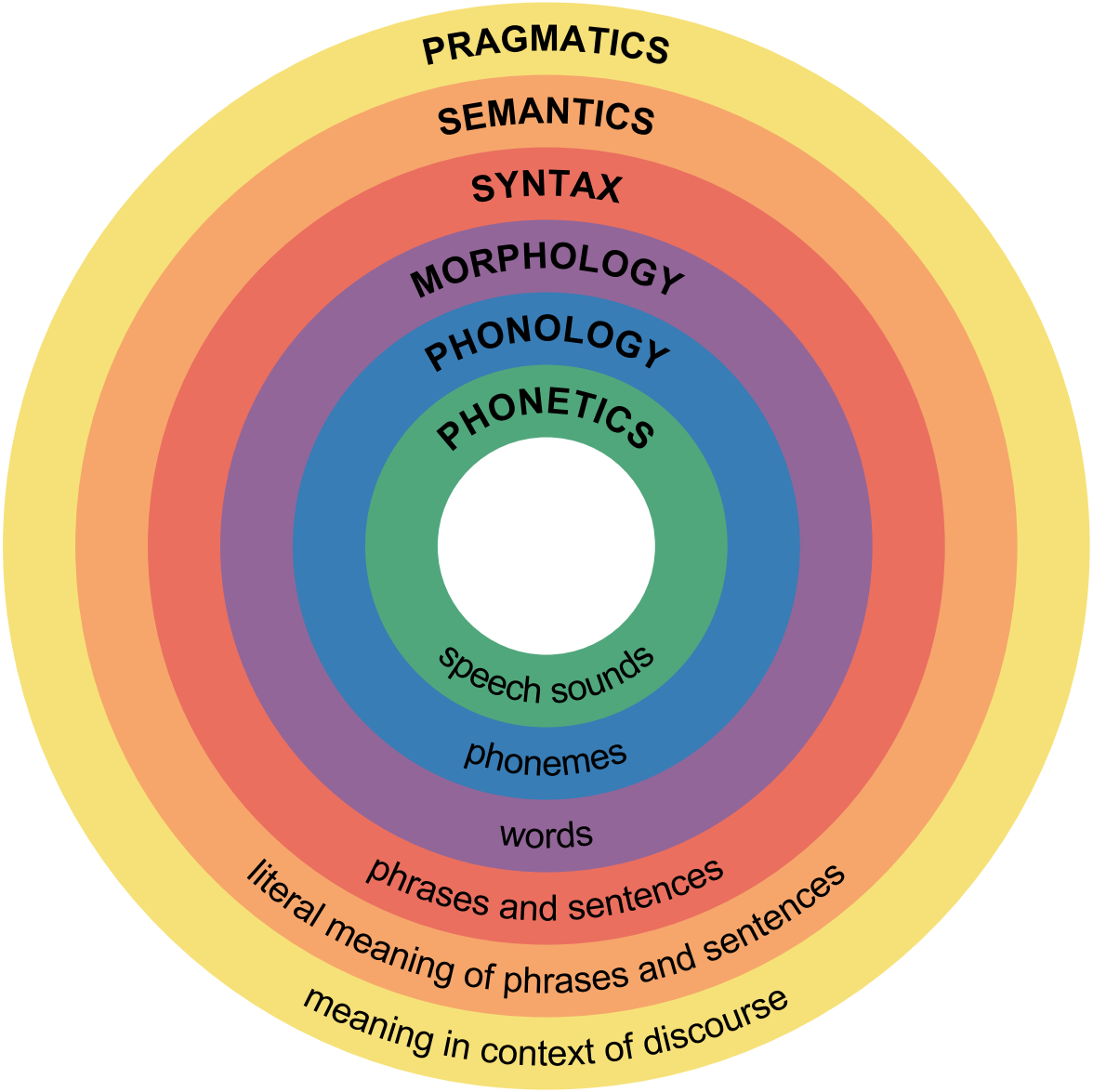
Phonetics and Phonology

(Natural Language Processing)

HND Thilini
hnd@ucsc.cmb.ac.lk

Major Levels of Linguistics



Phonetics

- The suffix -phone- is used in English, to denote words which are sound related
  - e.g. microphone, telephone
- The suffix -tics means ‘study of’, ‘art of’, ‘science of’ or technique
  - e.g. mathematics, optics, statistics

phonetics

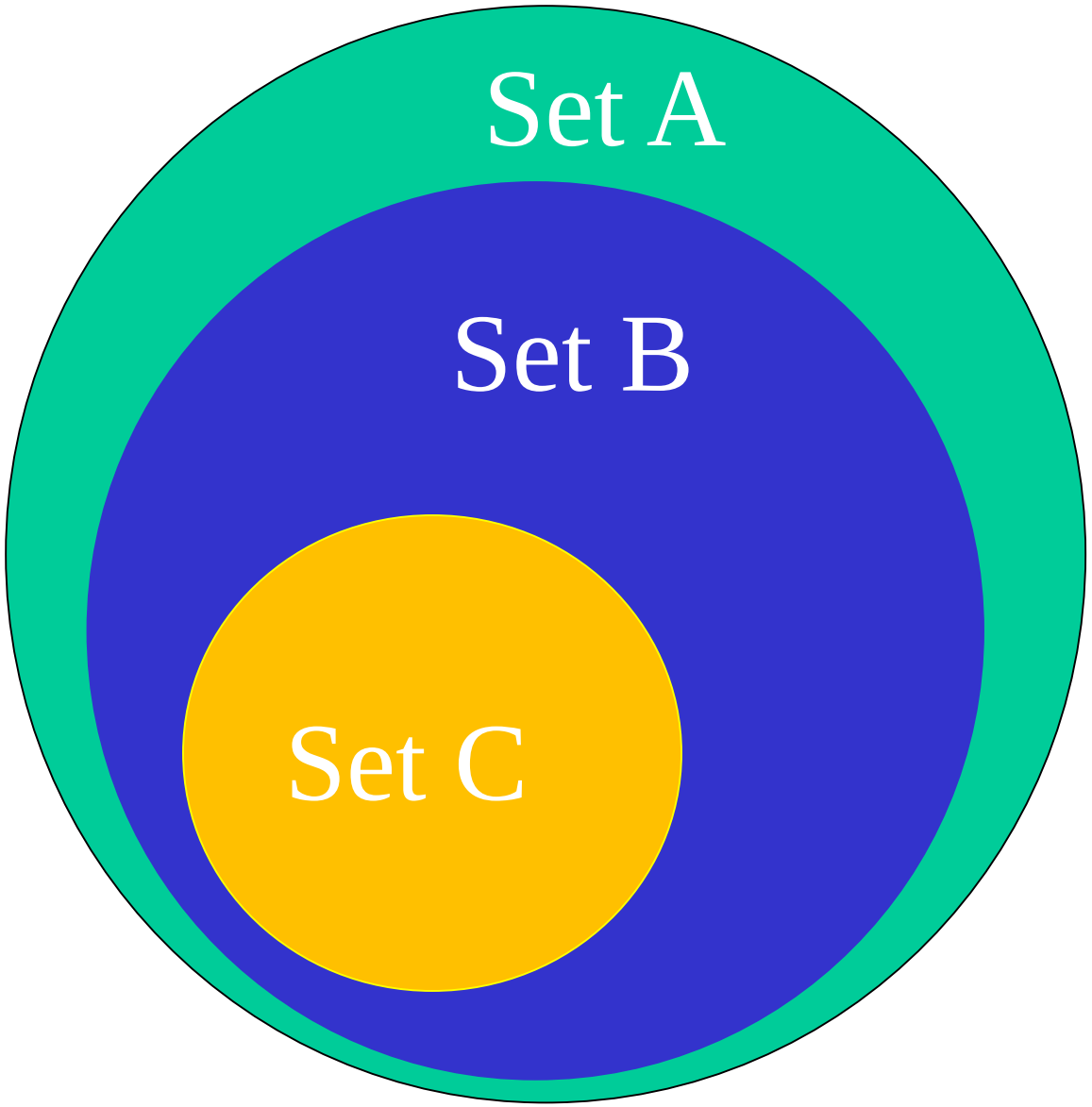
Definition

- Phonetics is the study and classification of speech sounds used in the languages of the world.

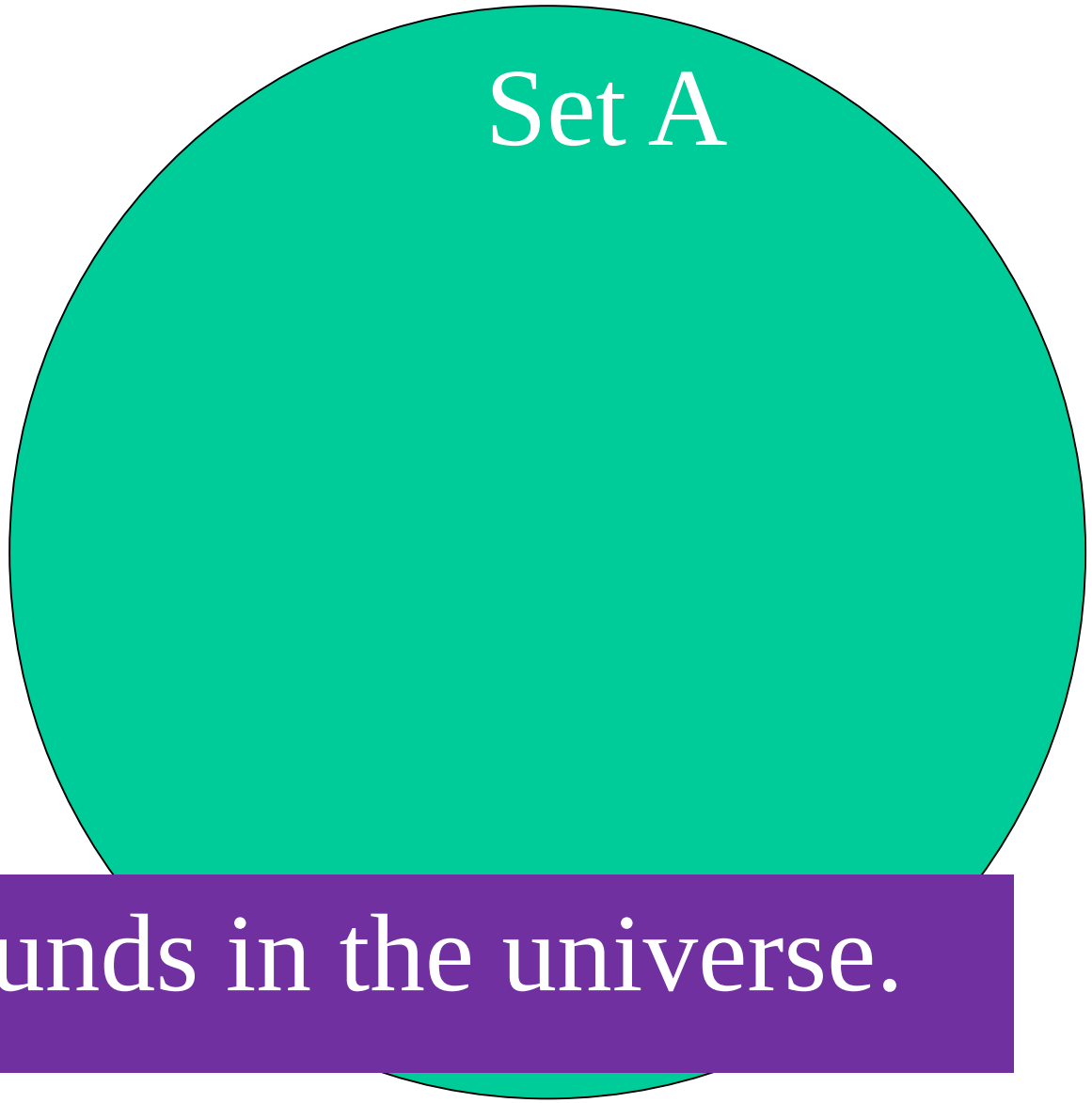
  - **What is Speech?**
  - **How speech is being produced?**

What are *Speech Sounds?*

What are Speech Sounds???

