

# COMP47700

## Speech and Audio

Dr Alessandro Ragano

Office: 3rd Floor, Science East, Insight Centre  
School of Computer Science  
University College Dublin



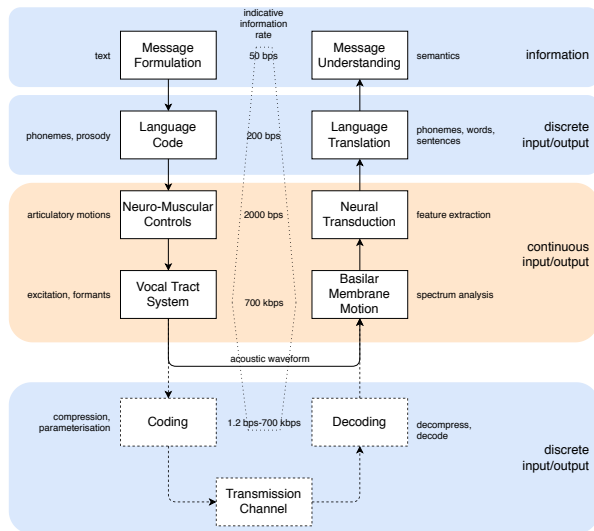
[alessandro.ragano@ucd.ie](mailto:alessandro.ragano@ucd.ie)

# 3.1.1

## Speech Production

# The Big Picture: The Speech Chain

## Speech Production, Transmission and Perception



Adapted from L. R. Rabiner and R. W. Schafer, Introduction to Digital Speech Processing, Foundations and Trend in Signal Processing, vol 1, no 1-2, pp 1-194, 2007

# The Speech Chain

## Simplified



## Production perspective

Look at speech from the **production perspective**: how generation effects the speech signal created and propagated in air and how the signal's components contribute to understanding

# Speech: Production – Characterisation – Understanding

## Speech production

### **How does the vocal tract create speech?**

Human speech production apparatus, mechanisms, and characteristics

## Characteristics of speech

### **Physical properties resulting from generation and propagation issues**

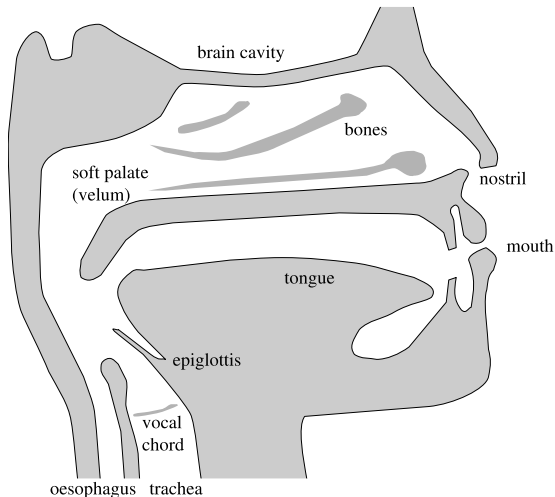
Loudness, frequency distribution, amplitude distribution, pitch rate, syllabic rate

## Speech understanding

### **Parts of speech (phones, phonemes, syllables etc.)**

Physical properties of speech resulting from generation mechanism

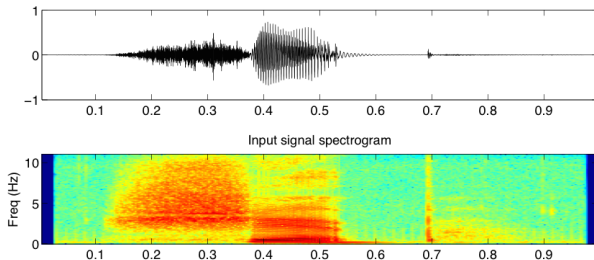
# Speech Production



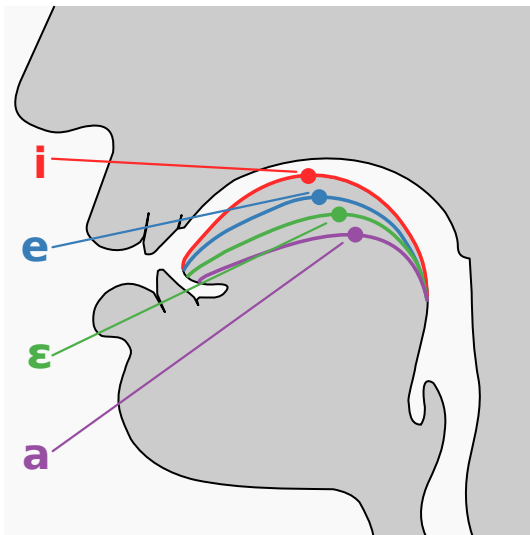
Sectional diagram of human vocal apparatus, showing major articulators, resonators and features of the vocal and nasal tracts.

# How is speech created?

- 1 **Lung power** mostly affects the **volume** of the sound, but **rapid variation** often **distinguishes a boundary** between syllables.
- 2 If the glottis is closed temporarily during speech, a **glottal stop** results such as the /t/ in a Yorkshire-accented reading of 'I went t' shops'. A plosive sound like the /d/ in 'dog', is a short **stop followed by an explosive release**.
- 3 **Vocal chord muscle tension** causes the chords to vibrate at different rates, forming the **pitch frequencies**. **Voiceless sounds** (e.g. /s/ in 'six'), where the vocal chords do not vibrate, have little or no pitch structure.
- 4 If the air is **diverted through the nose** by the velum closing, a nasal sound such as /m/ in 'mad' results. **Different timbre** also results from the slightly **different path length** from lungs to nose compared with lungs to mouth (imagine two different length organ pipes).
- 5 If the air travels through the mouth, a **humped tongue** and **opening then closing lower jaw** cause a vowel sound (e.g. /a/ in 'card'), if the lower jaw does not close, a glide (e.g. /w/ in 'won') is the result.
- 6 Different sounds also result if the **air is forced past the sides of a tongue** touching the roof of the mouth or the teeth (e.g. /l/ in 'luck', and the /th/ sound).



