Success and failure of new speech category learning in adulthood: A Hebbian account

Sara Parvanezadeh Esfahani

Overview

- Background
  - Perceptual narrowing of sounds
  - Bimodal vs unimodal distribution of sounds
- Testing /l/, /r/ contrast (work done by McCandliss, Feiz, Protopapas, Conway, & McClelland, 2002)
- Modeling McCandliss study (work done by Vallabha & McClelland, 2007)

Introduction

Two facts:

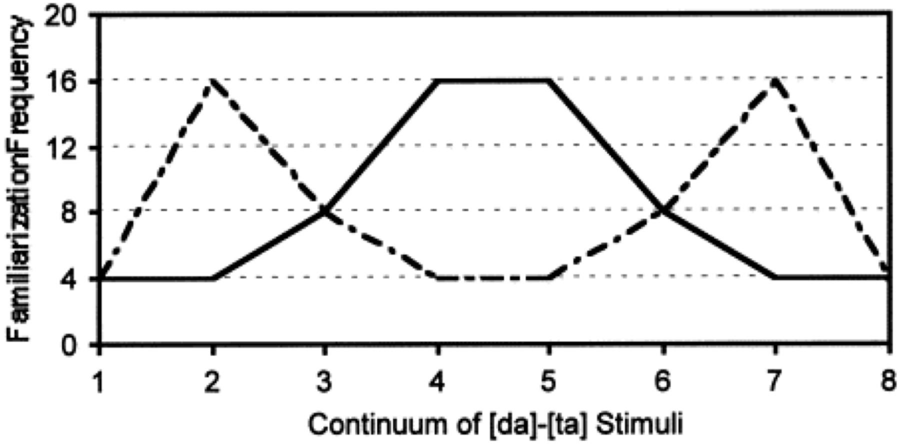
First, early in life, infants are capable of discriminating among many, if not all, the phonetic units of the world's languages, as established by the phenomenon of **categorical perception**.

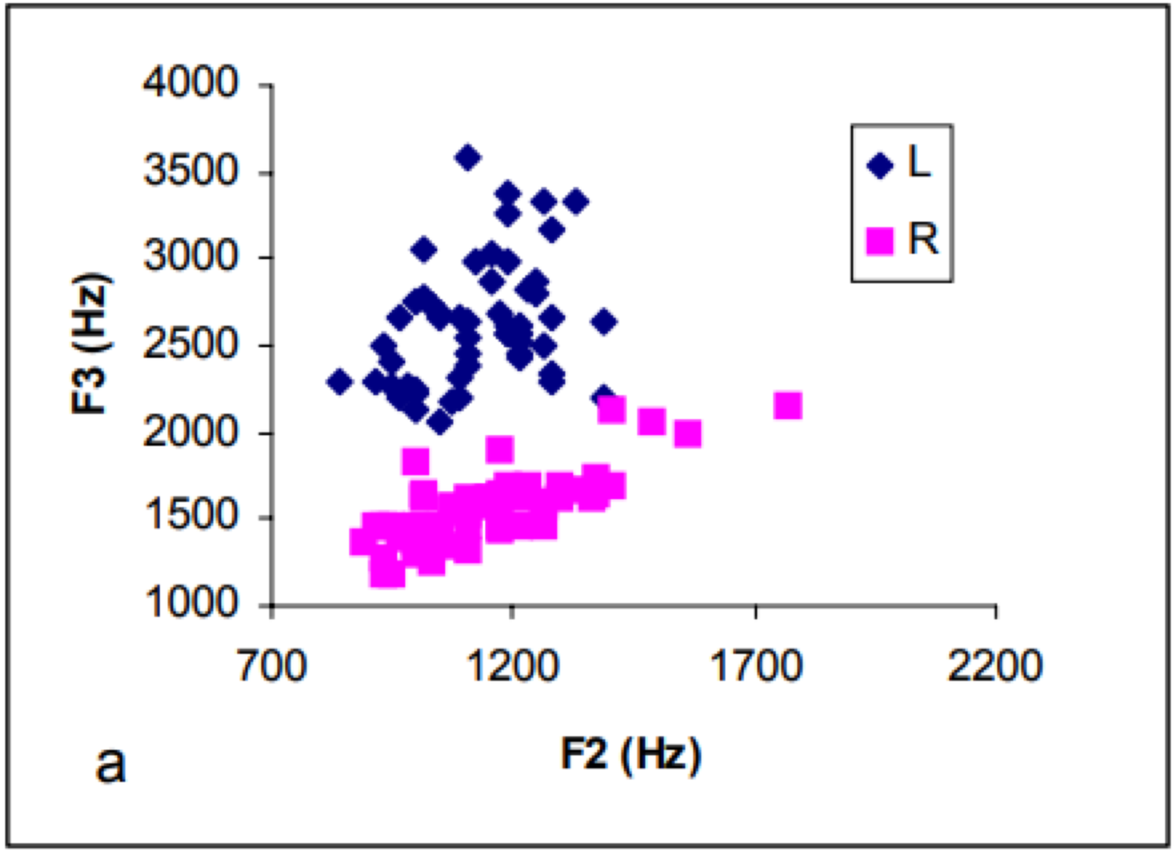
Second, by adulthood, universal phonetic capacity is no longer in place – nonnative phonetic discrimination can be very difficult.

