Listen Carefully!

Hindi:

A  B  X
Listen Carefully!

Hindi:

Answer: A

dental vs. retroflex contrast (A=dental, B=retroflex)
Listen Carefully!

Salish:

A  B  X
Listen Carefully!

Salish: A B X

Answer: A

velar vs. uvular contrast (A=velar, B=uvular)
Learning Sound Categories

age in months 6 7 8 9 10 11 12
Learning Sound Categories

6-8 months: discriminate non-native consonant contrasts

10-12 months: poor discrimination of non-native consonant contrasts

(Werker & Tees, 1984)
Learning Sound Categories

6 months: some language-specific perception of vowels

6-8 months: discriminate non-native consonant contrasts

10-12 months: poor discrimination of non-native consonant contrasts

(Werker & Tees, 1984)

(Kuhl et al., 1992)
How are sound categories learned?
An Inference Problem

Learner recovering linguistic structure

Hypotheses: possible linguistic analyses
Data: corpus (language input)

\[ p(h \mid d) \propto p(d \mid h)p(h) \]
An Inference Problem

Learner recovering linguistic structure

Hypotheses: possible linguistic analyses
Data: corpus (language input)

\[ p(h \mid d) \propto p(d \mid h)p(h) \]

What types of hypotheses should learners consider?
Outline

- Distributional learning
- Lexical-distributional learning
- Learning English vowels
- Dealing with systematic variability

(Joint work with Tom Griffiths, James Morgan, Sharon Goldwater)
Outline

- Distributional learning
- Lexical-distributional learning
- Learning English vowels
- Dealing with systematic variability

(Joint work with Tom Griffiths, James Morgan, Sharon Goldwater)
Distributional Learning
Distributional Learning
Distributional Learning

Bimodal group: good discrimination between endpoints
Unimodal group: poor discrimination between endpoints

(Maye, Werker, & Gerken, 2002)
A Generative Model

To create a corpus

Phonetic Categories

Corpus
A Generative Model

To create a corpus
1. Generate a phonetic category inventory
A Generative Model

To create a corpus

1. Generate a phonetic category inventory
   - Sample a mean, covariance, and frequency of occurrence for each Gaussian category

Phonetic Categories

Corpus
A Generative Model

To create a corpus

1. Generate a phonetic category inventory
   - Sample a mean, covariance, and frequency of occurrence for each Gaussian category
2. Generate a corpus
To create a corpus

1. Generate a phonetic category inventory
   - Sample a mean, covariance, and frequency of occurrence for each Gaussian category

2. Generate a corpus
   - For each sound, sample a phonetic category according to its frequency
A Generative Model

To create a corpus

1. Generate a phonetic category inventory
   - Sample a mean, covariance, and frequency of occurrence for each Gaussian category

2. Generate a corpus
   - For each sound, sample a phonetic category according to its frequency
   - Generate an acoustic value from the Gaussian distribution associated with that category
A Generative Model

Dimension 1

Dimension 2
A Generative Model
A Generative Model

\( \mu_c, \Sigma_c \) : parameters of category \( c \)

\( z_i \) : category of sound \( i \)

\( x_i \) : acoustics of sound \( i \)
A Generative Model

\[ \mu_c, \Sigma_c : \text{parameters of category } c \]
\[ z_i : \text{category of sound } i \]
\[ x_i : \text{acoustics of sound } i \]

Need to infer hidden variables:
- Parameters for each category
- Category label for each point

Can use Expectation Maximization, Gibbs sampling, online gradient descent, etc.
Distributional Learning

Voiced and Voiceless Stops

(Toscano & McMurray, 2008; McMurray, Aslin, & Toscano, 2009)
Distributional Learning

Vowel Categories (Single Speakers)

(Vallabha, McClelland, Pons, Werker, & Amano, 2007)
Overlapping Categories
Overlapping Categories
A Difficult Problem

(Hillenbrand, Getty, Clark, & Wheeler, 1995)
A Difficult Problem

(Hillenbrand, Getty, Clark, & Wheeler, 1995)
A Fancier Generative Model

$\alpha$: concentration parameter

$\mu_c, \Sigma_c$: parameters of category $c$

$z_i$: category of sound $i$

$x_i$: acoustics of sound $i$
Training Corpus

Corpus of 6,409 vowel tokens generated from Gaussian categories from Hillenbrand et al. (1995); frequencies match corpus frequencies
Distributional Learning

