Welcome

Intro to the class

Homework: Read class policies on Catalyst site,
Enroll in Moodle, Do Font & Audio Support Quiz

Course Materials

- Catalyst Website: https://catalyst.uw.edu/workspace/valerief/32238
  - Policies, course info, syllabus
  - Calendar, assignments
  - Grade Book, Drop Box, Discussion board
  - Slides, worksheets, practice materials, links
- Moodle (linked from Catalyst): http://wefli4.lang.washington.edu/moodle2/course/view.php?id=51
  - Online homework, quizzes, final exam, transcriptions

Workload

- Readings – 10-30 pages per class
  - Do before class on due date (the day the topic is covered in class)
- HW on Moodle – 1-2 per class, 5-15 questions each
  - These open before we cover the material in class and are due before class on the due date
- Transcriptions on Moodle – 7 total
- Quizzes on Moodle – 3 total, 15 min each
- Labs – 2 total (more details later)
- Project – 2-3 parts (more details soon)
- Final exam on Moodle – during finals week, 2 one-hour parts

First Homework

- Register for Moodle, Enroll in our course
  - Use your UW NetID and email
  - Course enrollment key (also shown on Catalyst): Ladefoged6e
- Do Font & Audio Support Quiz on Moodle
  - Download Charis SIL font
- Read Ladefoged & Johnson ch. 1
- Fill out Language Experiences & Preferences WebQ on our Catalyst site

Basic Phonological Terms

Concepts you should know before beginning this class

If necessary, review *Language Files* ch. 3

(posted to Catalyst)

Phonology Defined

- Phonology is the study of *sound systems* in language. This includes:
  - describing what constitutes “the same sound” versus “different sounds”
  - describing the way in which neighboring sounds influence each other.

Contrastive Sounds

- The first stage in a phonological analysis of a language is accurately describing the set of *contrasts*. Contrastive sounds serve to differentiate words in a language.
- Examples:
  - the first sound in “wake” and “rake” differentiates the two words
  - the last sound in “his” and “hiss” differentiates the two words
  - the middle sound in “leather” and “lever” differentiates the two words

Phonemes and Minimal Pairs

- Sounds that contrast belong to different *phonemes*. Two words that differ by a single contrastive sound (phoneme) are called a *minimal pair*. Larger groups of words that all differ by a single phoneme are called *minimal sets*.
  - Minimal pairs:
    - “wake” / “rake”
    - “his” / “hiss”
    - “leather” / “lever”
  - Minimal set:
    - “by” / “dye” / “guy”

Allophones

- A single phoneme may be pronounced in a variety of ways that don't change the meanings of words. These non-contrastive sounds are called *allophones*.
- You can think of a phoneme as an abstract mental unit comprising the set of all its allophones.
- Example:
  - in English, /t/ is sometimes pronounced with a puff of air (as in "tie") and other times without the puff of air (as in "bat"). These two different pronunciations are two of the allophones of the /t/ phoneme. (Note: the puff of air is called *aspiration*. In some languages, aspiration is a *phonemic contrast* — it can signal a change in word meaning).

Overview of Articulatory Phonetics

Preview of topics to be covered in this class

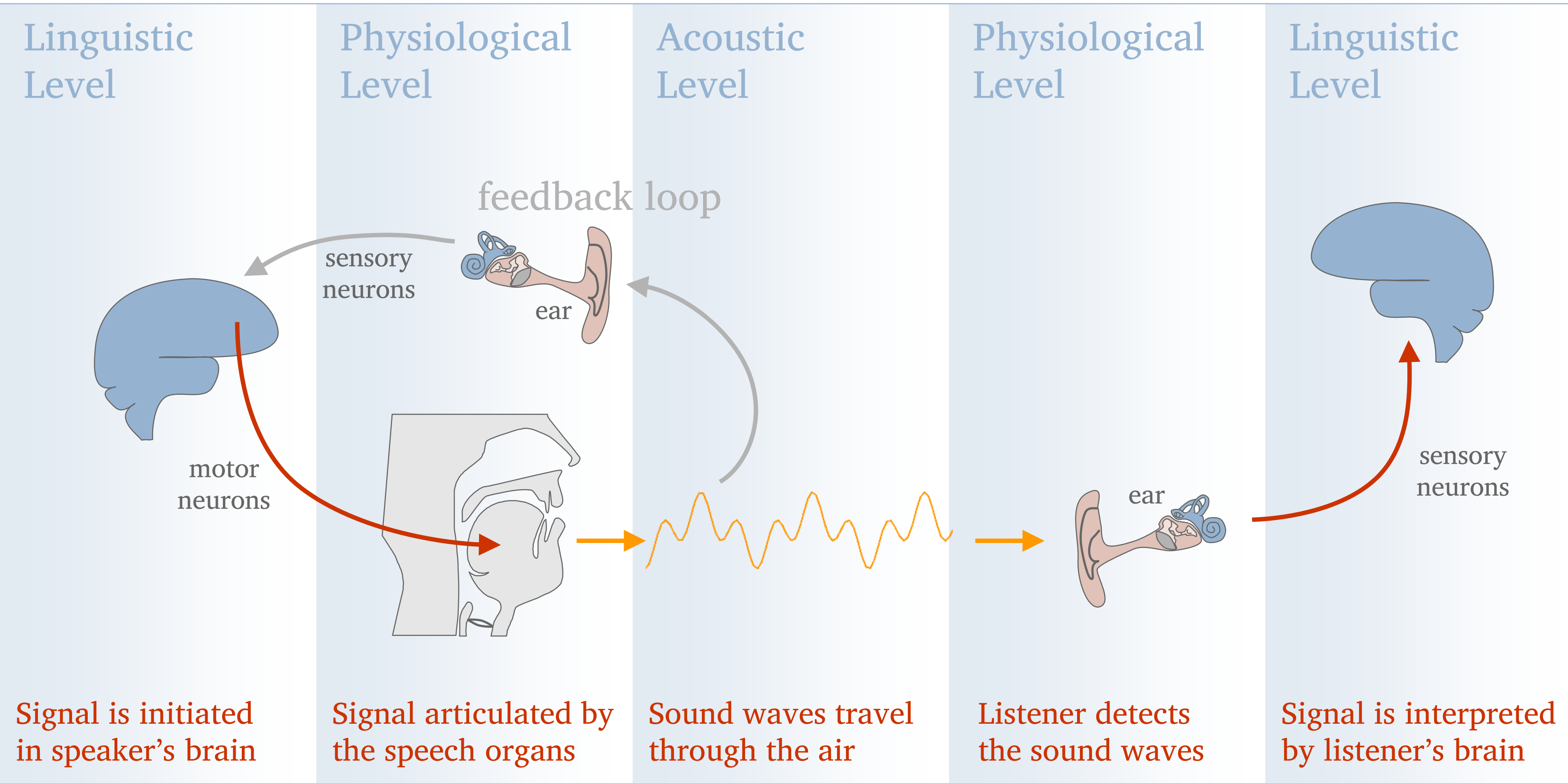
Read: Ladefoged & Johnson ch. 1 (LJ 1)

HW: Vocal Tract Anatomy

Three Kinds of Phonetics

- Articulatory phonetics
  - Describes the way speech sounds are produced anatomically
- Acoustic phonetics
  - Analysis of the sound waves (in air) produced by speech; typically analyses *recorded speech* using specialized software
- Auditory phonetics
  - Concerned mostly with speech perception; typically involves *synthesized speech* in tightly controlled experiments

The Speech Cycle

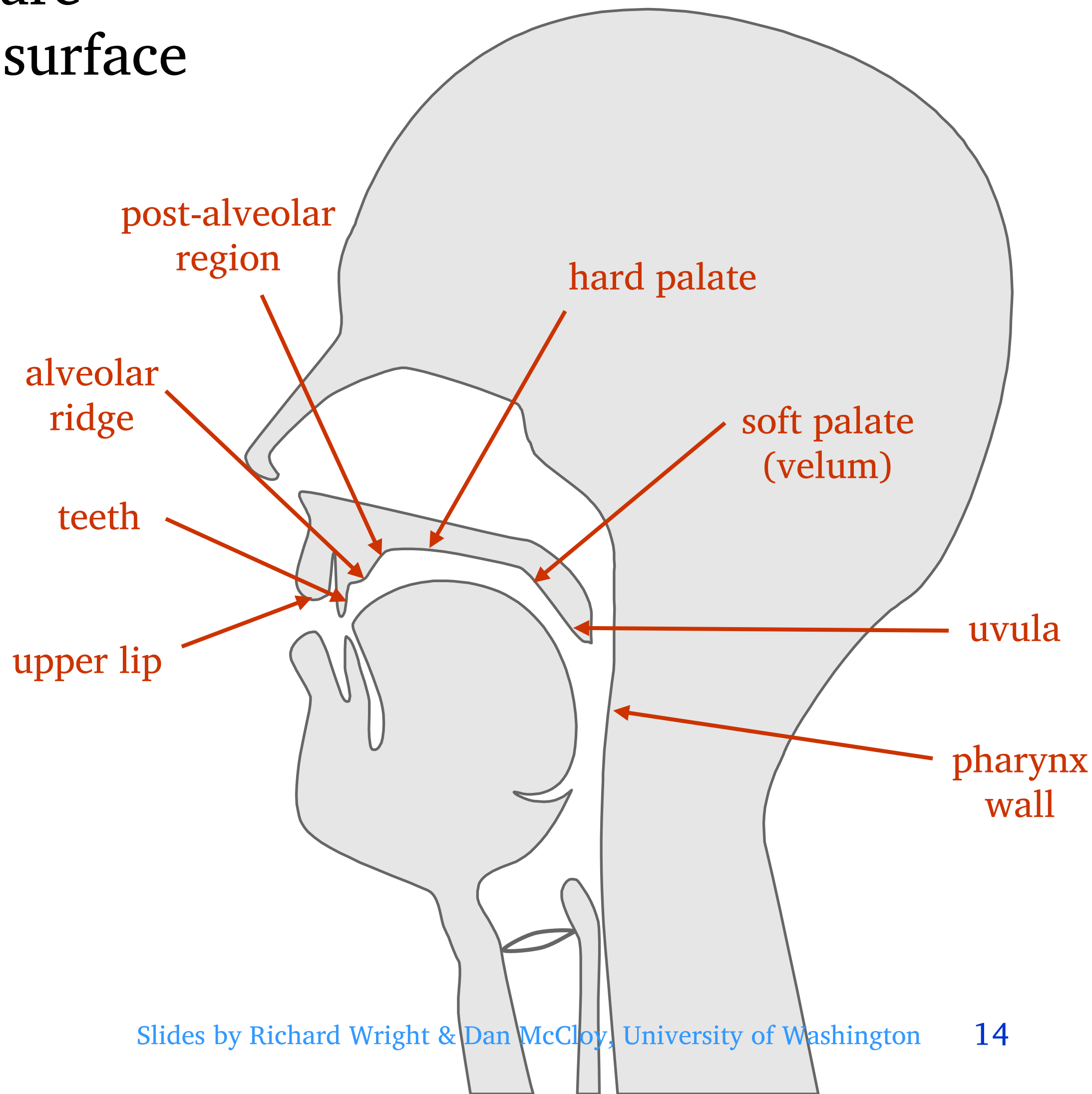


Articulatory Phonetics Defined

- Articulatory phonetics is the study of how speech sounds are produced, and involves detailed and precise descriptions of tongue position, lip rounding, vocal fold vibration, and many other articulatory factors.
- The varieties of sound made during human speech are the result of changes in the *vocal tract*. Most of these changes are the result of movements of the lips, tongue, and jaw, and involve the creation of some degree of *constriction* in the vocal tract.
  - When constriction is created in the vocal tract, it is typically the result of an *active articulator* (like the tongue) moving toward a *passive articulator* (like the hard palate).

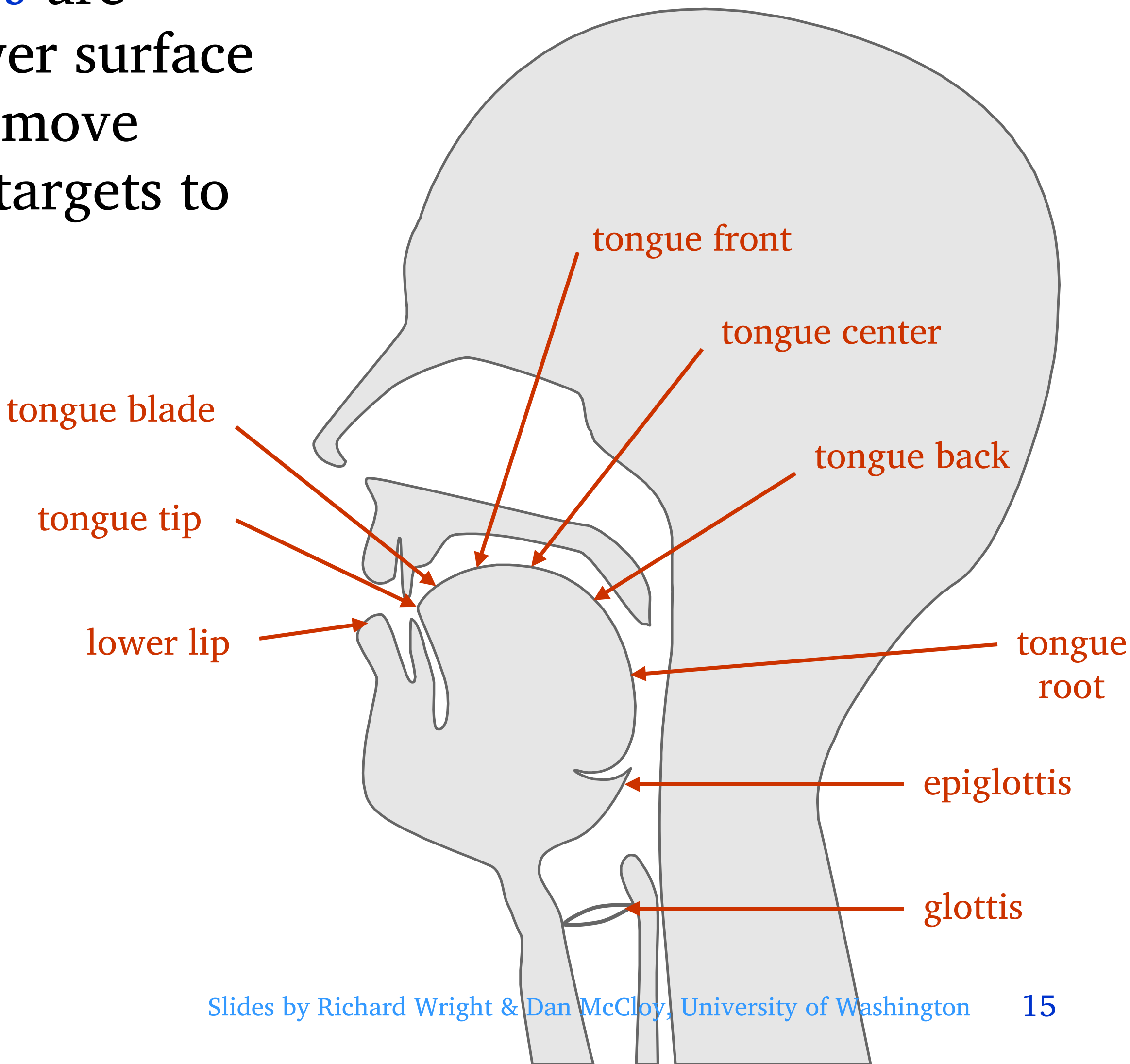
Passive Articulators

- Most *passive articulators* are located along the upper surface of the oral tract.



Active Articulators

- Most *active articulators* are located along the lower surface of the oral tract, and move toward their passive targets to create a constriction.



Describing Articulations

- There are many details that must be specified to give a complete description of a speech articulation:
  - *Place of articulation*
  - *Manner of articulation*
  - *State of the velum* (specified as *oral* vs *nasal*)
  - *State of the glottis* (also called *laryngeal setting* or *voicing*)
  - *Laryngeal timing* (a special measure for stop consonants, usually called *voice onset time* or *VOT*)
  - *Airstream mechanism*
  - *Sound duration*
  - *Pitch*

Describing Articulations: Place of Articulation

- *Place of articulation* describes which active and passive articulators are creating constriction. Examples:
  - *Bilabial* (lower lip to upper lip)
  - *Labiodental* (lower lip to upper incisors)
  - *Dental* (tongue tip/blade to upper incisors)
  - *Alveolar* (tongue tip/blade to alveolar ridge)
  - *Palatoalveolar* (tongue blade to region between alveolar ridge and palate)
  - *Retroflex* (tongue tip or underblade to region between alveolar ridge and palate)
  - *Palatal* (tongue center to palate)
  - *Velar* (tongue back to velum)
  - *Uvular* (tongue back to uvula)
  - *Pharyngeal* (tongue root to pharynx wall)
  - *Epiglottal* (epiglottis to pharynx wall)
  - *Glottal* (vocal folds to each other)

Describing Articulations: Manner of Articulation

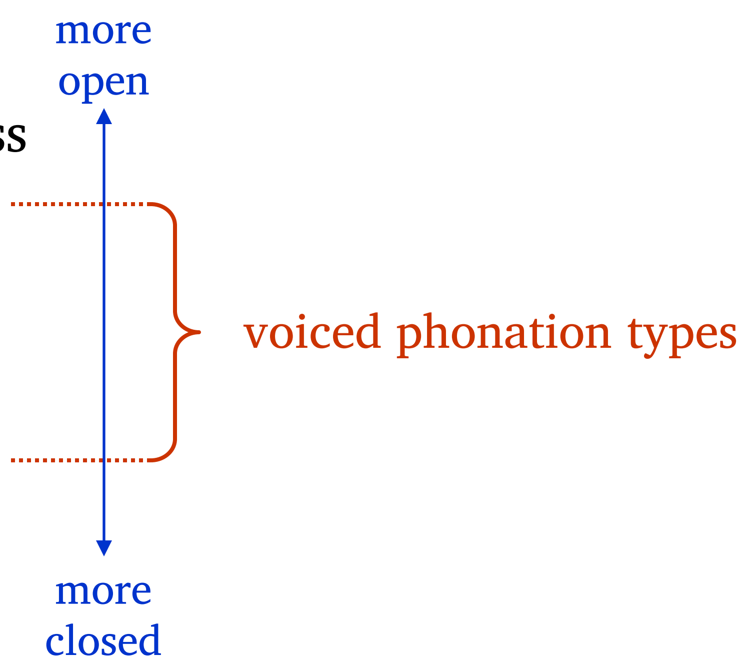
- *Manner of articulation* describes the degree and quality of constriction in the vocal tract. Examples:
  - *Stop*: complete closure of the oral tract airstream
  - *Fricative*: near-complete closure causing turbulence in the airstream
  - *Approximant*: closure broad enough to prevent turbulence
  - *Affricate*: a stop that is released only slightly, causing an immediately following fricative at the same place of articulation
  - *Tap/Flap*: an extremely brief stop closure (the active articulator is flung toward the passive articulator and bounces off it)
  - *Trill*: a rapid series of flap-like closures

Describing Articulations: Velic Closure

- During most articulations, it is possible to have the velum either raised or lowered, thereby closing or opening the velopharyngeal port.
  - Velic closure can affect the quality of both consonants and vowels.
  - The most common distinction is between *oral stops* (velum up/velopharyngeal port closed) and *nasal stops* (velum down/velopharyngeal port open).

Describing Articulations: State of the Glottis

- *State of the glottis* (also called *laryngeal setting*, *phonation type*, or *voicing*) describes the degree of vocal fold closure and the quality of vibration.
  - Voiceless
  - Breathy
  - Modal
  - Creaky
  - Closed



Describing Articulations: Laryngeal Timing

- *Voice onset time* (*VOT*) is a measure of time (typically in milliseconds) between the release of a stop closure and the onset of voicing.
  - When voicing has already begun before the stop is released, the VOT is negative; when voicing begins after the stop is released, the VOT is positive.
  - Positive VOT values lead to *aspiration noise* (a burst of fast-moving air exiting past the stop closure as it is released). Different languages divide the VOT continuum in different ways:
    - English stops: “voiced” = VOT near zero; “voiceless” = long VOT
    - French stops: “voiced” = negative VOT; “voiceless” = short VOT
    - Thai stops: “voiced” = negative VOT; “voiceless” = short VOT; “aspirated” = long VOT

Describing Articulations: Airstream Mechanisms

- All articulations require the movement of air to generate sound. The *airstream mechanism* describes the way in which air is caused to move.
  - The vast majority of speech sounds rely on the *pulmonic egressive* airstream mechanism (air pushed out from the lungs)
  - The *glottalic egressive* airstream mechanism causes air movement *out of* the oral cavity by closing the vocal folds and moving the larynx *up*, thereby *compressing* the air between the larynx and the oral closure. When the closure is released, air moves past the constriction, generating sound. Such sounds are called *ejectives*.

Describing Articulations: Airstream Mechanisms

- All articulations require the movement of air to generate sound. The *airstream mechanism* describes the way in which air is caused to move.
  - The *glottalic ingressive* airstream mechanism causes air movement *into* the oral cavity by closing the vocal folds and moving the larynx *down*, thereby *rarefying* the air between the larynx and the oral closure. When the closure is released, air moves past the constriction, generating sound. Such sounds are called *implosives*.
  - The *velaric ingressive* airstream mechanism causes air movement *into* the oral cavity by making two oral closures and moving the tongue *down*, thereby *rarefying* the air between the two closures (one is always at the velum). When the front closure is released, air moves past the constriction, generating sound. Such sounds are called *clicks*.

