# 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.

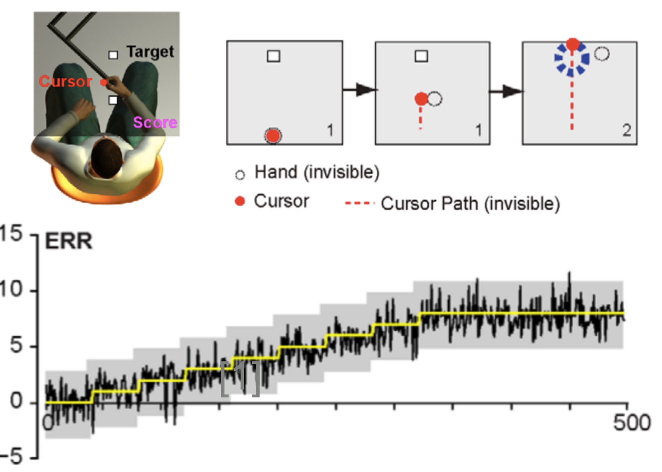
**Can adult speakers learn a new production target based only on reinforcement learning?**

## Mechanisms of motor learning

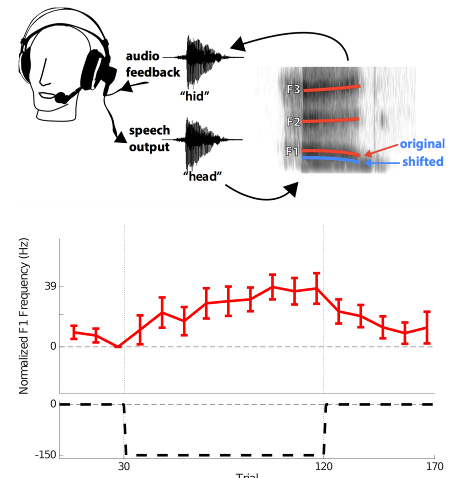
### Sensory-prediction error (SPE) learning:

*Learning from a mismatch between expected and perceived sensory outcomes of one's actions.*
*(≈ learning how to do a particular action)*

<u>Reaching [4]</u>
*visuomotor rotation*



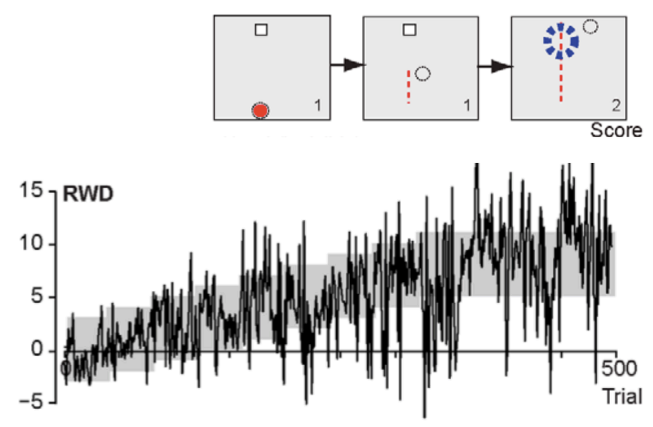

<u>Speech [5,6…]</u>
*F1 perturbation*



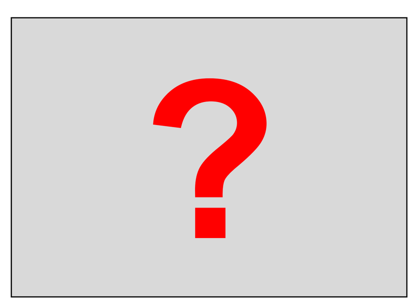

### Reinforcement or reward learning:

*Learning from whether an action was judged to be successful or unsuccessful, separate from whether it matches sensory expectations.*
*(≈ learning which actions to perform)*

