

Anchors Aweigh: The Impact of Overlearning on Entrenchment Effects in Statistical Learning

Federica Bulgarelli and Daniel J. Weiss
Pennsylvania State University

Previous research has revealed that when learners encounter multiple artificial languages in succession only the first is learned, unless there are contextual cues correlating with the change in structure or if exposure to the second language is protracted. These experiments provided a fixed amount of exposure irrespective of when learning occurred. Here, the authors presented learners with 2 consecutive artificial languages testing learning after each minute of familiarization. In Experiment 1, learners received fixed input, and the authors replicated the primacy effect. In Experiment 2, learners advanced to the second language immediately following robust learning of the first language (thereby limiting additional exposure past the point of learning). Remarkably, learners tended to acquire and retain both languages, although contextual cues did not boost performance further. Notably, there was no correlation between performance on this task and a flanker task that measured inhibitory control. Overall, the findings suggest that anchoring effects in statistical learning may be because of overlearning. We speculate that learners may reduce their attention to the input once they achieve a low level of estimation uncertainty.

Keywords: statistical learning, primacy effect, overlearning, multiple representations, anchoring, entrenchment

Over the past 2 decades, the field of language acquisition has been profoundly influenced by the discovery that infant and adult learners are capable of tracking distributional information in speech and using it to infer discrete structures from continuous input (Saffran, Aslin, & Newport, 1996; Saffran, Newport, & Aslin, 1996). Subsequent to this discovery, statistical learning abilities have been found to contribute to many levels of language learning, spanning the very early stages of acquisition such as phonetic learning (e.g., Maye, Weiss, & Aslin, 2008; Maye, Werker, & Gerken, 2002) through later stages such as grammar learning (Gómez & Gerken, 2000; Reeder, Newport, & Aslin, 2013). Although these studies have documented an impressive array of learning abilities, the majority of research in this field has reduced the task of the learner to acquiring a single uniform input. That is, learners are presented with a familiarization period during which they can extract the same information at any point in time by sampling the input statistics. Although such demonstrations are important for understanding the types of statistics that learners are capable of tracking (see Aslin, 2014), it is natural to wonder how these learning abilities would fare when confronted with conditions affording greater variability in the input.

This issue holds particular significance for language learning, as real-world environments are often characterized by significant variability. The input to language learners often includes speaker changes, topic shifts, dialectal differences, and even the introduction of new languages, in the case of bilingual input (Gebhart, Aslin, & Newport, 2009; Gonzales, Gerken, & Gómez, 2015; Weiss, Gerfen, & Mitchel, 2009). Thus, a central challenge for statistical learning is to determine the underlying source of the variance in the input. Essentially, learners must infer how many causal structures are generating the observed statistics (see Qian, Jaeger, & Aslin, 2012). Much like Piaget and Cook's (1952) description of the processes of assimilation and accommodation, learners must decide when the variance is meaningful and signals a true change to the underlying structure (such as a new language being spoken) versus when to expand the current model to incorporate the new input (such as a change in speaker).

Consequently, a fundamental task confronting the language learner is to detect when a change has occurred to the underlying structure of the input. Studies of change detection, such as *n*-armed bandit tasks, suggest that learners are quite adept at detecting changes to probability structures (e.g., Behrens, Woolrich, Walton, & Rushworth, 2007). However, this type of study includes trial-by-trial feedback that provides learners with valuable information that is not always present in real-world situations. For example, learners may use prediction error (a subtraction of the actual outcome from the expected outcome) to infer when a shift in the underlying causal structure has occurred (Behrens et al., 2007; Qian et al., 2012; Rescorla & Wagner, 1972). However, the problem for statistical learning in the context of language is that this type of immediate feedback is typically unavailable.

Insights into this issue have come from several studies of artificial language learning. For example, Weiss, Gerfen, and Mitchel (2009) presented participants with repeated alternations of two artificial

This article was published Online First March 7, 2016.

Federica Bulgarelli and Daniel J. Weiss, Department of Psychology and Program in Linguistics, Pennsylvania State University.

We thank Cara Theoret, Maura Petrusky, Lindsay Rauch, Sarah Miller, Nick Andereeg, Paris Atabek, and Nicolette Khosa for their assistance in conducting experiments. This research was supported by National Institutes of Health Grant NIH R01 HD067250 and a National Science Foundation Graduate Research Fellowship.

Correspondence concerning this article should be addressed to Federica Bulgarelli, Department of Psychology and Program in Linguistics, Pennsylvania State University, 107 Moore Building, University Park, PA 16802. E-mail: fub113@psu.edu

languages that were interleaved for several successive iterations. Each block of language lasted 2 min and could be parsed only if learners tracked the transitional probabilities between words. In the congruent condition, in which the language elements could combine across streams without interference (and thus learners could collapse across streams without penalty), learners were capable of acquiring both streams even in the absence of a contextual cue corresponding to the change. In the incongruent condition in which the language structures overlapped in a manner that interfered with the underlying statistics of both languages, learners required a contextual cue (i.e., a change in speaker) in order to learn both languages. Without the contextual cue, learners acquired neither structure. A subsequent study demonstrated that learning of both languages can occur even if the context change occurs in another modality (Mitchel & Weiss, 2010).

Notably, learning in the absence of contextual cues may result in a primacy effect in which the first encountered structure is learned but the second is not. Gebhart, Aslin, and Newport (2009) presented adult participants with a speech segmentation task in which the underlying structure changed midstream (i.e., only a single switch from one structure to the next). The first 5 min of the speech stream contained a set of words from one artificial language [Language 1 (L1)], whereas the second 5 min contained a new set of words [Language 2 (L2)]. In the absence of an overt contextual cue to signal to the learner that the underlying structure had changed, learners were able to acquire L1 but not L2. However, when learners were cued to the presence of the second structure by means of explicit instruction along with a 30-s pause, they were able to learn both languages at above-chance levels (see also Franco & Destrebecqz, 2012). Likewise, if L2 was tripled in length (15 min instead of 5 min), learners also acquired both languages. Similarly, in visual cuing tasks, it has been suggested that initial hypotheses about the statistical regularities in the environment may block learners from acquiring new regularities (Jungé, Scholl, & Chun, 2007). In sum, the emerging picture is that contextual cues may facilitate the detection of change in statistical learning and formation of multiple representations in a shorter timeframe than might otherwise be necessary (see Weiss, Poepsel, & Gerfen, 2015, for further discussion).

In the absence of any overt cues to change, however, learners often acquire the first language better than the second, which may not be learned at all. Gebhart et al. (2009) proposed that this primacy effect may arise because of a stationarity bias (Aslin, 2014). If learners approach the task with prior expectations that inputs do not undergo a rapid change in the absence of a strong contextual cue, they may underadjust initial estimates when confronted with new information, as evidenced in the results of judgment and contingency learning experiments (such as Dennis & Ahn, 2001; Yates & Curley, 1986). For example, Marsh and Ahn (2006) have argued that primacy effects tend to occur when learners have already formed a hypothesis about the relationship between events and underadjust even as they gain additional evidence.

In statistical learning tasks, this tendency toward underadjustment may be related to the level of estimation uncertainty achieved by learners in accounting for the observed statistical regularities (see Qian, Jaeger, & Aslin, 2012). When learners lack confidence in their ability to accurately represent the causal model generating the statistics of their environment, they are less likely to posit a

change in underlying structure when encountering variance in the signal. In that vein, Qian et al. (2012) suggested that the results of Gebhart et al. (2009) might be attributable to the fact that learners never achieved a low level of estimation uncertainty in the first language prior to switching to the second.