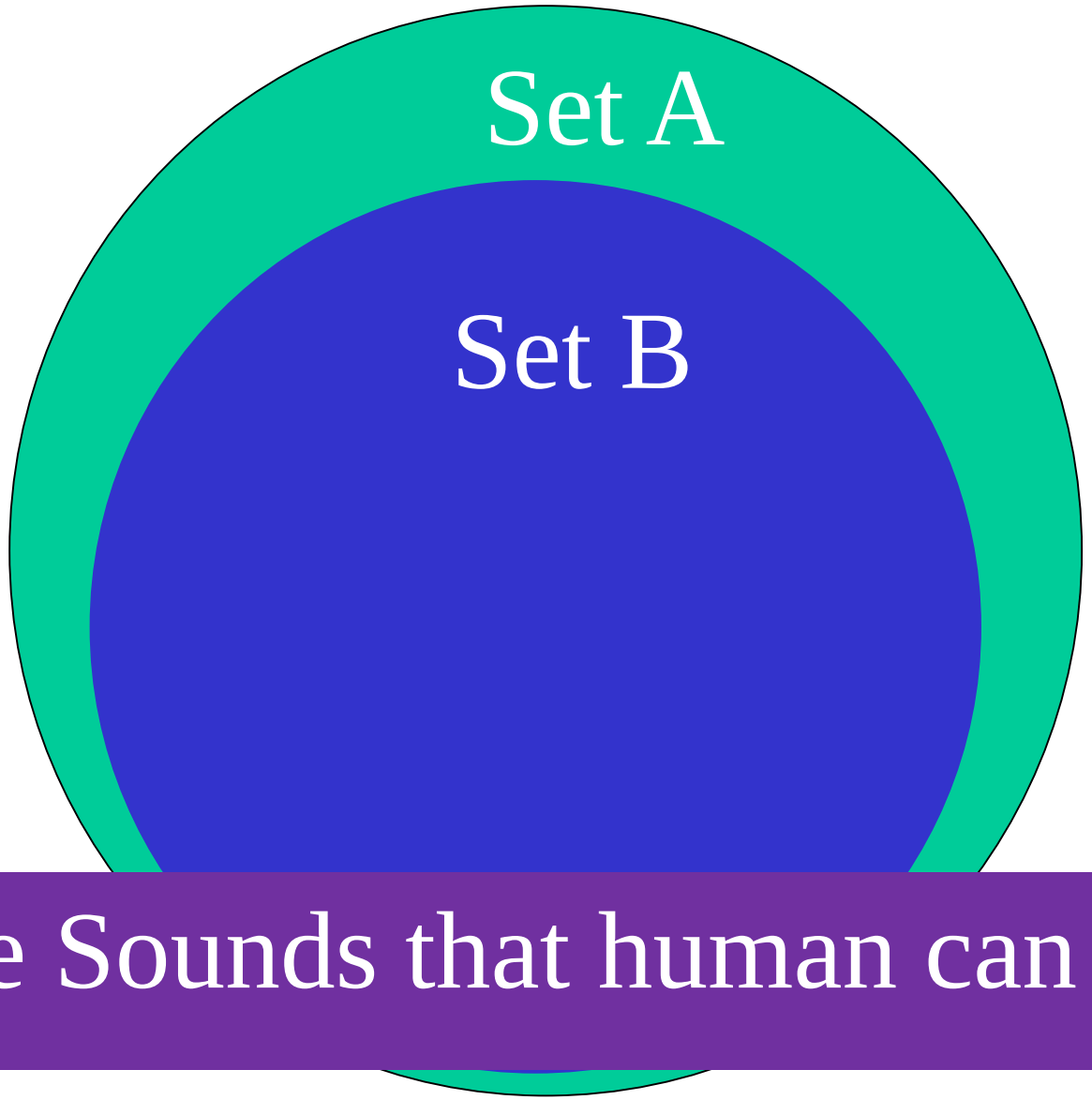


What are Speech Sounds???



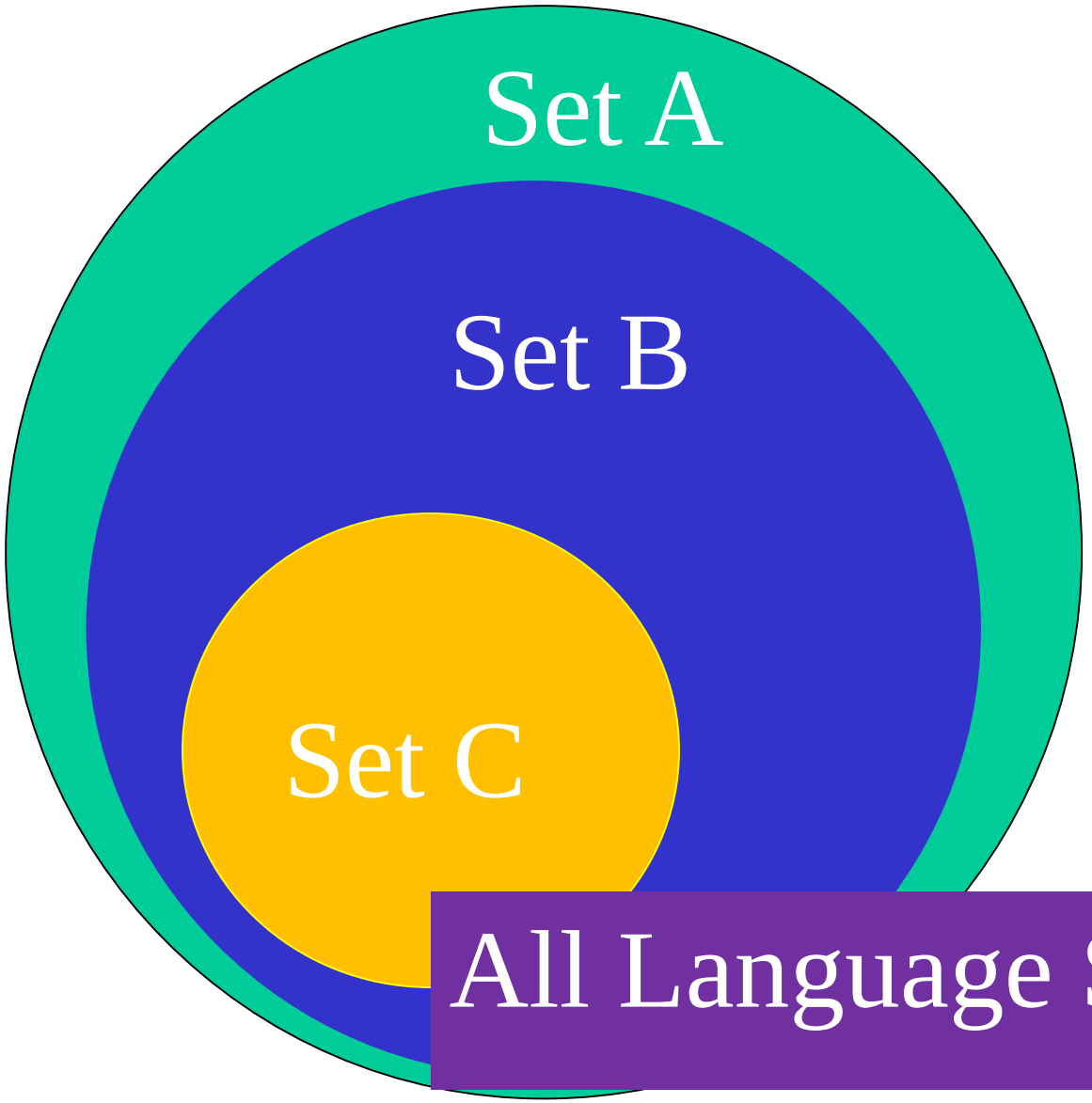
All sounds in the universe.

What are Speech Sounds???