Vowel Categories (All Speakers)

Infinite Mixture Model

Accuracy: 37%
Completeness: 57%
Distributional Learning

Vowel Categories (All Speakers)

Gradient Descent Algorithm
(Vallabha et al., 2007)

Accuracy: 34%
Completeness: 89%
To create a corpus

1. Generate a phonetic category inventory
   - Sample a mean, covariance, and frequency of occurrence for each Gaussian category

2. Generate a corpus
   - For each sound, sample a phonetic category according to its frequency
   - Generate an acoustic value from the Gaussian distribution associated with that category
Outline

- Distributional learning
- Lexical-distributional learning
- Learning English vowels
- Dealing with systematic variability

(Joint work with Tom Griffiths, James Morgan, Sharon Goldwater)
Word Segmentation Task

Familiarization:

Test:

familiar   unfamiliar

“Success”: Difference in looking times between familiar and unfamiliar words in fluent speech
Word Learning
Word Categorization
Phonetic Category Learning

6 months: some language-specific perception of vowels

6-8 months: discriminate non-native consonant contrasts

10-12 months: poor discrimination of non-native consonant contrasts

(Werker & Tees, 1984)

(Kuhl et al., 1992)
Phonetic Category Learning

6 months: some language-specific perception of vowels

6-8 months: discriminate non-native consonant contrasts

7.5 months: segment monosyllables and strong-weak bisyllables

10.5 months: segment weak-strong bisyllables

10-12 months: poor discrimination of non-native consonant contrasts

(Werker & Tees, 1984)
(Kuhl et al., 1992)
(Jusczyk & Aslin, 1995)
(Jusczyk et al., 1999)
Phonetic Category Learning

6 months: some language-specific perception of vowels
6-8 months: discriminate non-native consonant contrasts
10-12 months: poor discrimination of non-native consonant contrasts

6 months: segment words next to “Mommy” or baby’s own name
7.5 months: segment monosyllables and strong-weak bisyllables
10.5 months: segment weak-strong bisyllables

(Werker & Tees, 1984)
(Kuhl et al., 1992)
(Jusczyk & Aslin, 1995)
(Jusczyk et al., 1999)
(Bortfeld et al., 2005)
A Generative Model

To create a corpus

1. Generate a phonetic category inventory
   - Sample a mean, covariance, and frequency of occurrence for each Gaussian category

2. Generate a corpus
   - For each sound, sample a phonetic category according to its frequency
   - Generate an acoustic value from the Gaussian distribution associated with that category
A Better Generative Model

To create a corpus

Phonetic Categories

Lexicon

Corpus
A Better Generative Model

To create a corpus

1. Generate a phonetic category inventory
To create a corpus

1. Generate a phonetic category inventory
   - Sample a mean, covariance, and frequency of occurrence for each Gaussian category
A Better Generative Model

**To create a corpus**

1. Generate a phonetic category inventory
   - Sample a mean, covariance, and frequency of occurrence for each Gaussian category
2. Generate a lexicon
A Better Generative Model

To create a corpus

1. Generate a phonetic category inventory
   - Sample a mean, covariance, and frequency of occurrence for each Gaussian category
2. Generate a lexicon
   - Sample a length and frequency of occurrence for each lexical item
To create a corpus

1. Generate a phonetic category inventory
   - Sample a mean, covariance, and frequency of occurrence for each Gaussian category

2. Generate a lexicon
   - Sample a length and frequency of occurrence for each lexical item
   - For each phoneme slot, sample a phonetic category from the phonetic category inventory
To create a corpus

1. Generate a phonetic category inventory
   - Sample a mean, covariance, and frequency of occurrence for each Gaussian category

2. Generate a lexicon
   - Sample a length and frequency of occurrence for each lexical item
   - For each phoneme slot, sample a phonetic category from the phonetic category inventory

3. Generate a corpus
To create a corpus

1. Generate a phonetic category inventory
   - Sample a mean, covariance, and frequency of occurrence for each Gaussian category

2. Generate a lexicon
   - Sample a length and frequency of occurrence for each lexical item
   - For each phoneme slot, sample a phonetic category from the phonetic category inventory

