

Intro to Speech Processing

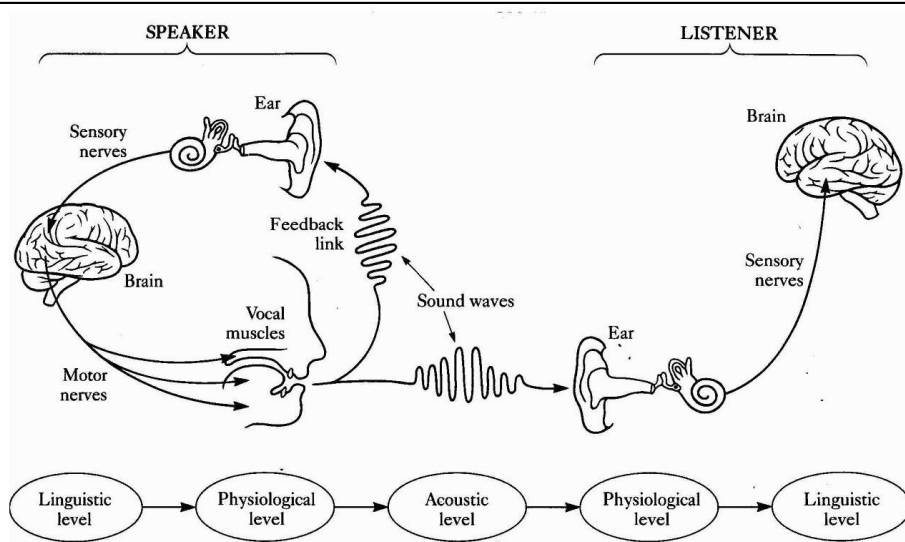
Arian Shamei

- Goals:
 - Understand how speech is produced
 - Understand basic components of speech
 - Understand source/filter theory

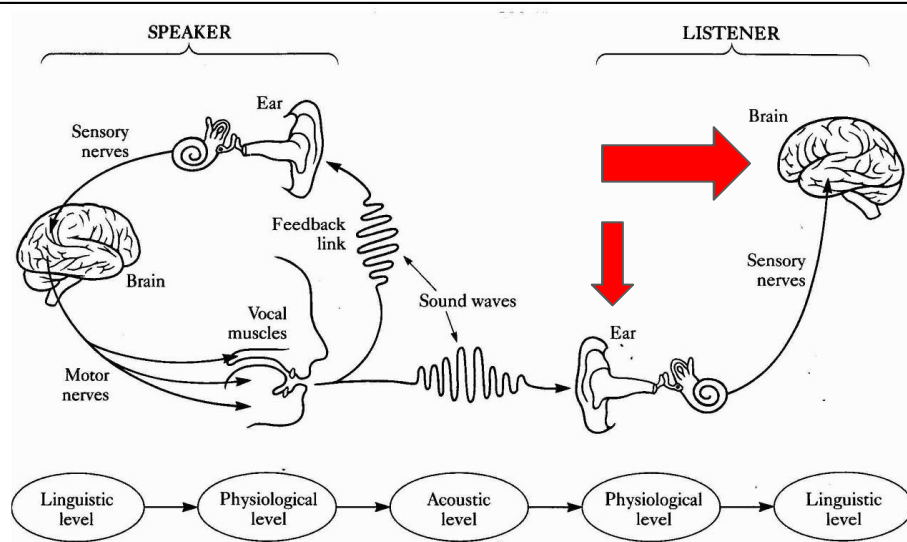
Background

Traditionally, phoneticians view speech as an acoustic signal consisting of a complex wave composed of many individual sine waves; the speaker's movements drive air molecules via compression and rarefaction, this in turn propagates energy to other air molecules which eventually influence movement along a sensitive membrane inside the head of surrounding humans. This concept is known as the speech chain..... Blah blah blah listener is important