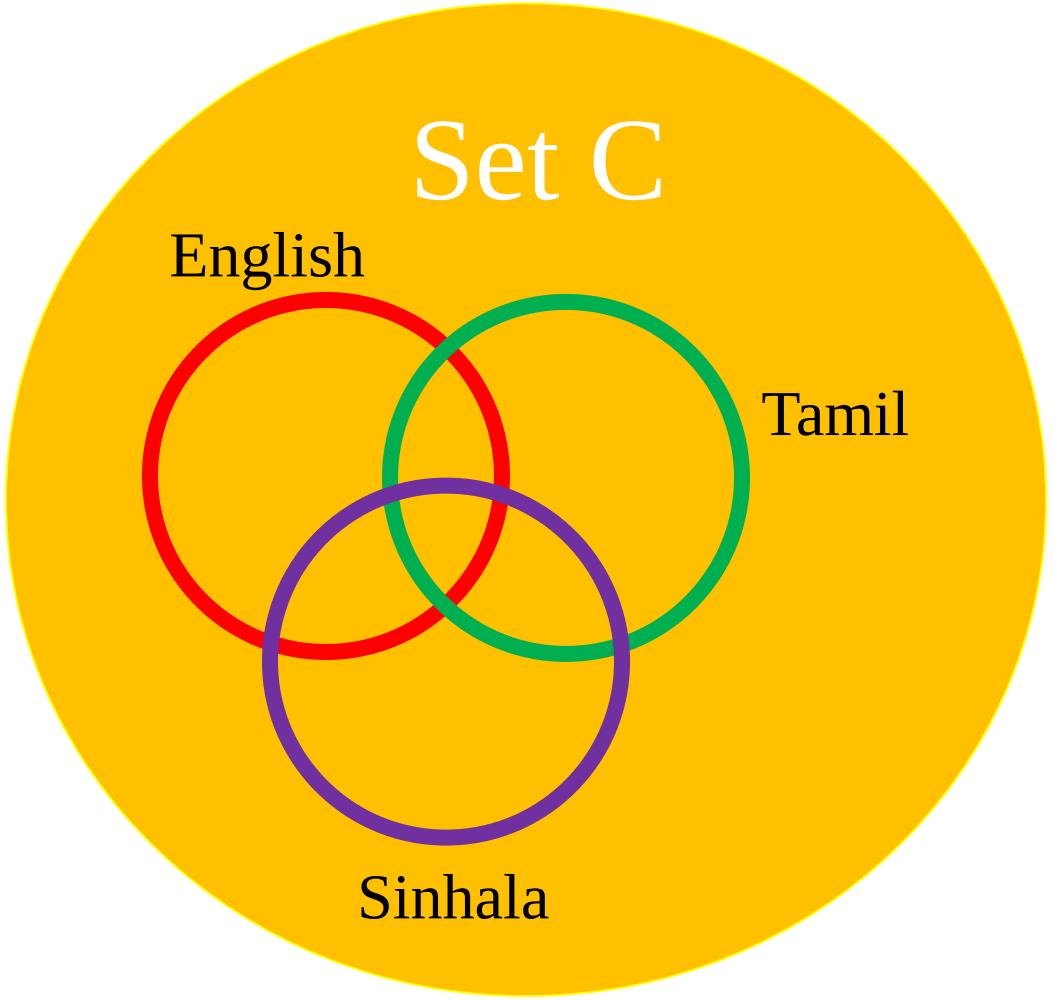
All the Sounds that human can hear

What are Speech Sounds???



All Language Sounds

What are Speech Sounds???



Phonology?

- The suffix -logy means the ‘systematic study of’
  - e.g. biology, zoology

phonology

Phonology

- How we understand?
- How our brain process speech?
- Grammar of speech!

- Definition
  - Phonology is the area of linguistics that describes the systematic way that sounds are differently realized in different context/environments and how this system of sound is related to the rest of the grammar

Grammar of Speech

- Eg: ලිප් ගල්

ල්= /l/
ඉ = /i/
ප් = /p/

ග්= /g/
අ = /a/
ල්= /l/

Grammar of Speech

- Eg: ලිප් ගල්

ල්= /l/
ඉ = /i/
ප් = /p/

ග්= /g/
අ = /a/
ල්= /l/

උක් ගස්, අත් දිග

Phonetics vs Phonology

- **Phonetics** deals with the production of speech sounds by humans, often without prior knowledge of the language being spoken.
  - What are the sounds? How are they made in the mouth?

- **Phonology** is about patterns of sounds, especially different patterns of sounds in different languages, or within each language, different patterns of sounds in different positions in words etc.
  - How do sounds combine? How are they memorized?

Speech Sounds

ல், ல் = /l/

க், க் = /g/

இ, இ = /i/

ப், ப் = /p/

அ, அ = /a/

Speech Sounds

ல், ல் = /l/

க், க் = /g/

இ, இ = /i/

ப், ப் = /p/

அ, அ = /a/

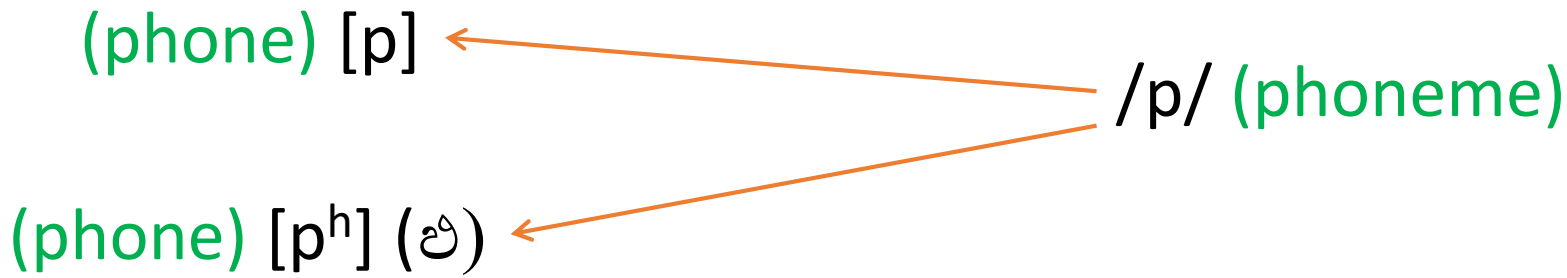
- Bath
- Advertisement
- Chance
- Class

Speech Sounds

- Phone:
  - The smallest sound unit of speech

    ప[p] (phone)

- Phoneme:
  - The smallest unit of speech that affects the meaning of a word.

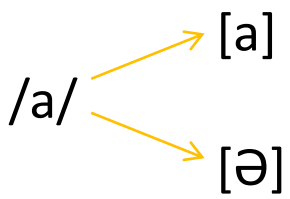


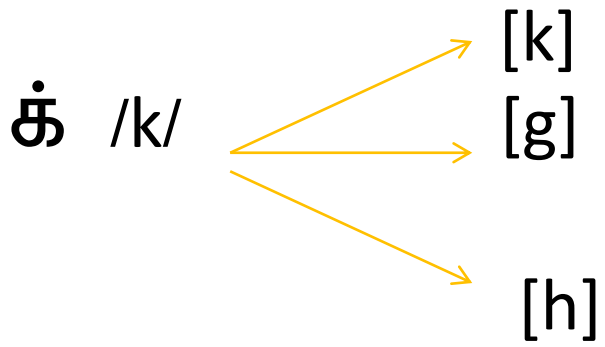
Phone vs Phoneme

- In phonetics and linguistics, a **phone** is any distinct speech sound or gesture, regardless of whether the exact sound is critical to the meanings of words.
- In contrast, a **phoneme** is a speech sound **in a** given language that, if swapped with another **phoneme**, could change one word to another.

Speech Sounds

- Allophone
  - is one of a set of multiple possible spoken sounds (or phones) or signs used to pronounce a single phoneme in a particular language





Minimal Pairs

- A minimal pair is a pair of words that have different meanings and which differ in only one sound.
  - Examples from English:
    - [but] and [cut] - /b/ and /k/
    - [at] and [it] - /æ/and /i/
    - [sip] and [zip] -/s/ and /z/

  - Examples from Sinhala:
    - [මල] and [ගල] = [malə] and [galə] - /m/ and /g/
    - [කර] and [කරා] = [karə] and [kara:] - /a/ and /a:/
    - [සංඳා] and [සංකා] = [saɲɳa:] and [saɲka:] - /ɳ/ and /k/

Speech Sounds

- Number of sounds differs from language to language
  - English
    - 44 sounds (20 vowels and 24 consonants)
  - Sinhala
    - 40 sounds (14 vowels and 26 consonants)
  - Tamil
    - 33 sounds (10 vowels and 23 consonants)

Transcribing Speech

- How to represent speech sounds?

Transcribing Speech

- How to represent speech sounds?
- What is the sound of letter “a” in following words?

Face

Bat

Hall

https://tophonetics.com/

Transcribing Speech

- How to represent speech sounds?
  - Can you pronounce following

    Ghoti

Transcribing Speech

- How to represent speech sounds?
  - Can you pronounce following

Ghoti => Fish

Transcribing Speech

- How to represent speech sounds?
  - Can you pronounce following

    Ghoti => Fish

    Why?
    - *gh*, pronounced /f/ as in *enough* /ɪˈnʌf/ or *tough* /tʌf/;
    - *o*, pronounced /ɪ/ as in *women* /ˈwɪmɪn/;
    - *ti*, pronounced /ʃ/ as in *nation* /ˈneɪʃən/ or *motion* /ˈmoʊʃən/.

Transcribing Speech

- How to represent speech sounds?
  - Can you pronounce following

    Ghoti => Fish

    Why?
    - *gh*, pronounced /f/ as in *enough* /ɪˈnʌf/ or *tough* /tʌf/;
    - *o*, pronounced /ɪ/ as in *women* /ˈwɪmɪn/;
    - *ti*, pronounced /ʃ/ as in *nation* /ˈneɪʃən/ or *motion* /ˈmoʊʃən/.
- So, need a one to one mapping.

https://en.wikipedia.org/wiki/Ghoti

Transcribing Speech

- How to represent speech sounds?

International Phonetic Alphabet
(IPA)

IPA

- Set of symbols intended as a universal system for transcribing speech sounds
- Provides,
  - Notational standard for the phonetic representation of all languages
  - Accurate and unique way of representing the sounds of any spoken language

IPA

THE INTERNATIONAL PHONETIC ALPHABET (revised to 2015)

