Carnegie Mellon University
Language
Technologies
Institute

Articulatory Phonetics

David R. Mortensen

August 31, 2022

Language Technologies Institute
Carnegie Mellon University

What color is this dress? This question shook the Internet to its foundations.

Then came another question:

Laurel? or **Yanny**?

What is the source of these ambiguities? Psychophysics (the study of sensation and perception tells us why. When applied to speech sounds—like the sounds in Laurel/Yanny—this is an aspect of PHONETICS. Phonetics is the study of the physical (including psychophysical) aspects of speech.

1

There are three main subfields of phonetics

Phonetics has **three** subfields, each with an associated field of physics:

- **Articulatory and descriptive phonetics** (biophysics, aerodynamics)
- **Acoustic phonetics** (acoustics)
- **Auditory phonetics** (psychophysics)

Learning objectives

1. Speech sounds can be characterized with a small finite number of dimensions
   - **Place**—where the vocal tract is most constricted
     - **Primary**—tightest constriction
     - **Secondary**—second-tightest constriction
   - **Manner**—what type of constriction is involved
   - **Voicing**—the configuration of the larynx during the constriction
   - **Airstream mechanism**—the means by which air is made to move in order to produce the sound
2. Knowing the basic anatomy of the vocal tract can make these dimensions easier to understand
3. This system has been used to construct a universal transcription system, the International Phonetic Alphabet, that can be used to represent the sounds of any language
4. The IPA is useful in some speech and language technologies

The plan: learn about articulatory phonetics, then acoustic phonetics

We will start with articulatory phonetics:

- Place of articulation
- Manner of articulation
- Vocing
- Airstream mechanism
- Coarticulation
- The International Phonetic Alphabet

In a linguistics class, you might be expected to memorize all or part of the IPA.

In this course, we just want you to know a few core symbols, how the IPA works, and how it relates to HLT. We teach articulatory and descriptive phonetics as a **foundation** for other concepts:

- Vowel acoustics
- Consonant acoustics
- The perception of speech
- Phonetics in speech technologies

ARTICULATORY PHONETICS: the study of the mechanisms by which humans produce speech sounds.
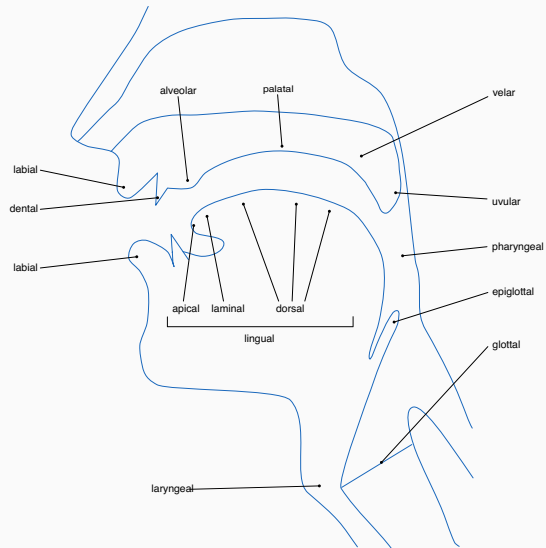
Learning objectives: place of articulation

At the end of this lesson, students will know the following things:

- What is intended by the term PLACE OF ARTICULATION
- What the main places of articulation are for consonants, and what terms are used to describe them
- How place of articulation is characterized for vowels
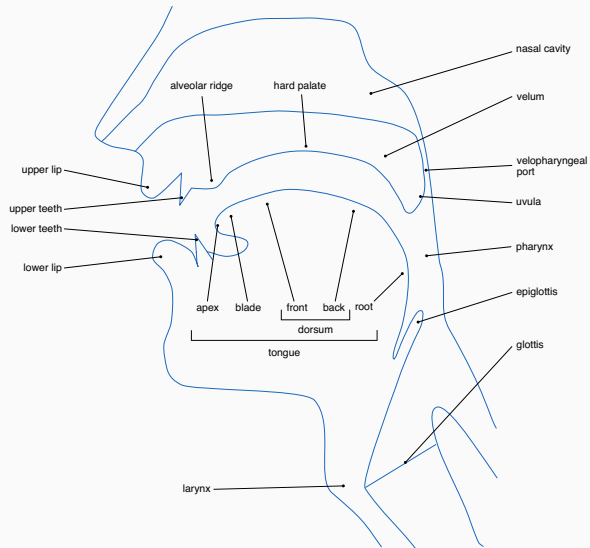
- Labial Constriction made with the lips
  - **Bilabial** Constriction made by bringing the two lips together, like the ⟨b⟩ in *ban*
  - **Labiodental** Constriction by bringing the bottom lip together with the upper teeth, like the ⟨f⟩ in *fine*
- Coronal Constrictions made with the tip or blade of the tongue
  - **Interdental** Constriction made by bringing the front of the tongue into the proximity of the upper teeth, like the ⟨th⟩ in *think*
  - **Dental** Constriction made by bringing the blade (or less often, the apex) of the tongue together with the upper teeth
- **Alveolar** Constriction made by bringing the apex (or less often, the blade) of the tongue together with the alveolar ridge, like the ⟨s⟩ in American English *suspect*
- **Post-alveolar** Constriction made by bringing the apex or blade of the tongue together with region directly behind the alveolar ridge, like the ⟨j⟩ in *judge*
- **Retroflex** Constriction made with tongue tip on hard palate, like ⟨t⟩, ⟨d⟩, ⟨s⟩, ⟨l⟩ and ⟨r⟩ in Indian English