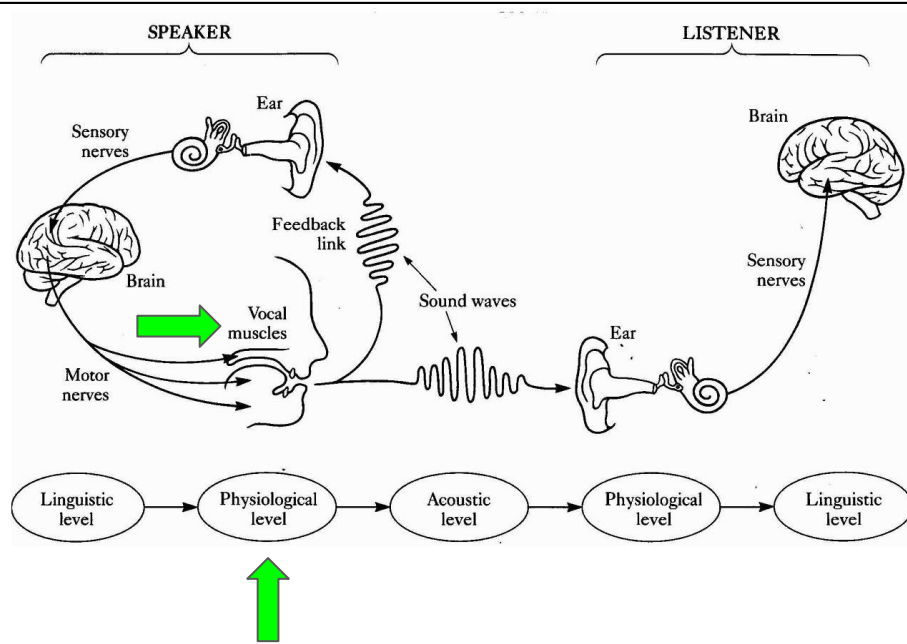
The Speech Chain



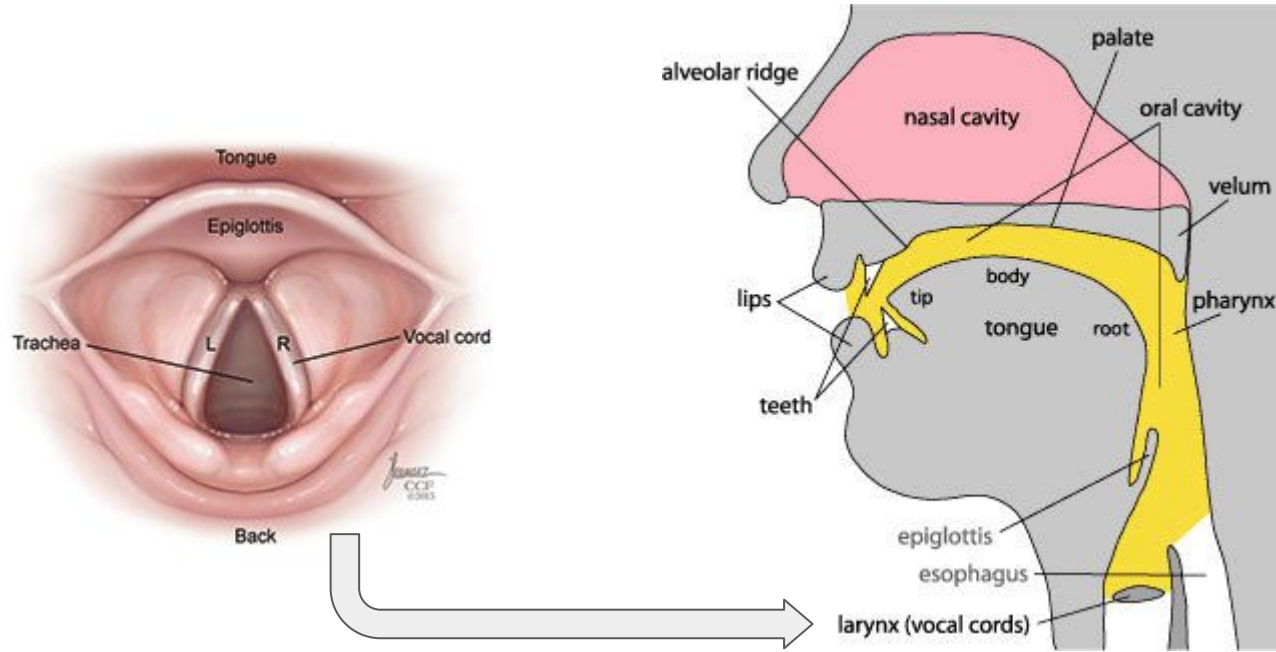
The Speech Chain



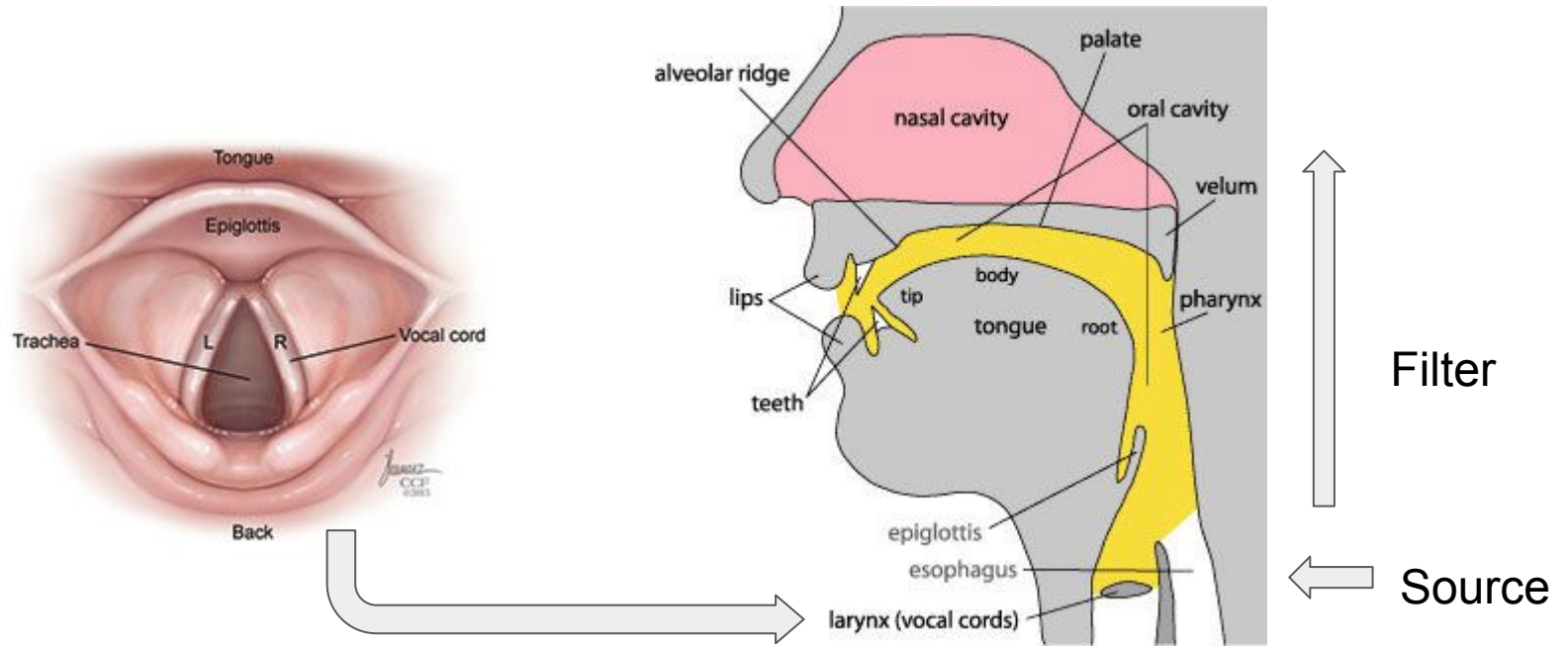
The Speech Chain



The vocal tract



The vocal tract



The Source

Size of the larynx
determines the cycling rate.

Cycling rate determines the
pitch (F_0).



What information does a speaker convey with F0?

