UC San Diego

# Distributional learning of L2 phonological categories by listeners with different language backgrounds

**Bożena Pająk & Roger Levy**

[bɔˈʒɛna ˈpajɔ̃k] [ˈɹɑdʒəˠ ˈlivi]

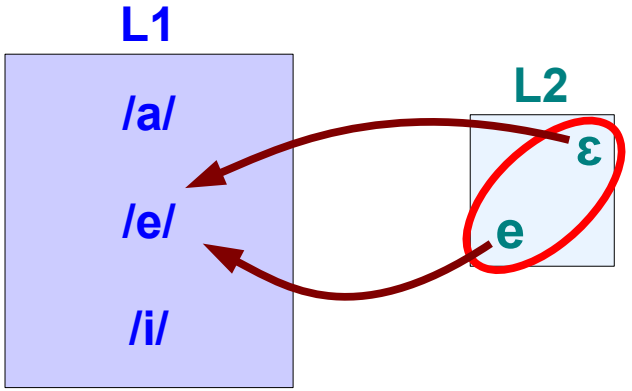
UC San Diego

*BUCLD :: Nov 6, 2011*

# Background

**Question**: How do people learn a second language?

- Generally hard, success depends on many factors
- Previous research:
  - focus on explaining *difficulties* in L2 learning
  - well-studied: L1 as a source of *interference*
- E.g., non-native speech perception & phonetic category learning
  (Best, 1995; Best & Tyler, 2007; Flege, 1995; Kuhl & Iverson, 1995)




known result:

hard to discriminate

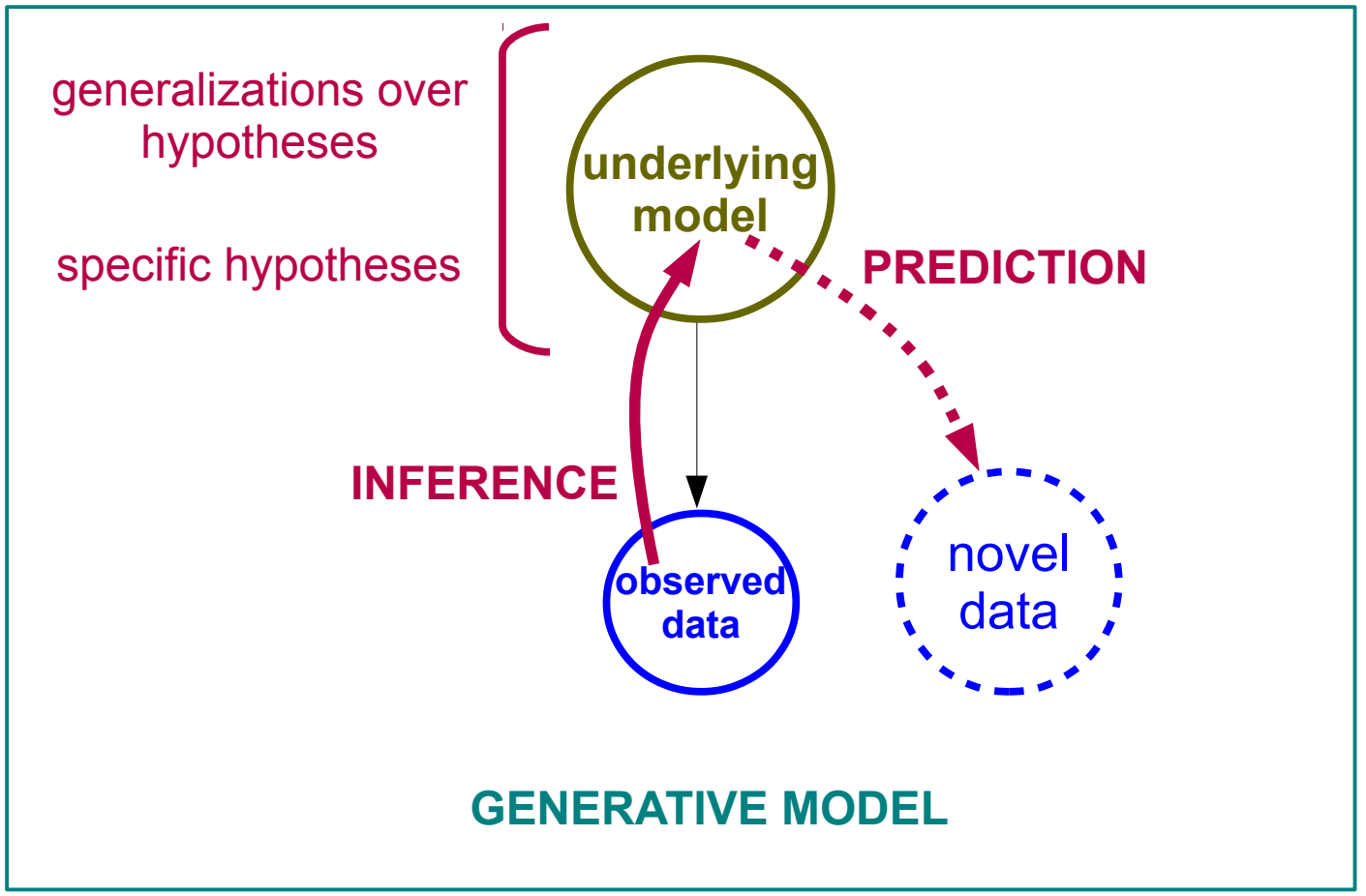
Explanation:
both e and ɛ map onto a single /e/ category in L1

# Questions

- **Less studied**: potential *benefits* of L1
  - known direct benefits: when L1 & L2 have very similar properties, e.g. cognates (e.g., Tréville, 1996)
- General learning literature:
  - people are very good at making generalizations and predictions based on current knowledge (e.g., Tenenbaum et al., 2011)
- **Our questions:**
  - Is there any evidence of generalization from L1 about L2 phonetic categories?
  - If so, how do these generalizations interact with the distributional cues from L2 input?

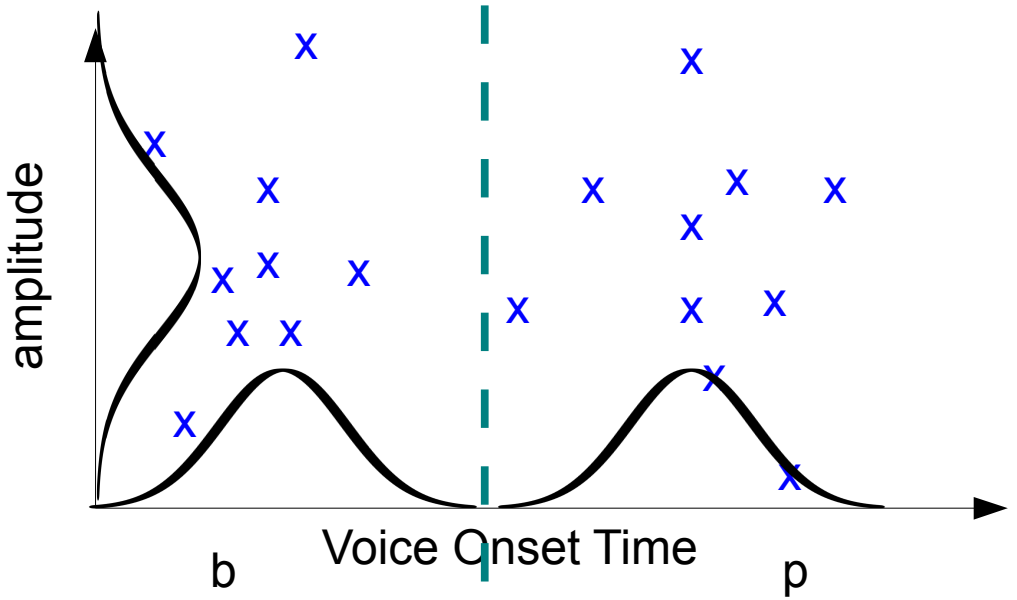
# Framework

- Learning as hypothesis construction and testing
  (e.g., Gerken 2010, Tenenbaum & Griffiths 2001, Tenenbaum et al. 2011, Xu & Tenenbaum 2007)



