Language acquisition, perception and production

Lecture 3 – Speech comprehension

Basic elements of speech

- Cat = /c/ + /ae/ + /t/

- Rough = /r/ + /^/ + /f/

Outline

- Properties of speech

- How do we understand it

Properties of speech

- Two levels of describing speech sounds
  1. Phonetics
     - Acoustic detail of speech sounds (physically)

  2. Phonology
     - Sound categories in each language

- /p/ in spin and pin

Properties of speech

- /p/ in pin is aspirated
- /p/ in spin is unaspirated

- Physical sounds are different, does not matter in meaning
  - Saying pin unaspirated does not change meaning
 - But it does in other languages (e.g., Thai)

Properties of speech

- **Phoneme is basic unit of sound in a language**
  - In English p in pin and p in spin are same phoneme
  - In Thai, p in paa and p^{h}aa are different phonemes

- **In Thai, the two p sounds are *phones***
  - Phonetic difference creates meaning difference

- **In English, the two p sounds are *allophones***
  - Phonetic difference does not create meaning difference
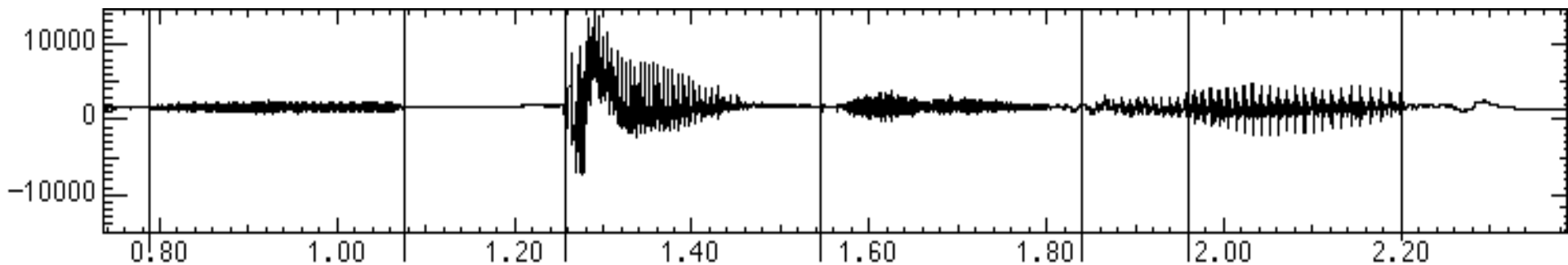
Minimal pairs