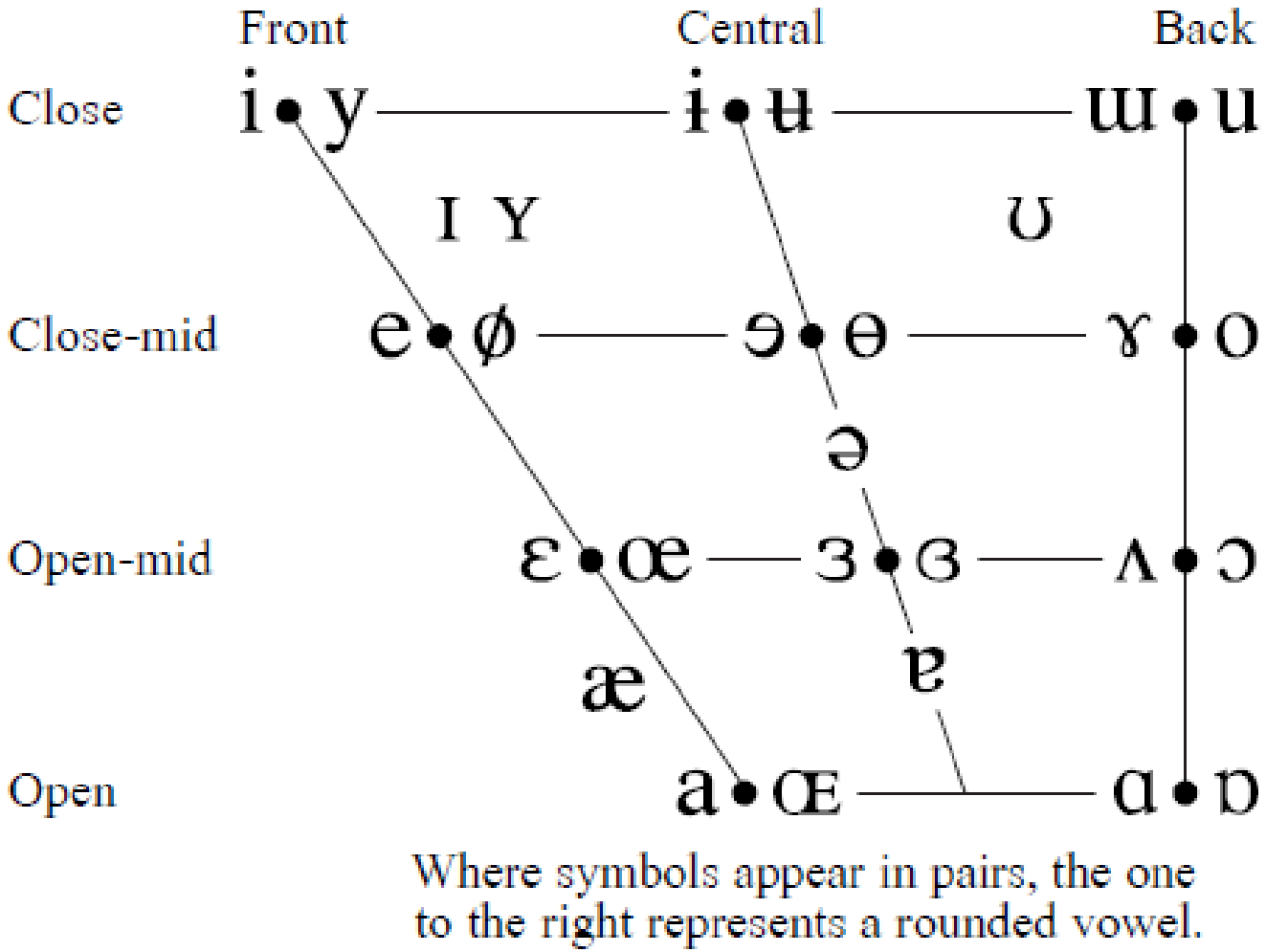
CONSONANTS (PULMONIC)

	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	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			

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

IPA

VOWELS



Where symbols appear in pairs, the one to the right represents a rounded vowel.

IPA for Sinhala and Tamil

Vowels

Sound	IPA	Sound	IPA
අ/அ	a	ආ/ஆ	a:
ඇ	æ	ඈ	æ:
ඉ/இ	i	ඊ/ஈ	i:
උ/உ	u	ඌ/ஊ	u:
එ/எ	e	ඒ/ஏ	e:
ඔ/ஒ	o	ඕ/ஓ	o:
schwa	ə		ə:

Consonants

Sound	IPA	Sound	IPA	Sound	IPA
ක/ඛ/க	k	ත/ථ/த	t	ල/ළ/ல	l
ග/ඝ/க	g	ද/ධ/த	d	ව/வ	w
ඟ	g̃	ඳ	d̃	ස/ச	s
ච/ஜ	c	ප/ඵ/ப	p	හ/க	h
ජ/ඣ	ɟ	බ/භ/ப	b	ශ/ෂ/ஶ	ʃ/ʂ
ට/ඨ/ட	ʈ	ම/ம	m	ෆ/ஃப்	f
ඩ/ඪ/ட	ɖ	ඹ	b̃	ඬ/ங/ங	ŋ
ඬ	ɖ̃	ය/ய	j	ඤ/ஞ	ɲ
න/ණ/ன	n	ර/ඍ/ற	r	ள	ɭ
ண	ɳ	ඦ	ɲ̃		

https://en.wikipedia.org/wiki/Help:IPA/Tamil

IPA for Sinhala

- Exercise:

අනුරාධපුරය	அனுராதபுரம்
ප්‍රාථමික	முதன்மை
ආරම්භක	
ශුෂ්ක	
කෘති	
ශ්‍රැති	

IPA for Sinhala

- Exercise:

අනුරාධපුරය	[anura:dəpurəjə]
ප්‍රාථමික	[pra:təmikə]
ආරම්භක	[a:rambəkə]
ශුෂ්ක	[ʃuʃkə]
කෘති	[kruti]
ශ්‍රැති	[ʃruti]

Types of Phonetics

1. Articulatory Phonetics

2. Acoustic Phonetics

3. Auditory/Perceptual phonetics

1. Articulatory Phonetics

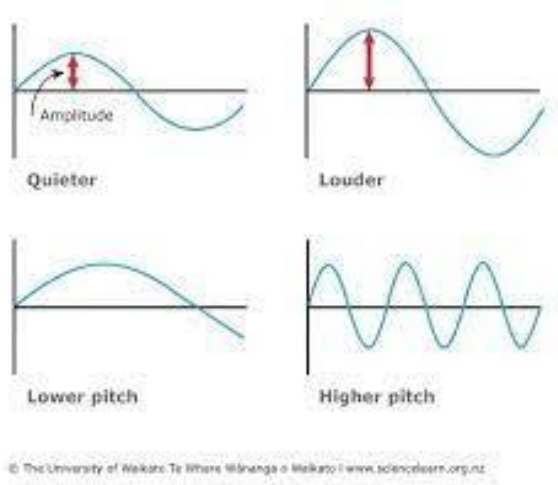
- Depending on the generated idea, our brain starts sending messages to articulators, to produce different sounds. The study of this process is known as Articulatory Phonetics

articulators:

In speech production, the moveable and non-moveable structures used to produce speech sounds (tongue, lips, jaw, palate).

2. Acoustic Phonetics

- Once speech is articulated, sound waves are generated. The properties of sound waves are worth studying for Text to Speech (TTS) systems and Automatic Speech Recognition (ASR) systems. This is known as, Acoustic Phonetics



3. Auditory/Perceptual phonetics

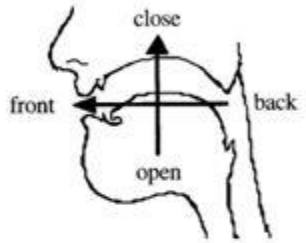
- It has been found that, what we hear is not exactly what is spoken

- Therefore, the study area of what we hear and how we hear is called

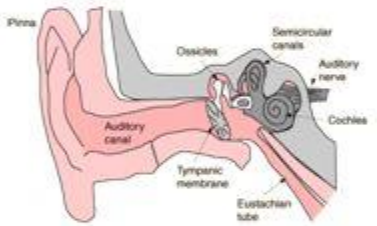
Auditory/Perceptual phonetics

Describing the sounds of language

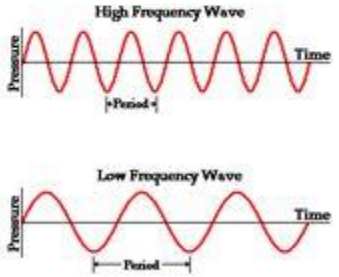
- articulatory phonetics



- auditory phonetics



- acoustic phonetics



Types of Phonetics

1. Articulatory Phonetics

2. Acoustic Phonetics

3. Auditory/Perceptual phonetics

Articulatory Phonetics

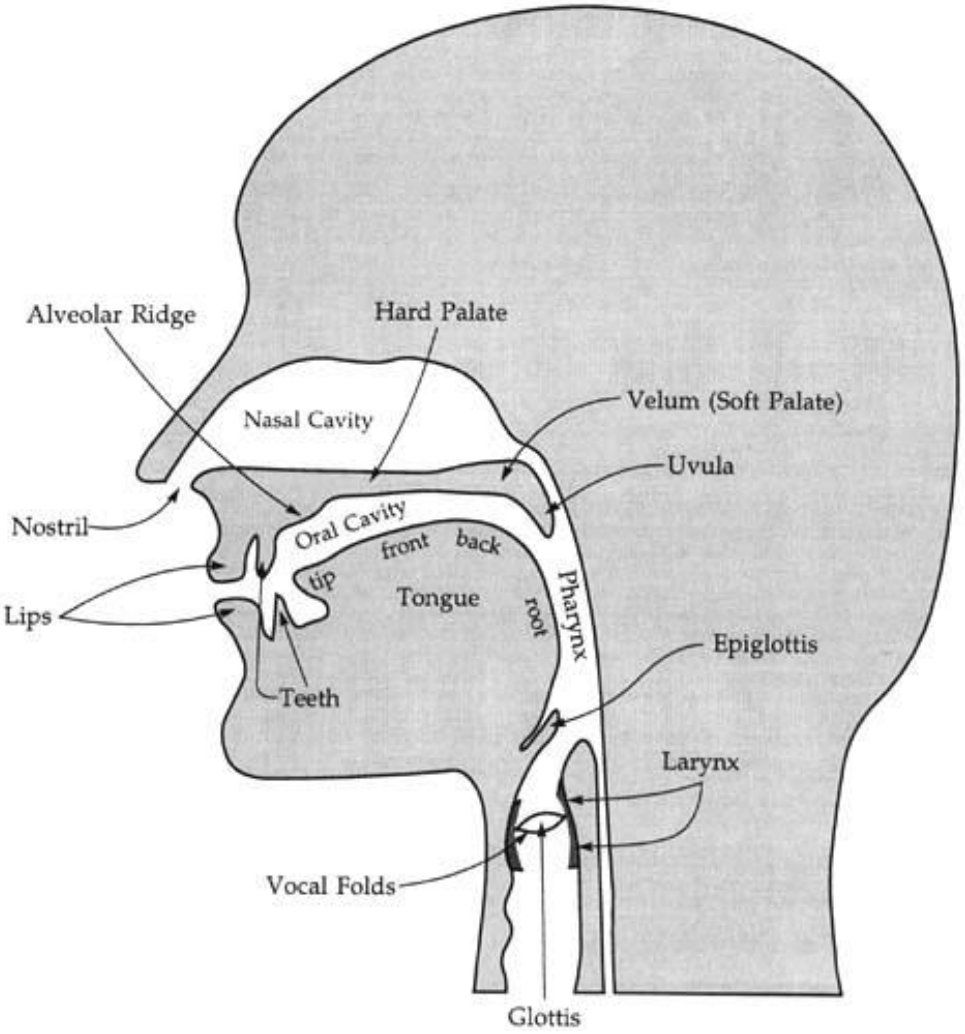
- How different **articulators** interact to create different sounds
- Anatomy of Vocal Organs
- Consonants
  - Places of Articulation
  - Manners of Articulation
  - Phonation
- Vowels

Anatomy of Vocal Organs

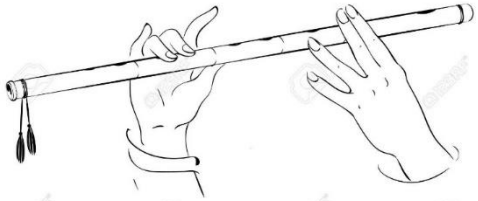
- Study of how speech sounds are produced by human vocal apparatus