Consonant Places II

- **Dorsal** Constrictions made with the body of the tongue and the palate (the roof of the mouth)
  - **Palatal** Constriction made by bringing the body of the tongue together with the hard palate
  - **Velar** Constriction made by bringing the body of the tongue together with the soft palate (velum), like the ⟨g⟩ in *glamorous*

- **Uvular** Constriction ma1de by bringing the body of the tongue together with the uvula

- **Pharyngeal** Constriction made by bringing the base of the tongue toward the back wall of the pharynx

- **Laryngeal/glottal** Constriction made with the vocal folds, like the ⟨h⟩ in *hellish*

Consonant Place in English as Taught in US Schools

| Place	Examples
BILABIAL	**p**in, **b**in, **m**in
LABIODENTAL	**f**in, **v**im
ALVEOLAR	**t**in, **d**in, **n**ine, **s**in, **z**ip
PALATO-ALVEOLAR	**ch**in, **j**inn, **sh**in, a**z**ure
VELAR	**k**ing, **g**ain
GLOTTAL	**h**it
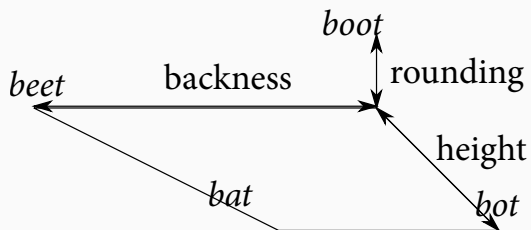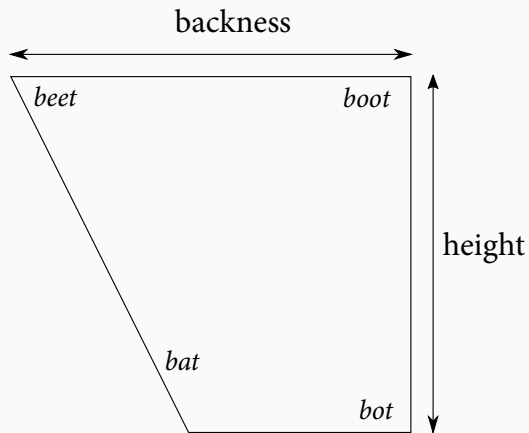
Vowels have three place dimensions: height, backness, and rounding. Rounding is the least important of these.

Vowel place rules of thumb

- A mid central vowel has minimal constriction

- A high front vowel typically has a PALATAL constriction

- A high back vowel typically has a VELAR constriction

- A low back vowel typically has a PHARYNGEAL constriction

13
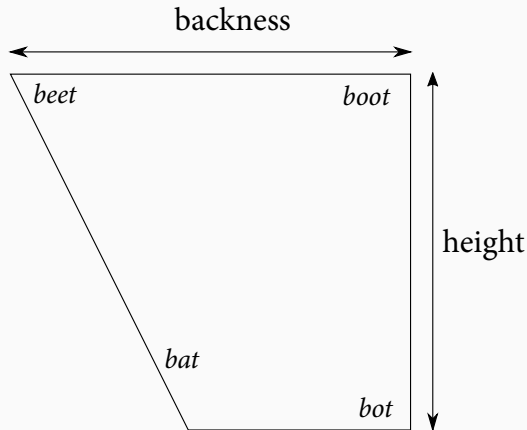
## Learning objectives: manner of articulation, voicing, and airstream mechanisms