- **Two words that differ by only one sound**
  - Bat – pat
  - Dog – cog

- **If sound difference = phoneme then difference in meaning**
- **If sound difference = phone then sometimes no difference in meaning (pin – p^hin).**
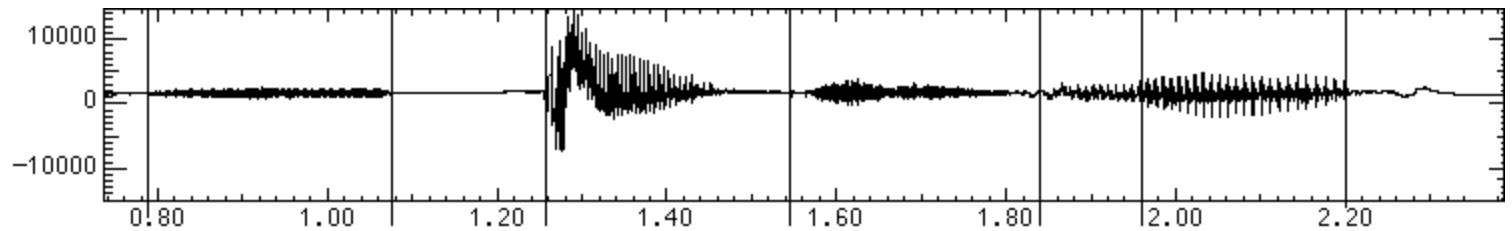
How are the sounds made?
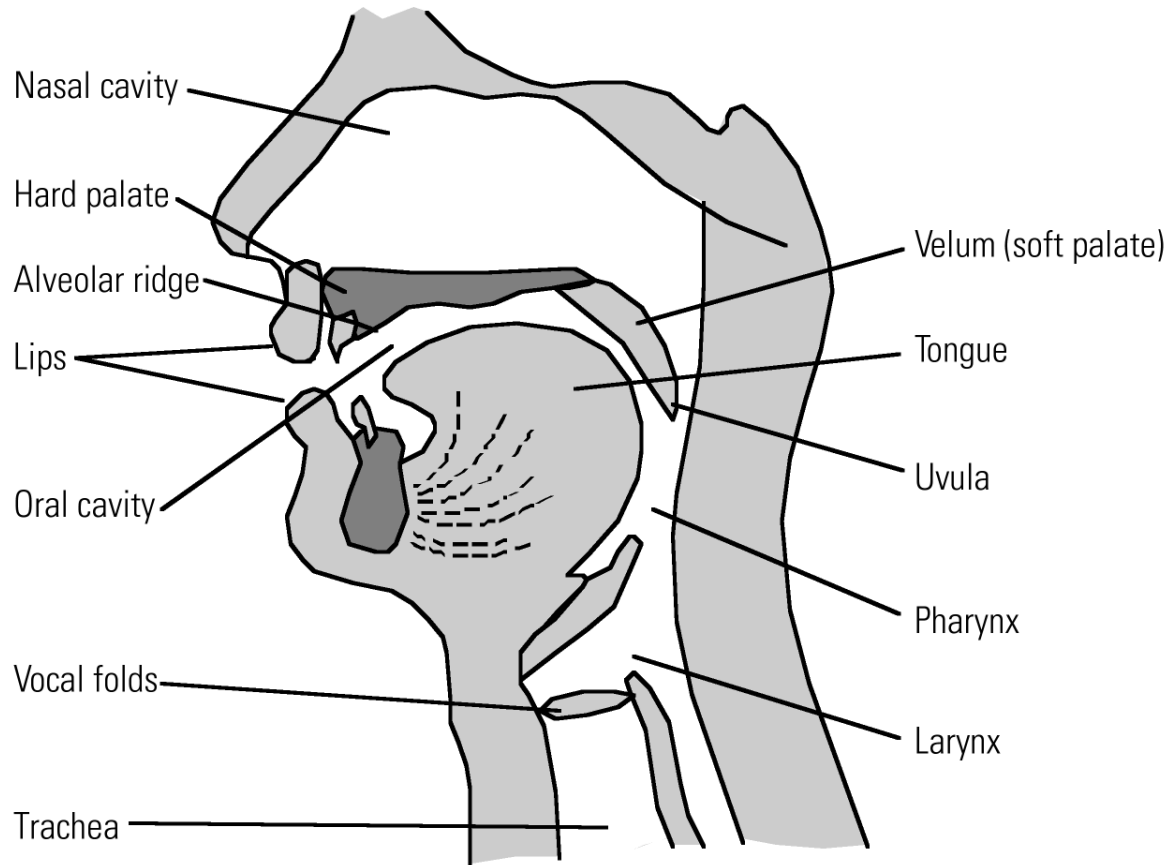
Properties of speech

- **How do we visualize speech?**
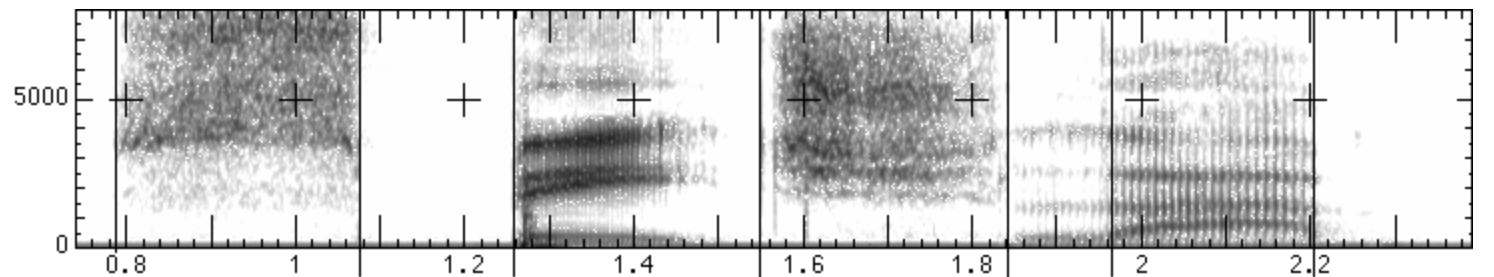 - Pressure waves
  - Spectograms

