Speech perception: Categorizing experience

HG 4015/7015

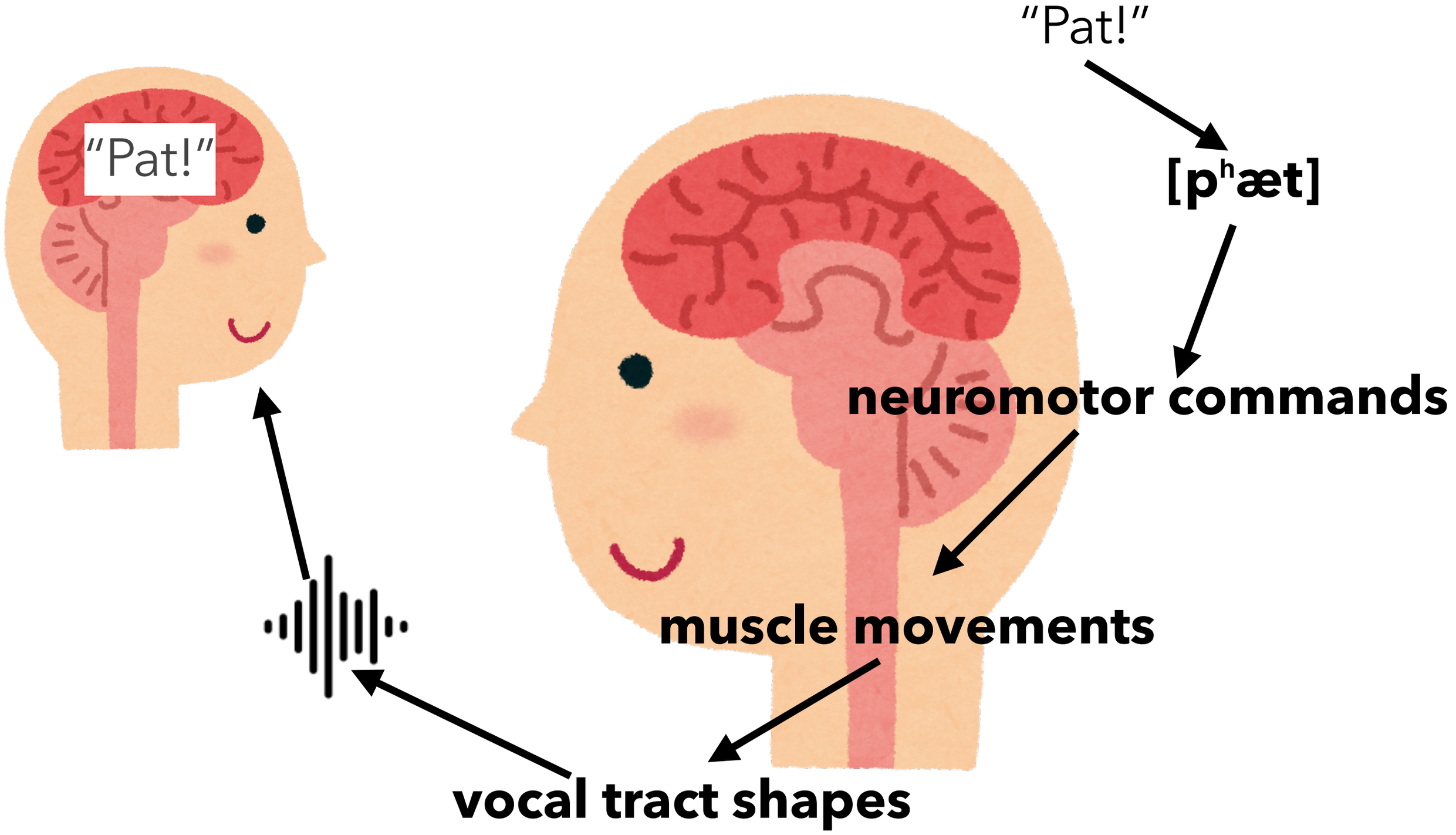
AY25-26 Sem 2

In this segment

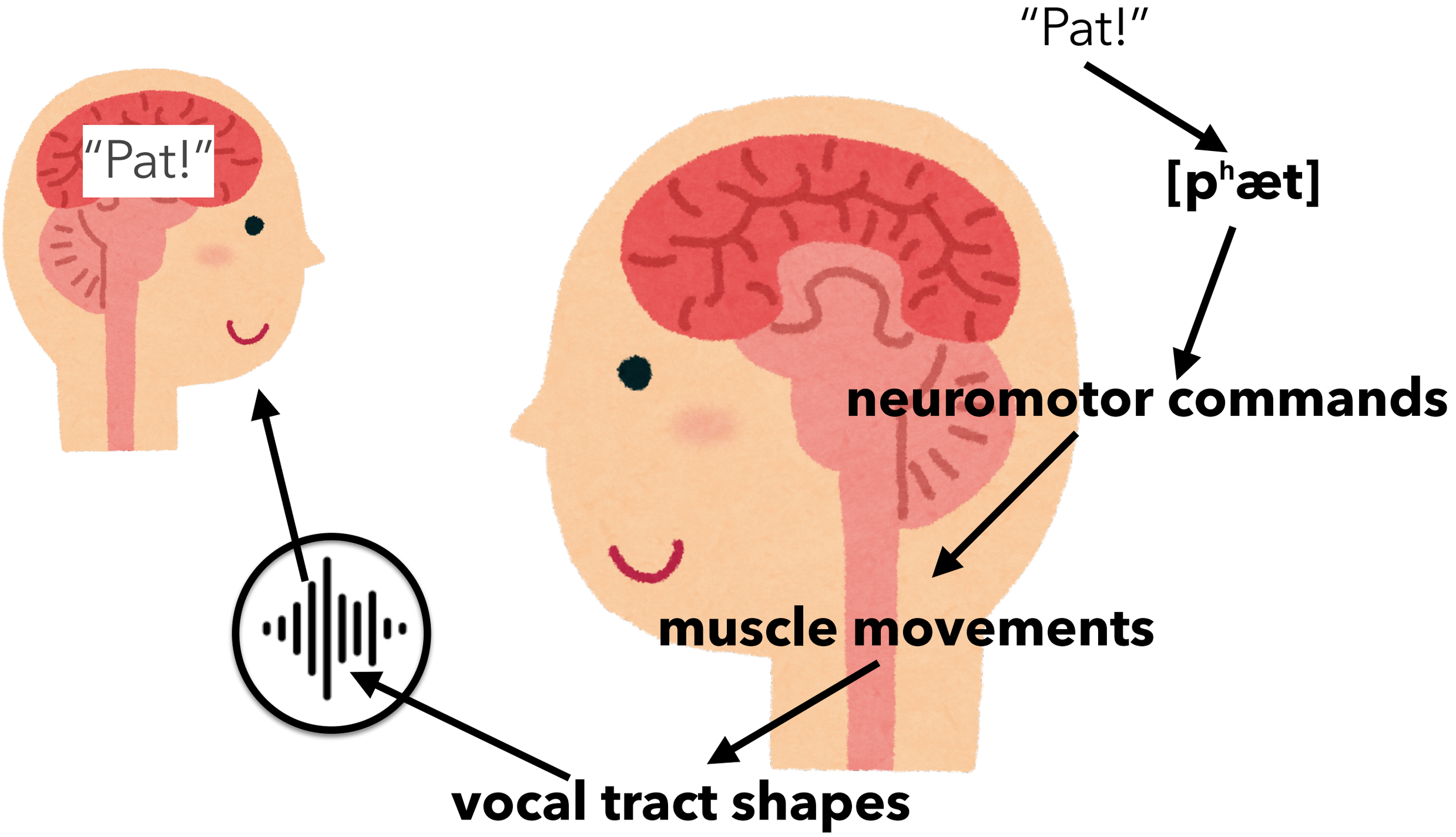
Introduction to speech perception:

- **Key concepts:** perceptual constancy, Motor Theory of Speech Perception, Direct Realism, categorical perception, perceptual cue, McGurk effect

