Timecourse of Bilingual Phonologies
Melissa Baese-Berk, University of Oregon
mbaesebe@uoregon.edu
Introduction

► I speak really quickly
  ▪ Please slow me down

► Office hours today & Monday ~11-12:30 in the library Starbucks
What makes someone bilingual?
What makes someone a bilingual vs. a non-native?
**Second Language Phonetics & Phonology**

- What comes into contact with what?
  
  What do you know when you know a language?
  (What do you **not** know when you encounter a non-native language?)

- What is the nature of non-native language sound structure acquisition?
  
  How does the target language (L2) “map onto” the native language (L1)?

- What training facilitates non-native language sound structure acquisition?

- Why is foreign-accented speech hard to understand? Or, is it?
Second Language Phonetics & Phonology

• What comes into contact with what?  
  What do you know when you know a language?  
  (What do you not know when you encounter a non-native language?)

• What is the nature of non-native language sound structure acquisition?  
  How does the target language (L2) “map onto” the native language (L1)?

• What training facilitates non-native language sound structure acquisition?

• Why is foreign-accented speech hard to understand?  
  Or, is it?
Category Learning

• Two types of perception:

![Continuous Perception Chart]

![Categorical Perception Chart]
Categorical Perception

- Consonants are tend to be perceived categorically (Liberman et al., 1957)
- The difference between /ba/, /da/, and /ga/ is the frequency of the onset of the second formant
- Here’s a continuum of 14 steps from /ba/ to /da/ to /ga/ made from synthesized speech
- Even though these sounds vary from one another minimally and continuously, listen and see whether you hear the changes between the syllables occur suddenly

![Graph showing frequency and time comparison for [ba], [da], and [ga]]
Not so simple!

Lots of things influence category boundaries!

- Sequencing, range, and number of stimuli presented (see also Schouten et al. 2003)
- Selective adaptation
- Cue weighting/trading relations
- Segmental context effects
- Speaking rate effects
- Speaker normalization
- Semantic and syntactic effects
Side note

- Categorical perception may not be so categorical after all ...
- McMurray et al. 2002
Questions?
Sound inventories: phoneme categories in the system

Norwegian

Afrikaans

Zulu

English

Mandarin
Sound inventories: phoneme categories in the system

How do language users **produce** sounds that are not in their native language inventory? (e.g. in English borrowings such as Bach, Chanukah, junta, etc.)

How do language users **perceive** sounds that are not in their native language inventory? (e.g. French ‘tu’ vs. ‘tous’ for English listeners, English ‘r’ vs. ‘l’ for Japanese listeners, etc.)
The UCLA Phoneme Segment Inventory Database:

Language inclusion criteria:
- 317 languages selected to cover the full range of language families

Segment analysis:
- Unit or sequence interpretation? E.g. diphthongs = 1 or 2 segments?
## Allophony 1: context induced within-category variability (Example 1: voiceless stops)

<table>
<thead>
<tr>
<th>English</th>
<th>Korean</th>
</tr>
</thead>
<tbody>
<tr>
<td>['pʰuːl]</td>
<td>[pʰul]</td>
</tr>
<tr>
<td>['spə:t]</td>
<td>['spə:t]</td>
</tr>
<tr>
<td>['tʰɒp]</td>
<td>[tʰ'al]</td>
</tr>
<tr>
<td>['stɒp]</td>
<td>[kʰ'cda]</td>
</tr>
<tr>
<td>['kʰɪlɪŋ]</td>
<td>['kʰɪlɪŋ]</td>
</tr>
<tr>
<td>['skəʊld]</td>
<td>['skəʊld]</td>
</tr>
</tbody>
</table>

| 'pool'   | [ə'pʰɪə]   |
| 'spurt'  | [də'spərt] |
| 'top'    | [ə'tʰæk]   |
| 'stop'   | [də'stɔɾ]  |
| 'killing' | [ə'kʰɪlɪŋ] |
| 'scold'  | [dɪ'skɔld] |

| 'appear' | 'despite' |
| 'attack' | 'destroy' |
| 'accrue' | 'discover' |

From *Introduction to English Phonetics and Phonology*, by Robert Carr
Allophony 2: context induced within-category variability (Example 2: vowel nasalization in CVN forms)