# Proposed model: assumptions

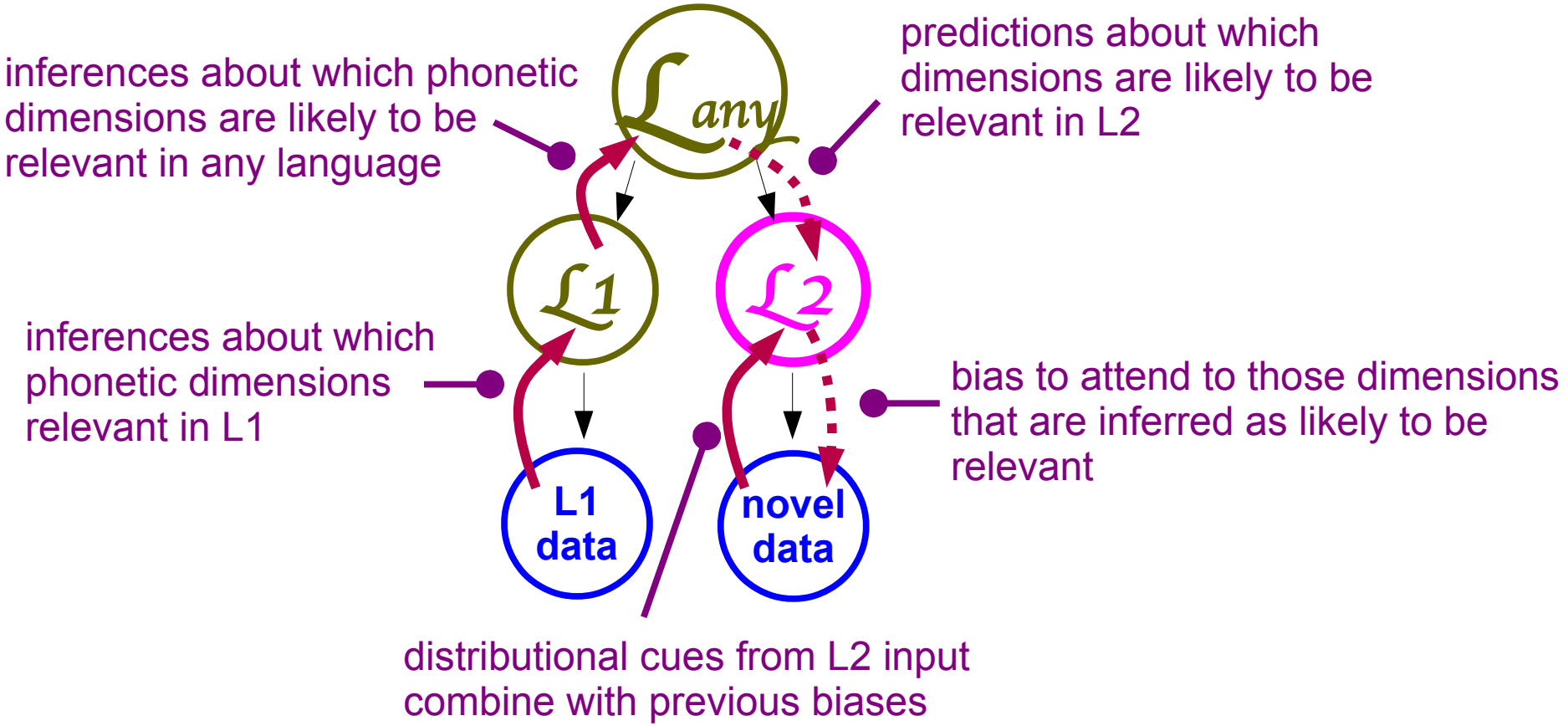
- Sounds vary along multiple acoustic-phonetic dimensions




- Categorization of sounds achieved through learning from distributional information to:
  - attend to relevant phonetic dimensions
  - disregard other dimensions

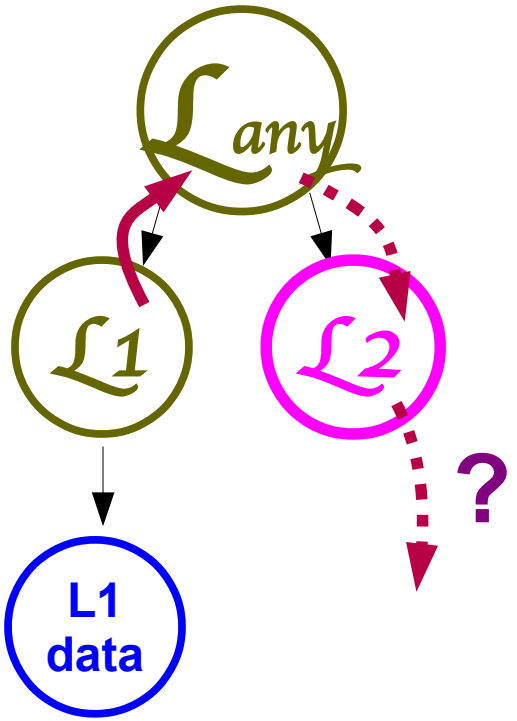
  (Kuhl et al., 1992; Jusczyk, 1992; Maye, Werker, & Gerken, 2002; Nittrouer & Miller 1997)

# Proposed model: phonetic category learning



# Pająk, 2010; in prep.

- Is there generalization from L1 in non-native speech perception for naïve listeners?

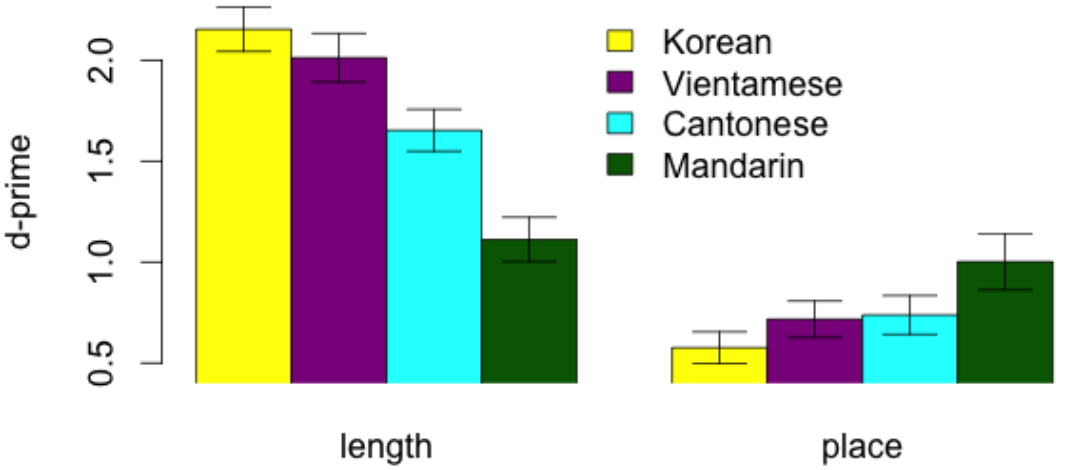


# Pająk, 2010; in prep.

- Task: discriminating between words in a novel language
  minimal pairs differing in:
  - consonant **length** kena - kenna
  - alveolo-palatal vs. retroflex **place** (w/o following vowel transition cue) keɕa - keʂa

| Participants (n=96) | **length** short and long sounds | **place** alveolo-palatal and retroflex sounds |
|---|---|---|
| **Korean** speakers | √ | |
| **Vietnamese** speakers | √ | |
| **Cantonese** speakers | √ | |
| **Mandarin** speakers | | √ |

only for vowels!