Describing Articulations: Duration and Pitch

- Both vowels and consonants can vary in their duration; in some languages the length of a sound can signal *lexical contrast* (i.e., change the meaning of a word).
  - Contrastive consonant length: Italian, Swedish, Finnish, Japanese (ex: Japanese 来た *kita* “came/arrived” vs. 切った *kitta* “cut/sliced”)
  - Contrastive vowel length: Danish, Korean, Swedish, Finnish, Japanese (ex: Croatian *hrom* [hrôm] “lame” vs. *krom* [hrô:m] “chrome”)
- Pitch can signal lexical contrast in some languages (e.g., Mandarin Chinese, Vietnamese, Hausa, Zulu). This is called *lexical tone*. Example:
  - ma 55 (媽) “mother” ma 35 (麻) “hemp”
  - ma 214 (馬) “horse” ma 51 (罵) “scold”
- In other languages, pitch marks other aspects of meaning (e.g., emotion or emphasis). This is called *intonation*.

International Phonetic Alphabet

(Yes, you need to memorize it)

HW: IPA (English consonants – articulation)

HW: IPA (English vowels – articulation)

Phonetic Alphabets

- Each symbol represents exactly one sound/articulation (one-to-one sound-symbol correspondence)
  - Can be used to transcribe any sound of any spoken language
- There are different systems; we will focus on the IPA
  - International Phonetic Alphabet, devised and maintained by the International Phonetic Association
  - See the front and back covers of the Ladefoged & Johnson text

Main IPA Consonant Chart

- Place of articulation, front → back of mouth

more closed

↓

less

	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			

- Manner of articulation
- Voicing: pairs shown in each cell Voiceless Voiced

Main IPA Consonant Chart

	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			

- Note: "Plosive" = pulmonic egressive oral stop
- Velic closure: Nasal stops are *nasal* – air flows out the nose (but is stopped in the mouth); all other consonants are *oral*
- Lateral fricatives/approximants are *lateral* – air flows around the sides of the tongue; all others are *central*

The Consonants of English

	Bilabial	Labio-dental	Dental	Alveolar	Palato-alveolar	Palatal	Velar	Glottal
Stop	p b			t d			k g	ʔ
Nasal	m			n			ŋ	
Tap				ɾ				
Fricative		f v	θ ð	s z	ʃ ʒ			h
Affricate					ʧ ʤ			
Approximant	w			ɹ		j		
Lateral				l				

Other IPA Consonant Charts

- Non-pulmonic stops

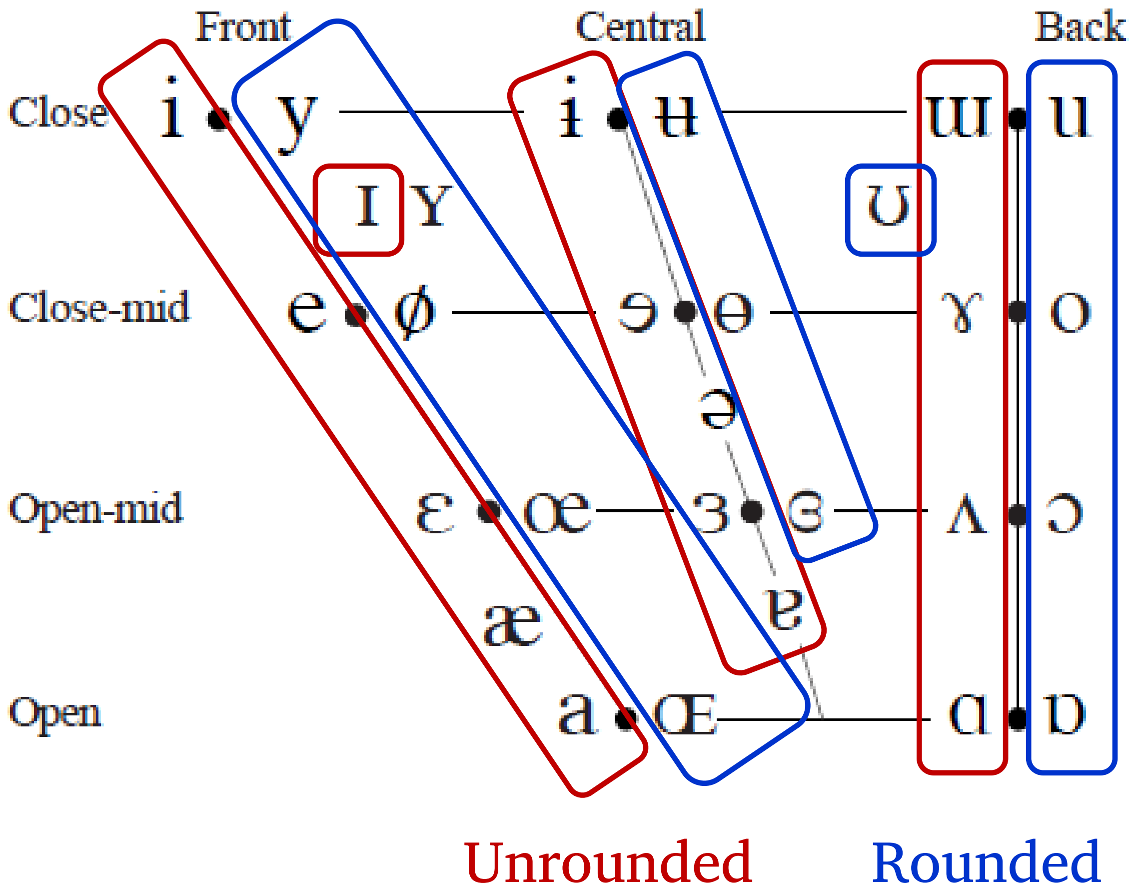
Clicks		Voiced implosives		Ejectives	
ʘ	Bilabial	ɓ	Bilabial	ʼ	Examples:
ǀ	Dental	ɗ	Dental/alveolar	pʼ	Bilabial
ǃ	(Post)alveolar	ʄ	Palatal	tʼ	Dental/alveolar
ǂ	Palatoalveolar	ɠ	Velar	kʼ	Velar
ǁ	Alveolar lateral	ʛ	Uvular	sʼ	Alveolar fricative

- Other

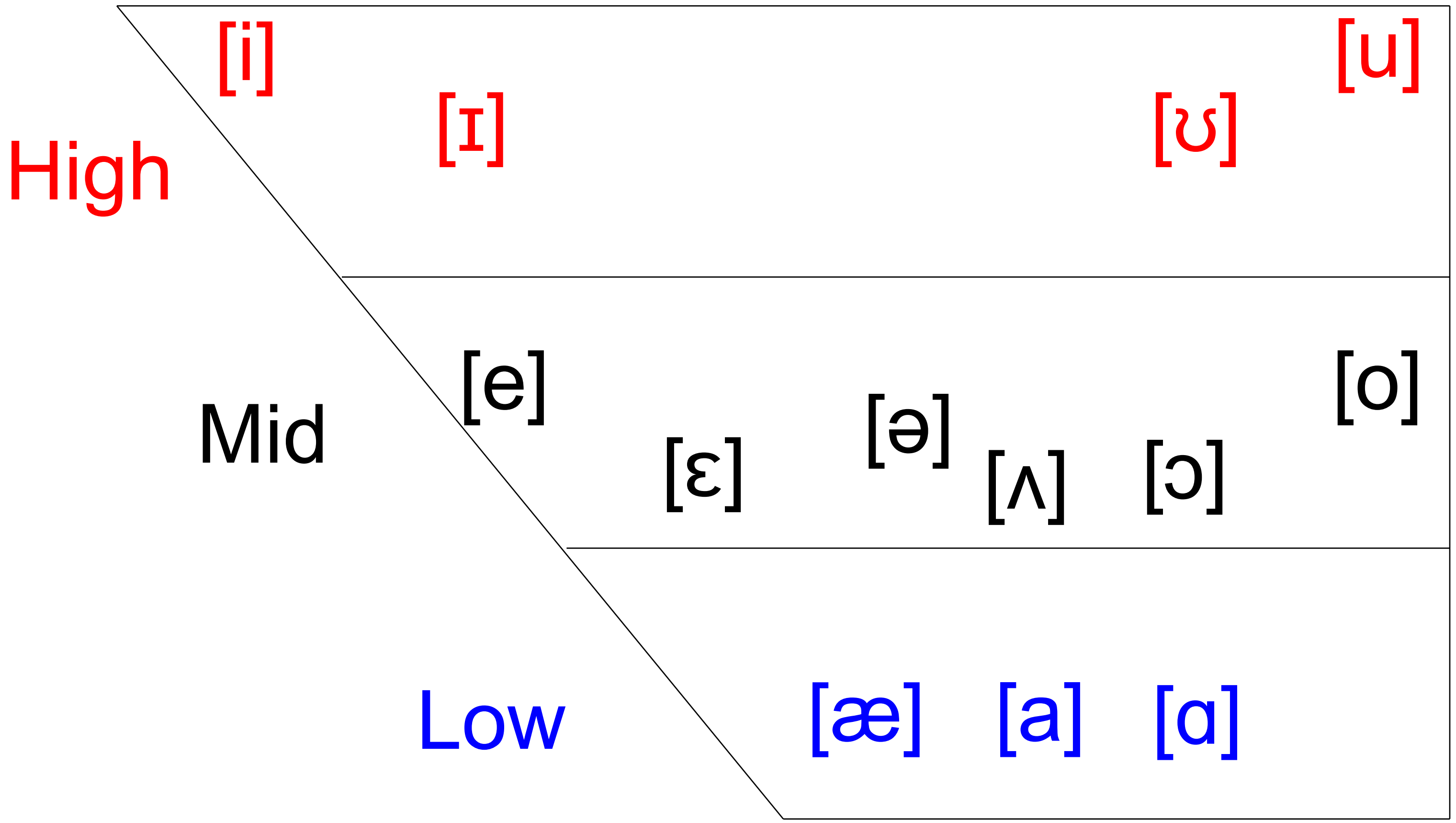
ʍ	Voiceless labial-velar fricative	ɕ ʑ	Alveolo-palatal fricatives
w	Voiced labial-velar approximant	ɺ	Voiced alveolar lateral flap
ɥ	Voiced labial-palatal approximant	ɧ	Simultaneous ʃ and x
ʜ	Voiceless epiglottal fricative		
ʢ	Voiced epiglottal fricative		Affricates and double articulations can be represented by two symbols joined by a tie bar if necessary. k͡p t͡s
ʡ	Epiglottal plosive		

IPA Vowel Chart

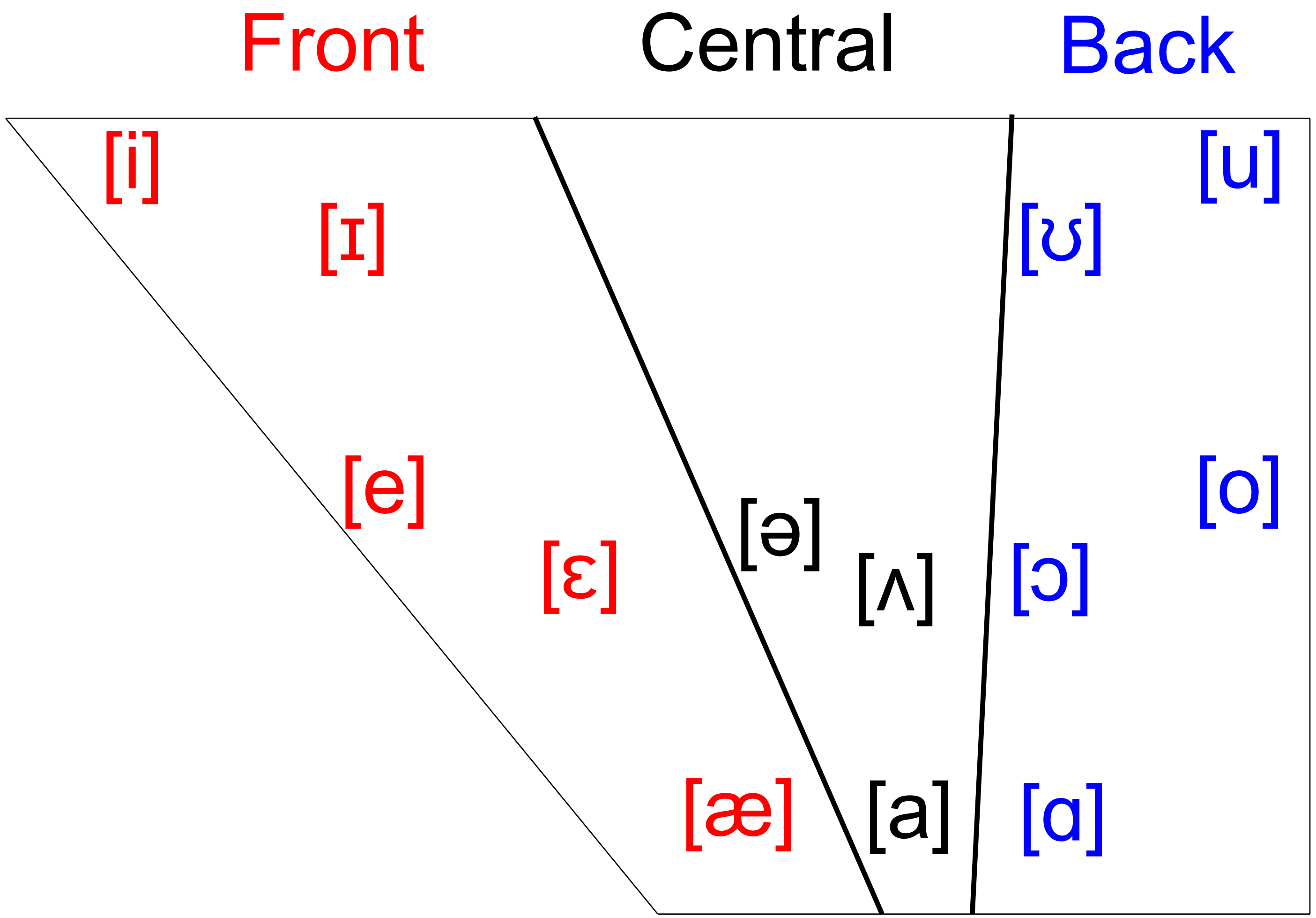
- High
- Mid
- Low



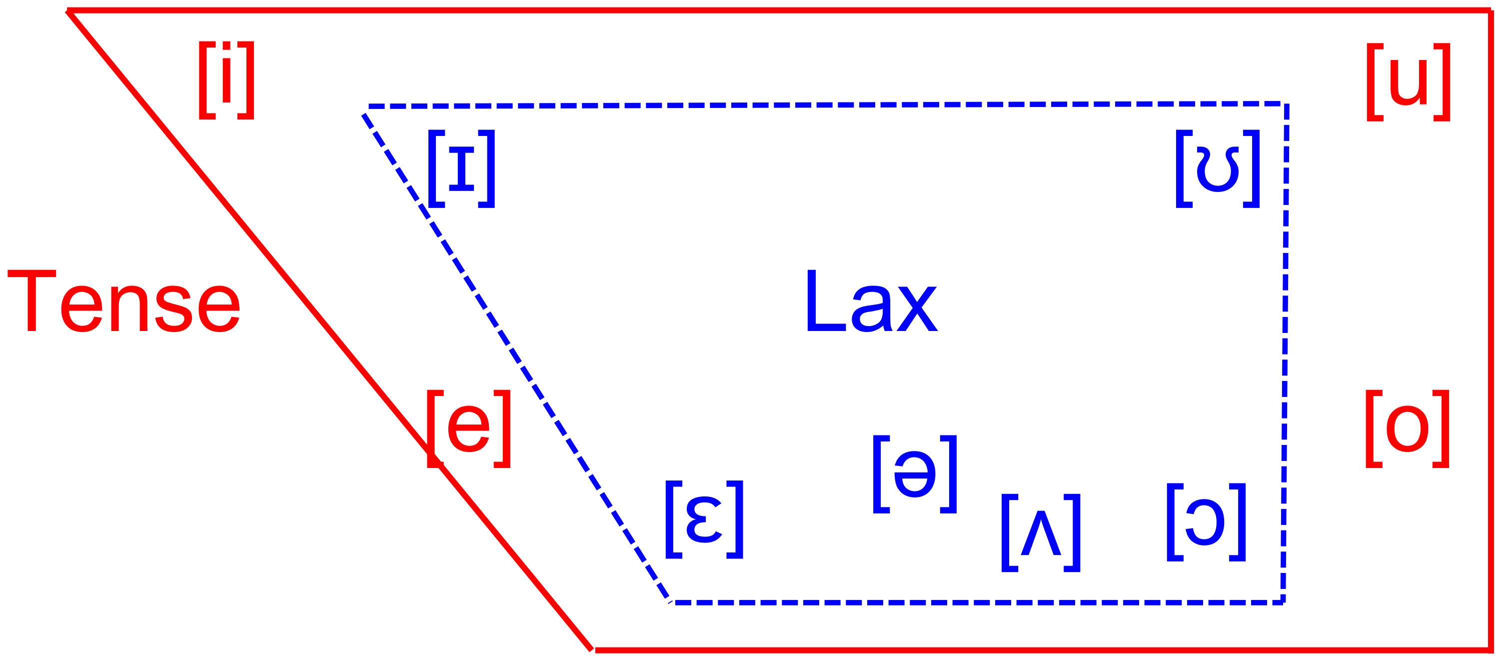
American English Vowels – Height



American English Vowels – Backness

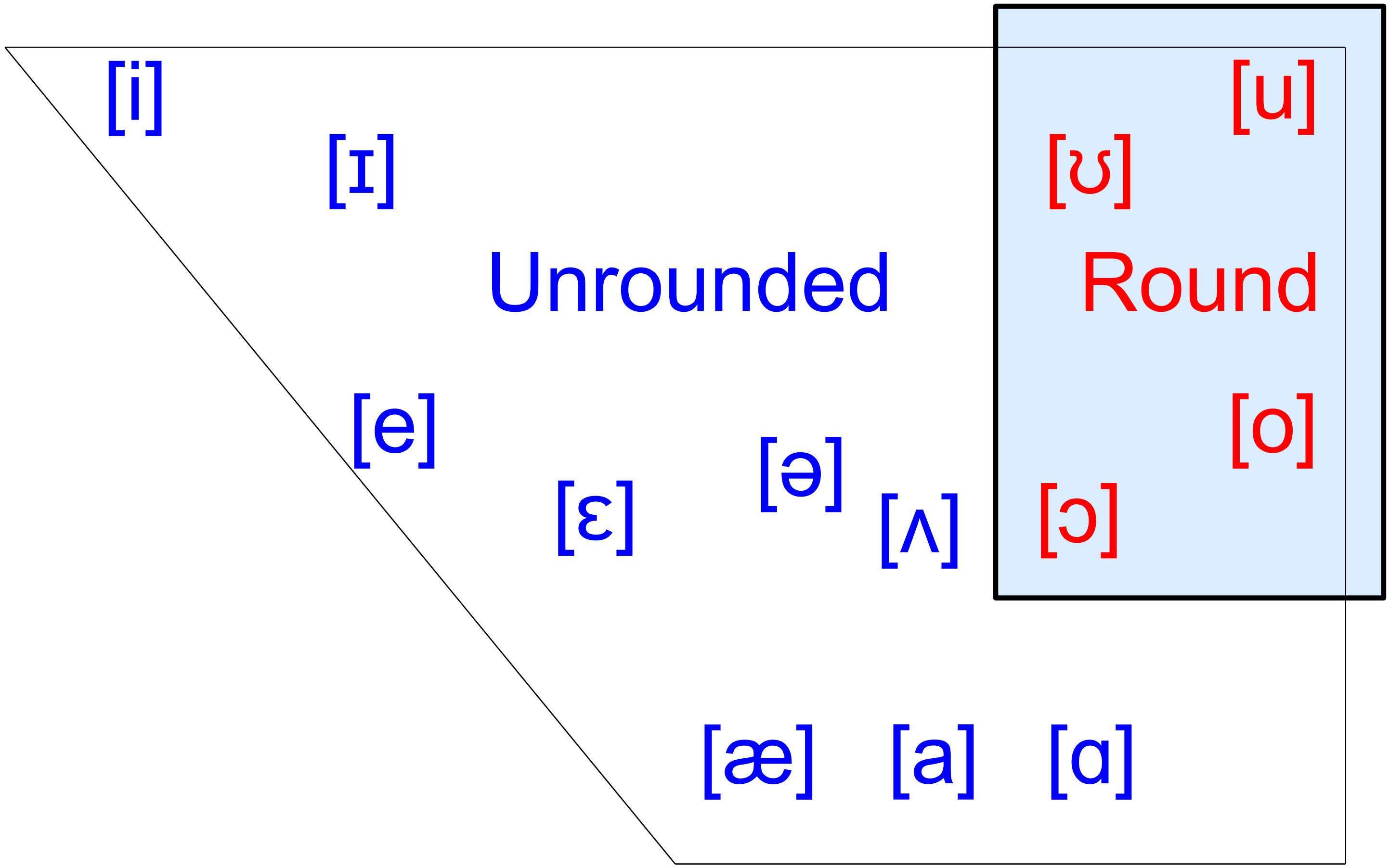


American English Vowels – Tenseness

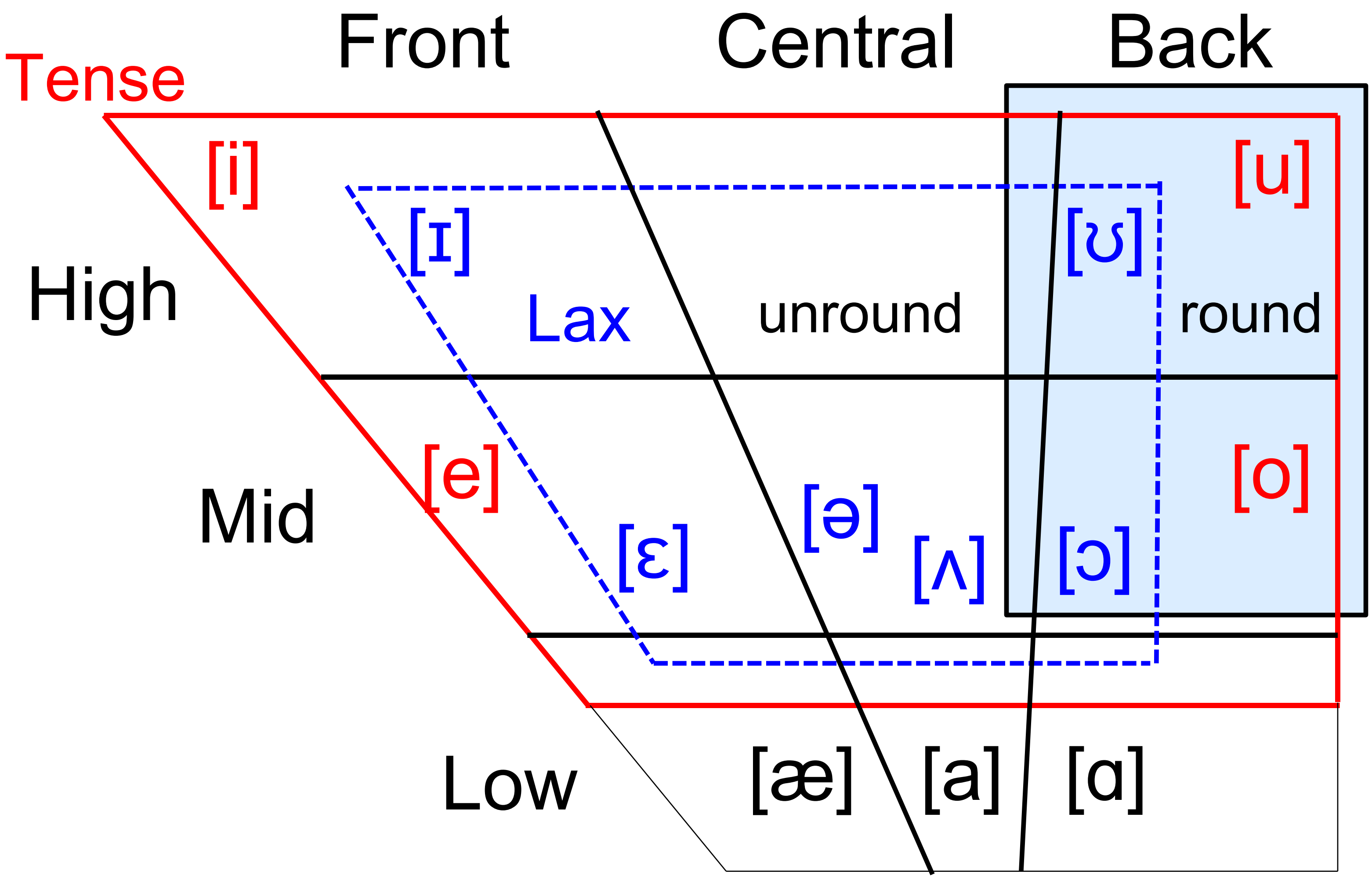


(n/a) → [æ] [a] [ɑ]