An alternative account, explored here, is that learners actually acquire the L1 relatively quickly, rapidly reaching a low level of estimation uncertainty and then receiving additional exposure to the same set of statistics after learning. Accordingly, this additional exposure after learning may reinforce the initial stationarity bias, making changes to the input harder to detect. This would accord with the observation that when contextual cues highlight the change of stream, the second stream can be acquired (Gebhart et al., 2009). Thus, the failure to acquire the L2 may be directly related to the initial length of exposure to L1. This hypothesis finds some support in the literature on overlearning (also known as overtraining), which can be defined as continued training on a set of contingencies after a threshold of learning has already been met (Ebbinghaus, 1885/1964). Under some conditions, particularly when the initial input is easily acquired, reversals may become harder to detect when participants are overtrained (see Mackintosh, 1969). To adjudicate between these two theoretical accounts, it is necessary to determine how quickly the statistical information is being acquired and whether learning remains consistent throughout familiarization.

Consequently, in this study we tested whether the length of exposure to L1 impacts learning of the L2, by presenting learners with the same two languages used by Gebhart et al. (2009). However, rather than providing two fixed 5-min blocks followed by a test, we presented learners with 1-min blocks of the language and tested after each minute. In Experiment 1, we mirrored the amount of exposure used by Gebhart et al. (2009) by providing learners with five 1-min blocks of each language followed by tests. This afforded a test of whether we could replicate the primacy effect observed by Gebhart et al., using our modified methodology. In Experiment 2, we advanced learners to the L2 immediately after they had achieved a high level of performance on any intermediate test of L1 (and consequently exposure to L1 was shorter for most participants) to determine whether performance is affected by additional exposure to the same language subsequent to learning (i.e., after reaching a low level of uncertainty regarding the first structure). In Experiment 3, we modified the paradigm used in Experiment 2 to include contextual cues that signaled a change in underlying structure. As noted above, in prior studies, contextual cues have facilitated the learning of multiple languages (e.g., Franco & Destrebecqz, 2012; Gebhart et al., 2009; Mitchel & Weiss, 2010; Weiss et al., 2009). Here, we wanted to test whether learners would still benefit from these cues if they were advanced to the new language immediately after reaching a low level of estimation uncertainty.

In addition, we sought to test whether differences in inhibitory control abilities might predict participants' ability to overcome the primacy effect in this paradigm. It has long been thought that inhibitory processes might facilitate the learning of new inputs by suppressing the effects of prior learning (e.g., Atwater, 1953; Underwood, 1949). Further, research in our lab has discovered that in statistical learning tasks involving colliding cues (i.e., multiple incongruent cues to word boundaries), individual differences in inhibitory control correlate with successful learning (see Weiss,

Gerfen, & Mitchel, 2010). Because the languages used in the original study (Gebhart et al., 2009) have a 50% overlap in syllable inventories, we predicted that participants might benefit from inhibitory control abilities by suppressing the first set of learned statistics to acquire the second set. Consequently, we predicted that learners who were able to acquire both L1 and L2 might exhibit a smaller Flanker effect (i.e., greater inhibitory control abilities) relative to those who learned only L1.

Experiment 1

In this experiment, we sought to replicate and extend the findings of Gebhart, Aslin, and Newport (2009) using a modified presentation method in which participants received an eight-item two alternative forced-choice (2AFC) test following each minute of language presentation. Learners encountered a total of 5 min of an L1 followed by 5 min of an L2 (each presented in 1-min blocks). This paradigm allowed not only for information regarding whether the first and second languages were learned but also afforded an estimate of how quickly they were acquired. In addition, participants received a flanker task to determine whether the ability to acquire both L1 and L2 might be related to the ability to inhibit the statistics learned during L1 presentation.

Method

Participants. Participants consisted of 69 Penn State undergraduate students (mean age = 18.75 years, $SD = 1.08$; 51 females and 18 males) recruited from an introduction to psychology course. All participants received course credit for their participation in the study. An additional 10 participants were recruited but were excluded due to not following instructions.

Stimuli. The two artificial languages presented in this experiment were drawn from previous studies (Gebhart et al., 2009; Newport & Aslin, 2004). Each language consisted of 16 trisyllabic words constructed from consonant–vowel (CV) syllables. The languages were created using Speaker, a text-to-speech application found in the speech synthesizer, MacInTalk. The synthesizer was adjusted to remove any acoustic cues to the word boundaries and to produce equivalent levels of coarticulation across all syllables. Both speech streams were created using the same female voice (Victoria). Syllable duration was edited to fall between .20 and .22 s in length, thereby assuring that duration could not cue learners to the position of the syllable within the word (see Newport & Aslin, 2004, for further details).

The languages were assembled from an inventory consisting of six consonants (d, k, b, p, g, and t) and six vowels (a, ae, e, i, o, u). Each language was composed of 16 possible words. The words were structured such that the vowels formed a consistent frame in one language, while the consonants formed a consistent frame in the other. There were two consonants and two vowels that could occur for each position (see Table 1 for the language structure and the subset of words and part words that were used at test). The words were concatenated to create a continuous speech stream with each word repeated the same number of times and with the constraint that no word was ever reduplicated. The speech stream was looped to create a 67-s stream, with a production rate of 284 syllables per minute. The transitional probability of the vowels within a word was 1.0, as each vowel was always followed by

Table 1
Artificial Language Structure and Test Items

| Language | Frames | Fillers | Test words | Test partwords |
|----------|-----------------------------|---------------|------------|----------------|
| A | <u>a</u> <u>u</u> <u>e</u> | [d][k][b] | dakube | kubepo |
| | <u>o</u> <u>i</u> <u>ae</u> | [p][g][t] | pagute | tedoki |
| | | | dokibae | gitaedo |
| B | b_p_g | [_ae][_a][_u] | pogitae | baepagu |
| | t_d_k_ | [_e][_o][_i] | bepogi | dokibae |
| | | | tedoki | dakutae |
| | | | baepagu | gubepo |
| | | taedaku | gibaepa | |

exactly one other vowel. Across the words, the vowel–vowel transitional probability was .5, as words composed from either frame could follow any word.

During test, participants received a two alternative forced-choice task (2AFC), with each trial consisting of a statistically defined word paired with a part word that was assembled by concatenating either the last syllable of one word followed by the first two syllables of a different word (i.e., 3–1–2), or the last two syllables of one word followed by the first syllable of a different word (i.e., 2–3–1). The test stimuli were created in isolation with a falling intonation imposed at the end of every token. Four words and part words were used at test. For the eight-item tests, each word and part word (see Table 1) occurred twice with a 1-s ISI between items and no pairings repeated. The order of the words and part words was counterbalanced.

Procedure. Participants were seated in a sound-attenuated chamber and instructed to attend to a recording played through headphones. They were informed that they would subsequently be asked questions about what they heard but did not receive any further details regarding the experimental design.

Participants listened to one 67-s block of the first language (L1). The ordering of the languages was counterbalanced across participants. After this familiarization period, participants received an eight-item 2AFC test and were instructed to choose which of the two tokens (i.e., the word or part word) corresponded to the familiarization language. They were then presented with another 1-min block of familiarization followed by another test. The cycle of familiarization and test repeated five times. After the fifth test participants received exposure to the second language (L2). The presentation of L2 was identical to L1 (i.e., five blocks of familiarization test).