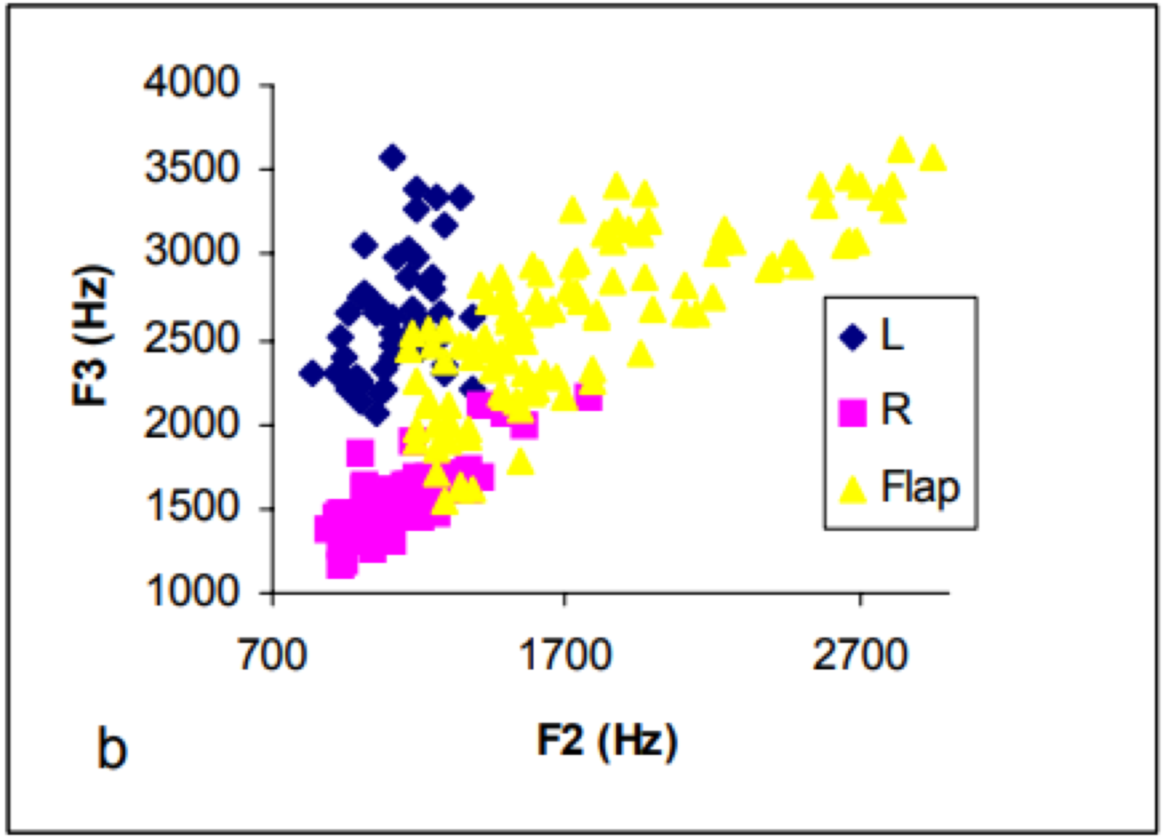
Perceptual Narrowing

- As infants gain experience with their native language they tend to show reduced sensitivity to many non-native phonemes and show an increased ability to be able to discriminate native-language phonemes
- This process has been referred to as **perceptual narrowing** or native language attunement.
- Evidence of perceptual narrowing is typically observed around 10 to 12 months of age for consonants and somewhat earlier for vowels.

- One source of information that infants might be able to exploit is how sounds are distributed in acoustic space.
- American English /r/ and /l/ are produced with a bimodal distribution along a particular acoustic dimension (i.e., the frequency of the third formant (F3)).
- In natural speech exemplars of /r/ cluster together at lower values of F3, whereas exemplars of /l/ clusters together at higher values of F3, with only a small degree of acoustic overlap between productions of /r/ and /l/. In contrast, in natural Japanese productions there is a unimodal cluster of sounds, called a flap, that overlaps with American English /r/ and /l/.







