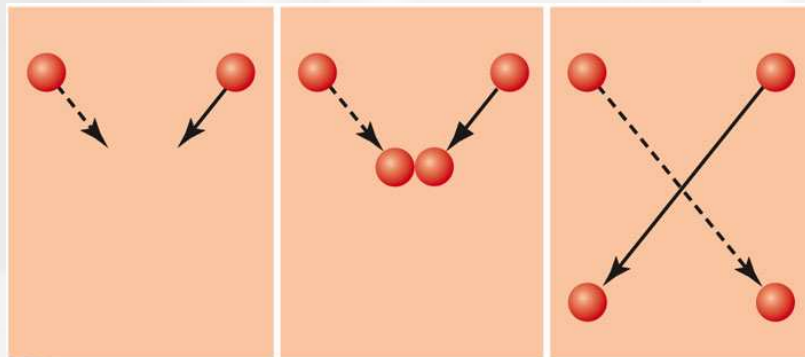


Interactions between sight and sound

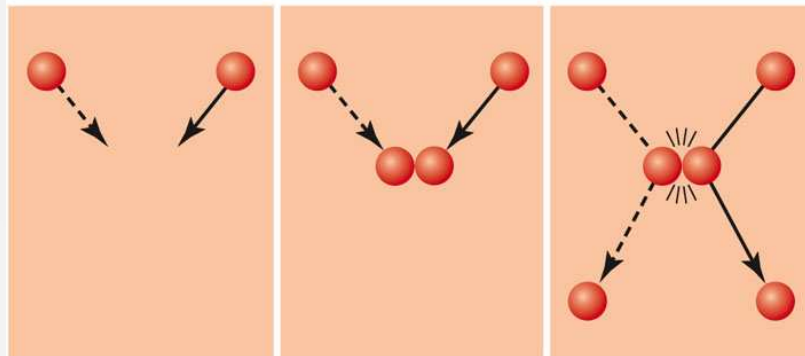
Experiment by Sekuler et al.

Balls moving without sound appeared to move past each other

Balls with an added “click” appeared to collide



(a) Objects appear to pass by each other



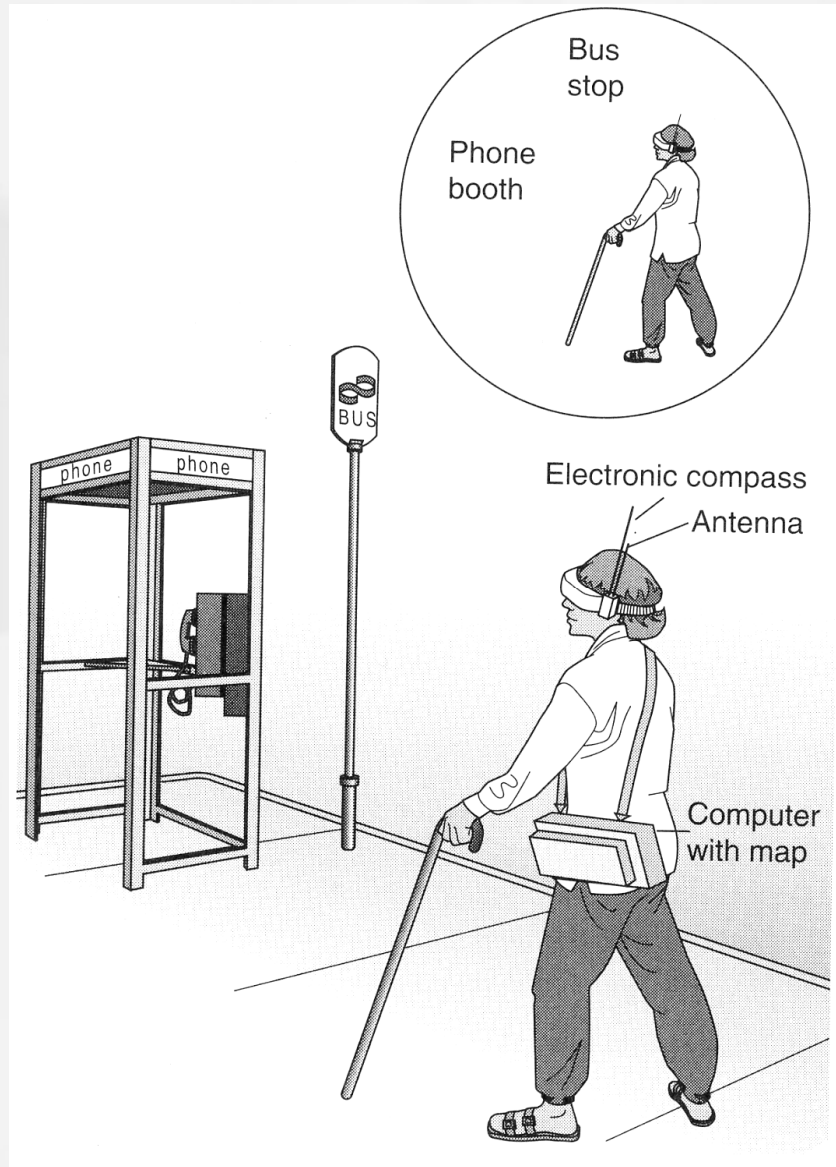
(b) Objects appear to collide

Sound-induced Illusory Flashing

Auditory clicks can influence perceived number of visual flashes.

<http://shamslab.psych.ucla.edu/demos/>

Using auditory stimuli to replace sight



Chapter 13: Speech Perception



Overview of Questions

- Can computers perceive speech as well as humans?
- Why does an unfamiliar foreign language often sound like a continuous stream of sound, with no breaks between words?
- Does each word that we hear have a unique pattern of air pressure changes associated with it?
- Are there specific areas in the brain that are responsible for perceiving speech?

Can computers perceive speech as well as humans?

What I talk now it will start adding text into my Microsoft PowerPoint document .

Sunday's that's pretty well in a quiet room with simple words in my own voice speaking pretty slowly.

The button often a silly mistakes to specify start speaking more more quickly to start using very complicated words like phenomenal the psychophysics on auditory cortex what a fine when stream January the cuts psychophysics in a while I probably say psychophysics lot more than most people doing an anxiety why he conceded that his media scene gloves (a) well into the wind is better than my dad in the ring was later stages 10 point 463 impressive but now it's doing it since hitting at all since when is the loss was doing some tenants find In a second save the quick brown fox typically do not some tenants find In the second save the queen brown fox typically do not something that's fine in the second save the queen brown fox site next line In the second save the queen brown fox and it's like not having a meeting of the amount of space at the time the delay and of all citizens of the game is morning I had a lot of my life and he started the year ended sounds and started making of words on the Internet has been at least in house on a one-on and so the

The Speech Stimulus

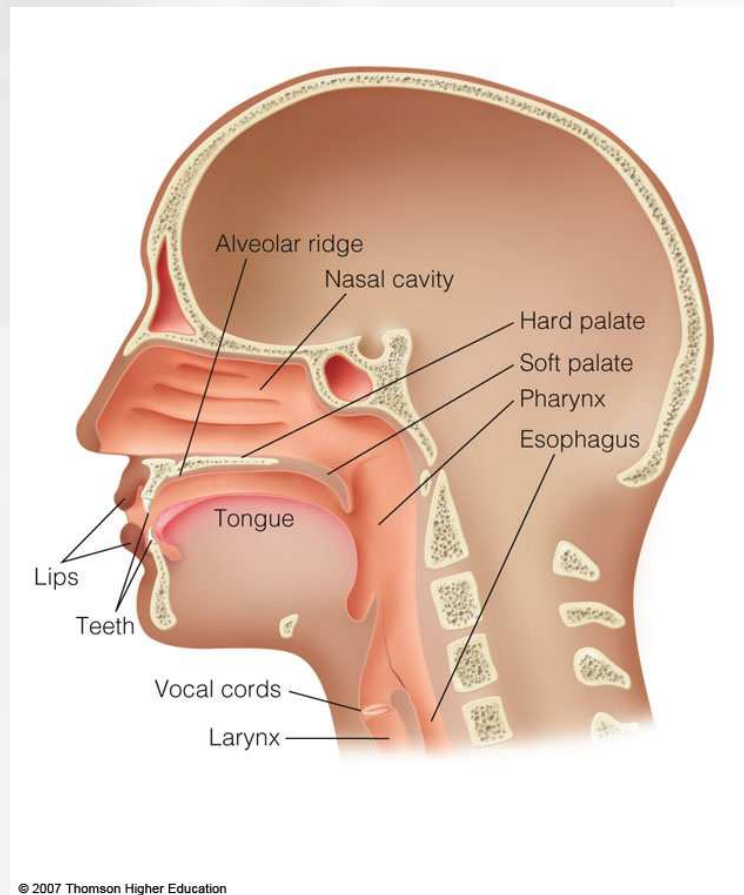
- Phoneme – the smallest unit of speech that changes meaning in a word
 - In English there are 47 phonemes:
 - 13 major vowel sounds
 - 24 major consonant sounds
 - Number of phonemes in other languages varied—11 in Hawaiian and 60 in some African dialects