Each participant also completed a flanker task (Eriksen & Eriksen, 1974) that was modified according to Bunge, Dudukovic, Thomason, Vaidya, and Gabrieli (2002; see also Emmorey, Luk, Pyers, & Bialystok, 2008). The stimuli consisted of red arrows flanked by four black arrows. Participants were asked to respond to the red arrow by clicking the left or right mouse button depending on the direction the arrow was pointing. There were three types of blocks during the experiment: control, conflict, and go/no-go. During the control blocks, participants saw a single arrow and were asked to respond with the direction of that arrow. During the conflict block, participants saw congruent and incongruent flanking arrows. During the go trials, the participants saw congruent or incongruent trials as in the second block. During the no-go trials, participants viewed the red arrow flanked by black Xs, which indicated that they should not respond to the stimuli.

Each trial began with a fixation cross in the middle of the screen for 250 ms. Participants then viewed the target stimuli for 2 s or until they produced a response. Before each block, participants received 12 practice trials. Each of the three block types was presented twice for a total of 84 trials per block type. In addition to these blocks, there were also two mixed blocks during which participants received intermixed congruent–incongruent and go/no-go trials. Both response times and accuracy were recorded.

Results

Across all blocks of the experiment, the average accuracy for L1 was 6.12 ($SD = 1.2$) out of 8 on the tests following familiarization. For L2, the average was 5 ($SD = 0.88$). These scores were significantly different: paired t test, $t(68) = 7.53$, $p < .001$, $d = 1.08$. Both were significantly above chance: L1, $t(68) = 14.83$, $p < .001$, $d = 3.59$; L2, $t(68) = 9.417$, $p < .001$, $d = 2.28$. Performance on the terminal test of each language (i.e., the eight-item test following the fifth familiarization period) was 5.87 ($SD = 1.62$) for L1, which was significantly above chance, $t(68) = 9.60$, $p < .001$, $d = 2.33$, and 4.87 ($SD = 1.58$) for L2, which was also significantly above chance, $t(68) = 4.57$, $p < .001$, $d = 1.11$. Terminal accuracy for L1 was significantly higher relative to L2, $t(68) = 3.65$, $p = .001$, $d = 0.62$.

A repeated-measures analysis of variance (ANOVA) testing the effect of language and block showed a main effect of language, $F(1, 67) = 57.16$, $p < .001$, $\eta_p^2 = .46$, such that L1 accuracy was significantly higher than L2 accuracy. When looking at performance after each block of familiarization, there was no significant linear, $F(1, 67) = 1.14$, $p = .289$, $\eta_p^2 = .017$, nor quadratic trend, $F(1, 67) = 2.00$, $p = .17$, $\eta_p^2 = .029$, suggesting that learning did not improve in a systematic way across the session for each language (see Figure 1). The interaction between language and block was also not significant, $F(1, 67) = 0.45$, $p = .51$, $\eta_p^2 = .007$.

Given our interest in determining whether learners achieve a low level of estimation uncertainty, we sought to determine a threshold of performance that would be indicative of robust learning (particularly for L1, because we were interested in exploring the primacy effect). Therefore, we tracked learners' performance after reaching their peak accuracy for L1 (for participants who

scored above chance). Participants were split into two groups depending on their peak accuracy, 16 participants reached a peak accuracy of 5 or 6 out of 8 (3 of them reaching a peak accuracy of 5, 13 reached a peak accuracy of 6), whereas 52 participants reached a peak accuracy of 7 or 8 out of 8 (27 of them reaching a peak accuracy of 7 and 25 reached 8). One participant in each group reached this peak on the last block, and thus these participants were not included in the following analyses. The average performance on tests after peak performance for the participants who scored a maximum of 5 or 6 out of 8 was 4.77 ($SD = 0.76$), whereas the group with peak accuracy of 7 or 8 out of 8 subsequently averaged 6.51 ($SD = 1.12$). The difference between these groups was significant, $t(64) = 5.53$, $p < .001$, $d = 1.81$. Moreover, after reaching a peak accuracy of 5 or 6 out of 8, 66% of participants (10 out of 15) performed at chance on a subsequent test. By contrast, only 21% ($n = 11$ out of 52) of participants who reached a peak accuracy of 7 or 8 out of 8 performed at chance on a subsequent test.

We also compared 6- and 7-peaking participants to ensure that this accurately represented the boundary for consistent learning. The average performance on subsequent tests for participants reaching a peak accuracy of 6 was 4.80 ($SD = 0.98$), which was significantly lower than those reaching a peak of 7 (6.33, $SD = 1.62$); $t(36) = 3.91$, $p < .001$, $d = 1.14$. Eight out of 12 six-peaking participants performed at chance on a subsequent test of L1, whereas only 5 out of 26 seven-peaking participants regressed to chance. In sum, reaching a minimum of 7 out of 8 on an interim test appeared to be the threshold for robust learning, as subsequent test performance for participants reaching this level was significantly higher relative to those who reached a peak of 6 or below.

As noted above, 52 out of 69 participants scored 7 or 8 on the first language, and on average, it took them 1.56 ($SD = 0.94$) blocks to reach this level of performance. Of these learners, only 20 subsequently reached threshold on L2 (38%). Participants who only reached this criterion on L1 did so after an average of 1.40 blocks ($SD = 0.76$) of exposure. Participants who reached criterion on both languages attained it significantly faster for L1 (1.80 blocks, $SD = 1.15$) relative to L2 (2.7 blocks, $SD = 1.38$); paired $t(19) = -2.35$, $p = .03$, $d = 0.71$. Overall, there was no significant difference in the number of blocks it took to reach criterion on L1 for the participants who scored 7 or above on both languages relative to those who reached this criterion only for L1, $t(29) = -1.36$, $p = .185$; note Levene's test indicated unequal variances ($F = 6.41$, $p = .015$, so degrees of freedom were adjusted from 50 to 29).

The flanker effect was calculated by subtracting RTs for accurate congruent trials from the RTs for the accurate incongruent trials during the mixed blocks of the task. Participants whose flanker effect was more or less than 2.5 SD s from the mean were excluded from this set of analyses ($n = 4$). Given our interest in determining the relationship between the flanker effect and the primacy effect, we first tested for a correlation between these two tests. However, there was no significant correlation between the flanker effect and participants' highest score on a test of L2, $r(65) = .11$, $p = .41$. Likewise, there was no significant difference between participants who reached criterion on both languages (93.18 ms, $SD = 40.17$ ms) and those who only did so for L1 (93.01 ms, $SD = 48.50$ ms); $t(47) = 0.013$, $p = .99$. Further, participants who did not reach criterion on L1 exhibited a flanker effect of 78.91 ms ($SD = 46.63$ ms), which was not significantly different when compared with participants who reached crite-

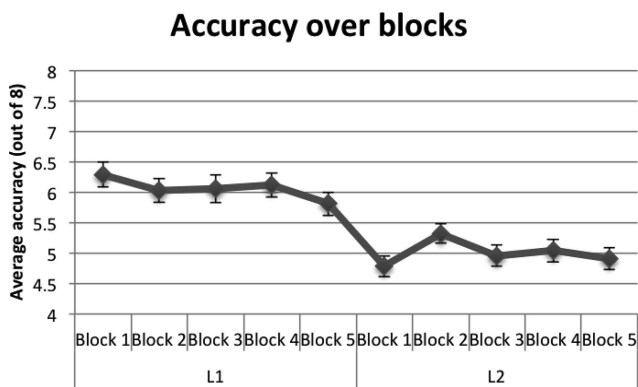


Figure 1. Average accuracy on eight-item intermittent tests over the course of exposure for Language 1 (L1) and Language 2 (L2).

tion on both, $t(32) = .86$, $p = .40$, or only on L1, $t(45) = 1.09$, $p = .28$.