Comparison of English speakers' productions for /L/ and /R/ (a) overlaid with Japanese speakers' production of the flap consonant (b), from Lotto, Sato, & Diehl, 2004.

McCandliss et al. (2002)

- Speech perceptual learning may depend on an unsupervised Hebbian learning process.
- According to Hebbian account, the acquisition of one's native language can lead to a strong tendency to treat certain nonnative stimuli as the same – may account for the difficulty Japanese adults have in acquiring the English contrast between /r/ and /l/.

Participants – Adults native speakers of Japanese

Method

Stimuli - Recorded /r/ and /l/ words by native speaker of English- acoustically manipulated to generate continua from exaggerated /l/ to exaggerated /r/

Four conditions

- Fixed vs adaptive training stimuli
- Presence vs absence of the feedback

Fixed - heard the same two intermediate tokens of /r/ and /l/ and had to label them as “r” or “l”.

Adaptive - first heard exaggerated tokens that were easily identifiable
After identifying 8 successive tokens correctly, /r/ and /l/ tokens moved closer to each other

Feedback conditions - subjects received immediate visual feedback on the correctness of each response.

No-feedback - the visual feedback was omitted.

- Eight subjects in each condition, 4 trained with a *road-load* continuum, and 4 with *rock-lock* continuum.

- Training was conducted over 3 days with daily sessions of 500 trials, with half of the subjects given an additional 3 sessions of training.

Learning Measures

1. Performance on probe stimuli that were periodically presented to the subject during the training,
2. Categorization on trained and untrained R/L continua (identification task)
3. Same–different discrimination of pairs with fixed interstimulus distance (“slide test”)
4. Same–different discrimination of pairs with increasing interstimulus distance (“expand test”).

Results

1. Adaptive training without feedback was effective in inducing perceptual learning.
2. Feedback: fixed training was better.   No-Feedback: adaptive training was better.
3. The learning was not stimulus specific. After training, subjects showed improved categorization on the untrained R/L continuum.
4. The improvement in R/L classification was paralleled by an increase in discriminability for stimulus pairs.
5. Some individual differences in no-feedback training
6. A strong bias toward labeling the stimuli as “L”
7. Learning in the fixed-without-feedback condition was very slow, hardly improving between the pre- and post-training categorization curves.

Two prediction based on Hebbian account:

1. If subjects treat non-native stimuli as native stimuli, they should NOT benefit from training.

   Evidence: group with fixed stimuli and no feedback showed no evidence of learning after 3days of training.

2. If subjects have the ability to discriminate non-native stimuli, they should benefit from training even with no feedback.

   Evidence: group with adaptive stimuli and no feedback showed considerable gain in both identification and discrimination.

- Results supported Hebbian account that eliciting separate representations can induce category learning.
- Although learning can take place without feedback, the speech perceptual system can take advantage of feedback when it is available.

Vallabha & McClelland (2007)

- Developed a network in which speech categories are attractors implemented through interactive activation and Hebbian learning.

- A multilayer network with
  - bidirectional inhibitory connections within a layer,
  - bidirectional excitatory connections between layers, and
  - Hebbian update of the between-layer connections.