Properties of speech

Properties of speech

- Formants

Spectrogram

formants

Frequency

f3

f2

f1

Time
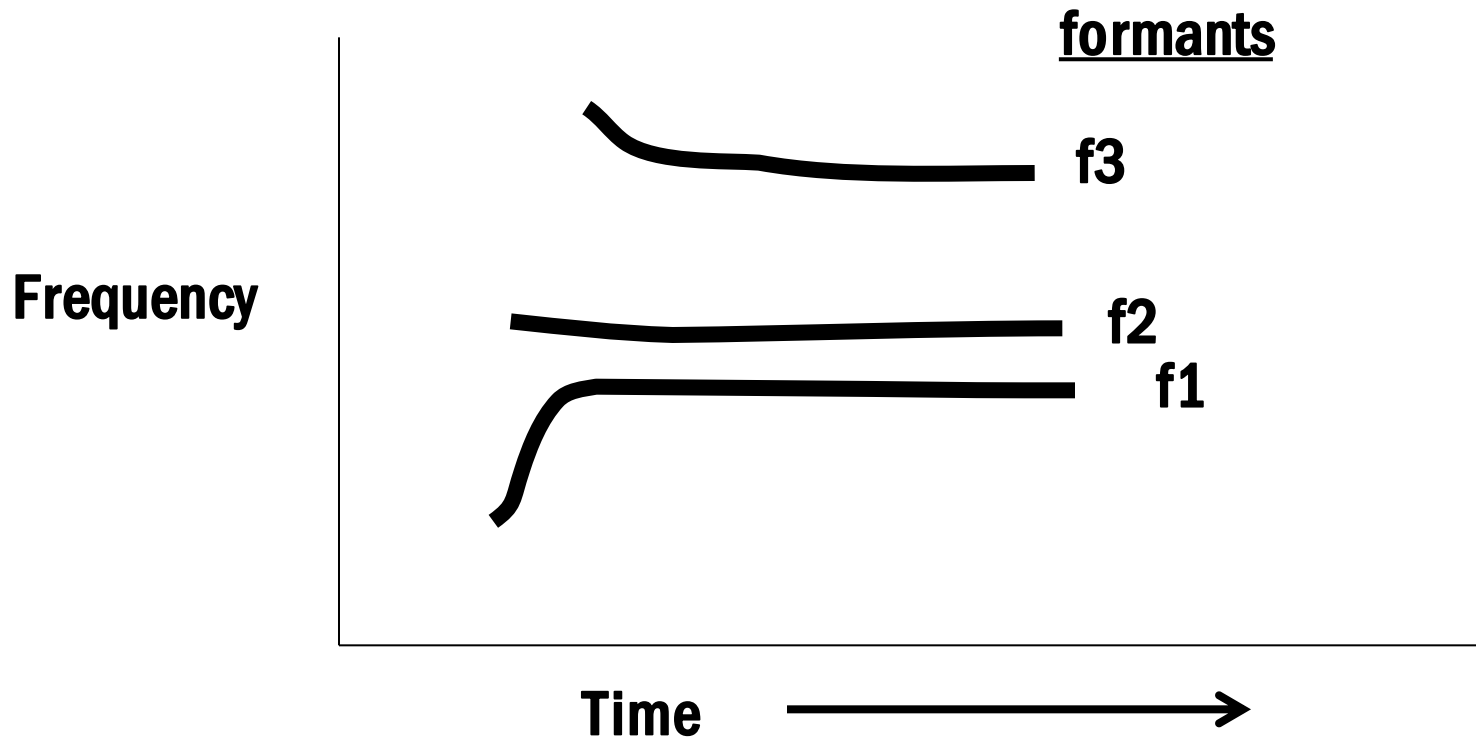
Phonemes of the world

- English has 40 phonemes

- Polynesian has 11, Khoisan has 141 ("bushman")
- In total there are thousands of phonemes

- Some are found in many languages (e.g., /m/, /n/, /t/, /d/, /k/, /g/, /s/, /z/)

- Easy to produce, easy to distinguish

Articulatory features

- Consonants are produced by restricting vocal tract
  - Place of articulation (dental, velar)
  - Manner of articulation (stop vs nasal vs fricative)
  - Voicing (voiced, unvoiced)

- /b/: voiced, labial, stop
- /p/: unvoiced, labial, stop

Phonology - International Phonetic Alphabet

Consonants

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k g	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	N		
Trill	B			r					R		
Tap or Flap				ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	L			

Articulatory features

- **Vowels: Unrestricted vocal tract**
  - Part of tongue (front vs back)
    - Beet vs boot
  - Position of tongue (high vs middle vs low)
    - beet vs bat

