Toward a goal of understanding the mind

Naomi Feldman
University of Maryland

December 4, 2014

Language in the mind

- How to learners acquire speech sound categories?
- How do listeners perceive speech in noisy situations?
- How do listeners represent the speech they hear?

- Higher-level questions about grammar, discourse, etc.

A common approach

- Behavioral and neural data from humans in the lab
- Cognitive model of the phenomenon being studied

→ measurements from the lab

A recurring theme in cognition

Well-designed systems are tuned to fit their environment

“emergent”	“optimal”	“data-driven”
(neural modeling)	(cognitive science)	(computer science)

A recurring theme in cognition

Well-designed systems are tuned to fit their environment

“emergent”	“optimal”	“data-driven”
(neural modeling)	(cognitive science)	(computer science)

To understand the mind, we need to study the environment

Understanding the mind

From cognitive/brain science

- Behavioral and neural data from humans in the lab
- Cognitive model of the phenomenon being studied

→ measurements from the lab

Understanding the mind

From computing/engineering

- Collections of data from the environment (e.g., corpora)
- Features that help systems generalize from those data

From cognitive/brain science

- Behavioral and neural data from humans in the lab
- Cognitive model of the phenomenon being studied

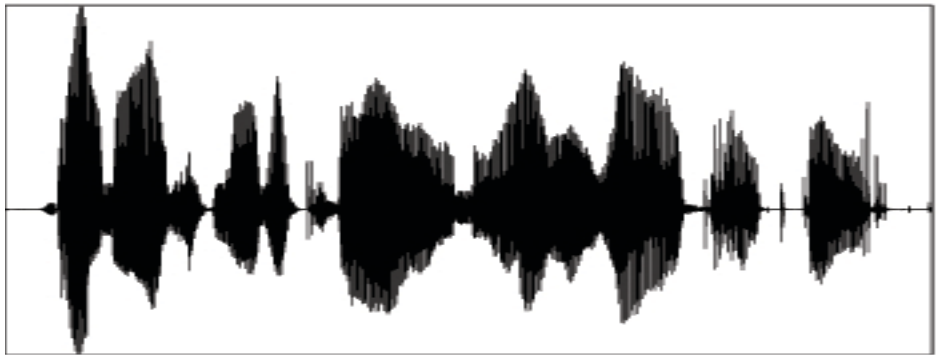
characteristics of the environment →  → measurements from the lab

“Big data for cognitive science”

- Jimmy Lin

An example from language

How is speech represented?



- Dynamic signal that's changing continuously
- Contains information in both frequency and time