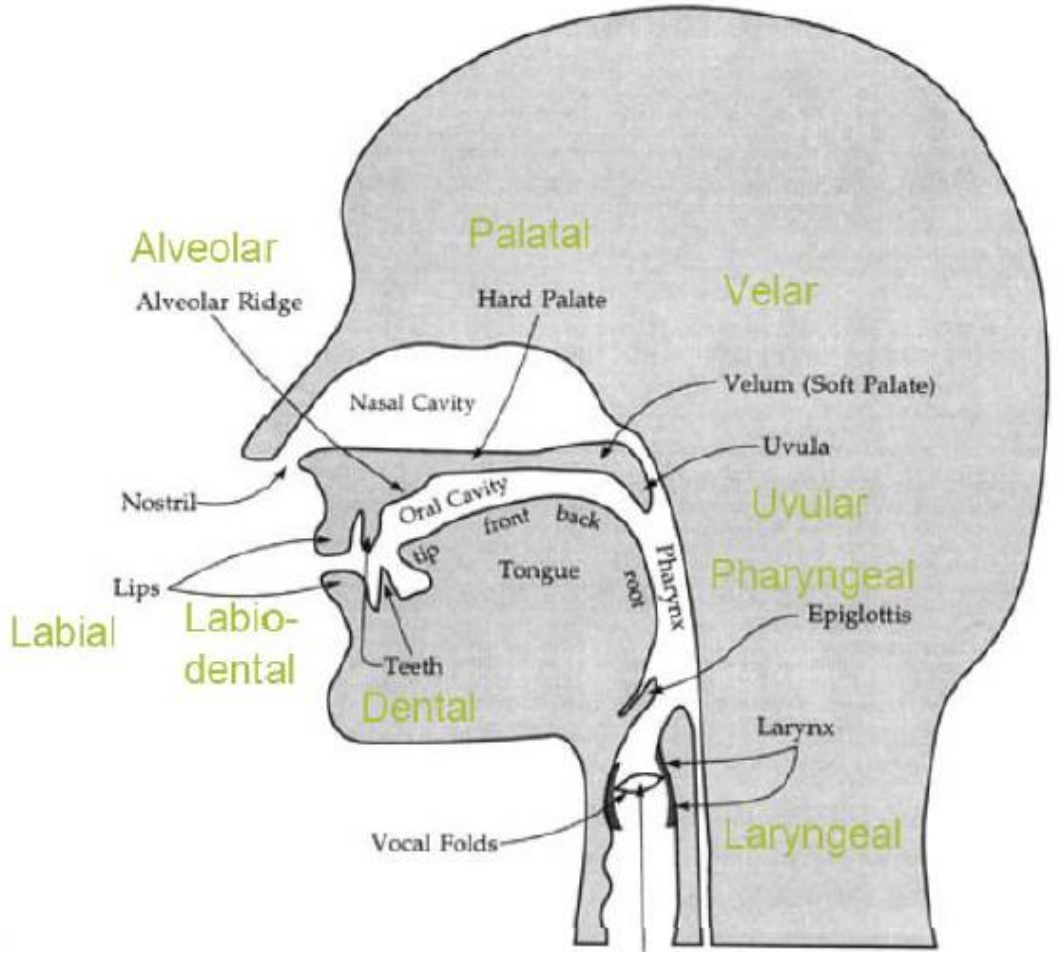
Vocal Tract

The instrument we have to generate sounds!



Vocal Tract




Alveolar
Palatal
Velar
Alveolar Ridge
Hard Palate
Nasal Cavity
Velum (Soft Palate)
Uvula
Nostril
Oral Cavity
front
back
tip
Tongue
root
Pharynx
Uvular
Pharyngeal
Lips
Labial
Labio-
dental
Teeth
Dental
Epiglottis
Larynx
Vocal Folds
Laryngeal

Consonants & Vowels

- Producing a speech sound involves making the vocal tract narrower at some location than it usually is.

  narrowing/making a constriction

- Consonants can be said to have a greater degree of constriction than vowels

Consonants

- Which consonant you're pronouncing depends on,
  - Where the constriction is made in the vocal tract
  - How narrow it is (the degree of constriction)

- To precisely define a consonant,
  - describe whether the vocal folds are vibrating and whether air is flowing through the nose

Consonants

- Consonants can be classified along these three major dimensions:

1. Place of Articulation

   where the narrowing/constriction occurs
2. Manner of Articulation

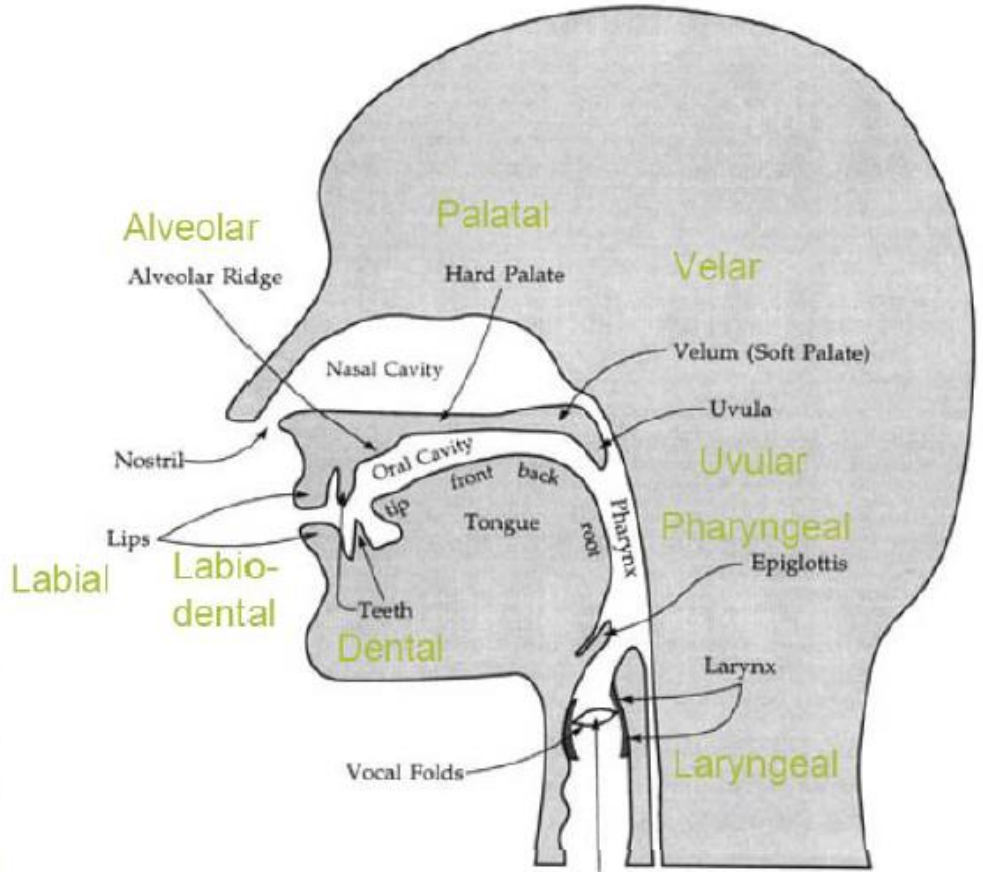
   how close they get;

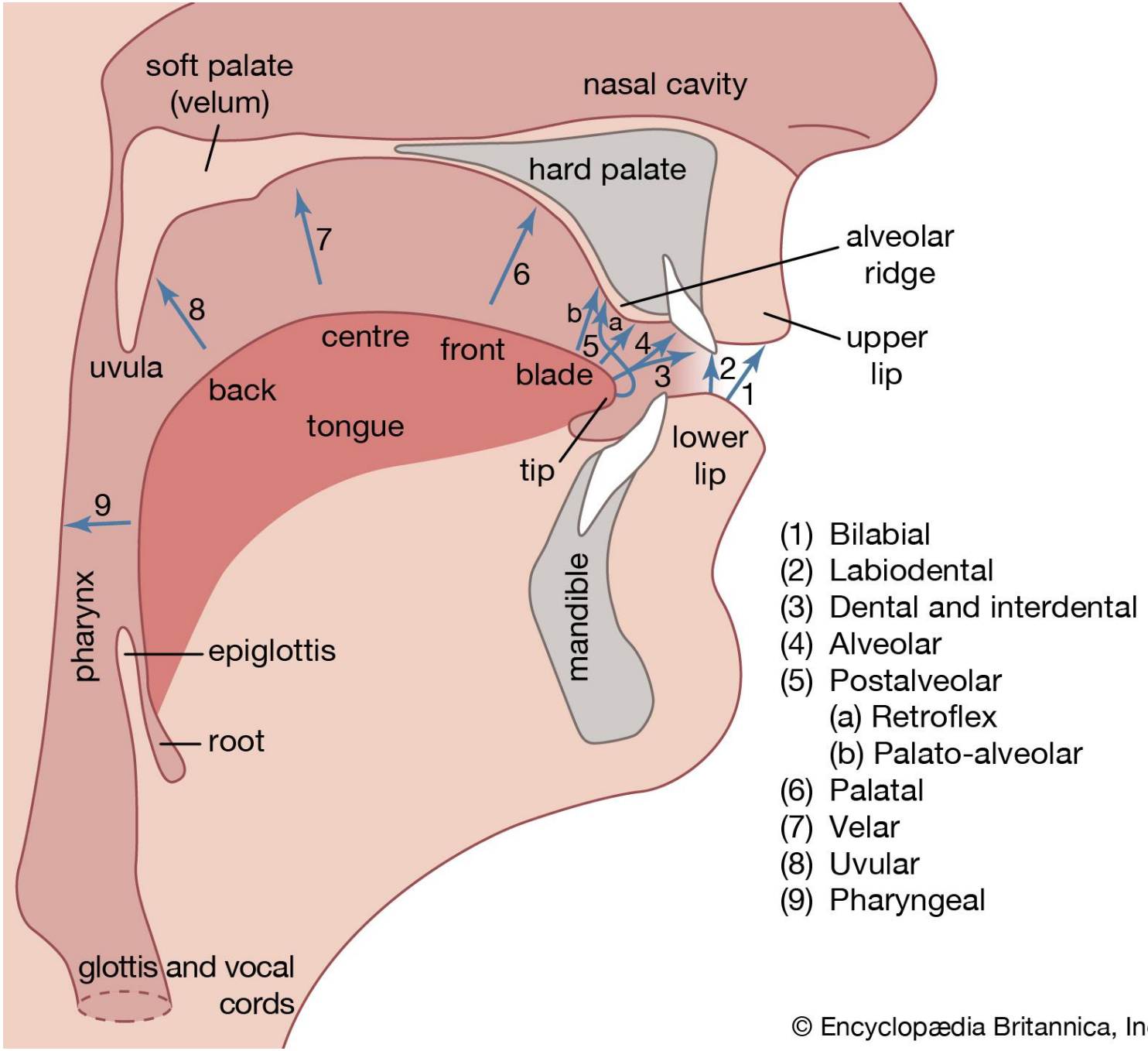
   how narrow the vocal tract constriction is
3. Voicing (Phonation)