Redrawn from Marie-Josep Sole, Language and Speech, 1992 (1 talker from each language)

English: variation in nasal portion => intentional nasalization
Spanish: variation in oral portion, constant duration of nasalization => unintended speech production constraint
Phontactics: phoneme combinations and syllable shapes

Some English phoneme combinations:

‘pl’ = legal word beginning (e.g. play)
‘tl’ = illegal word beginning
‘pt’ = legal word ending (e.g. apt)
‘pt’ = illegal word beginning in English

What do non-native speakers do with these sequences in borrowings?
(For English speakers: Ptolemy, Ptah (from Greek), Tlatelolco (from Nahuatl); For Spanish speakers: special, sprite (from English).

Syllable shapes:

English: everything up to CCCVCCCCC
Japanese: V, CV, CVN
The process of nativization:
A window into linguistic sound structure across languages

• What happens if an English word with a cluster is borrowed into a language with more restricted phonotactics?

  Simplify the cluster by:
  • inserting a vowel to break up the cluster
  • deleting one (or more) of the segments

  E.g. Japanese:
  con.trol -> kon.to.roo.ru

  E.g. Finnish:
  glass - > la.si
Questions?
Second Language Phonetics & Phonology

- What comes into contact with what?
  What do you know when you know a language?
  (What do you not know when you encounter a foreign language?)

- What is the nature of non-native language sound structure acquisition?
  How does the target language (L2) “map onto” the native language (L1)?

- What training facilitates non-native language sound structure acquisition?

- Why is foreign-accented speech hard to understand? Or, is it?
What comes into contact with what?

*What do you know when you know a language?*

(*What do you **not** know when you encounter a foreign language?*)

What is the nature of non-native language sound structure acquisition?

*How does the target language (L2) “map onto” the native language (L1)?*

What training facilitates non-native language sound structure acquisition?

Why is foreign-accented speech hard to understand?

Or, is it?
Does your L1 sound system influence how you perceive and produce your L2?
Native Language Interference

- Interference in lexical categories
  - Trabajar vs. funcionar
- False friends
  - embarazada ≠ embarrassed
Native Language Interference

► Sound systems

► Languages differ in their sound systems
  - Spanish has “rr” vs. “r”
  - English has “ch” vs. “sh”

► Have to learn both what are possible sounds and what are not!
  - How sounds combine, etc.

► Perception of non-native contrasts
  - Influenced by similarities and differences between sound systems of L1 and L2
Same or Different?
Same or Different?
Key observations

• Decline over the life span in the ability to perceive and produce non-native speech in a native way

• Variable initial abilities and ease of acquisition of non-native sounds
Models of cross-language and second-language perception and production

- Current accounts
  - Knowledge of a first language system constrains the acquisition of a non-native system ("change resistance")
  - Adult learners interpret non-native sounds with reference to native categories ("L1 filter")
Models of cross-language and second-language perception and production

1. Native Language Magnet Model (NLM): Patricia Kuhl & Paul Iverson
2. Perceptual Assimilation Model (PAM): Catherine Best
3. Speech Learning Model (SLM): James Flege
The Native Language Magnet Model (NLM): Patricia Kuhl and Paul Iverson

The perceptual magnet effect

The vowel space “initial state”:

- The most frequently heard sounds modify perception.

- The distribution of speech sounds in the ambient language changes the “basic cuts” in the initial state of the auditory perceptual space.
The Native Language Magnet Model (NLM):
Patricia Kuhl and Paul Iverson

/r/ and /l/ ‘Perceptual maps’ for American English and Japanese speakers

Japanese

American

Physical Spacing of Stimuli

Iverson et al., 2003
1. A model that reflects the change from a universal perception to a language-specific perception.

2. Mapping of ambient language warps the acoustic dimensions underlying speech which produces a filter through which language is perceived.