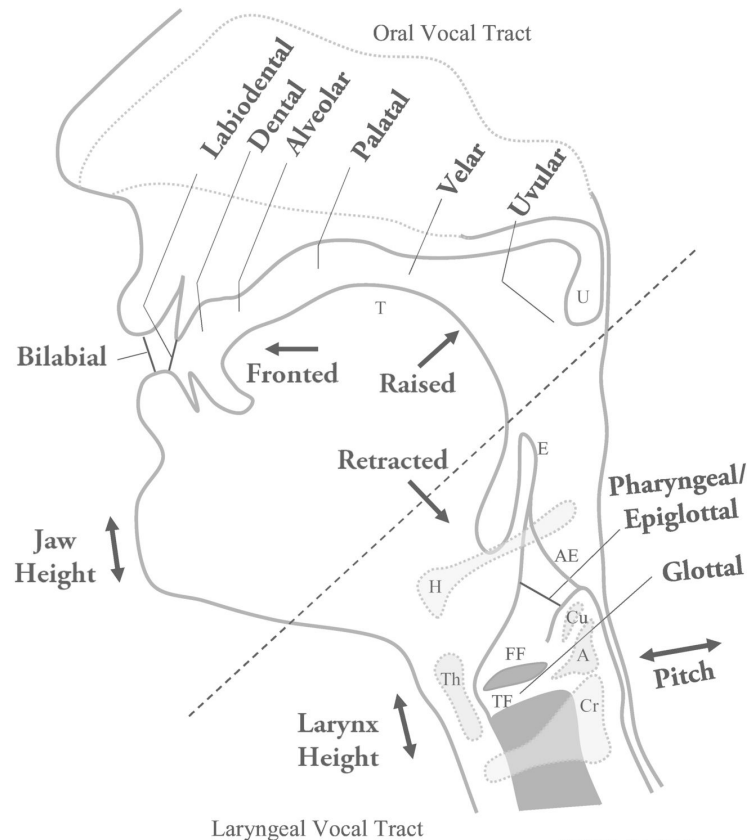
The Kazoo Test!

The Filter

Resonances of the source are modified by changing the shape of the filter.

Specific shapes and constrictions produce distinct speech sounds.

Think of this like a guitar string.
The hollow body produces a sound,
manipulation of the string alters it.



Speech Sounds

- Vowels
- Consonants

Speech Sounds

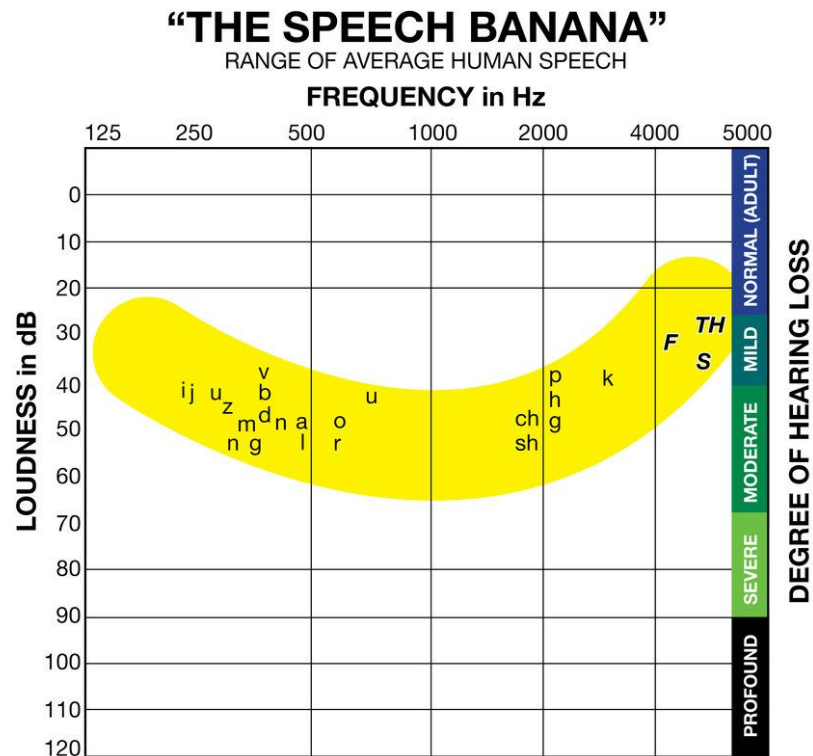
- Vowels

- sonorous, syllabic sounds made with the vocal tract “unobstructed”
- lower frequencies
 - /a/, /i/, /o/

- Consonants

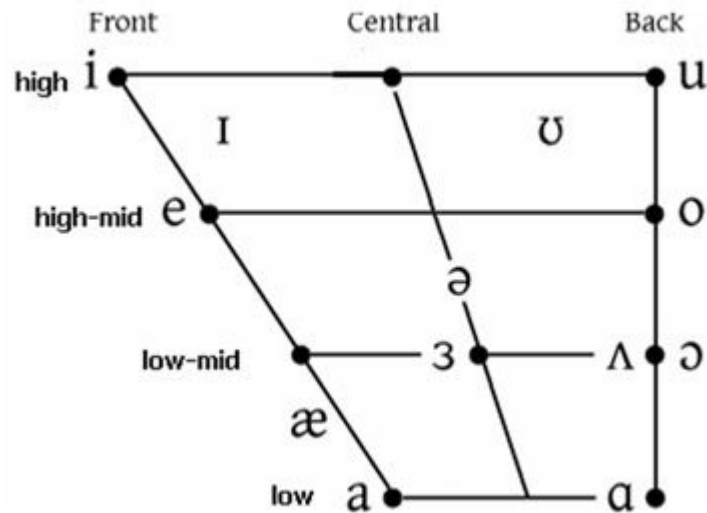
- articulated with complete or partial closure of the vocal tract
- higher frequencies
 - /p/, /t/, /k/