<u>Reaching [4]</u>




<u>Speech</u>

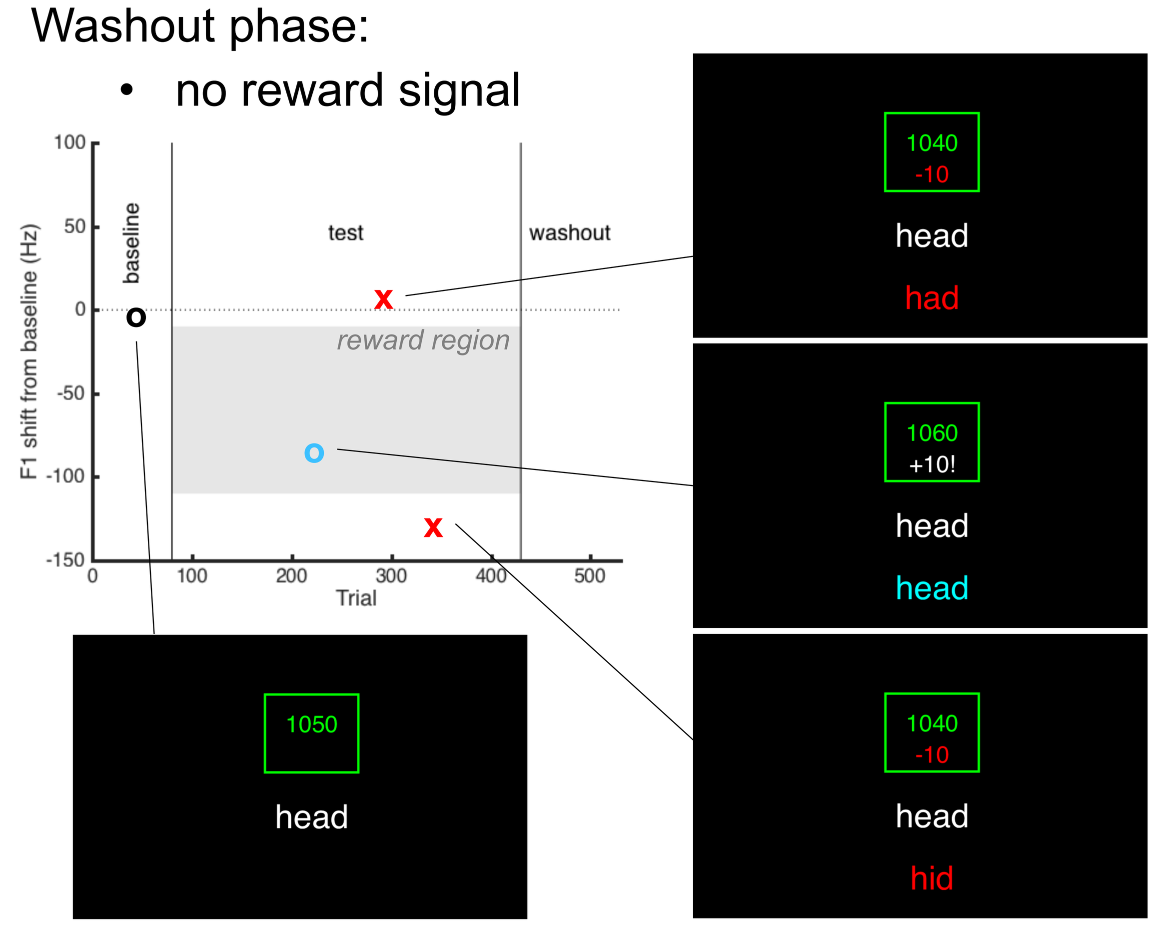



## Experiment design

Baseline phase:
- measurement of baseline F1 for /ɛ/ vowel (EH, e.g. "head") and /ɪ/ 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?*

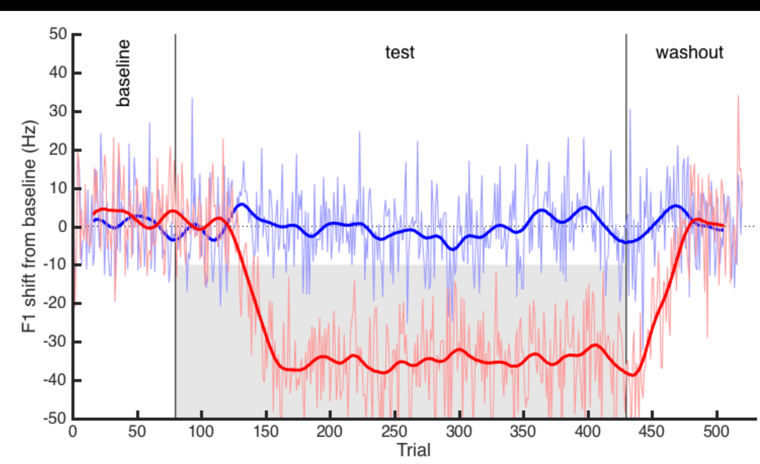
### Reinforcement signal:

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?*

## Predictions

**H0:**
**No learning**

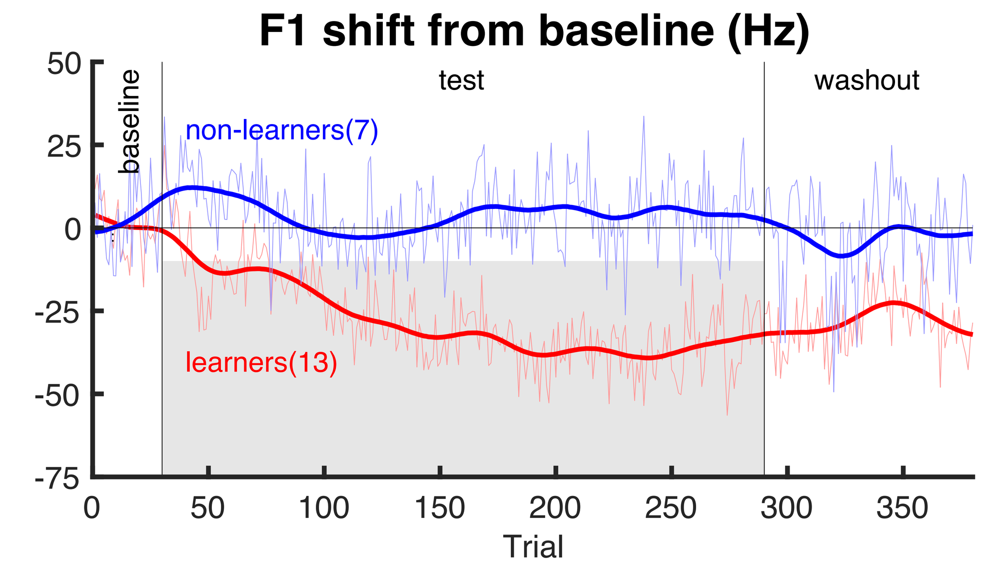
**H1:**
**Learning as in SPE learning**




## Results

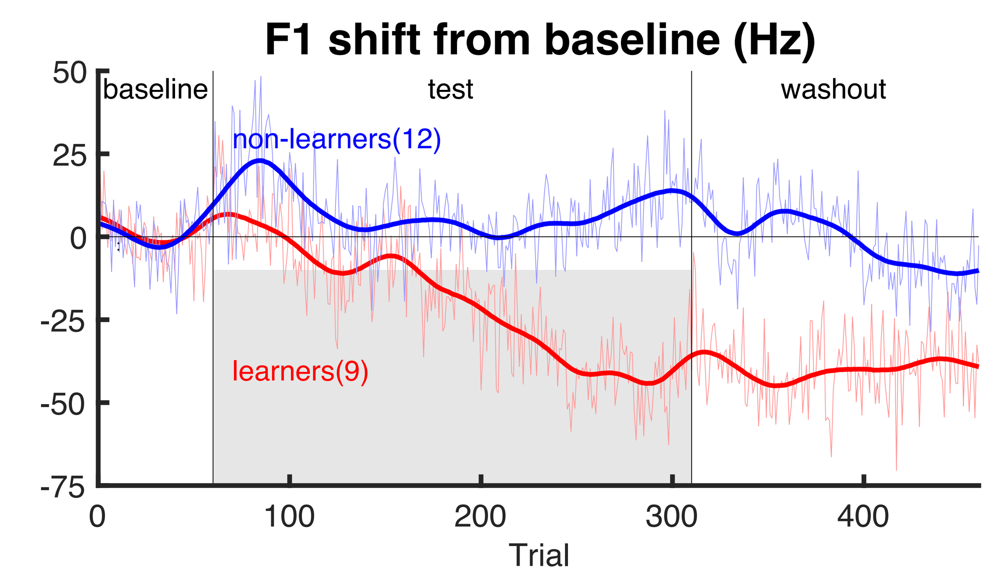
### <u>EXPERIMENT 1 (13/20 learners, 65%)</u>