Table 13.1 ■ Major consonants and vowels of English and their phonetic symbols

Consonants		Vowels
p pull	s sip	i heed
b bull	z zip	ɪ hid
m man	r rip	e bait
w will	ʃ should	ɛ head
f fill	ʒ pleasure	æ had
v vet	č chop	ʊ who'd
θ thigh	ʝ gyp	ʊ put
ð thy	ɣ yip	ʌ but
t tie	k kale	o boat
d die	g gale	ɔ bought
n near	h hail	a hot
l lear	ŋ sing	ə sofa
		ɹ many

The Acoustic Signal

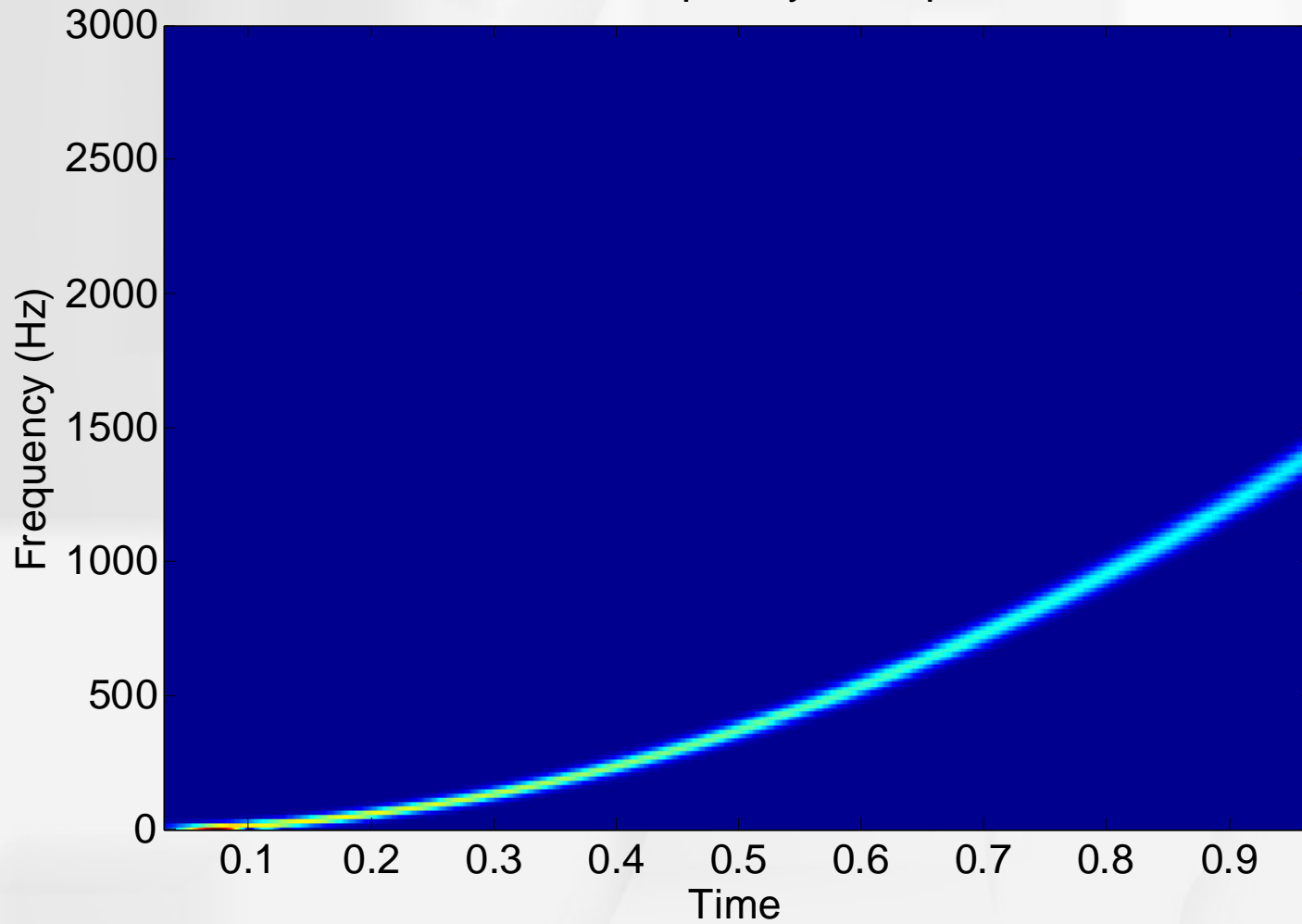
- Produced by air that is pushed up from the lungs through the vocal cords and into the vocal tract
- Vowels are produced by vibration of the vocal cords and changes in the shape of the vocal tract



The Sound Spectrogram



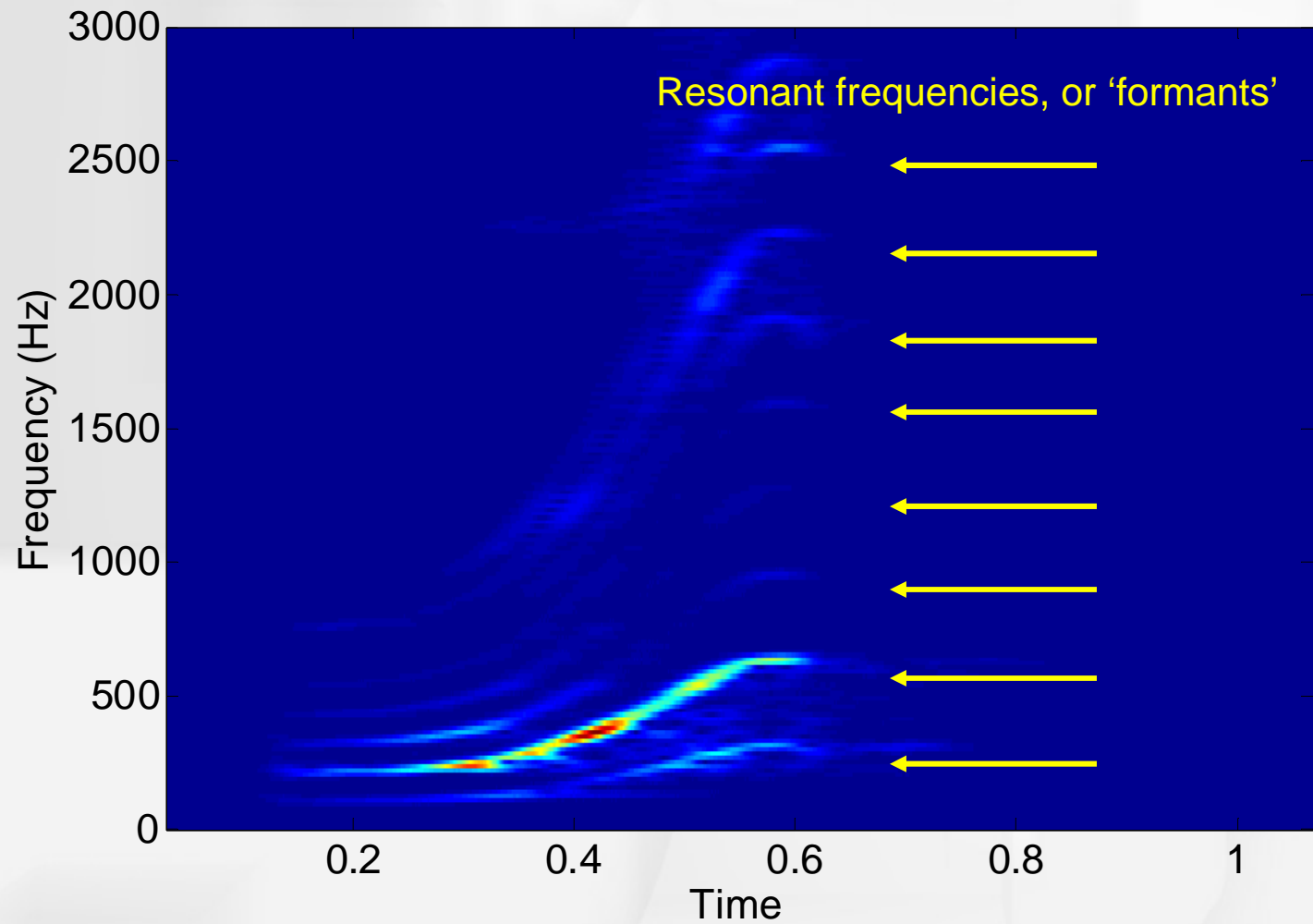
'frequency sweep'