Vowels Spread Throughout Mouth
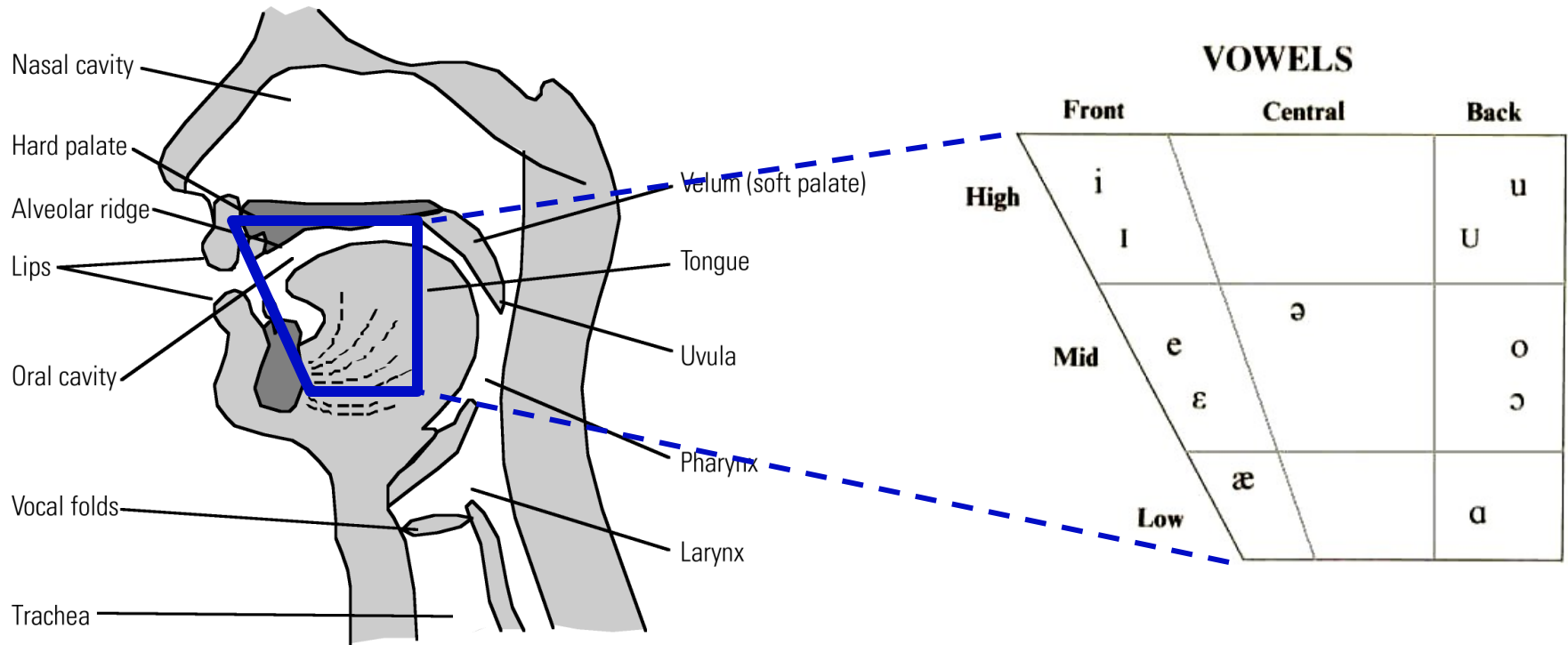
Phonemes are not produced serially

- How is speech produced?

- "cat" is not just /c/ + /ae/ + /t/
- "rough" is not just /r/ + /^/ + /f/

- 1940s and 50s reading machines for blind
- Speech perception is not like reading: Sticking sounds together sounds bad!

Phoneme speech

- Video of phoneme concatenation

Properties of continuous speech

- Coarticulation = adjust pronunciation of current sound to take into account preceding and following sounds

- Segments overlap, we can get out more in a shorter amount of time

- Fast (15 sounds/s), articulators not always in ideal position so we need to cheat