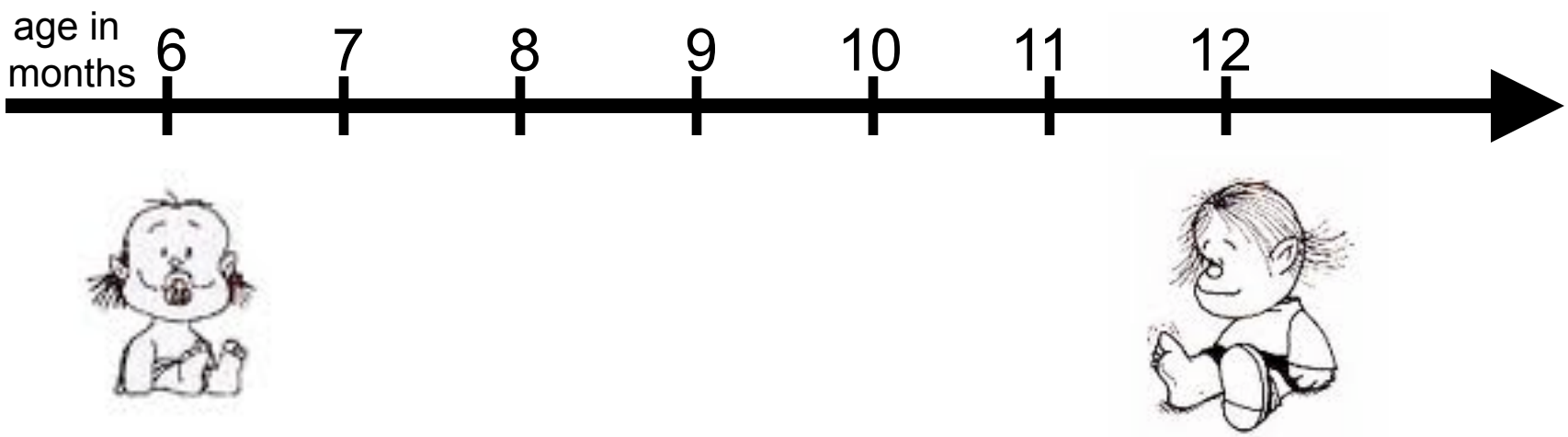
Caitlin Richter

Aren Jansen

Supported by NSF BCS-1320410

Developing representations

Speech perception becomes tuned to the native language



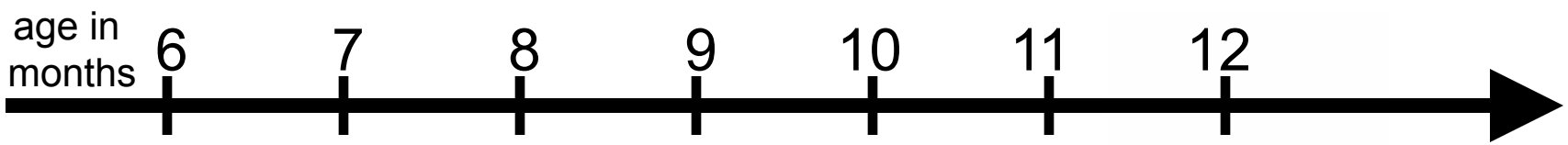
(Werker & Tees, 1984; Kuhl et al., 1992)

Developing representations

Speech perception becomes tuned to the native language

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)

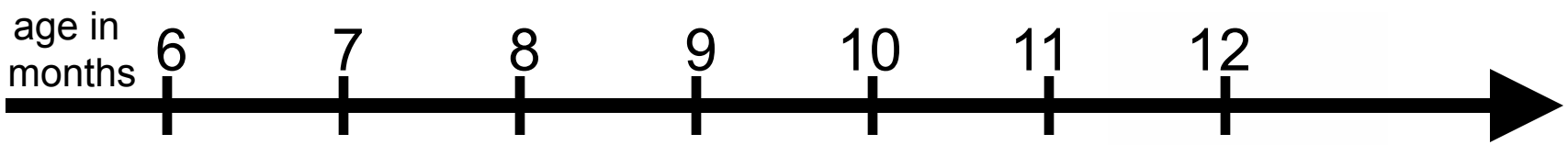
Developing representations

Speech perception becomes tuned to the native language

6 months: some language-specific perception of vowels

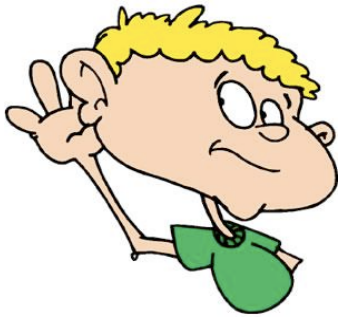
6-8 months: discriminate non-native consonant contrasts

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



(Werker & Tees, 1984; Kuhl et al., 1992)

Representation matters



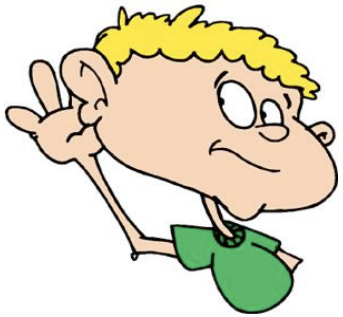
Bob

vs.