Discussion

Overall, the findings from Experiment 1 were largely consistent with the primacy effect reported in Gebhart et al. (2009). When accuracy was calculated for each language by collapsing across all tests, it was significantly higher for L1 relative to L2. Likewise, terminal accuracy was also significantly higher at the end of the L1 exposure relative to L2 exposure. Despite the changes to the experimental procedure, the results conceptually replicate the findings reported in the original study. Unlike the previous study, however, both L1 and L2 were learned above chance levels when averaging across the intermittent tests (though we note our number of participants was larger than the original study, and this likely accounts for the discrepancy, see also Zinszer & Weiss, 2013). It is important to note that despite testing after each minute of exposure, participants did not appear to learn the structure of the languages from the intermittent tests, as there were no significant positive trends in performance.

Despite our prediction that the flanker task, a measure of inhibitory control, might correlate with performance on the statistical learning task, we found no evidence of this trend. Thus, we found no evidence to suggest that the decrement in performance in L2 relative to L1 is a product of individual differences in inhibitory control. Rather, the primacy effect may arise due to other factors, such as overlearning, or perhaps a decrement in attention (see General Discussion).

Given our interest in determining the ability to acquire L2 after reaching a low level of estimation uncertainty, we sought to find a criterion of performance on the intermittent tests that would be indicative of robust (sustained) learning. We set this criterion to scoring 7 or 8 on any of the five intermittent tests of L1, given that 79% of participants who achieved this mark continued to perform at above-chance levels on all subsequent tests. By contrast, 66% of the participants whose highest accuracy reached only 5 or 6 out of 8 did not maintain above-chance performance throughout the course of familiarization.

Applying this criterion to Experiment 1, 75% of participants demonstrated strong learning of L1. Of the 75%, only 38% reached the same criterion for L2, again suggesting a significant primacy effect. Because most participants reached this threshold of learning for L1 within the first 2 min of exposure, they subsequently received lengthy exposure to the same set of statistics before being advanced to the L2. Consequently, in Experiment 2, we investigated whether this additional exposure after learning negatively impacted the ability of learners to acquire L2.

Experiment 2

In Experiment 2, we modified the design of Experiment 1 to allow participants to advance to the L2 as soon as they reached the criterion of 7 or above out of 8 on the test of L1 that followed each block. Participants received up to five 1-min blocks of L1 followed by up to five blocks of L2. The goal of this experiment was to determine whether allowing participants to advance as soon as they exhibited strong learning of L1 would reduce the primacy effect. We were further interested in determining whether they would maintain their L1 learning despite (potentially) shorter exposure.

Method

Participants. Participants consisted of 72 Penn State undergraduate students (58 females, 14 males; mean age = 18.9; $SD = 1.01$) from an introduction to psychology course who participated in this experiment for credit. None had previously participated in a statistical learning experiment. An additional 8 participants were recruited but were excluded from analysis because of experimenter error ($n = 7$) or sleeping during the study ($n = 1$).

Stimuli. The stimuli were the same as those used in Experiment 1.

Procedure. There was one primary difference between Experiments 1 and 2. In Experiment 2, participants received exposure to L2 immediately after they scored a 7 or above out of 8 on one of the tests following the 1-min familiarization periods of L1. Participants received up to five familiarization-test blocks for the first language, and if they did not reach our criteria of 7 out of 8 correct during those five blocks, the experiment was terminated. If participants scored 7 out of 8, they then received up to five familiarization-test blocks of L2. Because participants could advance to the L2 after as little as 1 min of exposure, we wanted to test whether they maintained their learning after exposure to the L2. Thus, after completing the L2 exposure and tests, participants completed a 16-item posttest on the languages they learned (if they learned both, they received posttests on both languages; if they only learned the first language, they received a posttest on just the first language), followed by the flanker task. The 16-item posttest was created in the same way as the eight-item tests, but each word was exhaustively paired with each part word, with no pairings repeated.

Results

Twenty-four of the 71 participants did not reach the criterion of seven out of eight correct on L1, and thus their data were not included in further analyses. Of the 47 participants who reached criterion on L1, 70% ($n = 33$) also reached this criterion for L2. This finding represented a significant increase in the proportion of learners who acquired both languages relative to Experiment 1, in which only 38% reached criteria for both languages (Barnard's exact test = 3.16, $p < .001$).

For learners who reached criterion on both languages, they were marginally faster to learn L1 (1.85 blocks, $SD = 1.2$) than they were to learn L2 (2.58 blocks, $SD = 1.23$); paired t test, $t(32) = -1.99$, $p = .055$. Learners who only reached threshold for L1 did so in 2.43 blocks ($SD = 1.6$), a difference that was not significantly slower than those who acquired both L1 and L2 [the Levene's test for equality of variances was violated, $F = 4.36$, $p < .05$, thus, an independent samples t test not assuming homogeneity of variance was computed, $t(19) = 1.22$, $p = .24$].

The posttest data was analyzed to determine whether participants were able to retain what they had learned. For learners who reached criterion on both languages, the average accuracy for the L1 posttest was 9.60 ($SD = 3.40$) out of 16, which is significantly above chance, $t(32) = 2.72$, $p = .01$, $d = 0.96$, and 11.37 ($SD = 2.54$) for L2, which is also significantly above chance, $t(32) = 7.59$, $p < .001$, $d = 2.68$. There was a significant difference between the posttests for the different languages, such that the second language was retained significantly better than the first language: paired t test, $t(32) = -2.106$, $p = .043$, $d = 0.59$.

Participants whose flanker effect was more or less than 2.5 *SDs* from the mean were excluded from flanker analyses ($n = 2$). Participants' flanker effect was not significantly correlated with their highest score on L2, $r(47) = .028$, $p = .85$. Further, participants who reached threshold on both languages exhibited a flanker effect of 94.37 ms ($SD = 55.35$ ms), whereas L1-only participants exhibited a flanker effect of 96.71 ms ($SD = 88.02$ ms), which was not a significant difference, $t(45) = 0.11$, $p = .91$. Likewise, participants who did not reach criterion on L1 exhibited a flanker effect of 86.23 ms ($SD = 64.37$ ms), which was not significantly different from participants who reached threshold on both languages, $t(54) = -0.51$, $p = .61$, nor participants who only did so on L1, $t(35) = -0.42$, $p = .68$.

Discussion

In Experiment 2, participants advanced to L2 as soon as they reached criterion on one of the intermittent tests of L1. We found that 70% of participants who reached criterion on L1 also reached it for L2. This represents a significant increase in the learning of L2 relative to Experiment 1 (see Figure 2) and suggests a drastic reduction in the primacy effect as measured by overall performance. Thus, the amount of additional exposure after robust learning has occurred appears to be an important factor contributing to entrenchment effects in statistical learning (see General Discussion).

Note that strong learners of both languages did acquire the L1 marginally faster than the L2, suggesting a primacy effect in the time course of learning. The lengthier time course for L2 might be due to a delay in noticing the change into L2 or perhaps in structuring the second representation. We also note that despite most participants reaching criterion on L1 in <2 min, they were able to retain above-chance learning of that language at the post-test. Nonetheless, the posttest revealed a recency effect in retention, as L2 posttest scores were significantly higher than L1. It is possible this arose because of the lengthier average exposure to L2 relative to L1 or perhaps just the proximity between exposure and test (with less intervening memory demands relative to L1). Notwithstanding, our results suggest that learners did not delete the first statistical representation to acquire the second. Rather, learners in Experiment 2 were not only able to detect

and acquire the second structure but were also able to maintain both sets of statistical representations at above-chance levels.

The lack of any significant correlation between learning and the flanker task (here and in Experiment 1) further suggest that the ability to overcome the primacy effect may not be related to individual differences in inhibitory control. Future work might profit from using other tasks that measure different components of inhibitory control, such as the AX Type Visual Continuous Performance Task (Morales, Gómez-Ariza, & Bajo, 2013) to ensure that our null findings reflect truly uncorrelated constructs.

