

Functional differences between statistical learning with and without explicit training

Laura J. Batterink, Paul J. Reber, and Ken A. Paller

Department of Psychology, Northwestern University, Evanston, Illinois 60208, USA

Humans are capable of rapidly extracting regularities from environmental input, a process known as statistical learning. This type of learning typically occurs automatically, through passive exposure to environmental input. The presumed function of statistical learning is to optimize processing, allowing the brain to more accurately predict and prepare for incoming input. In this study, we ask whether the function of statistical learning may be enhanced through supplementary explicit training, in which underlying regularities are explicitly taught rather than simply abstracted through exposure. Learners were randomly assigned either to an explicit group or an implicit group. All learners were exposed to a continuous stream of repeating nonsense words. Prior to this implicit training, learners in the explicit group received supplementary explicit training on the nonsense words. Statistical learning was assessed through a speeded reaction-time (RT) task, which measured the extent to which learners used acquired statistical knowledge to optimize online processing. Both RTs and brain potentials revealed significant differences in online processing as a function of training condition. RTs showed a crossover interaction; responses in the explicit group were faster to predictable targets and marginally slower to less predictable targets relative to responses in the implicit group. P300 potentials to predictable targets were larger in the explicit group than in the implicit group, suggesting greater recruitment of controlled, effortful processes. Taken together, these results suggest that information abstracted through passive exposure during statistical learning may be processed more automatically and with less effort than information that is acquired explicitly.

Our environment is governed by structure. Objects in the world are not organized randomly, but most often appear in predictable locations (e.g., toasters are typically found in kitchens but not bathrooms). Sounds in the environment such as music and bird-song consist of repeating motifs (Kalcounis-Rueppel et al. 2006; Lipkind and Tchernichovski 2011). Human speech is also structured, with certain sounds co-occurring more frequently than others (such as *sl-* versus *tl-* in English). The general process of extracting this type of structure from the environment is referred to as *statistical learning*.

Statistical learning is thought to play an important role in language learning (e.g., Thompson and Newport 2007; Yu 2008), particularly in speech segmentation (Saffran et al. 1996a,b, 1997). Natural speech consists of a stream of sound with no reliable pauses between words, and one of the challenges facing language learners is to discover word boundaries. Listeners may accomplish this task by extracting statistical relationships between syllables, as syllables that occur within words have higher transitional probabilities than syllables occurring across words. In the first demonstration of statistical learning, human infants were exposed to a continuous stream of repeating three-syllable nonsense words (e.g., *babupudutabapidabu...*). Following exposure, infants showed sensitivity to the difference between the three-syllable sequences and foil sequences made up of the same syllables recombined in a different order, demonstrating that they were able to use the statistics of the input stream to discover word boundaries in connected speech (Saffran et al. 1996a). Subsequent studies indicated that older children and adults also have this ability, becoming capable of discriminating between nonsense words and foil sequences after relatively short periods of exposure to input (Saffran et al. 1996b, 1997). Although statistical learning was initially implicated in language acquisition, it is

a domain-general mechanism, also operating across nonlinguistic stimuli. For example, tracking the relationship between objects and locations helps perceivers to parse complex visual scenes (Fiser and Aslin 2001, 2002), and exposure to tones following a probabilistic pattern enables listeners to recognize novel tone sequences following the same structure (Durrant et al. 2011, 2013). Overall, the function of statistical learning is to enable people to more accurately predict and prepare for incoming input, facilitating the processing of complex stimuli.

Statistical learning typically occurs automatically and incidentally, in the absence of explicit instruction or conscious attempts to extract the pattern. Statistical learning has been demonstrated both when stimuli are presented passively without any explicit task (e.g., Saffran et al. 1999; Fiser and Aslin 2001, 2002; Toro et al. 2005), and when participants are engaged in a cover task unrelated to the underlying pattern (Saffran et al. 1997, Turk-Browne et al. 2005, 2009). In addition, statistical learning seems to be unaffected by the precise instructions given to participants (Arcuili et al. 2014; Batterink et al. 2015). Thus, statistical learning appears to emerge as an obligatory consequence of exposure to input, without requiring explicit guidance to learn. Nonetheless, these findings do not rule out the possibility that the same information normally abstracted through passive statistical learning may be acquired through alternative routes. This idea is consistent with evidence from learning in other domains. For example, abstract visual categories can be acquired either intentionally or incidentally, with differential patterns of neural activation observed under different learning modes; the hippocampus and prefrontal cortex, among other regions, are recruited

© 2015 Batterink et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first 12 months after the full-issue publication date (see <http://learnmem.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-Non-Commercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

Corresponding author: lbatterink@northwestern.edu

Article is online at <http://www.learnmem.org/cgi/doi/10.1101/lm.037986.114>.

as a consequence of intentional learning, whereas the occipital cortex is recruited under incidental learning conditions (Reber et al. 2003). Similarly, probabilistic cue-outcome relations can be acquired through tasks emphasizing either declarative or non-declarative memory (Poldrack et al. 2001). Engagement of the medial temporal lobe (MTL) and striatum is modulated by whether the task encourages the use of declarative or nondeclarative strategies, with the MTL preferentially implicated under declarative memory conditions and the striatum more strongly involved under nondeclarative conditions. Thus, it is possible that the same or similar statistical representations may be either abstracted through passive exposure or acquired through more explicit training methods, with these two learning routes mediated by dissociable neural substrates.

In addition to evaluating whether it is possible to enhance statistical learning, examining potential explicit influences on statistical learning may also yield insight into which learning systems mediate statistical learning as it typically proceeds. Demonstrations that explicit knowledge improves performance on tasks requiring statistical knowledge would support the notion that statistical learning can potentially draw upon both implicit and explicit memory representations. This evidence would also suggest a parallel between statistical learning and other types of learning just described, such as categorical learning. In contrast, if explicit instruction does not enhance statistical learning functioning, such evidence would show that statistical learning depends upon representations or processes that operate independently of explicit memory (cf. Sanchez and Reber 2013).

A priori, the possibility of enhancing the function of statistical learning through supplementary explicit training, resulting in the optimization of online processing, appears to be a tenable one. First, there is evidence that although statistical learning occurs incidentally or implicitly, it produces knowledge that is at least partially explicit in nature, at least in adult learners (Franco et al. 2011; Bertels et al. 2012, 2013; Batterink et al. 2015). Whereas implicit knowledge, such as abstract grammar knowledge (Reber 1967) or perceptual-motor skills (Cohen and Squire 1980) must be acquired over time through direct experience, one feature of explicit knowledge is that it can be provided to learners through explicit instruction (i.e., by verbally transmitting information from one person to another). Thus, in principle, it should be possible to explicitly train learners on these representations. A second point in favor of this possibility is that the level of the explicit knowledge produced by statistical learning is typically rather impoverished. Some studies find that performance on explicit recognition measures does not reliably exceed chance (Sanders et al. 2002; McNealy et al. 2006; Turk-Browne et al. 2008), and the upper range of average performance rarely exceeds 70%–75% accuracy (e.g., Saffran et al. 1996b, 1997, 1999). Even when overall group-level recognition accuracy is relatively high, explicit knowledge often varies considerably between individual participants. For example, an event-related potential (ERP) study of statistical learning found that overall recognition performance was 74% correct, with the top third of participants performing at 90% accuracy and the bottom third of participants achieving only 59% accuracy (Abla et al. 2008). Thus, there is considerable potential for explicit training methods to enhance learners' explicit knowledge beyond the level normally achieved through passive statistical learning. Whether or not the enhancement of this explicit knowledge will ultimately alter the function of statistical learning remains to be determined.

A separate though related question is whether statistical learning mechanisms continue to operate when explicit knowledge of the underlying structure is acquired. It is possible that explicit training could prevent learners from relying upon statistical cues to extract the underlying units, because they have already

learned these units explicitly. Although this study does not directly test this idea, previous literature suggests that statistical learning mechanisms should operate regardless of whether learners have access to explicit knowledge of the underlying stimulus structure. First, statistical learning is generally thought to be an automatic and obligatory process (e.g., Saffran et al. 1997, 1999; Fiser and Aslin 2001, 2002; Turk-Browne et al. 2005), implying that concomitant explicit knowledge should not interfere with these mechanisms. More broadly, it has also been shown that implicit and explicit learning can occur in parallel. For example, a number of SRT studies have shown that implicit learning of underlying sequences occurs to the same extent whether or not learners have explicit knowledge of the pattern (Willingham and Goedert-Eschmann 1999; Willingham et al. 2002; Song et al. 2007; Sanchez and Reber 2013). In addition, neural activation representing implicit learning occurs in a common neural network regardless of whether subjects are aware of the sequence during performance, with additional regions activated when subjects are aware of the sequence (Willingham et al. 2002). The same phenomenon has been also demonstrated using contextual cueing, a completely different type of implicit learning paradigm that refers to the facilitated ability to locate a visual target in a scene because of prior exposure to the scene (Westerberg et al. 2011). Participants showed reductions in neural activity to repeat relative to novel stimuli in regions involved in visual perception and attention, regardless of whether they were informed about scene repetition. Explicitly instructing participants about the presence of repeating scenes caused additional activation of the medial temporal lobe without interfering with this neural signature of implicit learning. Taken together, these results indicate that statistical learning is likely to occur irrespective of explicit knowledge and in parallel to explicit learning. However, it is important to note that the conclusions we draw in this study, which concern differences between knowledge acquired explicitly and knowledge acquired passively through statistical learning, hold regardless of whether this is the case.

