Exploring the Neural Mechanisms Supporting Structured Sequence Processing and Language Using Event-Related Potentials: Some Preliminary Findings

Gretchen N.L. Smith (gsmith50@student.gsu.edu)
Department of Psychology, Georgia State University, P.O. Box 5010
Atlanta, GA 30302-5010 USA

Gerardo E. Valdez (gvaldez2@student.gsu.edu),
Department of Psychology, Georgia State University, P.O. Box 5010
Atlanta, GA 30302-5010 USA

Anne M. Walk (amcclur3@illinois.edu)
Department of Kinesiology and Community Health, University of Illinois, 405 North Mathews Avenue
Urbana, Illinois 61801 USA

John D. Purdy (jdprdy@gmail.com)
Department of Psychology, Saint Louis University, 3700 Lindell Blvd.
St. Louis, MO 63108

Christopher M. Conway (cconway@gsu.edu)
Department of Psychology, Georgia State University, P.O. Box 5010
Atlanta, GA 30302-5010 USA

Abstract

Structured sequence processing (SSP) refers to the neurocognitive mechanisms used to learn sequential patterns in the environment. SSP ability seems to be important for language (Conway, Bauernschmidt, Huang, & Pisoni, 2010); however, there are few neural studies showing an empirical connection between SSP and language. The purpose of this study was to investigate the association between SSP and language processing by comparing the underlying neural components elicited during each type of task. Healthy adult subjects completed a visual, non-linguistic SSP task incorporating an artificial grammar and a visual morpho-syntactic language task. Both tasks were designed to cause violations in expectations of items occurring in a series. Event-related potentials (ERPs) were used to examine the underlying neural mechanisms associated with these expectancy violations. The results indicated the P3a component elicited by the SSP task and the P600 component elicited by the language task shared similarities in their topographic distribution. These preliminary analyses suggest that the P3a and P600 may reflect processes involving detection of sequential violations in non-language and language domains, which is consistent with the idea that language processing relies on general-purpose SSP mechanisms.

Keywords: Structured Sequence Processing; Sequence Learning; Statistical Learning; Artificial Grammar Learning; Language Processing; Syntax; Event Related Potentials; P3a; P600

Introduction

Structured sequence processing (SSP), also termed sequential learning or statistical learning, is a core cognitive mechanism used to learn patterns of information from the environment over time. SSP emerges early in development (Aslin, Saffran, & Newport, 1998) and is largely automatic and implicit (Cleeremans, Destrebecqz, & Boyer, 1998), though explicit processes likely occur in parallel (Sun, Slusarz, & Terry, 2005). SSP involves learning complex embedded patterns in which each item that occurs next is determined probabilistically based on what item occurred previously (see Conway & Christiansen, 2001 for a more detailed discussion on different types of sequential learning).

A key facet of SSP is that it serves as a tool for making predictions about which elements will occur next in a sequence (Christiansen, Conway, & Onnis, 2012). When sequential patterns are learned, this information can be used not only to generate expectancies about upcoming stimuli in the sequence, but also to detect when stimuli deviate from expectation (Ferdinand, Mecklinger, & Kray, 2008). Bar (2007) suggests that a “circular mechanism” occurs in which the brain limits processing of stimuli that are predictable, while allotting cognitive resources to stimuli that are novel and/or unexpected. These “predictive processing” operations are generally beneficial to many aspects of cognition, including perception, movement, decision-making (Bubic, von Cramon, & Schubotz, 2010) and language (Federmeier, 2007).