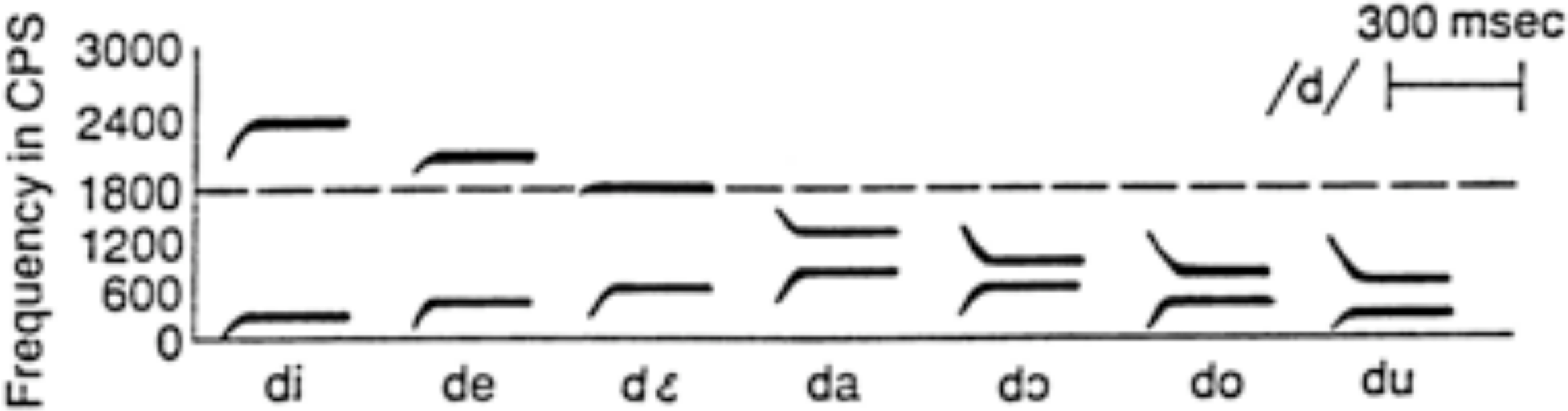
Producing speech



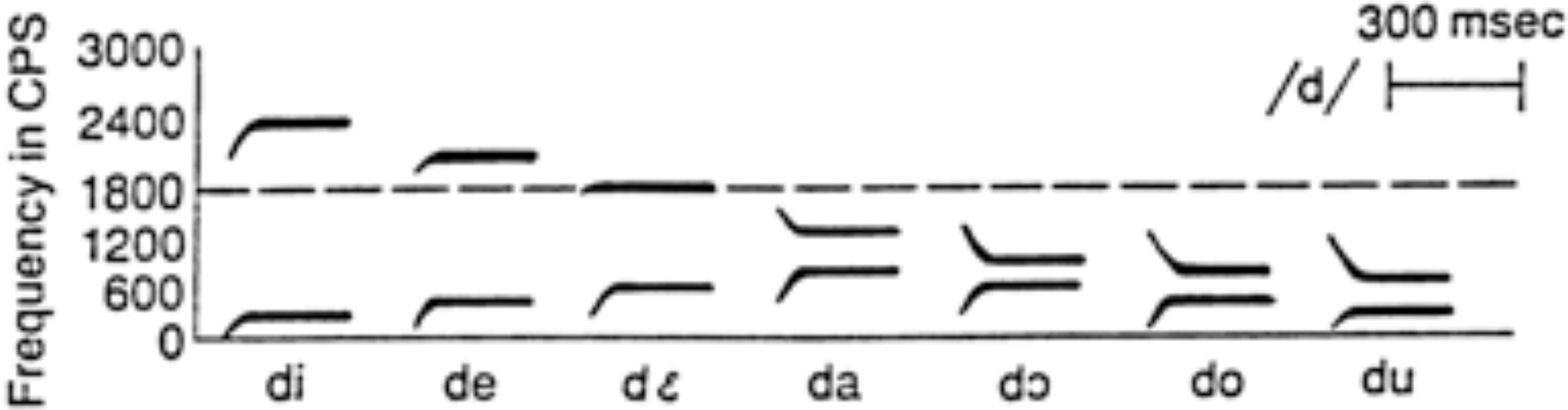
Producing speech



Perceiving speech

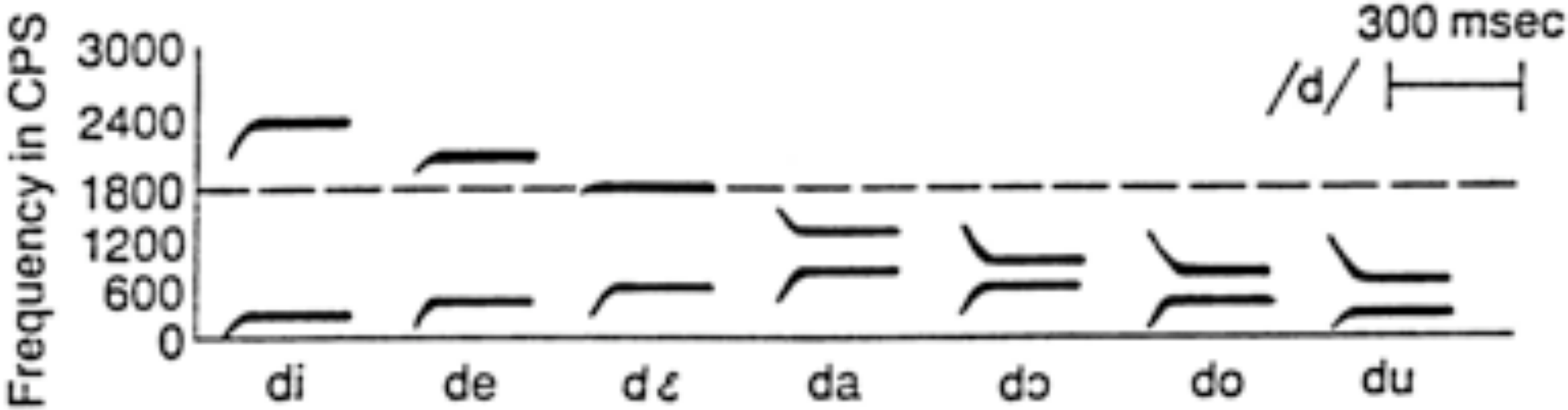


Perceiving speech



Formant: Local maxima of energy in the frequency spectrum that reflect the resonances of the human vocal tract. Can be seen as ‘dark bands’ in a spectrogram.

Perceiving speech



Perceptual constancy: perception remains constant despite variation - random or characteristic-in the sensory input.

Perceiving speech

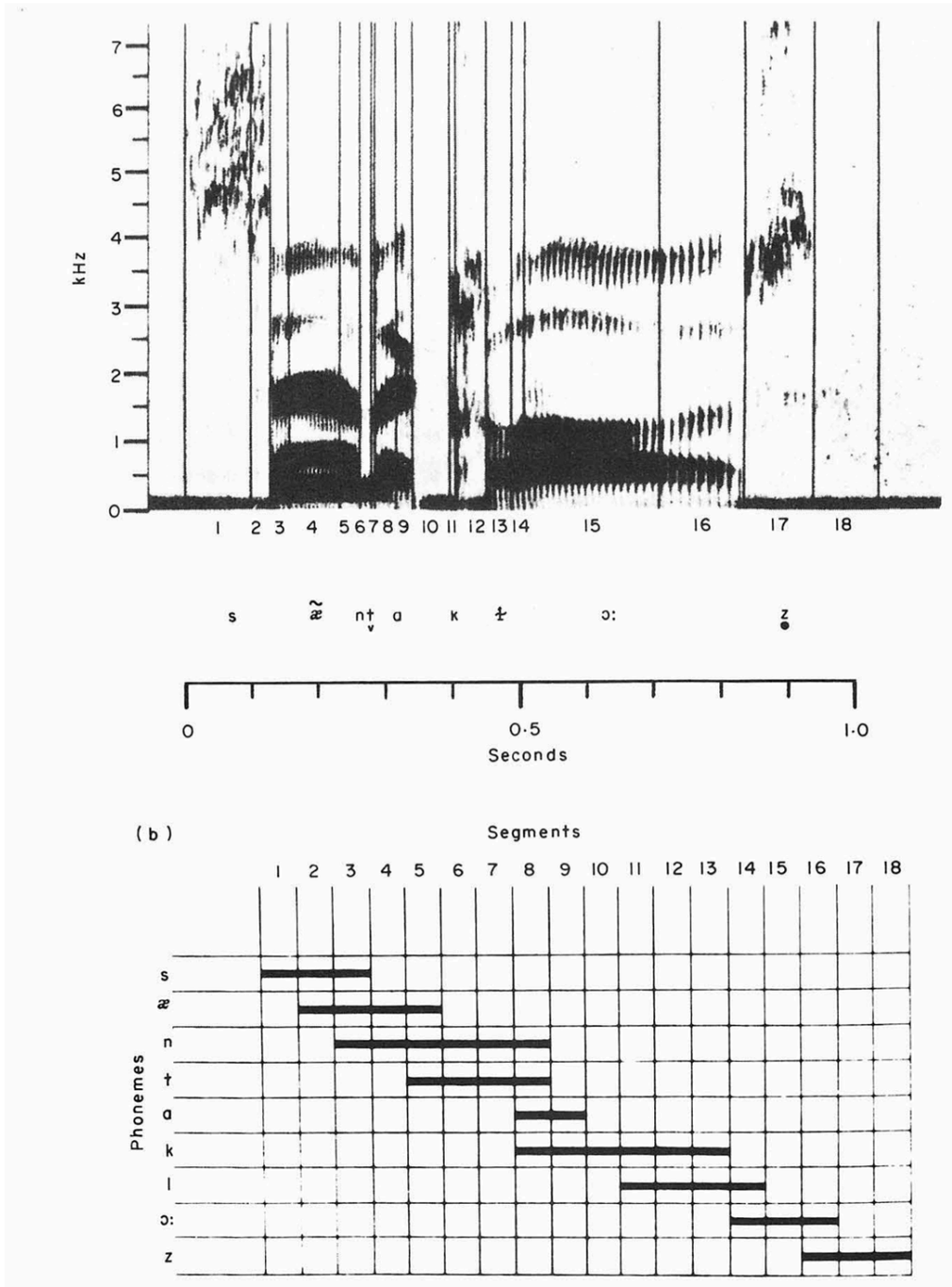
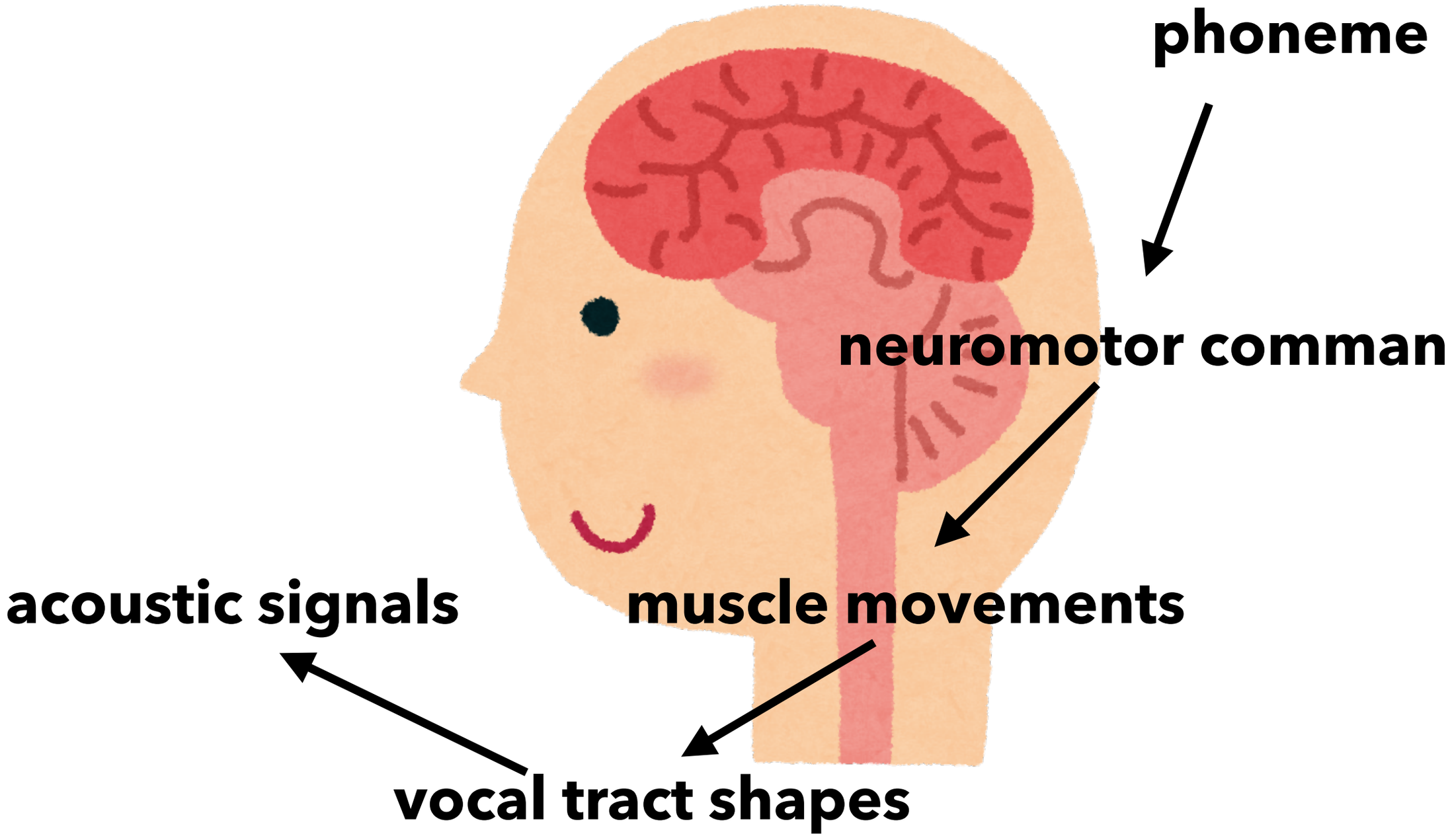