Coarticulation

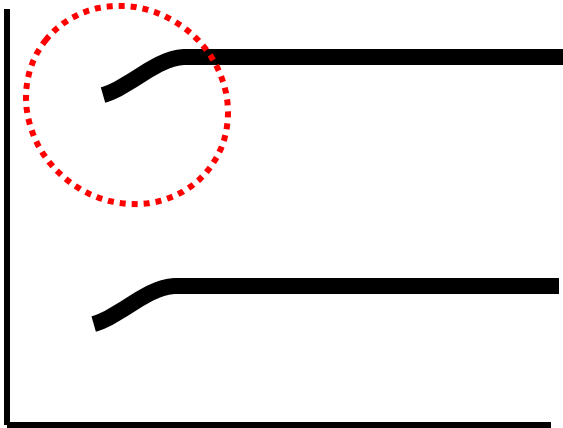- **Coarticulation**
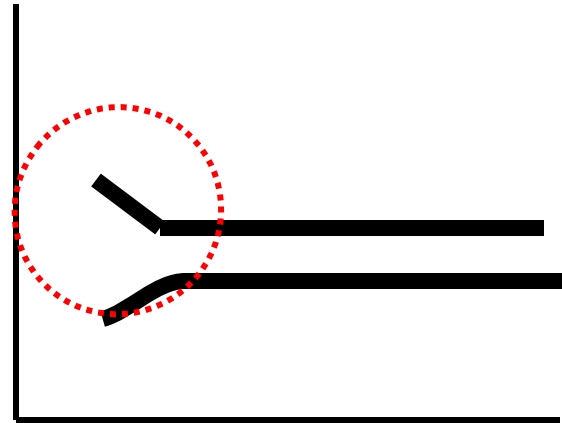 - Parallel Transmission
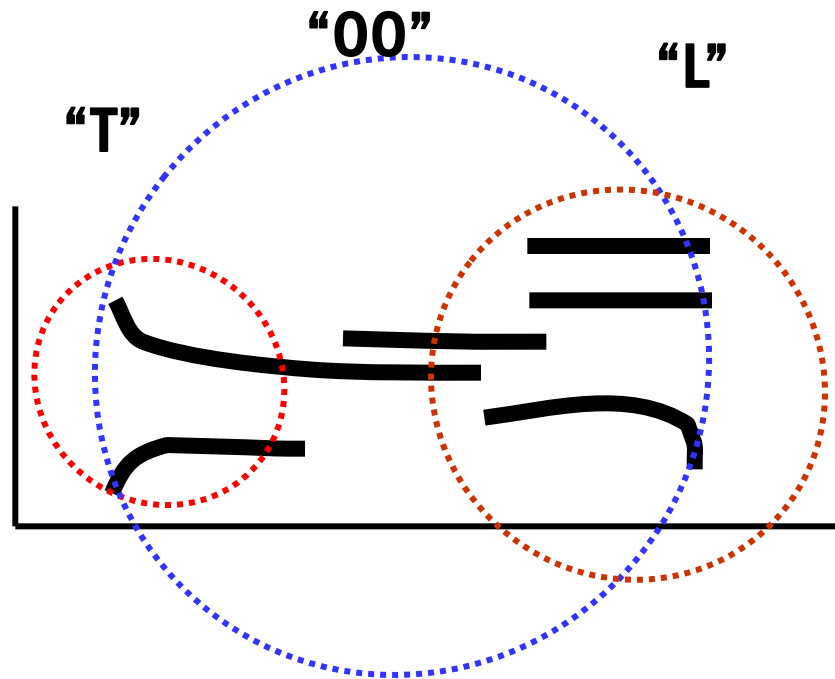
Coarticulation

- ## Context Dependence

"di"   **"du"**

Not segments, but features

- Speech is a trajectory through a sequence of articulatory targets

- You are anticipating upcoming sounds, this makes things easier to pronounce

- This is why coarticulation arises

Plan for today

- Properties of speech

- How do we understand it

How do we understand it?

- Fast: 15 sounds/sec, 30/s in fast speech

- Parallel transmission: sounds blend into each other
  - Each chunk of speech contains evidence of multiple phonemes

How we understand it?

- Prosody
  - Same word can be pronounced differently
    - "is that a car???" versus "look a car"
- Emotional state
  - Smiling
  - Frowning
  - stressed
- Different speakers
  - Female voices, male voices, etc

How do we understand it?

- Context conditioned variation

    – One to many variation: Same phoneme may be superficially realized in different ways

    – Many to one variation: Different phonemes can have the same sound in different contexts

How do we understand it?

- Problem of "invariance"