The Sound Spectrogram

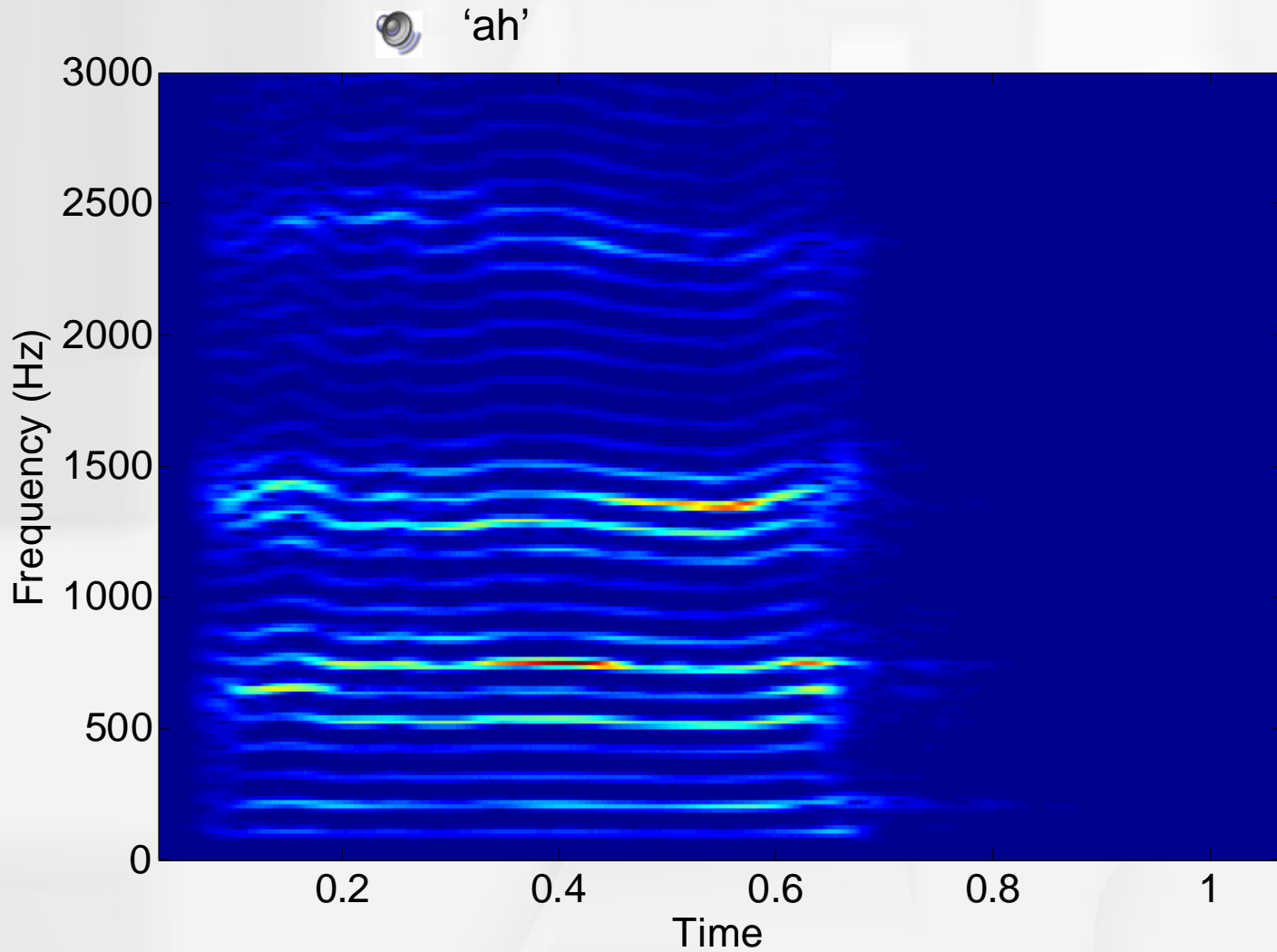


my (lame) attempt at a 'frequency sweep'



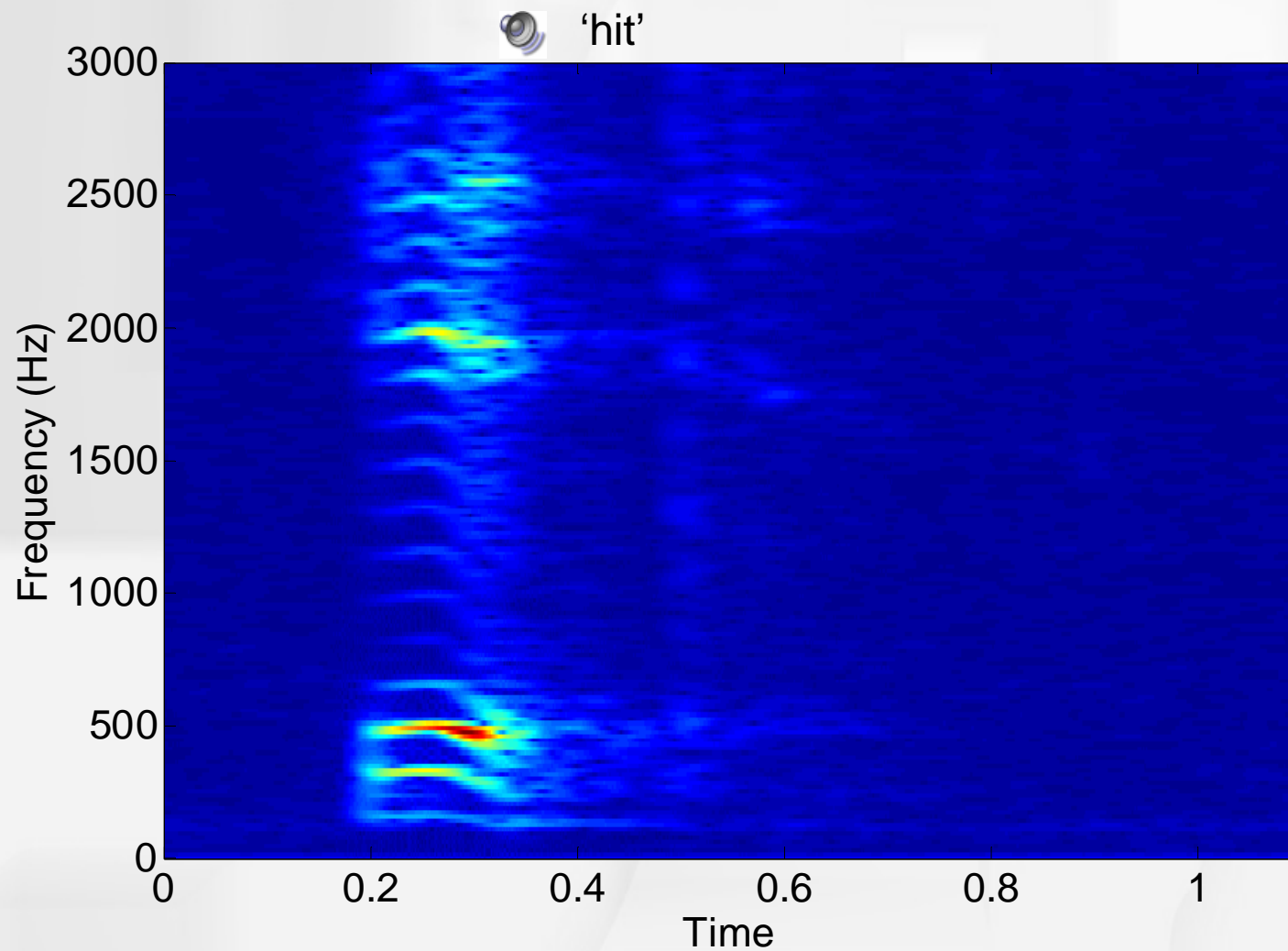
Vowel sounds are caused by a resonant frequency of the vocal cords and produce peaks in pressure at a number of frequencies called *formants*

The first formant has the lowest frequency, the second has the next highest, etc.