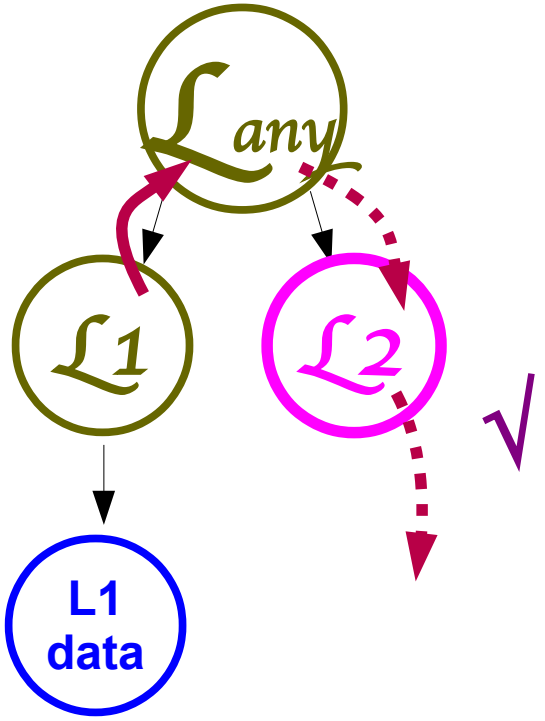



**There is generalization across segments – even from vowels to consonants!**

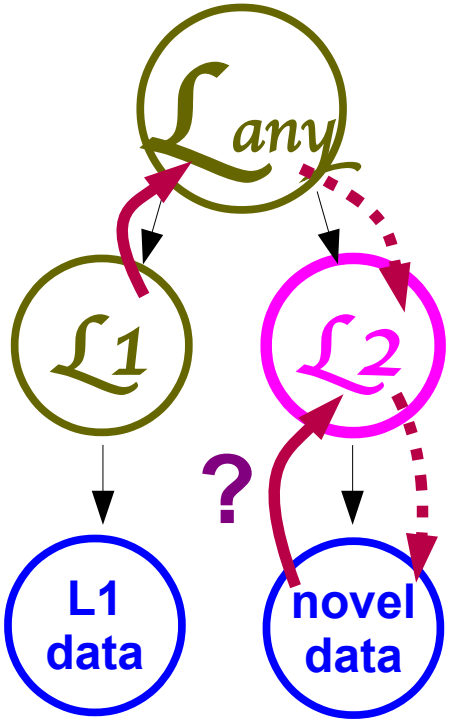
# Experiment

- **Pająk, 2010; in prep.:** evidence of generalized L1 biases in perception of novel sounds
- **Current experiment:** do L1 biases change learners' interpretation of distributional cues from novel language input?

**Pająk, 2010; in prep.**

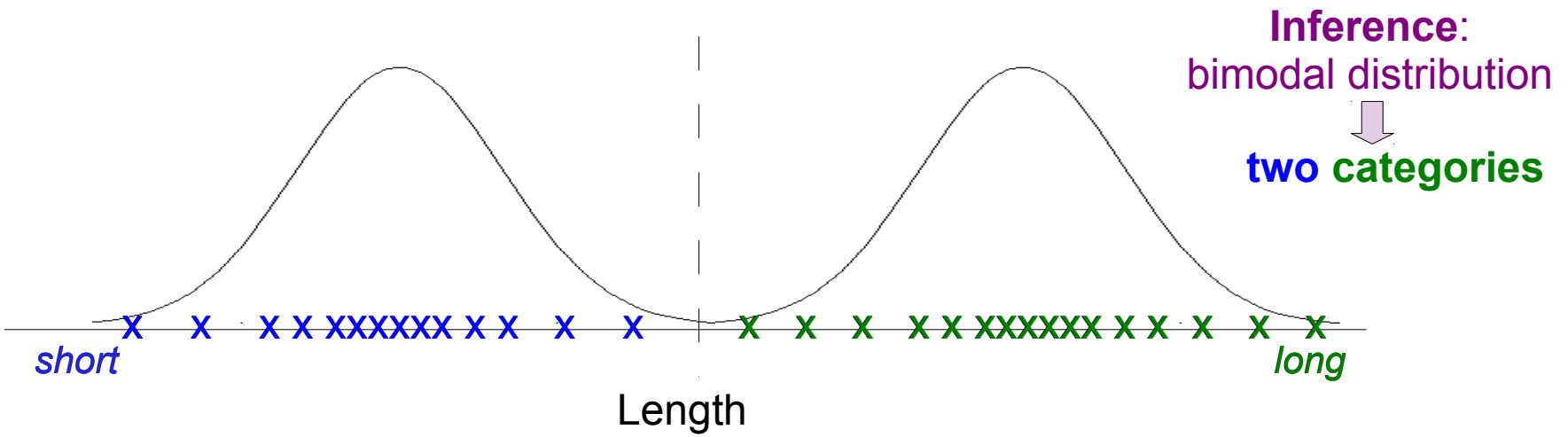


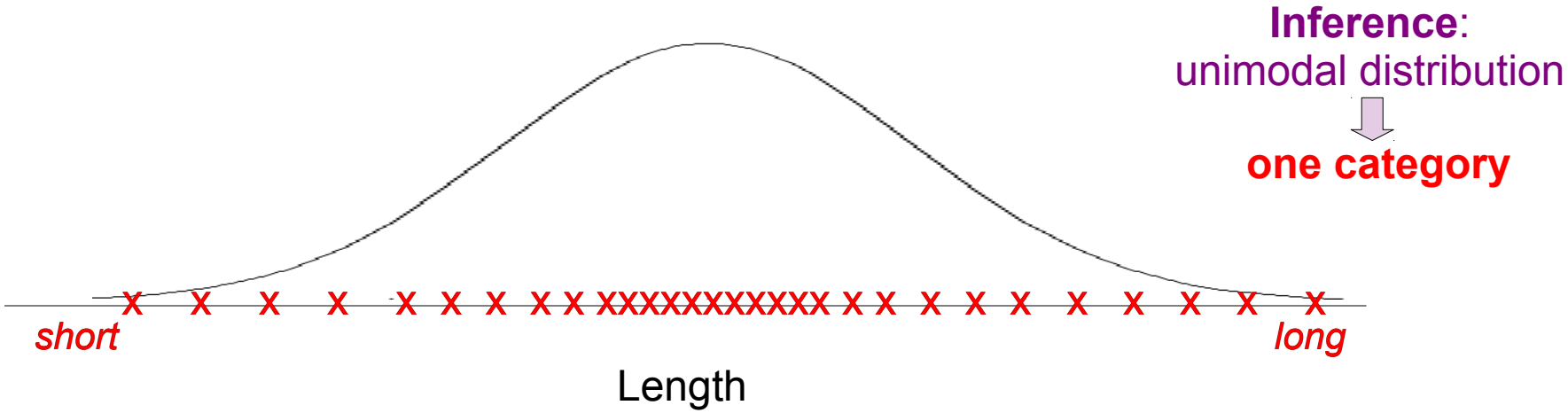

**Current experiment**



# Experiment: paradigm

- Distributional learning paradigm (Maye & Gerken 2000, Maye, Werker, & Gerken 2002)
  - a tool to probe learners' interpretation of distributional cues in novel input



# Experiment: paradigm

- Distributional learning paradigm (Maye & Gerken 2000, Maye, Werker, & Gerken 2002)
  - a tool to probe learners' interpretation of distributional cues in novel input

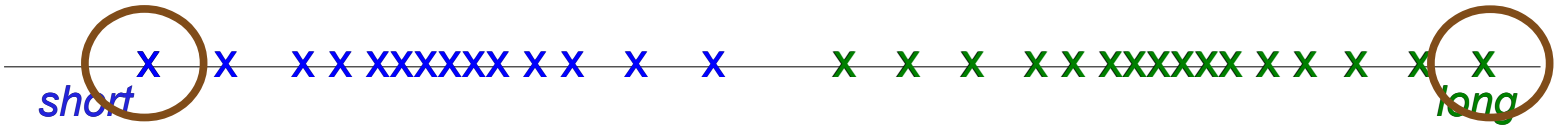
same or different?

→ **DIFFERENT**

**Inference**:
bimodal distribution

⇩

**two categories**




Length

---

same or different?

→ **SAME**

**Inference**:
unimodal distribution

⇩

**one category**


Length

# Experiment: instructions

- You'll be learning (sounds from) a new language.
- First, you'll listen to words from that language.
- Then, you'll hear pairs of words and, based on what you learned, decide whether these are two different words or the same word repeated twice.
- The same word can be pronounced a bit differently (e.g., with different intonation) – follow your intuition in deciding what counts as 'different' in this language.