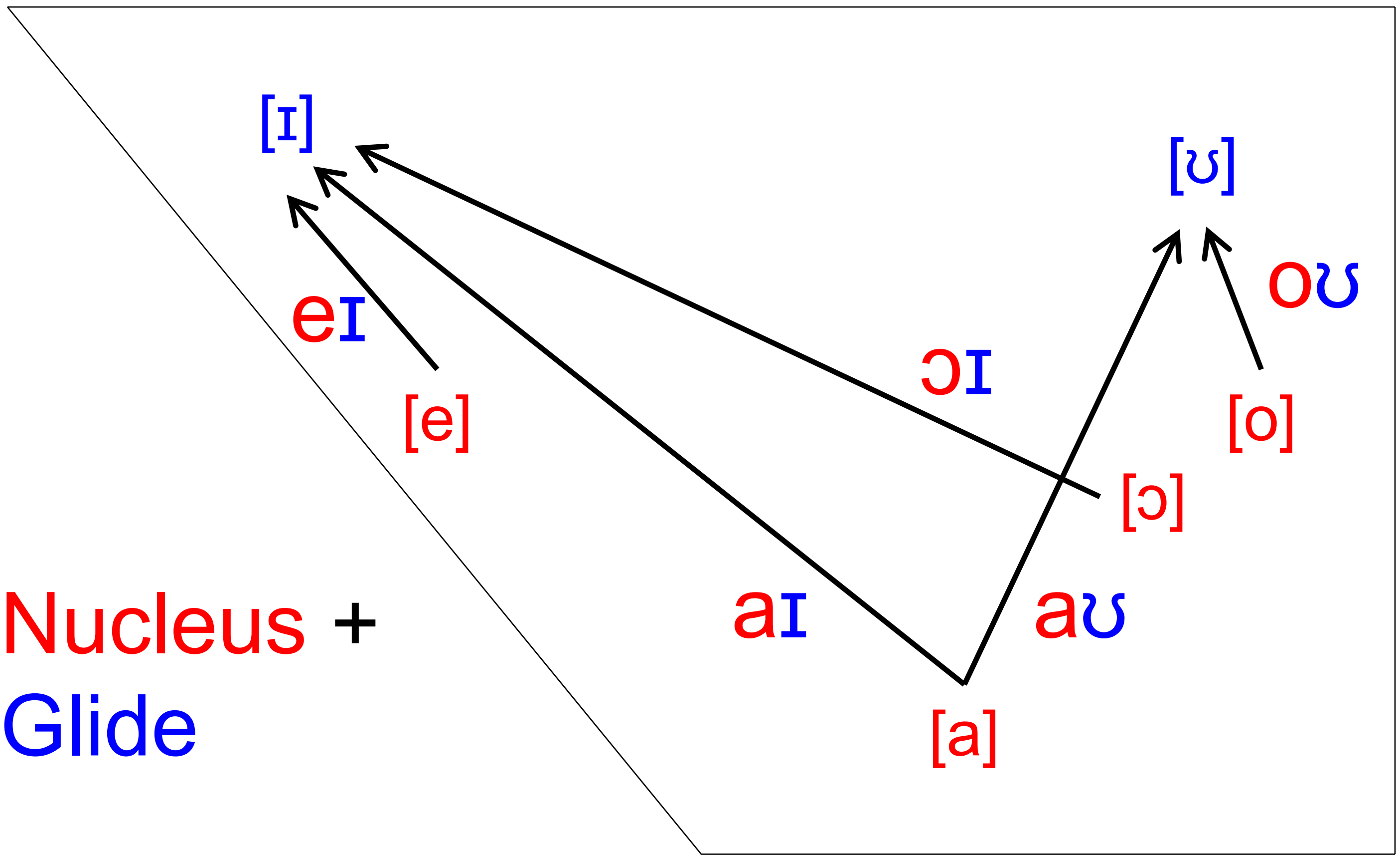
American English Vowels – Rounding



American English Vowels – Summary

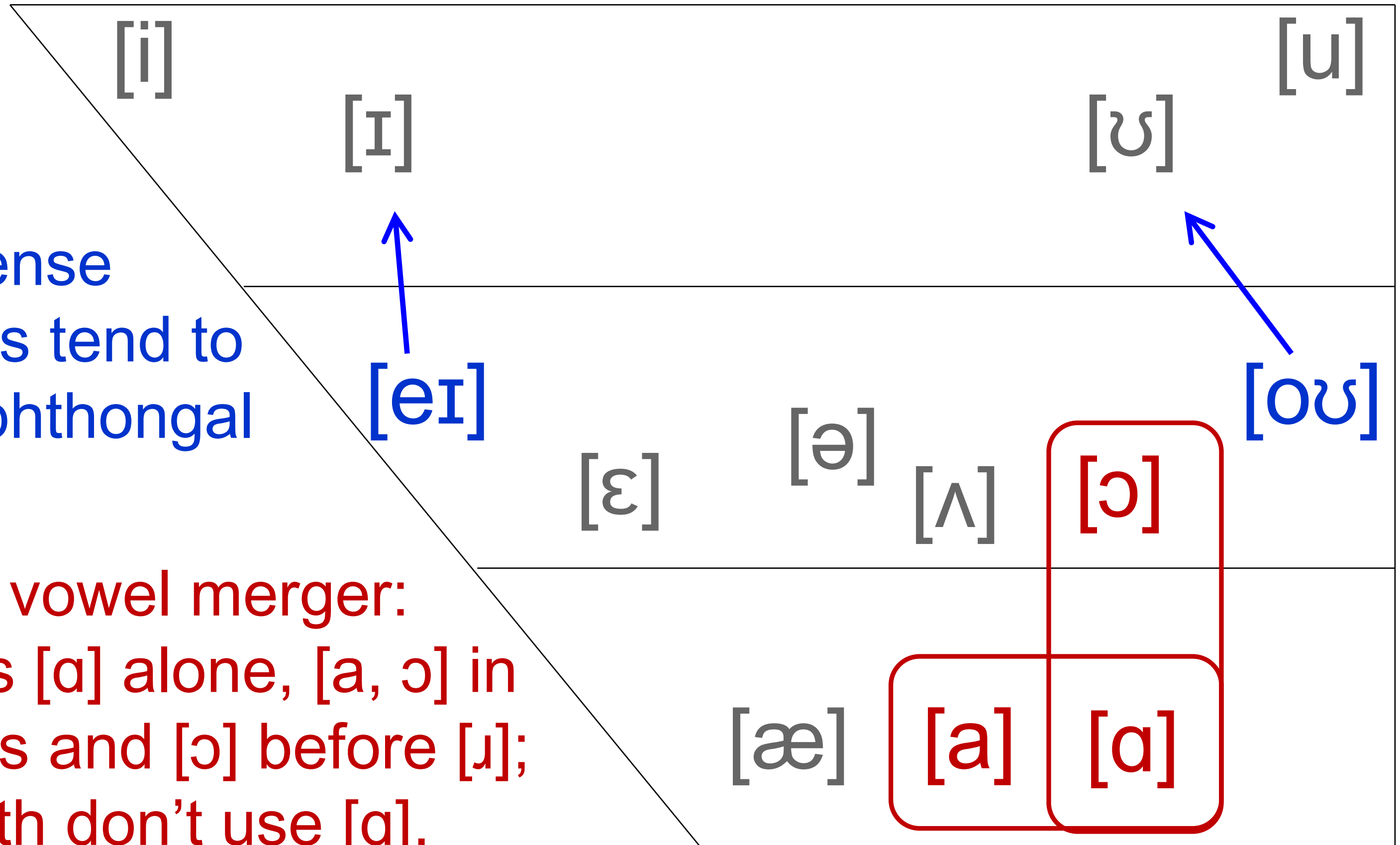


American English Vowels – Diphthongs



Nucleus + Glide

American English Vowels – Notes



Mid tense vowels tend to be diphthongal

Low-back vowel merger:
West uses [ɑ] alone, [a, ɔ] in diphthongs and [ɔ] before [ɹ];
East, South don't use [ɑ], considered an allophone of [a]

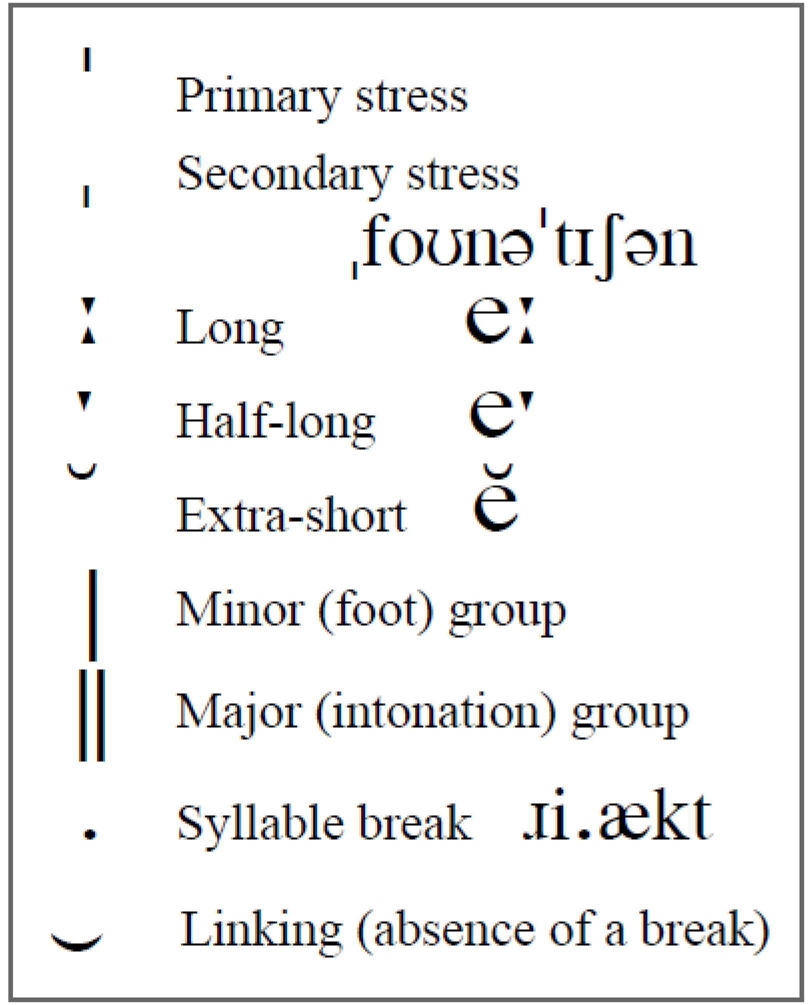
Diacritics

- Placed above or below a symbol to modify its description

̥	Voiceless	n̥ d̥	̤	Breathy voiced	b̤ a̤	̪	Dental	t̪ d̪
̬	Voiced	s̬ t̬	̰	Creaky voiced	b̰ a̰	̺	Apical	t̺ d̺
ʰ	Aspirated	tʰ dʰ	̼	Linguolabial	t̼ d̼	̻	Laminal	t̻ d̻
̹	More rounded	ɔ̹	ʷ	Labialized	tʷ dʷ	̃	Nasalized	ẽ
̜	Less rounded	ɔ̜	ʲ	Palatalized	tʲ dʲ	ⁿ	Nasal release	dⁿ
̟	Advanced	u̟	ˠ	Velarized	tˠ dˠ	ˡ	Lateral release	dˡ
̠	Retracted	e̠	ˤ	Pharyngealized	tˤ dˤ	̚	No audible release	d̚
̈	Centralized	ë	̴	Velarized or pharyngealized	ɫ			
̽	Mid-centralized	e̽	̝	Raised	e̝	(ɹ̝ = voiced alveolar fricative)		
̩	Syllabic	n̩	̞	Lowered	e̞	(β̞ = voiced bilabial approximant)		
̯	Non-syllabic	e̯	̘	Advanced Tongue Root	e̘			
˞	Rhoticity	ɚ a˞	̙	Retracted Tongue Root	e̙			

Suprasegmentals

- Used to indicate stress, length, syllable boundaries, intonation groups



Tone

- Used to indicate lexical tone (when different pitch levels or movements cause a change in the meaning of a word)
- Two systems: diacritics or lines (“tone letters”)

LEVEL			CONTOUR		
e̋ or	˥	Extra high	ě or	˩˥	Rising
é	˦	High	ê	˥˩	Falling
ē	˧	Mid	e᷄	˦˥	High rising
è	˨	Low	e᷅	˩˨	Low rising
ȅ	˩	Extra low	e᷈	˧˦˧	Rising-falling
↓		Downstep	↗		Global rise
↑		Upstep	↘		Global fall

Common Transcription Mistakes

	Correct IPA	Incorrect/non-IPA
palatoalveolar voiceless fricative	[ʃ]	[š]
palatoalveolar voiceless affricate	[ʧ] or [tʃ]	[č]
palatoalveolar voiced affricate	[ʤ] or [dʒ]	[j] [ǰ]
interdental voiceless fricative	[θ]	[th] [t^{h}]
alveolar (central) approximant	[ɹ]	[r]
palatal nasal	[ɲ]	[ñ]
long vowels	[oː]	[ō]
mid lax back unround vowel (stressed syllables)	[ʌ]	[ə]
palatal approximant	[j]	[y]
high front off-glide diphthong	[ɑɪ]	[ɑy] [ɑj]
alveolar apical vowel	[z̩]	[ɿ]
retroflex apical vowel	[ʐ̩]	[ʅ]

Links - Charts

- Interactive IPA charts
  - Ladefoged & Johnson online materials (uses QuickTime): http://www.phonetics.ucla.edu/course/chapter1/chapter1.html
  - Paul Meier Dialect Services (uses Flash, user-friendly): www.paulmeier.com/ipa/charts.html
  - U. of Iowa Flash Animation (includes visual aids): http://www.uiowa.edu/~acadtech/phonetics/
  - U. of Victoria (includes audio explanations, uses QuickTime): web.uvic.ca/ling/resources/ipa/charts/IPAlab/IPAlab.htm
  - Williams College (multiple examples per sound, uses QuickTime): http://wso.williams.edu/~jdowse/ipa.html

Links - Fonts

- IPA character-picker sites
  - Select, copy and paste IPA symbols from here to Moodle or other sites/programs that have trouble with IPA fonts
  - Weston Ruter's: http://westonruter.github.com/ipa-chart/keyboard/
  - Richard Ishida's: http://rishida.net/scripts/pickers/ipa/
- Download IPA font Charis SIL: http://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&id=CharisSIL_download

Links - Keyboards

- Keyboard layouts configured for IPA Unicode fonts:
  - Google "IPA keyboard layout unicode" plus your operating system
  - How to set your keyboard once you've installed a layout (for Windows 2000 or later): http://www.rejc2.co.uk/ipakeyboard/
  - More complete layout for Windows: http://www.phon.ucl.ac.uk/resource/phonetics/.
    - Quirk/Tip: uninstall Charis and Doulos fonts before installing the keyboard layout, then reinstall them afterward. If you forget to uninstall them, the keyboard installation process will pop up an error window for each font. Click Ignore. When it finishes installing, go back and reinstall Charis and Doulos (or the fonts will cause quirks like permanent Italics in Word).

Acoustic Theory

Physical properties of sound

Read: LJ 8 up to "Acoustic Analysis" (p. 193)

Begin HW: Acoustic Theory

Pressure and Waves

- When objects move through a fluid medium such as air, they create localized changes in the *pressure* of the fluid.
  - Note: a *fluid* is anything that flows, and is not the same as a *liquid*.
- In a *homogeneous fluid* (one that has the same properties throughout), pressure changes begin at the source of movement and propagate outward spherically.
- When objects are *vibrating*, the resulting pressure changes are *periodic*, meaning the changes fluctuate in a pattern that repeats over time. Such repeating patterns are called *waves*.

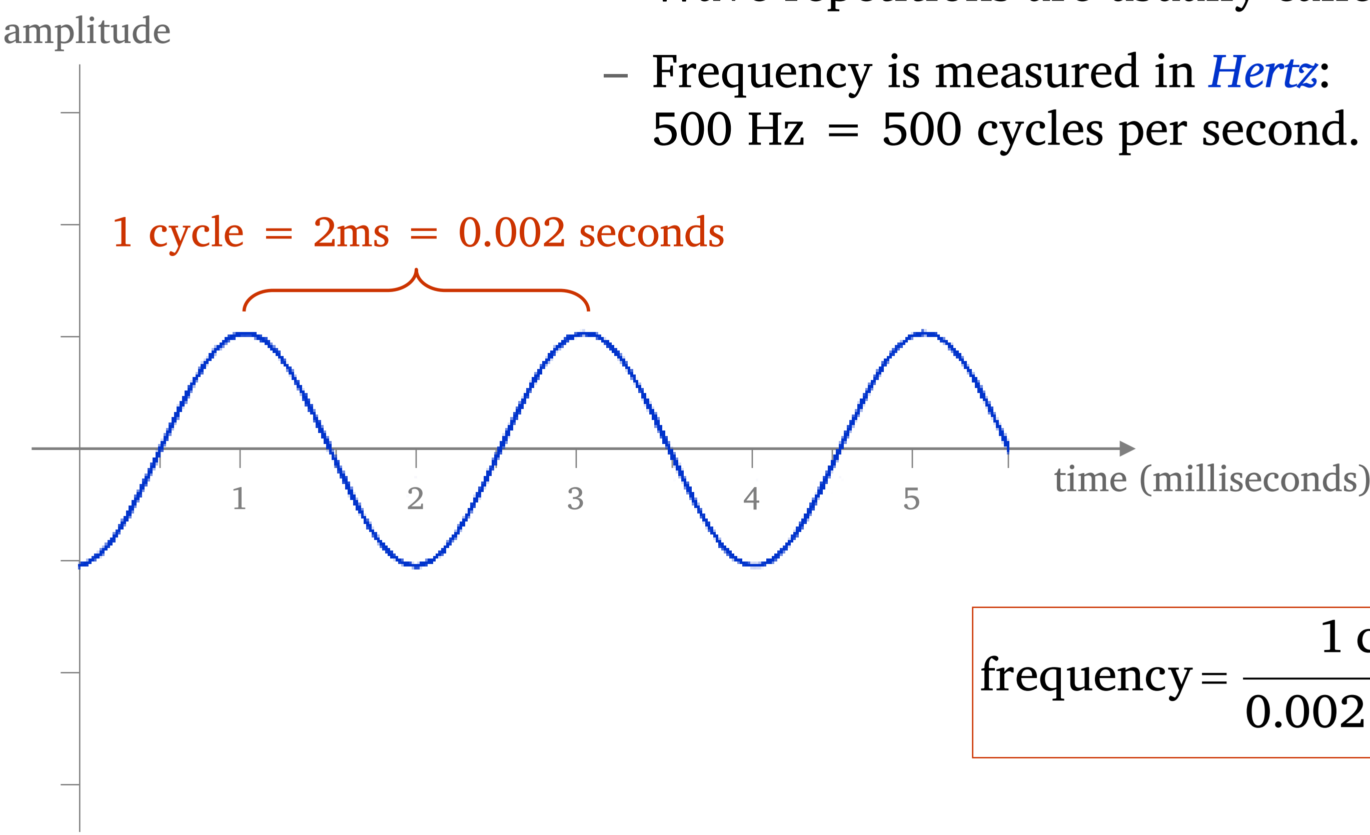
What is Sound?