- Solutions:
  - Acoustic features
  - Categorical perception
  - Context

How we understand it

- **Solutions: Acoustic features**
  - Some features are invariant!
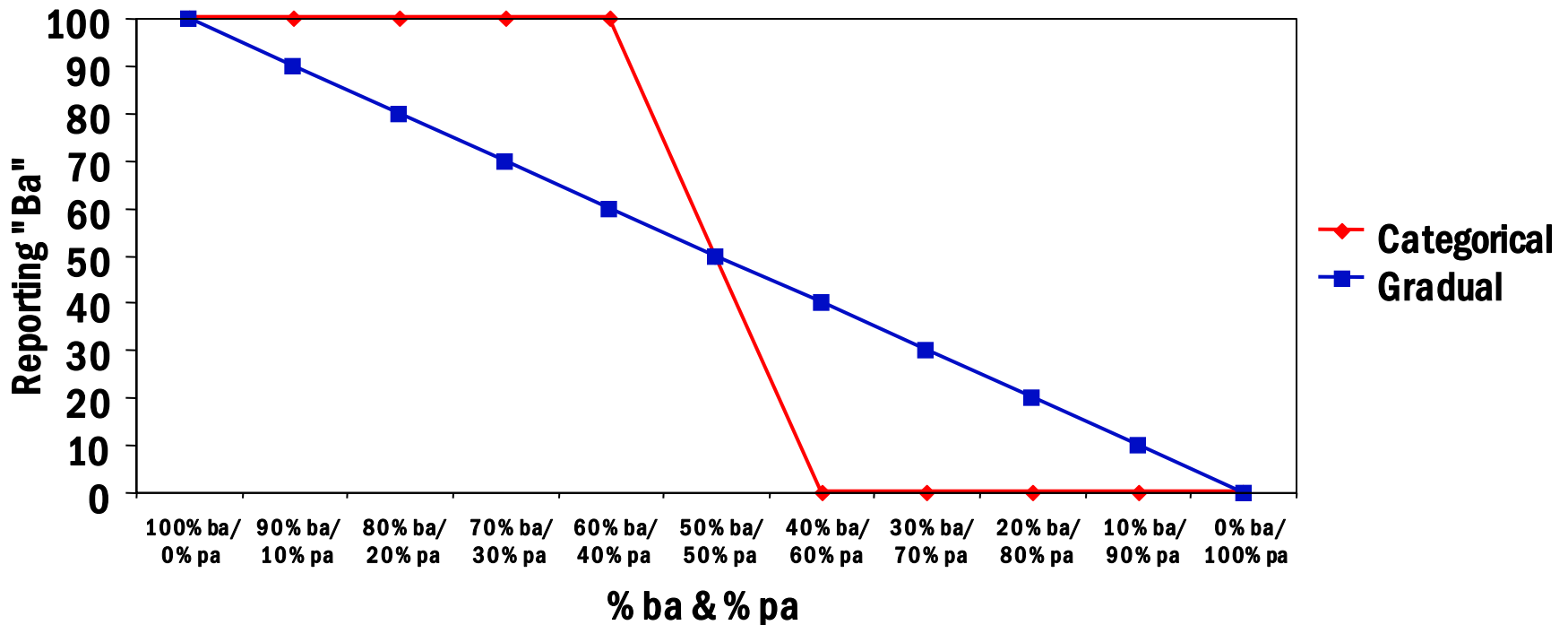
- **Stops, fricatives, vowels**

How we understand it

- Solutions: Categorical perception

- We impose categories on physically continuous stimuli

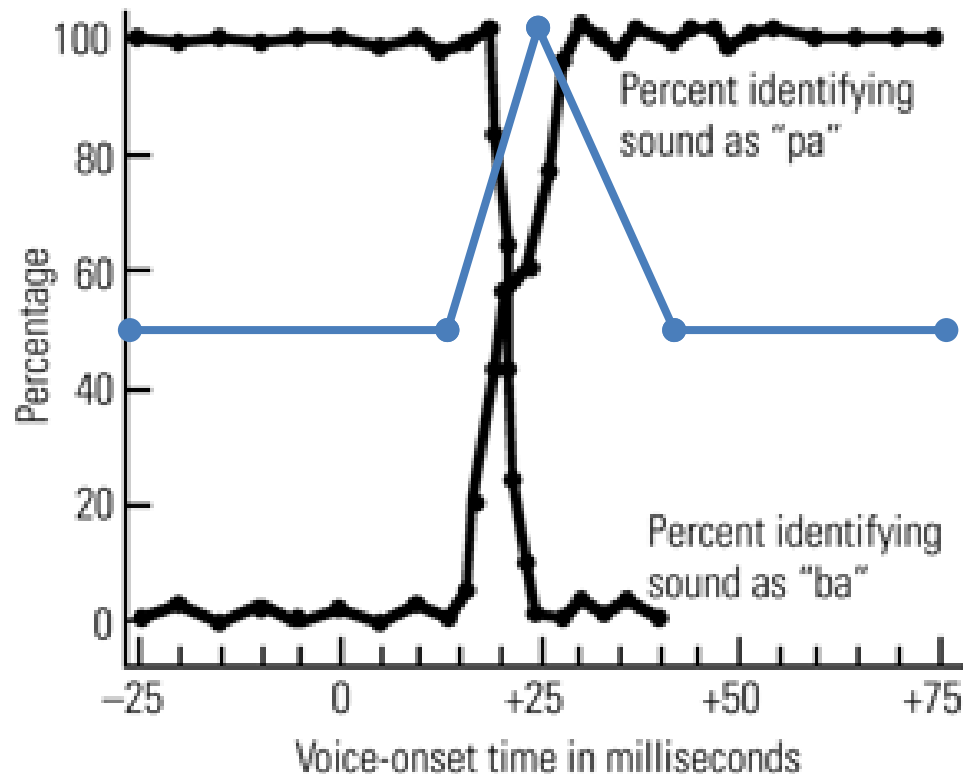- /ba/ to /pa/ continuum, by varying Voice Onset Time (VOT)

Varying Between Phonemes

Can vary 'Ba-ness' by altering Voice Onset Time

Predictions:

Categorical Perception

Phoneme distinction acts like 'classic categorization'

All-or-none category membership

Percent identifying sound as "pa"

Percent identifying sound as "ba"

Percentage

Voice-onset time in milliseconds

Categorical Perception

- What does Categorical perception do?
  - Ignore irrelevant information
  - Quickly classify transient events

Context - McGurk effect

How we understand it?