# Experiment: ppts & overview

- 58 Korean speakers & 53 Mandarin speakers

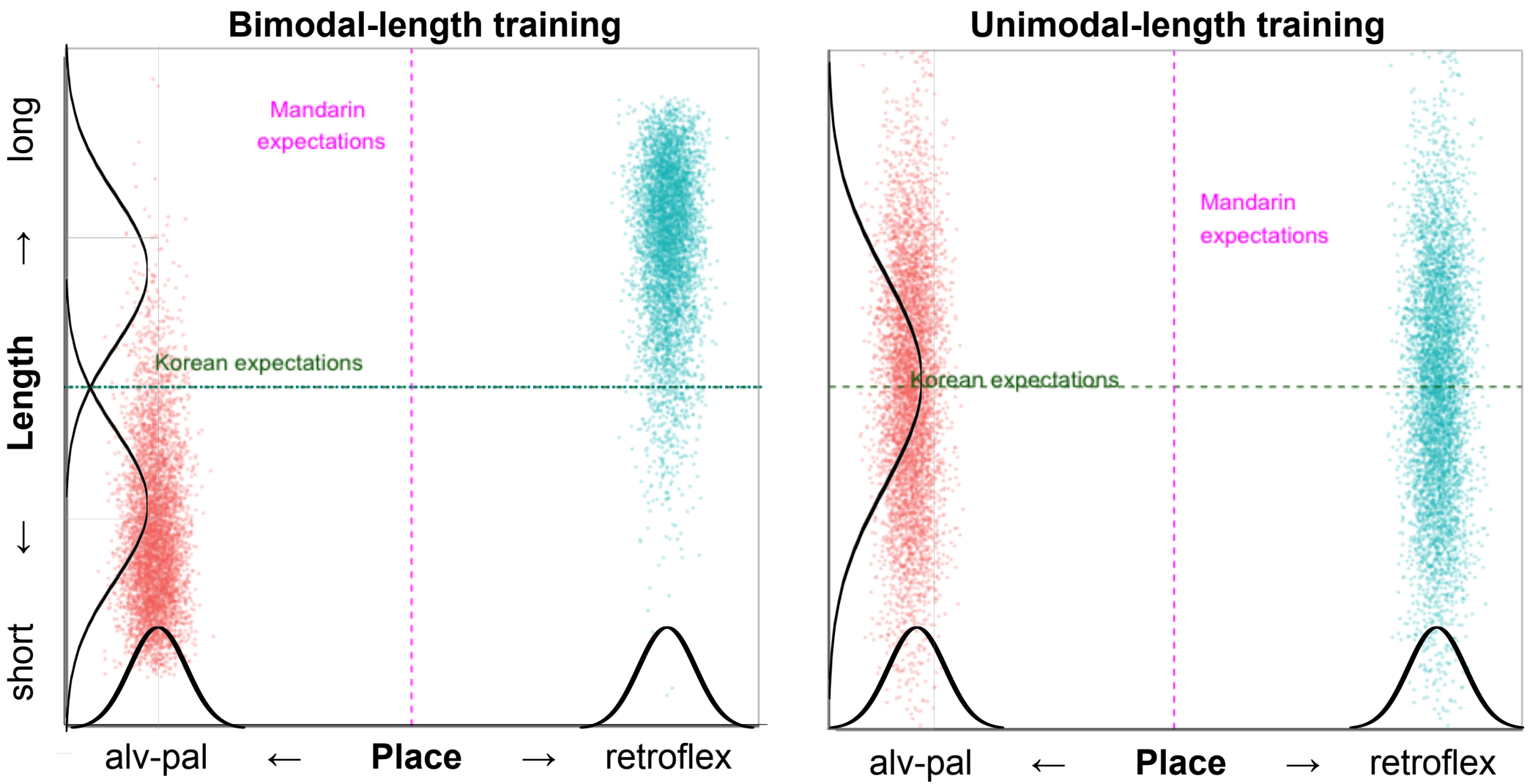
| | **length** short and long sounds | **place** alveolo-palatal and retroflex sounds |
|---|---|---|
| **Korean** | √ | |
| **Mandarin** | | √ |

**example** → [n] - [nn] [ɕ] - [ʂ]

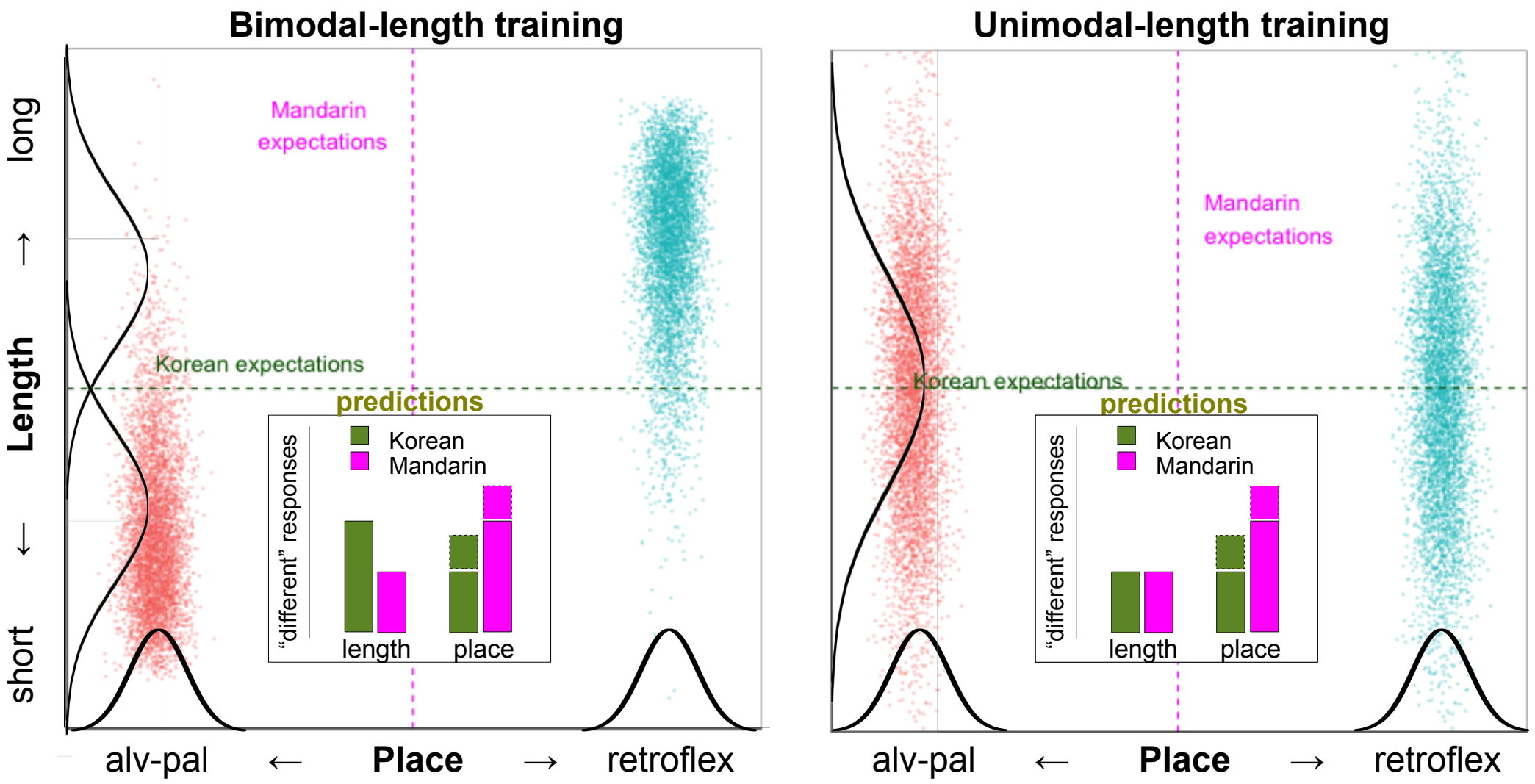
(Pinyin: **x** and **sh**)

- Exposing participants to evidence suggesting a novel place contrast, with a less reliable length cue

