Evidence for reward learning in speech production

Benjamin Parrell
University of Wisconsin–Madison
bparrell@wisc.edu

Reward learning in speech

Reward learning has been suggested to be a critical component of early speech motor plan acquisition [1,2,3]. There is little direct behavioral evidence for reward learning in speech motor control. Can adult speakers learn a new production target based only on reinforcement learning?

Mechanisms of motor learning

Sensory-prediction error (SPE) learning: Learning from a mismatch between expected and perceived sensory outcomes of one’s actions. (= learning how to do a particular action)

Reinforcement or reward learning: Learning from an action’s outcome and the presence or absence of reinforcement. (= learning which actions to perform)

Experiment design

Baseline phase:
- measurement of baseline F1 for /i:/ vowel (EH, e.g. “head”) and /u:/ vowel (IH, e.g. “hid”) – no reward signal

Test phase:
- participant starts with 1000 points
- tokens produced in “reward region” (~110 to ~10 Hz below baseline F1) earn +10 points
- productions outside this region lose -10 points

Washout phase:
- no reward signal

Conditions

Sensory feedback:
EXP 1/2: Masking noise (speech-shaped, 85 dB)
EXP 3: Normal auditory feedback
Q: Does the presence of sensory feedback in the unmasked condition interfere with reward learning?

Reinforcement signal:
EXP 1/3: Participant’s own speech, with F1 shifted to center of reward region
EXP 2: Arbitrary noise (chime)
Q: Do participants benefit from a “reformulation” of their own speech with an implicit auditory target?

References

Predictions

H0: No learning
H1: Learning as in SPE learning

Results

EXPERIMENT 1 (13/20 learners, 65%)
- Masking noise
- Negative reward: participant’s own production of incorrect word (closest to mean produced during the baseline phase)
- Positive reward: participant’s own production of the target word from the baseline phase (closest to mean), with F1 shifted down by 60 Hz to the center of the reward region

F1 shift from baseline (Hz)

EXPERIMENT 2 (9/21 learners, 43%)
- Masking noise
- Negative reward: external voice producing “heard” word
- Positive reward: chime

F1 shift from baseline (Hz)

EXPERIMENT 3 (14/18 learners, 78%)
- No masking noise
- Negative reward: participant’s own production of incorrect word
- Positive reward: shifted version of participant’s own production of the target word from baseline phase

F1 shift from baseline (Hz)

Key findings

Some adult speakers can learn new production targets based purely on external reinforcement.

Reinforcement learning resulted in long-term changes to production even after reward was no longer given.

This differs from SPE learning, where participants return to their baseline quickly but is similar to reinforcement learning in reaching [7].

Sensory feedback does not inhibit reward learning. This differs from reaching tasks, where the presence of SPE interferes with reward learning [8].

Participants were typically unable to adapt a useful explicit strategy to achieve this change. As assessed in follow-up surveys.

Work supported by the University of Delaware Research Foundation.