3. Generate a corpus
   - For each word, sample a lexical item according to its frequency
To create a corpus

1. Generate a phonetic category inventory
   - Sample a mean, covariance, and frequency of occurrence for each Gaussian category

2. Generate a lexicon
   - Sample a length and frequency of occurrence for each lexical item
   - For each phoneme slot, sample a phonetic category from the phonetic category inventory

3. Generate a corpus
   - For each word, sample a lexical item according to its frequency
   - Generate an acoustic value each phonetic category contained in that lexical item
A Better Generative Model
A Better Generative Model

Lexicon: /bæd/, /bɛd/
‘bad’ ‘bed’

phonetic dimension

b
ε
æ
d

time
A Better Generative Model

Lexicon: /bæd/, /bɛd/
‘bad’ ‘bed’

phonetic dimension
A Better Generative Model

Lexicon: /bæd/, /bɛd/
‘bad’ ‘bed’

phonetic dimension
A Better Generative Model

Lexicon: /bæd/, /bɛd/
‘bad’ ‘bed’

phonetic dimension
A Better Generative Model

Lexicon: /bæd/, /bɛd/
‘bad’ ‘bed’

phonetic dimension
A Better Generative Model

Lexicon: /bæd/, /bɛd/
‘bad’ ‘bed’

phonetic dimension
A Better Generative Model

Lexicon: /bæd/, /bɛd/
‘bad’ ‘bed’

phonetic dimension
A Better Generative Model

Lexicon: /bæd/, /bɛd/
‘bad’ ‘bed’

phonetic dimension
A Better Generative Model

\[ \mu_c, \Sigma_c : \text{parameters of category } c \]
\[ l_k : \text{form of lexical item } k \]
\[ z_i : \text{category of word } i \]
\[ w_i : \text{acoustics of word } i \]
\[ \alpha_c : \text{phonetic concentration parameter} \]
\[ \alpha_L : \text{lexical concentration parameter} \]
Models of Category Learning

**Distributional Model**
- Phonetic Categories
- Corpus
- data

**Lexical-Distributional Model**
- Phonetic Categories
- Corpus
- Lexicon
- hypotheses

---

The diagram illustrates the relationship between models of category learning and their components: Phonetic Categories, Corpus, Hypotheses, and Lexicon. It shows how data flows through the models and how hypotheses are generated.
Models of Category Learning

**Distributional**
- Assume sounds are generated independently of their neighbors
- Infer category parameters
- Phonetic categories characterize the types of variability found among sounds in the corpus

**Lexical-Distributional**
- Assume sounds are generated as parts of words
- Infer category parameters and forms of lexical items
- Phonetic categories are overhypotheses about the types of variability seen in lexical items
Qualitative Behavior

Compare lexical-distributional model’s behavior on two lexicons

- Informative lexicon: ‘add’, ‘ebb’

Qualitative Behavior

Compare lexical-distributional model’s behavior on two lexicons

- Informative lexicon: ‘add’, ‘ebb’

Minimal pairs:
add vs. Ed

Typically taken as evidence that sounds are different
Informative Lexicon

Lexicon: /æd/, /ɛb/
‘add’ ‘ebb’
Informative Lexicon

Lexicon: /æd/, /ɛb/
‘add’ ‘ebb’

phonetic dimension

æ
time

b

æ

ɛ

d
Informative Lexicon

Lexicon: /æd/, /ɛb/
‘add’ ‘ebb’

phonetic dimension
Informative Lexicon

Lexicon: /æd/, /ɛb/
‘add’ ‘ebb’

phonetic dimension
Informative Lexicon

Lexicon: /æd/, /ɛb/
‘add’ ‘ebb’

phonetic dimension
Informative Lexicon

Lexicon: /æd/, /ɛb/
‘add’ ‘ebb’
Informative Lexicon

Lexicon: /æd/, /ɛb/
‘add’ ‘ebb’

phonetic dimension
Informative Lexicon

Lexicon: /æd/, /ɛb/
‘add’ ‘ebb’

phonetic dimension
Informative Lexicon

Lexicon: /æd/, /ɛb/  
‘add’ ‘ebb’

phonetic dimension
Informative Lexicon

Lexicon: /æd/, /ɛb/
‘add’ ‘ebb’