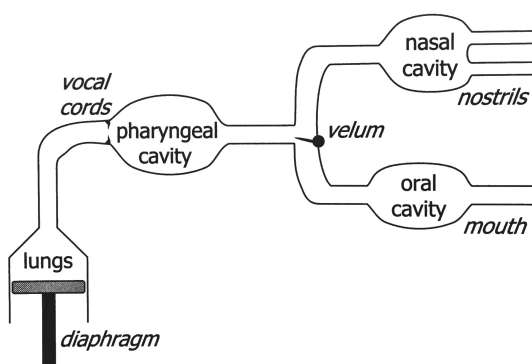
# Vowel Tongue Positions



Source: [https://en.wikipedia.org/wiki/International\\_Phonetic\\_Alphabet](https://en.wikipedia.org/wiki/International_Phonetic_Alphabet)



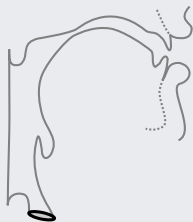
# Speech Production: Functional Diagram



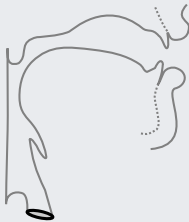
Source: McLaughlin (2016, p.56)

# Two Tube Model

## Creating Vowels



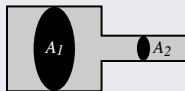
[ i ]



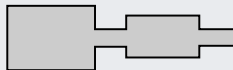
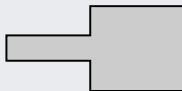
[ a ]



[ u ]

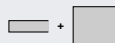
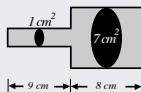


$l_1$   $l_2$



# Two Tube Model

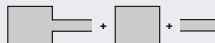
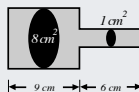
## Creating Intelligible Speech



972  
2917  
.  
.  
.

1093  
.  
.  
.

Formant	Actual	Estimated
F1	789	972
F2	1276	1093
F3	2808	2917
.	.	.
.	.	.



268

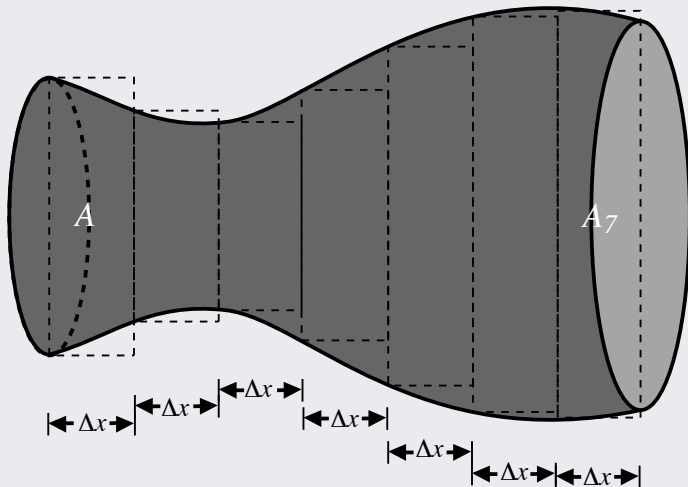
1944

2917

Formant	Actual	Estimated
F1	256	268
F2	1905	1944
F3	2917	2917
.	.	.
.	.	.

# Multi Tube Model

## Creating Realistic Speech



# Two Tube Model Demo

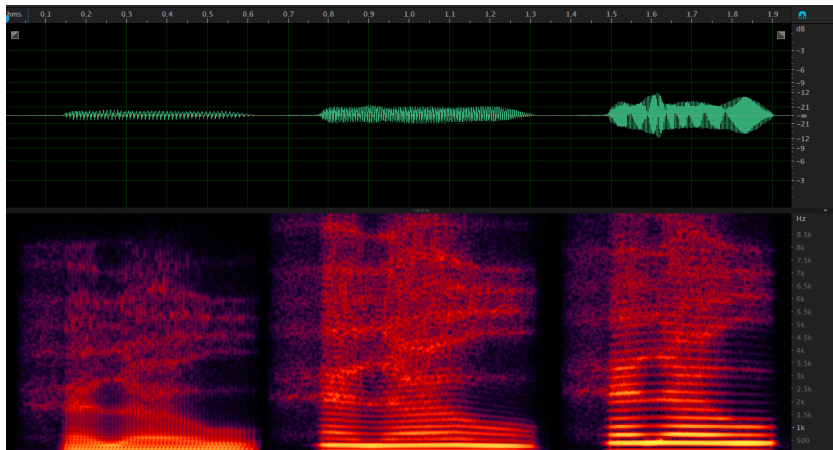
Potential Project Topic: Tube Modelling

GitHub repo for a Python Demo of tube synthesis:

<https://github.com/shun60s/Vocal-Tube-Estimation>

# Synthesis of Hello

The same greeting but different sounds and appearances giving use 'other' information beyond the word "hello"



Source: Signal adapted from <http://pages.cpsc.ualgary.ca/~hill/helloComparison/helloComparison.wav>