   Vibration of the vocal folds

1. Place of Articulation

- Labials
- Dental
- Alveolar
- retroflex
- Palatal
- Velum
- Uvular
- Pharynx
- laryngeal





THE INTERNATIONAL PHONETIC ALPHABET (revised to 2015)

CONSONANTS (PULMONIC)

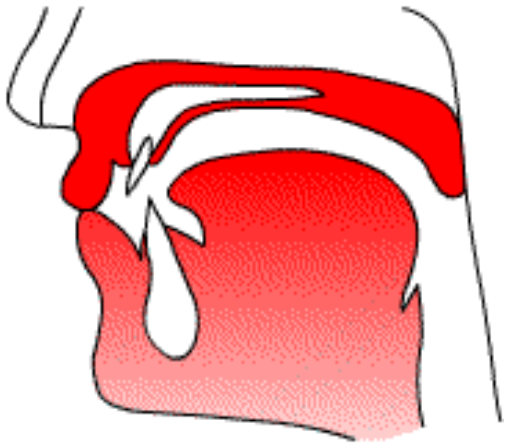
	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k g	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r					ʀ		
Tap or Flap		ⱱ		ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant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

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

Place of Articulation

- Labial
  - pronounced with active involvement of one or both lips as primary articulators
  - Bilabials
    - pronounced with the lips closed or nearly closed

p b m

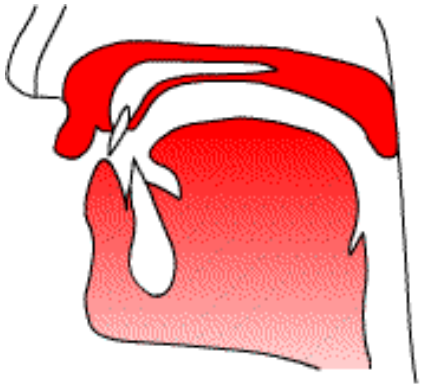


Place of Articulation

- Labial
  - pronounced with active involvement of one or both lips as primary articulators
  - Bilabials
    - pronounced with the lips closed or nearly closed

      p b m
  - Labiodentals
    - pronounced with the lower lip touching the upper teeth

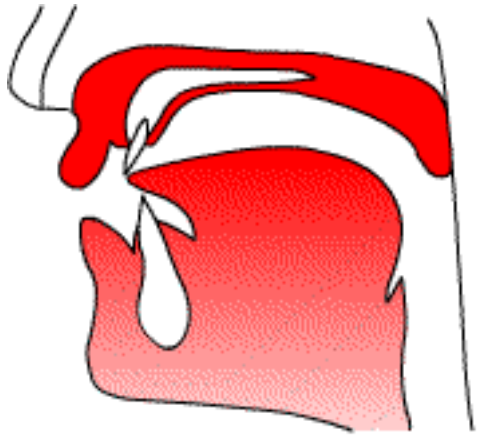
      f v



Place of Articulation

- Dental
  - the tip or blade of the tongue approaches or touches the upper teeth

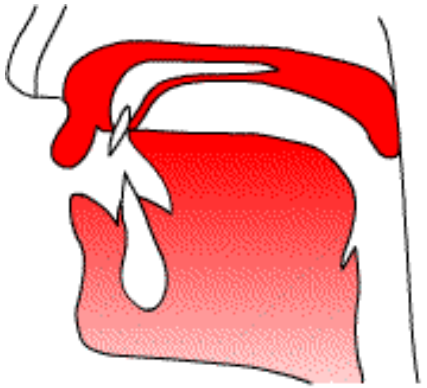
t d l



Place of Articulation

- Alveolar
  - the tongue tip/blade approaches or touches the alveolar ridge

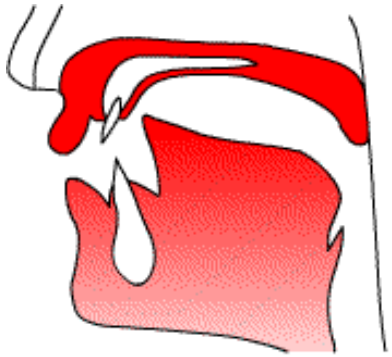
n s r



Place of Articulation

- Retroflex
  - pronounced with the tongue tip curled up and back

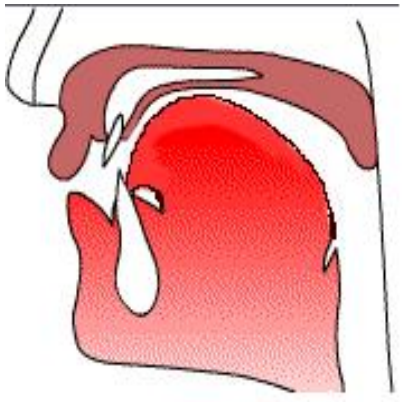
ʈ ɖ



Place of Articulation

- Palatal
  - pronounced with the blade of the tongue against or near the hard palate of the mouth

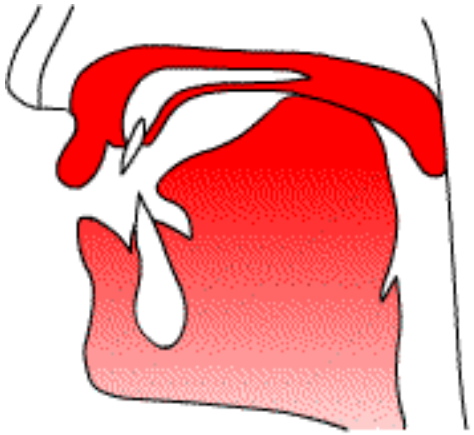
c ɟ j



Place of Articulation

- Velar
  - the body of the tongue approaches or touches the soft palate, or velum

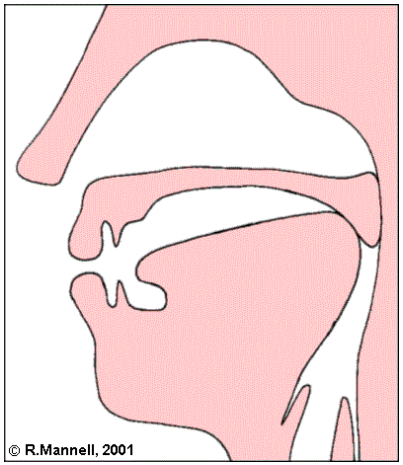
k g



Place of Articulation

- Uvular
  - tongue body is raised far enough back to approach the soft palate near the uvula

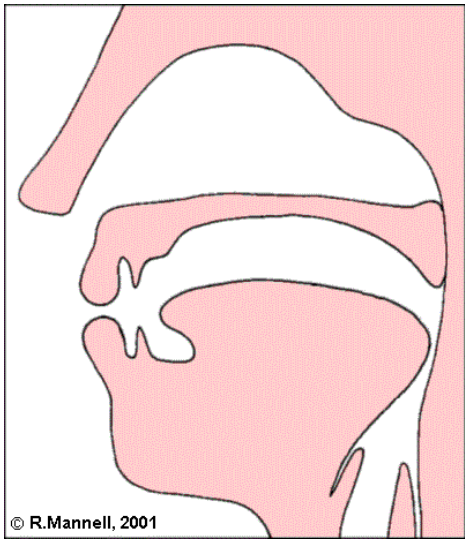
q G N



Place of Articulation

- Pharynx
  - Root of the tongue is retracted towards the back wall of the pharynx

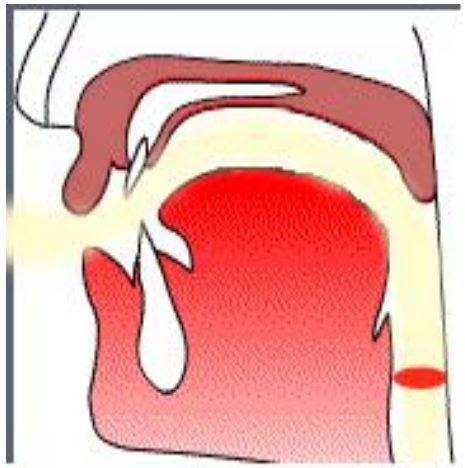
ħ



Place of Articulation

- Larynx / Glottal
  - Formed by a constriction in the glottis

h



https://www.mimicmethod.com/ft101/place-of-articulation/

https://seeingspeech.ac.uk/

Consonants

- Consonants can be classified along these three major dimensions:

1. Place of Articulation
   where the narrowing/constriction occurs
2. Manner of Articulation
   how close they get;
   how narrow the vocal tract constriction is
3. Voicing (Phonation)
   Vibration of the vocal folds

2. Manner of Articulation

- Completely Closed – Stops /occlusive/ Plosive
- Slightly Open – Fricatives
- Slightly More Open – Affricates
- Open – Approximants

THE INTERNATIONAL PHONETIC ALPHABET (revised to 2015)

CONSONANTS (PULMONIC)

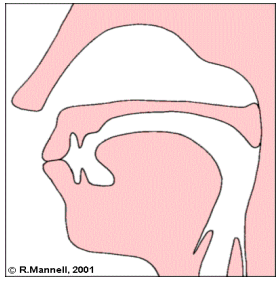
	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k g	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r					ʀ		
Tap or Flap		ⱱ		ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant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

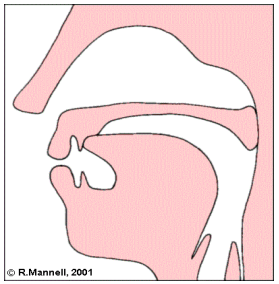
Symbols to the right in a cell are voiced, to the left are voiceless. Shaded areas denote articulations judged impossible.