3. Language specific filters make learning a new language as an adult more difficult: the mapping for the new language is different than the native language.
The Perceptual Assimilation Model (PAM): Catherine Best

Patterns of perceptual assimilation of non-native segments:

1. Assimilated to a native category (e.g. Norwegian /y/ to English /u/)

2. Assimilated as an uncategorizable speech sound (e.g. German /x/ for English listeners)

3. Not assimilated to speech (e.g. Zulu clicks for listeners from non-click languages)
The Perceptual Assimilation Model (PAM): Catherine Best

Patterns of perceptual assimilation of non-native contrasts:

<table>
<thead>
<tr>
<th>Pattern type</th>
<th>Expected discrimination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-Category Assimilation</td>
<td>Excellent</td>
</tr>
<tr>
<td>(TC Type)</td>
<td></td>
</tr>
<tr>
<td>Category-Goodness Difference</td>
<td>Moderate to very good</td>
</tr>
<tr>
<td>(CG Type)</td>
<td></td>
</tr>
<tr>
<td>Single-Category Assimilation</td>
<td>Poor</td>
</tr>
<tr>
<td>(SC Type)</td>
<td></td>
</tr>
<tr>
<td>Both Uncategorizable</td>
<td>Poor to very good</td>
</tr>
<tr>
<td>(UU Type)</td>
<td></td>
</tr>
<tr>
<td>Uncategorized versus Categorized</td>
<td>Very good</td>
</tr>
<tr>
<td>(UC Type)</td>
<td></td>
</tr>
<tr>
<td>Nonassimilable</td>
<td>Good to very good</td>
</tr>
<tr>
<td>(NA Type)</td>
<td></td>
</tr>
</tbody>
</table>
The Perceptual Assimilation Model (PAM): Catherine Best

Native phonological space

Single category assimilation

Non-assimilable

Two-category assimilation

Both uncategorizable
Empirical support from Zulu contrast discrimination by American English listeners:

Non-assimilated (NA):
Excellent discrimination of click contrasts (heard as non-speech)

Two category (TC):
Excellent discrimination of voiced vs. voiceless lateral fricatives
(assimilated to “shla” vs. “zhla”)

Category goodness difference (CG):
Good discrimination of velar voiceless aspirated vs. ejective stops
(variable assimilation to /k/)

Single category (SC):
Poor discrimination of voiced bilabial plosive versus implosive (both assimilated to /b/)
Questions?
The Speech Learning Model (SLM):
James Flege

**SLM Postulates**

- L1 learning mechanisms remain intact over the life span and apply to adult L2 acquisition.

- Language-specific aspects of speech sounds are represented in phonetic categories in long-term memory.

- Phonetics categories evolve over the life span in response to heard instances.

- Bilinguals strive to maintain contrast between L1 and L2 categories in a common phonetic space.
The Speech Learning Model (SLM):  
James Flege

SLM Hypotheses

Relevant unit size:
- L1 and L2 sounds are related to each other at the level of position-sensitive allophones.

New category formation:
- A new category can be established if some phonetic difference is discerned (between sounds in an existing category and new sounds).
- The likelihood of phonetic differences being discerned decreases as age-of-acquisition increases.
- Category formation may be blocked by equivalence classification.
The Speech Learning Model (SLM): James Flege

SLM Hypotheses

Monolingual versus bilingual phonetic space:
• Bilingual categories may be “deflected” away from monolingual categories in order to maintain contrast in the common phonetic space.
• Bilingual representations may be based on different features or feature weights.

Production and perception:
• Production eventually corresponds to properties represented in its phonetic category representation.
The Speech Learning Model (SLM): James Flege

Empirical support:
• L1 learning mechanisms remain intact over the life span and apply to adult L2 acquisition.

CRITICAL EVIDENCE:
A Test of the Critical-Period Hypothesis for Second-Language Acquisition

Kenji Hakuta,1 Ellen Bialystok,2 and Edward Wiley1
1Stanford University and 2York University, Toronto, Ontario, Canada

Psychological Science, 2004, 14: 31-38
Empirical support:
- Bilinguals strive to maintain contrast between L1 and L2 categories in a common phonetic space.
- Category formation may be blocked by equivalence classification.