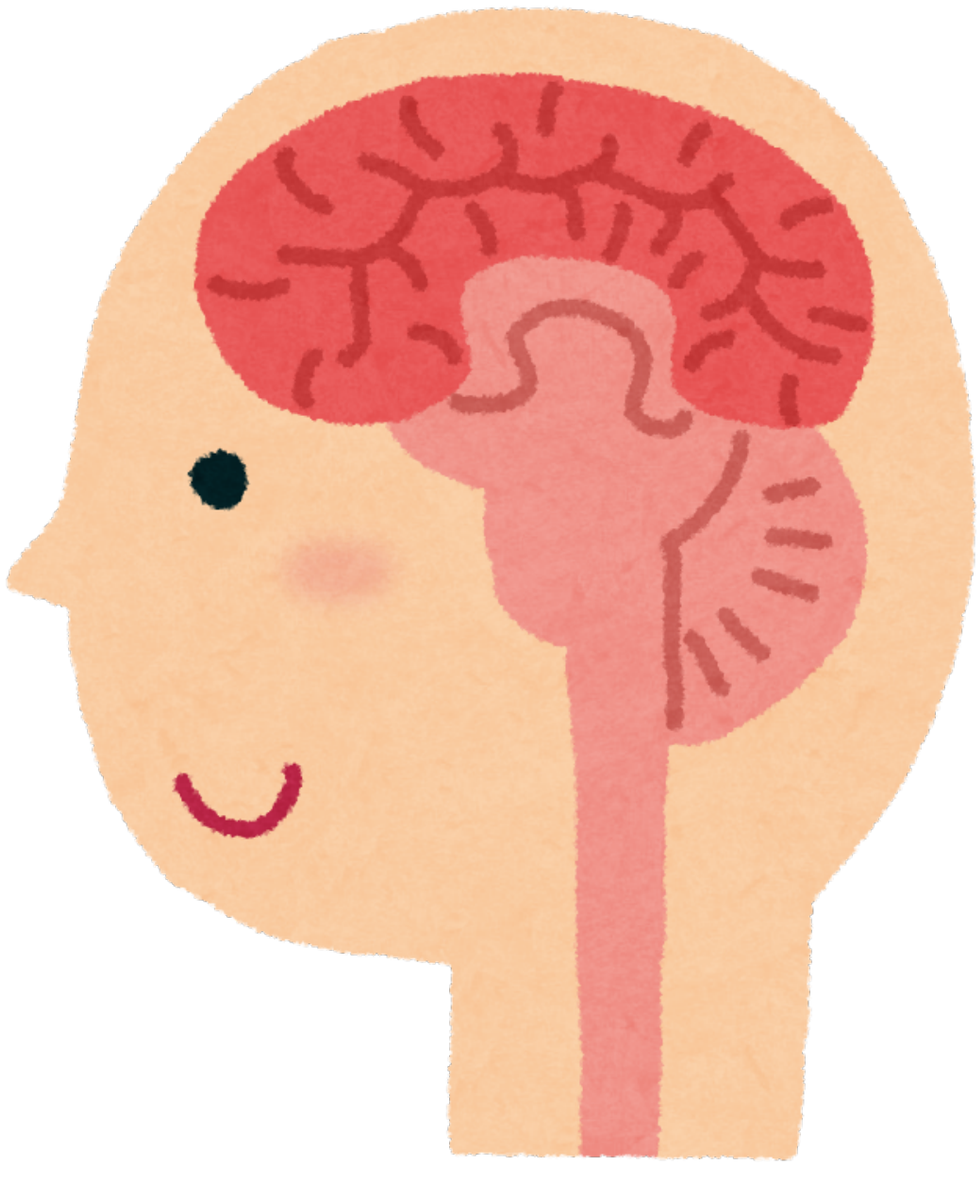
Figure 2. (a) Spectrographic display of "Santa Claus". (b) Schematic display of the relationship between acoustic and phonetic segments (reprinted with permission from Cutting & Pisoni (1978), and Fant & Lindblom (1961)).

From Fowler (1986)

Producing speech



Motor theory of speech perception

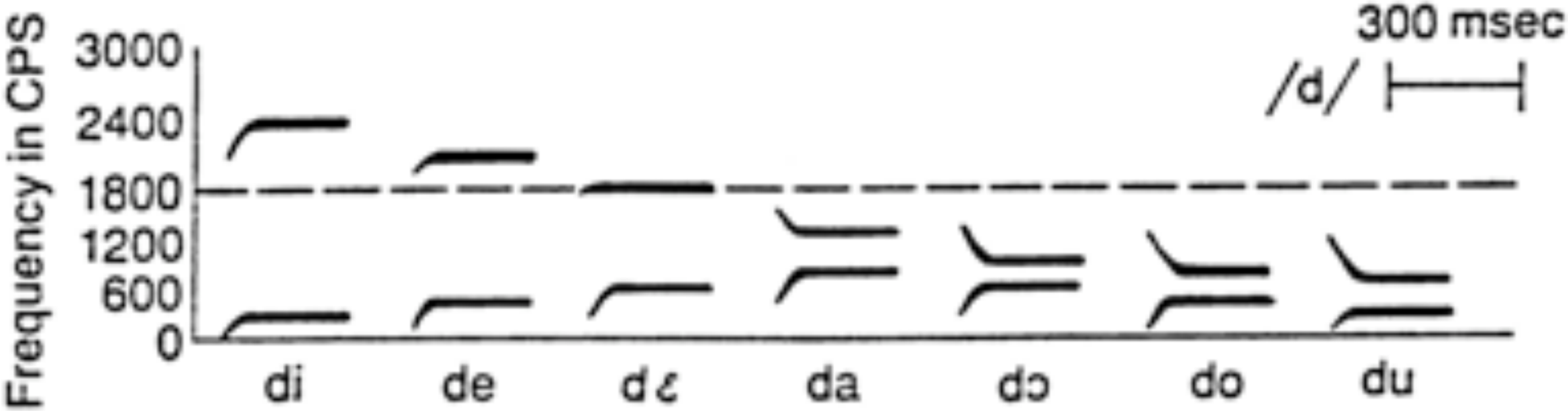


neuromotor commands

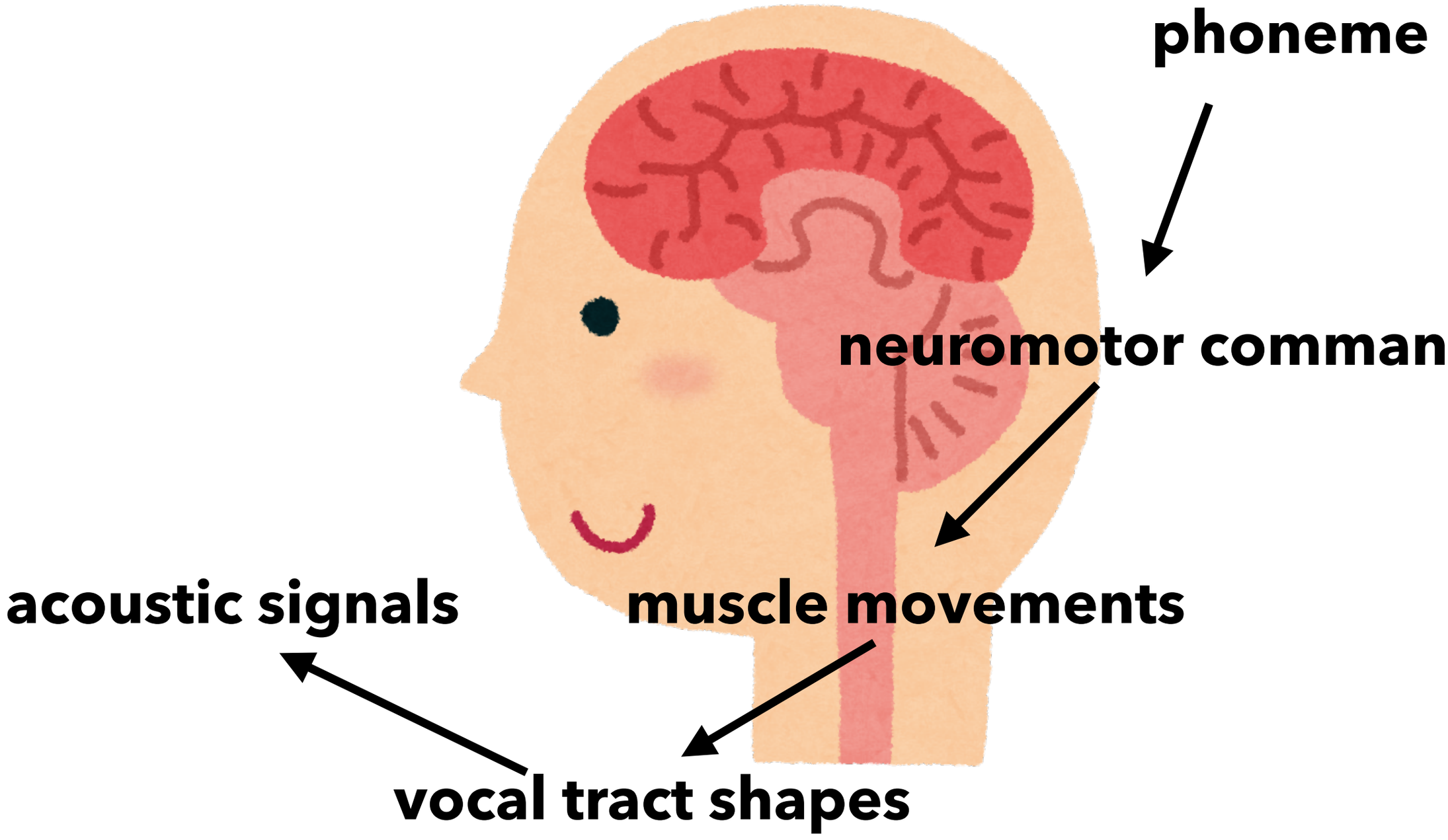
Motor theory of speech perception:

- Speech perception relies on a special decoder that infers the *intended gestures* (e.g. neuromotor commands) from the acoustic signal.
- "**Analysis by synthesis**": An internal synthesizer simulates the acoustic consequences of different neuromotor commands and matches them against the input.

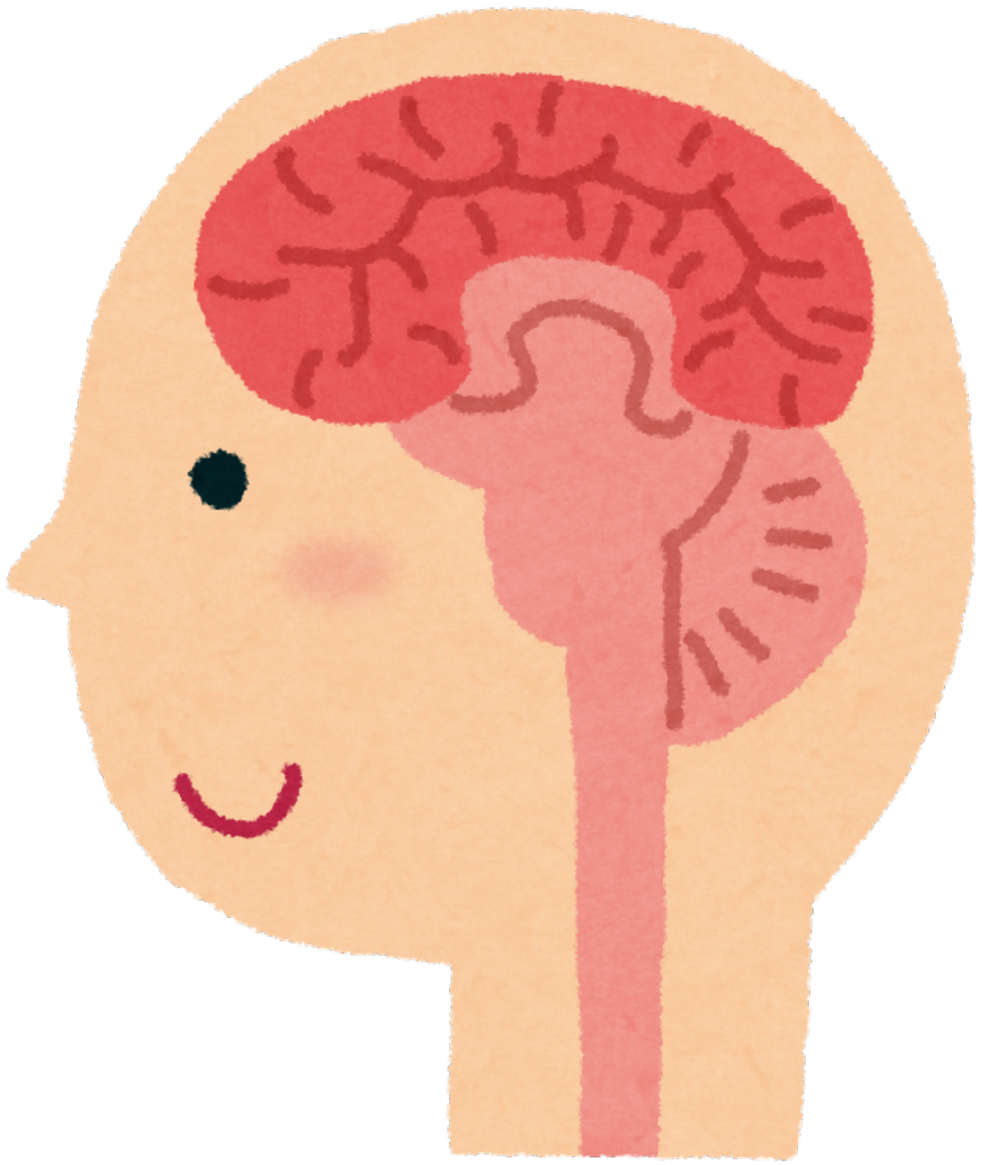
Motor theory of speech perception



Producing speech



Direct realism

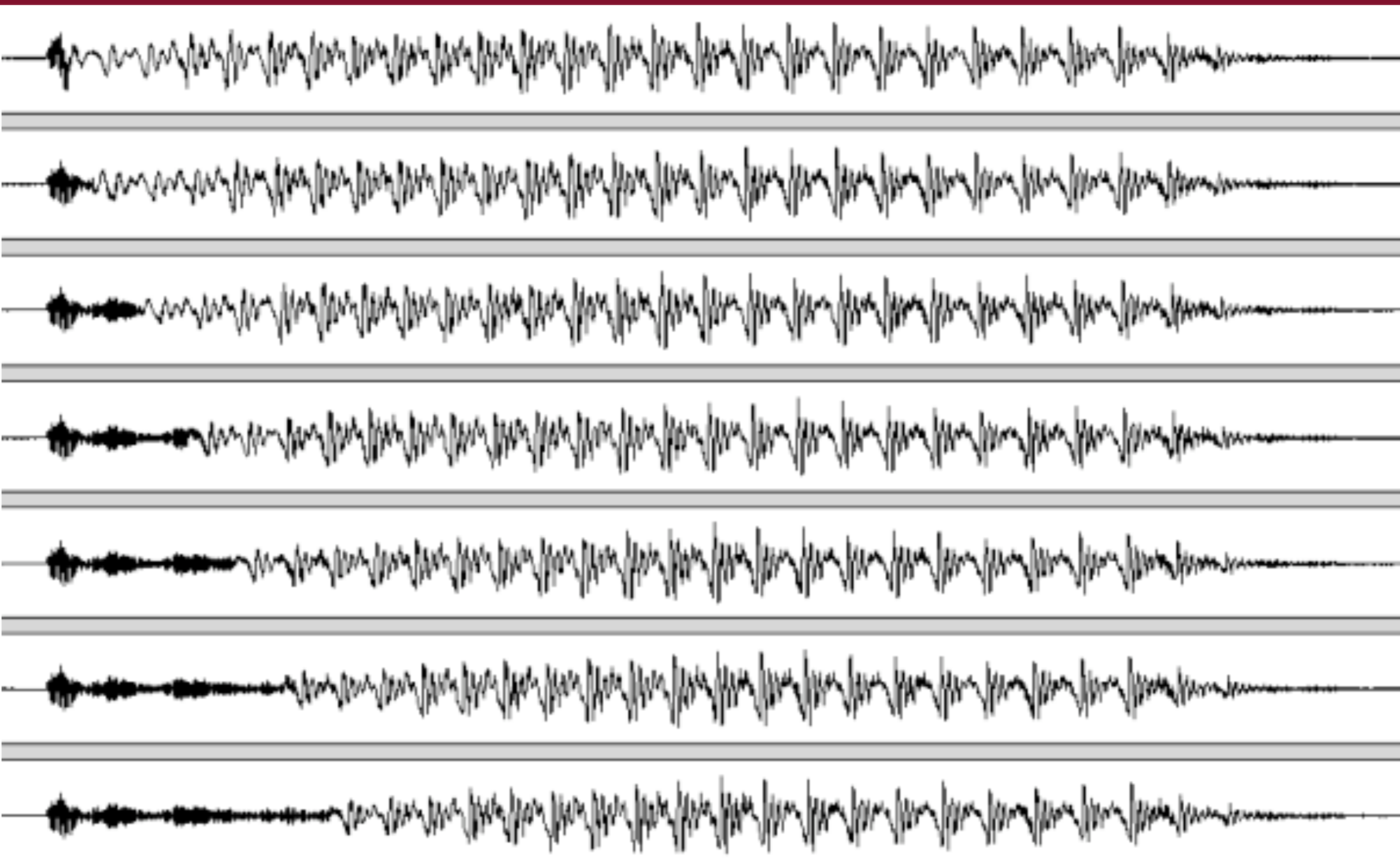


neuromotor commands

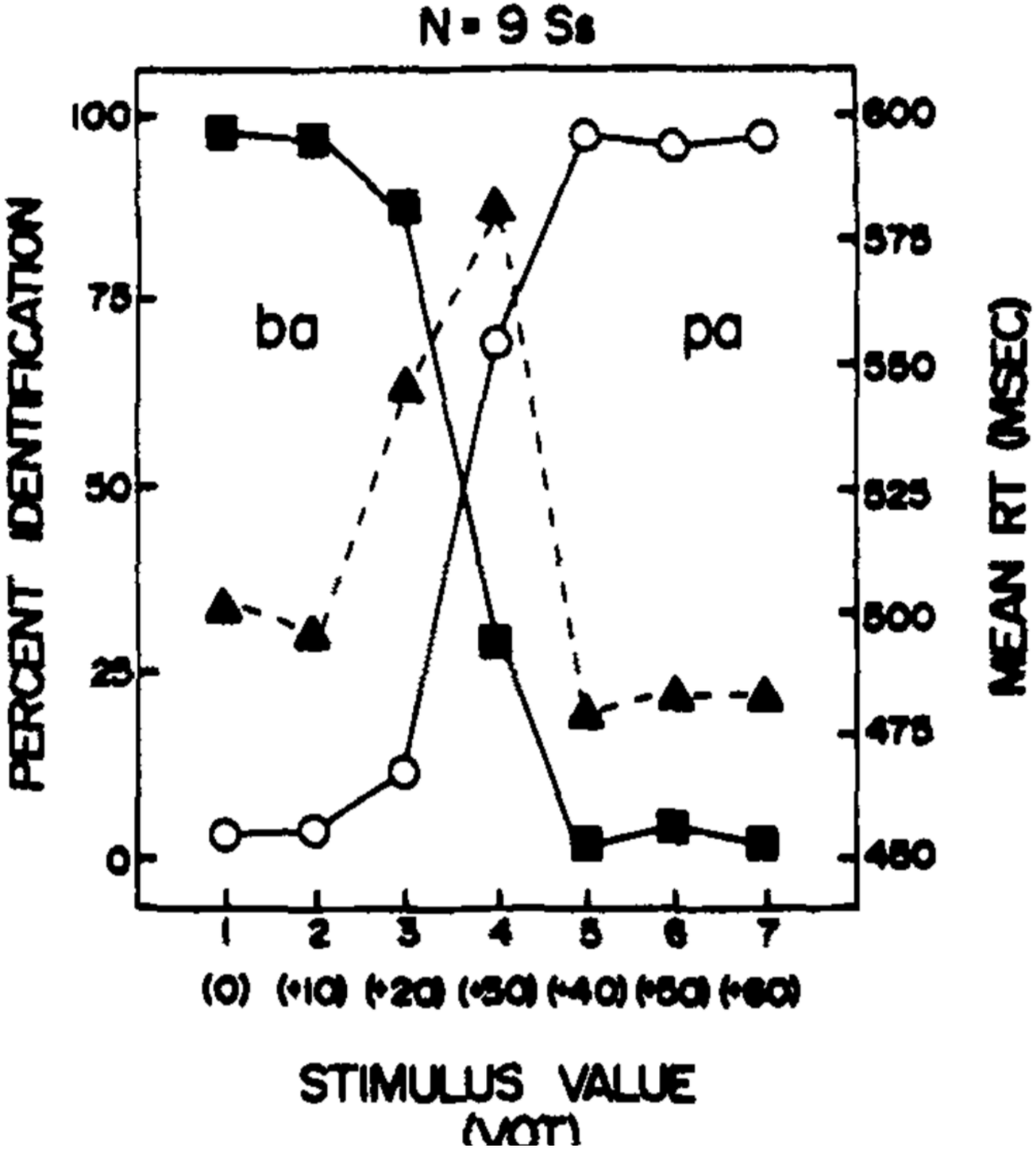
Direct realism in speech perception:

- Speech perception involves directly recovering the vocal tract gestures from the acoustic signal, which is rich enough to uniquely specify the vocal tract gestures.
- Does not rely on internal 'hypothesis testing': "The role of the perceptual system is not to create [phonetic segments], but to select [them]" (Fowler, 1986)
- No specialized mechanism is recruited for speech perception: General perceptual abilities suffice.

Categorical perception



Categorical perception

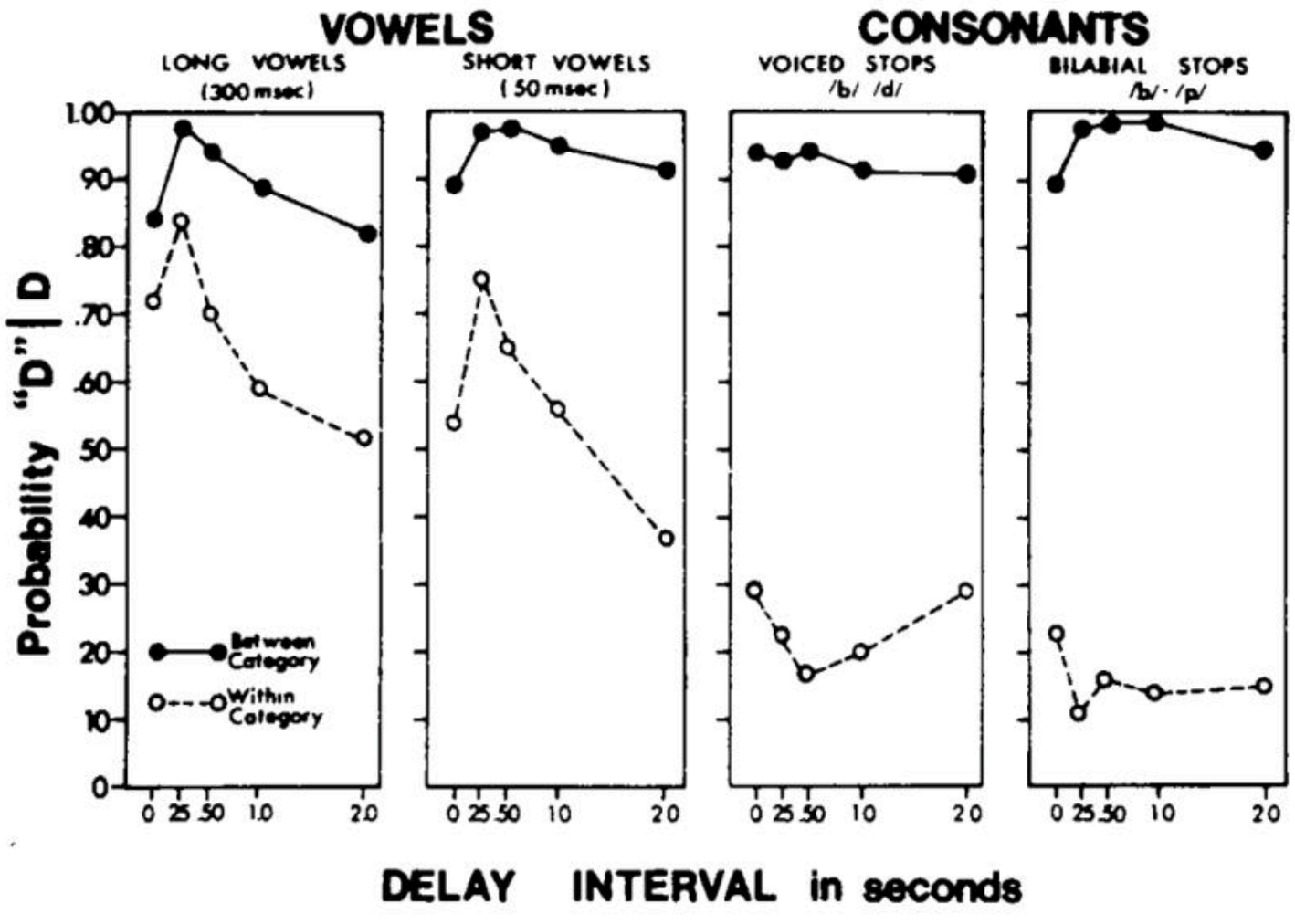


Categorical perception:

- Perception of phoneme categories is non-linear: Discriminability of speech sounds is enhanced around category boundaries, inhibited otherwise.

Pisoni & Tash (1974)