The goal of this study was to determine whether the function of statistical learning may be enhanced through supplementary explicit training, leading to improved performance on tasks where statistical knowledge is called upon. We used an auditory statistical learning paradigm, which involves exposing participants to a continuous stream of nonsense words (e.g., Saffran et al. 1996b, 1997). Learners were randomly assigned to an explicit condition in which they were explicitly instructed on the nonsense words prior to exposure to the speech stream, or to an implicit condition, involving only exposure to the stream without any explicit training component (Fig. 1). The function of statistical learning was then assessed using an online, speeded, performance-based measure. Known as the target-detection task, this task was originally used in visual statistical learning studies (Turk-Browne et al. 2005; Kim et al. 2009) and more recently has been adapted to assess statistical learning in the auditory domain (Batterink et al. 2015; Franco et al. 2015). The task requires participants to respond to target syllables occurring in a continuous syllable stream and provides an indirect measure of statistical learning (Fig. 1). The target-detection task assesses the extent to which learners use their acquired statistical knowledge to optimize online processing, with faster reaction times (RTs) indexing more efficient processing of predictable targets (e.g., Turk-Browne et al. 2005; Kim et al. 2009; Batterink et al. 2015; Franco et al. 2015).

ERPs were also recorded, providing an index of facilitation at the neural level. In particular, we focused on P300, a positive-going ERP component with a typical latency of ~250–500 msec post-stimulus that is elicited during stimulus discrimination (Polich 2007). Early studies using the two-stimulus oddball task have demonstrated that discriminating a target stimulus from

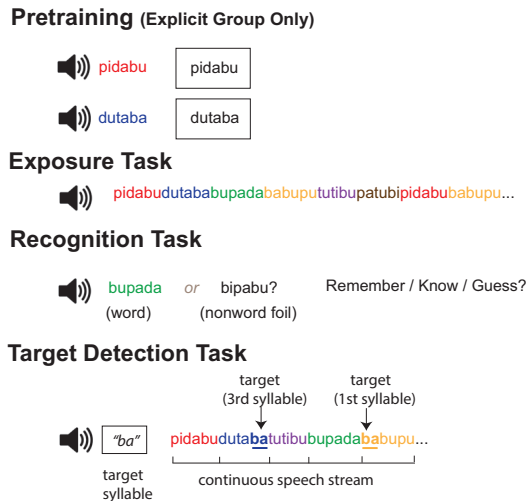


Figure 1. Summary of experimental design. Pretraining (for explicit group only) included auditorily presenting the six three-syllable nonsense words (*babupu*, *bupada*, *dutaba*, *patubi*, *pidabu*, *tutibu*) in isolation, with the written representation of each word shown visually. All participants (explicit group and implicit group) then completed the exposure, recognition, and target-detection tasks. The exposure task consisted of 21 min of continuous auditory exposure to the six repeating nonsense words. For the recognition task, each trial was composed of a word and nonword foil, presented auditorily. Participants indicated which item was more familiar, and provided a metamemory judgment. For the target-detection task, participants detected target syllables embedded in a continuous auditory speech stream composed of the six nonsense words. The syllable assigned as the target was rotated across trials.

a stream of standards elicits a robust P300, with P300 amplitude correlating inversely with target probability (Squires et al. 1976; Duncan-Johnson and Donchin 1977, 1982; Johnson and Donchin 1982). One widely accepted theory proposes that P300 reflects the allocation of attentional resources to the target, which are engaged in order to update the current neural representation of the stimulus environment (Polich 2007). When task demands and overall levels of attention and arousal are held constant, targets that are less probable or predictable should elicit larger P300 effects, as less predictable targets are more difficult to process and require greater attentional resources. We therefore hypothesized that participants should respond more quickly and elicit a reduced P300 to predictable syllable targets (i.e., those that occur in later syllable positions), reflecting a facilitation in processing due to statistical learning. This result would be consistent with our previous results from this task (Batterink et al. 2015). We further hypothesized that if explicit training enhances the function of statistical learning, learners in the explicit condition should show faster RTs and reduced P300s to target syllables compared with learners who have not benefited from explicit training. In contrast, if learners in the implicit condition show faster RTs and reduced P300s relative to explicit learners, this would suggest that the implicit abstraction of information through passive exposure, as occurs during statistical learning, confers certain advantages over more explicit modes of learning.

As a manipulation check to confirm that our training manipulation was successful, we also assessed learning directly using a forced-choice recognition task combined with a remember/know procedure (Fig. 1). In this task, participants were asked to discriminate between items presented during training and foil items, and to report on their awareness of memory retrieval for each trial. If explicit training successfully enhanced learners'

explicit representations of the statistical structure, explicitly trained participants should show higher accuracy and a greater proportion of “remember” responses on this task than implicitly trained participants, reflecting better explicit recollection. We also expected explicitly trained participants to show an enhanced late positive component (LPC) effect relative to implicitly trained participants. The LPC, a positive-going ERP modulation with an onset of ~400–500 msec post-stimulus, has been specifically linked to recollection (Paller and Kutas 1992; Rugg and Curran 2007). The LPC may reflect the amount of information recollected in response to a test item (Vilberg et al. 2006), although its precise functional significance continues to be debated (Rugg and Curran 2007). In our study, we hypothesized that explicitly trained participants should elicit a larger LPC to trained items compared with implicitly trained participants, as the former should have access to a stronger representation of the learned words. Results from the recognition task thus allowed us to establish whether supplementary explicit training functioned as intended to create stronger explicit memory representations. In summary, the present experiment was designed to determine whether explicit modes of learning relative to passive statistical learning confers advantages or disadvantages for online processing.

Results

Target-detection task

Behavior

As hypothesized, across all participants, RTs were faster for syllables occurring in later positions (Position effect: $F_{(2,78)} = 72.9$, $P < 0.001$; linear contrast: $F_{(1,39)} = 73.1$, $P < 0.001$). Planned contrasts revealed that there was a significant facilitation in RTs to third-position syllables as compared to second-position syllables ($F_{(1,39)} = 124.3$, $P < 0.001$), but that RTs to initial-position syllables versus second-position syllables did not significantly differ ($F_{(1,39)} = 0.33$, $P = 0.57$).

RTs over the three-syllable positions differed significantly as a function of group, reflecting a crossover interaction (Syllable Position \times Group: $F_{(2,78)} = 15.4$, $P < 0.001$) (Fig. 2A). Implicitly trained participants responded marginally faster to first-syllable targets than explicitly trained participants ($t_{39} = -1.86$, $P = 0.068$), whereas explicitly trained participants responded significantly faster to third-syllable targets than implicitly trained participants ($t_{39} = 2.39$, $P = 0.022$). For second-syllable targets, there was no significant difference in RTs between the two groups ($t_{39} = -0.009$, $P = 0.99$).

Because motor responses take ~200 msec to execute (e.g., Lakhani et al. 2011), we considered as hits only responses that occurred between 200 and 1200 msec after target onset. Responses that occurred after the onset of the target but prior to 200 msec (i.e., 0–200 msec post-stimulus) were considered to be early or anticipatory responses. This classification was based on the assumption that participants could have predicted upcoming targets and prepared their motor response prior to the actual onset of the target in order to execute a sub-200 msec response. Interestingly, the number of early or anticipatory responses differed as a function of group (Fig. 2B). Explicitly trained participants made a significantly greater number of anticipatory responses to targets (Group effect: $F_{(1,38)} = 6.25$, $P = 0.017$), an effect which differed as a function of syllable position (Syllable Position \times Group: $F_{(2,76)} = 3.85$, $P = 0.050$). Follow-up tests revealed that there was no significant difference in the number of anticipatory responses between groups for first- and second-syllable targets (for both $P > 0.13$), whereas a significantly greater number of anticipatory responses to third-syllable targets were made by explicitly trained

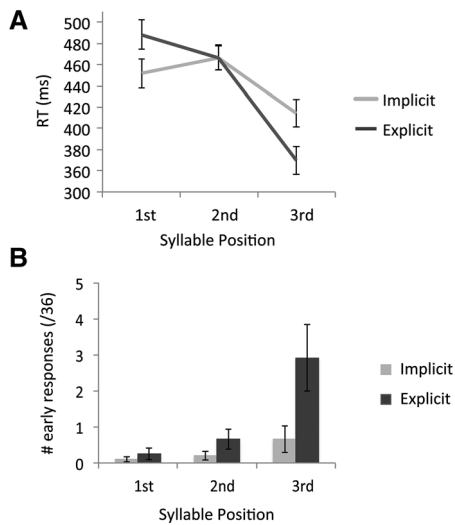


Figure 2. Behavioral results from the target-detection task. (A) RTs as a function of syllable position, by group. A significant crossover interaction was found. Error bars represent SEM. (B) Number of early responses (0–200 msec) to each target as a function of syllable position, by group. Error bars represent SEM.

participants relative to implicitly trained participants (Syllable 3: $t_{38} = 2.27$, $P = 0.032$) (Fig. 2B). Furthermore, for implicitly trained participants, the number of anticipatory responses did not reliably exceed 0 for first- or second-syllable targets ($P > 0.1$) and was only marginally > 0 for third-syllable targets ($t_{19} = 1.78$, $P = 0.091$). In contrast, explicitly trained participants made a significant number of anticipatory responses for both second ($t_{19} = 2.37$, $P = 0.028$) and third ($t_{19} = 3.14$, $P = 0.005$) syllable targets, but not for first-syllable targets ($t_{19} = 1.56$, $P = 0.14$).

ERPs during target detection

ERPs recorded during the target-detection task are shown in Figure 3. ERPs were analyzed over three intervals as follows.

400–800 msec. The interval from 400 to 800 msec was selected to quantify P300 on the basis of previous findings (Polich 2007; Batterink et al. 2015) and visual inspection of the waveforms. Owing to the posterior distribution of the P300, only amplitudes from posterior electrodes were included in this analysis. Across all participants, there was a linear effect of syllable position on P300 amplitude (Position effect: $F_{(2,78)} = 3.10$, $P = 0.056$; linear contrast: $F_{(1,28)} = 4.87$, $P = 0.033$) (Fig. 3A). Initial-position targets elicited the largest P300, medial-position targets elicited a moderate P300, and final-position targets elicited the smallest P300. The syllable position effect did not significantly differ as a function of group (Position effect \times Group: $F_{(2,78)} = 0.87$, $P = 0.41$).