SSP appears to be especially important in the domain of language (Conway, Bauernschmidt, Huang, & Pisoni, 2010). In particular, SSP may support knowledge and use of grammatical language, such as word order (Conway et al., 2010), phonology (Saffran, 2003), morphology and syntax

(Ullman, 2004). However, the association between SSP and language has largely been assumed. Only recently have behavioral studies demonstrated empirical links between SSP and natural language processing (e.g., Conway et al., 2010). Additionally, even fewer studies have empirically compared these two mechanisms at a neural level (e.g., Patel, Gibson, Ratner, Besson, & Holcomb, 1998). In a neural-based investigation using a within-subject design, Christiansen et al. (2012) examined the electrophysiological responses elicited during a visual SSP task and a visual syntactic natural language processing task. The findings indicated both the SSP task and the natural language task elicited a late positive-going deflection in voltage potential—a P600 component—that has been linked with the processing of syntactic violations (Lelekov, Dominey, & Garcia-Larrea, 2000). Furthermore, topographic maps of the P600 effects showed similar distribution between conditions (Christiansen et al., 2012). Overall, these results provided some of the earliest direct, within-subject empirical evidence that the same neural mechanisms may be used for SSP and syntactic natural language processing (Christiansen et al., 2012).

However, more direct neural evidence of a link between SSP and natural language processing is needed, using different types of SSP and language tasks. Therefore, the purpose of this study was to investigate the relation between SSP and natural language processing by comparing the electrophysiological profiles elicited during each type of task, using a within-subject design. The SSP task was designed to resemble an artificial grammar learning (AGL) paradigm (Reber, 1967), in which complex statistical regularities are embedded in the sequences. One key aspect of the SSP task used in this study is that it is more purely non-linguistic in nature than previous SSP tasks using language-like stimuli (e.g., Christiansen et al., 2012) or stimuli that are readily verbalizable and easily mapped onto vocalizations (e.g., Patel et al., 1998). Evidence showing that similar neural responses are elicited during a more fundamentally non-linguistic SSP task and a natural language task would provide additional—and possibly more compelling—support that language processing is based in part on mechanisms utilized to extract and encode structured sequential information in a general-purpose manner.

It is possible our visual non-linguistic SSP task and our morpho-syntactic visual natural language processing task would both elicit a P600, similar to the Christiansen et al., (2012) findings. The authors of that study hypothesized that the P600 might reflect processing broadly involved with making predictions about upcoming items in a series, which is not confined solely to language (Christiansen et al., 2012). It is also possible our SSP task might elicit ERP components that have been associated with extraction and encoding of non-linguistic structured sequential patterns. Previous studies have suggested that the N200 [negative-going deflection, occurring approximately 200 milliseconds (ms) after stimulus onset] and P3b (positive-going deflection occurring approximately 300 ms after stimulus onset) components are elicited in sequential learning paradigms and may reflect the processing of expectancy violations (e.g, Carrión & Bly, 2007). Additionally, the P3a (positive-going deflection occurring approximately 250 ms after stimulus onset) has been evoked from "novel" stimulus paradigms (Courchesne, Hillyard, & Galambos, 1975), has been linked with the recognition of grammatical violations in a second language (Jakoby, Goldstein, & Faust, 2011), and has been associated with focused attention (Comerchero & Polich, 1999).

Given the present study was exploratory, we expected to observe any of the ERP components mentioned above. Consequently, the central hypothesis was simply that violations in a non-linguistic SSP task and violations in a morpho-syntactic natural language task would elicit similar electrophysiological response profiles.

Method

Subjects

Forty-three subjects (ages 18-22; 25 female) participated. All subjects were recruited from Saint Louis University, were native speakers of English, with normal to corrected-to-normal vision and who, at time of testing, reported no history of hearing loss, difficulty with speech, or history of cognitive, perceptual, or motor disorder.

Experimental Paradigm

Measures of SSP and language were administered separately in a single test session. All subjects performed the measure of SSP first and the measure of language second.

Measure of SSP The measure of SSP was similar to the "Simon" visual-spatial SSP task used in previously published work (see Conway et al., 2010 for details). In this measure, subjects viewed sequences of 4 black squares appearing one at a time on a white background in 1 of 4 possible quadrants (upper left, upper right, lower left, lower right) (See Figure 1 below). The task was to reproduce each sequence immediately following presentation by touching the squares in the correct order on a touchscreen. Unknown to subjects, the measure of SSP consisted of two parts: a learning phase and a test phase, which differed in the types of sequences presented.