The Acoustic Signal

- Consonants are produced by a constriction of the vocal tract

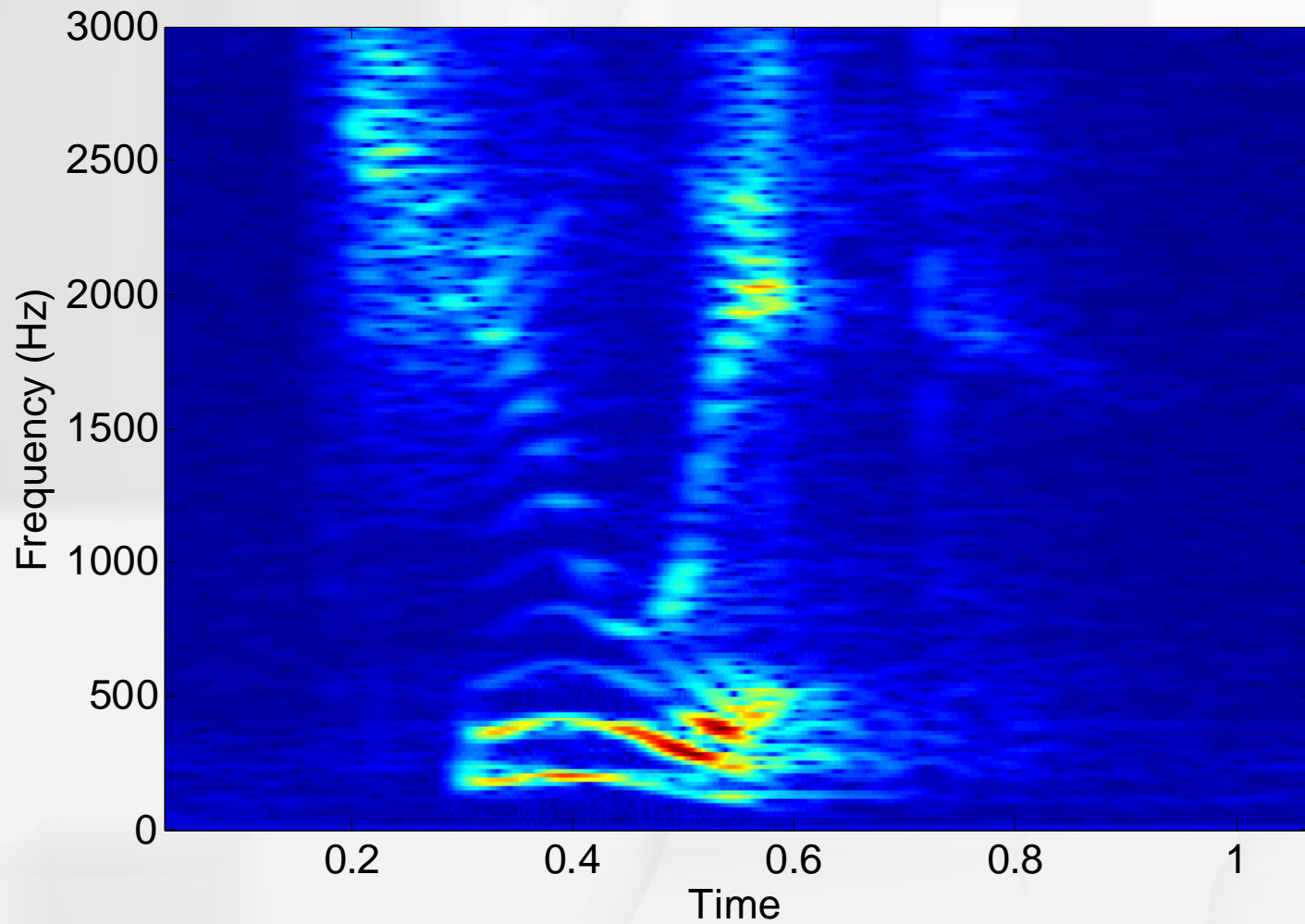


The segmentation problem:

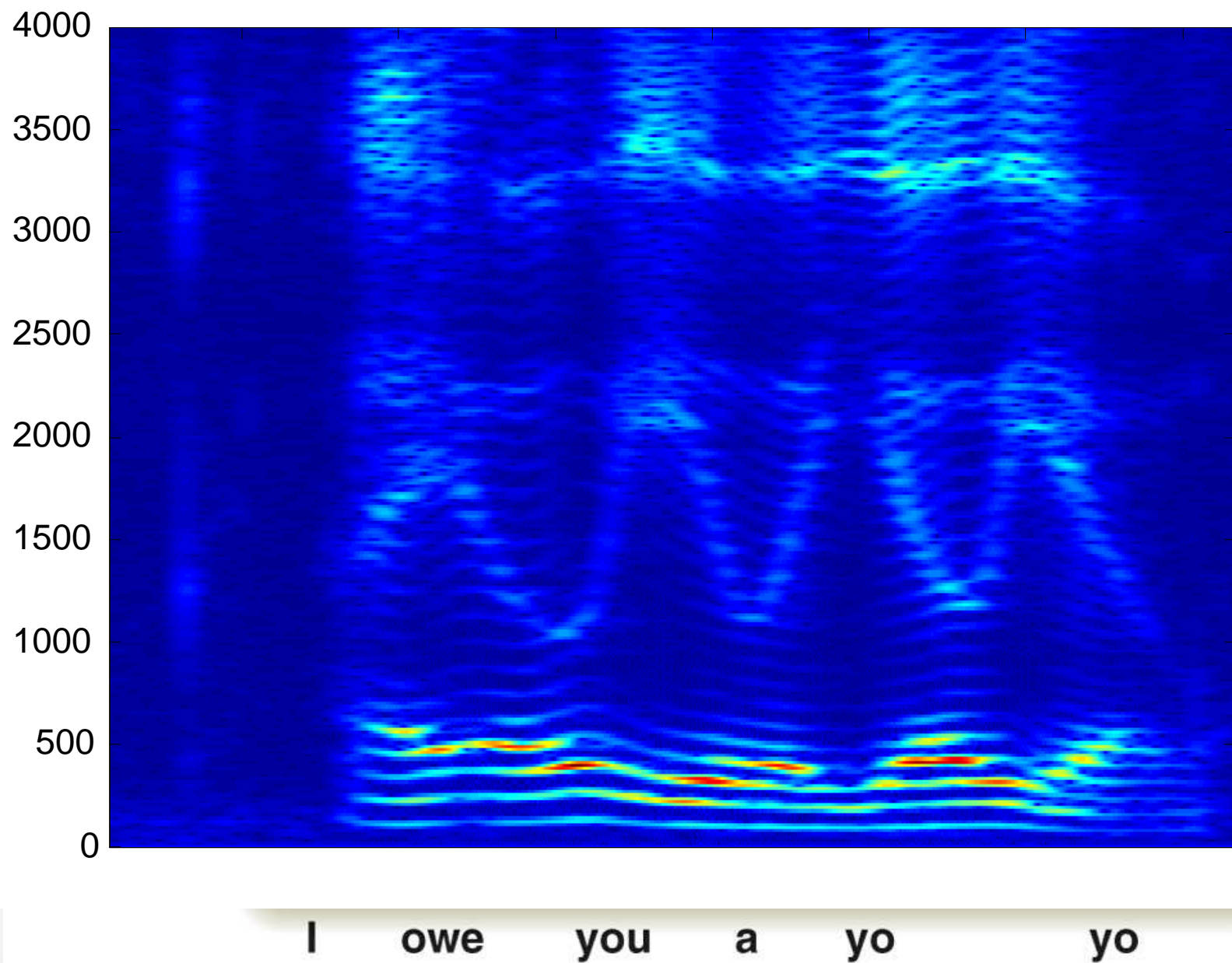
There are no physical breaks in the continuous acoustic signal.



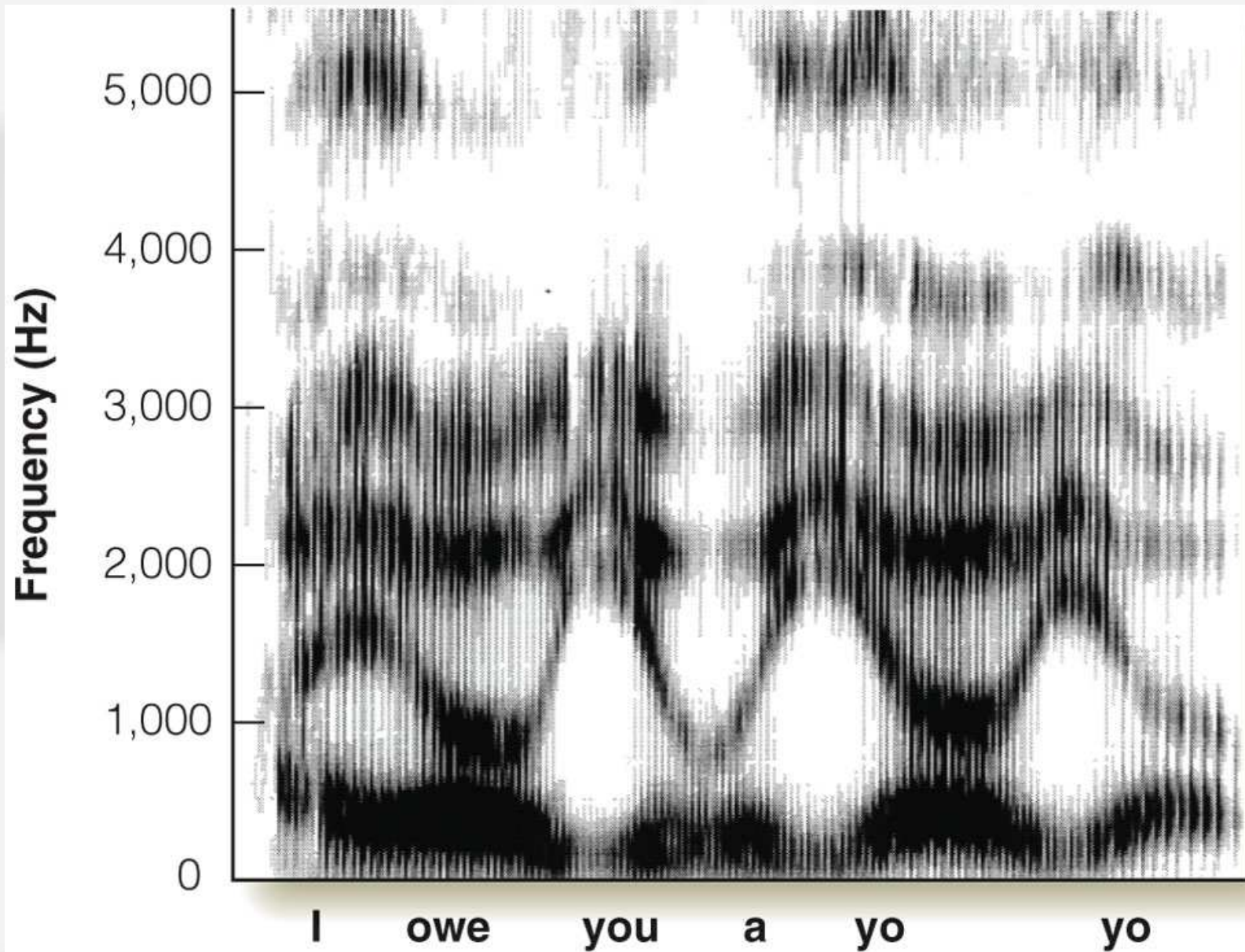
'chew it'