0–300 msec. This earlier interval was selected, prior to the peak of the P300, to evaluate whether training condition modulated predictive or anticipatory processing in any of the three-syllable conditions. Again, only posterior electrodes were included in the analysis. If explicit training results in greater anticipatory processing, a larger P300 effect to later syllables would be expected in the explicit group compared with the implicit group during this early interval. Consistent with this notion, the syllable position effect differed significantly between groups (Position effect \times Group: $F_{(2,78)} = 4.89$, $P = 0.012$). Follow-up

analyses indicated that ERPs did not differ between groups for first-position targets ($F_{(1,39)} = 0.27$, $P = 0.61$) or second-position targets ($F_{(1,39)} = 0.68$, $P = 0.41$), but that explicitly trained participants showed a significantly larger positivity than implicitly trained participants to third-position targets ($F_{(1,39)} = 12.8$, $P = 0.001$) (Fig. 3B).

0–1000 msec. Finally, a separate analysis across all electrodes using this very broad time interval was conducted to examine whether implicitly and explicitly trained participants showed differential positivities overall. Enhanced positive ERP amplitudes across the entire epoch may reflect a combination of P300, indexing target detection and evaluation, and LPC, reflecting extended retrieval of the novel pseudowords from long-term memory. Thus, greater positive ERP amplitudes can be taken as a general index of level of neural recruitment or cognitive effort. Because participants may anticipate second- and third-syllable targets prior to their actual occurrence, such effects may be temporally blurred, and thus a broad time interval is appropriate for this analysis. We hypothesized that explicit training may result in greater engagement of cognitive resources, as reflected by larger positive ERP amplitudes. Consistent with this hypothesis, implicitly and explicitly trained participants showed significantly different syllable position effects during this time interval (Position effect \times Group: $F_{(2,78)} = 3.56$, $P = 0.035$). There was no group difference for first ($F_{(1,39)} = 0.51$, $P = 0.48$) or second ($F_{(1,39)} = 0.74$, $P = 0.40$) syllable targets, but explicitly trained participants showed a significantly greater positivity overall than implicitly trained participants to third-syllable targets ($F_{(1,39)} = 5.16$, $P = 0.029$) (Fig. 3B).

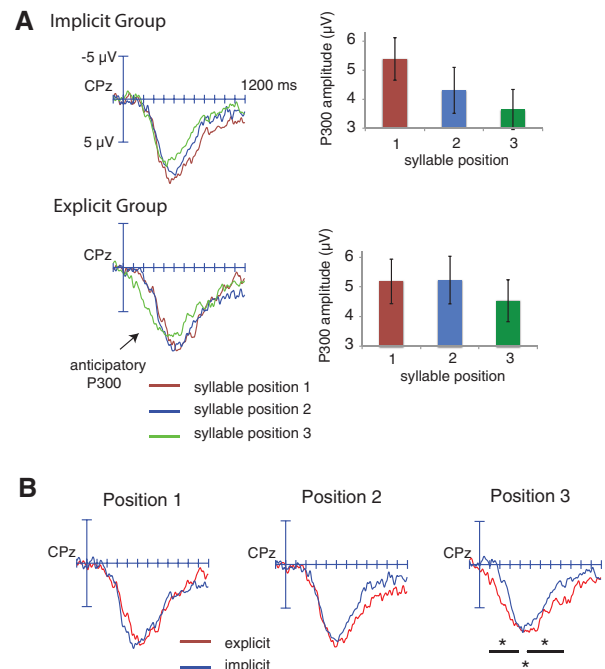


Figure 3. ERP results from the target-detection task. (A) ERPs time-locked to targets occurring in word-initial, word-medial, and word-final positions in the speeded target-detection task, as a function of group. Only correctly detected targets are included in these averages. The bar graph displays mean ERP amplitudes across all electrodes as a function of target position. Error bars represent SEM. (B) ERPs to each of the three-syllable targets are directly compared between implicit and explicit groups.

Recognition task

Behavior

Mean accuracy across all participants was significantly above chance (75%, $SD = 19\%$; $t_{40} = 8.39$, $P < 0.001$). Across both groups (including only participants who made responses in all three conditions; $n = 30$), “remember” responses were the most accurate followed by “familiar” responses, with “guess” judgments showing the lowest degree of accuracy (Memory Judgment effect: $F_{(2,56)} = 14.4$, $P < 0.001$; linear contrast: $F_{(1,28)} = 21.7$, $P < 0.001$) (Fig. 4A). This finding violates the zero-correlation criterion for implicit knowledge proposed by Dienes and Berry (1997), indicating that participants gained meta-knowledge of the underlying statistical structure. When participants claimed to be guessing, accuracy was not significantly above chance, neither across both groups (mean = 54.0%, $t_{32} = 0.82$, $P = 0.42$), nor when the two groups were considered separately (explicit group: mean = 56.7%, $t_{12} = 0.68$, $P = 0.51$; implicit group: mean = 52.3%, $t_{19} = 0.45$, $P = 0.66$). This result violates Dienes and Berry’s (1997) guessing criterion, again failing to provide evidence that recognition performance was driven by implicit knowledge.

As expected, explicitly trained participants performed significantly more accurately than implicitly trained participants on the recognition task ($t_{39} = 10.4$, $P < 0.001$). Mean accuracy for explicitly trained participants was 91.5% ($SD = 7.8\%$), whereas mean accuracy for implicitly trained participants was 59.3% ($SD = 11.7\%$). Recognition performance was significantly above chance for both groups (explicit: $t_{19} = 23.9$, $P < 0.001$; implicit: $t_{20} = 3.62$, $P = 0.002$). Proportion of “remember,” “familiar,” and “guess” responses also significantly differed by group (Memory Judgment \times Group: $F_{(2,78)} = 27.0$, $P < 0.001$) (Fig. 4B). Follow-up t -tests indicated that explicitly trained participants made significantly more “remember” responses than did implicitly trained participants ($t_{39} = -5.91$, $P < 0.001$), whereas implicitly trained participants made significantly more “familiar” and “guess” responses than did explicitly trained participants (familiar:

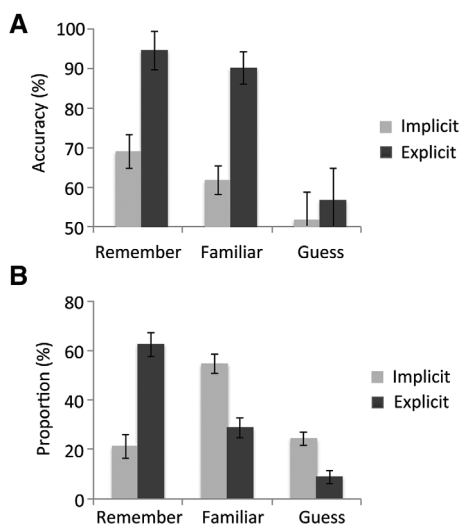


Figure 4. Behavioral results from the recognition judgment task. (A) Accuracy on the recognition judgment task as a function of metamemory judgment, by group. Only participants with responses in all three metamemory conditions are included in these averages, reflecting statistical analyses reported in the paper ($n = 17$ implicit group, $n = 13$ explicit group). Note that overall means from all participants are very similar ($< 2\%$ discrepancy per value). Chance performance on this task is 50%. Error bars represent SEM. (B) Proportion of each metamemory response by group. All participants are included. Error bars represent SEM.

$t_{39} = 4.49$, $P < 0.001$; guess: $t_{39} = 3.90$, $P < 0.001$). These results indicate that the pretraining manipulation successfully strengthened participants’ explicit memory for the nonsense words.

ERPs during recognition

ERPs recorded during the recognition task are shown in Figure 5. Two intervals were selected for the ERP analyses on the basis of previously published findings (Rugg et al. 1998; Friedman and Johnson 2000; Rugg and Curran 2007; Voss and Paller 2008) and visual inspection of the waveforms. An early interval was selected from 200 to 500 msec, corresponding to the general latency of previously identified ERP effects that have been linked to familiarity or perceptual priming in recognition tasks (e.g., Rugg et al. 1998; Paller et al. 2003). In the current data, this analysis captured an earlier, relatively broadly distributed positivity and included all nine electrode regions. A second interval was selected from 500 to 1000 msec, corresponding to the general latency of the LPC (e.g., Rugg and Curran 2007). Because this effect is more posteriorly distributed, this analysis included only central and posterior electrode regions. Two separate analyses were conducted, consisting of correct trials alone and incorrect trials alone. Incorrect analyses were designed to identify ERP memory effects that might be related to implicit memory based on previous experience with a stimulus but without leading to correct recognition.

200–500 msec. Correct trials only. Across all participants, words elicited a significantly larger positivity than nonword foils ($F_{(1,39)} = 17.85$, $P < 0.001$), an effect that was maximal over midline sites (Word Class \times Left/Mid/Right: $F_{(2,78)} = 8.97$, $P < 0.001$). Across all electrodes, this effect did not significantly differ as a function of group (Group \times Word Class: $F_{(1,39)} = 1.82$, $P = 0.19$) (Fig. 5A), but was marginally reduced in the explicit relative to the implicit group at right anterior and central sites (Group \times Word Class \times Left/Mid/Right \times Ant/Post: $F_{(4,156)} = 3.43$, $P = 0.017$; follow-up analysis over right central and anterior electrode groups: Word Class \times Group: $F_{(1,39)} = 3.09$, $P = 0.087$).

Incorrect trials only. Only participants with at least five incorrect artifact-free trials per condition were included in this analysis ($n = 25$; $n = 20$ implicit group, $n = 5$ explicit group). Similar to the ERP average for correct trials, words elicited a significantly larger positive ERP than nonwords ($F_{(1,23)} = 7.23$, $P = 0.013$) (Fig. 5B). This effect showed a maximal distribution over anterior and midline sites (Word Class \times Left/Mid/Right: $F_{(2,46)} = 3.53$, $P = 0.043$; Word Class \times Ant/Post: $F_{(2,46)} = 4.72$, $P = 0.027$).