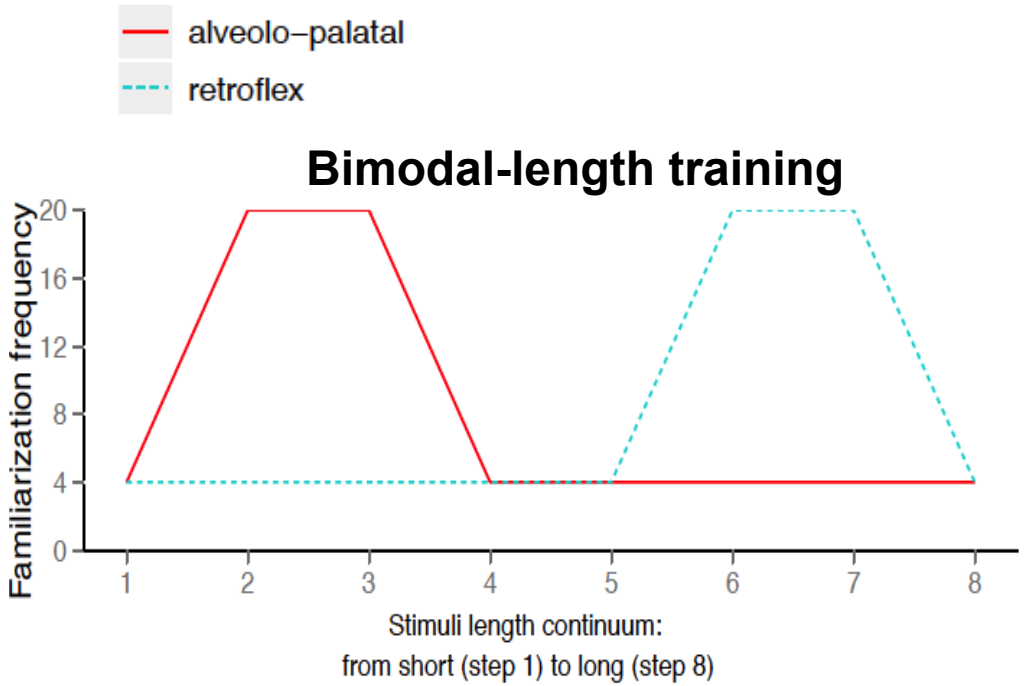
# Experiment: novel language input

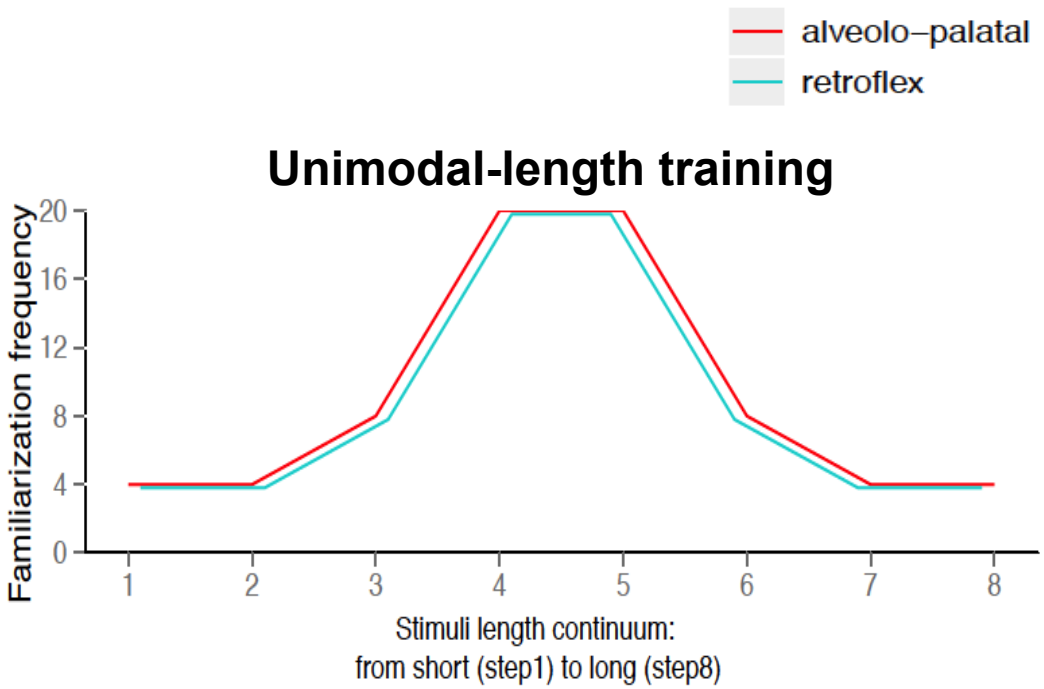


# Experiment: novel language input



# Experiment: materials & design

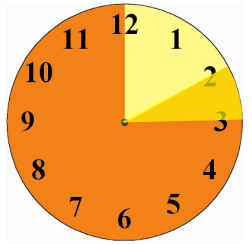



| **Training** | **x** | **Language** | **x** | **Contrast** |
|---|---|---|---|---|
| bimodal-length / unimodal-length | | Korean / Mandarin | | length / place |

**+ baseline condition**
(training on fillers with no variability in length)

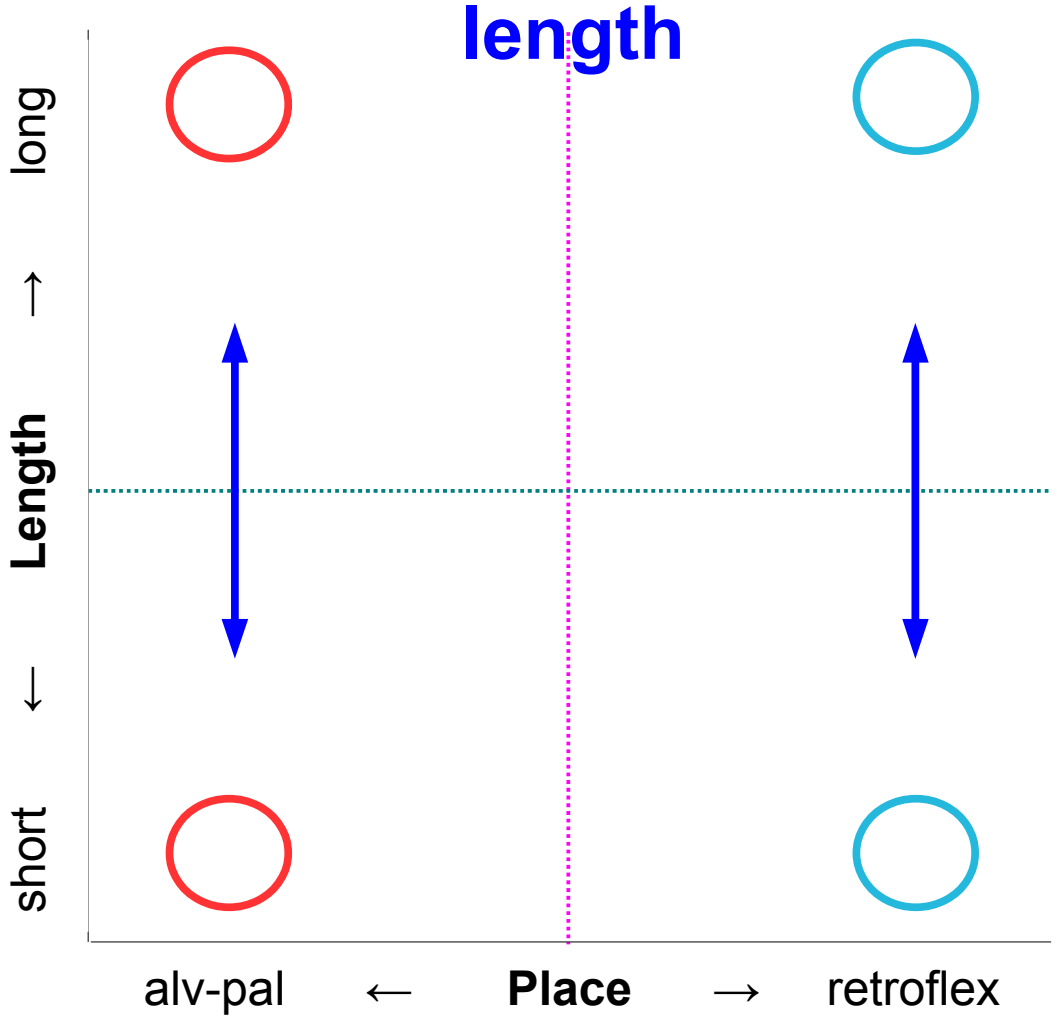
# Experiment: training

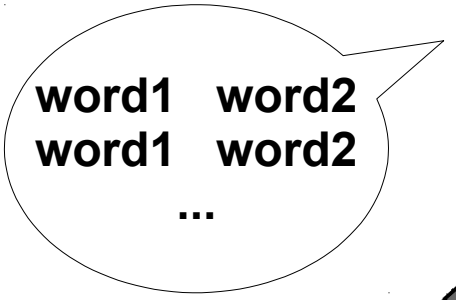


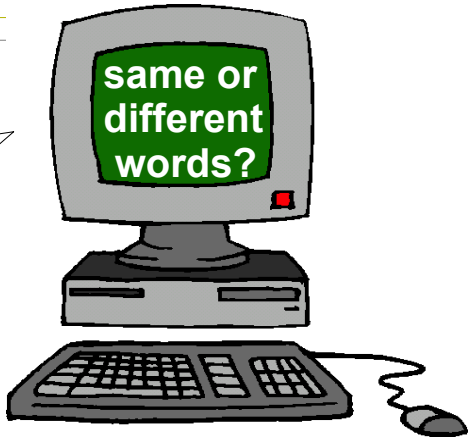

[aʂa]$_{step1}$
[aʂa]$_{step4}$
[aʂa]$_{step7}$
[aka]
...