Categorical perception

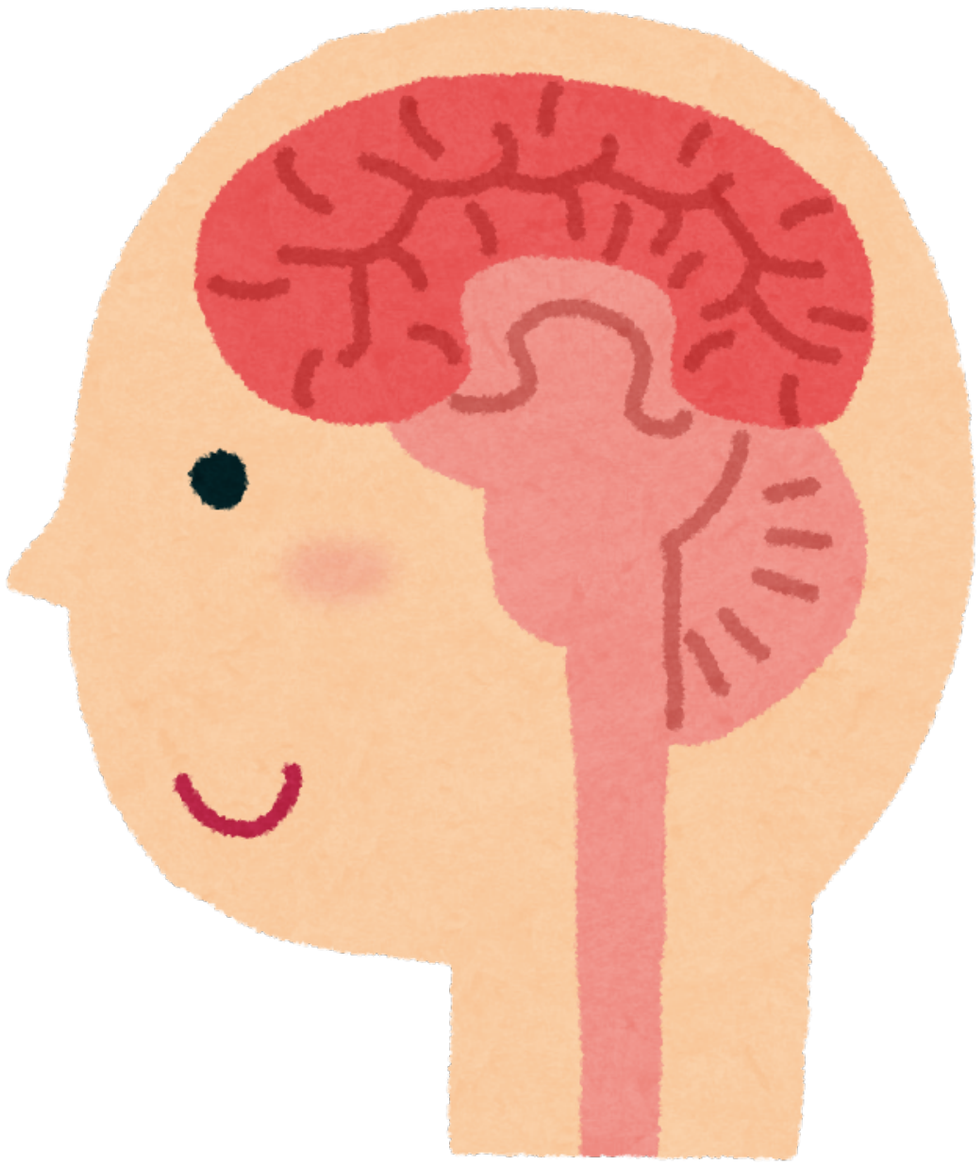


Categorical perception:

- Perception of phoneme categories is non-linear: Discriminability of speech sounds is enhanced around category boundaries, inhibited otherwise.

Pisoni (1973)

General approach



Phonemes

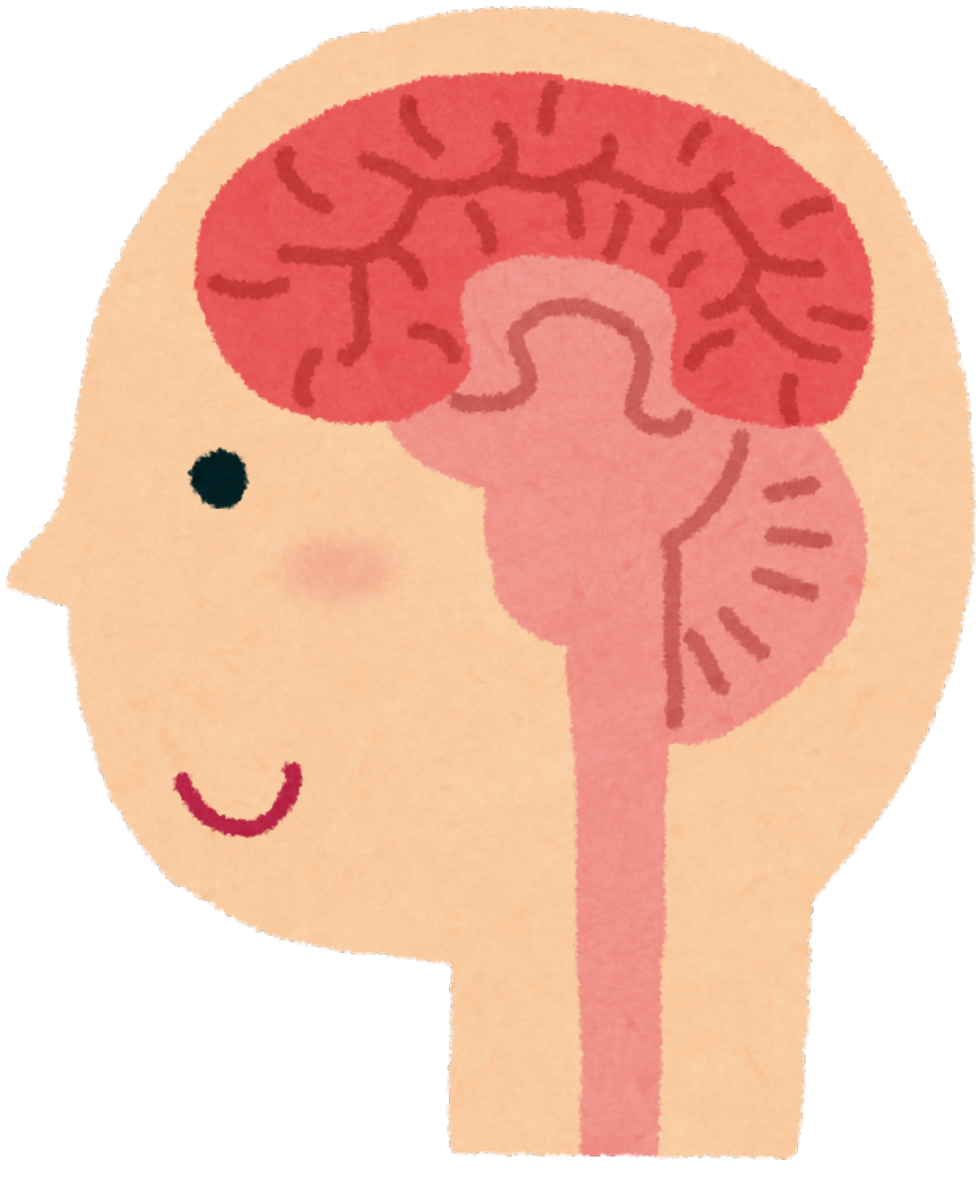
General approach to speech perception:

- Speech perception is not equivalent to recovering gestures.
- Categorization relies on multiple, imperfect **perceptual cues** (features of a stimulus) to categorize stimuli.
- Multiple cues allow for perceptual constancy if variation is predictable (structured).
- No specialized mechanism is recruited for speech perception: General perceptual abilities suffice.

Perceptual cues

McGurk effect

Differences in approach

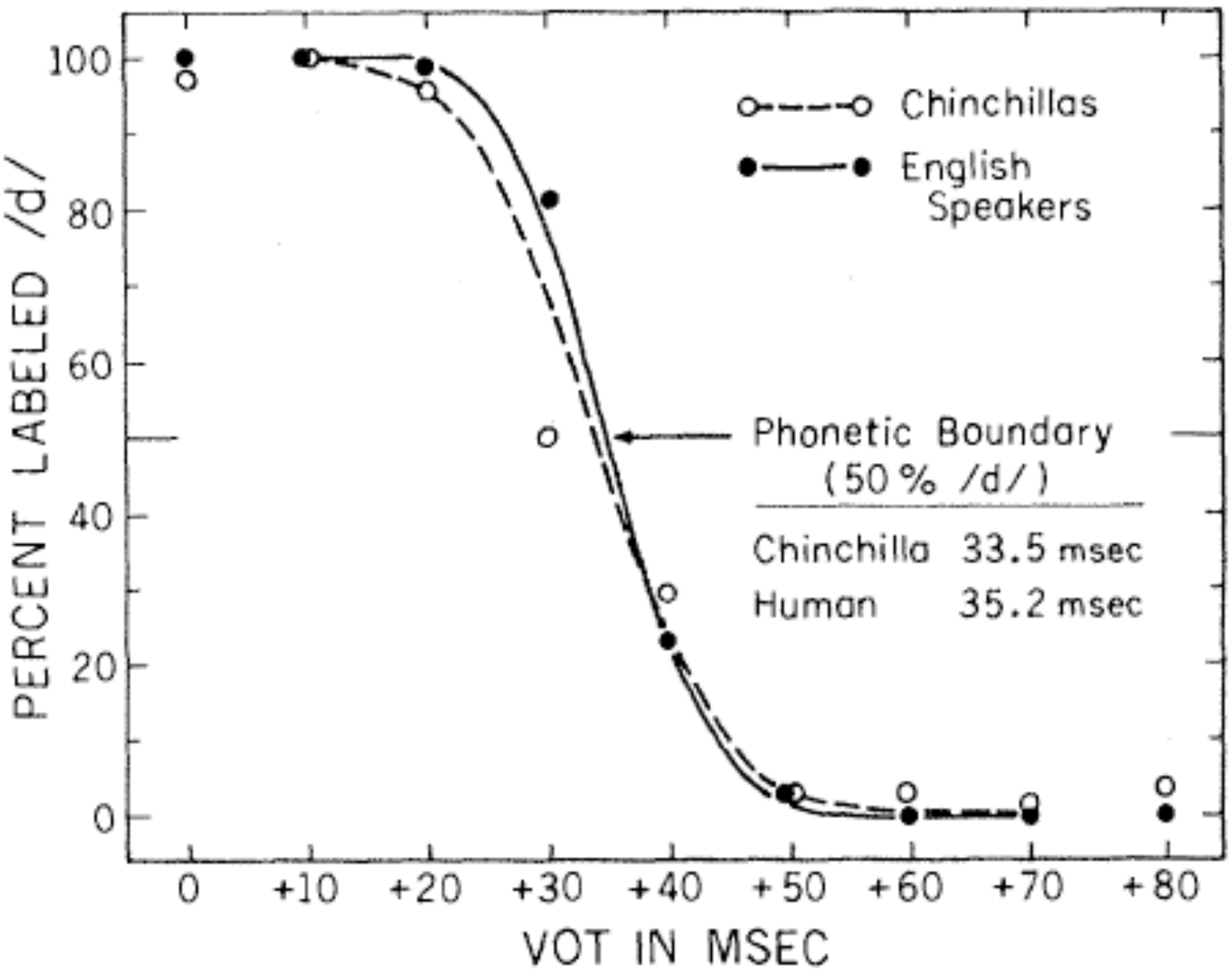


Theories vary in whether

- Speech perception involves perceiving real or intended gestures (Direct realism, Motor theory of speech perception), or not (General approach).
- Speech perception involves mechanisms unique to language (Motor theory of speech perception) or general perceptual mechanisms (Direct realism (contested), General approach).