- **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 *<u>baseline</u>* phase (closest to mean), with F1 shifted down by 60 Hz to the center of the reward region



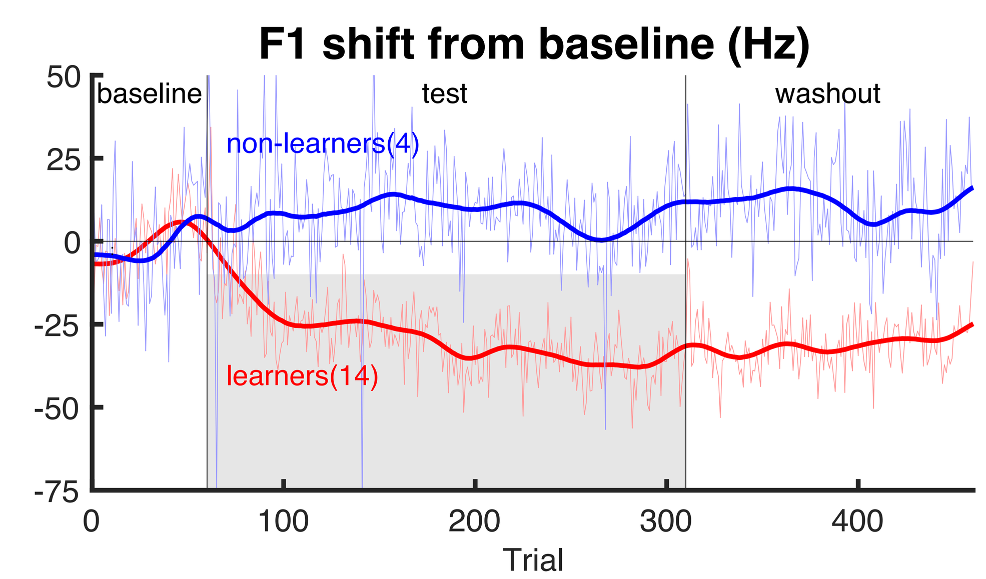

### <u>EXPERIMENT 2 (9/21 learners , 43%)</u>

- **Masking noise**
- **Negative reward**: external voice producing "heard" word
- **Positive reward**: chime