The segmentation problem



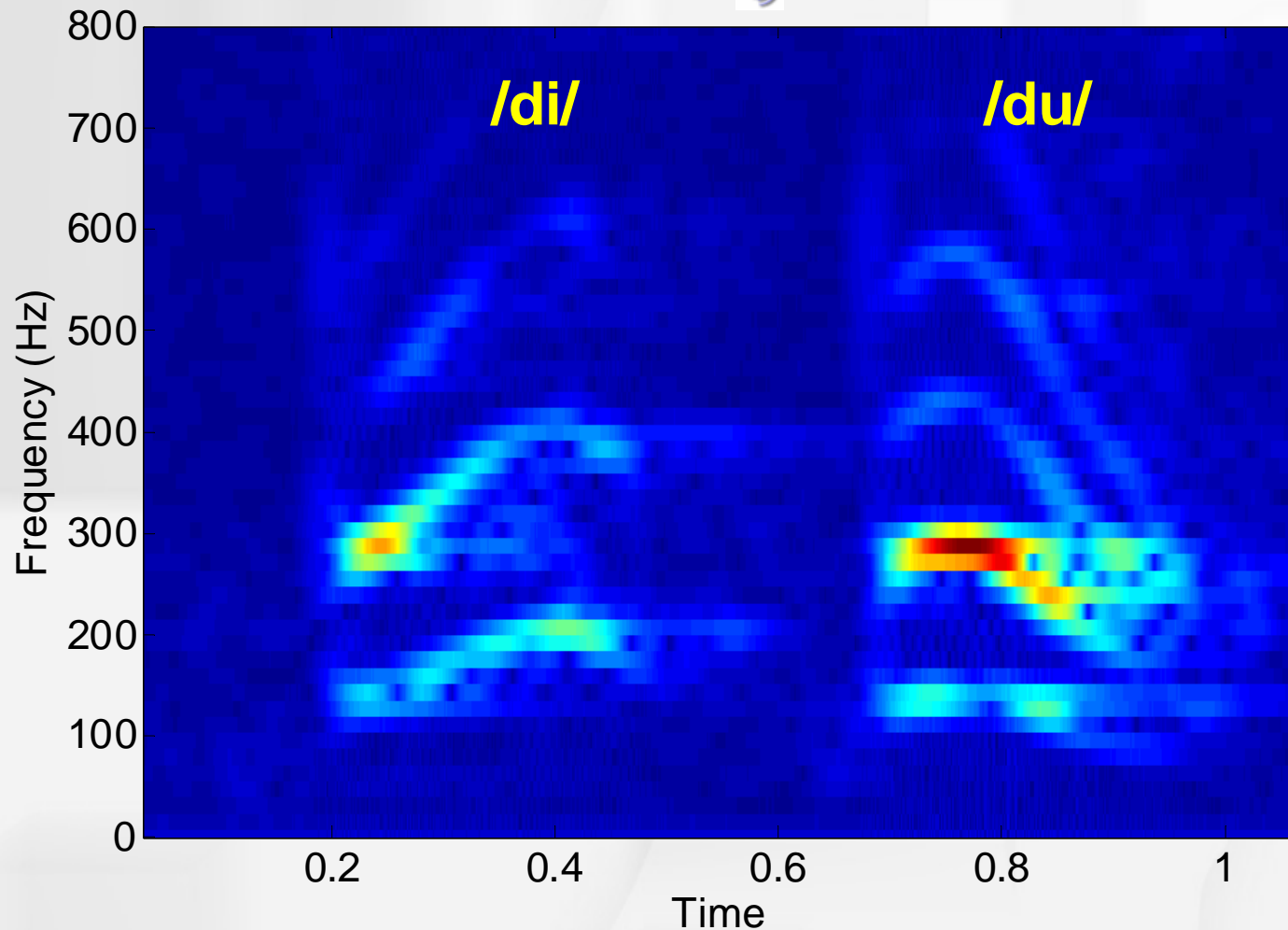
The segmentation problem



The variability problem

There is no simple correspondence between the acoustic signal and individual phonemes

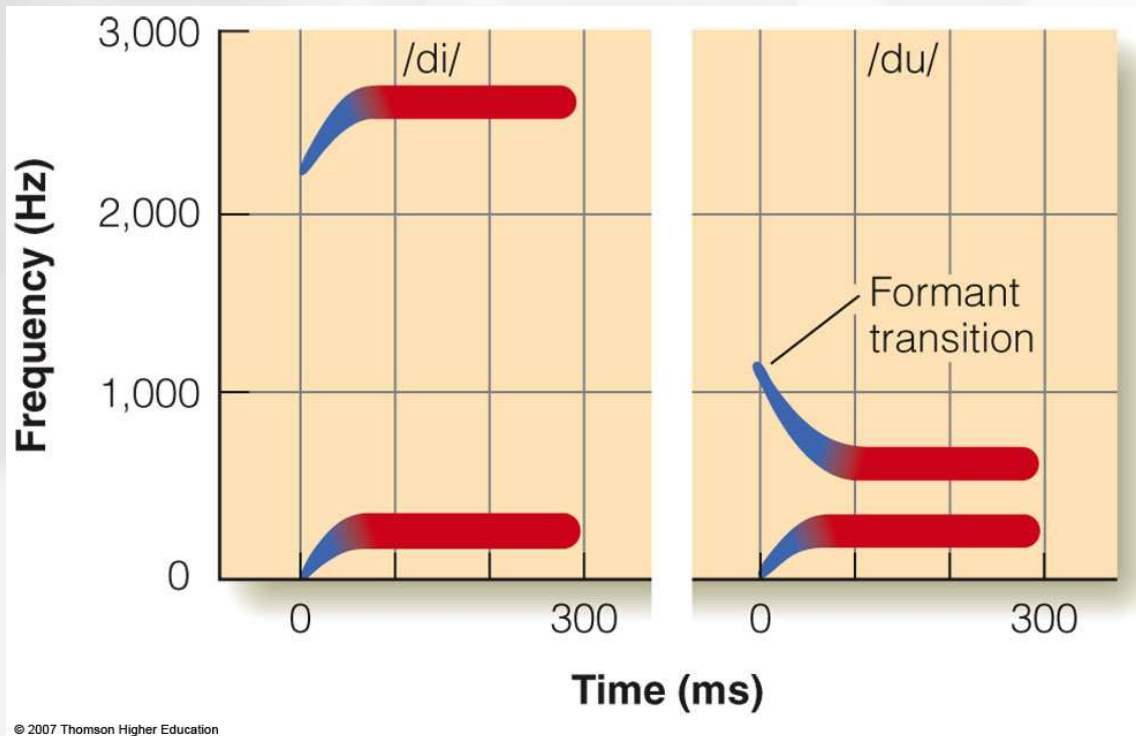
Coarticulation - overlap between articulation of neighboring phonemes:
'd' looks different depending on the vowel sound that follows it.



The variability problem

There is no simple correspondence between the acoustic signal and individual phonemes

- 1) Coarticulation - overlap between articulation of neighboring phonemes

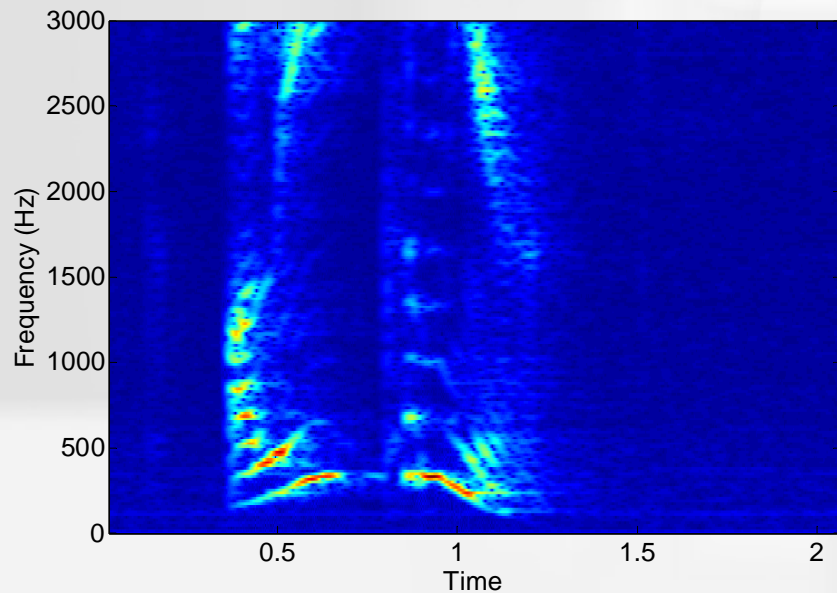


The variability problem

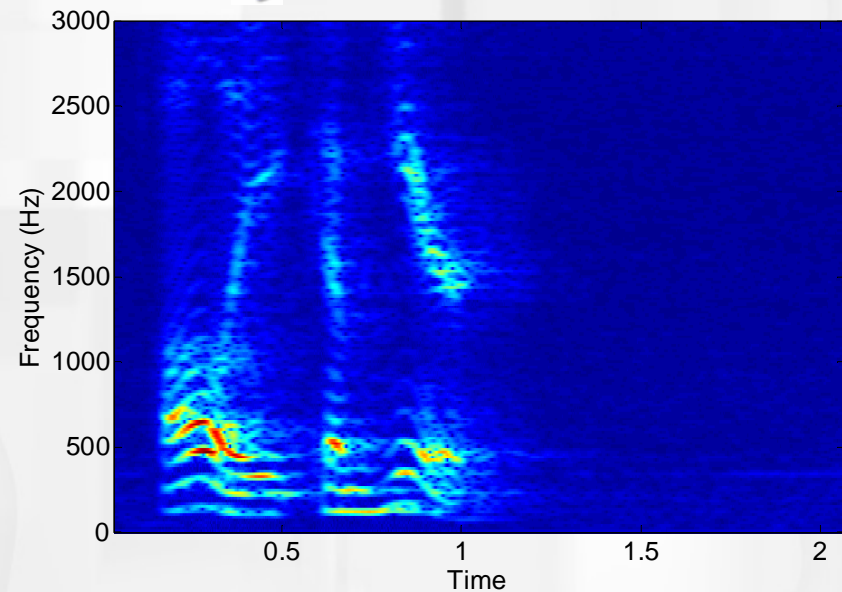
2) Variability across different speakers:

Speakers differ in pitch, accent, speed in speaking, and pronunciation

 'Ollie come here' (Ione)

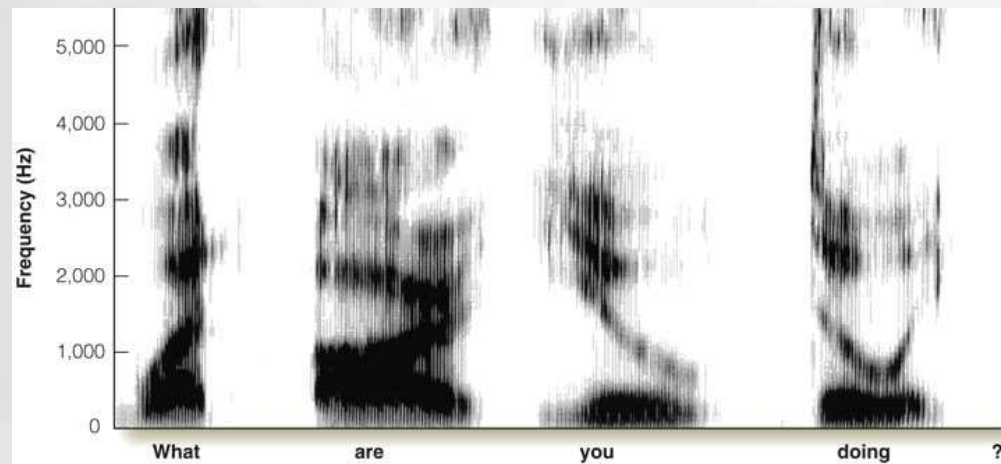


 'Ollie come here' (Geoff)

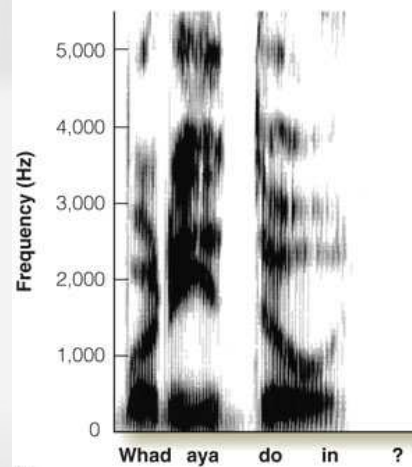


The variability problem

3) Different pronunciations have the same meaning, but very different spectrograms



(a)

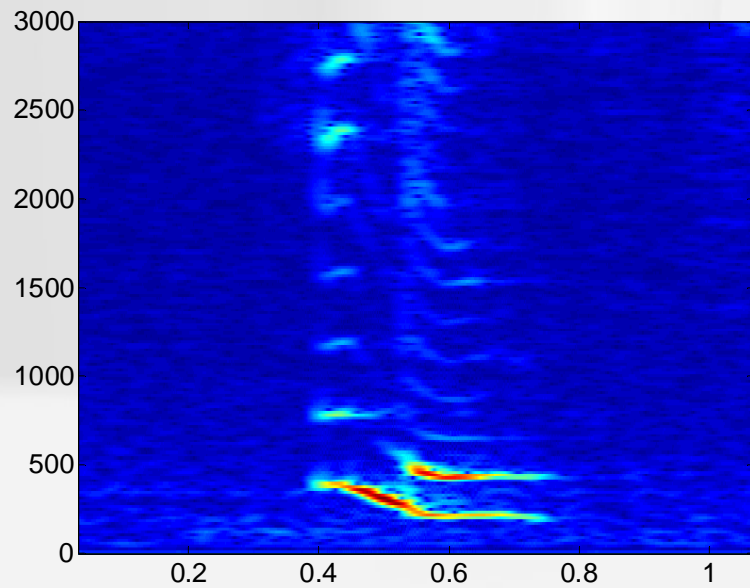


(b)

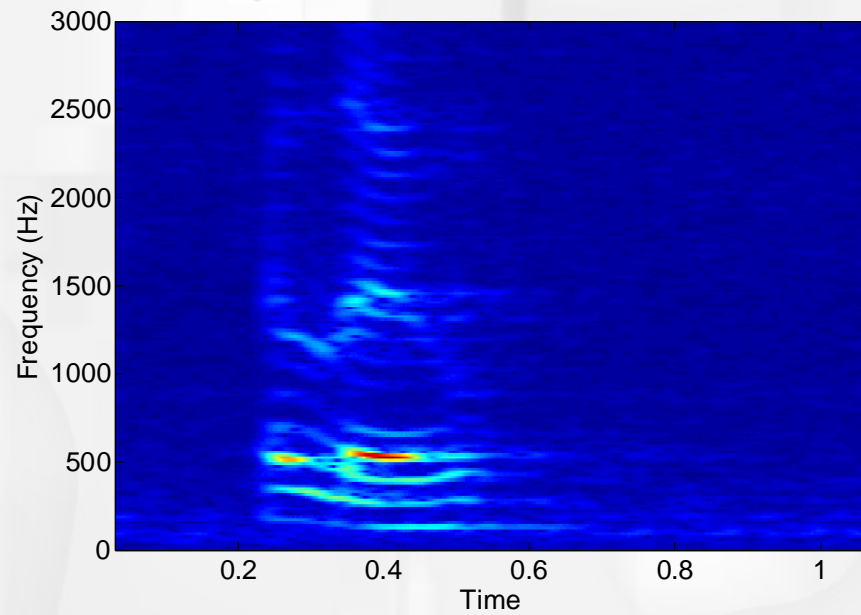
But there are some 'invariances' in speech perception.



'hello' (lone)



'hello' (Geoff)



These spectrograms look similar.

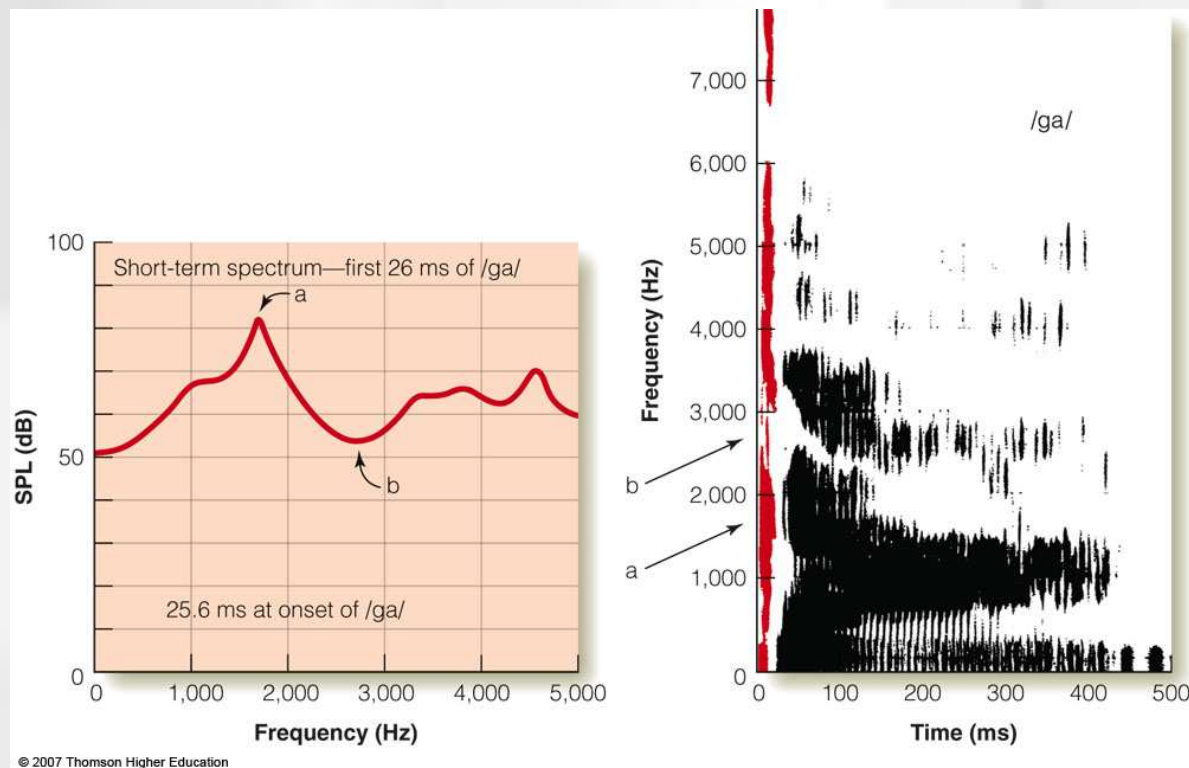
Invariant acoustic cues:

Some features of phonemes remain constant

Short-term spectrograms are used to investigate invariant acoustic cues.