Model architecture

Three layers

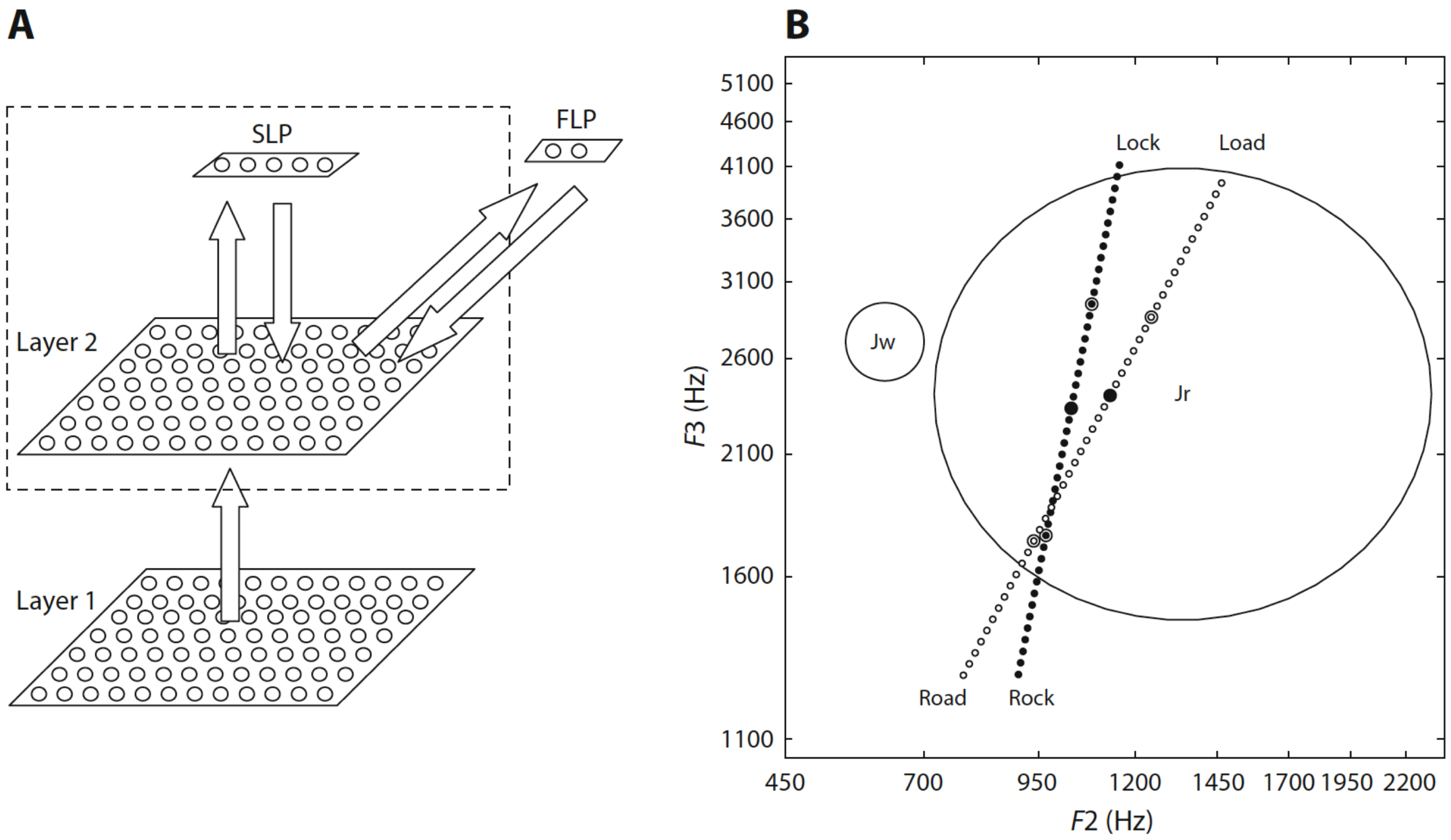
- L1 and L2 – implementing a topographic map over two acoustic features that are most relevant for distinguishing /r/ and /l/
- L3 – detecting clusters in the L2 activity distribution

Two phases

1. "Japanese environment" – the network is exposed to unlabeled instances of the Japanese alveolar flap and velar approximant
2. "R/L training" – the Japanese-trained network is trained to categorize AE /r/ and /l/ using the training conditions in McCandliss et al. (2002) - fixed stimuli with and without feedback and adaptive stimuli with and without feedback.

(B) The input space of the network

- The two large circles show distribution of the two Japanese categories.
- The small circles are the McCandliss training continua.