phonetic dimension
Distributional Model

Lexicon: /æd/, /ɛb/  
‘add’ ‘ebb’

phonetic dimension
Distributional Model

Lexicon: /æd/, /ɛb/
‘add’ ‘ebb’

phonetic dimension
Distributional Model
Lexical-Distributinal Model

Lexicon: /æd/, /ɛb/
‘add’ ‘ebb’
Lexical-Distribualional Model

Lexicon: /æd/, /ɛb/
‘add’ ‘ebb’

phonetic dimension
Lexical-Distributional Model

ebb
add
Minimal Pair Lexicon

Lexicon: /æd/, /ɛd/, /æb/, /ɛb/  
‘add’ ‘Ed’ ‘ab’ ‘ebb’
Minimal Pair Lexicon

Lexicon: /æd/, /ɛd/, /æb/, /ɛb/
‘add’ ‘Ed’ ‘ab’ ‘ebb’

phonetic dimension
Minimal Pair Lexicon

Lexicon: /æd/, /ɛd/, /æb/, /ɛb/
‘add’ ‘Ed’ ‘ab’ ‘ebb’

phonetic dimension
Minimal Pair Lexicon

Lexicon: /æd/, /ɛd/, /æb/, /ɛb/
‘add’ ‘Ed’ ‘ab’ ‘ebb’
Minimal Pair Lexicon

Lexicon: /æd/, /ɛd/, /æb/, /ɛb/
‘add’ ‘Ed’ ‘ab’ ‘ebb’
Minimal Pair Lexicon

Lexicon: /æd/, /ɛd/, /æb/, /ɛb/
‘add’ ‘Ed’ ‘ab’ ‘ebb’

phonetic dimension
Minimal Pair Lexicon

Lexicon: /æd/, /ɛd/, /æb/, /ɛb/
‘add’ ‘Ed’ ‘ab’ ‘ebb’
Minimal Pair Lexicon

Lexicon: /æd/, /ɛd/, /æb/, /ɛb/  
‘add’  ‘Ed’  ‘ab’  ‘ebb’

phonetic dimension
Minimal Pair Lexicon

Lexicon: /æd/, /ɛd/, /æb/, /ɛb/  
‘add’ ‘Ed’ ‘ab’ ‘ebb’

phonetic dimension
Minimal Pair Lexicon

Lexicon: /æd/, /ɛd/, /æb/, /ɛb/

‘add’ ‘Ed’ ‘ab’ ‘ebb’
Minimal Pair Lexicon

Lexicon: \(/æd/, /ɛd/, /æb/, /ɛb/\)

‘add’ ‘Ed’ ‘ab’ ‘ebb’

phonetic dimension
Minimal Pair Lexicon

ab
Ed  ebb
add
Lexical-Distributional Learning

- If the lexicon contains disambiguating information, the learner should use this information to disambiguate overlapping categories.
Lexical-Distributional Learning

- If the lexicon contains disambiguating information, the learner should use this information to disambiguate overlapping categories.

- Learner uses each level of structure to constrain the other:
Lexical-Distributional Learning

- If the lexicon contains disambiguating information, the learner should use this information to disambiguate overlapping categories.

- Learner uses each level of structure to constrain the other:
  - Distributional information helps determine which words are tokens of the same lexical item.
Lexical-Distributional Learning

- If the lexicon contains disambiguating information, the learner should use this information to disambiguate overlapping categories.

- Learner uses each level of structure to constrain the other:
  - Distributional information helps determine which words are tokens of the same lexical item.
  - Lexical information helps determine which sounds are part of the same phonetic category.
Empirical Evidence

Evidence from 15-month-olds in “switch” task (Stager & Werker, 1997)

(Thiessen, 2007)
Empirical Evidence

Evidence from 15-month-olds in “switch” task (Stager & Werker, 1997)

Habituation:

“daw”
Empirical Evidence

Evidence from 15-month-olds in “switch” task (Stager & Werker, 1997)

Habituation:
“daw”

Switch trial:
“taw”

(Thiessen, 2007)
Empirical Evidence

Evidence from 15-month-olds in “switch” task (Stager & Werker, 1997)

Habituation: “daw”

Switch trial: “taw”