Similar
/u/ in ‘tous’ or ‘two’

New
/u/ in ‘tous’ or /y/ in ‘tu’
Models of cross-language and second-language perception and production

• **Key observations**
  - Decline over the life span in the ability to perceive and produce non-native speech in a native way
  - Variable initial abilities and ease of acquisition of non-native sounds

• **Current accounts**
  - Knowledge of a first language system constrains the acquisition of a non-native system ("change resistance")
  - Adult learners interpret non-native sounds with reference to native categories ("L1 filter")
Questions?
Second Language Phonetics & Phonology

• What comes into contact with what?
  *What do you know when you know a language?*
  *(What do you *not* know when you encounter a foreign language?)*

• What is the nature of non-native language sound structure acquisition?
  *How does the target language (L2) “map onto” the native language (L1)?*

• What training facilitates non-native language sound structure acquisition?

• Why is foreign-accented speech hard to understand? Or, is it?
Second Language Phonetics & Phonology

• What comes into contact with what?
  *What do you know when you know a language?*
  (What do you **not** know when you encounter a foreign language?)

• What is the nature of non-native language sound structure acquisition?
  *How does the target language (L2) “map onto” the native language (L1)?*

• What training facilitates non-native language sound structure acquisition?

• Why is foreign-accented speech hard to understand?
  Or, is it?
Speech learning involves:

- Developing a system of contrasting categories

```
  pick  tick  stick  kit
  t
  p
  tick  stick  kit
```

- Generalization over category exemplars
Perceptual tuning to the native language => mistuning to non-native languages:

Goal of non-native speech sound training

• Re-tune the phonetic system

From:
   a system that is tuned (uniquely) to the distribution of sounds in the native language

To:
   a flexible system that can be tuned to the distribution of sounds in both the native and non-native language

Such that:
   novel words and sentences by novel talkers can be accurately classified and understood
How do listeners handle signal variability?

Classic models of categorization:

• Instance-specific information is separated from linguistic information.

• Linguistic categories are defined in terms of abstract features.
How do listeners handle signal variability?

Exemplar-based models of categorization:

• Instance-specific information is an integral part of the cognitive representation of the signal.

• Linguistic categories correspond to clusters of exemplars which are organized in the cognitive space based on acoustic or articulatory similarity.
Two approaches to phonetic training

A. Low-Variability Training:
• draw attention to essential difference(s) between contrasting categories
• develop sensitivity to dimensions of category contrast

B. High-Variability Training:
• expose to full extent of natural variability within categories
• encourage classification
Training Japanese listeners to identify English /r/ and /l/

rock brake dire lock Blake dial

R r l
The Native Language Magnet Model (NLM): Patricia Kuhl and Paul Iverson

/r/ and /l/ ‘Perceptual maps’ for American English and Japanese speakers

Physical Spacing of Stimuli

Iverson et al., 2003
The Perceptual Assimilation Model (PAM): Catherine Best

Native phonological space

Single category assimilation

Non-assimilable

Two-category assimilation

Both uncategorizable
The Speech Learning Model (SLM): James Flege

SLM Hypotheses

New category formation:
- A new category can be established if some phonetic difference is discerned (between sounds in an existing category and new sounds).
- Category formation may be blocked by equivalence classification.
Low variability training with adaptation
(E.g. McCandliss, Fiez, Protopapas, Conway & McClelland, 2002)

Two training stimulus sets:

1. Adaptive: begin with “exaggerated” examples and slowly move to more typical examples.

2. Fixed: train only on typical examples.

Train with and without feedback.

Test on trained and transfer continua.
Learning without feedback

Adaptive training results in a sharper category boundary on the trained continuum.