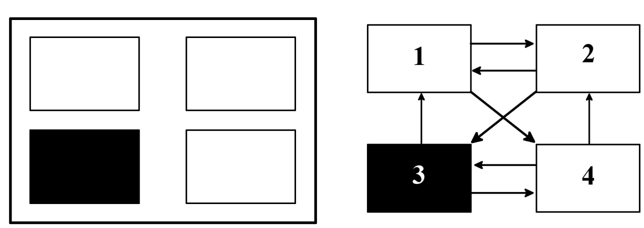

Figure 1: SSP Task (on left) with rule structure (on right).

In the learning phase, sequence elements were generated according to an underlying artificial grammar that specified the probability of a particular element in a sequence occurring given the preceding element (see Figure 1 above). For each sequence, the starting element (1-4) was randomly determined, and then the grammar was used to determine each subsequent element, until the full sequence length was reached. For example, given the starting element 3, the element 2 had a zero probability of occurring next, while the 1 and 4 elements had an equal (50%) chance of occurring. No element could follow itself in the sequence. The mapping of the rules to the locations was randomly determined for each subject; however, for each subject, the mapping remained consistent across all trials. All sequences were 5 elements long.

The learning phase consisted of 40 grammatical sequences. The phase began with a blank screen that appeared for 1 second. In the first part of the learning phase (20 sequences), each element in the sequence was displayed for 600 ms, followed by a 200 ms pause in which nothing was displayed on the screen. The final element in the sequence was followed by a 200 ms pause before the whole 2x2 grid of squares was displayed with a "Done" button in the middle. Using the touch screen, the subject then reproduced the sequence just presented, followed by pressing the word "Done". Immediate feedback was given as to the correctness of each response. The second part of the learning phase (20 sequences) was the same as the first part, except each element in the sequence was displayed for 400ms, followed by a 200 ms pause.

The test phase consisted of 64 sequences. One fourth of the sequences were "grammatical-trained" (i.e., the sequences were identical to grammatical sequences presented in the learning phase), one fourth were "grammatical-untrained" (i.e, new sequences sharing the same underlying structure as the other grammatical sequences), and one half were "ungrammatical" (i.e., the sequences violated the grammar). For the ungrammatical sequences, the starting element (1-4) was randomly determined, then any element could occur next in the sequence except repeating elements were not allowed. The timing of the test phase was identical to the timing used in in the second half of the learning phase. The subjects were not told that this was a test phase or that there were different types of sequences (grammatical-trained, grammatical-untrained, and ungrammatical). From the perspective of the subject, the test phase was the same reproduction task they had been doing all along.

Scoring for the Measure of SSP For the test phase, a sequence was scored as correct if a subject correctly reproduced it. A score of 5 was given for each correctly reproduced sequence and was based on the length of each sequence. As in previous studies (e.g., Conway et al., 2010) a learning score was then obtained by subtracting the total score for the ungrammatical sequences in the test phase from the grammatical sequences in the test phase. A higher learning score indicates better performance on structured sequences compared to ones that violate the structure, suggesting that successful SSP occurred.

Measure of Language Sixty-four sentences were presented in the measure of language. These sentences varied according to whether or not they contained grammar violations. Thirty-two sentences were grammatically correct, and 32 sentences contained morpho-syntactic violations pertaining to verb agreement with the subject. For example, "The famous singer walks onto the stage." was a grammatically correct sentence used in the task and "The famous singer walk onto the stage." was a sentence with a violation. Each sentence began with a white fixation point (+) and was presented 1 word at a time in white text on a black screen. Each element (word or fixation point) appeared for 400 ms and was followed by a blank screen for 400 ms. Thirty-two 7-word sentences and 32 8-word sentences were presented. The 32 8-word sentences each contained an auxiliary verb. The target words (grammatical or violation) always occurred at the 4th word in the 7-word sentences and at the 5th work in the 8-word sentences. Target words were always verbs. The ratio of grammatically correct sentences to sentences with morpho-syntactic violations was 1:1, for each sentence length.

