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 in adults: 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 intented meaning, this learning should be a lot easier. Learning with known category labels is called **supervised learning.**

Why?

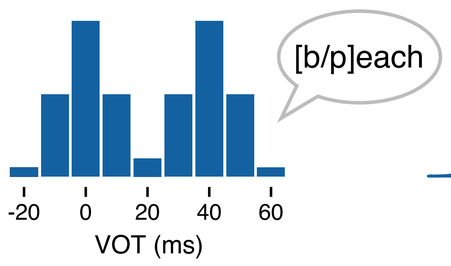
Categories are **distributions** of cues
- Productions vary within talker
- Productions vary across talkers

Distributional learning:
- Acquisition: learn language's distributions
- Adaptation: learn talker's distributions

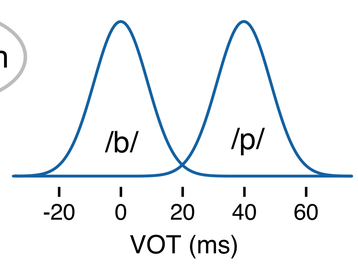
Same underlying process?
- Why is acquisition **slow** and adaptation **fast**?
- Adults have more information from experience
- Other cues **label** sounds with intended cateogry

What?

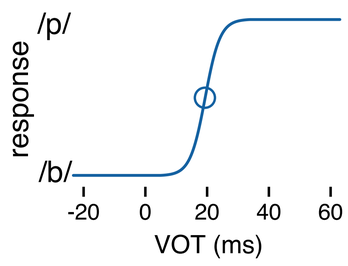
Distributional learning of /b/ and /p/



Listener hear words with VOT drawn from a **bimodal distribution**



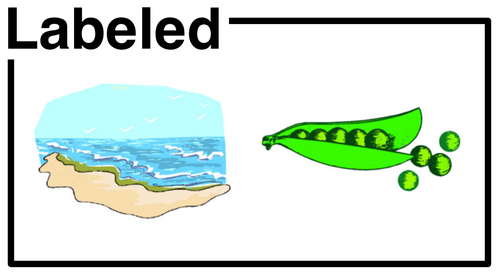
These samples imply /b/ and /p/ **clusters** with particular means and variances



Mean and variance of clusters **predicts** classification function, including **category boundary**

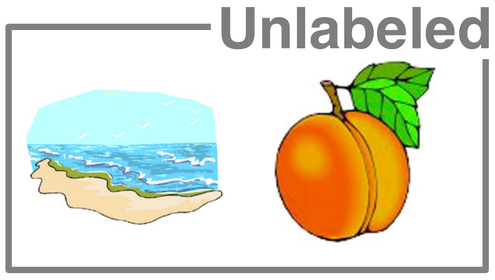
...with and without labels

Labeled



provide **teaching signal** (supervision)

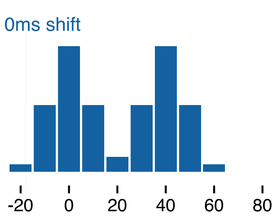
Unlabeled



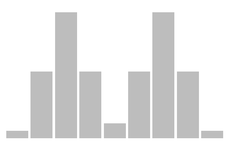
measure /b/-/p/ category boundary via responses to different VOTs

Experiment 1

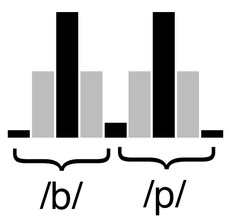
Distribution shift:



Unsupervised
All trials unlabeled

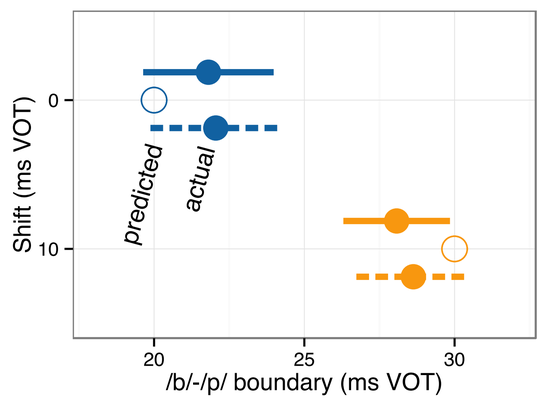


Supervised
Half unlabeled, half labeled

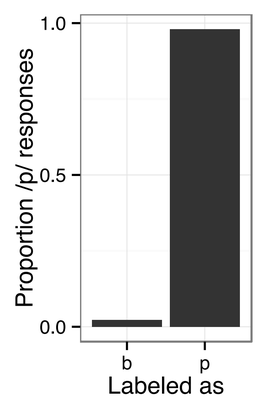


Results

Category boundaries measured from **unlabeled** trials



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



Discussion

1) Learning was good: category boundaries match distributions

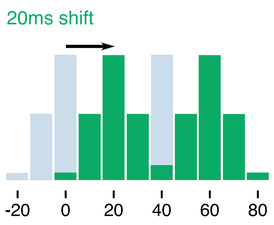
2) People used labels to **choose responses**

3) But **labels didn't change learning**

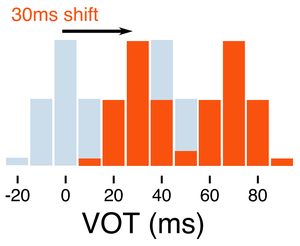
4) Was it **too easy**?

Experiments 2+3

20ms shift

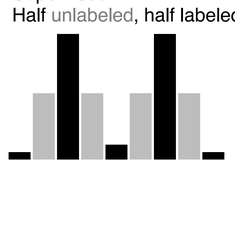


30ms shift



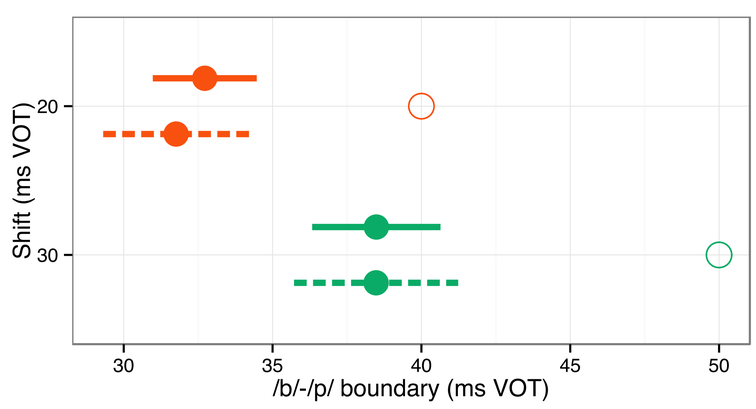
Unsupervised
All trials unlabeled

Supervised
Half unlabeled, half labeled



Results

Category boundaries



Discussion

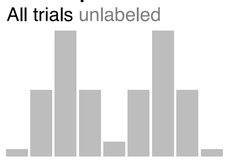
1) Learning was **worse** for large shifts

2) Still **no effect of labels** on learning speed or completeness/accuracy

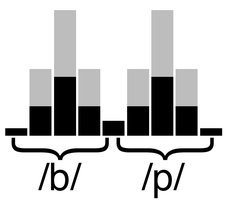
3) Were labeled trials distributed **too sparsely** over VOTs?

Experiment 4

Unsupervised
All trials unlabeled



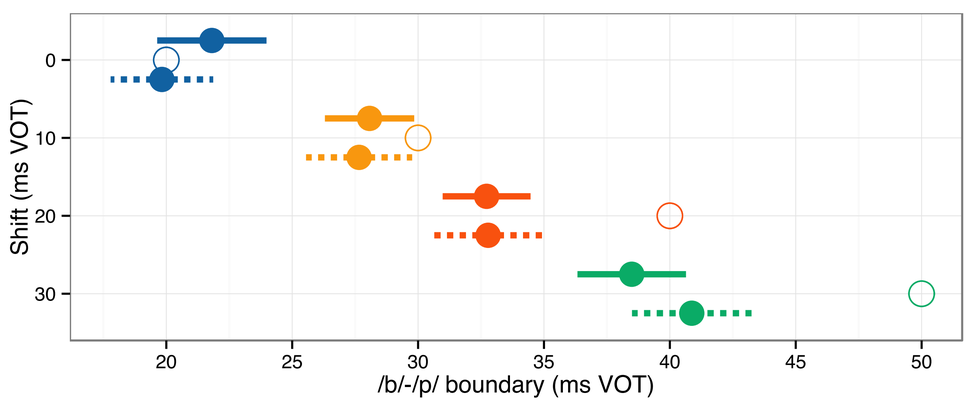
Supervised (mixed)
Half trials labeled



Results & Discussion

Category boundaries

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.