500–1000 msec. Correct trials only. Words elicited a significantly larger LPC than nonword foils ($F_{(1,39)} = 21.0$, $P < 0.001$), an effect that was largest over midline sites (Word Class \times Left/Mid/Right: $F_{(2,78)} = 7.36$, $P = 0.001$). The LPC effect was significantly larger in the explicit relative to the implicit group (Group \times Word Class: $F_{(1,39)} = 5.02$, $P = 0.031$) (Fig. 5A). Follow-up analyses demonstrated that the LPC word–nonword effect was significant in the explicit group ($F_{(1,19)} = 25.0$, $P < 0.001$), but was not significant in the implicit group ($F_{(1,20)} = 2.58$, $P = 0.12$).

Incorrect trials only. On trials where participants failed to correctly discriminate between the word and nonword foil, there was no significant LPC difference between words and nonwords ($F_{(1,23)} = 1.05$, $P = 0.32$, all distributional interactions $P > 0.19$) (Fig. 5B).

Differences between the two word class effects as a function of group

The above results suggest that the earlier positivity differentiates words and nonwords in both groups, whereas the later positivity

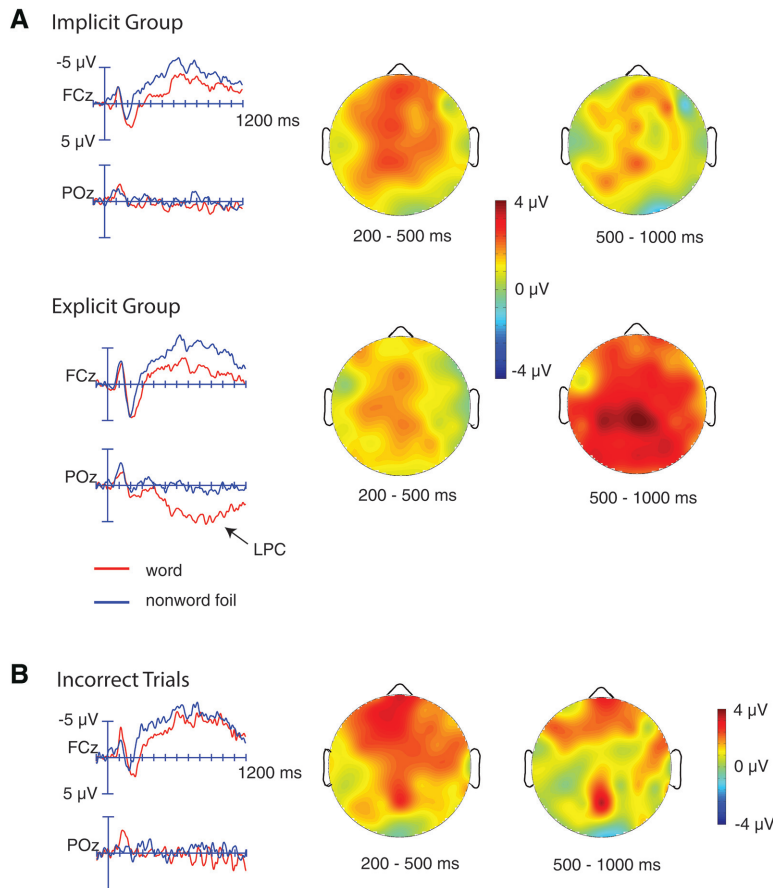


Figure 5. ERP results from the recognition judgment task. (A) ERPs time-locked to words and nonword foils in the recognition judgment task, by group. Only correct trials are included in these averages. Scalp topographies of the early positivity and LPC are shown on the right. (B) ERPs time-locked to words and nonword foils that were not correctly recognized. All participants with a sufficient number of incorrect trials are grouped together in this average (20 implicit-group participants and five explicit-group participants).

differentiates between words and nonwords only in the Explicit group. To confirm this dissociation statistically, an additional ANOVA was conducted, including all nine electrode regions, to examine whether the two groups showed significantly different profiles between the two word–nonword effects (the earlier positivity and the LPC). This analysis revealed a significant Time Interval \times Group \times Word Class interaction ($F_{(1,39)} = 13.4$, $P = 0.001$), indicating that there is a significant dissociation between the two ERP effects as a function of group.

ERP-behavioral correlations

To examine whether ERP effects observed during the recognition task could be classified as predominantly reflecting implicit versus explicit memory, we examined correlations between behavioral measures of explicit recognition and amplitude of these effects. Across subjects, explicit recognition significantly correlated with the LPC effect at left posterior ($r = 0.41$, $P = 0.008$) and right posterior regions ($r = 0.45$, $P = 0.003$), and correlated marginally with LPC amplitude at the midline posterior region ($r = 0.27$, $P = 0.083$). Explicit recognition did not show significant positive correlations with the early positivity effect (from 200 to 500 msec; for all $P > 0.8$ for $r > 0$). A direct comparison of the strongest positive correlation found between explicit recognition and the ERP effect at each time interval (200–500 msec versus 500–1000

msec), observed at the right posterior electrode group in both cases, showed that explicit recognition correlated significantly more strongly with the LPC effect than the early positivity ($z = 2.86$, $P = 0.0043$). This result provides evidence in support of a dissociation between these two ERP effects and the degree to which they relate to explicit memory.

Recall task

The recall task was administered twice to explicitly trained participants (once after pretraining and again after the exposure period) and once to implicitly trained participants (after the exposure period only). On average, explicitly trained participants perfectly recalled 2.45 (SD = 1.36) out of 6 words on the first recall task and 2.6 (SD = 1.85) out of 6 words on the second recall task. That these values are not higher attests to the difficulty of freely recalling arbitrary three-syllable sequences. One implicitly trained participant recalled 1 word; all other participants in the implicit group did not recall any words. Not surprisingly, explicitly trained participants recalled significantly more words than implicitly trained participants on the post-exposure test ($t_{39} = 8.01$, $P < 0.001$).

Discussion

In this study, we investigated whether explicit training on statistical regularities influences the function of statistical learning, as assessed through a speeded, incidental, performance-based task. To address this question, we compared behavioral performance and ERPs between two groups of participants who differed in whether or not they received explicit training on the statistical regularities hidden in the stimulus sequences.

Reaction-time task

The speeded, incidental, target-detection task revealed significant statistical learning effects in both groups of participants. RTs were faster for syllables occurring in later positions, and P300 amplitudes to syllable targets scaled linearly with syllable position, showing the largest amplitude to initial-position targets and the smallest amplitude to final-position targets. This graded P300 effect represents a neural index of statistical learning, and replicates our previous result (Batterink et al. 2015). These findings demonstrate that processing is progressively facilitated when targets occur in more predictable positions and confirm that the reaction-time task is sensitive to statistical learning.

Of more direct relevance to the main hypotheses of the study, implicitly and explicitly trained participants performed differently on the speeded, incidental target-detection task and also showed different ERP effects. Reaction-time patterns showed a significant crossover interaction across groups, with implicitly trained participants responding marginally more quickly to initial-position targets and explicitly trained participants

responding significantly more quickly to final-position targets. Although no ERP differences between groups were found for first- and second-position targets, explicitly and implicitly trained participants showed significant ERP differences when processing third-syllable targets. Relative to implicitly trained participants, the positivity to third-syllable targets in explicitly trained participants began earlier and was significantly larger, reflecting anticipatory processing and greater overall processing effort, respectively.

These results indicate that strengthening statistical representations explicitly has consequences for online usage of the knowledge gained through statistical learning, exerting both costs and benefits depending upon the availability of contextual information. For final syllable targets, explicitly trained participants responded significantly faster and made significantly more anticipatory responses (i.e., responses occurring within 200 msec after the onset of the target) relative to implicitly trained participants. Thus, when a target occurred near the end of a word, explicitly trained participants tended to prepare their motor response to the target prior to target onset. In addition, explicitly trained participants responded marginally less quickly to initial-position targets relative to implicitly trained participants. Based on these results, we speculate that explicitly trained participants performed the target-detection task while simultaneously attempting to explicitly recollect the learned words, essentially performing a dual task. In contrast, implicitly trained participants would not have had access to strong explicit memory traces of the learned words and would not have engaged in explicit recollection processes to the same extent, thus performing only a single task (auditory target detection). When a target occurred in an informative context (i.e., as the third syllable of a word), explicitly trained participants could make use of their explicit memory for the word, anticipating the upcoming target, which would enable them to prepare the appropriate motor response more quickly. At the output level, this confers a benefit for task processing. However, when contextual information was not available (i.e., when a target occurred as the first syllable of a word), efforts to explicitly recollect the learned words provided no useful information about the presence of an upcoming target and was counterproductive. According to this interpretation, the processing cost of recollecting this information is reflected by slower RTs to first-syllable targets.

A similar RT phenomenon was reported in a statistical learning study that presented participants with a sequence composed of both paired and unpaired images (Turk-Browne et al. 2010). RTs to the first item of a pair were delayed relative to RTs to unpaired trials, suggesting that the initial item of a pair engaged anticipatory processes that competed with the evaluation and response to that item. Our results are consistent with this interpretation, and further suggest that the cost of recalling a learned association and anticipating the next item is more pronounced when this knowledge is fully explicit. We hypothesize that this “recollection cost” is ongoing and contributes to responses for second- and third-syllable targets as well, but that the benefits conferred by being able to explicitly anticipate these more predictable targets outweigh this cost, resulting in faster overall RTs for third-syllable targets. For second-syllable targets recollection, costs and benefits are approximately equivalent, leading to no RT differences between the explicit and implicit groups. However, given that the RT difference to first-syllable targets between explicitly and implicitly trained participants did not reach statistical significance, this interpretation must remain somewhat speculative.