Although advancing participants to L2 as soon as they reached threshold on L1 significantly reduced the number of participants who demonstrated a primacy effect, 30% of L1 learners still failed to reach criterion on the L2. Given that previous research has found that the inclusion of contextual cues correlating with the change in structure benefits learning (e.g., Gebhart et al., 2009; Mitchel & Weiss, 2010; Weiss et al., 2009), in Experiment 3 we explored whether cuing learners to the change would further increase the percentage of L2 learners.

Experiment 3

Contextual cues are known to provide very accessible prompts for facilitating the detection of multiple structures (e.g., Alloy & Tabachnik, 1984). In several studies of statistical learning, the addition of contextual cues has been shown to allow learners to form multiple statistical representations, even for inputs that are otherwise statistically incompatible (Gebhart et al., 2009; Mitchel & Weiss, 2010; Weiss et al., 2009). Thus, Experiment 3 used the same method as Experiment 2 but added contextual cues to mark the switch into L2. Three cues were tested: a pitch-shift cue, an explicit cue, and a combination of these cues along with an additional insertion of a pause between streams.

Method

Participants. Two hundred and five Penn State undergraduate students from an introduction to psychology course participated in this experiment for credit. Seventy-nine participants (63 female; mean age = 18.81; $SD = 1.26$) were assigned to Condition 1, 63 participants (52 female; mean age = 18.55; $SD = 0.96$) were assigned to Condition 2, and 63 participants (43 females; mean age = 19.65; $SD = 2.86$) were assigned to Condition 3. None had previously participated in a statistical learning experiment. Across all conditions, an additional 16 participants were excluded because of not following experimenter instructions.

Stimuli. The stimuli for these conditions were identical to those used in Experiment 2, except for the following changes: In Conditions 1 and 3, one of the languages was pitch shifted down by 60 Hz¹ (using Audacity) to resemble a male speaker, whereas the other language remained in the original female voice.

Procedure. The procedure for the three conditions in Experiment 2 was roughly identical to Experiment 2. However, in Condition 1 (pitch-shift condition), one of the presented lan-

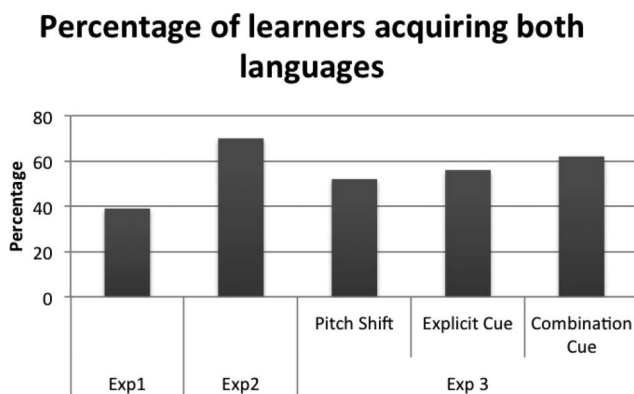


Figure 2. Percentage of participants who learned Language 2 (L2) after learning Language 1 (L1) in all experiments (Exp).

¹ The pitch-shifted language was tested in isolation and learned to the same accuracy level as the original language, $t(16) = 0.6356$, $p = .53$.

guages was pitch shifted. Thus, the order of the languages was counterbalanced, such that half the participants received the original voice first, whereas the remainder received the pitch-shifted voice first. In Condition 2 (explicit-cue condition), the instructions given to the participants changed, such that they were told that during the experiment they would be learning two languages. In Condition 3 (combined condition), the participants were explicitly told that they would be learning two languages, that the second language would start after a 30-s pause in the experiment, and that the second language would be presented by a novel speaker (the pitch-shifted language was used as in Condition 1). The order of the languages was again counterbalanced.

Results

Across all conditions, 42 participants did not reach criterion on L1. In the pitch-shift condition, 61 participants reached threshold on L1, and 51.6% of those also reached the threshold on L2 ($n = 31$), which trended toward being significantly lower than performance in Experiment 2 (Barnard's exact test = 1.69, $p = .07$). In the explicit-cue condition, 52 participants reached threshold on L1, and 55.7% of those also learned L2 ($n = 29$), which was not significantly different from Experiment 2 (Barnard's exact test = 1.07, $p = .18$). In the combination condition, 50 participants reached threshold on L1, of which 62% also reached the threshold on L2 ($n = 31$), which was not significantly different from Experiment 2 (Barnard's exact test = 0.37, $p = .41$; see Figure 2).

None of the contextual-cue conditions significantly differed from one another in the percentage of participants who reached threshold on L2 after reaching threshold on L1. The pitch-shift condition did not significantly differ from the explicit-cue condition (Barnard test exact statistics = 0.52, $p = .36$) nor the combined cue condition (Barnard test exact statistic = 1.18, $p = .14$). The explicit-cue condition and the combined cue condition were also not significantly different from each other (Barnard test exact statistic = 0.64, $p = .36$).

For learners who reached threshold on both languages, we compared how quickly each language was learned (see Table 2). In the pitch-shift condition, participants required 2.13 blocks ($SD = 1.06$) to reach criterion on L1 and 2.19 blocks ($SD = 1.30$) for L2, a difference that was not significant, $t(30) = -0.23$, $p = .82$. In the explicit-cue condition, participants who learned both languages were significantly faster at reaching criterion on the L1 ($M = 1.77$, $SD = 1.25$) relative to the L2 ($M = 2.60$, $SD = 1.57$); $t(29) = -2.05$, $p = .05$, $d = 0.58$. Similarly, in the combined cue

condition, participants were significantly faster at reaching criterion on L1 ($M = 1.52$, $SD = .68$) relative to L2 ($M = 2.42$, $SD = 1.46$); $t(30) = -2.80$, $p = .009$, $d = 0.79$. The length of time it took to reach criterion on each of the languages in Experiment 3 did not differ from Experiment 2 (all $ps > .05$).

The posttest data was analyzed to determine whether participants who had learned both languages were able to retain learning. For those that learned both languages, they retained both L1 and L2 above chance—L1: pitch-shift condition, $M = 10.32$, $SD = 2.40$, $t(30) = 5.39$, $p < .001$, $d = 1.97$; explicit-cue condition, $M = 9.77$, $SD = 3.25$, $t(29) = 2.98$, $p = .006$, $d = 1.11$; combination condition, $M = 10.13$, $SD = 2.71$, $t(30) = 4.36$, $p < .001$, $d = 1.59$; L2: pitch-shift condition, $M = 11.16$, $SD = 2.85$, $t(30) = 6.17$, $p < .001$, $d = 2.25$; explicit-cue condition, $M = 11.2$, $SD = 2.27$, $t(29) = 7.74$, $p < .001$, $d = 2.87$; combination condition, $M = 10.87$, $SD = 2.64$, $t(30) = 6.05$, $p < .001$, $d = 2.21$. Retention did not differ significantly between the languages across conditions (all $ps > .05$).

After excluding outliers ($n = 8$), the flanker effects were not significantly correlated with the highest L2 score for any of the conditions (all $ps > .05$). When comparing the flanker effect across learners, the only significant differences were found in the explicit-cue condition. In this condition, participants who did not reach criterion on L1 exhibited significantly smaller Flanker effects relative to those who reached criterion on both languages, $t(38) = -2.24$, $p = .031$, and those who only reached criterion on the first language, $t(39) = -2.07$, $p = .047$. The flanker effect in all other cases did not predict learning when contextual cues were added (all $ps > .05$).

