentation triage. *IEEE Trans. Biomed. Eng.* 58, 2513–2520. doi: 10.1109/TBME.2011.2158542
- Zachau, S., Rinker, T., Körner, B., Kohls, G., Maas, V., Hennighausen, K., et al. (2005). Extracting rules: early and late mismatch negativity to tone patterns. *Neuroreport* 16, 2015–2019. doi: 10.1097/00001756-200512190-00009

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 27 February 2014; accepted: 30 May 2014; published online: 18 June 2014.
Citation: Daltrozzo J and Conway CM (2014) Neurocognitive mechanisms of statistical-sequential learning: what do event-related potentials tell us? *Front. Hum. Neurosci.* 8:437. doi: 10.3389/fnhum.2014.00437

This article was submitted to the journal *Frontiers in Human Neuroscience*.
Copyright © 2014 Daltrozzo and Conway. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.