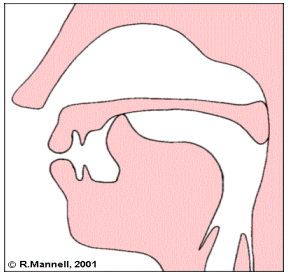
Manner of Articulation

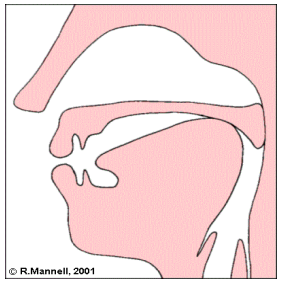
- Completely Closed – Stops
  - Complete blockage of ORAL tract
    - Oral Stops (Plosives)
      - the air-stream being stopped in the oral cavity and the soft palate is raised blocking off the nasal cavity

Bilabial-stops	Dental-stops	Retroflex-stops	Velar-stops
p b	t d	ʈ ɖ	k g





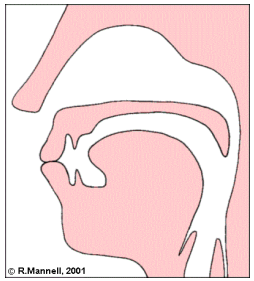


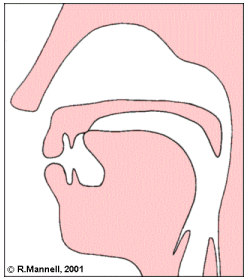


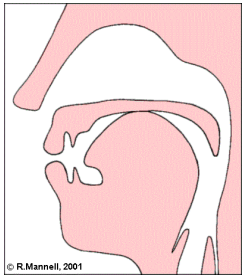
Manner of Articulation

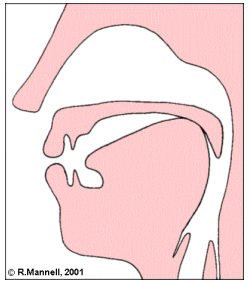
- Completely Closed – Stops
  - Complete blockage of ORAL tract
    - Nasal Stops
      - the air- stream being stopped in the oral cavity but the soft palate is down so that the air can go out through the nose

Bilabial-stops	Alveolar-stops	Palatal-stops	Velar-stops
m	n	ɲ	ŋ







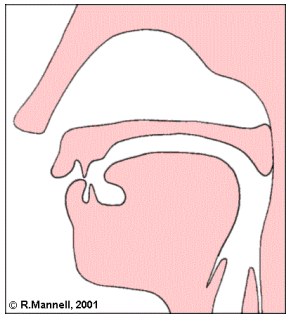


Manner of Articulation

- Slightly Open – Fricatives
  - articulators come close together, but there is a slight gap between them

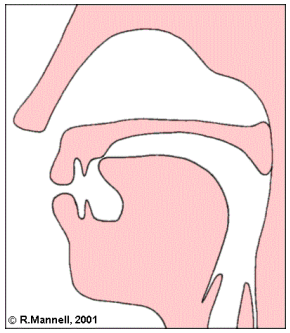
Labiodentals

f



Alveolar

s

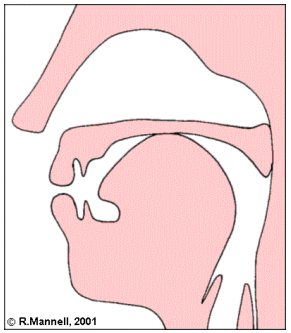


Manner of Articulation

- Slightly More Open – Affricates
  - Starts with a stop, and release like a fricative

Palatal

c ɟ

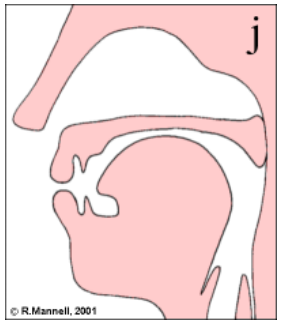


Manner of Articulation

- Open – Approximants
  - the air flows smoothly through the vocal apparatus

Palatal

j



Consonants

- Consonants can be classified along these three major dimensions:

1. Place of Articulation
   where the narrowing/constriction occurs
2. Manner of Articulation
   how close they get;
   how narrow the vocal tract constriction is
3. Voicing (Phonation)
   Vibration of the vocal folds

3. Voicing (Phonation)

- Voiced consonants: are produced when the vocal cords are vibrating

- Voiceless consonants: are produced when the vocal cords are not vibrating

Voiceless	Voiced
s	z

Voicing (Phonation)

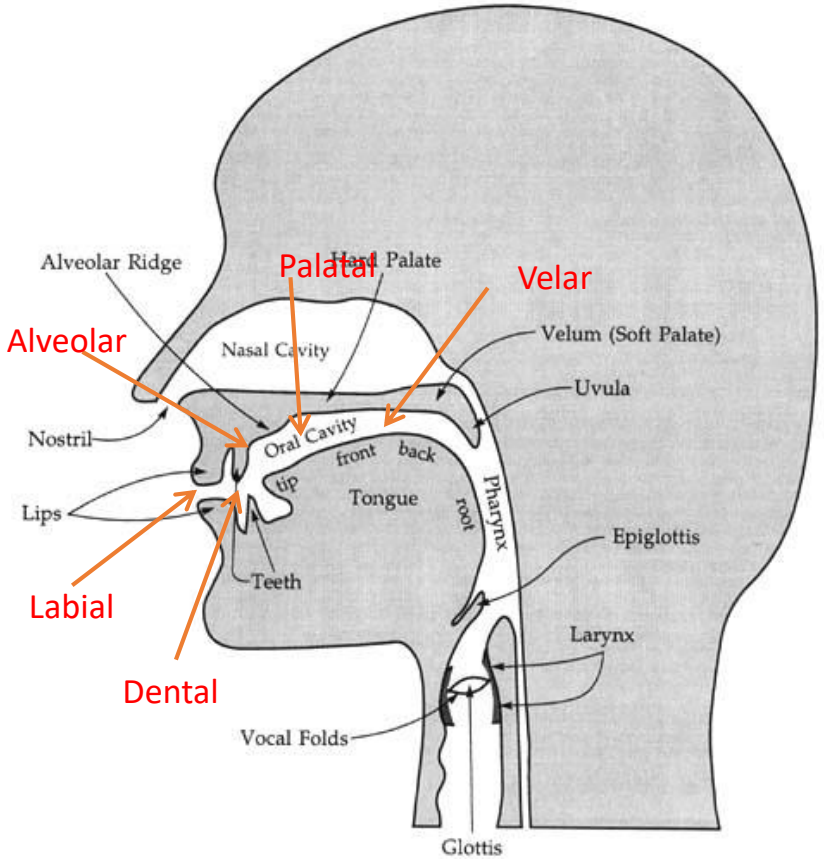
Voiceless		Voiced	
[k]	come	[g]	gum
[f]	fat	[v]	vat
[p]	pit	[b]	bit
[s]	sip	[z]	zip
[t]	ton	[d]	done

Articulatory Phonetics

Manner of Articulation

Completely Closed – Stops
Slightly Open – Fricatives
Slightly More Open – Affricates
Open – Approximants

Place of Articulation



Articulatory Phonetics

- Identify the Place and Manner of articulation of the following sounds

/s/
/n/
/f/
/t/
/j/
/g/

Articulatory Phonetics

- Identify the Place and Manner of articulation of the following sounds

	Place	Manner
/s/	alveolar	fricative
/n/	alveolar	nasal-stop
/f/	Labio-dental	fricative
/t/	dental	oral-stop
/j/	palatal	approximant
/g/	Velar	oral-stop

IPA

THE INTERNATIONAL PHONETIC ALPHABET (revised to 2015)

CONSONANTS (PULMONIC)

	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	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			

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

Vowels

- Monophthongs
  - One vowel
  - A monophthong is where there is one vowel sound in a syllable
- Dipthongs
  - Two vowels
  - A diphthong is where there are two vowel sounds in a syllable.

Syllables

- A syllable is a 'unit of pronunciation'.
  - *Mouse* (1 syllable)
  - *Rabbit* (2 syllables)
  - *Kangaroo* (3 syllables)
  - *Barracuda* (4 syllables)
  - *Hippopotamus* (5 syllables)
- Any word must have at least one syllable; even the word 'a' has one syllable.
- A word has two syllables when there are two vowel sounds divided by a consonant sound, or, to put it another way, two vowel sounds connected by a consonant sound.
- A word has three syllables when there are three vowel sounds divided/connected by two consonant sounds.

Vowels

- Monophthongs
  - One vowel
  - Examples: Funny - a sound and i sound
- Dipthongs
  - Two vowels
  - Examples: Guy and Behind - /ai/

Vowels

- Vowels are open sounds because they involve no obstruction to the flow of air
  - From lungs as it passes up through the windpipe (trachea), through the voice box (larynx) and out of the mouth
- Vowels are made by slight movements of tongue and lip postures

- All vowels are produced with the vocal folds vibrating and are said to be voiced sounds.

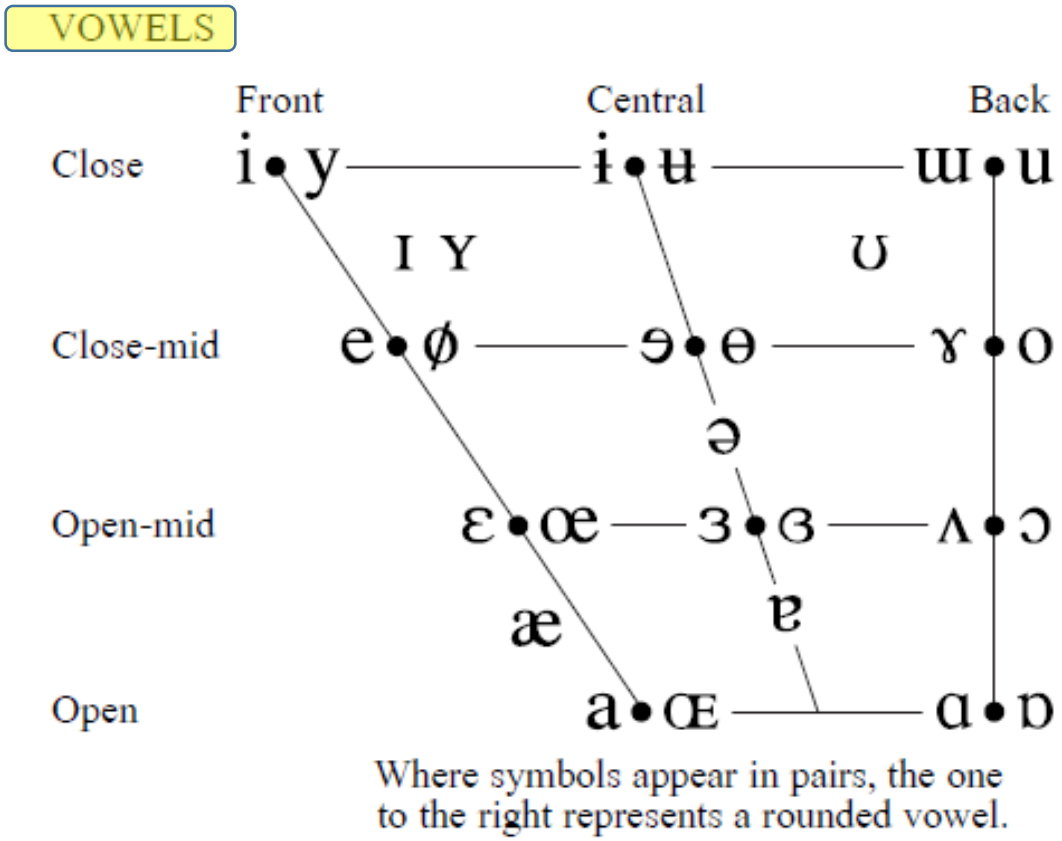
Vowels

- Vowels are commonly described according to the following characteristics:
  - The portion of the tongue that is involved in the articulation: front, central or back.
  - The tongue's position relative to the palate: high, mid or low.
  - The shape of the lips: rounded or unrounded (spread).
  - The length or duration of vocalization: long or short.