Discussion

Although previous research revealed that adding contextual cues improved performance on L2 (Gebhart et al., 2009), in our modified experimental paradigm, the cues neither increased the percentage of learners who successfully reached threshold on L2 after robustly learning L1 nor reduced the time needed to acquire L2. This finding suggests that advancing learners to L2 immediately after learning may facilitate the acquisition of L2 in a manner that overlaps with the benefits accrued by adding contextual cues.

As in Experiments 1 and 2, performance on the flanker task did not meaningfully correlate with performance on the statistical learning task. Surprisingly, the few significant findings correlating these tasks indicated that participants who failed to learn the L1 had smaller flanker effects than learners of either one or both languages. These findings do not lend any support to our original

Table 2
Mean Time to Learn in Minute-Long Blocks in Experiment 3

| Language(s) learned | Cond. 1: Pitch shift | | Cond. 2: Explicit knowledge | | Cond. 3: Pitch shift + explicit knowledge + pause | |
|---------------------|----------------------|-------------|-----------------------------|-------------|---|-------------|
| | L1 | L2 | L1 | L2 | L1 | L2 |
| L1 only | 1.42 (1.62) | — | 1.67 (0.84) | — | 1.60 (1.12) | — |
| L1 + L2 | 2.16 (1.01) | 2.42 (1.46) | 1.769 (1.28) | 2.50 (1.56) | 1.63 (0.68) | 2.73 (1.42) |

Note. Cond. = Condition; L1 = Language 1; L2 = Language 2.

hypothesis that learners may inhibit the statistics of the L1 to acquire the L2.

General Discussion

Previous research in statistical learning has reported a primacy effect when learners are presented with two successive artificial languages unless the change in language is signaled by a corresponding contextual cue or exposure to the second language is tripled (Gebhart et al., 2009). In the present study, we sought to determine whether learners become anchored in L1 due to the amount of continued exposure after robust learning has already occurred (i.e., overlearning). In Experiment 1, we replicated and extended the findings of Gebhart et al. (2009) by presenting learners with five 1-min blocks of exposure to the same two artificial languages used in the original study, with each block followed by an eight-item test. We found that learning of L1 could occur even within the first minute of exposure and quite often within the first 2 min. Remarkably, almost two thirds of these L1 learners failed to successfully acquire L2. In Experiment 2, we advanced participants to L2 immediately after they had reached our learning criterion on one of the intermittent eight-item tests for L1 to assess whether advancing to L2 very soon after robust learning of L1 took place might attenuate the primacy effect. In contrast to Experiment 1, 70% of participants were able to acquire the L2, which represented a significant increase relative to that initial condition. In Experiment 3, we sought to determine whether performance on this task could further be enhanced through the inclusion of contextual cues corresponding to the change of structure (as in Gebhart et al., 2009). However, unlike previous studies of multistream statistical learning, we did not find evidence of any advantage for learning when contextual cues were used in conjunction with advancing participants immediately after learning.

The central finding of this study is that the primacy effect observed in statistical learning of two structures is due, at least in part, to learners receiving additional exposure to the same set of statistics after learning. This was most evident in the significant increase in the percentage of learners acquiring L2 (based on our criteria of strong learning) in Experiment 2 relative to Experiment 1. As noted above, the critical difference in the design of these experiments was the length of time learners were exposed to L1 after reaching the established threshold of learning. Consequently, the results of Experiments 1 and 2 lend support to our hypothesis that learners may reach a low level of estimation uncertainty during exposure to the first language. This notion runs counter to the proposal by Qian, Jaeger, and Aslin (2012), who suggested that learners may never reach a low level of estimation uncertainty regarding the underlying structure of the input and, thus, encountering variability in the input becomes attributed to variance associated with the initial statistical representation. However, in Experiment 1, we observed that learners who achieved 7 or 8 out of 8 on the interim tests tended to sustain their learning across additional tests, suggesting that these participants had formed a relatively stable statistical representation of the input. Accordingly, we suggest that once learners reach a low level of estimation uncertainty, additional input serves to entrench them in the statistics of the

initial input, thereby reducing their ability to detect the change into L2.

Although our results shed light on at least one of the contributing factors that leads to an anchoring effect in statistical learning, the underlying cognitive processes remain a topic for continued research. The results of our flanker task did not suggest a relationship between learners' inhibitory control abilities and their ability to acquire the L2. Thus, we did not find supporting evidence that learners need to inhibit the statistics of L1 to acquire L2. Likewise, that L1 is maintained at posttest in Experiment 2 suggests the second representation is not being acquired at the expense of the first. One possibility that is consistent with our findings is that learners may reduce their attention to the stream after achieving a low level of estimation uncertainty. Given that several studies have reported a central role for attention during statistical learning (Campbell, Zimmerman, Healey, Lee, & Hasher, 2012; Fernandes, Kolinsky, & Ventura, 2010; Finn, Lee, Kraus, & Hudson Kam, 2014; Toro, Sinnett, & Soto-Faraco, 2011; Turk-Browne, Jungé, & Scholl, 2005), it may be the case that learners reduce attention to the input once they have achieved robust learning. Thus, when the switch to L2 occurs, learners are already in a less attentive state. This would also account for why switching learners to the L2 immediately after robust learning of L1 occurs is an effective means of tempering the primacy effect, as learners are still maintaining high levels of attention at the point of the switch.

An attentional account of the primacy effect is also consistent with the findings from previous studies that suggest learners are better able to acquire multiple streams when they are paired with a contextual cue to signal the change in structure, such as a pause and explicit knowledge (Gebhart et al., 2009) or a change in speaker voice (Weiss et al., 2009). Contextual cues are thought to help learners overcome their stationarity bias by alerting them to a possible change in structure corresponding to the detection of a change in some other aspect of the input (Gebhart et al., 2009; Mitchel & Weiss, 2010; Weiss et al., 2009). Correspondingly, the contextual cue may induce the learner to raise their level of attention to the input. Given the large number of potential contextual cues available in any learning situation (see Qian, Jaeger, & Aslin, 2012), intuitively some contextual cues may be more effective than others at informing learners about changes in the underlying structure of the input (see Gebhart et al., 2009; Mitchel & Weiss, 2010). Thus, in Experiment 3 we used three different kinds of contextual cues to explore whether there are any additive effects for learning. That learners in this experiment did not perform better relative to those in Experiment 2 is consistent with the idea that these cues function to refocus attention. Given that advancing learners immediately after robust learning has occurred may serve to introduce the switch during periods of high attention, the benefits of adding contextual cues may not apply to this paradigm. Future experiments should more directly manipulate attention to further understand its role during multistream statistical learning experiments.

Another effect of switching participants soon after learning is that it increases variability early in familiarization. Introducing variability may itself provide an advantage for learners, as has been reported in many domains of psychology. For example several discrimination-learning studies have demonstrated that

when learners (such as rats) receive repeated reversal training, they are more likely to reverse their choice when they encounter a new reversal (e.g., Dufort, Guttman, & Kimble, 1954; Krechevsky, 1932; see Gallistel, Mark, King, & Latham, 2001). Further, research on desirable difficulties in learning (Bjork, 1994) suggests that varying the conditions of learning and interleaving conditions of study, rather than keeping learning conditions constant and predictable, trigger encoding and processing that support learning (e.g., Bjork, Bjork, & McDaniel, 2011).