The phrase "Was that a good or a bad sentence? Press 1 for good, press 2 for bad." appeared on the screen immediately following presentation of the final word in the sentence. The task was to make a keypad response on button 1 if the sentence was "good" and to respond on button 2 if the sentence was "bad." Subjects were not given explicit instruction as to what "good" or "bad" meant (i.e., that grammatical sentences were "good" and sentences with violations were "bad), nor were they told that some sentences were grammatical and some had violations. No feedback was given. A 1-second pause was given between a response and the presentation of the next sentence.

Scoring for the Measure of Language A sentence was scored correct if the subject made the correct grammaticality judgment for that sentence (i.e., a button press on "1"/ "good" for grammatical sentences; a button press on "2"/ "bad" for sentences with morpho-syntactic violations).

Expectancy Violations for Both the Measure of SSP and the Measure of Language

Both measures were designed to cause violations in expectations of items occurring in a series (i.e., a violation of the learned sequence in the SL task and a violation of grammar in the language task). Event-related potentials (ERPs)—portions of ongoing electroencephalogram (EEG) time-locked to cognitive events of interest—were used to associated with these expectancy violations.

EEG/ERP Data Acquisition and Preprocessing

EEG was recorded during the test phase of the measure of SSP and throughout the measure of language using a 128-channel high-density sensor net with vertex recording reference (Electrical Geodesics, Eugene OR). Standard sensor net application techniques were followed. Recordings were made using NetStation acquisition software (Electrical Geodesics, Inc.), with a 0.1–100Hz bandpass filter and digitized at 250 Hz. Electrode impedances were kept below 50 kiloohms. Rest breaks were given as needed.

ERP for the measure of SSP was time-locked to the presentation of a stimulus that violated the artificial grammar and was compared to a stimulus in a similar position in a sequence that was grammatical. ERP for the measure of language was time-locked to the presentation of a word in the sentence that violated the morpho-syntactic grammar and was compared to a word in a similar position in a sentence that was grammatical.

Data was preprocessed using Netstation (Electrical Geodesics, Inc.). The continuous raw EEG recording was filtered through a 0.1 Hz high pass filter and a 30 Hz low pass filter. Channels were marked bad for a given trial if blinks or eye movements were detected, if amplitudes >150 μV, if the channel was flat (had zero variance), or if manual inspection suggested noise specific to that channel. Channels marked bad were interpolated in the raw EEG from data measured at nearby electrodes. After exclusion of artifacts, the continuous EEG was segmented into epochs in the interval -200 msec to +1000 msec with respect to the onset of the target stimulus (i.e., violation of grammar, for both the measure of SSP and the measure of language). Data were not re-referenced from the vertex channel.

Data from 3 subjects was excluded from analysis due to bad EEG channels that were either too high in number or too clustered together. Data from 8 subjects was excluded due to poor data quality or missing data. Therefore, data from a total of 32 subjects was analyzed.

Regions of Interest

Nine regions of interest (ROI) were defined for data analysis, with each containing 9 channels: frontal (FRz), central (CNz), posterior (POz), left anterior (LAn), left central (LCn), left posterior (LPo), right anterior (RAn), right central (RCn), and right posterior (RPo) (see Figure 2 below).