- *Sound* is the perceptual response to pressure fluctuations in the atmosphere.
  - Technically, *sound* does not exist in the atmosphere; *pressure waves* do. Sound exists in the awareness of the hearer. In practice, however, it is common to talk about *sound waves* and *sound pressure*.

- http://www.youtube.com/watch?v=Si-OYX20FRs
- http://www.youtube.com/watch?v=dbeK1fg1Rew

Frequency

- The *frequency* of a wave is a measure of how many times it repeats in a given amount of time.
  - Wave repetitions are usually called *cycles*.
  - Frequency is measured in *Hertz*:
    500 Hz = 500 cycles per second.



$$\text{frequency} = \frac{1\ \text{cycle}}{0.002\ \text{seconds}} = 500\text{Hz}$$

Frequency Sensitivity

- Different species respond to different frequencies of pressure fluctuations. Young healthy humans typically are sensitive to frequencies between 20 Hz and 20,000 Hz.
  - Pressure fluctuations within this range will register with humans as sound.
  - Speech sounds only go up to 10,000 Hz.
  - Pressure fluctuations outside the 20-20,000 Hz range will not be heard by humans (though vibrations in the air at frequencies lower than 20 Hz may sometimes be *felt* with the skin).

Frequency Sensitivity and Pitch Perception

- Human sensitivity to pressure wave frequency is *nonlinear* — sounds in the middle ranges are heard more easily than very high or very low frequency sounds.
- The human percept of frequency is called *pitch*. Because of the nonlinear nature of frequency sensitivity, frequencies that are measured to be equally spaced may not sound as though they are equally spaced pitches.

Amplitude

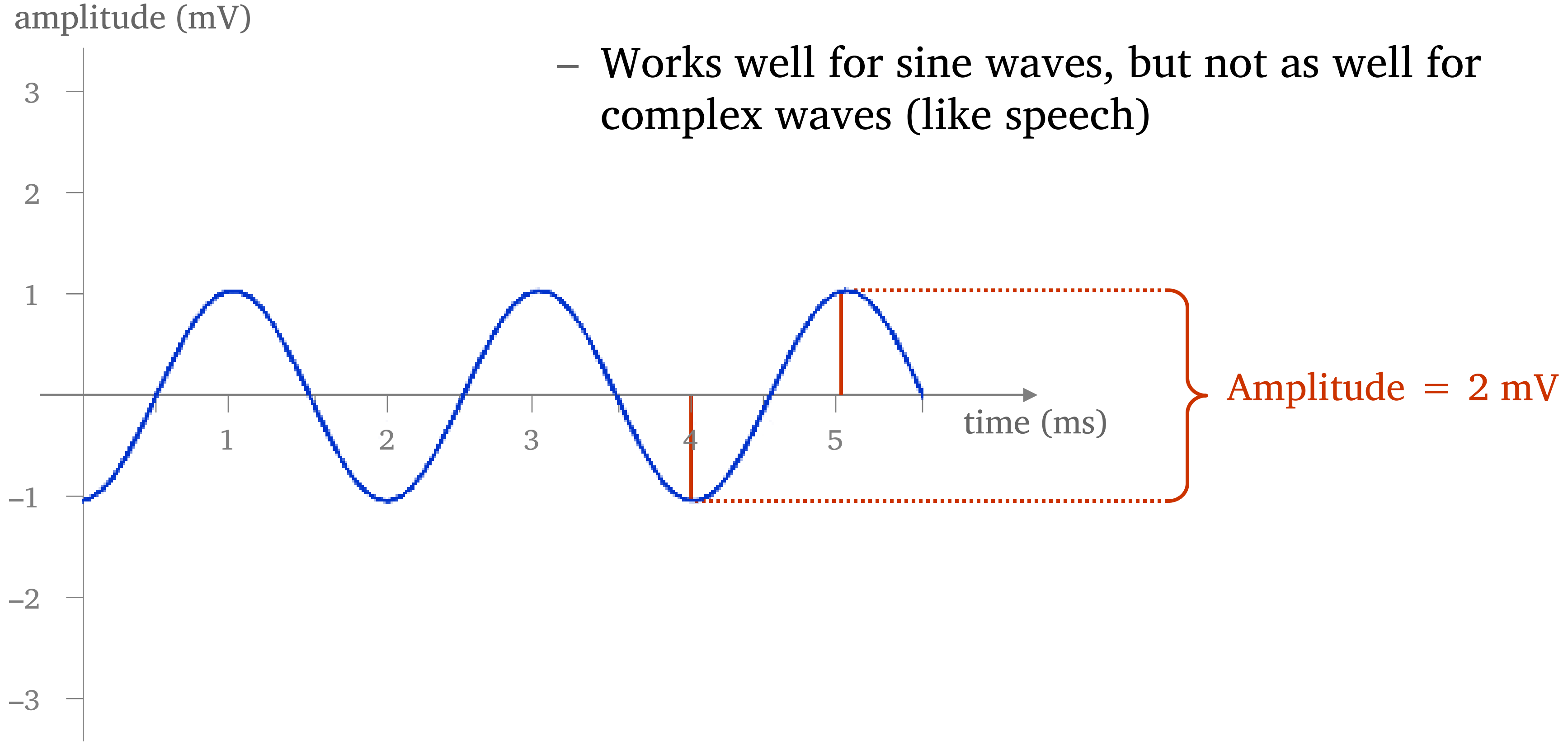
- Speech causes pressure fluctuations in the air. Air molecules get pushed together (*compression*), and then they spring apart (*rarefaction*), attempting to return to a state of rest.
- *Amplitude* is a measure of the magnitude of the pressure fluctuation (i.e., the difference between the wave pressure and the pressure of the medium at rest). For waves in air this is also called *sound pressure*.
  - The standard units for sound pressure are *Pascals* (Pa).
  - When sound waves are converted to electrical signals by a microphone, the amplitude is measured as difference in electrical potential (i.e., *voltage*; the unit is *Volts*).

Measuring Amplitude

- There are two ways that amplitude is usually calculated:
  - *Peak-to-peak*: the vertical difference between the highest point and lowest point of the wave.
  - *Root Mean Square (RMS)*: the average absolute distance from zero of the wave over the course of one cycle.

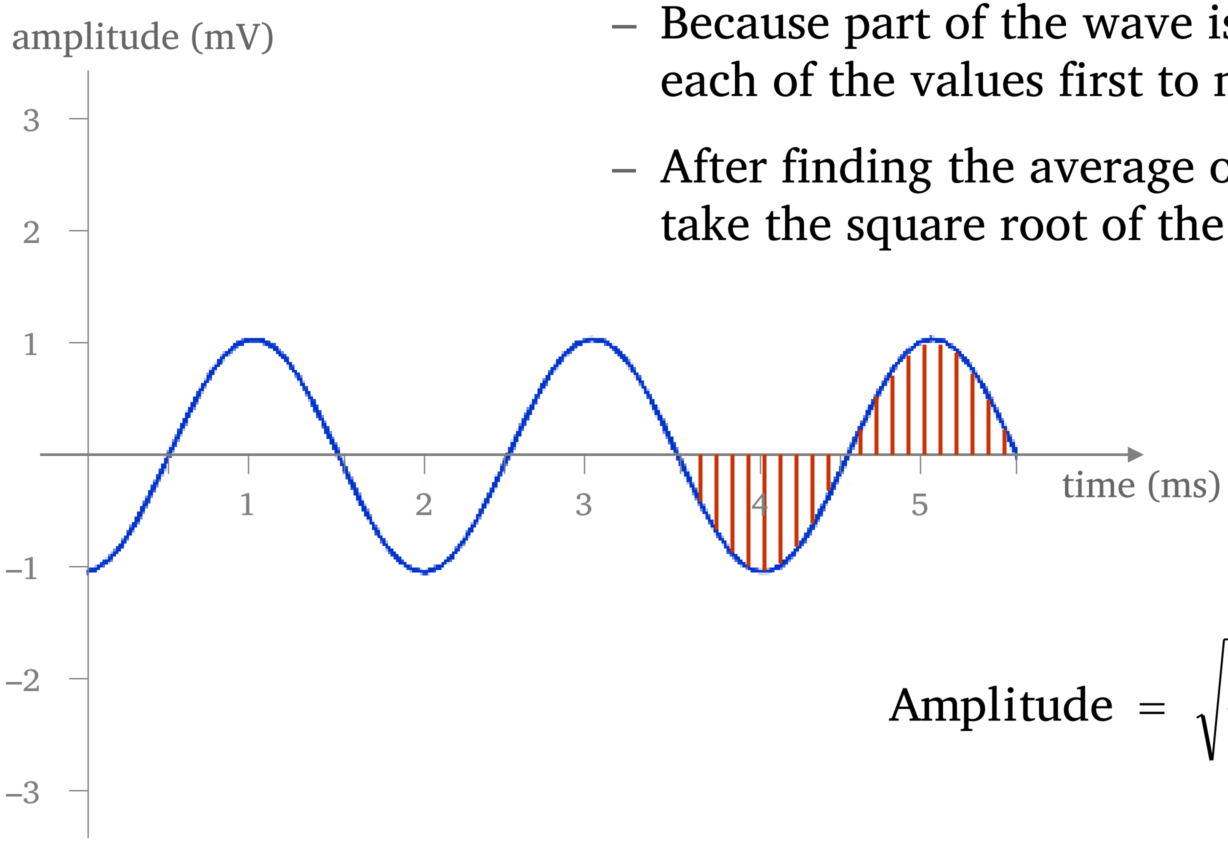
Peak-to-Peak Amplitude

- The peak-to-peak amplitude is calculated by finding the difference between the highest and lowest values of the wave.
  - Works well for sine waves, but not as well for complex waves (like speech)



Root Mean Square (RMS) Amplitude

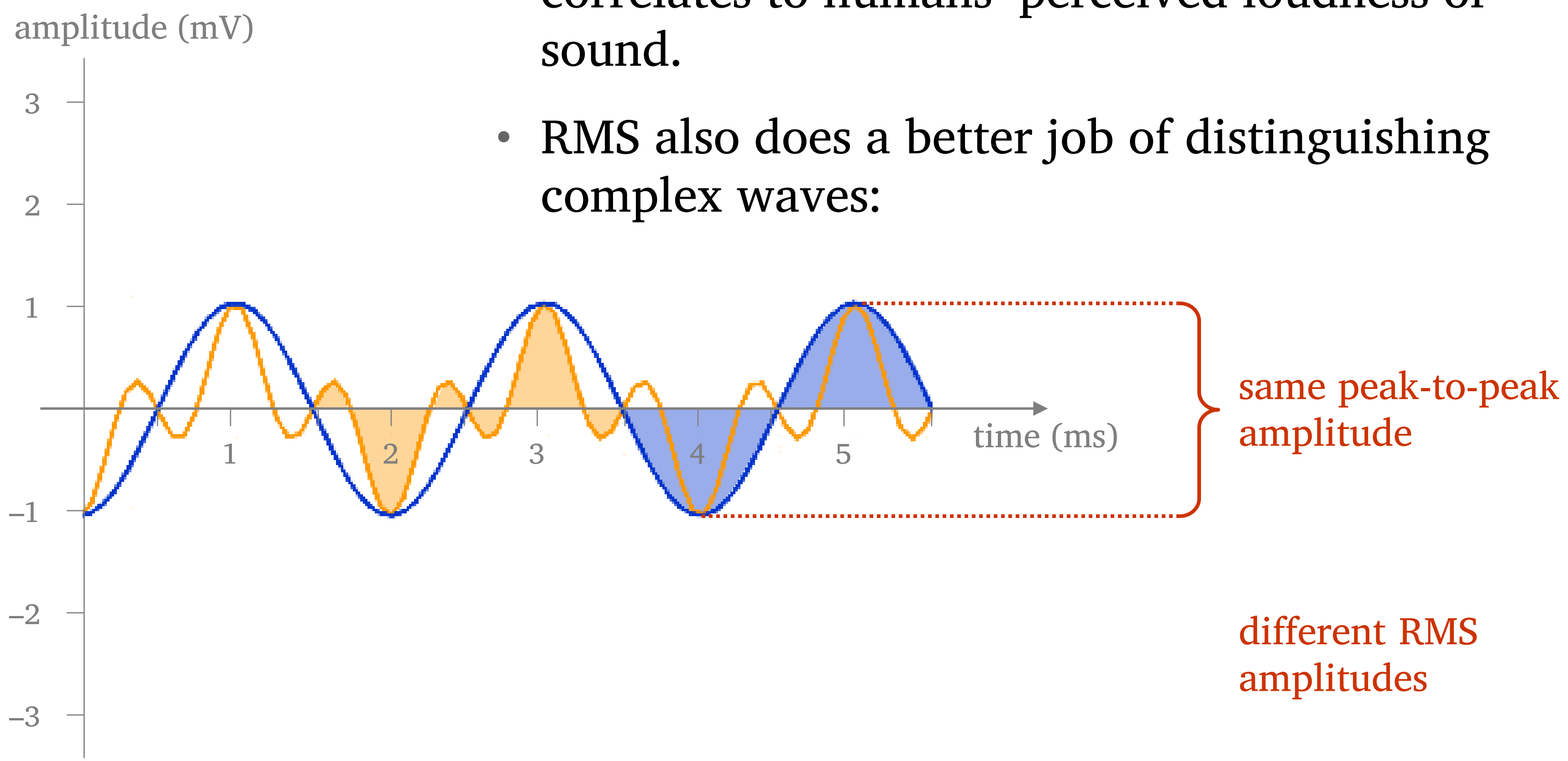
- The RMS is the average value of the wave over the course of one cycle.
  - Because part of the wave is below zero, we square each of the values first to make them all positive.
  - After finding the average of the squared values, we take the square root of the result.



$$\text{Amplitude} = \sqrt{\frac{a_1^{\,2} + a_2^{\,2} + a_3^{\,2} \ldots + a_n^{\,2}}{n}}$$

Calculating Amplitude: Which Method is Best?

- Phonetic researchers typically use RMS as a measure of amplitude because it more closely correlates to humans' perceived loudness of sound.
- RMS also does a better job of distinguishing complex waves:

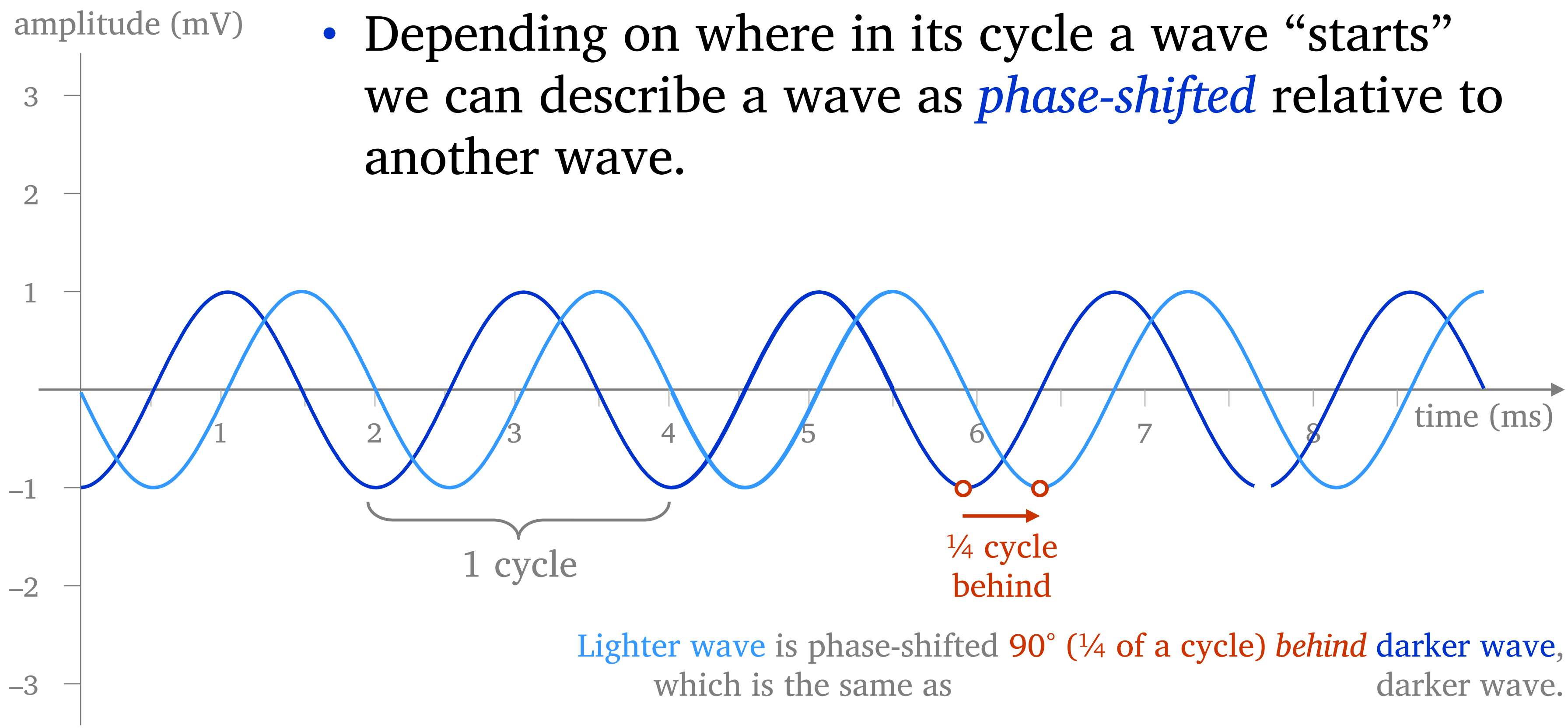


Intensity and Loudness

- The *intensity* of a wave is a measure of *average energy flux* (energy per unit area per unit time). It is proportional to the square of the wave's amplitude.
  - The official unit of intensity is W/m^2 (Watts per square meter), but in practice, intensity is usually expressed in dB (decibels), which is a dimensionless unit representing the logarithm of the ratio between the intensity of the measured sound and the intensity of a reference sound.
  - For our purposes, all you need remember is that a 1dB increase in intensity corresponds to a tenfold increase in RMS amplitude.
- *Loudness* is the human percept of intensity. Like pitch, loudness perception is nonlinear.
  - Measures of loudness (that attempt to account for the nonlinearities of human perception) include the *Bark*, the *sone*, and the *Phon*.

Phase

- Progress through a wave's cycle is measured in degrees. Like circles, waves have 360° per cycle.
- Depending on where in its cycle a wave "starts" we can describe a wave as *phase-shifted* relative to another wave.



Lighter wave is phase-shifted 90° (¼ of a cycle) *behind* darker wave, which is the same as darker wave.

Phase Perception

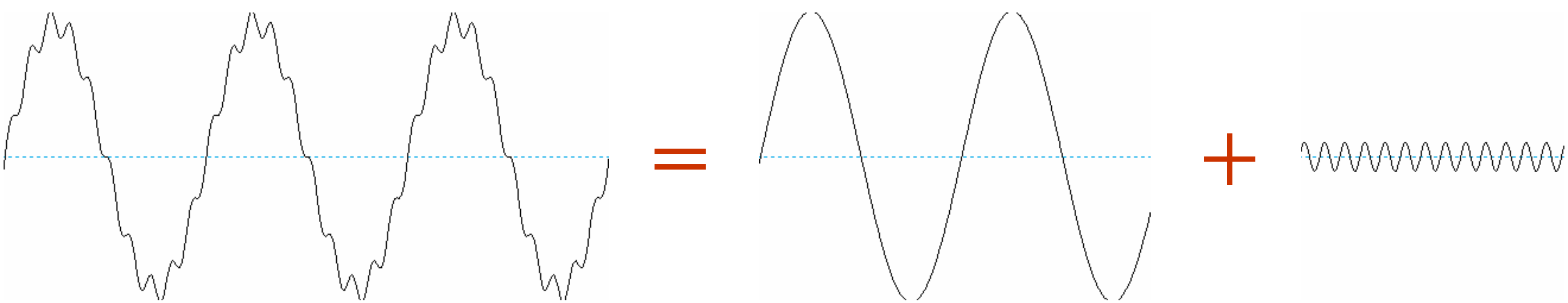
- Humans cannot hear the difference between two waves that differ only in their phase.
  - However, when wave travels different distances to reach each of our ears, the waves striking our two eardrums will be *out of phase* (i.e., phase-shifted relative to each other).
  - Comparing the phase of the two waves helps us perceive where a sound originated in space.

Summary

- *Frequency* is measured in Hz (cycles per second) and is a property of waves.
  - *Pitch* is the percept of frequency, measured in *Bark* or *Mels*.
- *Amplitude* is a measure of the magnitude of pressure fluctuations (measured in *Pa* or *mV*).
  - *Intensity* (measured in dB) is usually calculated by the root mean square method of measuring amplitude over a span of time.
  - *Loudness* is the percept of intensity, measured in *sones* or *Phon*.
- *Phase* does not affect our perception of sound *per se*, but it does help us identify a sound's directional source.

