

Informativity in adaptation: Supervised and unsupervised learning of linguistic cue distributions



Dave F. Kleinschmidt, Rajeev Raizada, and T. Florian Jaeger, University of Rochester
 dkleinschmidt@bcs.rochester.edu, raizada@bcs.rochester.edu, fjaeger@bcs.rochester.edu

Our question:

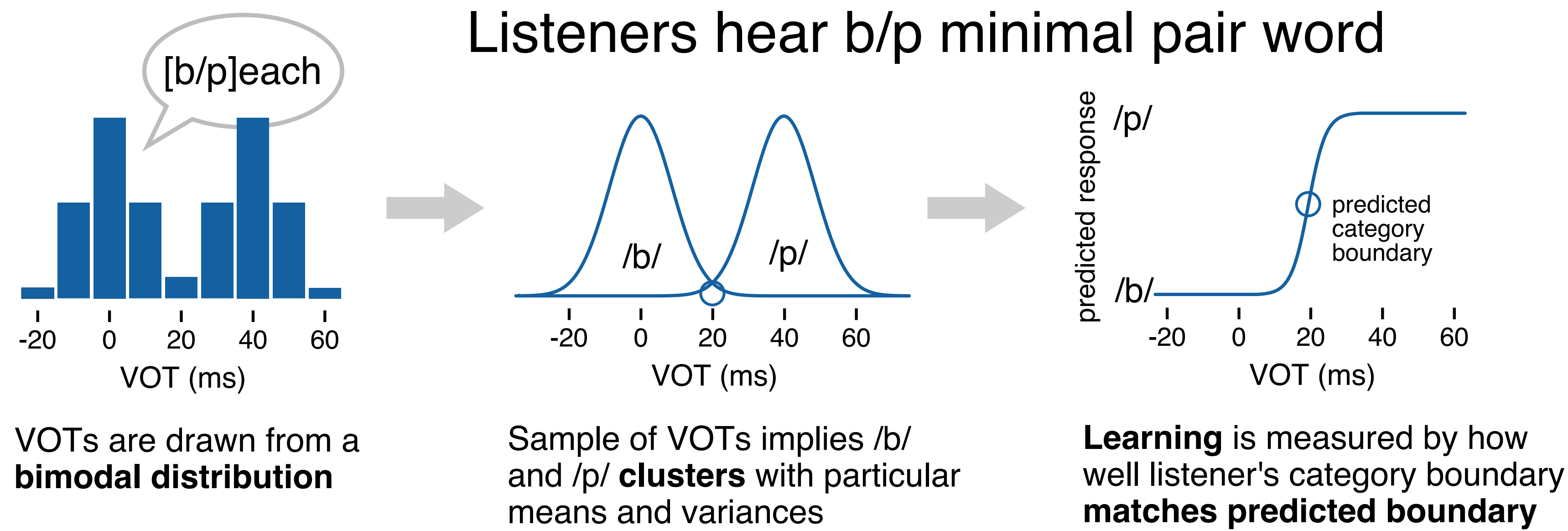
Do people use **category labels** during adaptation?

Language learning doesn't stop once you reach adulthood: talkers use linguistic cues to realize their intentions in different ways. To **adapt** to a new talker, you have to learn the way they use cues. If you know their intended meaning, this learning should be a lot easier. Learning with known category labels is called **supervised learning**, and learning from cues only is called **unsupervised learning**.

Why we ask:

- Categories are **distributions** of cues
- Productions vary within talker
- Productions vary across talkers
- Requires **distributional learning** for
- Acquisition: learn language's distributions
- Adaptation: learn talker's distributions
- Are they the same underlying process?
- Why is acquisition **slow** and adaptation **fast**?
- Adults have more information from experience
- Other cues **label** sounds with intended category

What we did: Distributional learning of /b/ and /p/
 Listeners hear b/p minimal pair word



...with and without labels and click on matching picture. Trial is either:

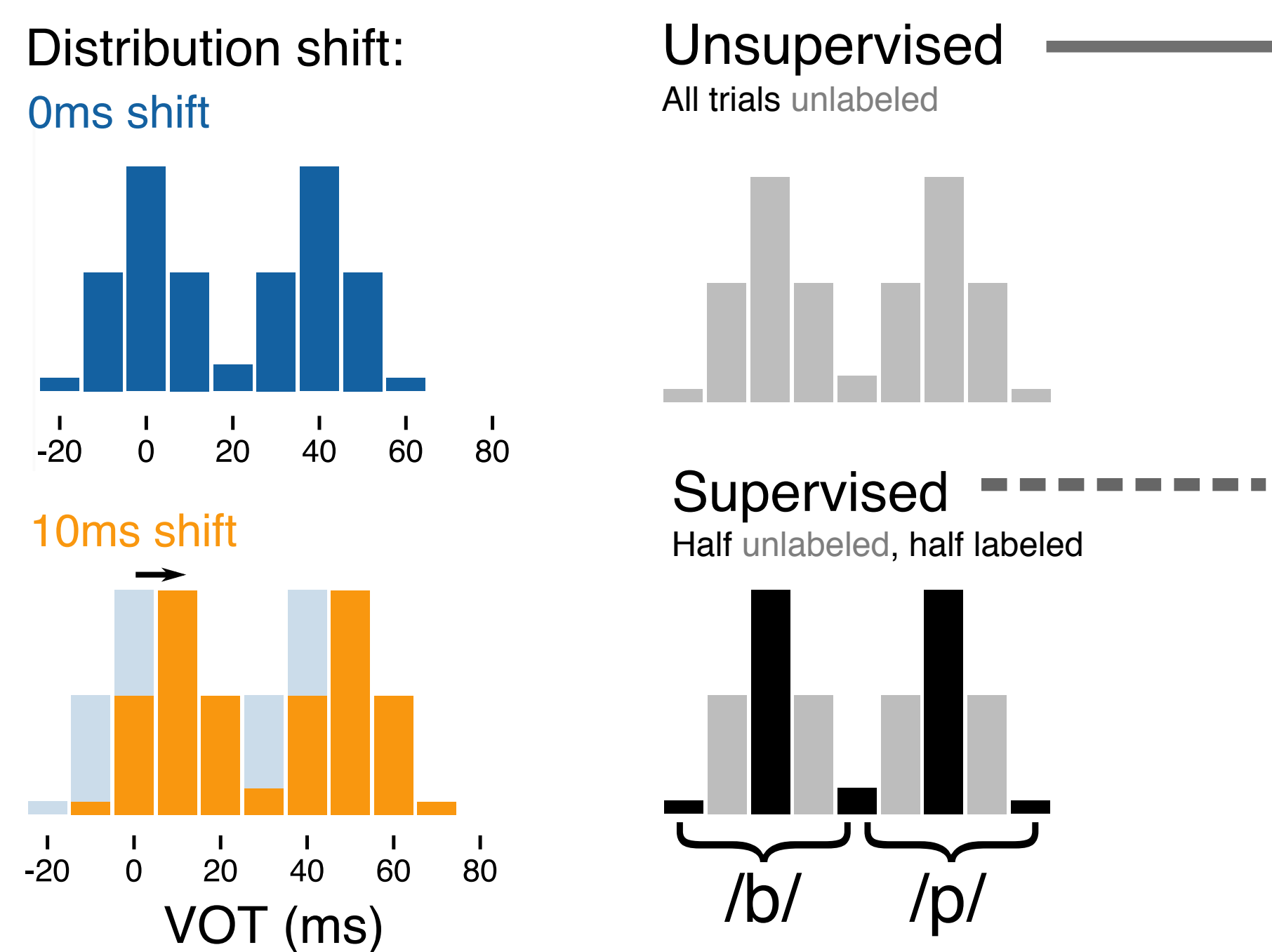
Labeled

provide **teaching signal** (supervision): only *beach* matches, talker **intended /b/**