(Thiessen, 2007)
Empirical Evidence

Evidence from 15-month-olds in “switch” task (Stager & Werker, 1997)

Habituation:

- “daw”

Switch trial:

- “taw”

(Thiessen, 2007)
Empirical Evidence

Evidence from 15-month-olds in “switch” task (Stager & Werker, 1997)

Habituation:

- “daw”
- “tawgoo”
- “dawbow”

Switch trial:

- “taw”
- “taw”

(Thiessen, 2007)
Empirical Evidence

Evidence from 15-month-olds in “switch” task (Stager & Werker, 1997)

Habituation:

- “daw”
- “tawgoog”
- “dawbow”

Switch trial:

- “taw”
- “tawgoog”
- “dawgoog”

(Thiessen, 2007)
Empirical Evidence

Evidence from 15-month-olds in “switch” task (Stager & Werker, 1997)

Habituation:

- “daw”
- “daw”
- “tawgo”
- “dawbow”

Switch trial:

- “taw”
- “taw”
- “dawgo”
- “dawgo”

(Thiessen, 2007)
Empirical Evidence

- 15-month-olds show better discrimination when lexicon provides disambiguating information (Thiessen, 2007)
  
  \[
  \begin{array}{cc}
  \text{“tawgoo”} & \text{“dawbow”} \\
  \text{“tawgoo”} & \text{“dawgoo”}
  \end{array}
  \]

- Adults show similar behavior in a non-referential task when learning about vowel categories (Feldman, Myers, White, Griffiths, & Morgan, 2011)
Outline

- Distributional learning
- Lexical-distributional learning
- Learning English vowels
- Dealing with systematic variability

(Joint work with Tom Griffiths, James Morgan, Sharon Goldwater)
Simulations

![Simulations](image-url)
Simulations

- Lexicon from CHILDES Parental Corpus (Li & Shirai, 2000)
  - Orthographic forms phonematized using Carnegie Mellon Pronouncing Dictionary
  - Lexical items sampled according to corpus frequency

- Corpus of 5000 word tokens, comprising 6,409 vowel tokens and 8,917 consonant tokens

- Acoustic values for vowels sampled based on Hillenbrand et al. (1995) data
  - Means, covariance matrices computed from speakers’ productions
  - Speech sounds generated from Gaussians
Distributional Model

Vowel Categories (All Speakers)

First Formant (Hz)

Second Formant (Hz)

Infinite Mixture Model

First Formant (Hz)

Second Formant (Hz)

Accuracy: 0.369
Completeness: 0.575
Lexical-Distributitional Model

Vowel Categories (All Speakers)

Lexical-Distributional Model

Accuracy: 0.709
Completeness: 0.710
Lexical-Distributioonal Model

Vowel Categories (All Speakers)

Lexical-Distributioonal Model

Accuracy: 0.778
Completeness: 0.736
Benefit of Using Words

F-Score

<table>
<thead>
<tr>
<th></th>
<th>Lexical-Distributional ((\alpha_L=10))</th>
<th>Lexical-Distributional ((\alpha_L=10,000))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distributional</td>
<td>0.45</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>0.76</td>
<td>0.74</td>
</tr>
</tbody>
</table>
Number of Categories

<table>
<thead>
<tr>
<th>Lexical-Distributional Model</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Distributional Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>
## Number of Categories

<table>
<thead>
<tr>
<th>$\alpha_C$</th>
<th>Lexical-Distributional Model</th>
<th>Distributional Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Number of Categories

<table>
<thead>
<tr>
<th>$\alpha_C$</th>
<th>$\alpha_L=1$</th>
<th>$\alpha_L=10$</th>
<th>$\alpha_L=100$</th>
<th>$\alpha_L=1,000$</th>
<th>$\alpha_L=10,000$</th>
<th>Distributional Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Number of Categories

<table>
<thead>
<tr>
<th>$\alpha_C$</th>
<th>$\alpha_L=1$</th>
<th>$\alpha_L=10$</th>
<th>$\alpha_L=100$</th>
<th>$\alpha_L=1,000$</th>
<th>$\alpha_L=10,000$</th>
<th>Distributional Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7</td>
</tr>
</tbody>
</table>