Speech Banana



Vowels

- Height
- Backness
- Rounding



Vowels & Formants

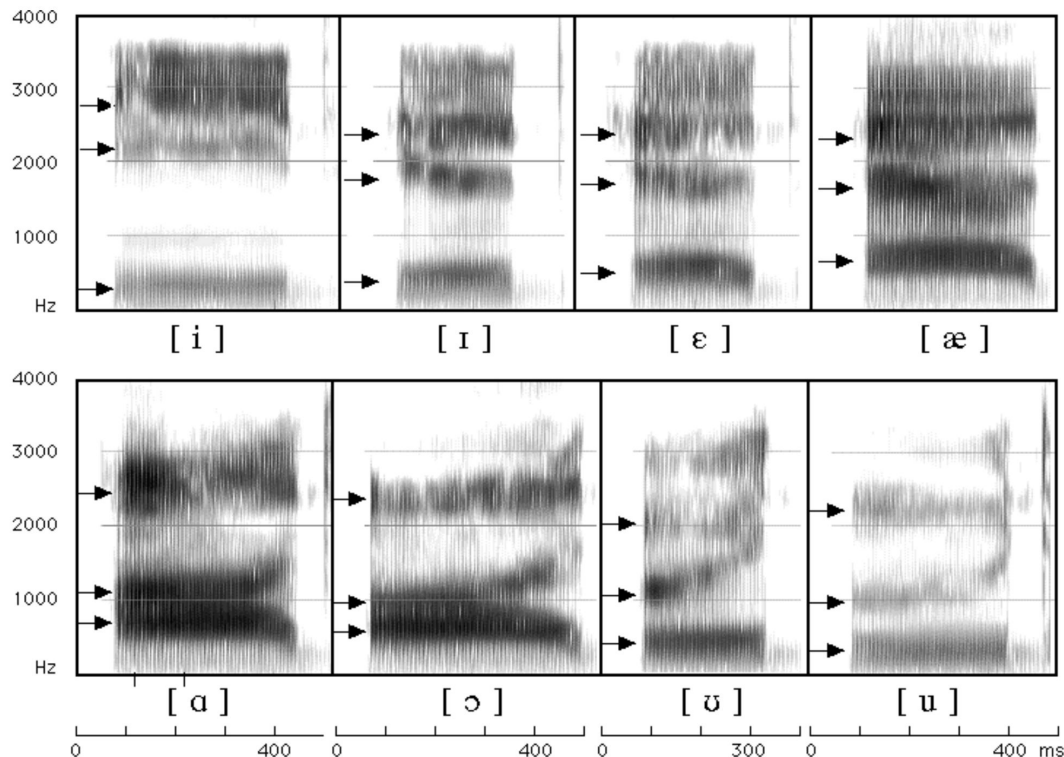
Formants:

Resonant frequencies of
the vocal tract.

F1: Height (inverse)

F2: Backness

F3: Rounding



Consonants

Systematic obstructions and manipulations of the vocal tract.

Described via three parameters:

1. Place of Articulation
2. Manner of Articulation
3. Voicing

/t/ vs /d/

/p/ vs /k/

/t/ vs /s/

Consonants

Place of Articulation													
CONSONANTS (PULMONIC)													
	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal		
Plosive	p b		t d			ʈ ɖ	c ɟ	k ɡ	q ɢ			ʔ	
Nasal	m	ɱ	n			ɳ	ɲ	ŋ	ɴ				
Trill	ʙ		r						ʀ				
Tap or Flap		ɸ	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	ʟ					


Symbols to the right in a cell are voiced, to the left are voiceless. Shaded areas denote articulations judged impossible.