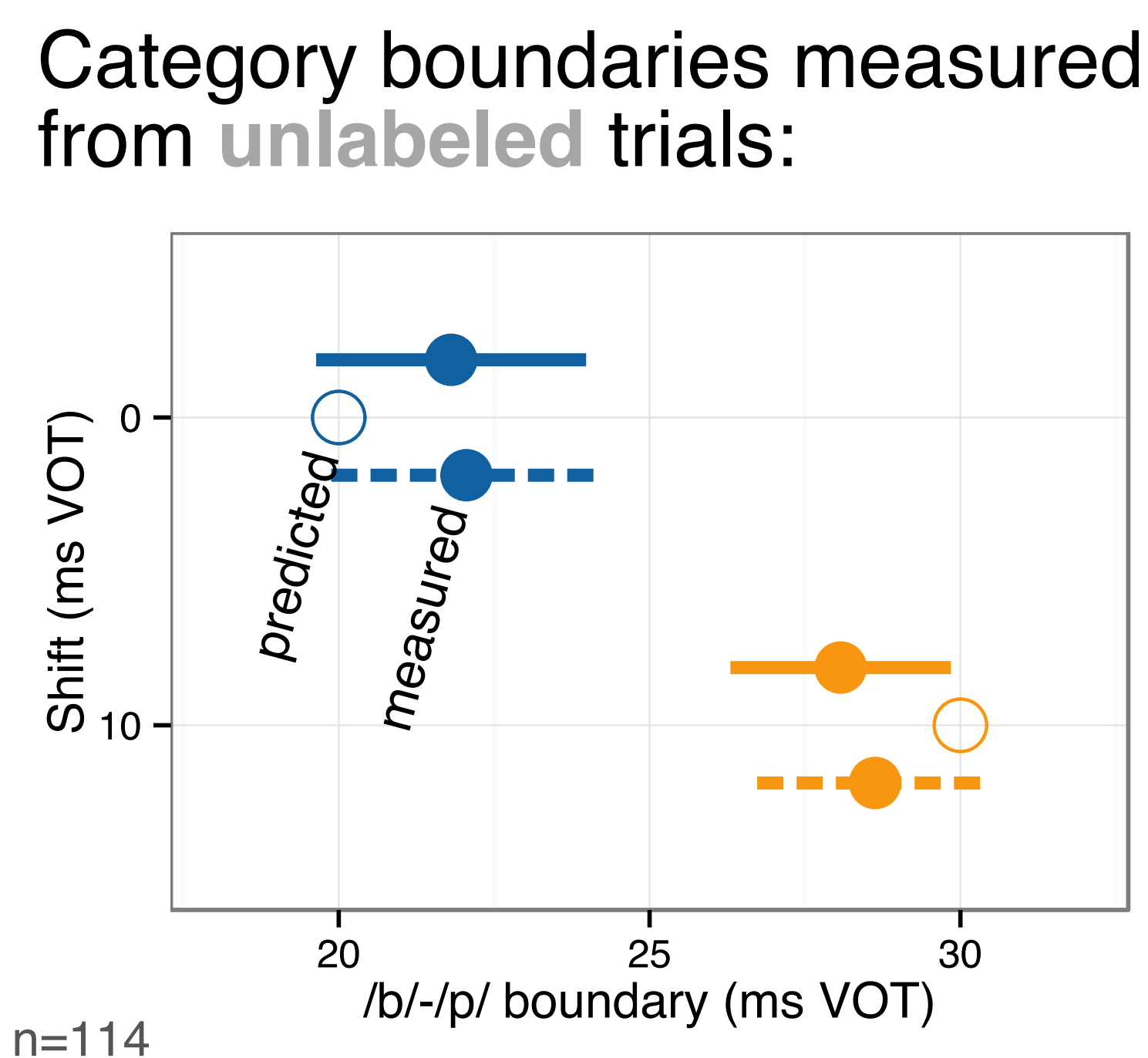
Unlabeled

measure /b/-/p/ category boundary: both *beach* and *peach* match, **ambiguous**

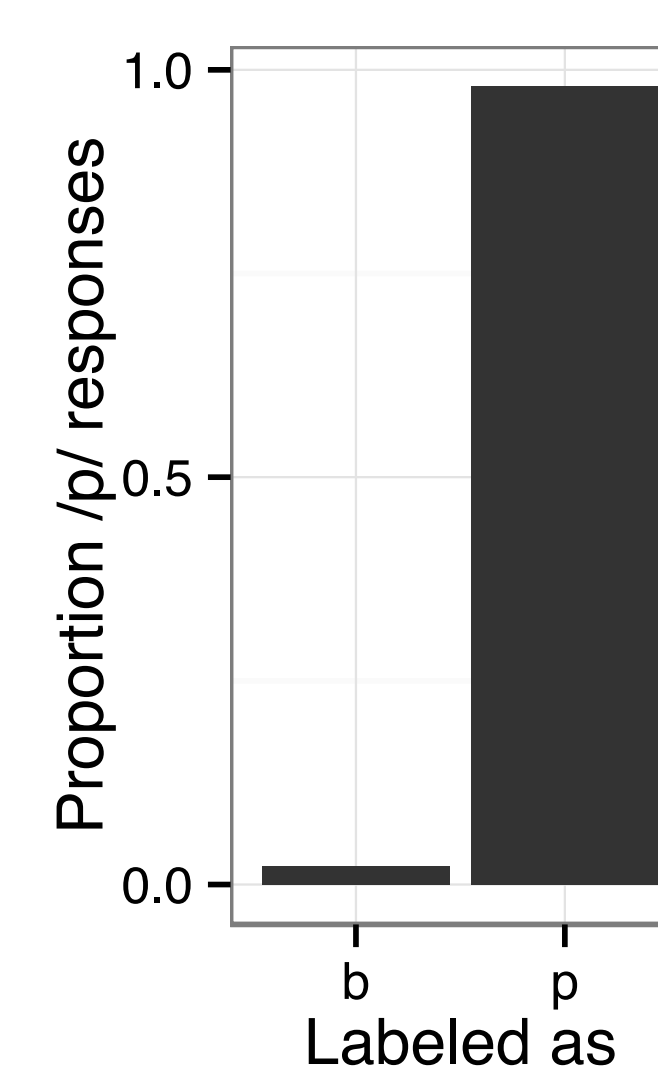
Experiment 1



Results



Accuracy on **labeled** trials: **98%**

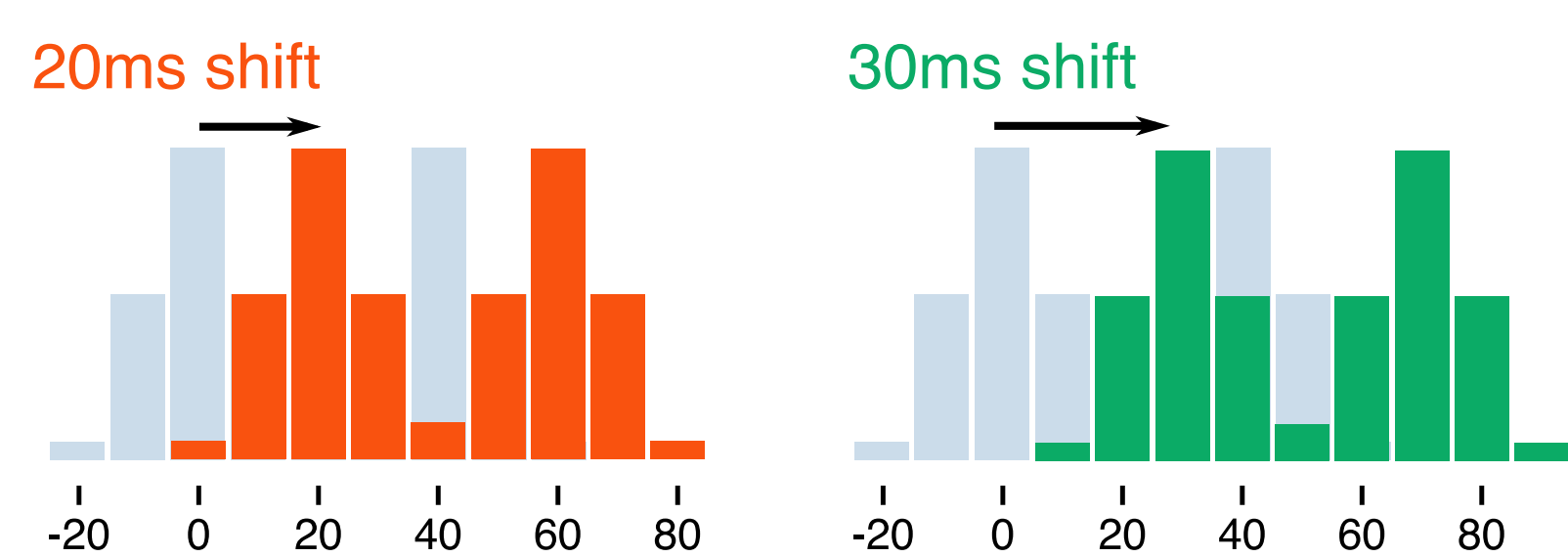


Discussion

- 1) Learning was good: category boundaries match distributions
- 2) People used labels to **choose responses** (98% accurate)
- 3) But **labels didn't change learning** (no difference between supervised and unsupervised)
- 4) Was it **too easy**?

Experiments 2+3

Bigger shifts to make learning harder:

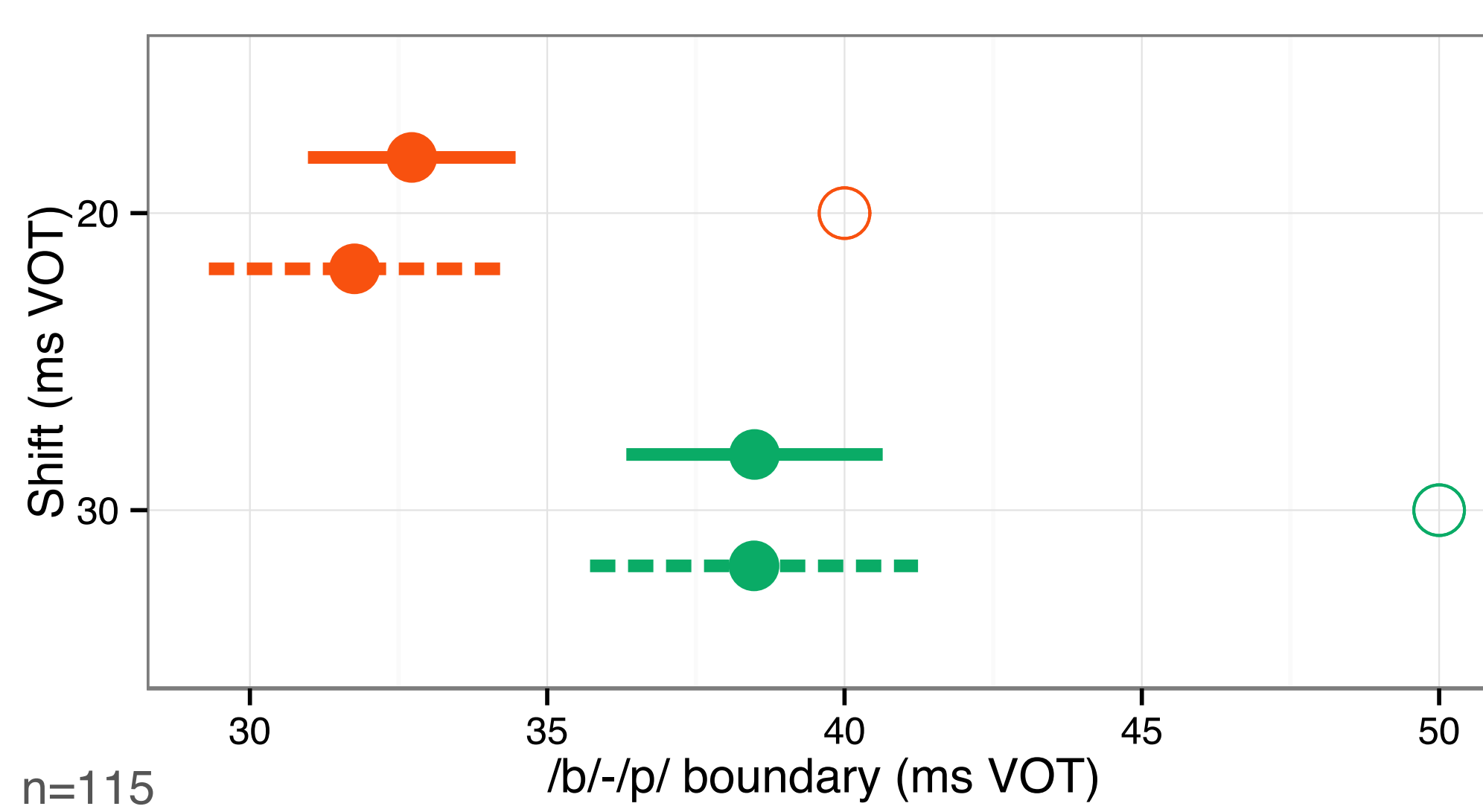


Same as Experiment 1:

Unsupervised: All trials unlabeled
 Supervised: Half unlabeled, half labeled

Results

Category boundaries:

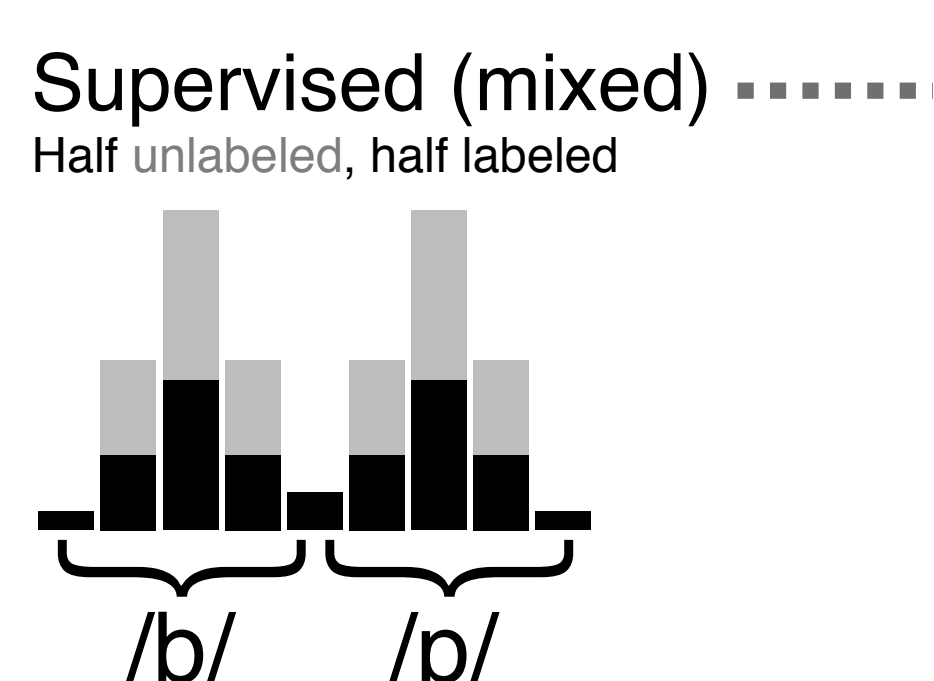


Discussion

- 1) Learning was **worse** for large shifts (category boundaries much lower than distributions predict)
- 2) Still **no effect of labels** (unsupervised = supervised).
- 3) Were labeled trials distributed **too sparsely** over VOT?