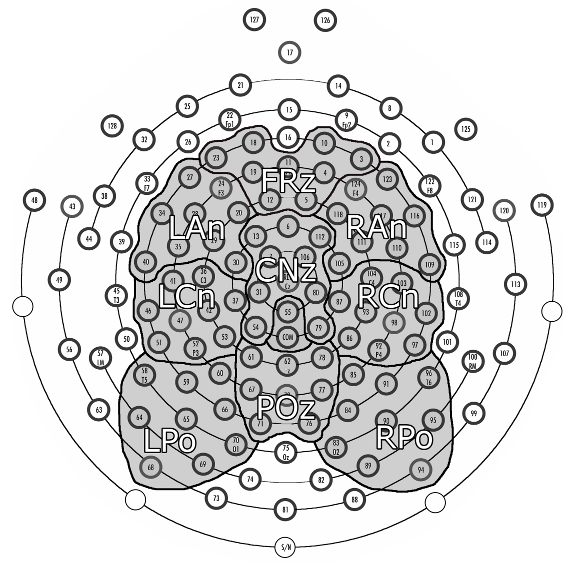

Figure 2: 2-D layout of the 128-channel sensor net (top is front). For data analysis, the channels were grouped into 9 regions of interest (outlined above), each consisting of 9 channels.

Results

Behavioral Average Task Performance for the Measure of SSP and the Measure of Language

Average accuracy given in percentage correct for reproduction of the three sequence types presented in the measure of SSP was as follows: 73% (grammatical-trained), 71% (grammatical-untrained), and 70% (ungrammatical). These three scores were not significantly different from one another (p=.999). Average accuracy given in percentage correct for the grammaticality judgment of the sentences presented in the measure of language was 93%.

Electrophysiological Response Elicited by the Measure of SSP

Visual inspection indicated a P3a component for ungrammatical sequences relative to both types of grammatical sequences in several ROI. Paired samples t-tests were conducted on the grand-averaged mean amplitude waveforms associated with the P3a component 270-330ms after the sequence violations, with significant effects in the central (CNz) [$t(31)=3.968$, $p<.001$], frontal (Frz) [$t(31)=3.321$, $p=.002$], and right anterior (RAn) [$t(31)=2.303$, $p=.028$] regions [See Figure 3 below for an example of the P3a effect in the frontal (FRz) region].

Correlations were computed between the grand-averaged mean amplitude waveforms associated with the significant P3a effects for ungrammatical sequences relative to both types of grammatical sequences and the averaged learning score on the measure of SSP (grammatical-ungrammatical). Results showed that learning score was significantly negatively correlated with the P3a effect in the frontal (FRz) region [$r(31)= -.300$, $p=.05$].

Electrophysiological Response Elicited by the Measure of Language

Visual inspection indicated a P600 component for ungrammatical sentences relative to grammatical sentence in several ROI. Paired samples t-tests were conducted on the mean amplitude waveforms associated with the P600

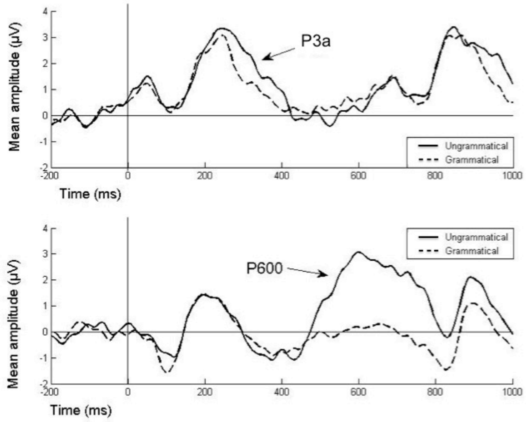

Figure 3: P3a component in the frontal region and P600 component in the central region.

component 525-925ms after the syntactic violations, with significant effects in the central (CNz) [$t(31)=6.616$, $p<.001$], right posterior (RPo) [$t(31)=2.880$, $p=.007$], left posterior (LPo) [$t(31)=5.360$, $p<.001$], right anterior (RAn) [$t(31)=7.610$, $p<.001$], left anterior (LAn) [$t(31)=2.902$, $p=.007$], and left central (LCn) [$t(31)=3.433$, $p=.002$] regions [See Figure 3 above for an example of the P600 effect in the central (CNz) region].