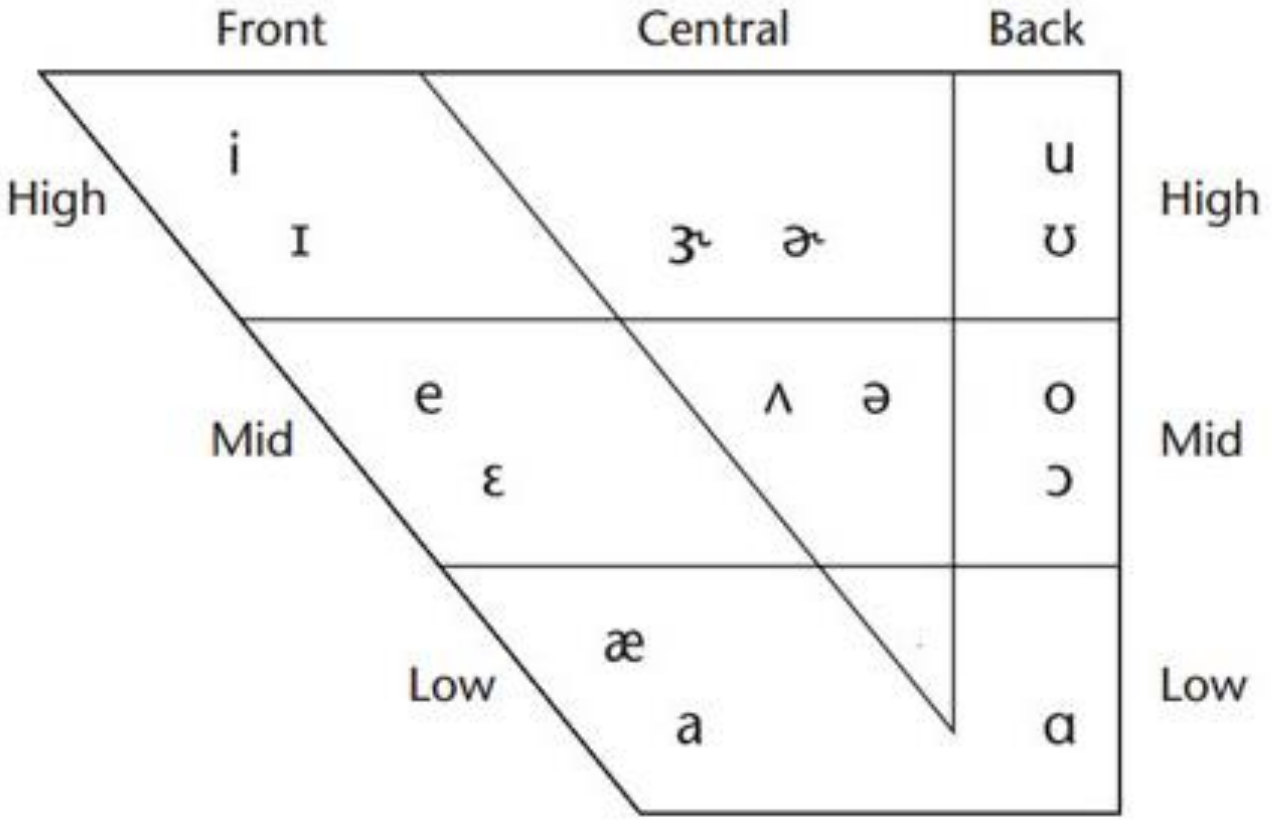
Vowels

- Tongue movement
  - High – Mid – Low /i/ - /e/ - /æ/
  - Front – Central – Back /i/ - /ə/ - /u/
- Lip postures
  - Rounded /u:/
  - Spread /i:/
  - Neutral /ə/

IPA


VOWELS
Front
Central
Back
Close
Close-mid
Open-mid
Open
i y ɨ ʉ ɯ u
ɪ ʏ ʊ
e ø ɘ ɵ ɤ o
ə
ɛ œ ɜ ɞ ʌ ɔ
æ ɐ
a ɶ ɑ ɒ
Where symbols appear in pairs, the one to the right represents a rounded vowel.

IPA



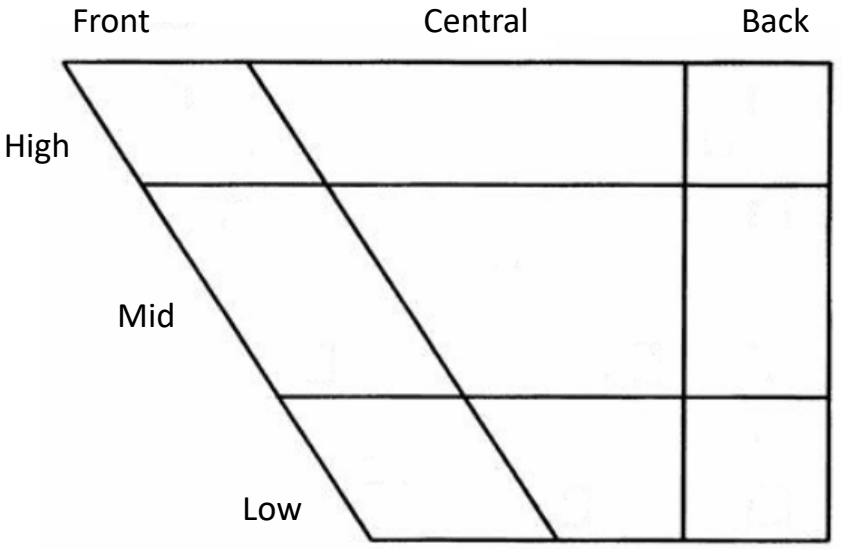
https://rose-medical.com//vowel-sounds.html

Vowels

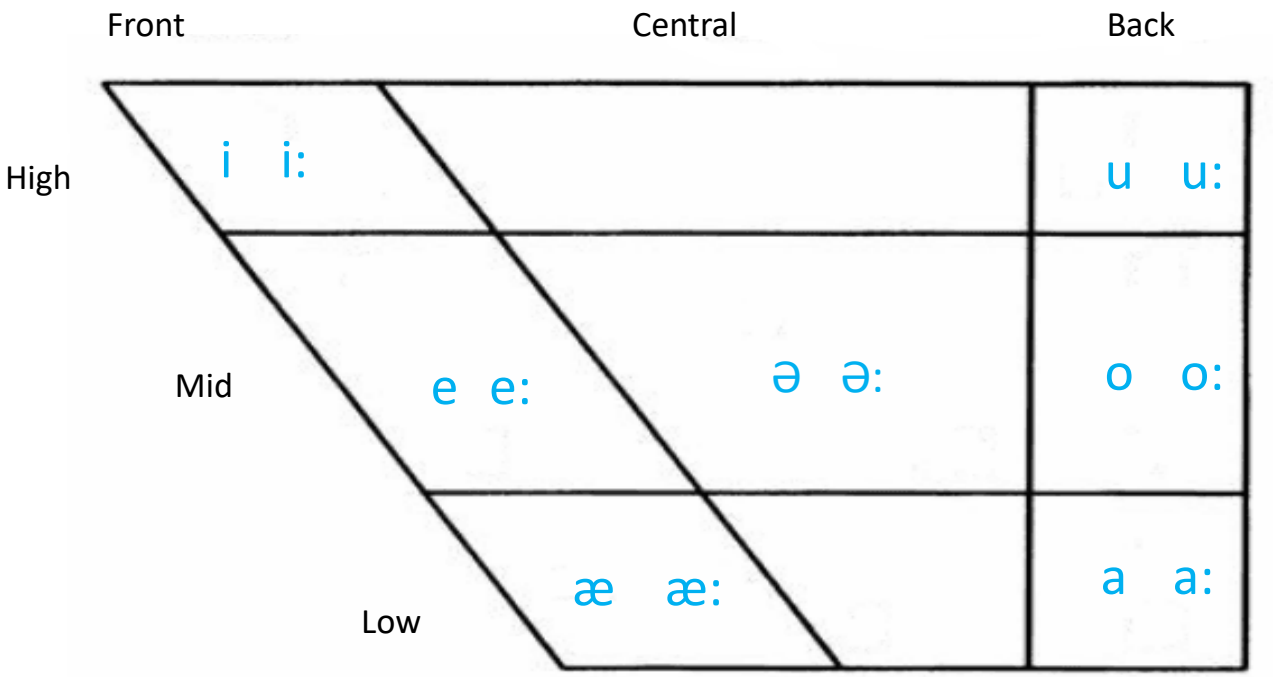
- Identify Sinhala vowel positions

Vowels

Sound	IPA	Sound	IPA
අ	a	උ	u
ආ	a:	ඌ	u:
ඇ	æ	එ	e
ඈ	æ:	ඒ	e:
ඉ	i	ඔ	o
ඊ	i:	ඕ	o:
schwa	ə		ə:



Vowels



Vowels

- Diphthongs
  - Vowels which have 2 vowel qualities

    - English: [ai] → eye /ai/
      [au] → cow /kau/

    - Sinhala: [ai] → අයියා /aiya:/
      [au] → ඖෂධ /auʃədə/
      [æi] → ඇයි /æi/

Types of Phonetics

1. Articulatory Phonetics
2. **Acoustic Phonetics**
3. Auditory/Perceptual phonetics

2. Acoustic Phonetics

- Study of the sound waves generated from the articulated speech
  - study of the **acoustic** characteristics of speech, including an analysis and description of speech in terms of its physical properties, such as frequency, intensity, and duration
- Use of Acoustic Phonetics in NLP
  - Automatic Speech Recognition (ASR)
  - Text to Speech Systems (TTS)

Automatic Speech Recognition (ASR)

- ASR (or STT?)
  - Understanding Speech
    - Train machines to understand human speech
      - Speech Recognition (Identify what is being said)
      - Speaker Recognition (Identify voice/speaker)
    - Acoustic Model
      - Model the acoustic (speech) features of a language
    - Language Model
      - Model the language (grammar) features of a language

Text To Speech (TTS)

- TTS
  - Generating Speech
    - Train machines to read text