Nonetheless, ERP results also support the interpretation that explicitly trained participants engaged in explicit recollection of the novel words throughout the task, and that this recollection

exerted an online processing cost. Explicitly trained participants exhibited a significantly larger P300 than implicitly trained participants in a very early time interval (0–300 msec), suggesting that they began recruiting neural mechanisms to facilitate the detection of these targets even prior to the physical onset of the stimuli. This finding converges with the behavioral anticipation effect observed in explicitly trained participants (Fig. 2B). Across the entire epoch, explicitly trained participants also showed a significantly greater positivity to third-syllable targets relative to implicitly trained participants. This effect likely reflects a combination of the P300, indexing target detection and evaluation, and the LPC, reflecting extended retrieval of the novel pseudo-words from long-term memory. If explicit knowledge facilitates processing, to the extent that ERP amplitude indexes processing efficiency (with lower amplitude ERPs indicating facilitated processing and less effort), we should observe a *decreased* positivity to third-syllable targets in the explicit relative to the implicit group. That an increased positivity was in fact observed suggests that explicitly trained participants recruited more neural resources when processing third-syllable targets than implicitly trained participants. In other words, explicitly trained participants’ faster RTs to third-syllable targets appears to be driven by greater engagement of controlled, effortful processes. One speculation is that explicitly trained participants subjectively experienced greater cognitive effort than implicitly trained participants, as they were constantly engaged in explicit recollection as an additional task. Relative to implicit knowledge, the processing of explicit knowledge may induce greater demands on controlled, limited-capacity resources.

Statistical learning resembles sequence learning in that both involve the acquisition of patterns of stimuli that unfold in time (Daltrozzo and Conway 2014). Our finding that enhanced ERP effects are elicited during stimulus processing when participants have explicit knowledge converges with a number of previous ERP studies on sequence learning, which have typically used the serial reaction-time (SRT) task. In a typical SRT-paradigm, learners respond to visual cues that follow a hidden repeating sequence. Learning is measured as the gradual reduction in RTs that takes place across the sequential trials, relative to random or deviant trials that do not follow a repeating pattern. In general, ERP studies on the SRT task have shown that explicit knowledge of the underlying sequence is associated with enhancements in ERP effects linked to learning such as the N200 and the P300. For example, Eimer et al. (1996) found that an N200 was elicited to rare deviant items that occurred within a repeating sequence of standards, and that this effect was larger for participants who showed post hoc explicit knowledge of the sequence, as assessed through verbal reports and a recognition test. Using a similar post hoc sorting method, Rüsseler and Rossler (2000) found both an N200 and P300 effect to deviant items in participants with explicit knowledge, whereas implicit learners did not show these effects. Similar results were reported by Schlaghecken et al. (2000), who identified explicitly learned sequence parts for each participant by using the process-dissociation procedure (Jacoby 1991). Deviant items in explicitly represented sequences elicited an enhanced N200 and P300 relative to standards, whereas deviant items in sequences that were not explicitly represented did not modulate these ERP components.

Other studies have examined the effect of explicit knowledge on sequence learning by manipulating the instructions given to participants at the time of learning, similar to the design used in this study. Baldwin and Kutas (1997) found that targets that violated an underlying grammar in a sequence-learning task elicited an enhanced P300 relative to grammatical targets. This effect was observed in both implicitly trained participants and participants who were taught the grammatical rules prior to the task, but

was larger in the explicit group. Similarly, using an SRT-paradigm, Rüsseler et al. (2003a) found an enhanced N200 and P300 to deviant stimuli that disrupted a repeating sequence for intentional learners—who were informed of the presence of the sequence—but not for incidental learners. The same group of authors (Rüsseler et al. 2003b) also compared ERPs between an intentional and incidental group of learners using a modified version of the SRT, in which a central target surrounded by flankers either followed a repeating sequence or was determined randomly. Erroneous responses to the central target elicited a larger Error Related Negativity in intentional versus incidental learners, indicating that explicitly searching for a sequential regularity led to a more intensive engagement of the error monitoring system. In summary, ERP effects associated with learning are typically more robust when participants have explicit knowledge of the underlying structure, and in some cases, ERP effects related to learning may be observed only under explicit conditions. Behaviorally, explicit awareness of the stimulus structure can often lead to faster or more efficient behavioral responding (e.g., Eimer et al. 1996; Baldwin and Kutas 1997). In principle, these faster responses could reflect facilitation at the neural level, or could be driven by greater recruitment of limited-capacity resources. ERPs are helpful in resolving this question, suggesting that relative to implicit knowledge, the processing of explicit knowledge produces greater neural activation, likely reflecting the recruitment of additional neural systems that are not activated under implicit conditions.

Although this study did not directly assess whether statistical learning occurred in explicitly trained learners to the same extent as in implicitly trained learners, the finding that explicitly trained learners were able to capitalize on their explicit knowledge to support performance during the target-detection task suggests that at least some degree of passive statistical learning occurred in this explicit group. This conclusion is based on behavioral findings from the SRT task, in which RT benefits associated with explicit knowledge appear to emerge only if some implicit knowledge is acquired in parallel. Reber and Squire (1998) found that when explicit knowledge of the repeating sequence was given in isolation, without allowing learners the opportunity to actually perform the SRT task, RTs on the subsequent performance test revealed no evidence of learning. In other words, these learners were unable to recruit their explicit knowledge to support online performance on the SRT task. In contrast, when learners acquire both explicit and implicit knowledge of the repeating sequence, as a consequence of being given the opportunity to perform the stimulus-response task, explicit knowledge leads to faster RTs compared with implicit knowledge alone (Willingham et al. 1989; Frensch and Miner 1994). These results suggest that explicit knowledge boosts performance only if learners have the opportunity to acquire some implicit knowledge of the sequence in parallel. By extension, it is likely that explicitly trained learners in this study would have been unable to recruit their explicit knowledge to help them perform the target-detection task if no statistical learning had occurred. However, this suggestion must remain speculative for now. Addressing this question directly will require future research specifically designed to examine whether preexisting explicit knowledge of the stimulus structure disrupts the passive statistical learning process.

Finally, it is worth noting that significant RT differences emerged only between second- and third-syllable targets. In our stimuli, the first two syllables predicted the final syllable deterministically, whereas the first syllable predicted the second syllable only probabilistically. That there were no significant RT differences between first- and second-syllable targets suggests that overt behavioral responses in our task rely primarily upon deterministic rather than probabilistic cues. Relying upon deter-

ministic cues may minimize errors, so that a behavioral response is executed only when the learner is certain that the next stimulus will be a target. According to this idea, although more contextual information is available prior to the onset of second-syllable targets relative to first-syllable targets, the learner may not prepare his or her response to these targets in advance because their identities cannot be predicted with certainty. This would lead to faster responding for the final-position target, and no difference in response times between the first and second target. This idea is consistent with two previous visual statistical learning studies with stimulus triplets in which both the second and third items were deterministically predicted by the first item (Turk-Browne et al. 2005; Kim et al. 2009), unlike the present stimulus streams. In contrast to the RT pattern we observed, both these studies found a graded RT effect as a function of syllable position. In addition, a recent auditory statistical learning study using syllable triplets in which the second and third syllables were deterministically predicted by the first syllable found the same pattern of results: RTs were slowest to target syllables occurring in the first position, intermediate to syllables occurring in the second position, and fastest to syllables occurring in the final position (Franco et al. 2015). Taken together, these results suggest that significant RT differences may emerge only when an item can be uniquely predicted from the preceding context.

Recognition task

Results from the recognition task serve as a manipulation check, confirming that our training manipulation had a powerful effect on the strength of learners' explicit memory representations. On the recognition task, explicitly trained participants showed much higher overall accuracy than implicitly trained participants (92% compared with 59%). In addition, explicitly trained participants had a greater proportion of "remember" responses, whereas implicitly trained participants showed a greater proportion of "familiar" and "guess" responses. Explicitly trained participants also showed a significantly larger LPC effect to words compared with nonword foils than implicitly trained participants, indicative of explicit recollection mechanisms. These results indicate that the pretraining manipulation successfully strengthened participants' explicit memory traces for the novel pseudowords. As expected, explicit pretraining created stronger explicit memory representations of the novel pseudowords, which in turn led to greater engagement of conscious recollection when participants in this group were required to discriminate between words and nonword foils.

Results from the recognition task also have implications for statistical learning in general. In both implicitly and explicitly trained participants, accuracy correlated with participants' subjective experience of recollection, with the highest level of accuracy observed for remember responses, a moderate level of accuracy for familiar responses, and the lowest accuracy for guess responses. In addition, accuracy was not significantly above chance for guess responses in either group. Thus, according to both the zero-correlation criterion and the guessing criterion proposed by Dienes and Berry (1997), recognition judgments were strongly influenced by explicit memory, with no evidence that implicit knowledge contributed to these judgments. These results support the idea that, although statistical learning involves passive, incidental exposure to structured input and is generally thought of as an implicit process, it nonetheless produces explicit knowledge. However, the level of explicit knowledge typically produced is only moderate, and/or variable across participants. Statistical learning may produce additional implicit knowledge that is not captured by the recognition task, with the recognition task underestimating the total amount of knowledge that is acquired.

ERP data shed light on the neural mechanisms used to support performance on this task. Two distinct ERP effects were revealed. The first effect was an early and relatively widespread positivity observed from 200 to 500 msec, greater to words compared with nonword foils. Several lines of converging evidence suggest that this first effect—in contrast to the later LPC—reflects implicit mechanisms that are dissociable from the ability to explicitly discriminate between words and nonword foils. First, the effect did not significantly differ between implicitly and explicitly trained participants, even though explicitly trained participants demonstrated a much better ability to explicitly discriminate between words and nonword foils than implicitly trained participants. Second, this effect did not correlate with explicit behavioral recognition across participants. Finally, the effect was observed even for incorrect trials. Taken together, these findings suggest that this earlier negativity is an obligatory response elicited by previously encountered words regardless of recognition accuracy, and that it reflects processes that operate independently of those that lead to correct discrimination. This effect appears to be somewhat similar in latency and distribution to a previously identified ERP correlate of implicit memory, described as an early (300–500 msec) positivity elicited to both correctly and incorrectly recognized studied words relative to new words (Rugg et al. 1998). The early negativity observed in this study may also represent a neural correlate of memory in the absence of conscious recollection, reflecting processes that emerge as a function of exposure and that are unrelated to individual recognition performance. These implicit memory processes appear to be operative in both groups of participants, regardless of training condition.