# Experiment: testing

**Identical testing for all participants:**
→ length & place contrasts



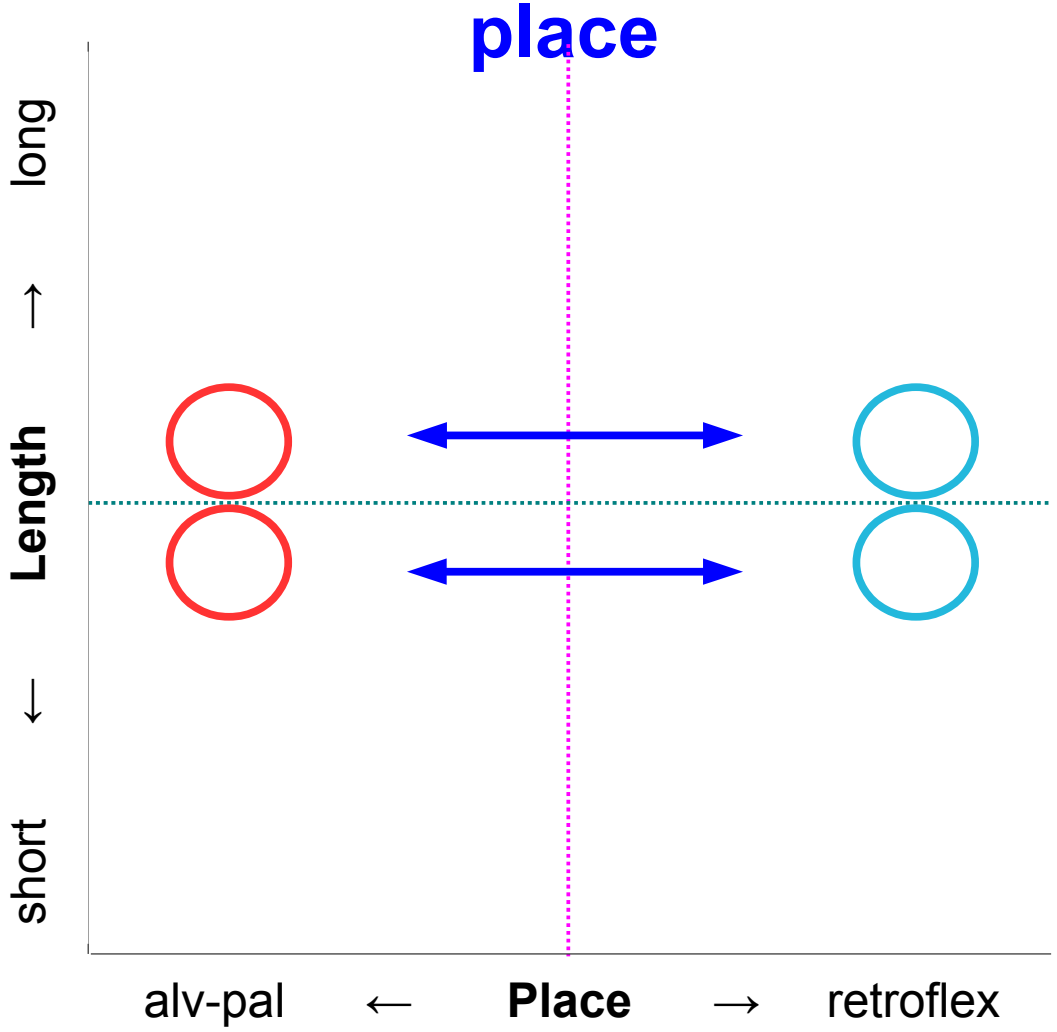



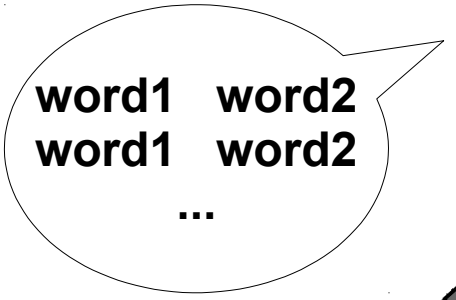



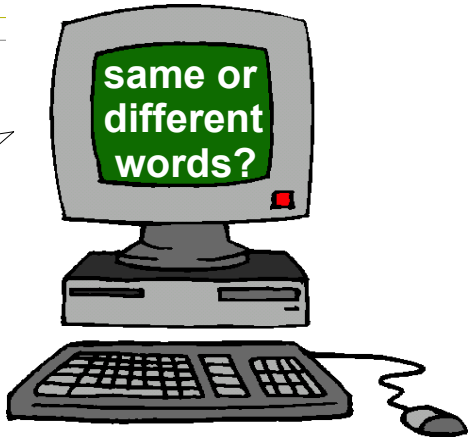

alveolo–palatal
retroflex

# Experiment: testing

**Identical testing for all participants:**
→ length & place contrasts



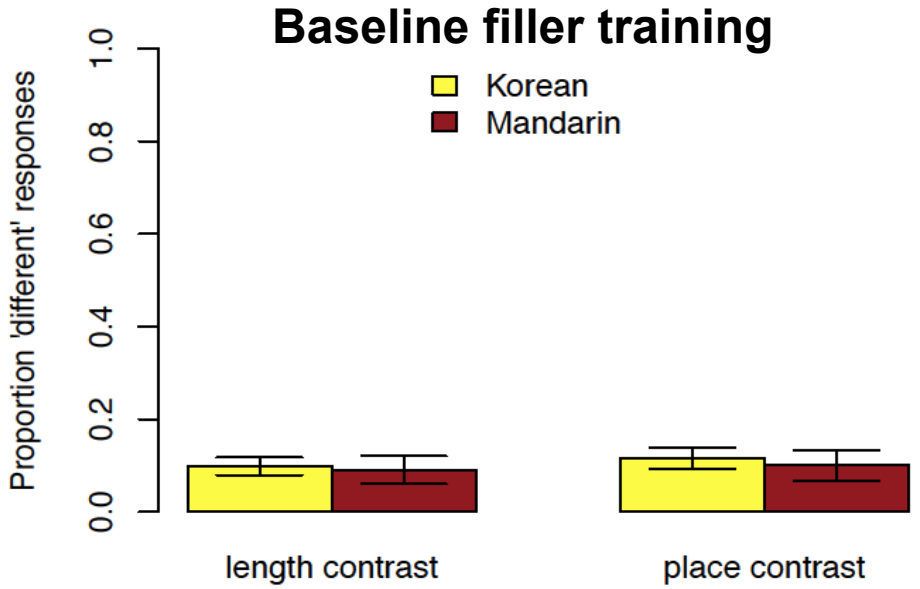



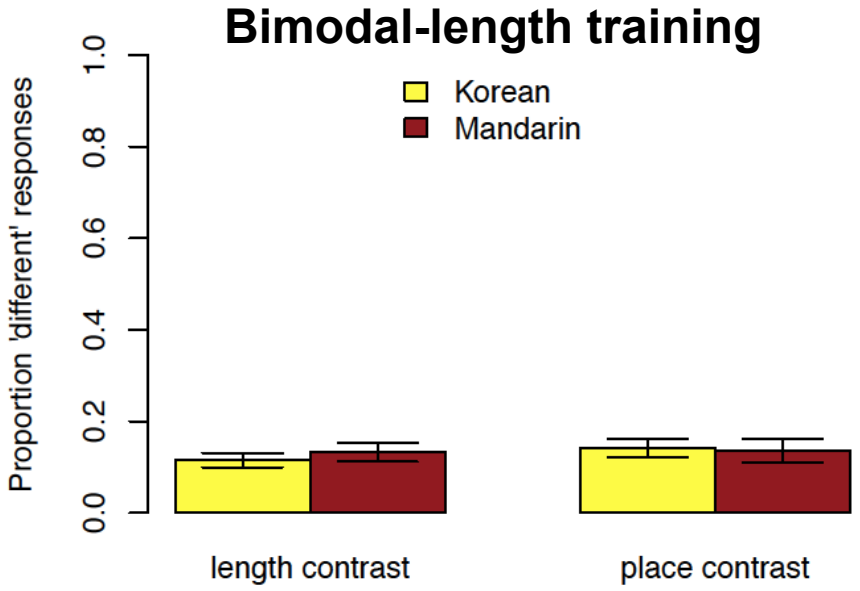



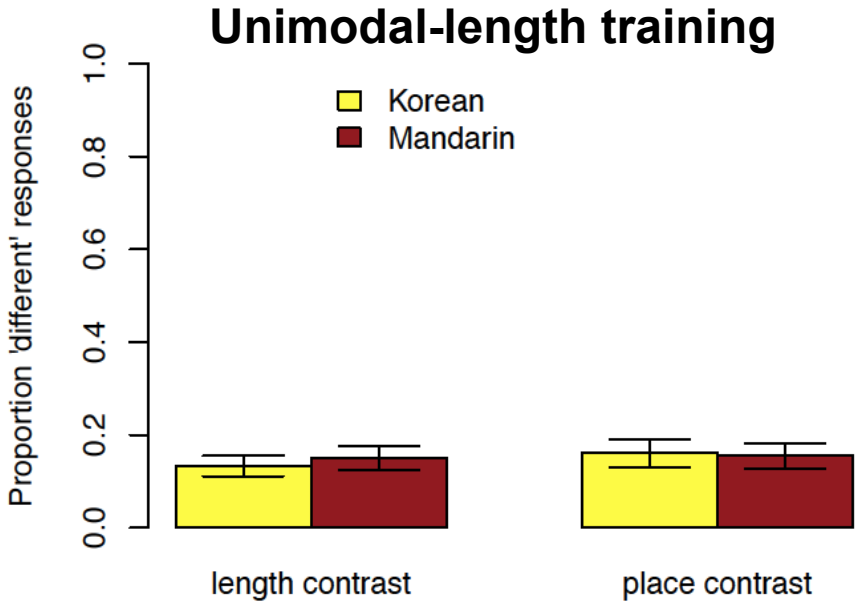

alveolo–palatal
retroflex

# Experiment: results

**SAME trials**

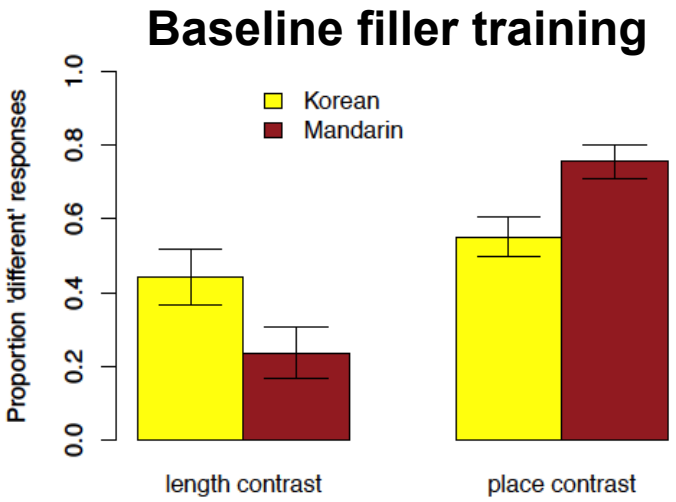