Consistent with these benefits of variability, Zinszer and Weiss (2013) found that the primacy effect in statistical learning could also be attenuated if the input contains more switches between the L1 and L2. Even when participants receive a 5-min block of L1 at the start of the familiarization period (i.e., the same amount that entrenched learners in the original Gebhart et al., 2009, study), both structures can be learned above chance if the initial block is followed by several transitions between L1 and L2. This finding suggests that the transition points themselves, much like contextual cues, may serve to reengage participants' attention to the stream. Encountering switches earlier in the familiarization period may similarly advantage learners in avoiding entrenchment in a single structure (see also Experiment 2 in Zinszer & Weiss, 2013). These results indicate that the initial bias toward learning the first of two structures can be overcome by increasing the variability of the input and the timing of this variability also plays a role in entrenchment. Viewed in this context, our manipulation of switching participants to the L2 as soon as learning occurs may have provided learners with increased variability while attention remained high, and in turn, facilitated their ability to notice the change in structure without requiring contextual cues or the need for additional exposure to the second structure.

Another possible implication of our study highlights the potential tradeoff between efficiency and change detection. For efficient learners, L1 may be acquired quickly, but as input continues, this may lead to a reduction in the ability to notice change due to the anchoring effects associated with overlearning. Whereas most studies of overlearning find benefits for long-term retention (e.g., Earhard, Fried, & Carlson, 1972; Krueger, 1929; reviewed in Rohrer, Taylor, Pashler, Wixted, & Cepeda, 2005), this may come at the expense of detecting and acquiring new structures (particularly for successive learning that occurs over brief periods of time). This tradeoff may be evident in the reversal learning literature, as noted earlier. When contingencies are overlearned, reversals can become harder to detect (see Mackintosh, 1969 for a review). This trade-off between efficiency versus change detection is consistent with preliminary results from a recent neuroimaging study that used fMRI to track activation patterns during a multistream statistical learning task. Learners who were most efficient (based on activation levels in regions that support statistical learning) in their L1 processing tended to become more anchored in that language relative to those who required more effort to acquire L1, who tended to subsequently acquire L2 (Karuza et al., 2016). Undoubtedly, the trade-off between efficiency and change detection must be affected by factors such as the timescale for learning, the ease of acquisition, the amount of overlap between languages, as well as a host of other factors.

We close by noting the similarities between the primacy effect observed in statistical learning and entrenchment effects in other domains of learning. For example, in acquiring a new language, late second language learners tend to be outpaced in overall proficiency by early learners (e.g., Johnson & Newport, 1989). Although age of acquisition has often been proposed as an explanation for this difference in learning outcomes, a complementary claim by Hernandez et al. (2005) suggests that entrenchment in the first language may present difficulties for late learners in acquiring the nuanced aspects of the new language. Similar to the conclusions drawn from visual learning studies (Junge et al., 2007), Hernandez et al. (2005) asserted that previous knowledge (i.e., linguistic knowledge pertaining to the first language) is resistant to change and may only be altered slowly, which interferes with the acquisition of new information regarding a different linguistic system (see also Monner, Vatz, Morini, Hwang, & DeKeyser, 2013). Our results suggest a parallel between anchoring effects that occur at short intervals and those with a much longer trajectory. The introduction of variability within the window of active learning may attenuate entrenchment effects and allow learners to more easily form multiple representations that correspond to multiple inputs.

References

- Alloy, L. B., & Tabachnik, N. (1984). Assessment of covariation by humans and animals: The joint influence of prior expectations and current situational information. *Psychological Review*, *91*, 112–149. <http://dx.doi.org/10.1037/0033-295X.91.1.112>
- Aslin, R. N. (2014). Infant learning: Historical, conceptual, and methodological challenges. *Infancy*, *19*, 2–27. <http://dx.doi.org/10.1111/inf.12036>
- Atwater, S. K. (1953). Proactive inhibition and associative facilitation as affected by degree of prior learning. *Journal of Experimental Psychology*, *46*, 400–404. <http://dx.doi.org/10.1037/h0059416>
- Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*, 1214–1221. <http://dx.doi.org/10.1038/nn1954>
- Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). New York: Worth Publishers.
- Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA: MIT Press.
- Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. E. (2002). Immature frontal lobe contributions to cognitive control in children: Evidence from fMRI. *Neuron*, *33*, 301–311. [http://dx.doi.org/10.1016/S0896-6273\(01\)00583-9](http://dx.doi.org/10.1016/S0896-6273(01)00583-9)
- Campbell, K. L., Zimmerman, S., Healey, M. K., Lee, M. M. S., & Hasher, L. (2012). Age differences in visual statistical learning. *Psychology and Aging*, *27*, 650–656. <http://dx.doi.org/10.1037/a0026780>
- Dennis, M. J., & Ahn, W. K. (2001). Primacy in causal strength judgments: The effect of initial evidence for generative versus inhibitory relationships. *Memory & Cognition*, *29*, 152–164. <http://dx.doi.org/10.3758/BF03195749>
- Dufort, R. H., Guttman, N., & Kimble, G. A. (1954). One-trial discrimination reversal in the white rat. *Journal of Comparative and Physiological Psychology*, *47*, 248–249. <http://dx.doi.org/10.1037/h0057856>

