Event-related potential evidence of abstract phonological learning in the laboratory

Claire Moore-Cantwell[1], Joe Pater[2], Robert Staubs[3], Benjamin Zobel[2], and Lisa Sanders[2]
[1]Simon Fraser University, [2]University of Massachusetts Amherst, [3]MITRE corporation

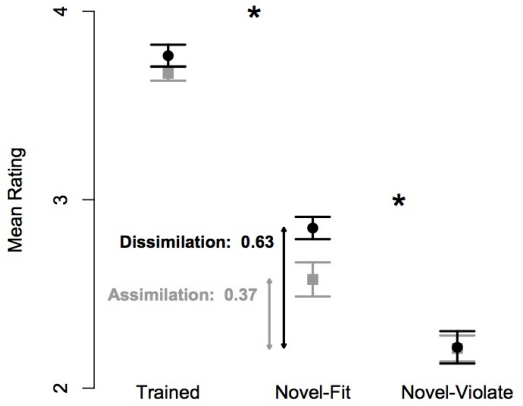
The experimental study of artificial language learning has become a widely used means of investigating the predictions of theories of phonology and of learning (review in Moreton and Pater 2012a,b). Although much is now known about the generalizations that learners make from various kinds of data, relatively little is known about how those generalizations are cognitively encoded. This paper presents an Event-related potential (ERP) study of brain responses to violations of lab-learned phonotactics. Novel words that violated a phonotactic constraint learned in the lab elicited a larger Late Positive Component (LPC) than novel words that satisfied it. Because an LPC has also been found in the study of naturalistically learned phonotactics, this new result provides support for the claim that phonotactic learning in the lab mimics naturalistic phonotactic learning, and can be used to study it. Furthermore, because the LPC is associated with violations of abstract "rules", such as syntactic violations and violations of musical expectations, this result provides evidence that generalizations acquired in the lab are cognitively abstract.

ERPs have been broadly applied in the study of language processing. An LPC, or P600, has been observed for a range of syntactic violations, including agreement, phrase structure, subcategorization, and constraints on long-distance dependencies (overview in Gouvea et al. 2010). It has also been observed for violations of naturalistically learned phonotactics, both first language (Domahs et al. 2009) and second (McLaughlin et. al. 2010). Beyond language, LPCs have been found for violations of musical structure (Patel et al. 1998; see Carrión and Bly 2008 for an overview), and for rule violations in arithmetic tasks (Núñez-Peña and Honrubia-Serrano 2004). It is also an indicator of abstract structural relations in these other cognitive domains: "an index of detection for any anomaly in rule-governed sequences" (Núñez-Peña and Honrubia-Serrano 2004, 130); [it] "reflects processes of knowledge-based structural integration" (Patel et al. 1998, 51). The N400, in contrast, is a component whose size is modulated by lexical processing difficulty. Recently observed words, expected words, or words that rhyme with a recently heard word all elicit smaller N400's than unexpected words, phonologically mismatched words, or nonwords (Rugg, 1984, Rossi et al, 2011).

We examined the learning of phonotactic generalizations over CVCV words, in which the consonants were drawn from the stops [d, g, t, k], and vowels from [i, æ, u, ɔ], as in Moreton (2008) *et seq*. Participants learned a set of words that conformed to one of two patterns (12 participants for each). In the Voice-Match condition, the stops were both voiceless, or both voiced. In the Voice-Mismatch condition, the stops disagreed in voicing. They heard the word over speakers, and were presented with a choice of four objects as the referent. After choosing one, they were presented with both the sound and the correct object. Eight words were presented five times each in a training block. After each training block, participants were asked to rate 24 words on a four-point scale for how likely it was that they were part of the language they were learning. Eight of the words were the ones they had learned – "Trained", 8 were novel words that fit the pattern – "Novel-Fit", and 8 were novel words that violated the pattern – "Novel-Violate". There 5 pairs of training-test blocks, for a total of 120 judgments for each participant. EEG was recorded continuously throughout the training and test trials. EEG from training trials

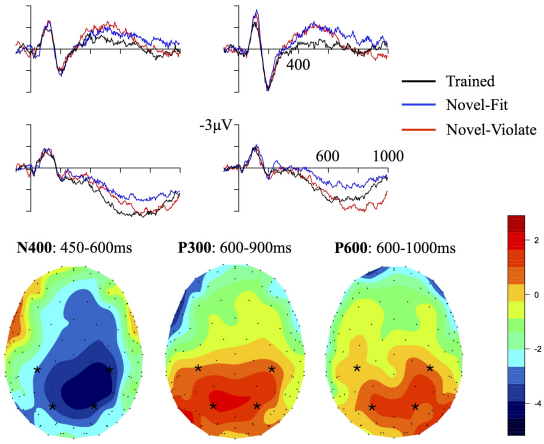
was averaged together by each block; EEG from test trials was averaged by condition (Trained, Novel-Fit, Novel-Violate) across all blocks.

During the first training block, participants were already well above chance performance of 25% on the word-picture matching task (M = 53.9%, SD = 2.4); in the four subsequent training blocks, performance was even better (M = 89.2%, SD = 1.8). During testing, participants rated the Trained words as more likely to be in the language (M = 3.72, SD = .18) than Novel-Fit words (M = 2.71, SD = .29) (t(23) = 14.77, p < .001). Novel-Fit words were rated as more likely than Novel-Violate (M = 2.21, SD = .26) (t(23) = 7.98, p < .001). No statistically significant differences were observed between Voice-Match and Voice-Mismatch in behavioral or electrophysiological data.



ERPs recorded during training revealed an N400 that decreased in amplitude over the course of the experiment, which is consistent with decreased lexical processing difficulty as the words were learned. The N400 was also smaller in Trained words than in Novel words. There was no evidence that amplitude 400-700 ms after onset differed between Novel-Fit and Novel-Violate items.

Novel-Fit and Novel-Violate items did differ in terms of the LPC, as illustrated in the figure below. The waveforms show averaged EEGs from four electrodes whose placement on the scalp is indicated by asterisks. Starting at about 600 ms, after word onset, the Novel-Violate waveform shows a deflection in a positive direction (down) relative to Novel-Violate. The P600 heatmap indicates that this difference is greatest in the right posterior region.



If the participants in this study were treating Novel-Fit words differently because of some direct calculation of similarity to the words they learned, we might expect an N400 difference with Novel-Violate (which was in fact found in early L2 acquisition by McLaughlin et. al. 2010). The finding of an LPC suggests instead that the phonotactic regularity is encoded as a structural "rule", like syntax, or L1 and late L2 phonotactics.

Selected references

Domahs et al. 2009. Event-related potentials reflecting the processing of phonological constraint violations. *Language and Speech*. *Language and Speech* 52: 415–435.

McLaughlin et al. 2010. Brain potentials reveal discrete stages of L2 grammatical learning. *Language Learning* 60: 123-150.