By comparison, the later LPC effect exhibited the reverse profile of effects. The LPC was significantly larger in explicit compared with implicitly trained participants, consistent with the idea that explicitly trained participants engaged in explicit retrieval for the novel words to a greater extent than did implicitly trained participants. In addition, significant correlations were found between explicit recognition and LPC amplitude, suggesting that participants who engaged in recollection to a greater extent achieved the highest recognition performance. Finally, no significant LPC effect was observed on incorrect trials, suggesting that this effect indexes processing associated with a correct decision. The LPC effect in this study is similar to old/new effects observed during explicit memory tasks, which have been linked to the recollection of specific information (Rugg et al. 1998; Rugg and Curran 2007; Voss and Paller 2008). This LPC result also nicely converges with a sequence-learning study conducted by Ferdinand et al. (2010), who had participants perform an SRT task and then tested their recognition of sequence triplets using a forced-choice old/new task. Participants who had greater explicit knowledge of the sequence (classified as “verbalizers” based on verbal report) showed a large posterior positivity to second and third items of sequence triplets in this recognition task, resembling the LPC effect in this study. In contrast, participants without any explicit knowledge of the sequence (“nonverbalizers”) did not show this effect. Thus, explicit awareness of the underlying stimulus structure appears to be necessary to elicit the LPC effect, consistent with the idea that the LPC effect indexes conscious memory retrieval. In this study, incorrect trials did not elicit an LPC effect, suggesting that explicit retrieval is required for successful discrimination between learned words and nonword foils in the recognition task; implicit intuition does not appear to be sufficient to consistently support correct responding. In sum, ERP data indicate that both groups formed implicit memory traces for the learned words, which did not necessarily lead to correct recognition. In contrast, explicitly trained participants engaged in conscious recollection of the novel words to a far greater degree

than did implicitly trained participants, a process that supported their superior behavioral performance on the recognition task.

Conclusions

The route through which learning occurs has consequences for the types of representations that are formed, and ultimately influences online processing. Our results suggest that the passive, gradual abstraction of information, as occurs in statistical learning, may confer certain advantages over explicit forms of learning. Statistical learning typically produces a relatively weak level of explicit knowledge. In addition to weak levels of explicit knowledge, dissociable representations that cannot be consciously retrieved and that are implicit in nature may also be accrued. When explicit knowledge is available at only a weak level, learners are unlikely to attempt to retrieve this knowledge during online processing, instead relying predominantly upon implicit representations. In contrast to explicit retrieval processes, these implicit representations can be recruited automatically, without conscious effort or awareness. Because attention and conscious control are limited-capacity resources (e.g., Schneider and Shiffrin 1977; Shiffrin and Schneider 1977), reliance upon implicit over explicit representations in this context should be advantageous, allowing observers to direct their attention more fully to other aspects of the environment and to other ongoing tasks. However, when explicit representations produced by statistical learning are strengthened, learners may be induced to rely more heavily upon this explicit knowledge, placing greater demands on limited-capacity resources such as attention and working memory. This dependence upon explicit knowledge exerts an overall cost for online processing, diverting resources away from other aspects of task processing.

These results clarify the function of statistical learning. Outside of the laboratory, the function of statistical learning is presumably to facilitate online processing and performance, not to enable learners to explicitly distinguish between old and novel stimuli. For example, statistical learning has been shown to trigger “perceptual adaptation,” which results in facilitated object processing after only two to three repetitions of a regularity but does not contribute to subsequent explicit recognition (Turk-Browne et al. 2010). This mechanism, mediated partially by anterior hippocampus, enables more efficient responses to predictable trials, while at the same time exerting costs for predictive trials relative to a nonpredictive baseline. Perceptual anticipation produces predictive potentiation of category-selective visual areas in anticipation of a stimulus from that category, along with suppression of the area when predictable stimuli come from a different category, highlighting the highly adaptive nature of this mechanism.

The idea that statistical learning is adaptive is clearly illustrated by several examples within the context of language processing. For example, repeated exposure to sentences with temporary syntactic ambiguities (often referred to as “garden path sentences”) can reduce or even reverse the processing disadvantage associated with these unexpected structures (Fine et al. 2013). These effects are thought to occur as language users rapidly compute and converge toward the statistical contingencies of the current linguistic environment, allowing them to anticipate linguistic events according to their probability and minimizing surprises. This adaptation can be very rapid, emerging within several trials. Exposure to relatively unusual syntactic structures through passive reading leads participants to judge novel sentences following the same structures as more acceptable (Luka and Barsalou 2005; Luka and Choi 2012), similar to the mere exposure effect (Zajonc 1968). This effect also occurs very rapidly, emerging after a single

prior exposure to a similar sentence, and can persist even after a full week of normal exposure to natural language (Luka and Barsalou 2012). These results indicate that incremental adjustments to the language processing system occur on a continual basis, allowing for the dynamic and flexible acquisition of novel syntactic representations based on the current environment. Similarly, exposure to nonnative, accented speech causes learners to rapidly adjust their reliance on particular acoustic dimensions, a process referred to as dimension-based statistical learning (Idemaru and Holt 2011). This mechanism allows learners to adapt to the substantial acoustic variability among different speakers, accents, and dialects.

Taken together, such findings indicate that statistical learning can be viewed as a continuous, rapid and incremental learning process, whereby computations of local statistics in novel environments enable people to predict and more efficiently process incoming input. Our results are in line with this view, as they demonstrate that information learned explicitly is not processed in the same way as information acquired through passive statistical learning. Whereas the processing of explicit knowledge requires greater neural activation and the recruitment of controlled, limited-capacity resources, implicit knowledge abstracted through statistical learning can lead to rapid, flexible, and effortless adaptation to the current environment.

Materials and Methods

Participants

Forty-one participants (20 women) were recruited at Northwestern University for this experiment. Participants were all fluent English speakers, between 18 and 36 yr old ($M = 21.6$ yr, $SD = 3.3$ yr) and had no history of neurological problems. Participants were paid \$10/h.

Stimuli

A visual summary of the experimental design is shown in Figure 1. For the learning phase, stimuli were modeled after previous auditory statistical learning studies (e.g., Saffran et al. 1996b, 1997). This language consists of 11 syllables combined to create six trisyllabic nonsense words (*babupu*, *bupada*, *dutaba*, *patubi*, *pidabu*, *tutibu*). Some members of the syllable inventory occur in more words than others in order to ensure varying transitional probabilities within the words themselves, as in natural language. A speech synthesizer was used to generate a continuous speech stream composed of the six trisyllabic nonsense words at a rate of ~250 syllables per minute. Each nonsense word was repeated 300 times in pseudorandom order, with the restriction that the same word never occurred consecutively. Because the speech stream contained no pauses or other acoustic indications of word onsets, the only cues to word boundaries were transitional probabilities, which were higher within words than across word boundaries (cf. Saffran et al. 1996b, 1997). The speech stream was edited to include a total of 31 pitch changes. Each pitch change represented either a 20-Hz increase or decrease from the baseline frequency (~160 Hz). As described in greater detail in the Procedure section, these pitch changes were introduced in order to provide participants with a cover task during the learning period, involving the detection of these infrequent pitch changes, in order to ensure adequate attention to the auditory stimuli. Finally, the speech stream was divided into three equal blocks, each one ~7 min in length.

For the recognition test phase, six nonword foils were created (*batabu*, *bipabu*, *butipa*, *dupitu*, *pubada*, *tubuda*). The nonwords consisted of syllables from the language's syllable inventory that never directly followed each other in the speech stream, even across word boundaries. Frequency of individual syllables across words and nonword foils were matched. We included a further check on the possibility of systematic differences between items

by running a group of control participants ($n = 11$), who completed the recognition task without prior exposure to the speech stream. The ability to discriminate words from nonwords was not reliably above chance (50.5%, $t_{10} = 0.17$, $P = 0.87$), indicating that above chance performance on the recognition task cannot be attributed solely to systematic differences between words and nonwords.

Finally, for the speeded target-detection task, 33 separate speech streams were created with the same speech synthesizer used to create stimuli in the learning phase. Each stream consisted of two repetitions of each of the six nonsense words, concatenated together in pseudorandom order. The speech streams for the target-detection task were produced at a somewhat slower rate than the original speech streams (~150 syllables per minute). This moderate rate was chosen to ensure that the task would be both feasible and challenging, providing an online measure of speeded processing. In order to compute RTs to target syllables, target syllable onsets were coded by three trained raters using both auditory cues and visual inspection of sound spectrographs. Any discrepancy >20 msec among one or more raters was resolved by a fourth independent rater.

Procedure

Participants were randomly assigned to an explicit ($n = 20$) condition or implicit ($n = 21$) condition. Participants assigned to the explicit condition were extensively pretrained on the six nonsense words prior to the exposure phase. During pretraining, participants listened to the words played individually and were required to identify each word from six presented options using a keypad. Participants were instructed to attempt to memorize the words during the pretraining phase, and were informed that later their memory for the words would be tested. Immediately after pretraining, participants completed a recall task in which they were asked to recall the six words that they had just learned. Participants in the implicit condition were not given any information about the nonsense words.

Participants in both groups were exposed to the same auditory speech stream. Both groups were informed that the speech stream contained occasional pitch changes and instructed to detect these pitch changes using the keypad, pressing one button for low pitch changes and another for high pitch changes. Participants in the explicit group were told that the language was made up of the six nonsense words that they had just learned. Participants in the implicit group were simply instructed to listen to the nonsense language and to detect the pitch changes. To increase interest in the task, participants earned a small amount of additional money (12 cents) for each successfully detected pitch change. Owing to technical issues, behavioral data for the pitch-detection task from three participants (one in the explicit condition and two in the implicit condition) could not be analyzed. Overall, the remaining participants performed well on the pitch-detection task, detecting 95% ($SD = 3.7\%$) of pitch changes. The performance of explicitly and implicitly trained participants (implicit = 94.6%; explicit = 96.1%) was not significantly different ($t_{36} = 1.30$, $P = 0.20$).