# Experiment: results

## DIFFERENT trials




**Baseline (initial biases)**

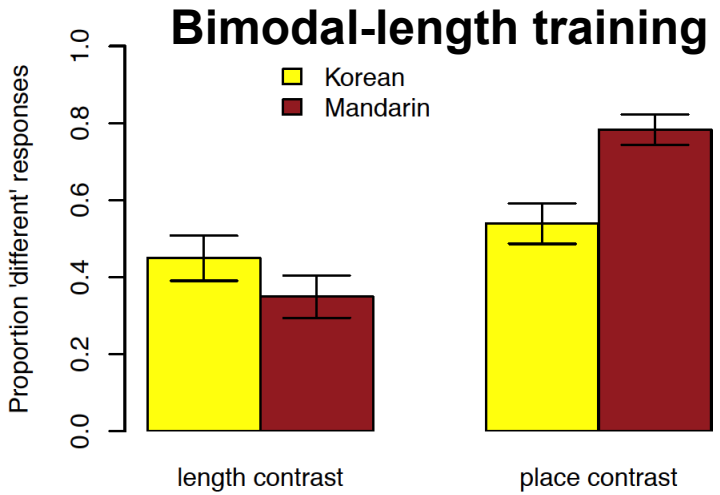
- Korean ppts biased in favor of **length**
- Mandarin ppts biased in favor of **place**

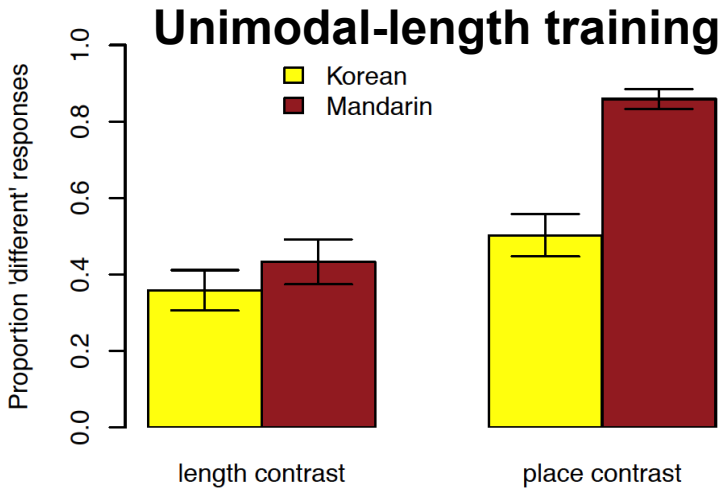
**Bimodal-length**

- relative biases the same as in baseline
- Mandarin ppts go up on **length**

**Unimodal-length**

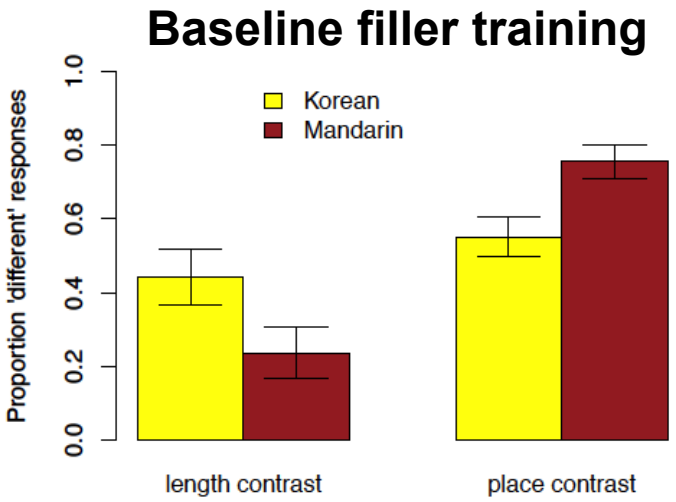
- the relative bias for **place** the same as in baseline
- on **length**, Korean ppts go down, and Mandarin ppts go up