Sequence of short-term spectra can be combined to create a running spectral display.

From these displays, there have been some invariant cues discovered



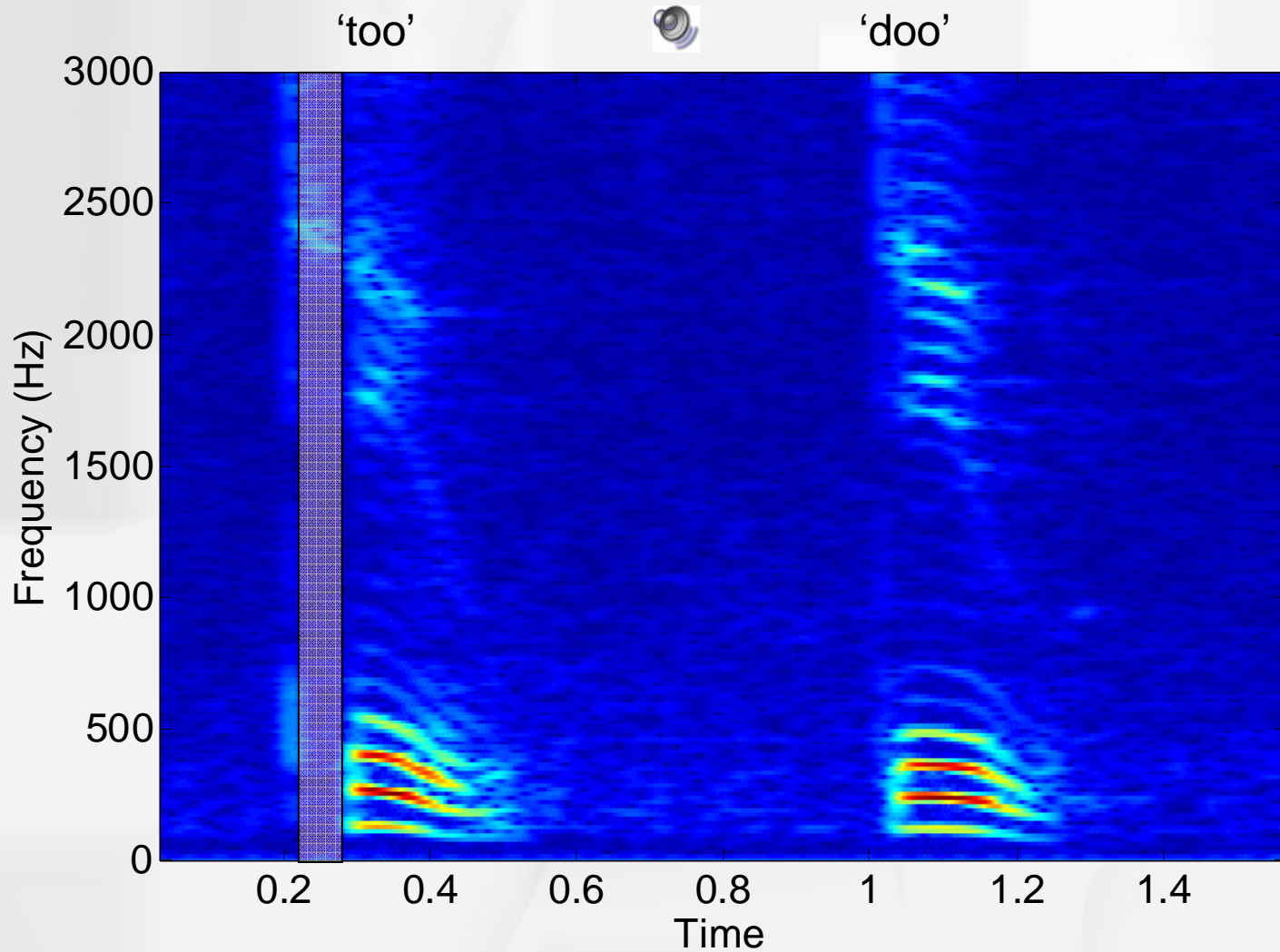
Categorical Perception

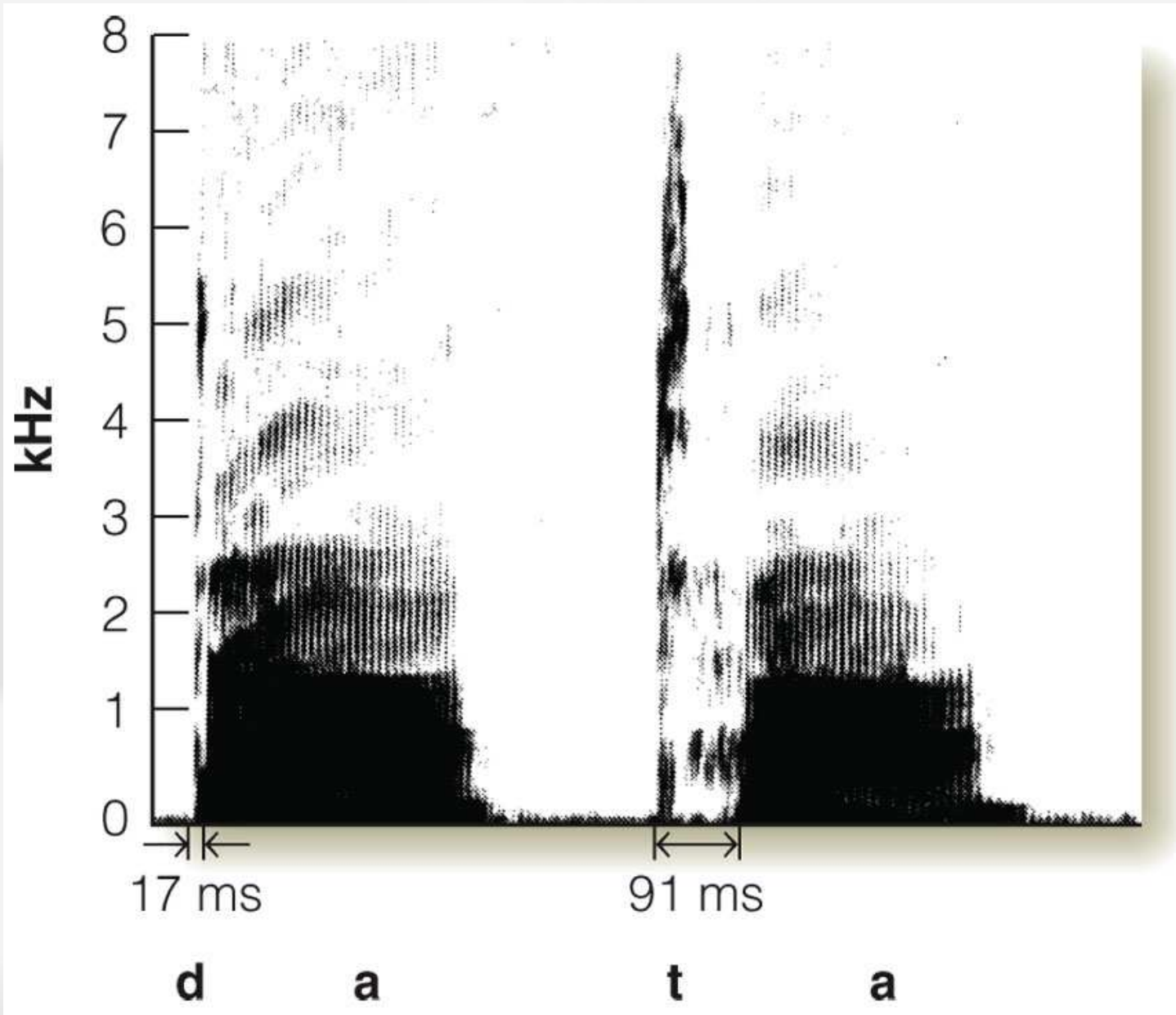
- This occurs when a wide range of acoustic cues results in the perception of a limited number of sound categories
- An example of this comes from experiments on voice onset time (VOT) - time delay between when a sound starts and when voicing begins
 - Stimuli are **da** (VOT of 17ms) and **ta** (VOT of 91ms)

Voice onset time (VOT)

Delay between when the sound begins and the onset of vocal cords.