After finishing the listening phase of the experiment, participants in the implicit condition were informed that the nonsense language that they had just listened to was composed of individual words. All participants were then given a recall task in which they were asked to recall the six words. This was the first and only recall task administered to participants in the implicit condition, and the second of such tasks for participants in the explicit condition.

Next, participants completed a forced-choice recognition judgment task. Each trial included a word and a nonword foil. Participants gave two responses for each trial, (1) indicating which of the two sound strings sounded more like a word from the language, and (2) reporting on their awareness of memory retrieval, with *remember* indicating confidence based on retrieving specific information from the learning episode, *familiar* indicating a vague feeling of familiarity with no specific retrieval, and *guess* indicating no confidence in the selection. Each of the six words and six

nonword foils were paired exhaustively for a total of 36 trials. In half of the trials the word was presented first while in the other half the nonword foil was played first; presentation order for each individual trial (whether presented first/second) was counterbalanced across subjects. Each trial began with the presentation of a fixation cross. After 1000 msec, the first word was presented. The second word was presented 1500 msec after the onset of the first word. Individual word duration ranged from 800 to 900 msec.

Finally, participants completed the speeded target-detection task. On each trial, they were instructed to detect a specific syllable within a continuous speech stream. Both RT and accuracy were emphasized. Each of the 11 syllables of the language's syllable inventory (ba, bi, bu, da, du, pa, pi, pu, ta, ti and tu) served as the target syllable three times, for a total of 33 streams. The order of the 33 streams was randomized for each participant. Each stream contained between 2 and 8 target syllables, depending upon which syllable was the target, providing a total of 36 trials in each of the three-syllable conditions (word-initial, word-medial, and word-final). It was expected that RTs would be fastest to syllable targets that occur in the final position of a word, with word-initial and word-medial targets eliciting slowest and intermediate RTs, respectively. At the beginning of each trial, participants pressed "Enter" to listen to a sample of the target syllable. The stimulus stream was then initiated. The duration of each stimulus stream was ~15 sec, with an average SOA between syllables of ~400 msec. The interval between individual syllables was jittered due to variability in the speech streams created by the speech synthesizer, which is designed to approximate timing variability in natural speech.

Behavioral data analysis

For the target-detection task, median RTs to detected targets ("hits") were calculated for each syllable condition (word-initial, word-medial, and word-final) for each participant. Responses that did not occur within 0–1200 msec of a target were considered to be false alarms. RTs were analyzed using a repeated-measures ANOVA with syllable position (initial, medial, final) as a within-subjects factor and instruction condition (implicit, explicit) as a between-subjects factor. Planned contrasts were used to examine whether RTs decreased linearly as a function of syllable position. The numbers of anticipatory response occurrences in each condition were also analyzed using a repeated-measures ANOVA with the same factors as described above. Follow-up tests determined whether the number of anticipatory responses for each syllable position differed between groups.

For the recognition task, accuracy and proportion of responses were measured using repeated-measures ANOVAs with metamemory condition (remember, familiar, guess) as a within-subjects factor and instruction condition (implicit, explicit) as a between-subjects factor. For all accuracy analyses involving metamemory condition as a factor, only subjects with responses in all three conditions were included ($n = 17$ implicit group, $n = 13$ explicit group).

EEG recording and analysis

EEG was recorded with a sampling rate of 512 Hz from 64 Ag/AgCl-tipped electrodes attached to an electrode cap using the 10/20 system. Recordings were made with the Active-Two system (Biosemi, Amsterdam, The Netherlands), which does not require impedance measurements, an online reference, or gain adjustments. Additional electrodes were placed on the left and right mastoid, at the outer canthi of both eyes, and below both eyes. Scalp signals were recorded relative to the Common Mode Sense (CMS) active electrode and then re-referenced off-line to the algebraic average of the left and right mastoid. Left and right horizontal eye channels were re-referenced to one another, and the vertical eye channel was re-referenced to FP1.

ERP analyses were carried out using EEGLAB (Delorme and Makeig 2004). Data were band-pass filtered from 0.1 to 40 Hz. Large or paroxysmal artifacts or movement artifacts were identi-

fied by visual inspection and removed from further analysis. Data were then submitted to an Independent Component Analysis (ICA), using the extended runica routine of EEGLAB software. Ocular and channel artifacts were identified from ICA scalp topographies and the component time series, and removed. ICA-cleaned data were then subjected to a manual artifact correction step to detect any residual or atypical ocular artifacts not removed completely with ICA. For a subset of subjects ($n = 25$), one or more channels were identified as bad, excluded from all ICA decompositions, and subsequently interpolated. Finally, time-series epochs timelocked to critical events were extracted and plotted from –100 to 1200 msec with baseline correction to a 100-msec prestimulus interval. In the recognition task, ERPs were time-locked to the onsets of word and nonword foils, whereas in the target-detection task ERPs were time-locked to the onset of target syllables.

For statistical analyses, amplitudes were averaged across neighboring electrodes to form nine channel groups of interest (left anterior region: AF7, AF3, F7, F5, F3; left central region: FT7, FC5, FC3, T7, C5, C3; left posterior region: TP7, CP5, CP3, P7, P5, P3, PO7, PO3; midline anterior region: AFZ, F1, FZ, F2; midline central region: FC1, FCZ, FC2, C1, CZ, C2; midline posterior region: CP1, CPZ, CP2, P1, PZ, P2, POZ; right anterior region: AF4, AF8, F4, F6, F8; Right central region: FC4, FC6, FT8, C4, C6, T8; right posterior region: CP4, CP6, TP8, P4, P6, P8, PO4, PO8). Mean amplitudes were analyzed using a repeated-measures ANOVA with word class (word, nonword foil), left/mid/right (left, middle, right), and anterior/posterior (anterior, central, posterior) as within-subjects factors and group (explicit, implicit) as a between-subjects factor. For the recognition task, time interval (200–500 and 500–1000 msec) was also included as an additional factor in a separate ANOVA designed to assess differences between the two word class effects as a function of group. Greenhouse–Geisser corrections were applied for factors with more than two levels.

For the target-detection task, analyses were designed to investigate whether the amplitude of the P300 elicited by target syllables during the target-detection task varied as a function of syllable position. Only trials in which participants made a correct response within 1200 msec were included in these analyses (i.e., from 0 to 1200 msec). ERP averages thus included both "hits" and "anticipatory" responses; it was not possible to separate these conditions due to low number of anticipatory responses. For all time intervals, mean amplitudes to target syllables in the three-syllable conditions (initial, medial, and final) were calculated for each participant.

For the recognition task, analyses were designed to examine whether words and nonword foils elicited different memory-related ERP effects. For incorrect analyses, only participants with at least five trials per condition were included in the grand average, resulting in a total of 25 participants (20 implicit group and 5 explicit group). Group analyses of incorrect trials were not conducted, due to the low number of explicitly trained participants with a sufficient number of incorrect trials.

Finally, to explore whether ERP effects reflected processes contributing to explicit recognition, we computed several correlations between recognition accuracy and ERPs in the recognition task. For each participant, a behavioral measure of explicit recognition was computed as the average accuracy across remember and familiar responses (excluding guess trials). We examined whether this behavioral measure of explicit recognition correlated with two ERP effects in the recognition task, an early positivity at 200–500 msec and the LPC. Pearson's correlations were computed between the word effect at each of the nine electrode regions and the behavioral measure of explicit recognition. Finally, we examined whether explicit recognition correlated significantly more strongly with the LPC versus the early positivity by selecting the strongest positive correlation coefficient from among the nine electrode regions for each time interval. Using the web utility provided by Lee and Preacher (2013), we converted this coefficient into a z -score using Fisher's r - to z -transformation. The asymptotic covariance of the estimates was then computed and an asymptotic z -test conducted in order to test whether the

correlation coefficients differed significantly from one another (Steiger 1980).

Acknowledgments

This work was supported by NIH grants T32 NS 047987 and F32 HD 078223.