# Experiment: results

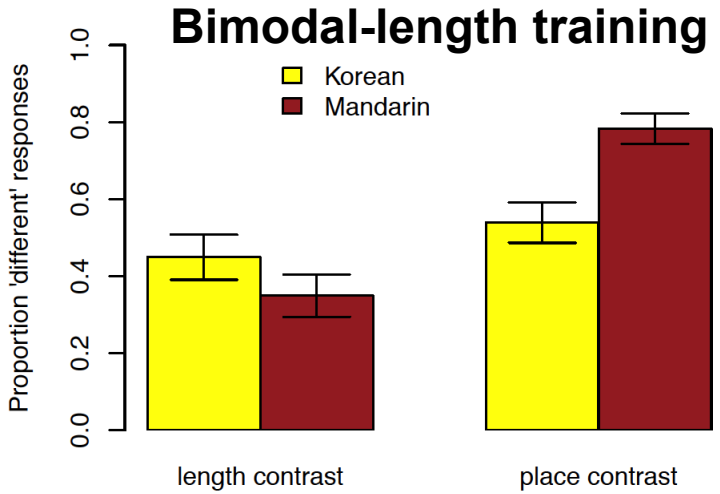
## DIFFERENT trials

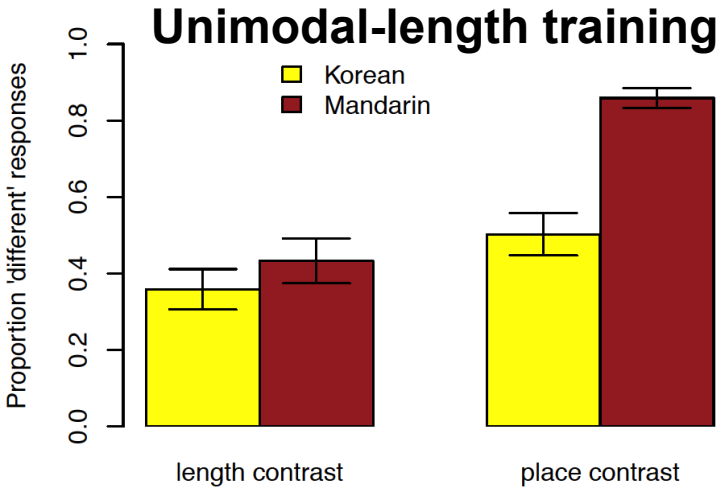



**What does it mean?**

**Data not entirely clear, but:**

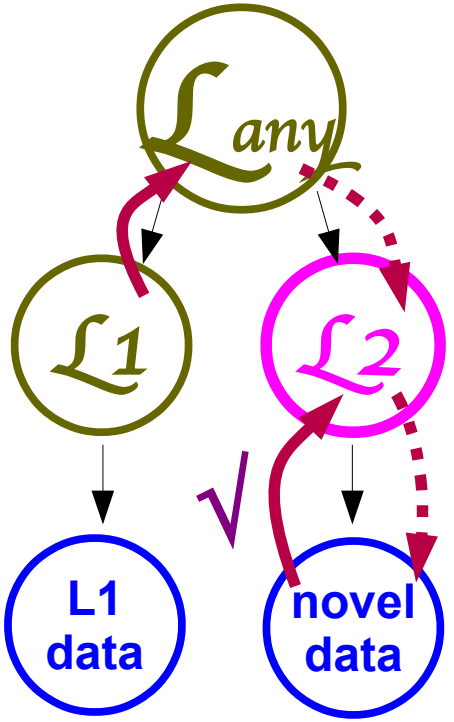
- initial biases – as predicted
- training didn't override these biases for **place**
- Korean ppts were able to properly interpret distributional cues for **length**
- for Mandarin ppts, mere exposure to variability along the length dimension (in either training) increased sensitivity to **length** ← not able to use distributional cues?



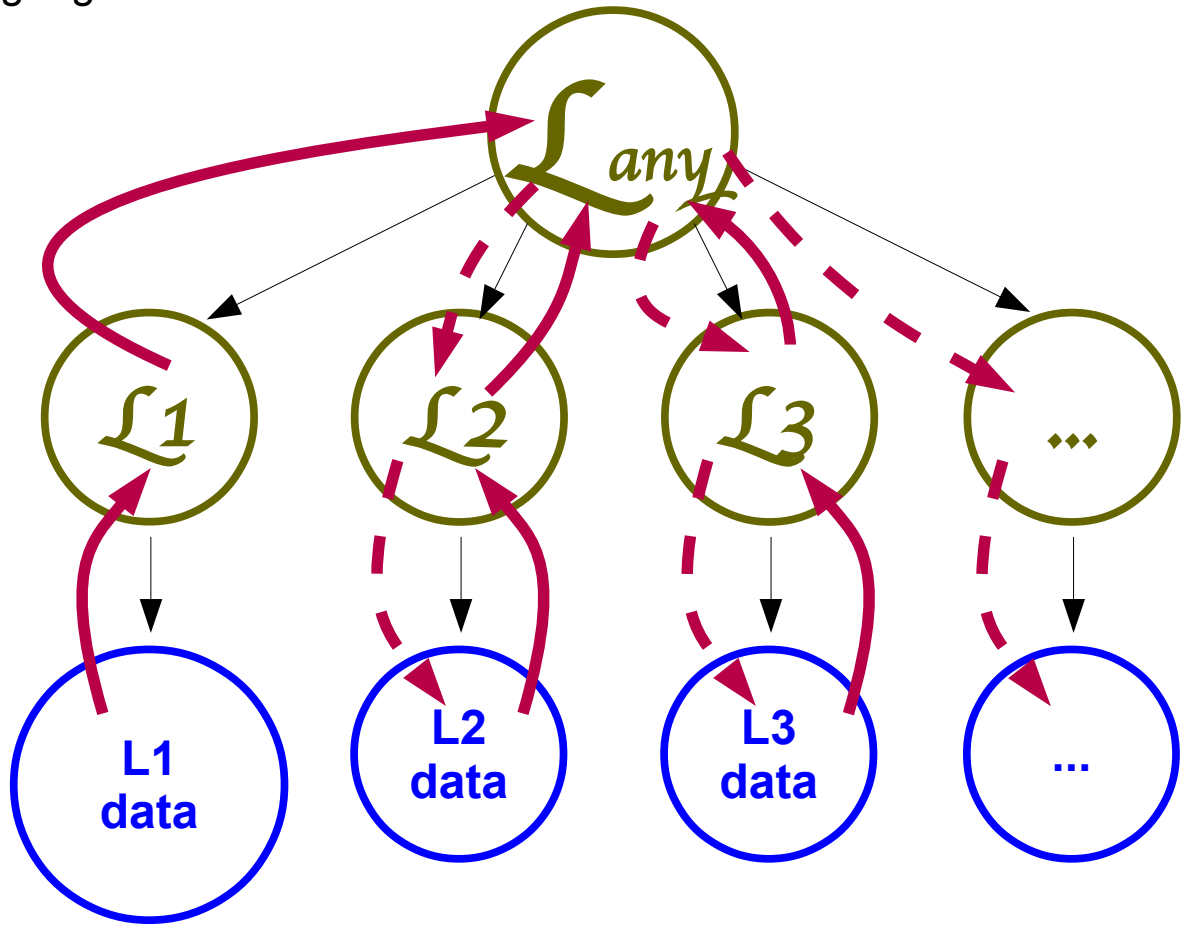
# Experiment: summary

- L1 biases were reflected in categorization of novel sounds
- But categorization was also affected by short exposure to distributional cues in a novel language
- Distinct L1 biases led to different interpretations of distributional cues



# Extending the model beyond L2

- The model provides a straightforward way of combining inferences from multiple languages
- This captures the intuition that learning each additional language becomes easier – because learners' $\mathcal{L}_{any}$ inferences become sharper
- The model can make predictions about how learners' inferences change after learning each additional language



UC San Diego

# Thank you


***Discussion***:

Eric Baković, Klinton Bicknell, Sarah Creel, Victor Ferreira, Tamar Gollan, Carolyn Quam, Sharon Rose, *Computational Psycholinguistics Lab, Language Acquisition & Sound Recognition Lab, San Diego Phonology Interest Group*

***Research assistants***:

K. Michael Brooks, Hannah Byers-Straus, Nicole Pyon

***Funding***:

NIH Training Grant T32-DC000041 from the Center for Research in Language at UC San Diego

UC San Diego Academic Senate Grant