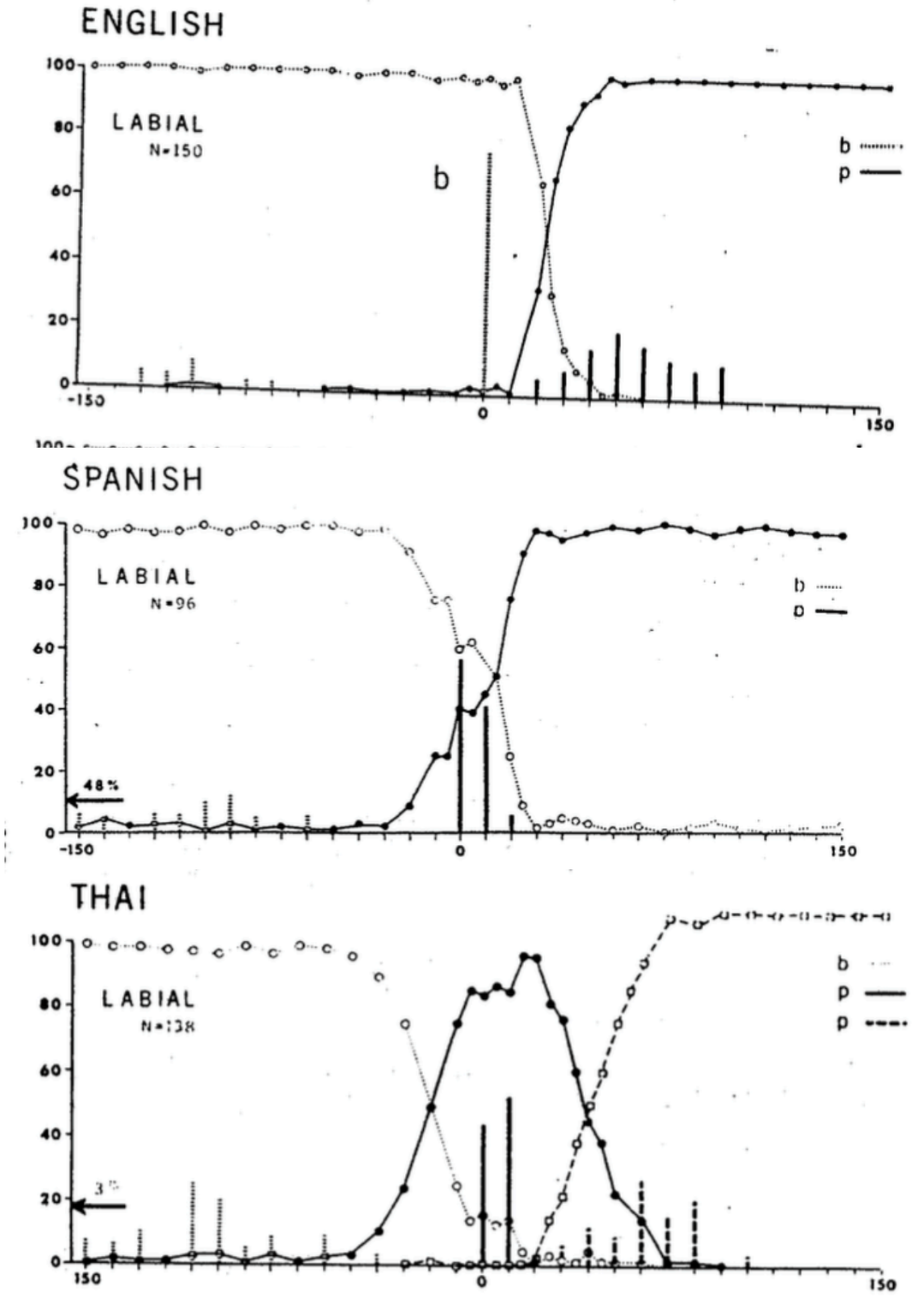
Phonemes

Unique to humans?



Kuhl & Miller (1975)

Perceptual learning



[pʰæt] - “pat”

[pæt] - “bat”

[pan] - “bread”

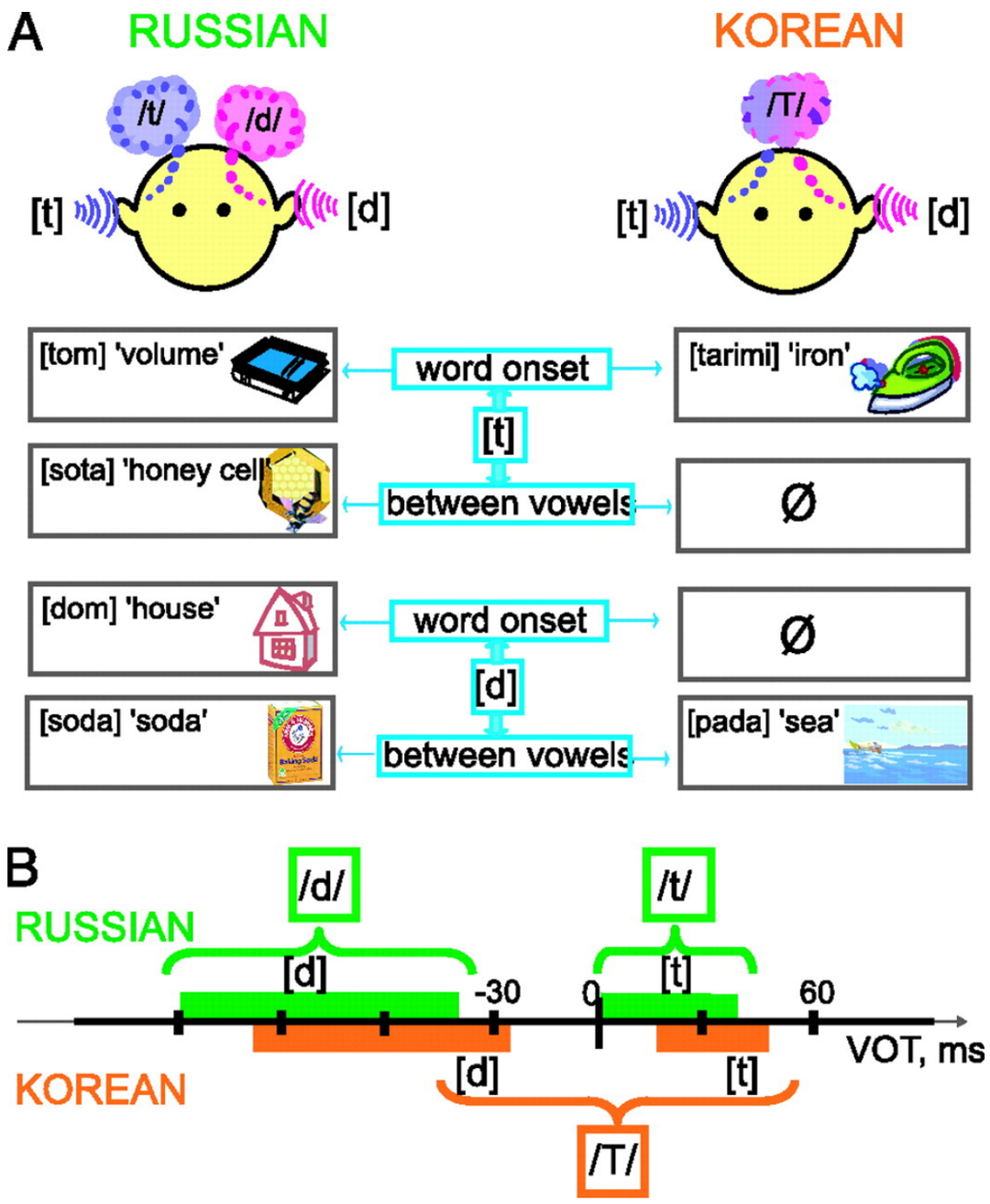
[ban] - “they go”

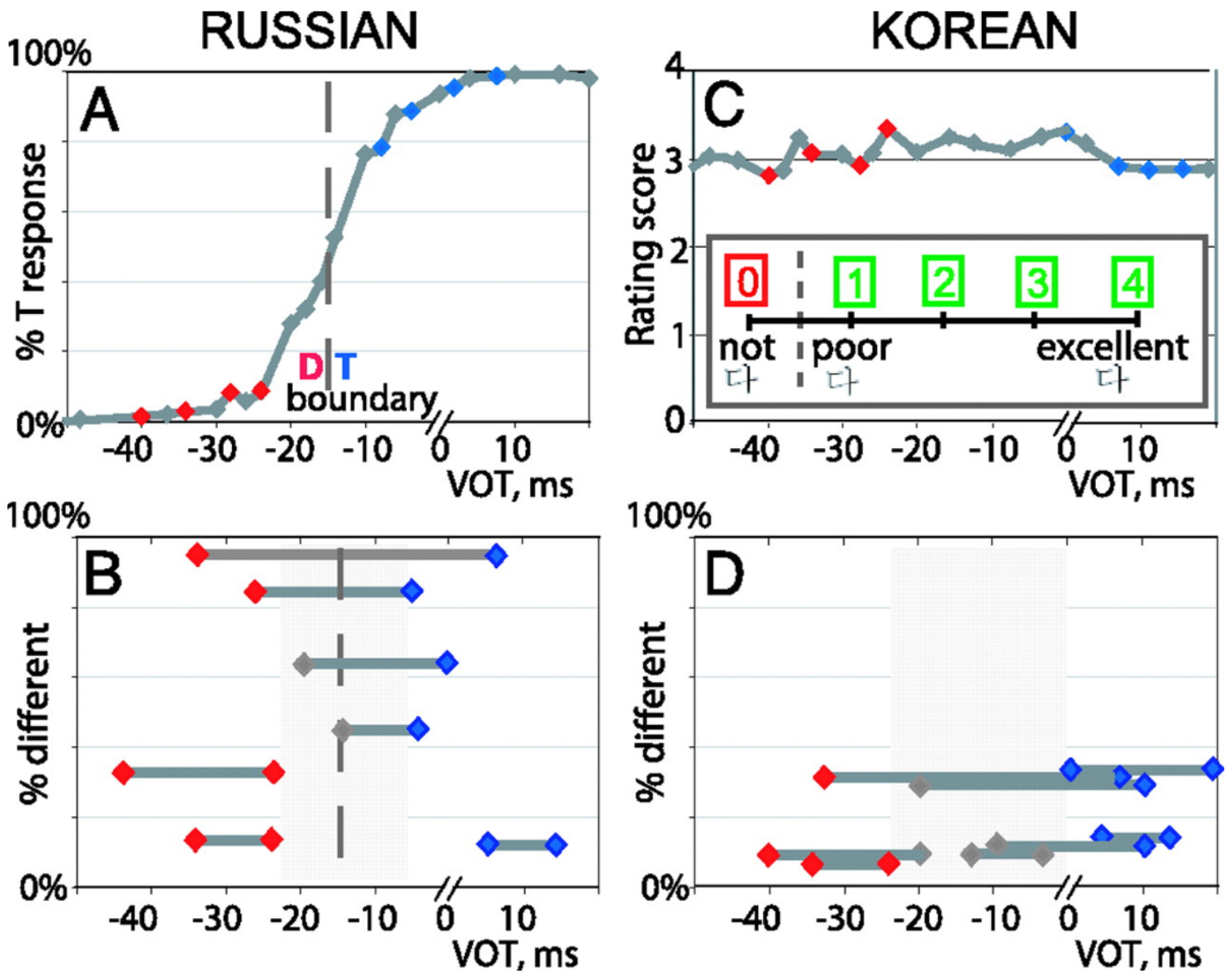
[pʰòt] - “blister”