Voicing = Activation of the vocal folds

Place of Articulation = The locus in the 'mouth' where constriction occurs

Manner of Articulation = The degree of pressure / constriction in the 'mouth'

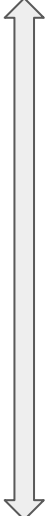
Consonants



 Locus of Constriction

CONSONANTS (PULMONIC) © 2016 IPA

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r					ʀ		
Tap or Flap		ⱱ		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	ʟ			



 Degree of Pressure

Symbols to the right in a cell are voiced, to the left are voiceless. Shaded areas denote articulations judged impossible.

Types of consonants (manner)

Stops / Plosives: /t/

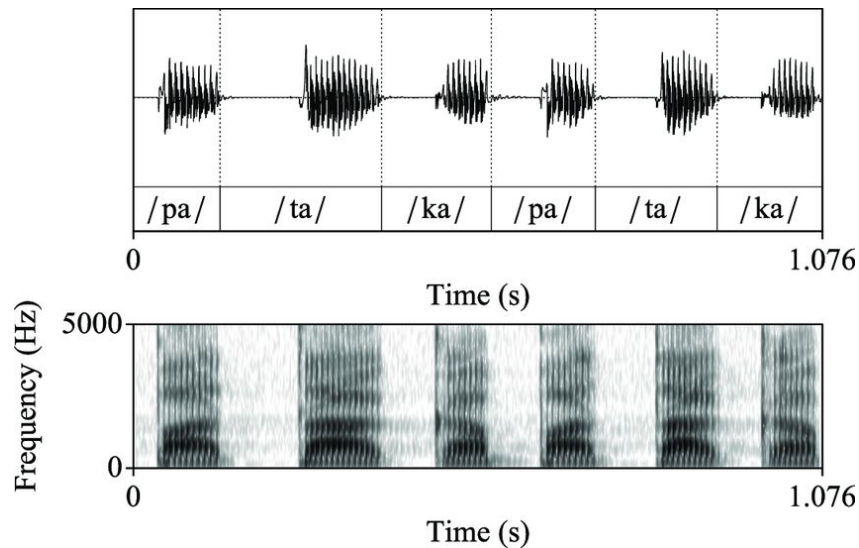
Fricatives: /s/

Nasals: /n/

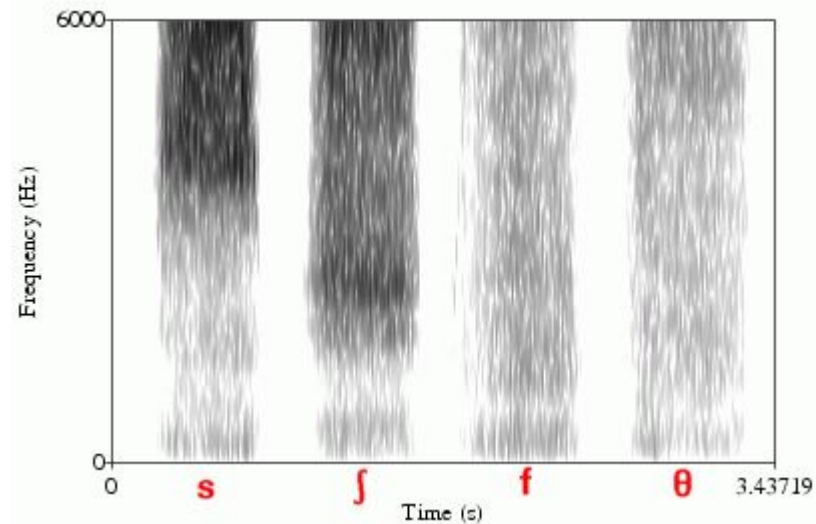
Liquids: /r/

Approximants / demi-vowels: /w/

Spectral properties (consonants)