- **Solutions: context**
  - Use knowledge of how surrounding segments are articulated to interpret ambiguous segments

- **Rate Normalization**
  - Correct for speaking rate

How we understand it?

- Use higher level information
  - Noisy perception (Miller, Heise, Lichten, 1951)
    - Grammatical: Accidents kill motorists on the highway.
    - Anomalous: accidents carry honey between the house.
    - Scrambled: around accidents country honey the shoot.

  - Shadowing – echo speech you hear (Marslen-Wilson, 1973)
    - Repeat words 200 ms behind normal speech
    - Errors respect semantic and syntactic structure!
    - "..heard at" → "…heard that…"
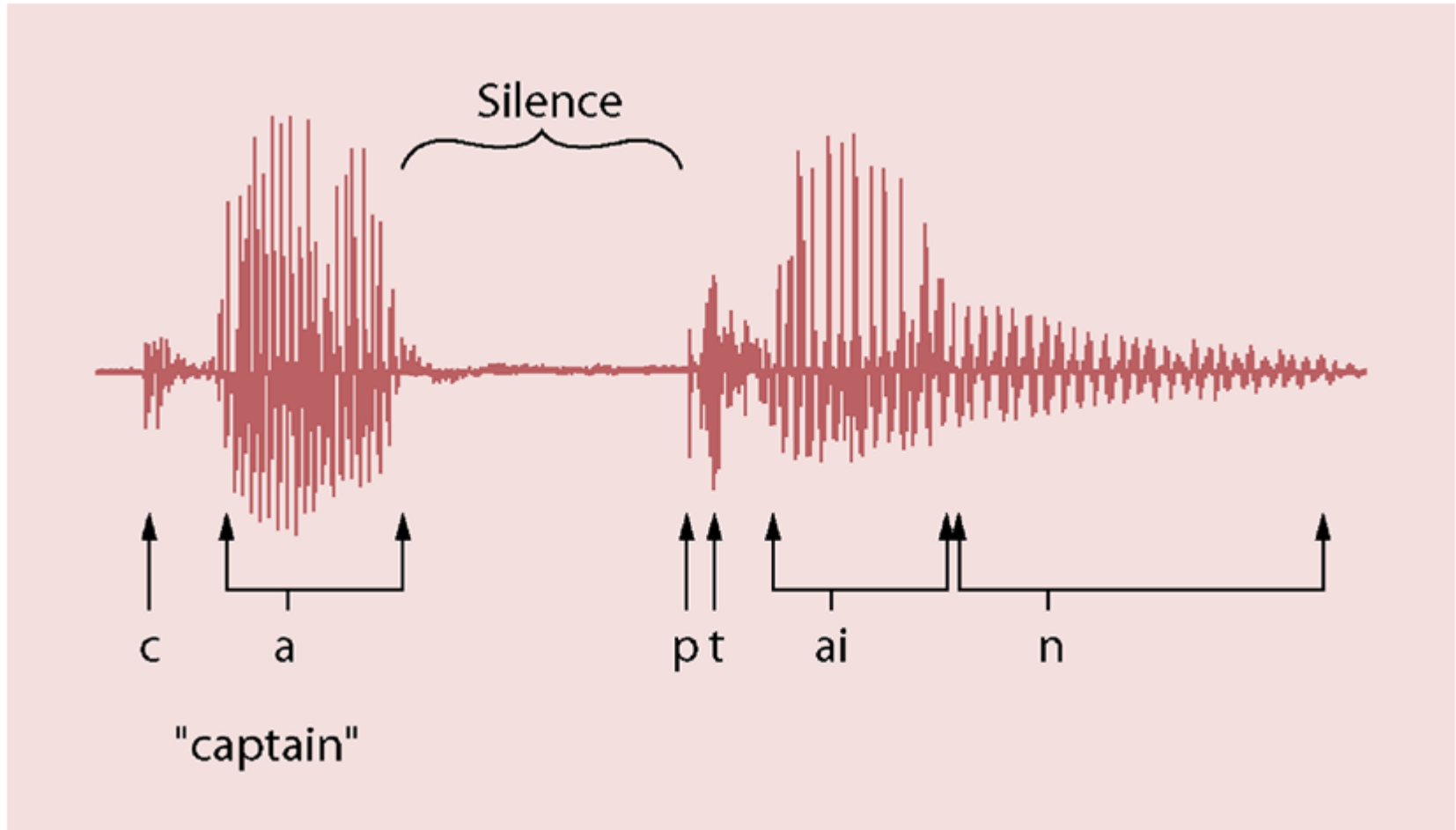
  - receive input not just bottom-up!