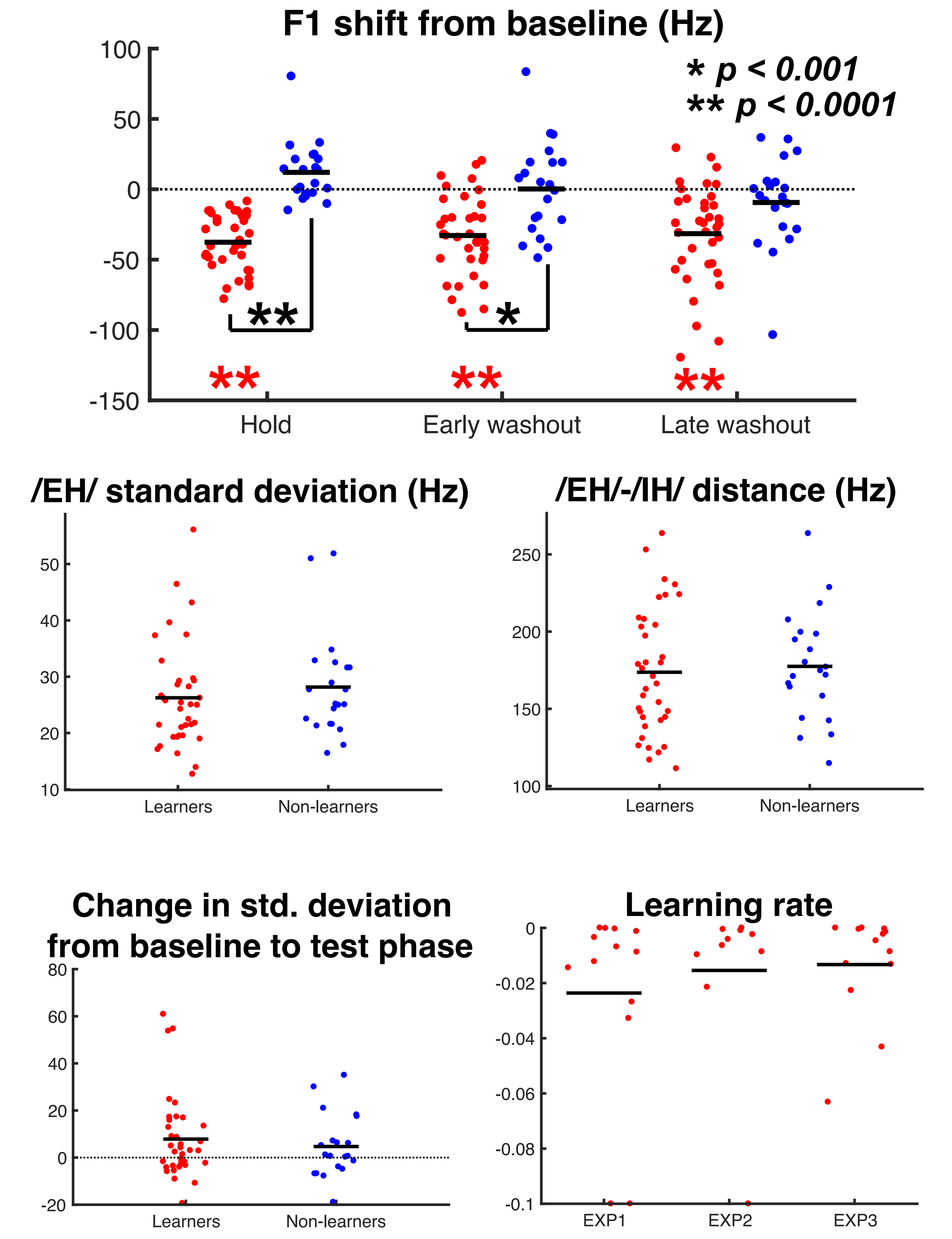



### <u>EXPERIMENT 3 (14/18 learners , 78%)</u>

- **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




### <u>COMBINED RESULTS</u>




## Key findings

**Some adult speakers can learn to alter their speech 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.
Printing courtesy of the Friends of the Waisman Center

## References

[1] Howard, I. S., & Messum, P. (2011). Modeling the development of pronunciation in infant speech acquisition. *Motor Control*, *15*(1), 85-117.
[2] Warlaumont, A. S., & Finnegan, M. K. (2016). Learning to produce syllabic speech sounds via reward-modulated neural plasticity. *PloS One*, *11*(1), e0145096.
[3] Rasilo, H., Räsänen, O., & Laine, U. K. (2013). Feedback and imitation by a caregiver guides a virtual infant to learn native phonemes and the skill of speech inversion. *Speech Communication*, *55*(9), 909-931.
[4] Izawa, J., & Shadmehr, R. (2011). Learning from sensory and reward prediction errors during motor adaptation. *PLoS Computational Biology*, *7*(3), e1002012.
[5] Houde, J., & Jordan, I. M. (1998). Sensorimotor adaptation in speech production. *Science*, *279*, 1213-1216.
[6] Purcell, D. W., & Munhall, K. G. (2006). Adaptive control of vowel formant frequency: Evidence from real-time formant manipulation. *The Journal of the Acoustical Society of America*, *120*(2), 966-77.
[7] Uehara, S., Mawase, F., & Celnik, P. (2017). Learning Similar Actions by Reinforcement or Sensory-Prediction Errors Rely on Distinct Physiological Mechanisms. Cerebral Cortex, 1–13. https://doi.org/10.1093/cercor/bhx214.
[8] Cashaback, J. G. A., McGregor, H. R., Mohatarem, A., & Gribble, P. L. (2017). Dissociating error-based and reinforcement-based loss functions during sensorimotor learning. *PLoS Comput Biol*, *13*(7), e1005623.