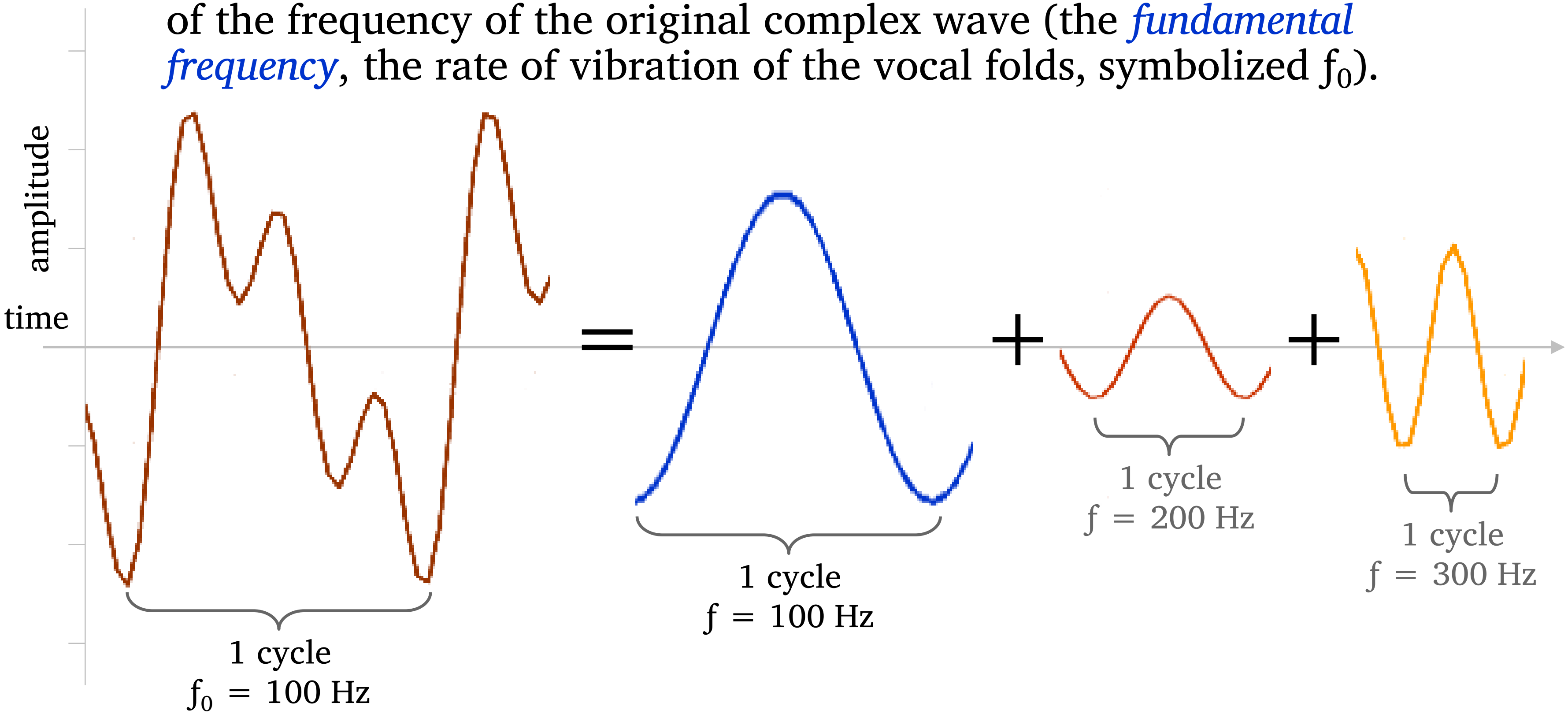
Complex Waves

- In reality, nearly every material vibrates irregularly, and thus the sound waves created by vibration are almost always *complex waves*.
- Pressure waves are *additive*, so that a complex wave can be said to be made up of several *component waves*.



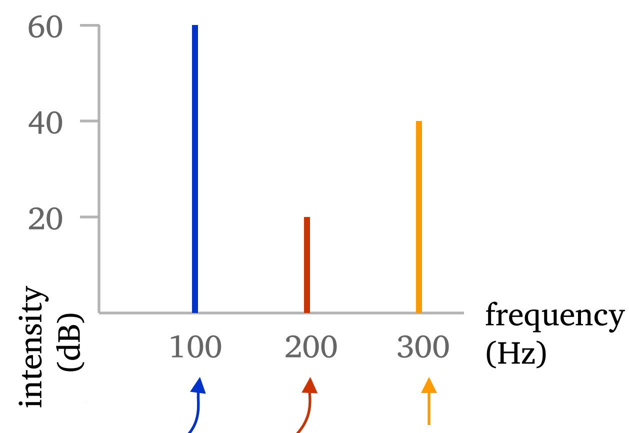
Fourier's Theorem

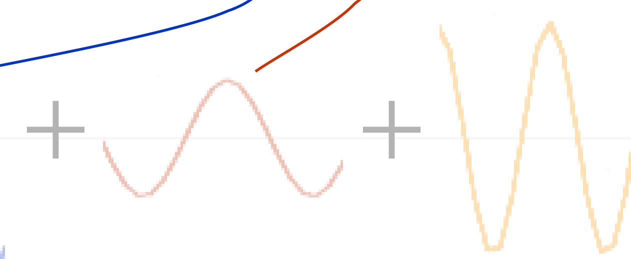
- Fourier's theorem states that all complex periodic waves can be broken down into a sum of simple *sine* waves of varying amplitudes, frequencies, and phases.
  - The frequencies of the component waves will always be integer multiples of the frequency of the original complex wave (the *fundamental frequency*, the rate of vibration of the vocal folds, symbolized f_0).



Power Spectra

- A *power spectrum* (plural *spectra*) is a graph showing the different frequencies of sine wave components that make up a complex wave, and the different magnitudes of each contributing component wave.
  - Spectral magnitudes are usually given in units of sound intensity (i.e., *decibels*).
  - The process of converting a complex wave into its power spectrum is called *Fourier transform*.





Quality

- When two complex waves have the same fundamental frequency but vary in the relative amplitudes of their components, we will perceive them to have the same pitch, but differ in *quality*.
  - Quality (e.g., *vowel quality*) is the human percept of a power spectrum (i.e., the perception of varied intensities at several frequencies simultaneously). In other words:
    - Changing the shape of the vocal tract (e.g., to make a different vowel quality) changes the amplitudes and frequencies of the component waves, so power spectra will differ between vowel qualities.
  - In the study of musical sound, quality is often called *timbre*.

The Source-Filter Theory of Speech Production

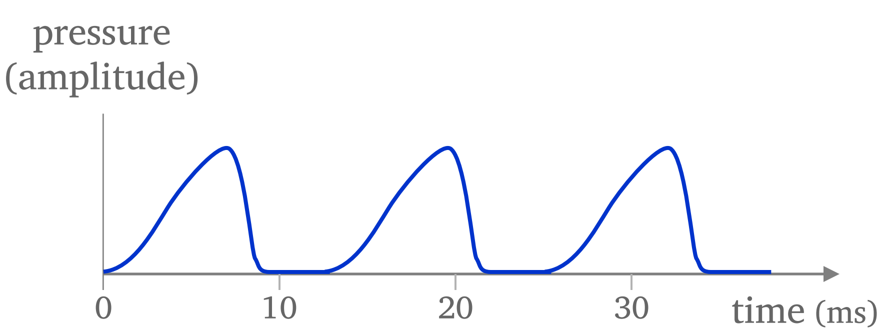
The effects of the vocal folds and vocal tract

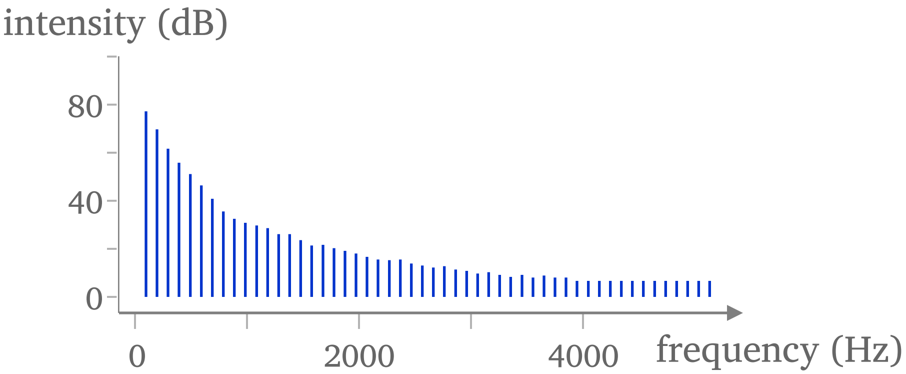
Read: LJ 8 up to “Acoustic Analysis” (p. 193)

Finish HW: Acoustic Theory

The Sound of the Larynx

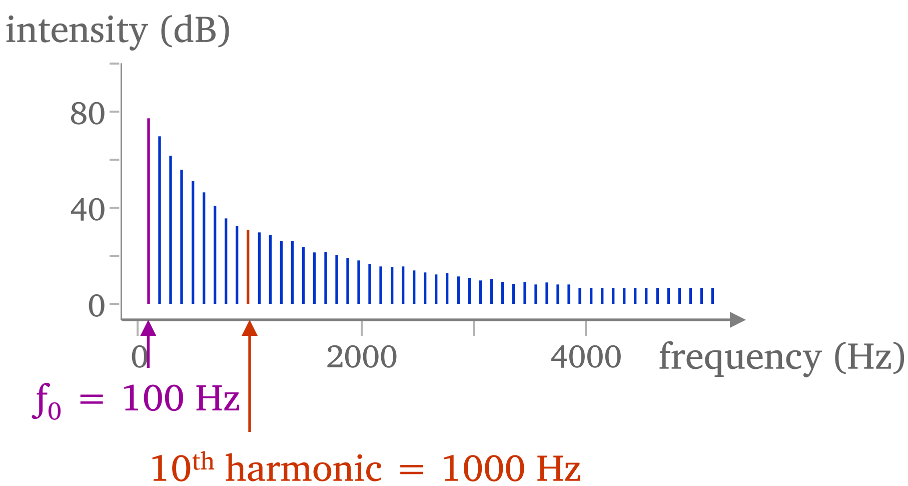
- Remember that the pressure waves generated by vibrating vocal folds are complex waves.
- In modal-voiced sounds, the vocal folds tend to blow open smoothly and then snap shut, giving the wave a sawtooth-like pattern, and a corresponding power spectrum as shown below.
  - Bernoulli's principle at work: air from the lungs pushes vocal folds open. Open = less friction = lower pressure → vocal folds snap shut. Repeat.

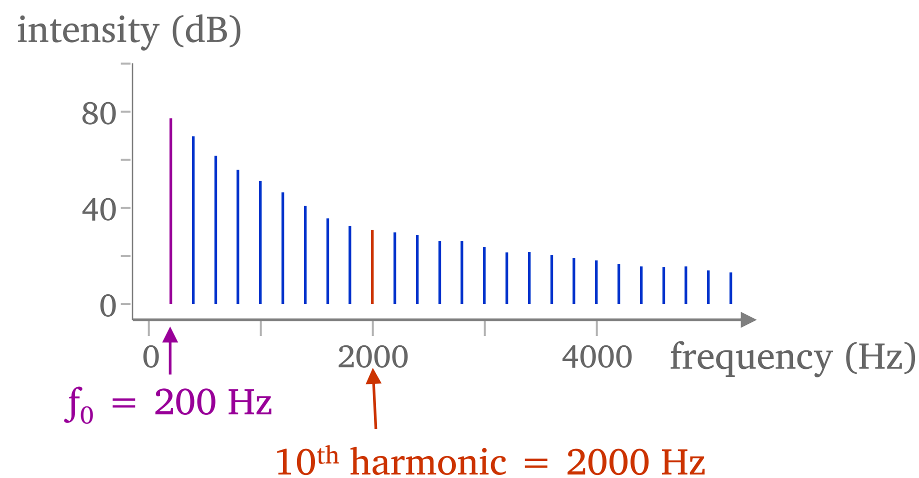




The Effect of f_0 on the Power Spectrum

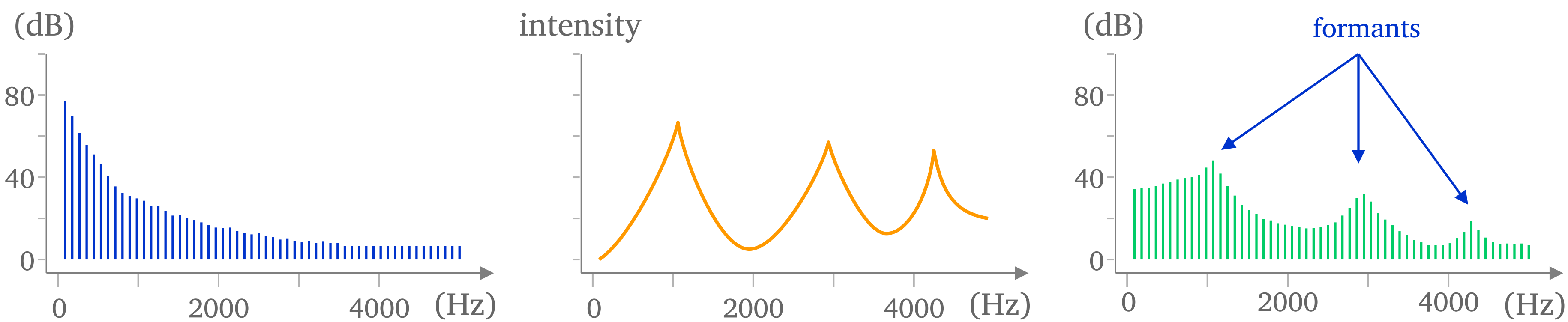
- Recall that the component waves of a complex wave are integer multiples of the fundamental frequency; thus, the peaks on a power spectrum are evenly spaced.
- Changing the fundamental frequency of the voice changes the width of the space between the harmonics on the power spectrum.





The Vocal Tract as Resonator

- As the "sawtooth" wave passes out through the vocal tract, the shape of the vocal tract acts as a set of *resonators*:
  - The resonator *reinforces* certain component frequencies (so they increase in amplitude) and *attenuates* other component frequencies (so they decrease in amplitude).
  - This creates broad peaks of higher energy at the frequencies where the filter graph peaked. These broad peaks are called *formants*.



vocal fold power spectrum (source) | vocal tract frequency curve (filter) | resulting power spectrum

The Source-Filter Theory

- The separation of the speech signal into the source wave of the vocal folds and the vocal tract resonators is called the *source-filter theory* of speech production.
  - The *source* is determined by fundamental frequency and phonation type.
  - The *filter* is determined by the size and shape of the vocal tract (including the position of the articulators).

Formants and Formant Frequencies

- The frequencies that are most strongly reinforced by the resonator are called *formants*.
  - Formants are typically the tallest bars on a power spectrum.
  - As the shape of the vocal tract changes (through lip, tongue, jaw, and velum movements) those *formant frequencies* also change.
- In the analysis of speech, formants are a very important means of distinguishing sounds from one another.
  - The first three formants (labeled F_1, F_2, F_3) are the most important formants for understanding the speech signal we hear.

Spectrograms

An essential tool for instrumental phonetics

Read: LJ 8 from “Acoustic Analysis” (p. 193–end)

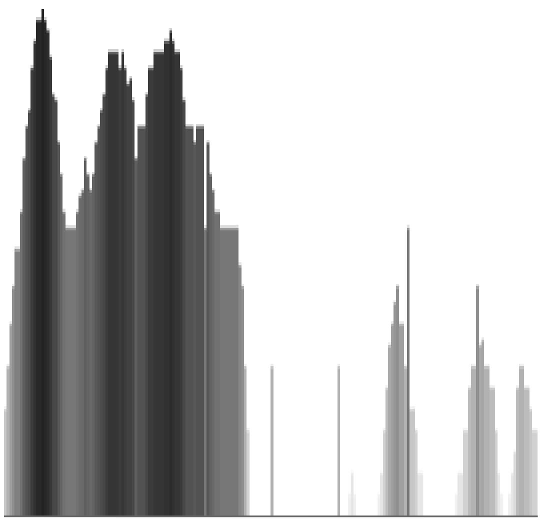
HW: Spectrograms

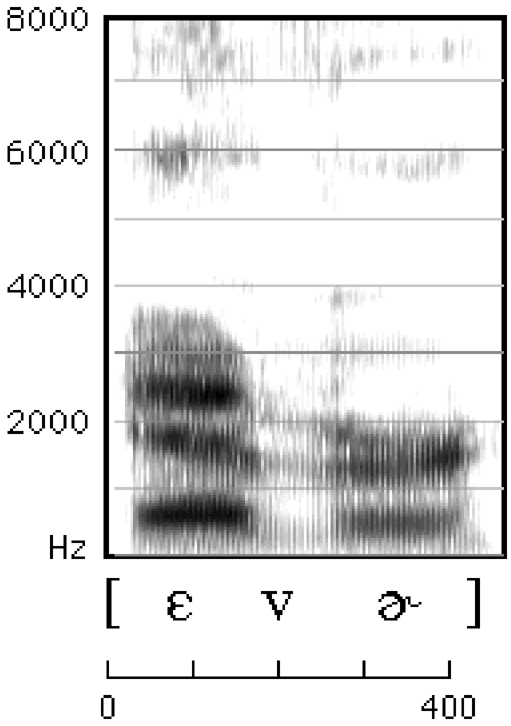
Voice Spectra and Spectrograms

- Recall that power spectra allow us to see the component frequencies and amplitudes that make up a complex wave.
- *Spectrograms* are a way of viewing hundreds of sequential power spectra, so we can see how the spectrum of a speech signal changes over time.
  - This is done by turning the spectrum sideways, reducing each bar on the spectrum to a dot, and using the color of the dot to symbolize the height of the bar (i.e., the amplitude or intensity of the component frequency).

Building a Spectrogram

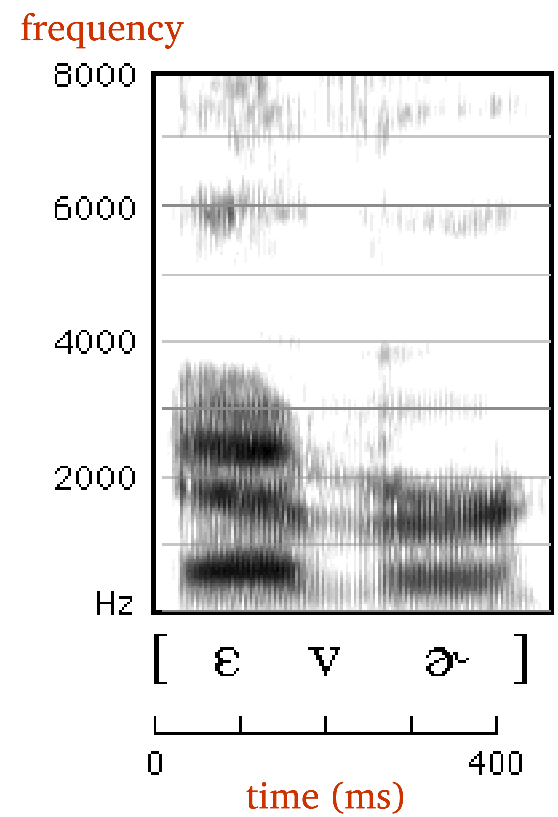
1. Power spectrum amplitudes are represented as different shades of grey (taller = darker)
2. Spectrum is rotated and collapsed to 1-dimension.
3. Hundreds of spectra are arrayed side by side, allowing us to see changes in formant frequencies over time.





Reading a Spectrogram

- A spectrogram plots three variables:
  - The horizontal axis is *time*, usually measured in milliseconds.
  - The vertical axis is *frequency* (Hz).
  - The third variable, *intensity*, is represented by the color of each point that is plotted.
- The dark bands show the *formants* of the voiced sounds.
  - The first three formants vary depending on the shape of the vocal tract and are important for understanding the speech signal.



Narrowband Spectrograms

- When calculating the frequency components that make up the speech signal, it is necessary to take a chunk of time to figure out the frequency of the wave.
- *Narrowband spectrograms* show calculations based on longer chunks of time, which allows more accurate calculation of the *frequencies* making up the signal and can reveal individual *harmonics* of the signal.
  - Because the chunks of time are longer, narrowband spectrograms are less precise about when a given speech event happened (e.g., a release burst or the onset of voicing). This lack of precision in the temporal dimension is called *time smear*.
  - Because of their detailed frequency information, narrowband spectrograms are useful for investigating changes in pitch, tone, and intonation.

Narrowband Spectrograms

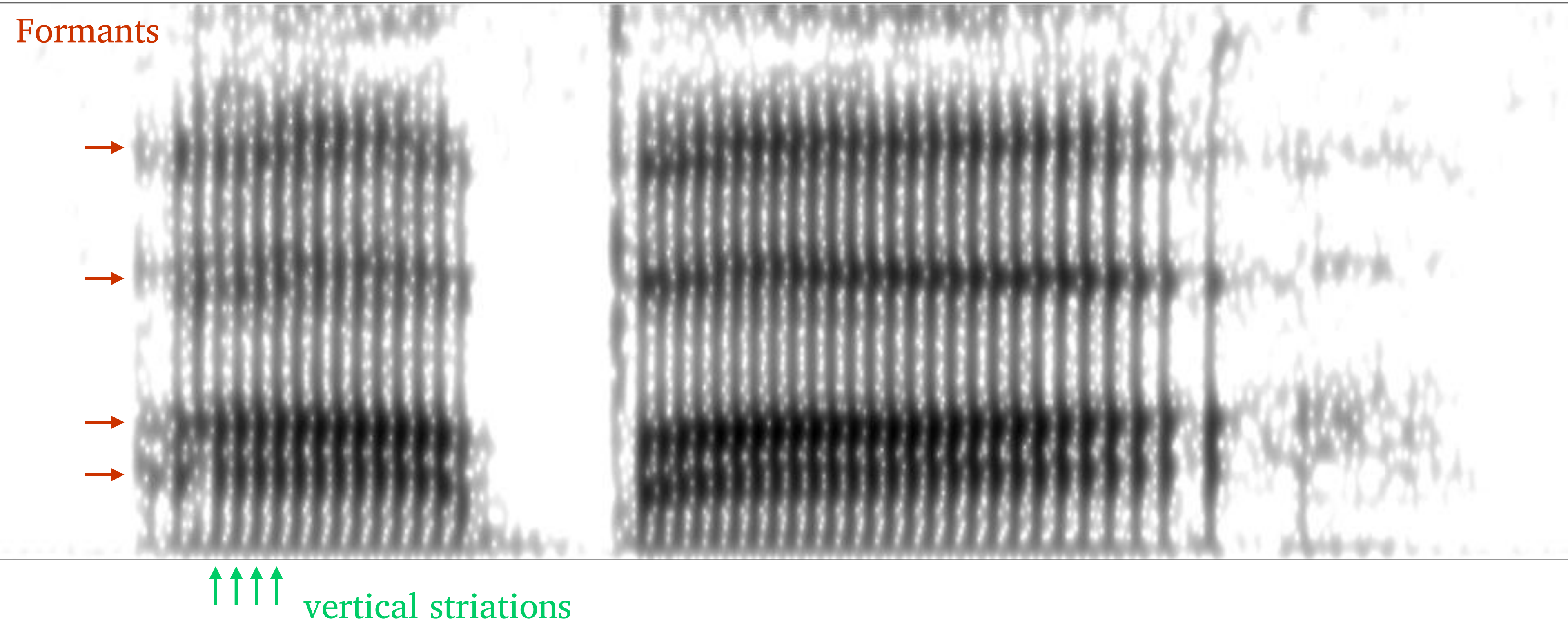
- [ɑbɑ]
  - Frequency range: 0 – 5000 Hz; duration: ≈0.75 s
  - Window size: 0.05 s

Wideband Spectrograms

- *Wideband spectrograms* show calculations based on short chunks of time, which allows accurate measurement of event timing at the expense of detailed frequency information. Wideband spectrograms are vertically *striated*; each vertical striation shows an individual vibration of the vocal folds.
  - Remember that formants are *broad peaks* of intensity that range over several neighboring frequencies. Since the wideband spectrograms "smear together" neighboring frequencies, this actually helps us to see the formants more clearly.
  - Vowel quality, place of articulation, and other segmental features are all aspects of the *filter* and thus cause broad-frequency changes to the signal. Thus they are best viewed using wideband spectrograms.
  - Because of their accurate timing calculations, wideband spectrograms are also useful for measuring VOT, vowel duration, and other time-based characteristics of speech.