Experiment 4

Mix labeled trials over whole VOT range

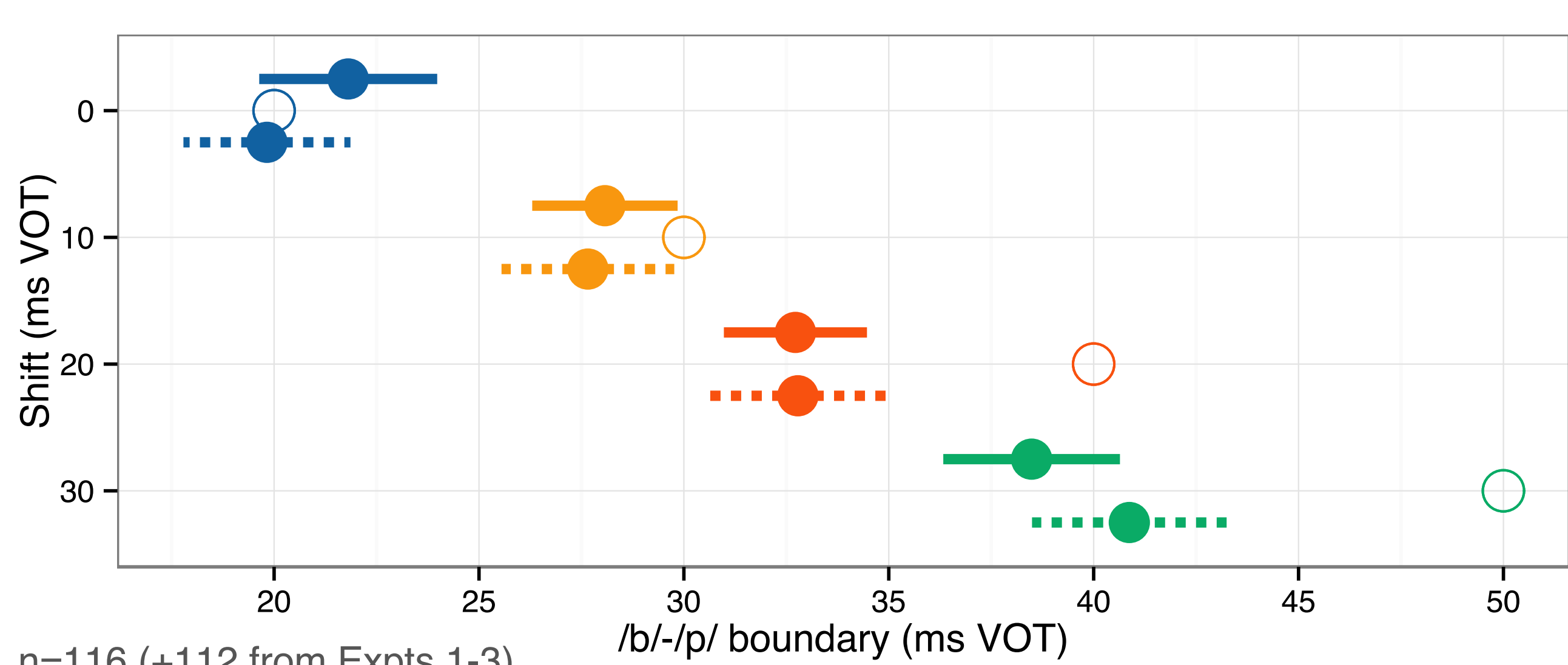


From Experiments 1-3:

Unsupervised: All trials unlabeled

Results

Category boundaries:



Discussion

- 1) Still **no effect of labels** (or only marginal if any)

Conclusions

Surprisingly, **people do not appear to use informative labels for adaptation**, even though they do for classification.

Two possible reasons why:

- 1) Other studies use **intrinsic labels** (lexical or audio-visual cues).
- 2) Goldilocks problem: **too easy or hard** for label use to be detectable. Unlabeled trials contain a lot of distributional information, and high-shift conditions are very unnatural