In this lesson, you should learn the following things:

- What is intended by the term MANNER OF ARTICULATION
- What the manners of articulation are
- What is intended by voicing/phonation and voice onset time (VOT)
- What the attested airstream mechanisms are for speech sounds

- **plosives** or **oral stops** Characterized by the complete obstruction of the vocal tract and the closure of the velopharyngeal port; like the ⟨p⟩ in *porpoise*

- **nasal stops** or NASALS Characterized by the complete obstruction of the vocal tract but with the velopharyngeal port open; like the ⟨m⟩ in *muddle*

- **trills** Produced with a "loose" closure so that the passage of air produces an oscillation

- **flap** or **tap** essentially a momentary plosive produced when an ACTIVE ARTICULATOR strikes a PASSIVE ARTICULATOR; like the ⟨t⟩ in *writer* or the ⟨d⟩ in *rider*

- **fricatives** Characterized by a tight constriction that produced turbulence when air is blown through it; like the ⟨s⟩ in *slither*

- **lateral fricative** A special kind of fricative in which the opening is on one or both sides of the tongue; common in exotic languages like Hmong and Welsh

- **approximant** Characterized by a loose constriction; includes glides like the ⟨w⟩ in *wand* and other sounds like the ⟨r⟩ in *raven*

- **lateral approximant** A special type of approximant in which there is an opening on one or both sides of the tongue; like the ⟨l⟩ in *leprechaun*
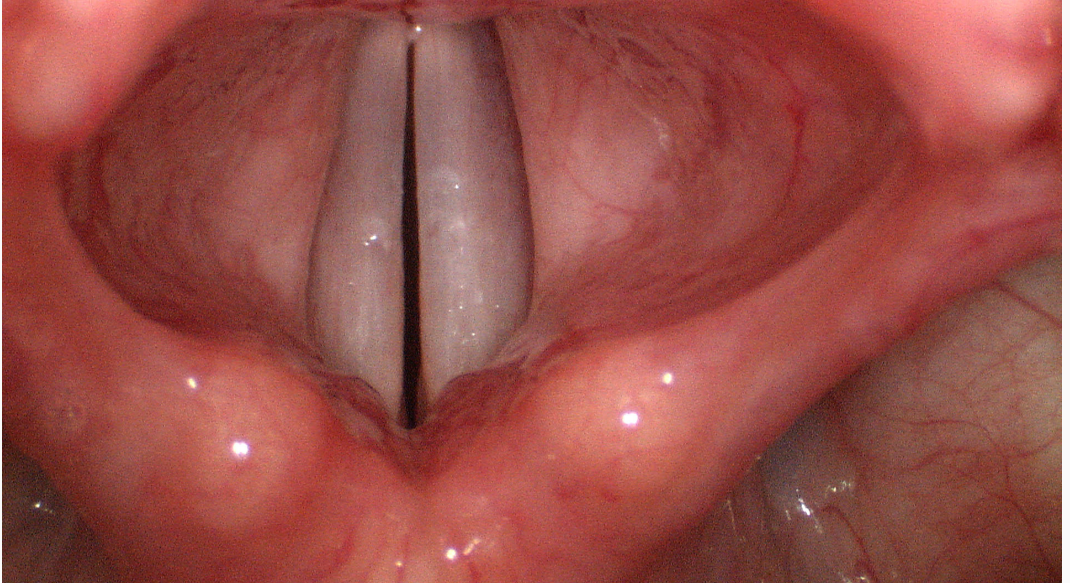
The LARYNX is a cartilaginous structure that contains the GLOTTIS. Part of the larynx protrudes from the throat of many males and is called the ADAM'S APPLE. The parts of the glottis that open, close, and vibrate are called the VOCAL FOLDS. Sometimes, they are also called the VOCAL CORDS.

Voice onset time (VOT) is the time between an articulatory landmark and the onset of phonation

- Usually, when linguistics students are introduced to voicing, they are given examples like English ⟨s⟩ and ⟨z⟩.
- If you put your fingers on your throat while saying *zzzzzz*, you can feel a buzzing that is not there when you produce *ssssss*.
- This buzzing is phonation, and it continues throughout ⟨z⟩; fricatives are either voiced or voiceless.
- However, for many speech sounds, voicing is a matter of degree. This is captured through voice onset time (VOT), the time interval between the release of a constriction and the onset of phonation/voicing:
    - The VOT of Russian [b] is somewhat negative.
    - The VOT of English [b] is slightly negative or near zero.
    - The VOT of English [p] in *spit* is slightly positive or near zero.
    - The VOT of English [pʰ] in *pit* is positive and relatively long.