Wideband Spectrograms

- [ɑbɑ]
  - Frequency range: 0 – 5000 Hz; duration: ≈0.75 s
  - Window size: 0.01 s

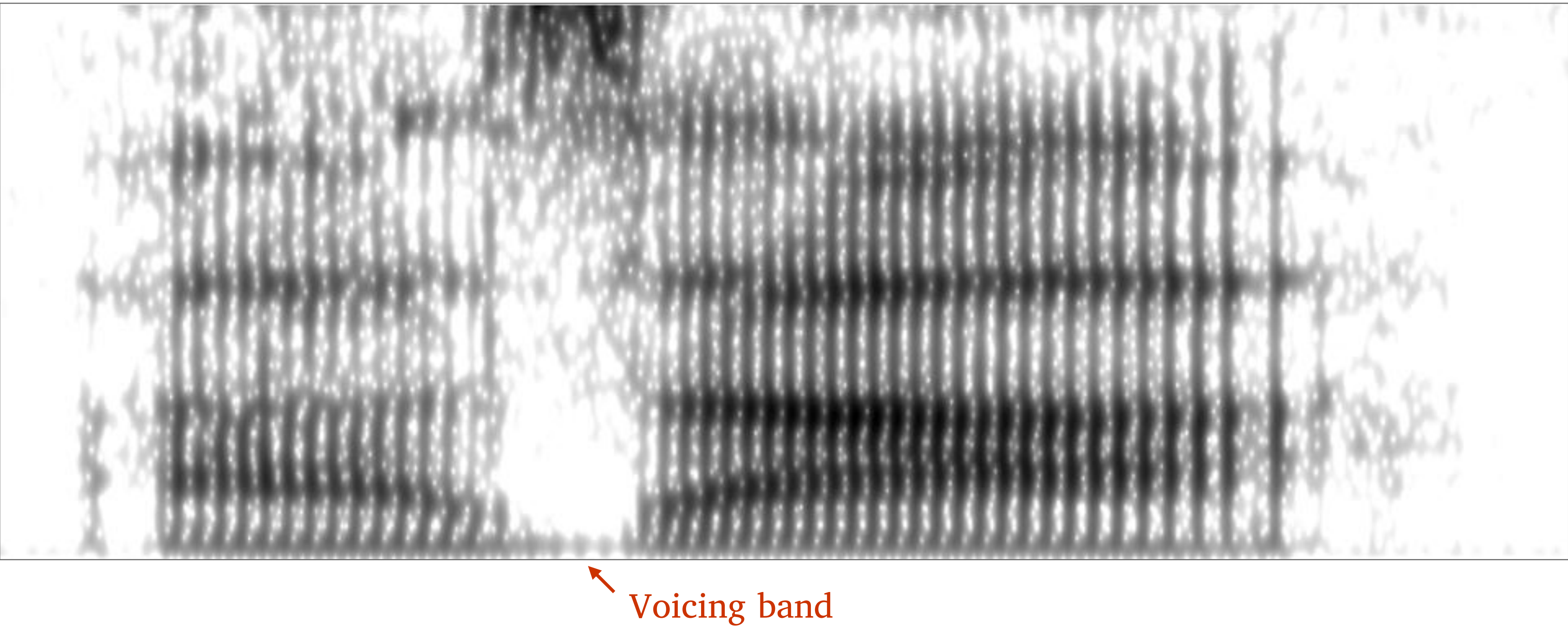


Reading a Spectrogram: Voicing

- *Voiced sounds* have a dark band very close to the bottom of the spectrogram, corresponding to the fundamental frequency of the speaker's voice.
  - Many "voiced" consonants are allophonically devoiced, so the absence of a voicing band does not guarantee that the sound is a voiceless phoneme.

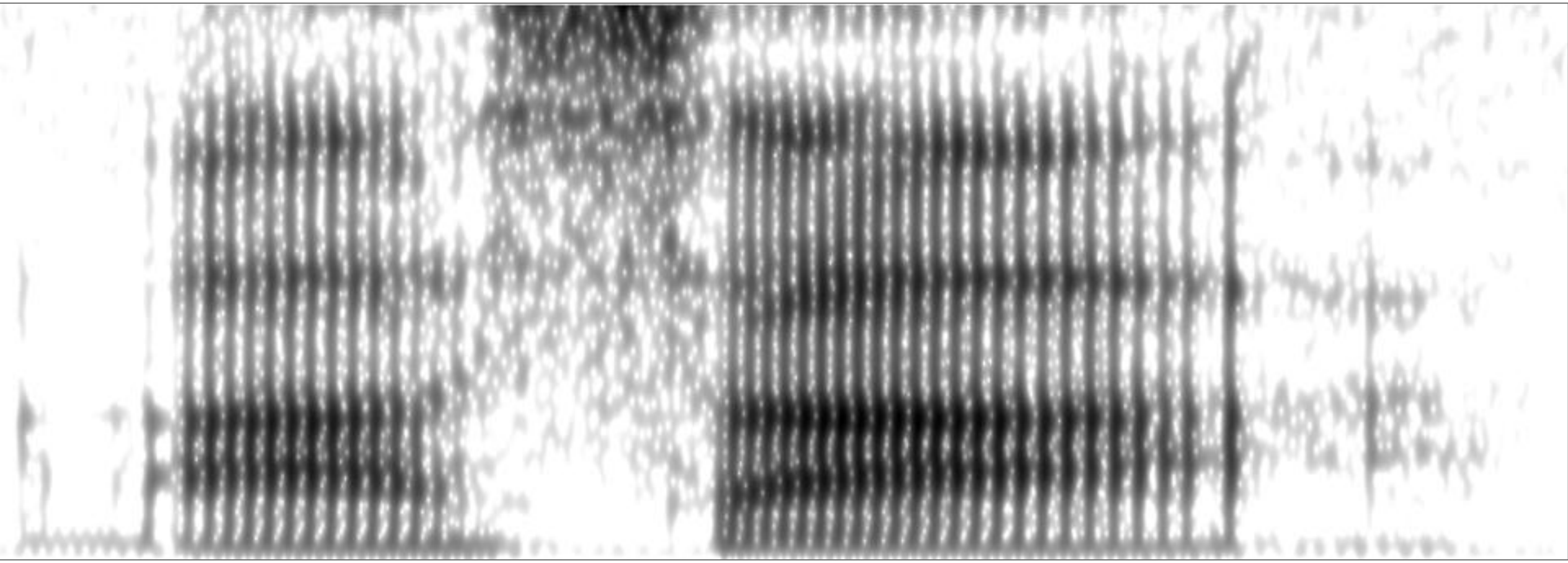
Reading a Spectrogram: Voicing

- [ɑzɑ]
  - Frequency range: 0 – 5000 Hz; duration: ≈0.75 s



Reading a Spectrogram: Voicing

- [asɑ]
  - Frequency range: 0 – 5000 Hz; duration: ≈ 0.75 s



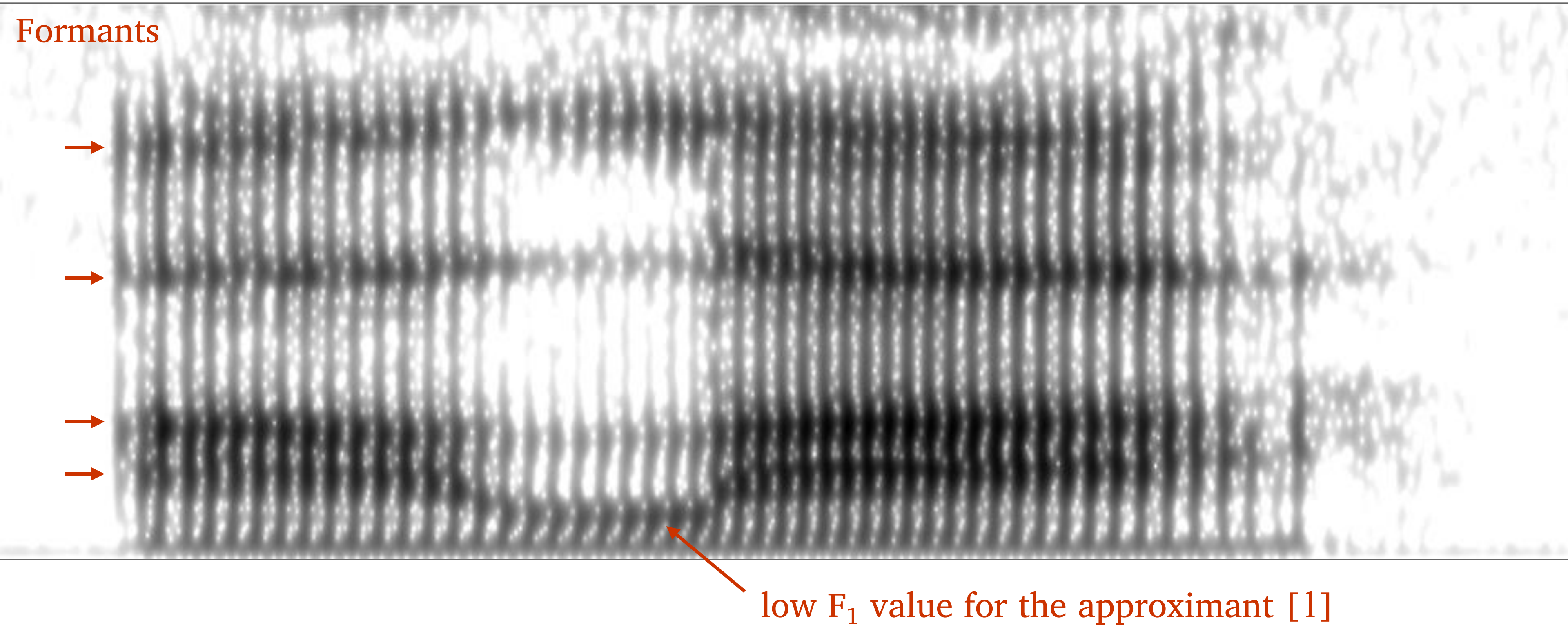
No voicing band

Reading a Spectrogram: Approximants

- *Vowels* and *approximants* show vertical striations in wideband spectrograms, and clear dark formant bands.
  - Vowels are identified on a spectrogram primarily by the values of their F_1 and F_2 formants.
  - Approximants involve constriction comparable to that of a high vowel, and so tend to have fairly low F_1 values. Approximants can often be distinguished by the movement of their F_2 and F_3 formants.

Reading a Spectrogram: Approximants

- [ɑlɑ]
  - Frequency range: 0 – 5000 Hz; duration: ≈0.75 s

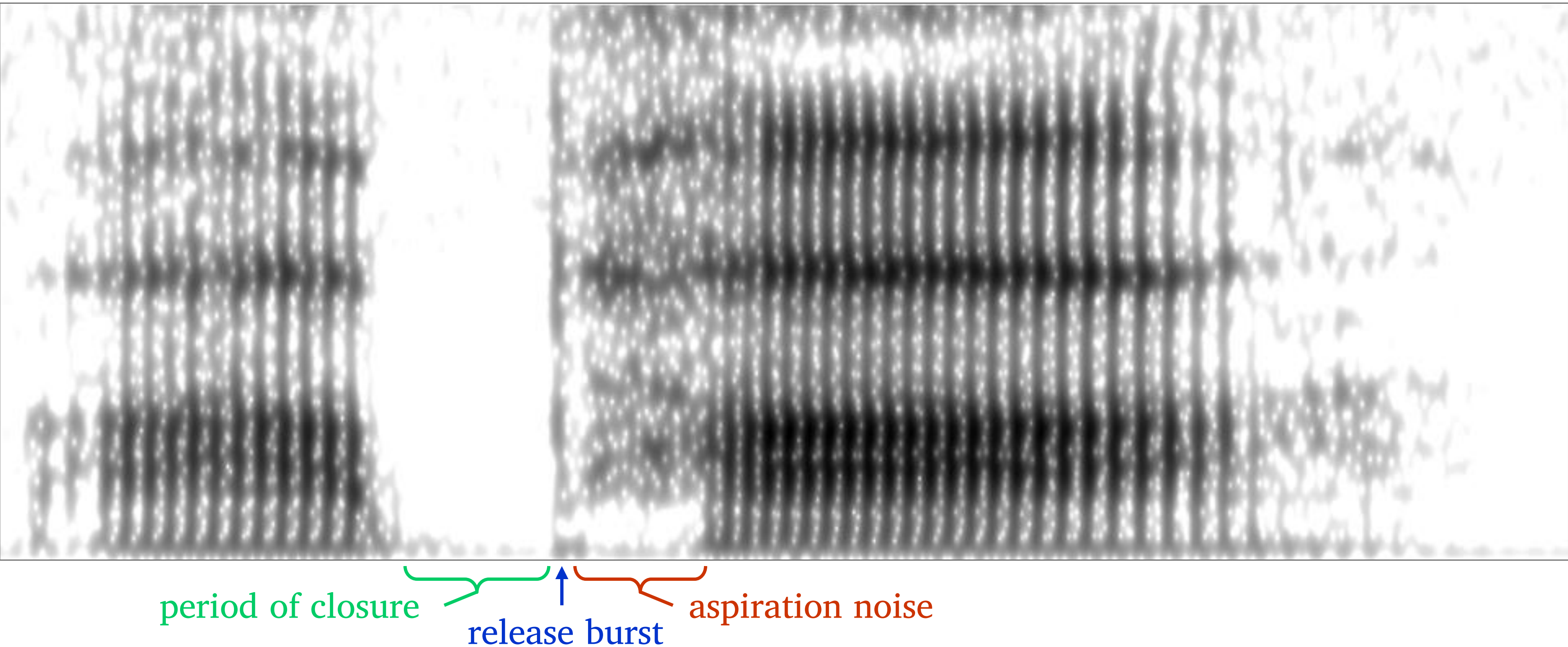


Reading a Spectrogram: Oral Stops

- *Oral stops* involve complete stoppage of airflow and thus show up as a “gap” in the spectrogram.
  - The formant transitions of the voiced sounds before or after a stop are the best way to determine the stop’s place of articulation.
  - The release burst of a stop may be continuous with a following voiced sound, or may stand alone at the end of a word as a brief vertical band of high energy. The distribution of that energy can also provide clues to the place of articulation of the stop.

Reading a Spectrogram: Oral Stops

- [ɑpɑ]
  - Frequency range: 0 – 5000 Hz; duration: ≈ 0.75 s

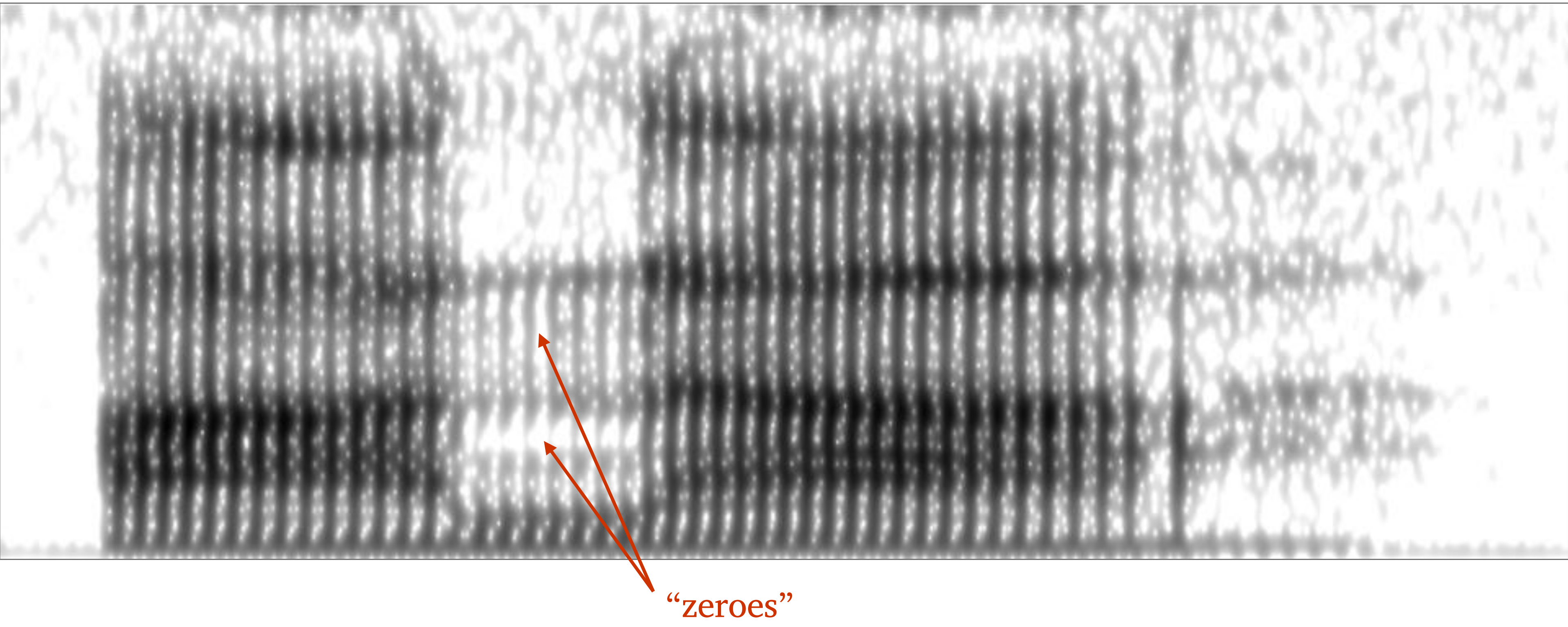


Reading a Spectrogram: Nasal Stops

- *Nasal stops* allow continuation of airflow and show formant structure much like vowels.
  - Like oral stops, nasal place of articulation is most easily determined by the formant transitions of the surrounding vowels.
  - Unlike vowels, nasals typically show “zeroes” (white areas) between the lower formants.
  - Nasals can be distinguished from vowels by a much lower amplitude in the waveform (not the spectrogram).