- Earhard, B., Fried, C., & Carlson, G. (1972). Interference, overlearning, and anticipation time. *Journal of Experimental Psychology*, *94*, 345–347. <http://dx.doi.org/10.1037/h0032765>
- Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology* (N. York, Trans.). Dover: H. A. Rugen & C. E. Bussenius. (Original work published 1885)
- Emmorey, K., Luk, G., Pyers, J. E., & Bialystok, E. (2008). The source of enhanced control in bilinguals: Evidence from bimodal bilinguals. *Psychological Science*, *19*, 1201–1206.
- Eriksen, B., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*, 143–149. <http://dx.doi.org/10.3758/BF03203267>
- Fernandes, T., Kolinsky, R., & Ventura, P. (2010). The impact of attention load on the use of statistical information and coarticulation as speech segmentation cues. *Attention, Perception, & Psychophysics*, *72*, 1522–1532. <http://dx.doi.org/10.3758/APP.72.6.1522>
- Finn, A. S., Lee, T., Kraus, A., & Hudson Kam, C. L. (2014). When it hurts (and helps) to try: The role of effort in language learning. *PLoS ONE*, *9*, e101806. <http://dx.doi.org/10.1371/journal.pone.0101806>
- Franco, A., & Destrebecqz, A. (2012). Chunking or not chunking? How do we find words in artificial language learning? *Advances in Cognitive Psychology/University of Finance and Management in Warsaw*, *8*, 144–154. <http://dx.doi.org/10.2478/v10053-008-0111-3>
- Gallistel, C. R., Mark, T. A., King, A. P., & Latham, P. E. (2001). The rat approximates an ideal detector of changes in rates of reward: Implications for the law of effect. *Journal of Experimental Psychology: Animal Behavior Processes*, *27*, 354–372. <http://dx.doi.org/10.1037/0097-7403.27.4.354>
- Gebhart, A. L., Aslin, R. N., & Newport, E. L. (2009). Changing structures in midstream: Learning along the statistical garden path. *Cognitive Science*, *33*, 1087–1116. <http://dx.doi.org/10.1111/j.1551-6709.2009.01041.x>
- Gómez, R. L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences*, *4*, 178–186. [http://dx.doi.org/10.1016/S1364-6613\(00\)01467-4](http://dx.doi.org/10.1016/S1364-6613(00)01467-4)
- Gonzales, K., Gerken, L., & Gómez, R. L. (2015). Does hearing two dialects at different times help infants learn dialect-specific rules? *Cognition*, *140*, 60–71. <http://dx.doi.org/10.1016/j.cognition.2015.03.015>
- Hernandez, A., Li, P., & MacWhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences*, *9*, 220–225. <http://dx.doi.org/10.1016/j.tics.2005.03.003>
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*, 60–99. [http://dx.doi.org/10.1016/0010-0285\(89\)90003-0](http://dx.doi.org/10.1016/0010-0285(89)90003-0)
- Jungé, J. A., Scholl, B. J., & Chun, M. M. (2007). How is spatial context learning integrated over signal versus noise? A primacy effect in contextual cueing. *Visual Cognition*, *15*, 1–11. <http://dx.doi.org/10.1080/13506280600859706>
- Karuza, E. A., Li, P., Weiss, D. J., Bulgarelli, F., Zinszer, B., & Aslin, R. N. (2016). *Sampling over non-uniform distributions: A neural efficiency account of the primacy effect in statistical learning*. Manuscript submitted for publication.
- Krechevsky, I. (1932). Antagonistic visual discrimination habits in the white rat. *Journal of Comparative Psychology*, *14*, 263–277. <http://dx.doi.org/10.1037/h0070494>
- Krueger, W. C. F. (1929). The effect of overlearning on retention. *Journal of Experimental Psychology*, *12*, 71–78. <http://dx.doi.org/10.1037/h0072036>
- Mackintosh, N. J. (1969). Further analysis of the overtraining reversal effect. *Journal of Comparative and Physiological Psychology*, *67*(2, Pt. 2), 1–18. <http://dx.doi.org/10.1037/h0026784>
- Marsh, J. K., & Ahn, W. (2006). Order effects in contingency learning: The role of task complexity. *Memory & Cognition*, *34*, 568–576.
- Maye, J., Weiss, D. J., & Aslin, R. N. (2008). Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science*, *11*, 122–134. <http://dx.doi.org/10.1111/j.1467-7687.2007.00653.x>
- Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*, B101–B111. [http://dx.doi.org/10.1016/S0010-0277\(01\)00157-3](http://dx.doi.org/10.1016/S0010-0277(01)00157-3)
- Mitchel, A. D., & Weiss, D. J. (2010). What's in a face? Visual contributions to speech segmentation. *Language and Cognitive Processes*, *25*, 456–482. <http://dx.doi.org/10.1080/01690960903209888>
- Monner, D., Vatz, K., Morini, G., Hwang, S.-O., & DeKeyser, R. (2013). A neural network model of the effects of entrenchment and memory development on grammatical gender learning. *Bilingualism: Language and Cognition*, *16*, 246–265. <http://dx.doi.org/10.1017/S1366728912000454>
- Morales, J., Gómez-Ariza, C. J., & Bajo, M. T. (2013). Dual mechanisms of cognitive control in bilinguals and monolinguals. *Journal of Cognitive Psychology*, *25*, 531–546. <http://dx.doi.org/10.1080/20445911.2013.807812>
- Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*, 127–162. [http://dx.doi.org/10.1016/S0010-0285\(03\)00128-2](http://dx.doi.org/10.1016/S0010-0285(03)00128-2)
- Piaget, J., & Cook, M. (Trans.). (1952). *The originals of intelligence in children*. New York, NY: International Universities Press. <http://dx.doi.org/10.1037/11494-000>
- Qian, T., Jaeger, T. F., & Aslin, R. N. (2012). Learning to represent a multi-context environment: More than detecting changes. *Frontiers in Psychology*, *3*(July), 228.
- Reeder, P. A., Newport, E. L., & Aslin, R. N. (2013). From shared contexts to syntactic categories: The role of distributional information in learning linguistic form-classes. *Cognitive Psychology*, *66*, 30–54. <http://dx.doi.org/10.1016/j.cogpsych.2012.09.001>
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II Current Research and Theory*, *21*, 64–99.
- Rohrer, D., Taylor, K., Pashler, H., Wixted, J. T., & Cepeda, N. J. (2005). The effects of overlearning on long-term retention. *Applied Cognitive Psychology*, *19*, 361–374. <http://dx.doi.org/10.1002/acp.1083>
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996, December 13). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928. <http://dx.doi.org/10.1126/science.274.5294.1926>
- Saffran, J. R., Newport, E. E., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*, 606–621. <http://dx.doi.org/10.1006/jmla.1996.0032>
- Toro, J. M., Sinnott, S., & Soto-Faraco, S. (2011). Generalizing linguistic structures under high attention demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 493–501. <http://dx.doi.org/10.1037/a0022056>
- Turk-Browne, N. B., Jungé, J., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*, 552–564. <http://dx.doi.org/10.1037/0096-3445.134.4.552>
- Underwood, B. J. (1949). Proactive inhibition as a function of time and degree of prior learning. *Journal of Experimental Psychology*, *39*, 24–34. <http://dx.doi.org/10.1037/h0059550>
- Weiss, D. J., Gerfen, C., & Mitchel, A. D. (2009). Speech segmentation in a simulated bilingual environment: A challenge for statistical learning? *Language Learning and Development*, *5*, 30–49. <http://dx.doi.org/10.1080/15475440802340101>
- Weiss, D. J., Gerfen, C., & Mitchel, A. (2010). Colliding cues in word segmentation: The role of cue strength and general cognitive processes. *Language and Cognitive Processes*, *25*, 402–422. <http://dx.doi.org/10.1080/01690960903212254>
- Weiss, D. J., Poepsel, T., & Gerfen, C. (2015). Tracking multiple inputs:

- The challenge of bilingual statistical learning. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages*, pp. 167-190. Amsterdam, the Netherlands: John Benjamins Publishing Company.
- Yates, J. F., & Curley, S. P. (1986). Contingency judgment: Primacy effects and attention decrement. *Acta Psychologica*, 62, 293-302. [http://dx.doi.org/10.1016/0001-6918\(86\)90092-2](http://dx.doi.org/10.1016/0001-6918(86)90092-2)
- Zinszer, B. D., & Weiss, D. J. (2013). When to hold and when to fold: Detecting structural changes in statistical learning. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 3858-3863). Austin, TX: Cognitive Science Society.

Received August 6, 2015

Revision received January 21, 2016

Accepted January 21, 2016 ■

Call for Nominations

The Publications and Communications (P&C) Board of the American Psychological Association has opened nominations for the editorships of *Clinician's Research Digest: Adult Populations and Child and Adolescent Populations*; *Journal of Experimental Psychology: Learning, Memory, and Cognition*; *Professional Psychology: Research and Practice*; *Psychology and Aging*; and *Psychology, Public Policy, and Law* for the years 2019 to 2024. Thomas Joiner, PhD; Robert L. Greene, PhD; Ronald T. Brown, PhD; Ulrich Mayr, PhD; and Michael E. Lamb, PhD, respectively, are the incumbent editors.

Candidates should be members of APA and should be available to start receiving manuscripts in early 2018 to prepare for issues published in 2019. Please note that the P&C Board encourages participation by members of underrepresented groups in the publication process and would particularly welcome such nominees. Self-nominations are also encouraged.

Search chairs have been appointed as follows:

- *Clinician's Research Digest: Adult Populations and Child and Adolescent Populations*, Chair: Pamela Reid, PhD
- *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Chair: Stephen Rao, PhD
- *Professional Psychology: Research and Practice*, Chair: Kate Hays, PhD
- *Psychology and Aging*, Chair: Pamela Reid, PhD
- *Psychology, Public Policy, and Law*, Chair: David Dunning, PhD

Candidates should be nominated by accessing APA's EditorQuest site on the Web. Using your browser, go to <http://editorquest.apa.org>. On the Home menu on the left, find "Guests/Supporters." Next, click on the link "Submit a Nomination," enter your nominee's information, and click "Submit."

Prepared statements of one page or less in support of a nominee can also be submitted by e-mail to Sarah Wiederkehr, P&C Board Editor Search Liaison, at swiederkehr@apa.org.

Deadline for accepting nominations is Monday, January 9, 2017, after which phase one vetting will begin.