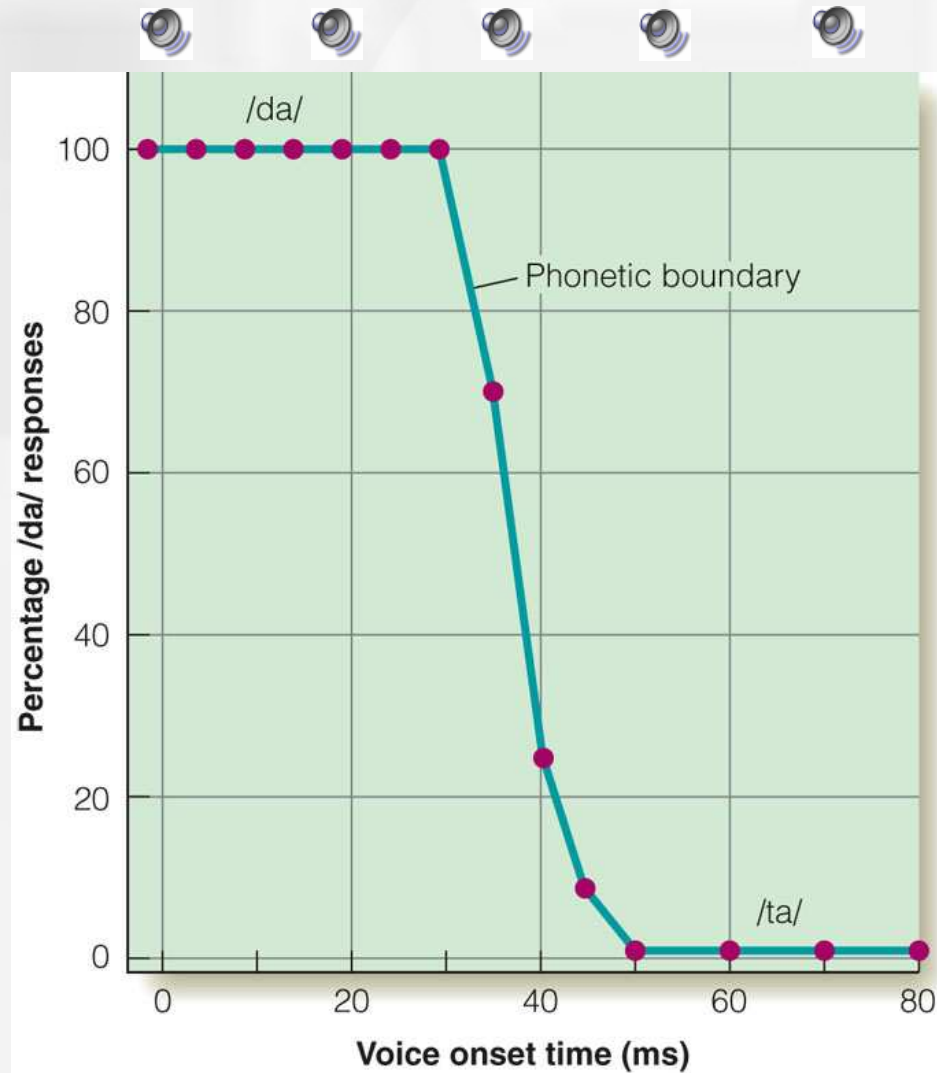
Distinguishes between 'ta' vs. 'da', and 'pa' vs. 'pa'.

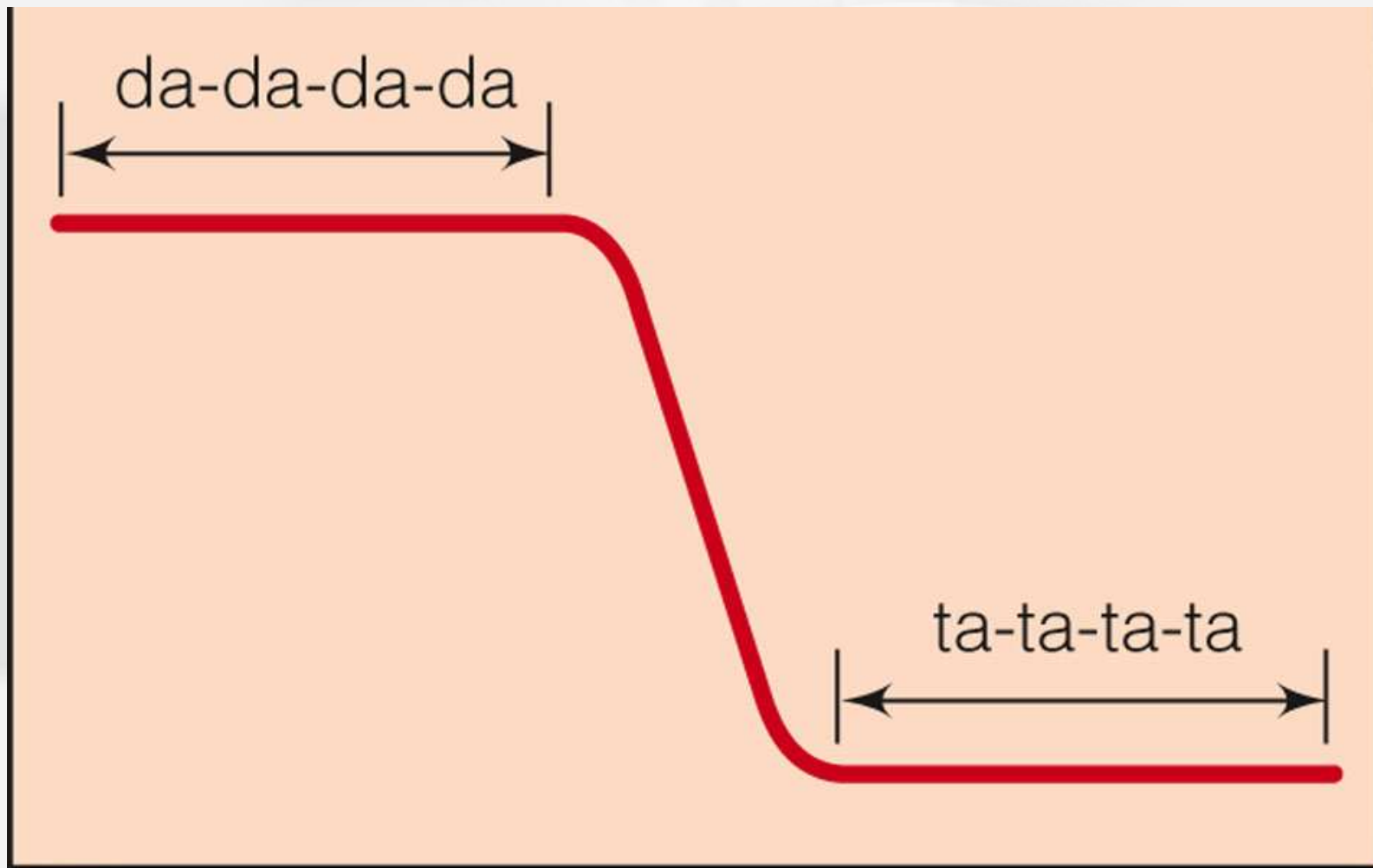




'Categorical perception'

Despite the continuous variation of VOT, we only hear one phoneme or the other.

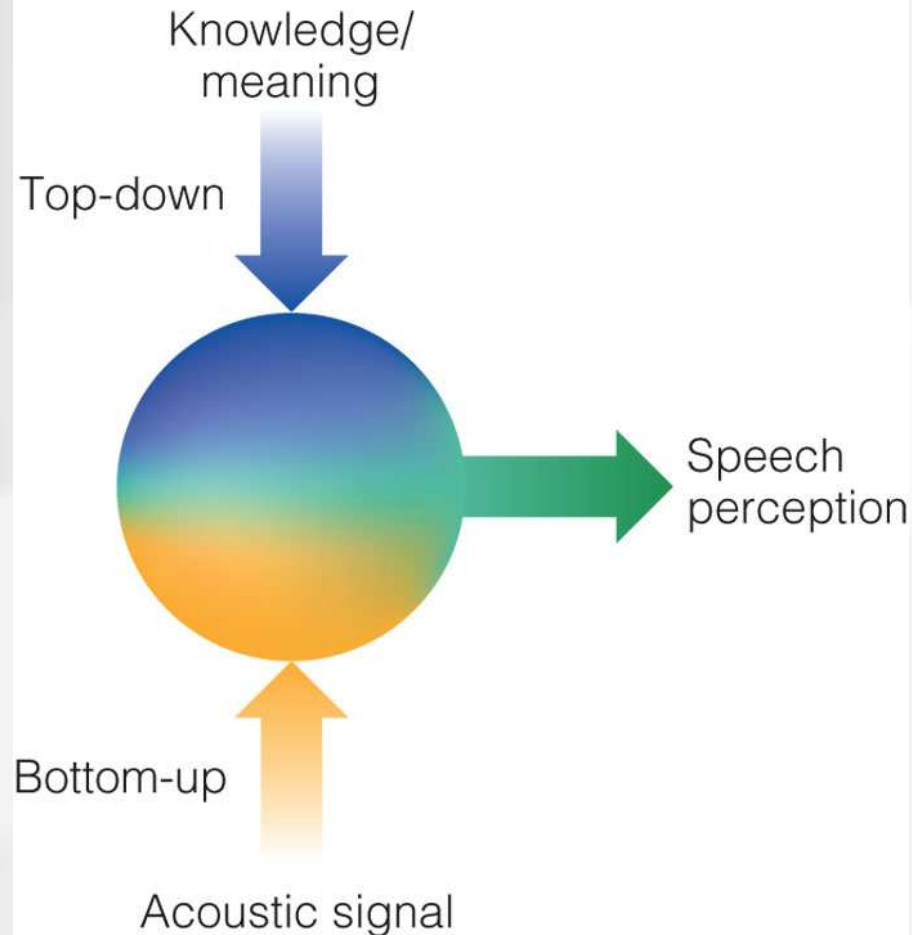




Voice onset time

Cognitive Dimensions of Speech Perception

- Top-down processing, including knowledge a listener has about a language, affects perception of the incoming speech stimulus
- Segmentation is affected by context and meaning
 - I scream you scream we all scream for ice cream



I screamed you
screen we all screen
for high screen