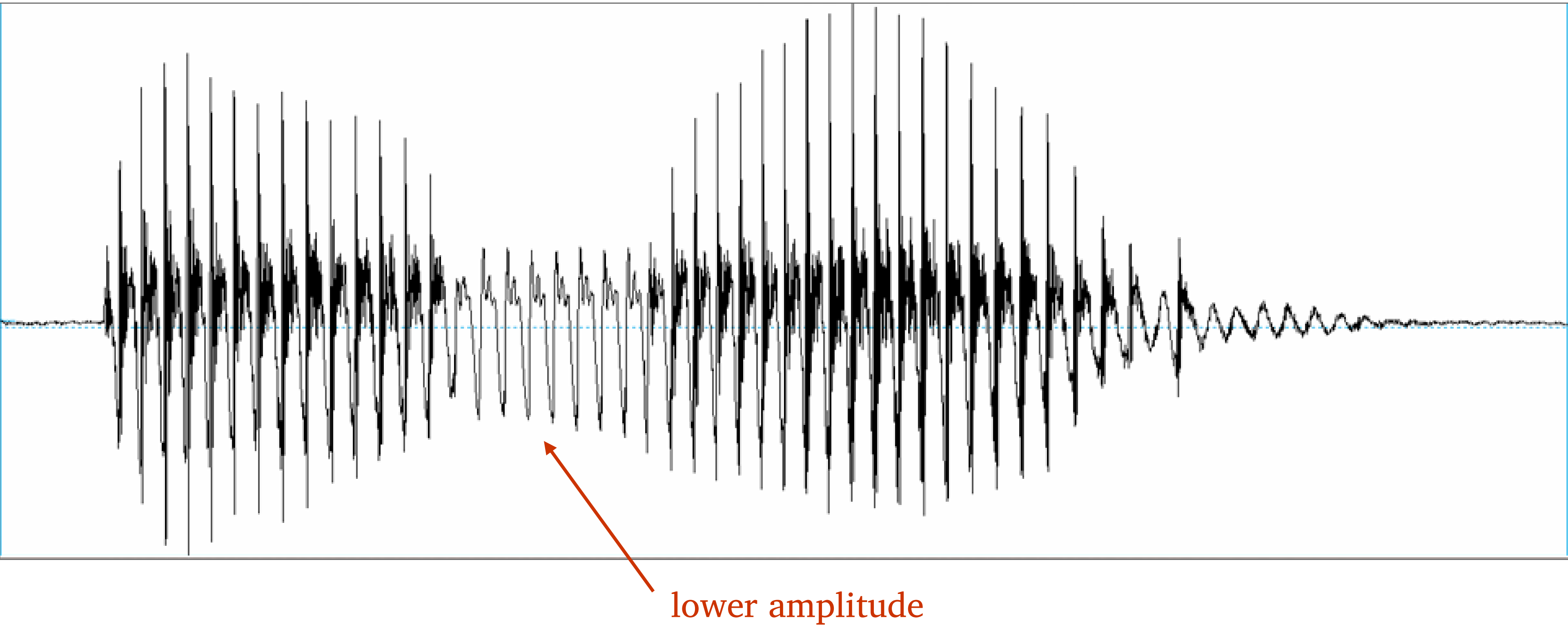
Reading a Spectrogram: Nasal Stops

- [ɑnɑ]
  - Frequency range: 0 – 5000 Hz; duration: ≈0.75 s



Reading a Spectrogram: Nasal Stops

- [ɑnɑ]

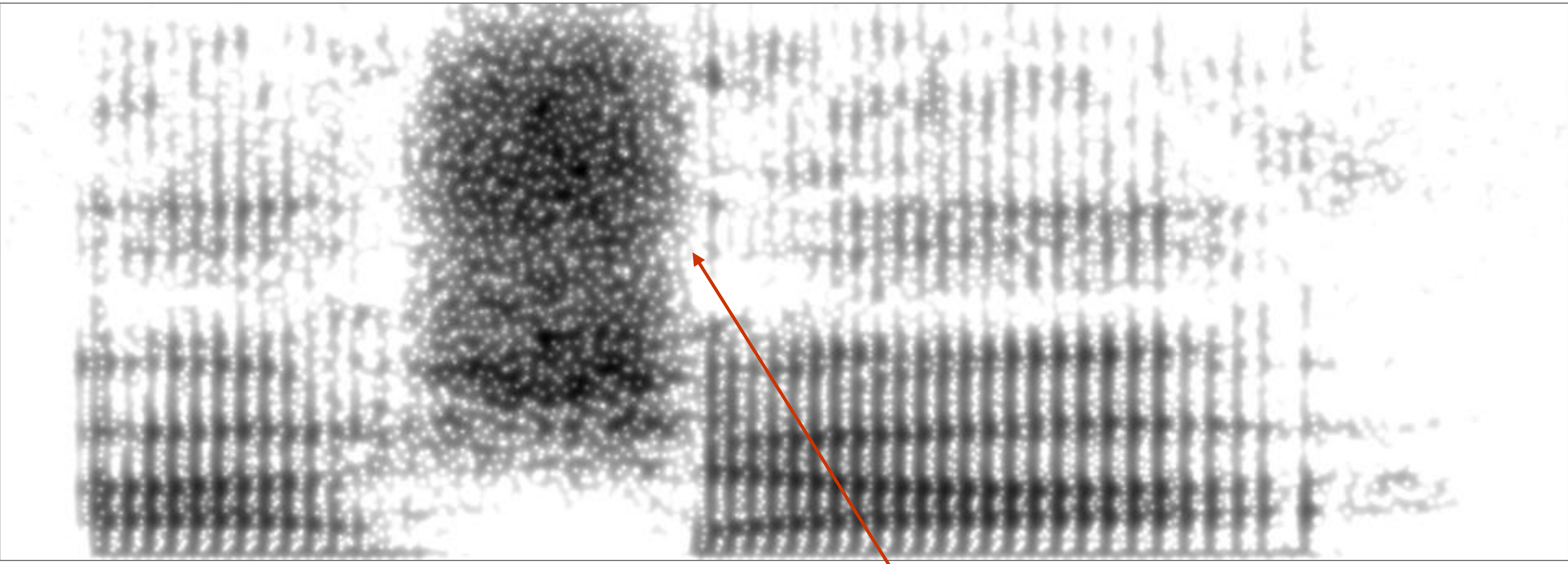


Reading a Spectrogram: Fricatives

- *Fricatives* are characterized by fairly high energy (darker color) especially in the higher frequencies, and (usually) a lack of formant bands.
  - [h] often does show formant bands. This is because [h] generates turbulence in the same place that voicing originates (the glottis), so the turbulence of [h] is subject to the same filter that vowels are (namely, the entire vocal tract). The spectrum of [h] can vary widely depending on the following vowel.

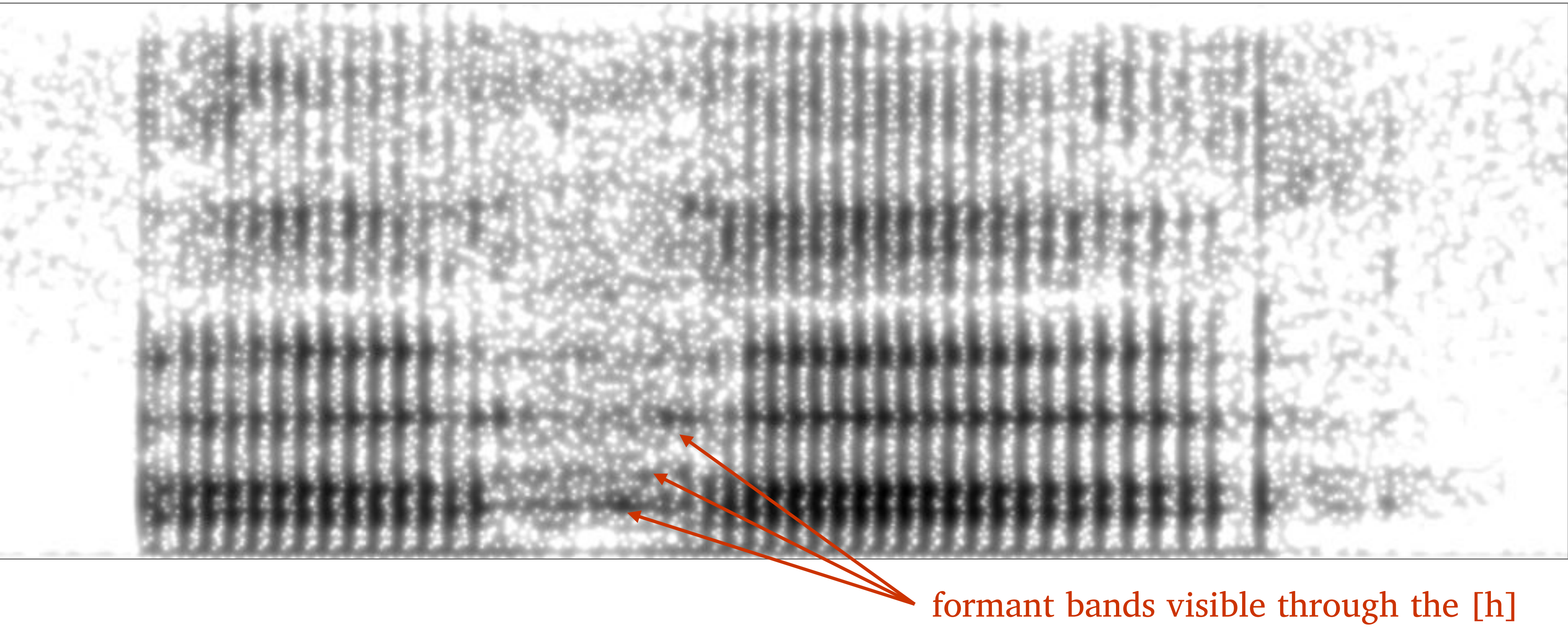
Reading a Spectrogram: Fricatives

- [ɑʃɑ]
  - Frequency range: 0 – 10,000 Hz; duration: ≈0.75 s



Reading a Spectrogram: Fricatives

- [ɑhɑ]
  - Frequency range: 0 – 10,000 Hz; duration: ≈0.75 s



Reading a Spectrogram: More info

- For more detailed information on spectrogram reading, see:
  - Ladefoged and Johnson ch. 8, esp. p. 193-212
  - Rob Hagiwara's "Mystery Spectrogram" website: http://home.cc.umanitoba.ca/~robh/howto.html#intro
  - Spectrogram Reading Practice worksheets on the class webpage

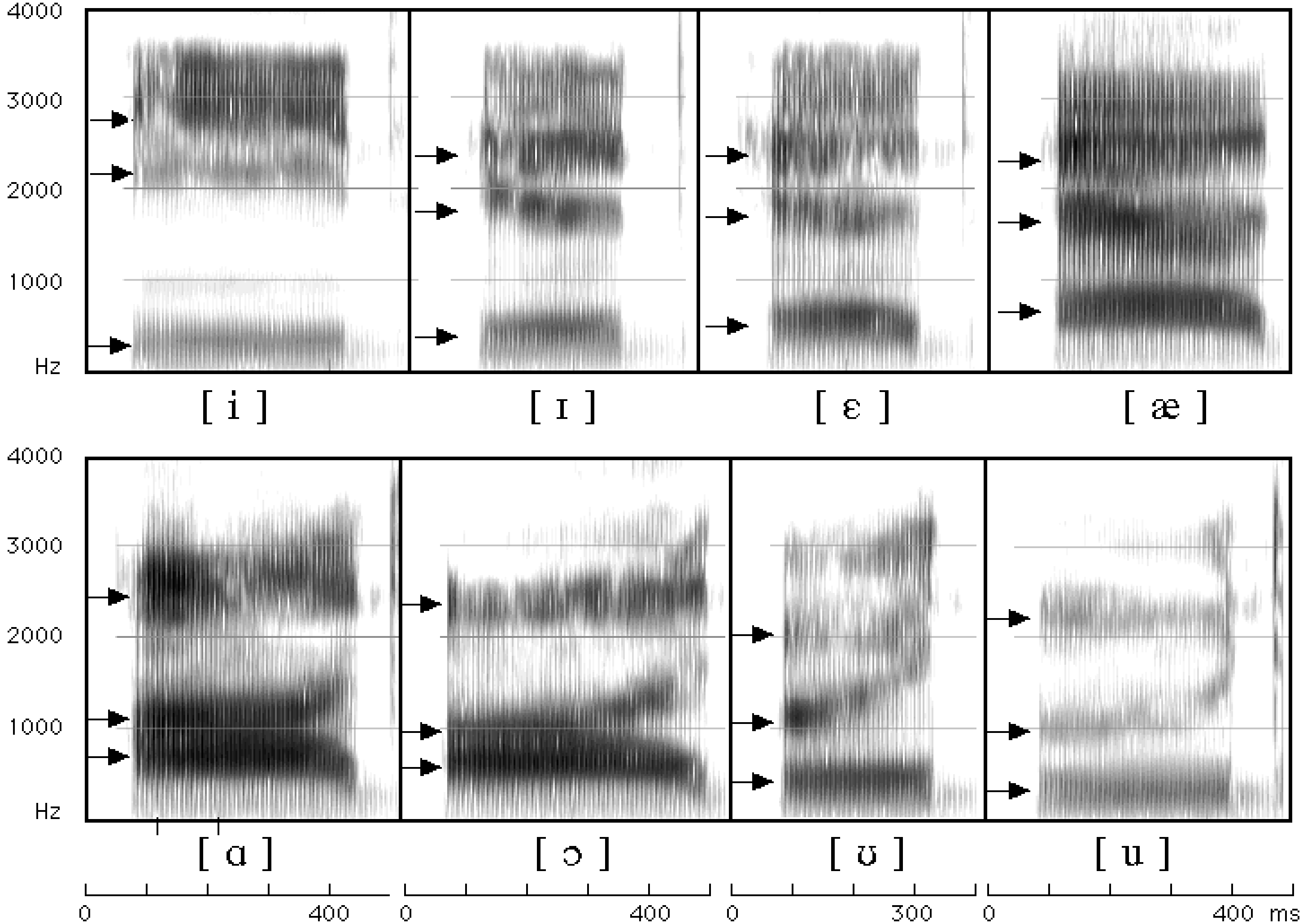
Reading Spectrograms

More examples and practice

See also: Spectrogram Reading Practice Handouts

HW: Spectrograms

Reading a Spectrogram: Vowels



from Ladefoged, P. & Johnson, K. (2011). *A Course in Phonetics* (6th ed.), p. 194.

Formant Transitions: Cues to Place of Articulation

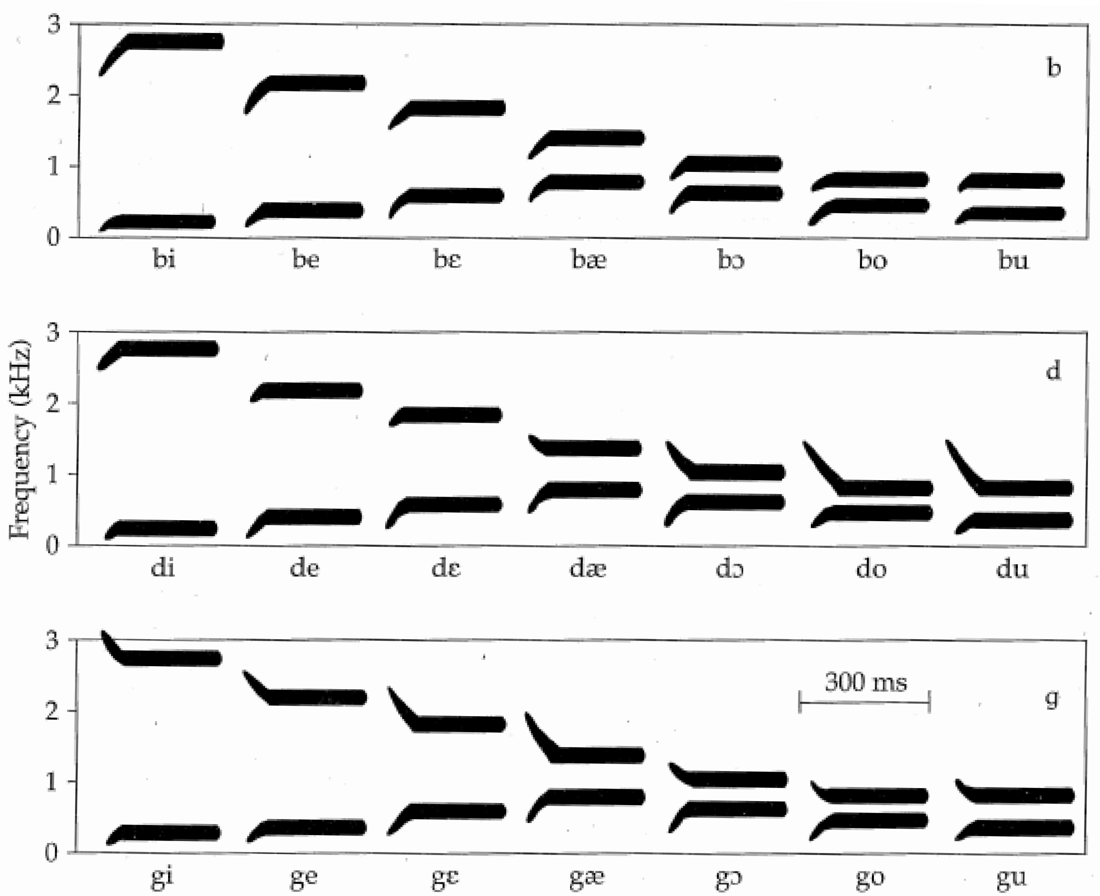
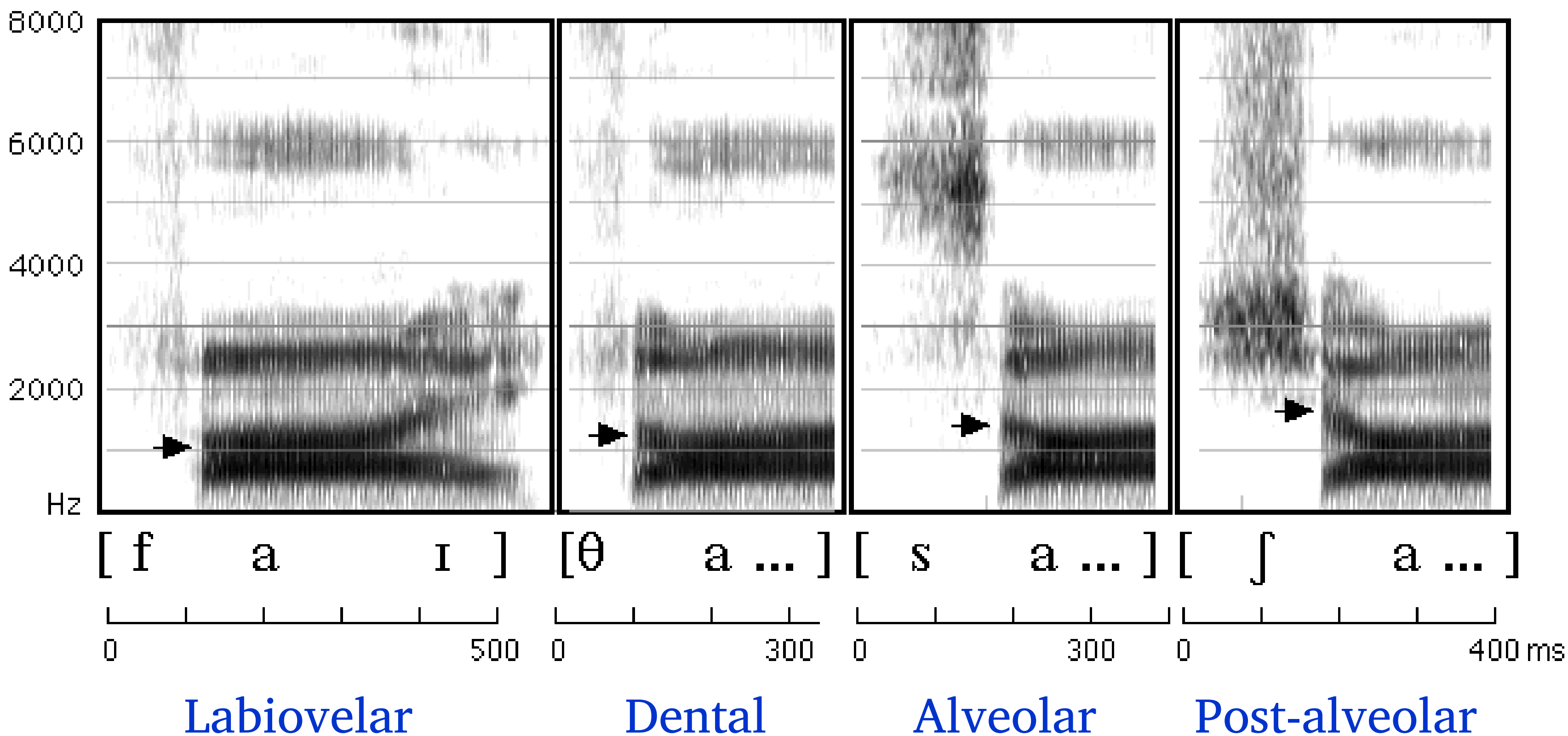


Figure 8.7 F_1 and F_2 transition patterns in stop release used to synthesize [b], [d], and [g] followed by various vowels. Adapted from Delattre et al., 1955, p. 770, and published with permission.

from Johnson, K. (2003). *Acoustic & Auditory Phonetics (2nd ed.)*. Malden, MA: Blackwell.

Reading a Spectrogram: Place

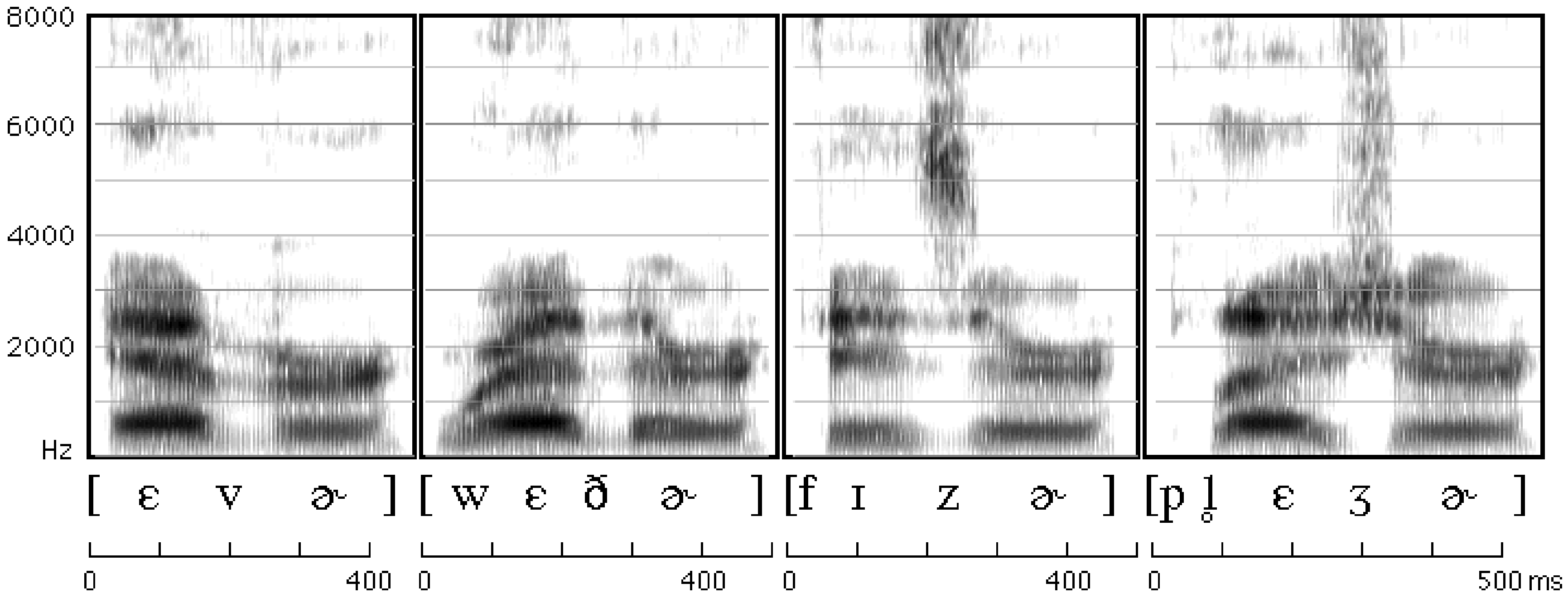
- F2 transition is higher the farther back the place of articulation:



from Ladefoged, P. & Johnson, K. (2011). *A Course in Phonetics* (6th ed.), p. 201.