References

- Abla D, Katahira K, Okanoya K. 2008. On-line assessment of statistical learning by event-related potentials. *J Cogn Neurosci* **20**: 952–964.
- Arciuli J, von Koss Torkildsen J, Stevens DJ, Simpson IC. 2014. Statistical learning under incidental versus intentional conditions. *Front Psychol* **5**: 1–8.
- Baldwin KB, Kutas M. 1997. An ERP analysis of implicit structured sequence learning. *Psychophysiology* **34**: 74–86.
- Batterink L, Reber PJ, Neville HJ, Paller KA. 2015. Implicit and explicit contributions to statistical learning. *J Mem Lang* **83**: 62–78.
- Bertels J, Franco A, Destrebecqz A. 2012. How implicit is visual statistical learning? *J Exp Psychol Learn Mem Cogn* **38**: 1425–1431.
- Bertels J, Demoulin C, Franco A, Destrebecqz A. 2013. Side effects of being blue: influence of sad mood on visual statistical learning. *PLoS One* **8**: e59832.
- Cohen NJ, Squire LR. 1980. Preserved learning and retention of pattern-analyzing skill in amnesia: dissociation of knowing how and knowing that. *Science* **210**: 207–210.
- Daltrozzo J, Conway CM. 2014. Neurocognitive mechanisms of statistical-sequential learning: what do event-related potentials tell us? *Front Hum Neurosci* **8**: 437.
- Delorme A, Makeig S. 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* **134**: 9–21.
- Dienes Z, Berry D. 1997. Implicit learning: below the subjective threshold. *Psychon Bull Rev* **4**: 3–23.
- Duncan-Johnson CC, Donchin E. 1977. Effects of a priori and sequential probability of stimuli on event-related potential. *Psychophysiology* **14**: 95.
- Duncan-Johnson CC, Donchin E. 1982. The P300 component of the event-related brain potential as an index of information processing. *Biol Psychol* **14**: 1–52.
- Durrant SJ, Taylor C, Cairney S, Lewis PA. 2011. Sleep-dependent consolidation of statistical learning. *Neuropsychologia* **49**: 1322–1331.
- Durrant SJ, Cairney SA, Lewis PA. 2013. Overnight consolidation aids the transfer of statistical knowledge from the medial temporal lobe to the striatum. *Cereb Cortex* **23**: 2467–2478.
- Eimer M, Goschke T, Schlaghecken F, 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.
- Ferdinand NK, Ringer D, Frensch PA, Mecklinger A. 2010. Event-related potential correlates of declarative and non-declarative sequence knowledge. *Neuropsychologia* **48**: 2665–2674.
- Fine A, Jaeger TF, Farmer TA, Qian T. 2013. Rapid expectation adaptation during syntactic comprehension. *PLoS One* **8**: e77661.
- Fiser J, Aslin RN. 2001. Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychol Sci* **12**: 499–504.
- Fiser J, Aslin RN. 2002. Statistical learning of higher-order temporal structure from visual shape sequences. *J Exp Psychol Learn Mem Cogn* **28**: 458–467.
- Franco A, Cleeremans A, Destrebecqz A. 2011. Statistical learning of two artificial languages presented successively: how conscious? *Front Psycho* **2**: 229.
- Franco A, Eberlein J, Destrebecqz A, Cleeremans A, Bertels J. 2015. Rapid Serial Auditory Presentation Segmentation: a new measure of statistical learning in speech segmentation. *Exp Psychol*. doi: 10.1027/1618-3169/a000295.
- Frensch PA, Miner CS. 1994. Effects of presentation rate and individual differences in short-term memory capacity on an indirect measure of serial learning. *Mem Cognit* **22**: 96–110.
- Friedman D, Johnson R Jr. 2000. Event-related potential (ERP) studies of memory encoding and retrieval: a selective review. *Microsc Res Tech* **51**: 6–28.
- Idemaru K, Holt LL. 2011. Word recognition reflects dimension-based statistical learning. *J Exp Psychol Hum Percept Perform* **37**: 1939–1956.
- Jacoby L. 1991. A process dissociation framework: automatic from intentional uses of memory. *J Mem Lang* **30**: 513–541.
- Johnson R, Donchin E. 1982. Sequential expectancies and decision making in a changing environment: an electrophysiological approach. *Psychophysiology* **19**: 183–200.
- Kalcounis-Rueppell MC, Metheny JD, Vohnhof MJ. 2006. Production of ultrasonic vocalizations by *Peromyscus* mice in the wild. *Front Zool* **3**: 3.
- Kim R, Seitz A, Feenstra H, Shams L. 2009. Testing assumptions of statistical learning: is it long-term and implicit? *Neurosci Lett* **461**: 145–149.
- Lakhani B, Van Ooteghem K, Miyasike-daSilva V, Akram S, Mansfield A, McIlroy WE. 2011. Does the movement matter? Determinants of the latency of temporally urgent motor reactions. *Brain Res* **1416**: 35–43.
- Lee IA, Preacher KJ. 2013. Calculation for the test of the difference between two dependent correlations with one variable in common [Computer software]. <http://quantpsy.org>.
- Lipkind D, Tchernichovski O. 2011. Quantification of developmental birdsong learning from the subsyllabic scale to cultural evolution. *Proc Natl Acad Sci* **108**: 15572–15579.
- Luka BJ, Barsalou LW. 2005. Structural facilitation: mere exposure effects for grammatical acceptability as evidence for syntactic priming in comprehension. *J Mem Lang* **52**: 436–459.
- Luka BJ, Choi H. 2012. Dynamic grammar in adults: incidental learning of natural syntactic structures extends over 48 h. *J Mem Lang* **66**: 345–360.
- McNealy K, Mazziotta JC, Dapretto M. 2006. Cracking the language code: neural mechanisms underlying speech parsing. *J Neurosci* **26**: 7629–7639.
- Paller KA, Kutas M. 1992. Brain potentials during memory retrieval provide neurophysiological support for the distinction between conscious recollection and priming. *J Cogn Neurosci* **4**: 375–392.
- Paller KA, Hutson CA, Miller BB, Boehm SG. 2003. Neural manifestations of memory with and without awareness. *Neuron* **38**: 507–516.
- Poldrack RA, Clark J, Paré-Blagoev EJ, Shohamy D, Creso Moyano J, Myers C, Gluck MA. 2001. Interactive memory systems in the human brain. *Nature* **414**: 546–550.
- Polich J. 2007. Updating P300: an integrative theory of P3a and P3b. *Clin Neurophysiol* **118**: 2128–2148.
- Reber A. 1967. Implicit learning of artificial grammars. *J Verb Learn Verb Behav* **6**: 855–863.
- Reber PJ, Squire LR. 1998. Encapsulation of implicit and explicit memory in sequence learning. *J Cogn Neurosci* **10**: 248–263.
- Reber PJ, Gitelman DR, Parrish TB, Mesulam MM. 2003. Dissociating explicit and implicit category knowledge with fMRI. *J Cogn Neurosci* **15**: 574–583.
- Rugg MD, Curran T. 2007. Event-related potentials and recognition memory. *Trends Cogn Sci* **11**: 251–257.
- Rugg MD, Mark RE, Walla P, Schloerscheidt AM, Birch CS, Allan K. 1998. Dissociation of the neural correlates of implicit and explicit memory. *Nature* **392**: 595–598.
- Rüsseler J, 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.
- Rüsseler J, Hennighausen E, Münte TF, 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.
- Rüsseler J, Kuhlicke D, Münte TF. 2003b. Human error monitoring during implicit and explicit learning of a sensorimotor sequence. *Neurosci Res* **47**: 233–240.
- Saffran JR, Aslin RN, Newport EL. 1996a. Statistical learning by 8-month-old infants. *Science* **274**: 1926–1928.
- Saffran JR, Newport EL, Aslin RN. 1996b. Word segmentation: the role of distributional cues. *J Mem Lang* **35**: 606–621.
- Saffran J, Newport E, Aslin R, Tunick R, Barrueco S. 1997. Incidental language learning: listening (and learning) out of the corner of your ear. *Psychol Sci* **8**: 101–105.
- Saffran JR, Johnson EK, Aslin RN, Newport EL. 1999. Statistical learning of tone sequences by human infants and adults. *Cognition* **70**: 27–52.
- Sanchez DJ, Reber PJ. 2013. Explicit pre-training instruction does not improve implicit perceptual-motor sequence learning. *Cognition* **126**: 341–351.
- Sanders LD, Newport EL, Neville HJ. 2002. Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. *Nat Neurosci* **5**: 700–703.
- Schlaghecken F, Stürmer B, Eimer M. 2000. Chunking processes in the learning of event sequences: electrophysiological indicators. *Mem Cognit* **28**: 821–831.
- Schneider W, Shiffrin RM. 1977. Controlled and automatic human information processing: I. Detection, search, and attention. *Psychol Rev* **84**: 1–66.
- Shiffrin RM, Schneider W. 1977. Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychol Rev* **84**: 127–190.

- Song S, Howard JH JR, Howard DV. 2007. Implicit probabilistic sequence learning is independent of explicit awareness. *Learn Mem* **14**: 167–176.
- Squires KC, Wickens C, Squires NK, Donchin E. 1976. The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science* **193**: 1142–1146.
- Steiger JH. 1980. Tests for comparing elements of a correlation matrix. *Psychol Bull* **87**: 245–251.
- Thompson SP, Newport EL. 2007. Statistical learning of syntax: the role of transitional probability. *Lang Learn Dev* **3**: 1–42.
- Toro JM, Sinnett S, Soto-Faraco S. 2005. Speech segmentation by statistical learning depends on attention. *Cognition* **97**: B25–B34.
- Turk-Browne NB, Jungé J, Scholl BJ. 2005. The automaticity of visual statistical learning. *J Exp Psychol Gen* **134**: 552–564.
- Turk-Browne NB, Scholl BJ, Chun MM, Johnson MK. 2009. Neural evidence of statistical learning: efficient detection of visual regularities without awareness. *J Cogn Neurosci* **21**: 1934–1945.
- Turk-Browne NB, Scholl BJ, Johnson MK, Chun MM. 2010. Implicit perceptual anticipation triggered by statistical learning. *J Neurosci* **30**: 11177–11187.
- Vilberg KL, Moosavi RF, Rugg MD. 2006. The relationship between electrophysiological correlates of recollection and amount of information retrieved. *Brain Res* **1122**: 161–170.
- Voss JL, Paller KA. 2008. Brain substrates of implicit and explicit memory: the importance of concurrently acquired neural signals of both memory types. *Neuropsychologia* **46**: 3021–3029.
- Westerberg CE, Miller BB, Reber PJ, Cohen NJ, Paller KA. 2011. Neural correlates of contextual cueing are modulated by explicit learning. *Neuropsychologia* **49**: 3439–3447.
- Willingham DB, Goedert-Eschmann K. 1999. The relation between implicit and explicit learning: evidence for parallel development. *Psychol Sci* **10**: 531–534.
- Willingham DB, Nissen MJ, Bullemer P. 1989. On the development of procedural knowledge. *J Exp Psychol Learn Mem Cogn* **15**: 1047–1060.
- Willingham DB, Salidis J, Gabrieli JD. 2002. Direct comparison of neural systems mediating conscious and unconscious skill learning. *J Neurophysiol* **88**: 1451–1460.
- Yu C. 2008. A statistical associative account of vocabulary growth in early word learning. *Lang Learn Dev* **4**: 32–62.
- Zajonc RB. 1968. Attitudinal effects of mere exposure. *J Pers Soc Psychol* **9**: 1–27.

Received December 19, 2014; accepted in revised form August 14, 2015.