Comparison Between the Electrophysiological Responses Underlying SSP and Language

Topographical maps were created from ungrammatical minus grammatical difference waves associated with the P3a and P600 described above, to compare the scalp distribution of electrical activity for both components in the two tasks (See Figure 4 below). Visual inspection showed some similarities in the topographic profile for the two components; specifically, the early phase of the P600 (525-580ms) resembles the full P3a component. A correlation was computed between the difference waves (ungrammatical-grammatical) between the early phase of the P600 and the full P3a component. The results showed a positive correlation in the right anterior (RAn) region between the P3a (270-330ms) and the early P600 (525-580ms) that approached significance [$r(32)=.291$, $p=.106$].

Discussion

The present study provided some initial analyses comparing the electrophysiological responses elicited in a visual, non-linguistic SSP task and a visual morpho-syntactic natural language processing task. Following exposure to sequences that followed an embedded artificial grammar, subjects showed a P3a to violations of the grammar, while showing a P600 to morpho-syntactic violations in a visual natural language processing task. The P600 elicited in the language task is consistent with findings from paradigms that involve the processing of syntactic violations. The P3a elicited in

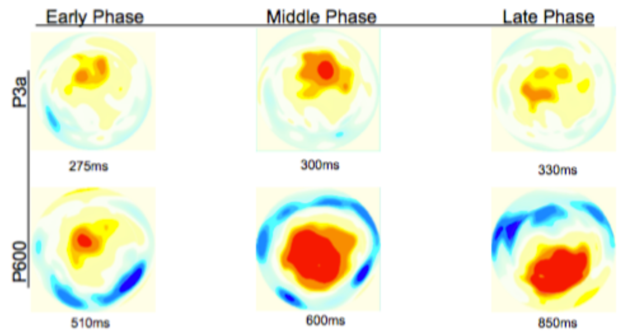

Figure 4: Topographic maps showing distribution of the P3a and P600 components, measured at the scalp.

the SSP task is consistent with previous AGL ERP studies showing a P300 component in response to violations of a grammar (e.g., Opitz, Ferdinand, & Mecklinger, 2011). The presence of the P3a could be due to violations of the artificial grammar drawing attentional resources, corresponding with Bar's (2007) "circular mechanism" of predictive processing previously described. The negative correlation between the SSP learning score—an indication of better performance on grammatical sequences—and the P3a effect for ungrammatical sequences relative to grammatical sequences supports this notion.

Although averaged behavior performance on the measure of SSP showed no learning effect as a group, the correlation between the SSP learning score and the ERP effect for the ungrammatical sequences suggests that even though an overall group learning effect was not observed, there is a distribution of learning scores that seem meaningful, with some individuals showing learning and others not showing learning. Future work will investigate whether individual differences in cognitive processes such as attention and working modulate learning of structured sequences.

One interpretation of the topomaps findings is that the P3a and P600 are distinct components, yet both reflect processes involving the detection of sequential violations in artificial and natural grammar processing tasks, respectively. It has been previously suggested that P300 and P600 components may both reflect the processing of incongruent information in different types of tasks (e.g., Christiansen et al., 2012). A similar role for both types of components, therefore, suggests some degree of overlap between SSP and language processing mechanisms.

The early phase of the P600 showed a distribution similar to the full P3a elicited in the SSP task, suggesting that they may have a common neural origin. Although the distribution of the full P3a is more confined than the full P600, this could be due to the relative unfamiliarity and limited exposure to the SSP task, whereas the wider scalp distribution of the P600 might be the result of the extensive and prolonged exposure humans have had with language. With more exposure to non-linguistic SSP tasks, the P3a elicited to violations in the artificial grammar might show robustness similar to the P600 elicited to violations in natural language grammar. On the other hand, the differences in the ERP correlates observed for the two tasks could instead be due to differences in the types of violations inherent to each (something akin to word order violations in the AGL task but subject-verb agreement violations in the natural language task).