Plan for today

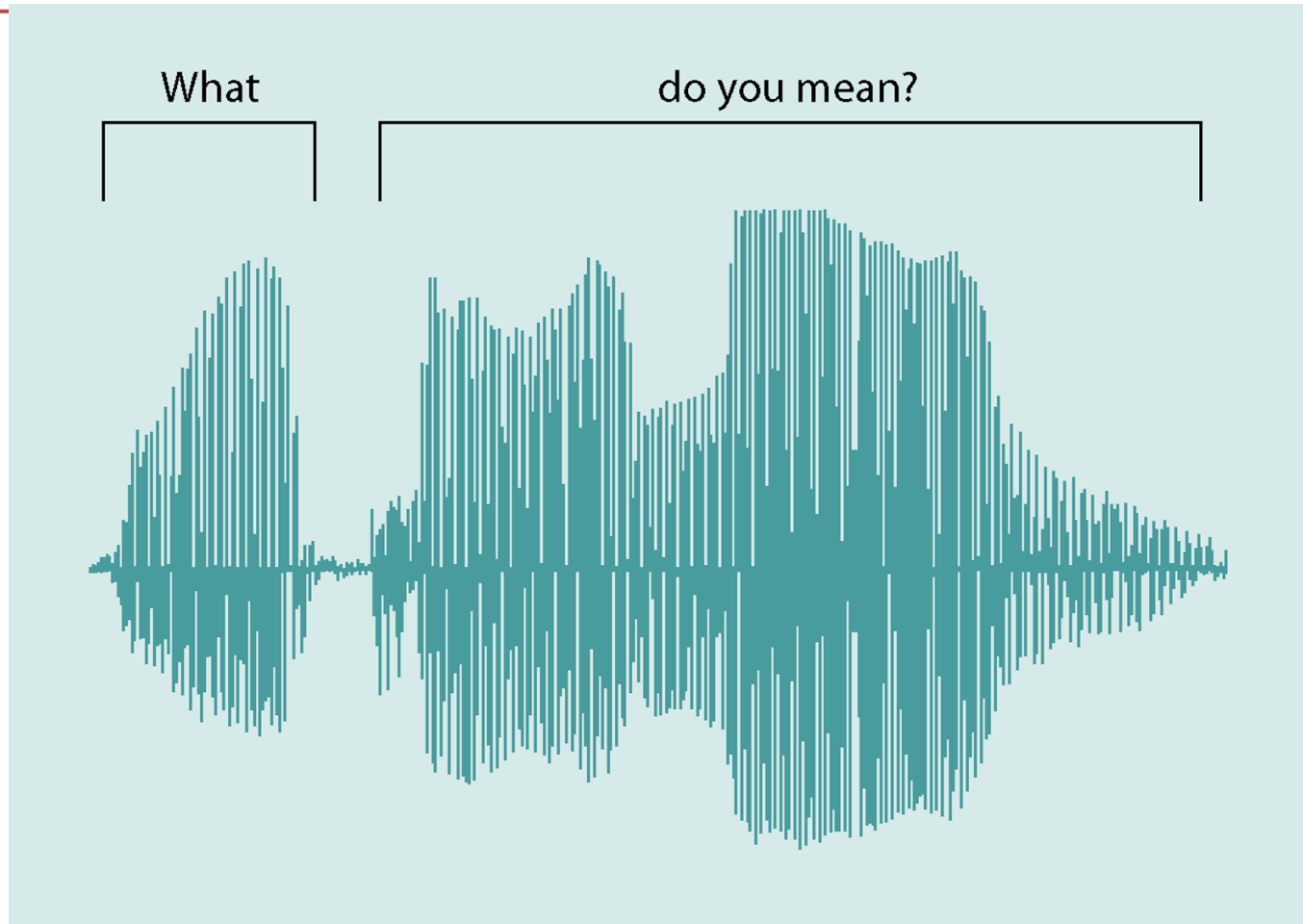- What is speech

- How do we understand it

Perception of continuous speech

- We have been talking about single words

- What about continuous speech?

Perception of continuous speech

Perception of continuous speech

Perception of continuous speech

- How do we segment speech?
  - Use probabilities
  - Use stress rules (greenhouse versus green house)
  - Use context

Summary

- **Properties of speech**
  - Phonemes, articulatory features
- **Problems with understanding speech sounds**
  - Coarticulation
- **Solutions**
  - Categorical perception
  - Context