Number of Phonetic Categories (gold standard = 12)
Number of Categories

<table>
<thead>
<tr>
<th>$\alpha_C$</th>
<th>$\alpha_L=1$</th>
<th>$\alpha_L=10$</th>
<th>$\alpha_L=100$</th>
<th>$\alpha_L=1,000$</th>
<th>$\alpha_L=10,000$</th>
<th>Distributional Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>14</td>
<td>13</td>
<td>13</td>
<td>12</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>1</td>
<td>14</td>
<td>14</td>
<td>13</td>
<td>12</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td>14</td>
<td>13</td>
<td>13</td>
<td>12</td>
<td>12</td>
<td>7</td>
</tr>
</tbody>
</table>

Number of Phonetic Categories (gold standard = 12)
## Number of Categories

<table>
<thead>
<tr>
<th>$\alpha_C$</th>
<th>Lexical-Distributional Model</th>
<th>Distributional Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha_L=1$</td>
<td>$\alpha_L=10$</td>
</tr>
<tr>
<td>0.1</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>1</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>10</td>
<td>14</td>
<td>13</td>
</tr>
</tbody>
</table>

Number of Phonetic Categories (gold standard = 12)
Number of Lexical Items (gold standard = 1019)
Number of Categories

<table>
<thead>
<tr>
<th>$\alpha_C$</th>
<th>$\alpha_L=1$</th>
<th>$\alpha_L=10$</th>
<th>$\alpha_L=100$</th>
<th>$\alpha_L=1,000$</th>
<th>$\alpha_L=10,000$</th>
<th>Distributional Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>14</td>
<td>13</td>
<td>13</td>
<td>12</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>900</td>
<td>916</td>
<td>969</td>
<td>1145</td>
<td>1601</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>14</td>
<td>13</td>
<td>13</td>
<td>12</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>899</td>
<td>912</td>
<td>968</td>
<td>1138</td>
<td>1605</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>14</td>
<td>13</td>
<td>13</td>
<td>12</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>900</td>
<td>926</td>
<td>958</td>
<td>1164</td>
<td>1602</td>
<td></td>
</tr>
</tbody>
</table>

Number of Phonetic Categories (gold standard = 12)
Number of Lexical Items (gold standard = 1019)
Lexical-Distributional Model

Extra category includes:

- find, found
- think, thank
- will, we’ll, well
- give, gave
- made, mad, mid
- big, bag
- way, we

as well as lexical items that were not minimal pairs
Minimal Pairs

- Phonologists use minimal pairs to identify contrastive categories

- Minimal pairs make it *more* difficult to distinguish between phonemes if no meanings are known: items in the pair could be the same word

- Model can overcome minimal pair problem with certain parameter values, but children may use other strategies
More Interactions in Learning?

- **Phonotactics**
  - Sensitivity to phonotactics at 9 months could make a learner more willing to accept multiple lexical items with a common consonant frame (Jusczyk et al., 1994)

- **Semantics**
  - Semantic information may help pull apart minimal pairs (Yeung & Werker, 2009; but see Thiessen, 2007)
  - Semantic information may help a learner recognize redundant lexical items
Simulations

- Lexicon generated from the model
  - Words composed only of vowels
  - Structure of the lexicon matches the learner’s expectations

- Corpus of 5000 word tokens, comprising 22,397 vowel tokens

- Acoustic values sampled based on Hillenbrand et al. (1995) data
  - Means, covariance matrices computed from speakers’ productions
  - Speech sounds generated from Gaussians
Distributional Model

Accuracy: 59%
Completeness: 62%
Summary

- Using information from words can help disambiguate overlapping categories, even if the forms in the lexicon are not given explicitly to the learner.

- Qualitative behavior mimics human data.

- Interactive learning poses different challenges than learning each domain in isolation:
  - Disambiguating overlapping categories is difficult in isolation.
  - Similar-sounding words are difficult for interactive learner.
Outline

- Distributional learning
- Lexical-distributional learning
- Learning English vowels
- Dealing with systematic variability

Work by Ewan Dunbar, Brian Dillon, & Bill Idsardi
More information: http://ling.umd.edu/~emd/ or emd@umd.edu
Phonological Alternations

Lexical-distributational model assumes a single Gaussian distribution for a phonetic category, regardless of context

What about phonological alternations?
Phonological Alternations

[kʰ] at the beginning of a stressed syllable

[k] in an ‘sk’ cluster
Phonological Alternations

“Kate”