Siri

People are much better than speech recognition systems at generalizing from the data they hear

Representation matters



Bob

vs.

Siri

People are much better than speech recognition systems at generalizing from the data they hear

Non-native listeners often fail to perceive unfamiliar phonetic distinctions

How is speech represented?

From computing/engineering

- Speech corpora from many different languages
- Effective methods for representing the speech signal

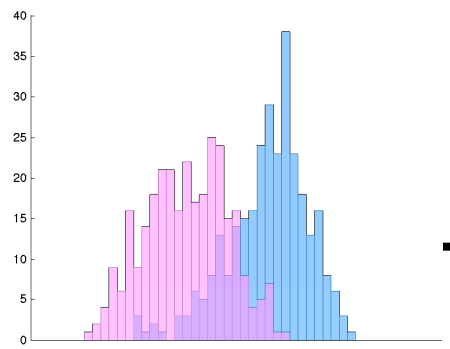
From cognitive/brain science

- Extensive data from human listeners
- Cognitive models of language acquisition and processing

characteristics of the environment →  → measurements from the lab

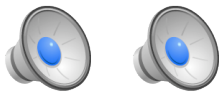
How is speech represented?

Cognitive model that connects distributions of sounds in the input to performance on a laboratory task



distribution of sounds in the input

“same” or “different”?



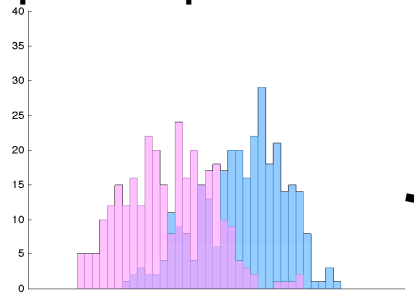
listeners’ performance on a discrimination task

(Feldman et al., 2009)

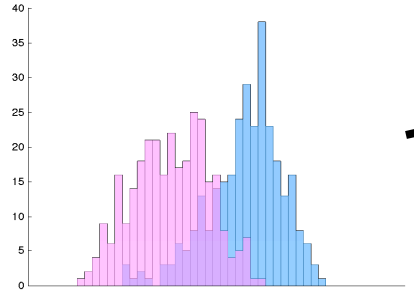
How is speech represented?