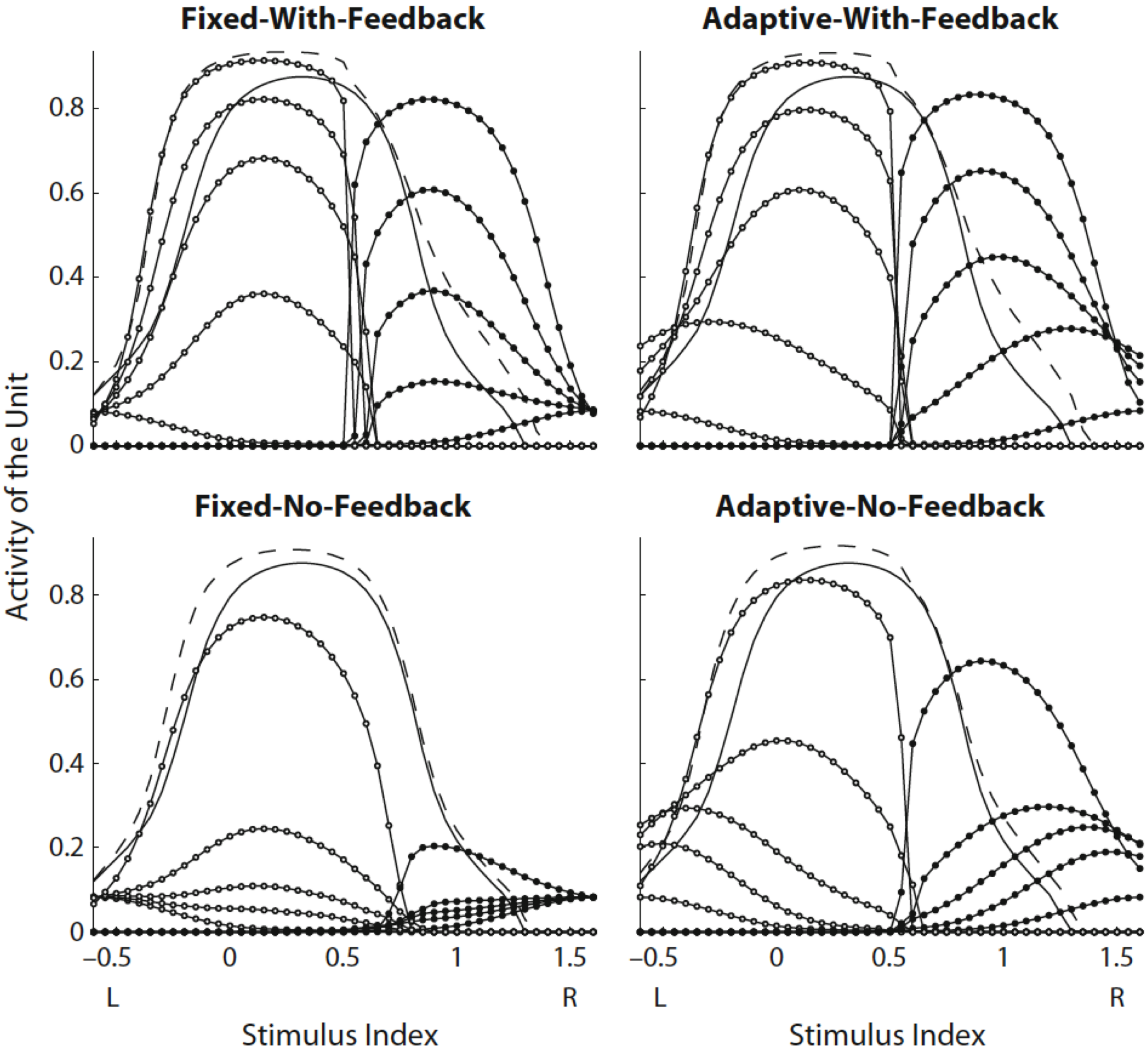
(A) The network architecture

- L1 and L2 are 20 x 20 grids of units
- L2 is bidirectionally connected to the slow learning pool (SLP; with 5 unit, used in phase 1 training) and to the fast learning pool (FLP; with two units, used in phase 2 R/L training).
- All between-layer connections are excitatory. Each unit in L2, SLP and FLP excites itself and inhibits all the other units in its layer.

Results & Discussion

The activity of the FLP units for stimuli on the *lock-rock* continuum at four different stages of training (immediately after pretraining and after 1000, 1500, 2000, and 3000 updates), showing the response of

- the L unit (empty circles),
- the R unit (filled circles),
- the Japanese flap unit before R/L training (unmarked solid line), and
- after R/L training (dashed line)



Results & Discussion

- The L-side stimuli are more prone to fall into the Japanese flap attractor than are the R-side stimuli.
- An L-side stimulus, therefore, activates the flap attractor, which recurrently excites L2, which increases L2 activity levels, which causes faster Hebbian learning.
- Thus, the L-unit in the FLP tends to
    1. learn faster than the R-unit,
    2. develop the same mapping as the Japanese flap unit, and
    3. take over parts of the stimulus space corresponding to the English /r/, causing an overall bias toward L-labeling.

- Overall, the model is able to capture many of the results from McCandliss et al. (2002).

1. The effectiveness of adaptive-without-feedback training
2. The relative time course of learning in different feedback and stimulus conditions
3. A strong bias toward L-labeling in posttraining categorization
4. Transfer of training to a new stimulus continuum
5. An acquired increase of the ability to discriminate at the category boundary

- The model provides a possible basis for some aspects of the individual differences seen in learning.
- It gives a mechanistic account of how perception is influenced by prior experience, how difficulty in discrimination may be related to difficulty in second language learning, and how outcome information and exaggerated input may be used to facilitate learning.

Thank You!