[kʰ] at the beginning of a stressed syllable

“skate”

[k] in an ‘sk’ cluster
Phonological Alternations

[k] and [kʰ] are allophones of the same phoneme
- Complementary distribution: [k] and [kʰ] appear in different phonological contexts
- No minimal pairs involving [k] and [kʰ]
- Speakers and listeners think of [k] and [kʰ] as “the same sound”

Typically characterized by a rule:

\[ k \rightarrow kʰ \text{ at the beginning of a stressed syllable} \]
Learning Phonemes: Option 1

Two stages:

1. Learn separate phonetic categories for [k] and [kʰ]

2. In a separate learning process, notice that the [k] and [kʰ] occur in complementary distribution, and infer that they are allophones of a single phoneme
Learning Phonemes: Option 2

Give up the assumption that sound categories are Gaussian distributions

(categories are Gaussians)

(categories are linear models)

(Dunbar, Dillon, & Idsardi, in preparation)
Linear Models

\[ Y_i \sim N(\beta_0 + \beta_1 X_{1i}) \]

\( Y_i \) is a random variable that follows a normal distribution with mean \( \beta_0 + \beta_1 X_{1i} \) and variance \( \sigma^2 \).

\( \beta_0 \) and \( \beta_1 \) are the intercept and slope parameters, respectively.

Control Test

0 \( \beta_0 \) \( \beta_0 + \beta_1 \)

t-test/ANOVA

(Dunbar, Dillon, & Idsardi, in preparation)
Mixture of Linear Models

(Dunbar, Dillon, & Idsardi, in preparation)
Mixture of Linear Models

VOT

categories are Gaussians

categories are linear models

(Dunbar, Dillon, & Idsardi, in preparation)
Mixture of Linear Models

Inuktitut: Vowels change before uvular consonants

\[ \begin{align*}
i & \rightarrow e \\
u & \rightarrow o \\
a & \rightarrow a
\end{align*} \]

E.g., /aŋiʃuk/ \rightarrow \text{[aŋiʃuk]} \quad \text{“older sibling”}

/aŋiʃuq/ \rightarrow \text{[aŋiʃoq]} \quad \text{“big”}

(Dunbar, Dillon, & Idsardi, in preparation)
How are sound categories learned?
An Inference Problem

Learner recovering linguistic structure

Hypotheses: possible linguistic analyses
Data: corpus (language input)

\[ p(h \mid d) \propto p(d \mid h)p(h) \]
An Inference Problem

Learner recovering linguistic structure

Hypotheses: possible linguistic analyses
Data: corpus (language input)

\[ p(h \mid d) \propto p(d \mid h)p(h) \]

What types of hypotheses should learners consider?
An Inference Problem

Distributional Model
Phonetic Categories

Lexical-Distributional Model
Phonetic Categories

Mixture of Linear Models
Phonemes

Corpus

Lexicon

Corpus

?
Phonetic Category Learning

6 months: some language-specific perception of vowels
6-8 months: discriminate non-native consonant contrasts
10-12 months: poor discrimination of non-native consonant contrasts

(Werker & Tees, 1984)
(Kuhl et al., 1992)
Phonetic Category Learning

- **6 months:** some language-specific perception of vowels
- **6-8 months:** segment monosyllables and strong-weak bisyllables
- **7.5 months:** segment words next to “Mommy” or baby’s own name
- **10.5 months:** segment weak-strong bisyllables
- **10-12 months:** poor discrimination of non-native consonant contrasts

**References:**
- Werker & Tees, 1984
- Kuhl et al., 1992
- Jusczyk & Aslin, 1995
- Jusczyk et al., 1999
- Bortfeld et al., 2005
Acknowledgements

Word model joint work with Tom Griffiths, James Morgan, Sharon Goldwater; allophone model by Ewan Dunbar, Brian Dillon, Bill Idsardi

Thanks to:
Sheila Blumstein, Adam Darlow, Mark Johnson, Joseph Williams

Metcalf Infant Research Lab (Brown)
Computational Cognitive Science Lab (Berkeley)
Computational Modeling Reading Group (Brown)

NSF grants BCS-0924821, BCS-0631518
NIH grant HD032005
AFOSR grant FA9550-07-1-0351