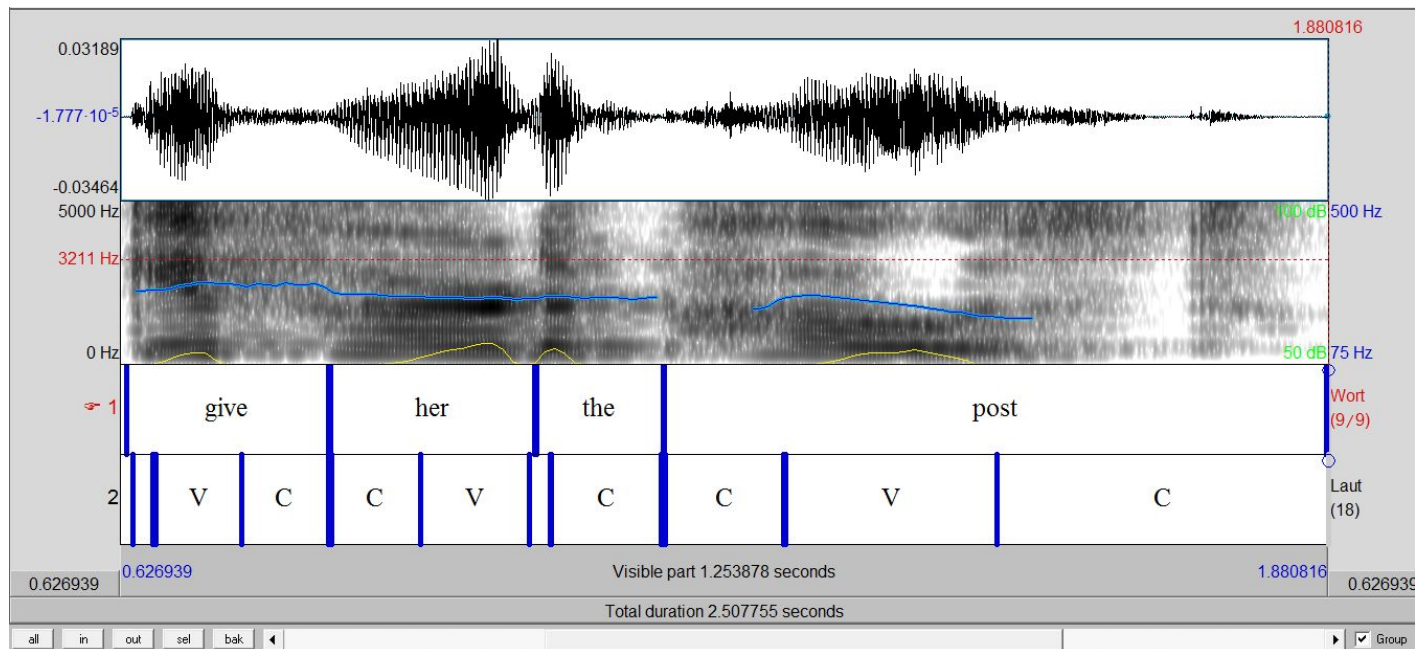


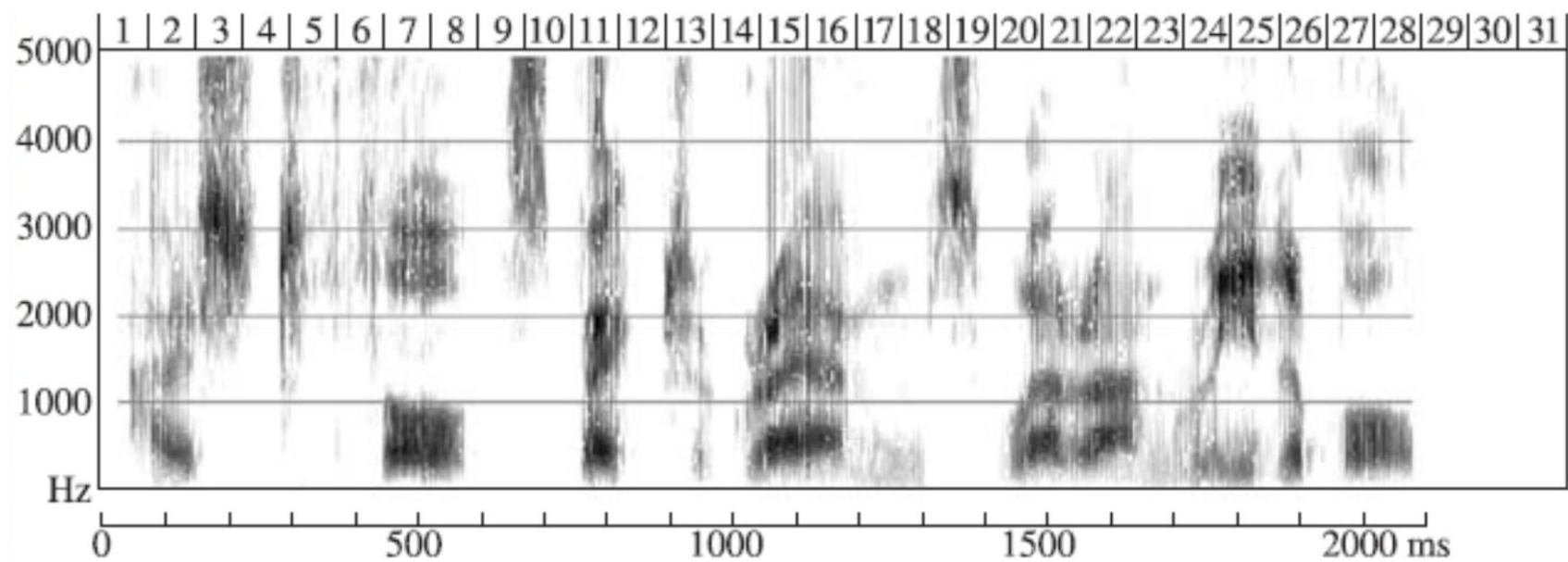
Stops

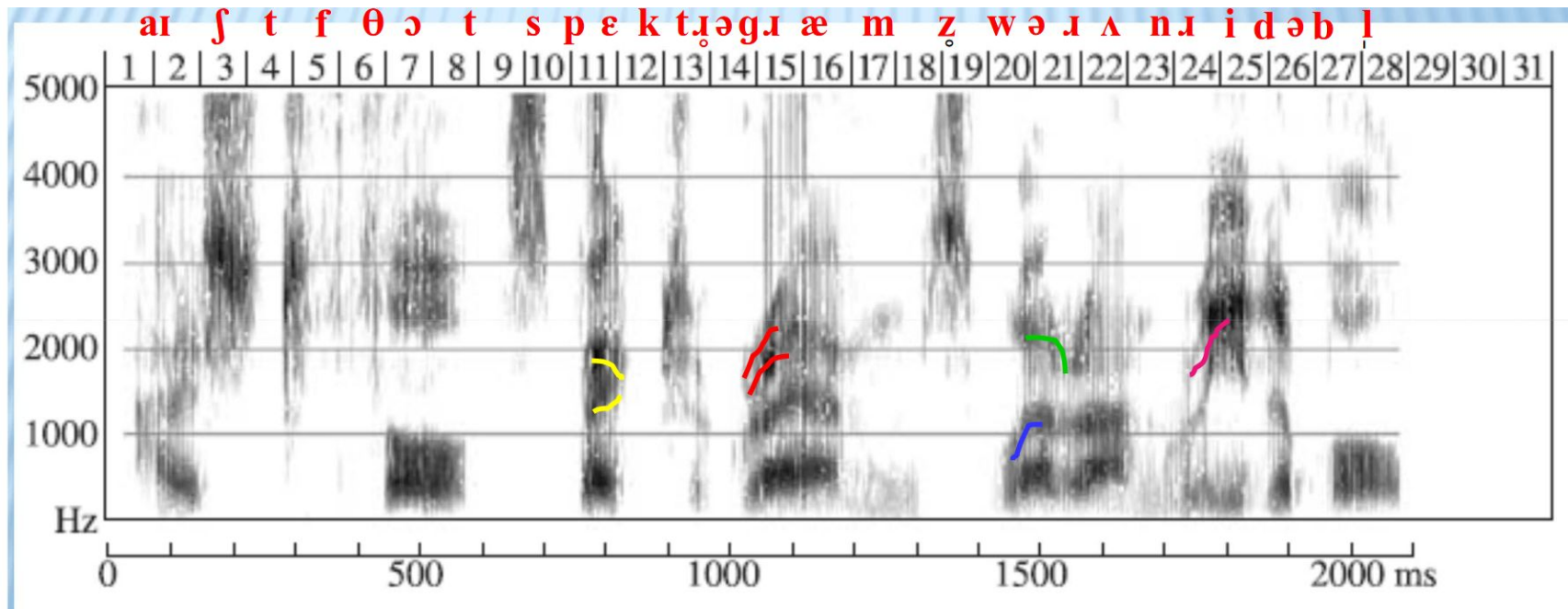


Fricatives

Spectrogram of continuous speech







Pop Quiz

What type of filter should you apply if you wanted to study:

1. Only fricatives
2. Only vowels
3. Only F0