Airstream mechanisms are the ways in which air is forced through the vocal tract

Making speech sounds requires airflow. There are three different organs that can be used to create airflow (for speech sounds in normal speech) and there are two possible directions of airflow (inward, or INGRESSIVE and outward, or EGRESSIVE):

Organ	Term	Egressive	Ingressive
lungs	**pulmonic**	plosives, etc.*	**
larynx	**glottalic**	ejectives	implosives
velum and tongue	**velaric**	***	clicks

*Most speech sounds in most languages use this airstream mechanism.
**No language uses this airstream mechanism for speech sounds, but in some cultures (e.g. rural Sweden) there are speech styles that are pulmonic ingressive.
***No language uses this airstream mechanism.

Learning objectives: the International Phonetic Alphabet

By the end of this lesson, you should know:

- What the IPA is.
- How the IPA relates to place, manner, voicing, and airstream mechanism
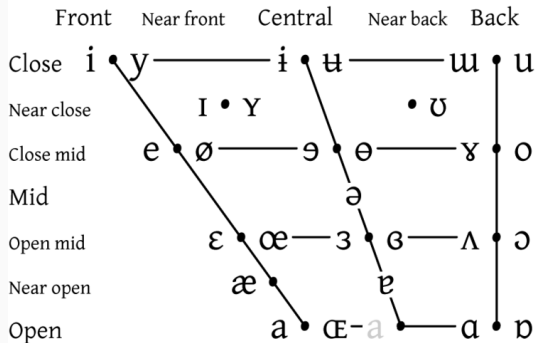- What the IPA is good for

The Pulmonic Consonants

	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		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
Lat. Fric.							ɬ	ɮ														
Approximant				ʋ				ɹ				ɻ		j		ɰ						
Lat. Approx.								l				ɭ		ʎ		L						

VOWELS

Vowels at right & left of bullets are rounded & unrounded.

Why not just use orthography?

ORTHOGRAPHY refers to the conventional writing system used to write a language. Literally, it means "right-writing".

Some orthographies faithfully represent the pronunciation of a language (Spanish, Hungarian, Vietnamese); others (English, Chinese) show only a tenuous relationship between sound and symbol.

Each orthography has its own conventions. The same symbol might be pronounced in different ways and the same sound is often represented by different symbols. In IPA, in contrast, there is a near-perfect mapping between sound and symbol.

Some IPA Symbols for English as Taught in US Schools

[pʰ]	pit	[b]	bit
[p]	spit		
[tʰ]	tick	[d]	dot
[t]	stick		
[kʰ]	cot	[g]	got
[k]	Scot		
[m]	mit	[m]	sim
[n]	nit	[n]	sin
[ŋ]	—	[ŋ]	sing
[l]	lip	[ɫ]	pill
[w]	wack	[j]	yack
[ɹ]	rack		

[f]	fan	[v]	van
[θ]	thigh	[ð]	thy
[s]	sink	[z]	zinc
[ʃ]	assure	[ʒ]	azure
[h]	hat		

[i]	beet			[u]	boot
[ɪ]	bit			[ʊ]	book
[ej]	bait	[ə]	but	[ow]	boat
[ɛ]	bet	[ʌ]	butt	[ɔ]	bought
[æ]	bat			[ɑ]	bot

There are computational means for obtaining IPA transcriptions

For some applications, it may be necessary to have data transcribed in IPA (or some equivalent system). However, for many applications, IPA can be obtained from orthographic text. This is called grapheme-to-phoneme transduction (or G2P). Two types of G2P systems:

- **Rule-based:** Unitran, Epitran. Cover many languages where training data is not available. Only work well when orthographies are "shallow"
- **ML-based:** Based on WFSTs or seq2seq models. Require extensive training data in the form of pronouncing dictionaries, but work well for languages with "deep" orthographies like English and Arabic.

What is the IPA good for?

- What is the IPA good for in linguistics?
  - Documenting languages
  - Annotating speech data for many kinds of analysis
  - Providing a universal reference point for comparisons of sounds within and between languages
  - Characterizing disordered speech
- What is the IPA good for in speech and language technologies?
  - Textually representing pronunciation for both speech recognition and speech synthesis (X-SAMPA, which is an ASCII representation of the IPA, is also used for this)
  - Projecting data from different languages into a single representational space for transfer, identifying names, etc.
  - Others yet to be discovered.

Questions?