Coulson, King, and Kutas (1998) noted similarities between the P600 and P3b elicited by manipulations to probability and saliency in an oddball paradigm. They concluded the P600 might be a part of the P300 family of components (Coulson et al., 1998). Still, similarity in distribution may not reflect similarity in neural generators [see Osterhout & Hagoort, 1999 (response to Coulson et al., 1998) for a more detailed discussion on cautions of attempting to determine similarity of neural mechanisms from EEG recorded at the scalp]. To help address this limitation, we are using source localization analyses to examine whether the components elicited in our two tasks share a common neural origin. Although preliminary, these findings suggest the possibility that the neurocognitive mechanisms involved in detecting sequential violations in a non-linguistic AGL task are similar to those involved in detecting morpho-syntactic violations in natural language, with the P3a possibly being an earlier version of the P600 or a reduced variant of it.

Acknowledgements

N. Buchholz, J. Daltrozzo, J. Deocampo, G. Frishkoff, E. Hilvert, S. Pardasani, B. Rickles, K. Ross, G. Signiski, S. Singh, and R. Town. Funding: NIH R01DC012037.

References

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *9*, 321 324.

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, *11*, 280 289.

Bubic, A., von Cramon, D. Y. & Schubotz, R. I. (2010) Prediction, cognition and the brain. *Frontiers in Human Neuroscience*, *4*, 1-15.

Carrión, R.E. & Bly, B.M. (2007). Event-related potential markers of expectation violation in an artificial grammar learning task. *NeuroReport, 18,* 191-195.

Christiansen, M.H., Conway, C.M., & Onnis, L. (2012). Similar neural correlates for language and sequential learning: Evidence from event-related brain potentials. *Language and Cognitive Processes*, *27*, 231-256.

Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences*, *2*, 406 416.

Comerchero, M.D. & Polich, J. (1999). P3a and P3b from typical auditory and visual stimuli. *Clinical Neurophysiology 110*(1), 24–30.

Conway, C.M., Bauernschmidt, A., Huang, S.S., & Pisoni, D.B. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition*, *114*, 356-371.

Conway, C.M., & Christiansen, M.H. (2001). Sequential learning in non-human primates. *Trends in Cognitive Sciences*, *5*, 529-546.

Coulson, S., King, J.W., Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes*, *13*(1), 21-58.

Courchesne, E., Hillyard, S.A., Galambos, R. (1975). Stimulus novelty, task relevance and the visual evoked potential in man. *Electroencephalography and Clinical Neurophysiology*, *39*(2), 131-43.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, *44*, 491-505.

Ferdinand, N. K., Mecklinger, A., & Kray, J. (2008). Error deviance processing in implicit and explicit sequence learning. *Journal of Cognitive Neuroscience*, *20*, 629 642.

Jakoby, H., Goldstein, A., & Faust, M. (2011). Electrophysiological correlates of speech perception mechanisms and individual differences in second language attainment. *Psychophysiology*, *48*(11), 1517-1531.

Lelekov, T., Dominey, P. F., & Garcia-Larrea, L. (2000). Dissociable ERP profiles for processing rules vs instances in a cognitive sequencing task. *NeuroReport*, *11*(5), 1129-1132.

Opitz, B., Ferdinand, N. K., & Mecklinger, A. (2011). Timing matters: the impact of immediate and delayed feedback on artificial language learning. *Frontiers in Human Neuroscience, 5*(8), 1-9.

Osterhout, L., & Hagoort, P. (1999). A superficial resemblance does not necessarily mean you are part of the family: Counterarguments to Coulson, King, and Kutas (1998) in the P600/SPS-P300 debate. *Language and Cognitive Processes*, *14*, 1-14.

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, *10*, 717-733.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Behavior*, *6*, 855-863.

Saffran, J. R. (2003). Statistical learning: Mechanisms and constraints. *Current Directions in Psychological Science*, *12*(4), 110-114.

Sun, R., Slusarz, P., & Terry, C. (2005). The interaction of the explicit and the implicit in skill learning: A dual process approach. *Psychological Review*, *112*(1), 159-192.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*, 231-270.