Learning with feedback

Both adaptive and fixed training result in a sharper category boundary on both the trained and transfer continua.
High variability training
(E. g. Bradlow, Pisoni, Akahane-Yamada and Tohkura, 1997)

Training: Identify a spoken word as a member of a minimal pair
Variable training stimuli: multiple talkers, positions in word

Test: Same as training task
Generalization 1: new words by a familiar talker
Generalization 2: new words by a new talker

FIG. 1. Percent correct perceptual identification performance for trained (left panel) and control (right panel) subjects at pretest, post-test, and the two tests of generalization. The error bars represent one standard error from the mean.
High variability training

- Highly robust learning can be achieved with high-variability training.
- Signal variability is informative.
High variability training on other contrasts

<table>
<thead>
<tr>
<th>Contrast(s)</th>
<th>Listeners</th>
<th>Task (researchers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mandarin lexical tones</td>
<td>American English</td>
<td>Identification</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Wang, Spence, Jongman and Sereno, 1993)</td>
</tr>
<tr>
<td>Hindi dental vs. retroflex stops</td>
<td>English and Japanese</td>
<td>Identification, fading technique</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Pruitt, 1995)</td>
</tr>
<tr>
<td>Japanese vowel lengths</td>
<td>English</td>
<td>Identification</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Yamada, Yamada and Srange, 1996)</td>
</tr>
<tr>
<td>English word-final / t/-/d/</td>
<td>Chinese</td>
<td>Identification or categorial discrimination</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Flege, 1995)</td>
</tr>
<tr>
<td>German vowel contrasts</td>
<td>English</td>
<td>Identification</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Kingston, 2003)</td>
</tr>
</tbody>
</table>
The Speech Learning Model (SLM): James Flege

SLM Hypotheses

New category formation:
- A new category can be established if some phonetic difference is discerned (between sounds in an existing category and new sounds).
- Category formation may be blocked by equivalence classification.

Production and perception:
- Production eventually corresponds to properties represented in its phonetic category representation
Effect of training in perception on production

English listener evaluations of Japanese productions

![Graph showing percent correct identification over time for trained and control groups. Pre-Test, Post-Test, and 3 months data are presented.]
Relationship between learning in perception and production

- Production leads perception(!?)
- Perception leads production
- Too poor to improve
- No room for improvement
But … This is complicated
Implicit Learning

• Distribution of tokens along continuum dictates category structure.

- Unimodal participants should not discriminate between tokens
- Bimodal group should distinguish between tokens across the two categories, but not within.
Training – Perception-Only Training

• Perception-only Training
  – Training: Hear a token from continuum, press a button; 2 consecutive days
  – Tokens presented in either bimodal or unimodal distribution
Training – Perception+Production

• Perception+Production Training
  – Training: Hear a token from continuum, repeat the token aloud, press a button; 2 consecutive days.
  – Tokens presented in either bimodal or unimodal distribution
Testing

• To assess learning participants were tested in perception and production.
• Testing identical for all participants
• Perception test – Discrimination
• Production test – Repetition

• Discrimination test – Are these tokens the same or different?
Testing

- 3 types of pairs
  - Same
  - Within Category
  - “Across” Category (NOTE: within for unimodal)
- Measurement is $d'$ – helps correct for bias
Predictions – Perception-Only

Predicted Discrimination Pattern

![Graph showing predicted discrimination pattern with categories Bimodal and Unimodal.](chart.png)
Perception-Only Training

![ Discrimination (Post-Test) graph showing disparities between across and within category perceptions for bimodal and unimodal perceptions. ]
Predictions – Perception+Production

Production does not affect perceptual learning

- Across Category
- Within Category

Bimodal  Unimodal

Training Type
Predictions – Perception+Production

![Diagram showing production helps perceptual learning with bars for Bimodal and Unimodal training types, with 'Across Category' and 'Within Category' markers.]
Predictions – Perception+Production

Production hinders perceptual learning

- Bimodal
- Unimodal

Training Type

- Across Category
- Within Category
Predictions – Perception+Production

Production does not affect perceptual learning

Production helps perceptual learning

Production hinders perceptual learning
Perception+Production Training

Discrimination (Post-Test)

- Bimodal Perception
- Unimodal Perception
- Bimodal Perception+Production
- Unimodal Perception+Production

- Across Category
- Within Category
Questions?