PREDICTION ERROR AND CUE COMPETITION IN SPEECH SOUND ACQUISITION

Given the continuous, noisy and highly complex nature of the speech signal, listeners must not only attend to important cues, but also ignore a vast number of irrelevant cues. This ability to ignore unimportant cues starts to emerge in the first few months of life, coincides with enhanced perception of critical cues (Werker & Tees, 1984) and increases over the lifetime (e.g. Flege et al., 1995). Adult listeners become so adept at ignoring irrelevant cues that this often poses problems for learning a second language (L2; Miyawaki et al., 1975). But exactly how we learn the cues of our languages and how we adjust to new cues in new languages have been subjects of much research and debate. Over the last two decades, the proposal that speech sounds are learned through statistical clustering mechanisms (e.g. unimodal vs. bimodal distributions) has become perhaps the most common explanation for speech sound acquisition (e.g. Maye et al., 2002). However, computational models suggest that clustering alone is not sufficient for learning speech sounds (e.g. McMurray Aslin & Toscano, 2009).

Independently of the language acquisition literature, there is a long history of research on learning in psychology, from which error-driven learning models have emerged (e.g. Rescorla & Wagner, 1972). These models and discriminative learning theory propose that learning occurs through prediction and prediction error. Potentially any perceivable cues may be used to predict important event outcomes. Prediction accuracy is used to adjust expectations about future events.

In two simulations and artificial language learning experiments, the present study tests two predictions of discriminative learning models for L2 speech sound acquisition: cue competition (blocking; Exp 1) and predictive structure (Exp 2). Simulations using Rescorla-Wagner equations first trained a single acoustic cue (e.g. tone) to predict an outcome (object size). Later, a second cue was added: nasal+tone predicted size. Due to the nonlinear 'learning curve' and cue competition, the initial cue gained strong connection weight, while the second cue weight remained weak – Kamin's famous (1969) *blocking effect*. In simulations in which two cues were trained from the beginning, they reached equal cue weighting. Model predictions were tested in Experiment 1. The blocking group was pre-trained with a single cue (e.g. tone) predicting an outcome (size). Next, two cues (tone+nasal) were presented with the outcome. Finally, participants were tested on the second cue (nasal). A control group had the same training and test except that the initial pre-training was with a control cue instead of tone. Accuracy data was analysed with glmer. Results showed that accuracy was significantly lower in the blocking condition than the control condition ($p=.031$) – learning one acoustic cue had *blocked* later learning of a second cue. Because the pre-trained cue provided sufficient *information* to predict the outcome, the second cue added *no further predictive value*. This result may help explain many L1 transfer effects in L2 learning.

The Hebbian notion 'neurons that fire together wire together' is a common conceptualisation of associative learning. Discriminative learning proposes instead that there is an asymmetric temporal structure where earlier cues predict later outcomes (e.g. Ramscar et al., 2010). That is, learning results from prediction, not simply co-occurrence. Experiment 2 tests whether speech sound acquisition occurs from co-occurrence or the temporal predictive structure of learning events. Participants were trained either with rich variable acoustic cues followed by simple images or the reverse order. Accuracy was significantly higher when the variable acoustic cues preceded the images, compared to the reverse ($p=.018$). When cues occurred early, irrelevant cues could be downweighted and relevant cues strengthened according to outcomes relative to expectations. When the images occurred first, there was no cue competition and only statistical co-occurrence frequencies were learned.

Together, the results of these experiments suggest that speech sound acquisition involves prediction and prediction error. The results suggest two conclusions: 1) the history of learning affects current learning and 2) the predictive structure of learning events affects what is learnt.

Kamin, L. (1969). Predictability,surprise, attention & conditioning. Punishment & aversive behavior.

Rescorla, R. and Wagner, A., (1972). A theory of Pavlovian conditioning. Black, A. H., Prokasy, W. F. (Eds.), Classical conditioning II: Current research theory. Ap.-Cent-Crofts, New-York,. 64– 99.

McMurray, B., Aslin, R. N., Toscano, J. C., 2009. Statistical learning of phonetic categories: insights from a computational approach. Developmental Science 12 (3), 369–378.