Different representations imply different input distributions

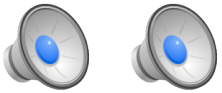
Speech Representation 1



Speech Representation 2



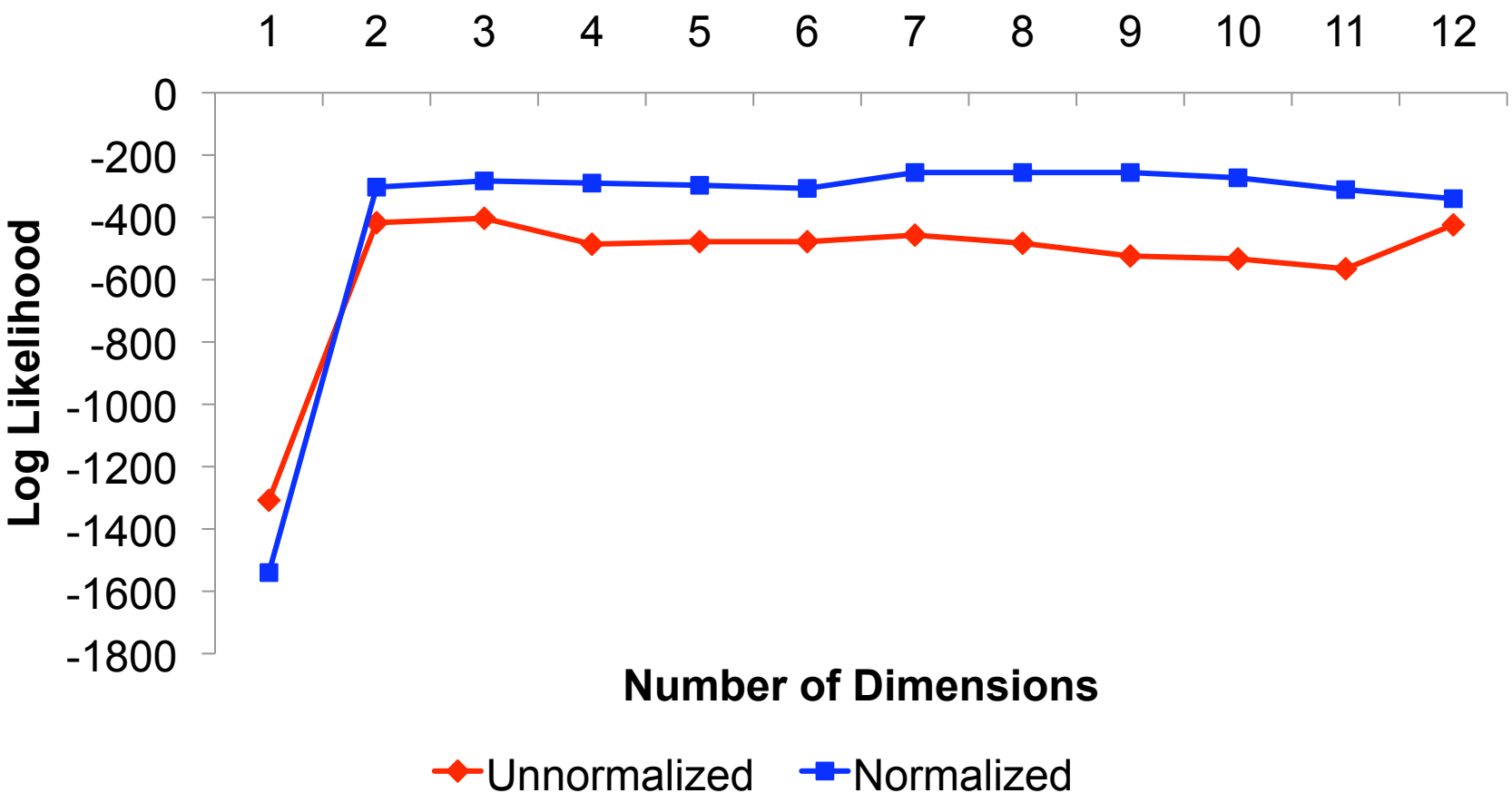
"same" or "different"?



listeners' performance on
a discrimination task

Which representations best predict human discrimination data?

Speaker normalization



(Richter et al., in prep)

A central role for the mind

From computing/engineering

- Collections of data from the environment (e.g., corpora)
- Features that help systems generalize from those data

From cognitive/brain science

- Behavioral and neural data from humans in the lab
- Cognitive model of the phenomenon being studied

characteristics of the environment →  → measurements from the lab

Benefit to cognitive science

- Ecological validity for evaluating hypotheses about cognitive representations of speech

- Engineering tools provide hypotheses and insights regarding cognitive representations
  - Methods for normalizing across speakers (Wegmann et al., 1996)
  - RASTA is essentially an edge detector for speech (Hermansky & Morgan, 1994)

Benefit to engineering

- Speech representations that yield good performance on speech recognition tasks also predict human data best (Richter et al., in prep)

- Can cognitive models of phonetic learning improve zero-resource speech recognition systems that learn representations from unlabeled data?

Connections to neuroscience?

- Existing data on neural activity in when listening to speech (e.g., Mesgarani et al., 2014; Näätänen et al., 1997; Toscano et al., 2010)

- Use neural data to investigate relationships between neural activation patterns and cognitive representations

  → How do feature representations relate to neural activations computed from the same stretch of speech?