[pòt] - “to lie”

[bòt] - “chapter”

The importance of contrast





Kazanina et al. (2006)

Categorical perception

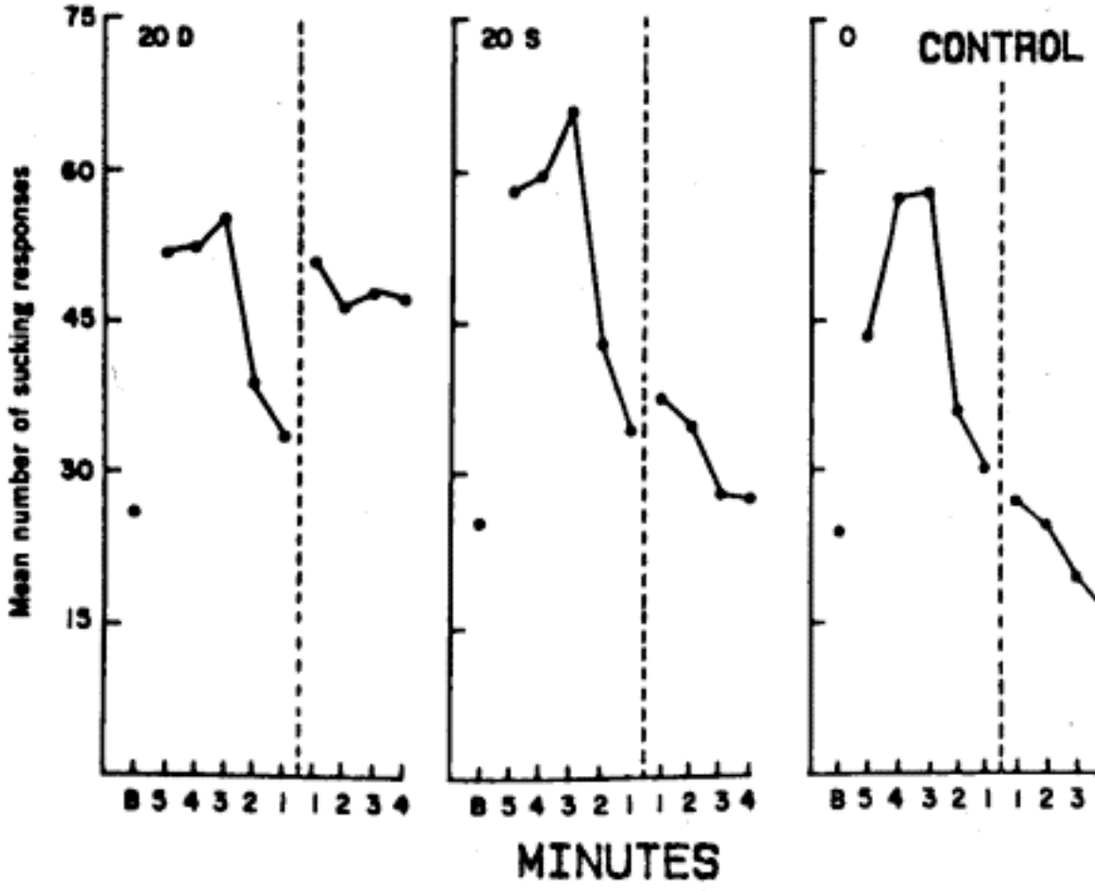


Figure 4.7
Mean number of sucking responses for 4-month-old infants as a function of time and experimental condition. The dashed line indicates the occurrence of the stimulus shift, or, in the case of the control group, the time at which the shift would have occurred. Adapted from P. D. Eimas, E. R. Siqueland, P. W. Jusczyk, and J. Vigorito (1971). Speech perception in infants. *Science* 171, 303–306. © 1971 by the AAAS.

Eimas et al. (1971)

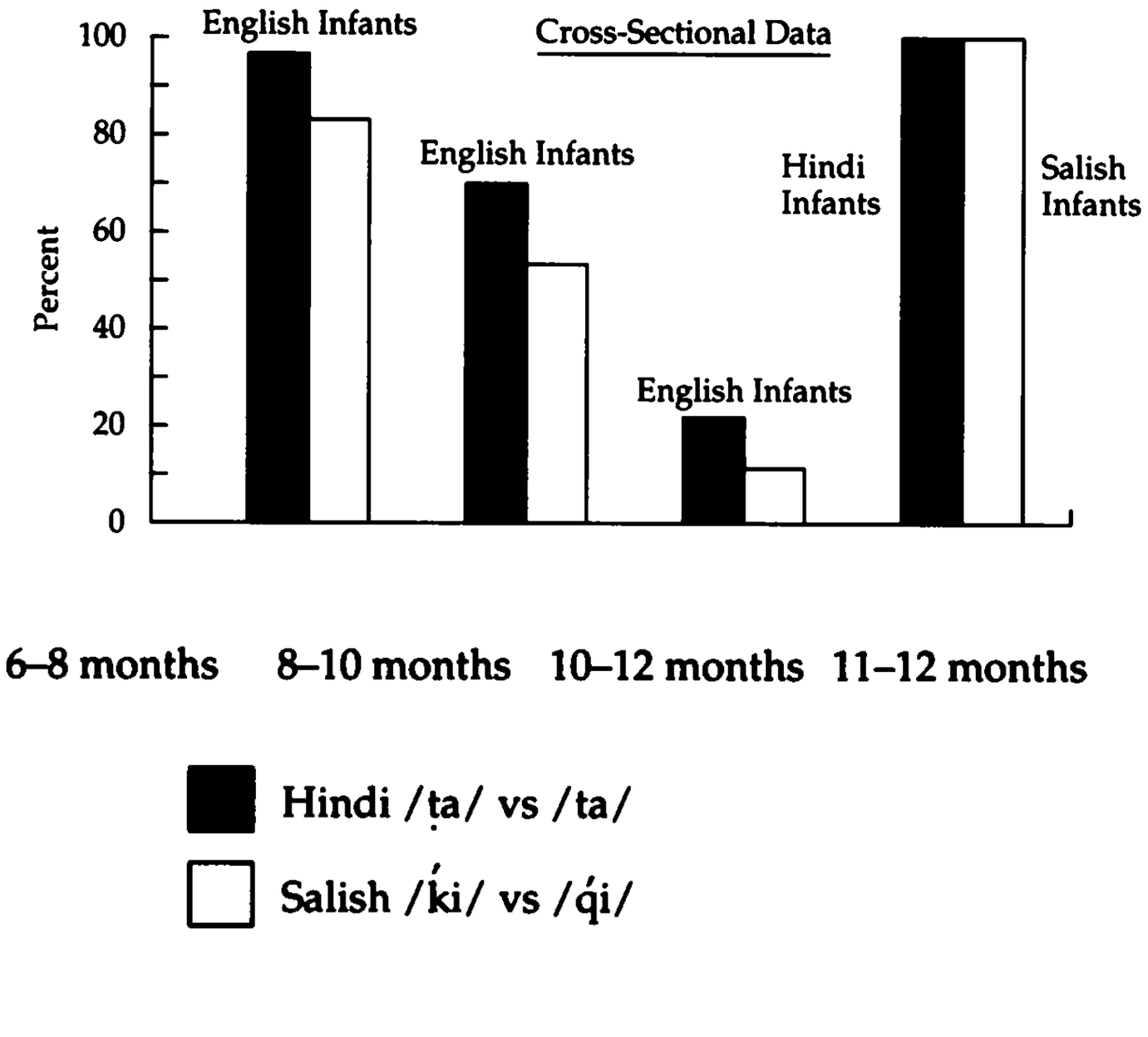
- Used *high amplitude sucking* procedure to test for categorical perception in infants.
- Infants were played sounds until they got 'bored' (slowed down rate of sucking on pacifier).
- The stimulus sounds were then changed to a different sound, within category or across category.
- Dishabituation occurred across the category boundaries.

Eimas et al. (1971)

Categorical perception

JANET WERKER, THOMPSON SALISH & CONDITIONED HEAD TURN PROCEDURE

Learning sound contrasts



- Infants show sensitivity to contrasts not in their native language, but sensitivity declines over the first year of life.
- Experience with a contrast (e.g. retroflex versus dental place of articulation) allows for perceptual maintenance of contrast.

Summing up:

- Speech perception is characterized by perceptual constancy in the face of substantial acoustic and physical variation in speech sounds.
- Classical theories of speech perception proposed that to perceive speech is to perceive real or intended vocal tract gestures.
- The general approach instead holds that speech perception relies on integrating multiple cues to speech perception.
- Preverbal infants and animals show sensitivity to some category boundaries - suggests that languages exploit ‘natural’ acoustic boundaries for phonetic contrast.
- Linguistic experience can modify categorical boundaries, or lead to loss of perceptual contrast.