Reading a Spectrogram: Distinguishing Fricatives

- Look at the spread and density of frication

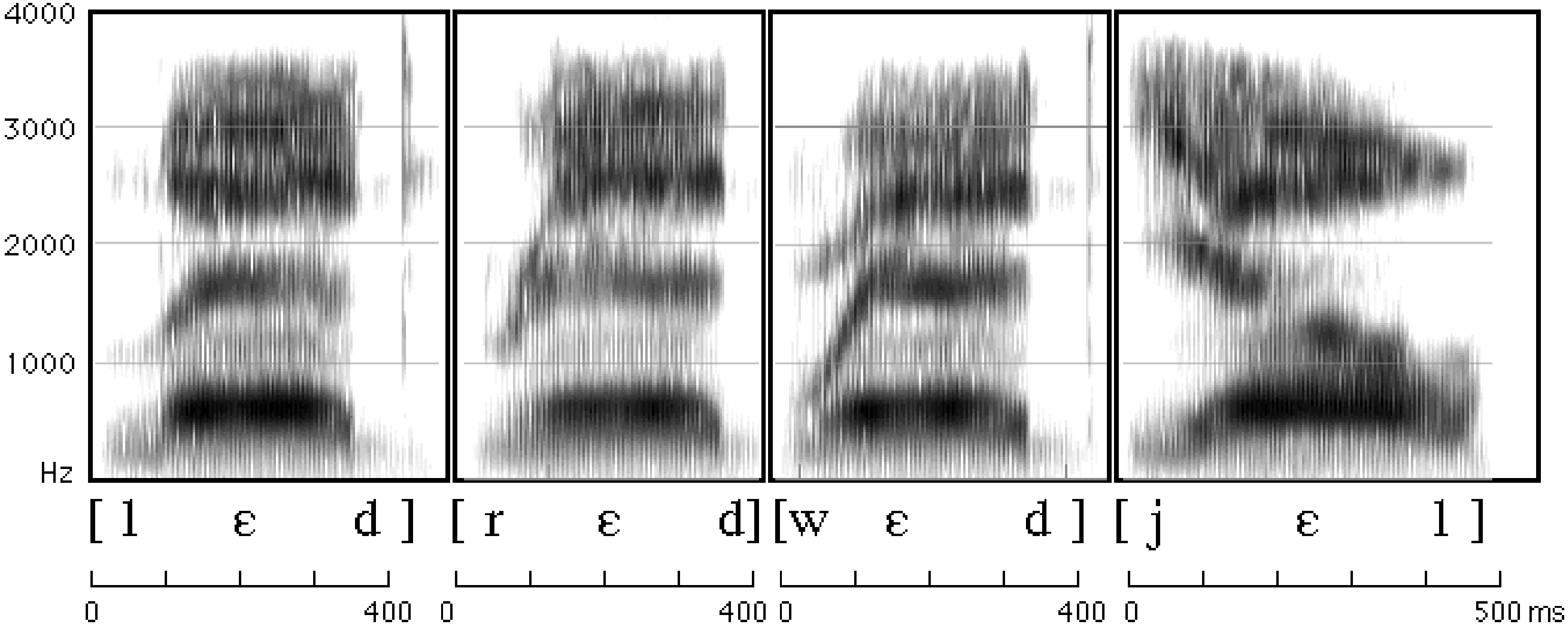


- Quiet: Frication light, spread out, present in lower frequencies
- Loud (e.g. sibilants): Dense, concentrated in high frequencies.

from Ladefoged, P. & Johnson, K. (2011). *A Course in Phonetics* (6th ed.), p. 202.

Reading a Spectrogram: Distinguishing Approximants

- Approximants flow smoothly into vowels

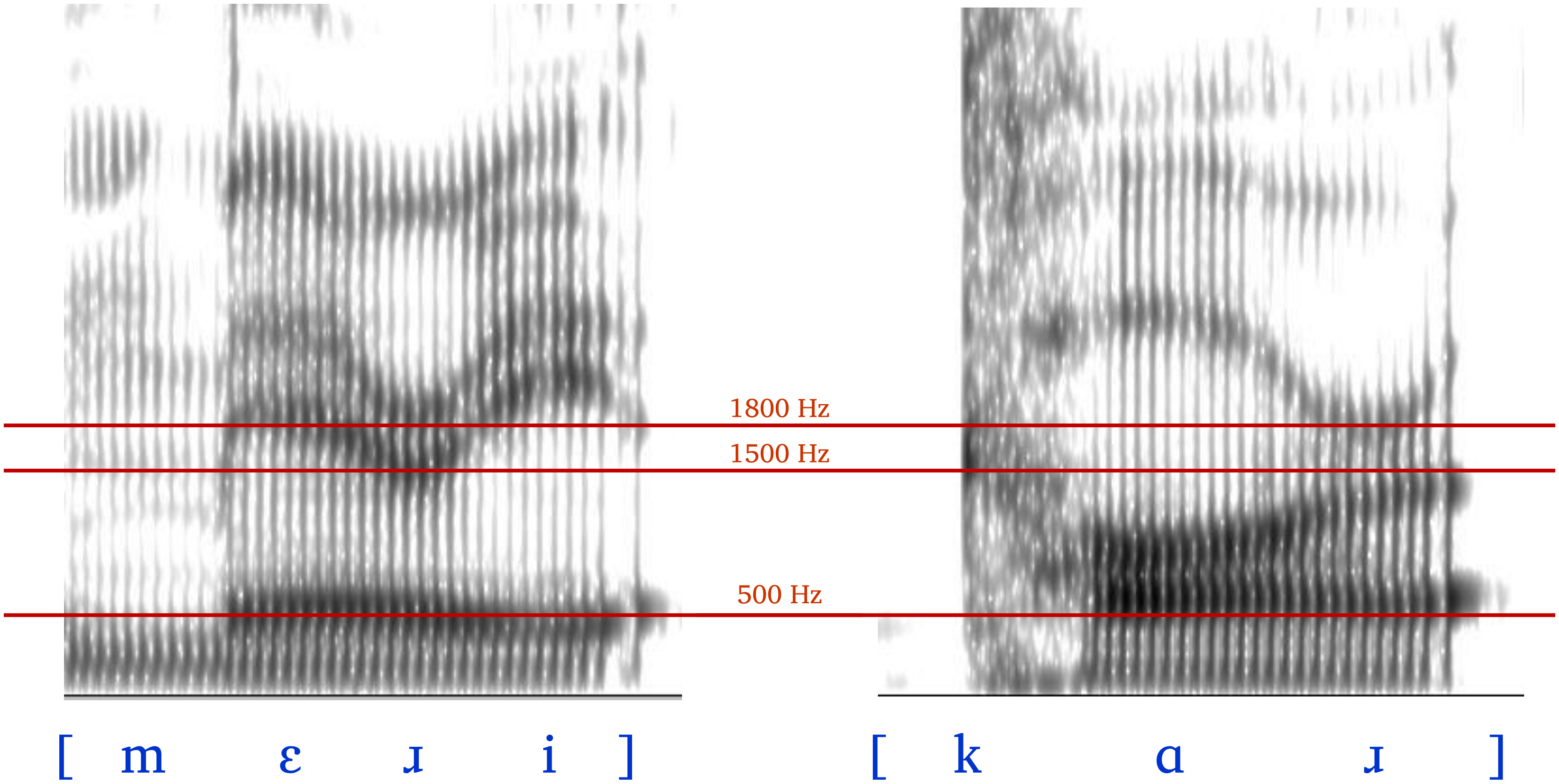


- [l] – dip in F2, slight rise in F3
- [ɹ] – central F2, scoop in F3
- [w] – like high back vowel [u/ʊ]
- [j] – like high front vowel [i/ɪ]

from Ladefoged, P. & Johnson, K. (2011). *A Course in Phonetics* (6th ed.), p. 202.

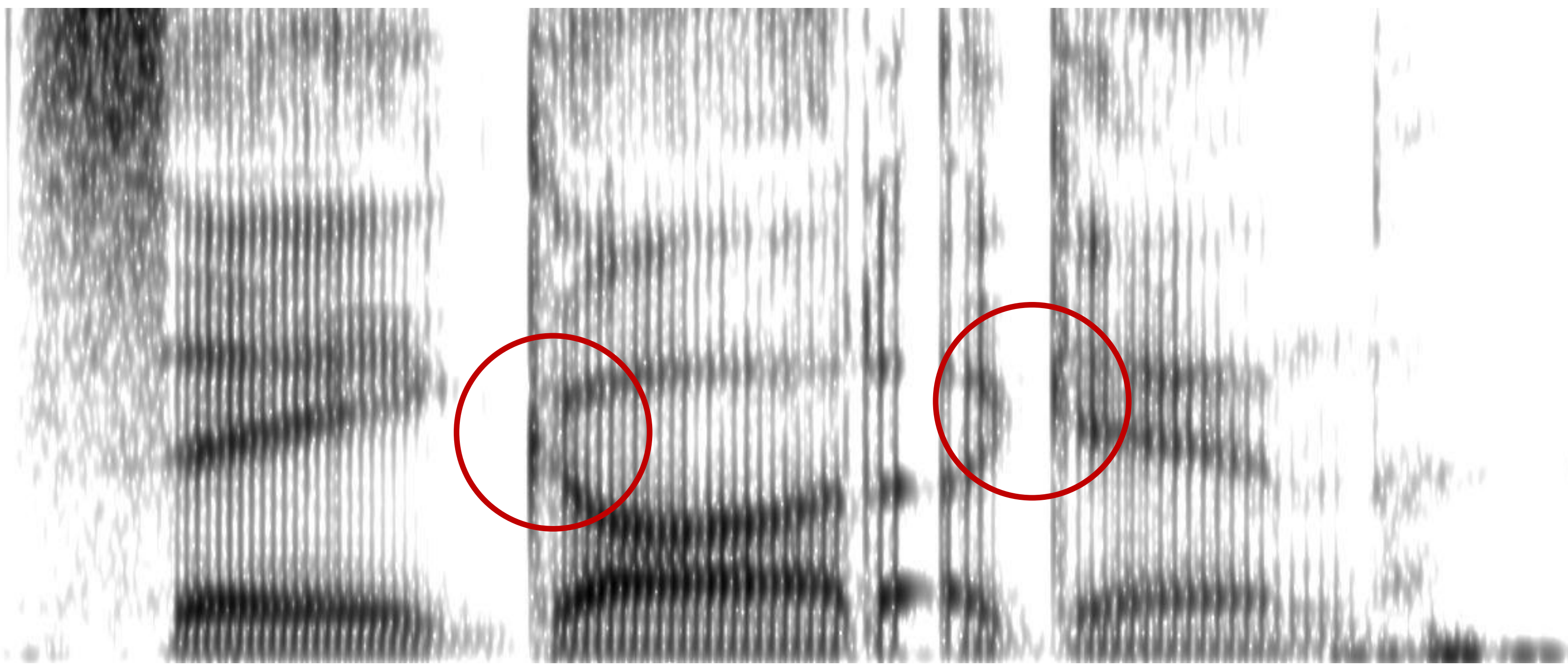
Spectrogram Examples: [ɹ]

- Scoop shape to F3
- F2 central (between front & back vowels)



Spectrogram Examples: Velar Pinch

- Velar pinch = Raised F2, lowered F3 near velars



[s e ɪ ˈg ɑ d ə ˈg ɛ n]

“Say god again.”

Vowels and Vowel-Like Articulations

Features and descriptions

Read: LJ 9

HW: IPA (All vowels - articulation)

Transcription: Non-English Vowels

Vowel Quality and Vowel Features

- Recall that we distinguish consonants based on *articulatory features* like *voicing*, *place*, and *manner* of articulation.
- Similarly, we distinguish vowels based on a combination of features, but some of these are *acoustic features* (characteristics of the vowel's wave properties) or *auditory features* (characteristics of how the vowel sounds) rather than articulatory ones.
- *Vowel quality* is a cover term for the auditory basis on which we distinguish one vowel from another.
  - Recall that *quality* is a term for our perception of spectral differences (what musicians call *timbre*).
  - Vowel quality is usually described with reference to four features: *height*, *backness*, *roundness*, and *tenseness*.

Vowel Height and Backness

- The primarily classification of vowels occurs along two dimensions described by two features: front–central–back (*backness*) and high–mid–low (*height*).
  - Our perception of vowel height is correlated with the frequency of the first vowel formant (F1).
  - Our perception of vowel backness is correlated with the frequency of the second vowel formant (F2).
- Some linguists prefer the terms *close* and *open* instead of high and low (respectively).

Vowel Roundness

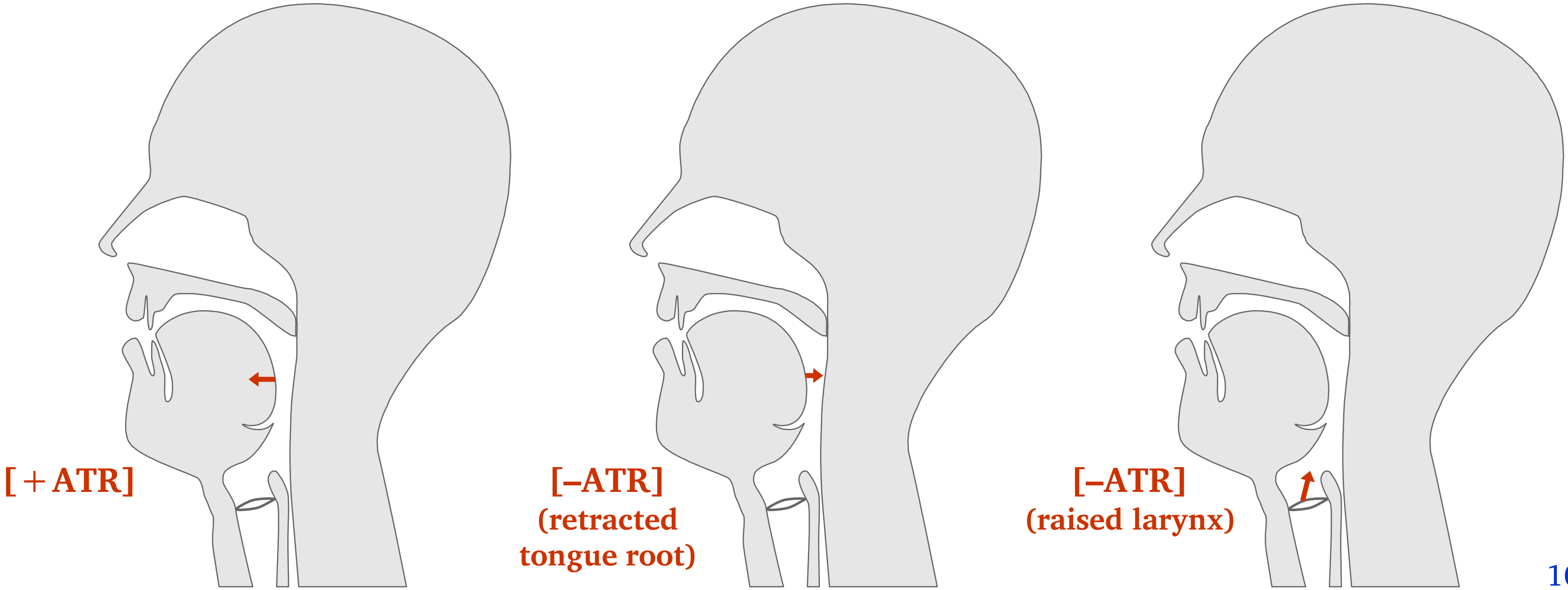
- The articulatory gesture of *roundness* or *lip rounding* affects several formants simultaneously, and its effect depends on the resonant properties of the rest of the vocal tract (i.e., where in the vowel space its unrounded counterpart lies).
  - Generally speaking, the effect of lip rounding is to lower F2, and sometimes F3 and F4 as well.

Vowel Tenseness / Laxness

- An additional acoustic feature (*tense–lax* or *tenseness*) is needed to fully distinguish some pairs of vowels.
  - The tense–lax distinction is effectively a shorthand for “more peripheral/more central”, and as such is a conglomerate of F1/F2 changes.
  - Like roundness, the effect of tenseness depends on where in the vowel space its lax counterpart lies.

Advanced Tongue Root

- Some languages have an additional vowel quality contrast that comes from expanding the pharyngeal cavity.
  - Normally this is done by advancing or retracting the tongue root (represented as [+ATR] and [–ATR] respectively).
  - Alternatively the pharyngeal cavity can be contracted by raising the larynx during the vowel.



Advanced Tongue Root (cont.)

- Because speech sounds are the result of the resonating characteristics of the vocal tract (which includes the pharyngeal cavity), changing the volume of the pharyngeal cavity using either method has a corresponding effect on the vowel quality.
  - Note however, that the pitch and voice quality may change in the larynx raising method, even if the effect on vowel quality is equivalent to a tongue root retraction.
- Some phonologists use [±ATR] to specify differences between tense and lax English vowels, but there is no physiological basis for this classification.

Akan

+ATR

"wash" [sí̘]

"break" [bu̘]

–ATR

"say" [si̙]

"get drunk" [bu̙]

http://archive.phonetics.ucla.edu/

Monophthongs, Diphthongs and Triphthongs

- Vowels which have a single consistent quality are called *monophthongs*, vowels which move between two qualities are *diphthongs,* and three qualities, *triphthongs*.
  - Monophthongs are written with a single IPA symbol, diphthongs are written with two symbols, triphthongs with three.
- Diphthongs are usually described as having a *nucleus* (the main or primary quality) and an *onglide* (before the nucleus) or an *offglide* (after the nucleus).
  - Triphthongs often have both on- and off-glides, with the nucleus in the middle.

Diphthongs in English

- In English diphthongs, onglides and offglides are almost always realized as lax vowels.
  - Consequently many linguists consider English onglides and offglides to be **phonemically** lax (e.g., “pie” /paɪ/, instead of /pai/).
  - Regardless, English onglides and offglides are nearly always **phonetically** lax (e.g., [pʰaɪ]).
- Many linguists transcribe diphthongs as a vowel-glide or glide-vowel sequence (e.g., “cute” /kjut/).
  - Strictly speaking this is incorrect because transcribing /j/ instead of /i/ implies that the glide is behaving phonologically like a *consonant*, and only *vowels* can form diphthongs.
  - This “on-glide” /j/-sound is more accurately transcribed as a true diphthong (/kiut/ or /kɪut/), or as a palatalization of the preceding consonant: [kʰʲut].

Rhotic Vowels

- A *rhotic vowel* is a vowel produced with an "r-like" quality.
  - Rhotic vowels are not necessarily the result of coarticulation with a following /ɹ/ sound; in some languages certain vowels are always rhotic regardless of the following sound.
  - Rhotic vowels are usually transcribed by a small "wing" attached to the vowel symbol: "purr" [pʰɝ], "heard" [hɝd], "murder" [ˈmɝɾɚ]. Another possible analysis is to consider them syllabic consonants: "murder" [ˈmɹ̩ɾɹ̩]
- Acoustically, [ɹ] (or [ɚ]) is very similar to [l], but may have distributional differences, e.g.:
  - "burr" ([bɝ] or [bɹ̩]) vs "bull" ([bʊɫ] but not *[bɫ̩])
  - "her" ([hɝ] or [hɹ̩]) vs "hull" ([hʌɫ] but not *[hɫ̩])

R-coloring

- In contrast to rhotic vowels, *r-coloring* (sometimes called *rhotacization)* is caused by overlap between a normal vowel and a following /ɹ/-like sound.
- Some linguists treat vowels followed by /ɹ/ as diphthongs or triphthongs, rather than vowel-consonant sequences.
  - Transitions are very smooth between vowels and approximants, and r-coloring may extend far into adjacent vowels, making it difficult to identify boundaries.
  - Examples: “are” /ɑɹ/, “err”/ɛɹ/, “or”/ɔɹ/, “ear” /iɹ/, “ire” /aɪɹ/

Stressed and Unstressed Vowels

- In many languages some syllables are more prominent than others. They are often louder and have a different pitch than neighboring syllables, and are usually longer in duration. These are called *stressed syllables*.
- In *unstressed syllables* vowel quality contrasts may become diminished or may merge into a single quality. This is called *vowel reduction*. The quality of reduced vowels is often central, similar to the vowel /ʌ/ found in the word “cup”.
- By convention, reduced vowels are transcribed with the symbol [ə] (called *schwa*). Depending on the language or dialect there may be a variety of reduced vowel qualities that actually occur in unstressed syllables: the lax high front vowel [ɪ] and the high central vowel [ɨ] are common in unstressed syllables for many speakers of English.

Vowel Length

- Some languages have vowels that can be phonemically long or short (i.e., changing vowel length can change word meaning).
  - The difference in duration varies from language to language but long vowels are usually about 1.5-1.7 times as long as their corresponding short vowels.
  - Many languages (e.g., most dialects of English) use vowel length differences as a secondary feature of vowel quality differences.
  - A few languages (e.g., Luiseño) are claimed to have phonemic three-way vowel length contrasts, though it is controversial whether such claims are accurate.
- Length contrast is transcribed by using a symbol resembling a colon made of facing triangles after the vowel ([aː]). A regular colon may be substituted if necessary ([a:]).
  - Example: Estonian [saːta] “to get” vs. [sata] “hundred”.
  - It is common in orthography to use vowel-doubling (e.g., *saata*) or macrons (e.g., *sāta*) to mark long vowels. This is non-standard and should not be done in phonetic transcriptions.

Nasal Vowels and Nasalization

- Because vowels do not rely on oral cavity air pressure (as stops, fricatives, and affricates do), the velum may be open or closed during their production. When the velum is open, nasal cavity resonances are added to oral cavity resonances. Vowels pronounced with the velum open are called *nasal vowels*.
  - Nasal vs. non-nasal vowel contrasts are fairly common. Languages with phonemic nasal/non-nasal contrasts include: French, Polish, Portuguese, Navajo, Bengali, Irish, Southern Min, and Yorùbá.
  - Many languages that do not have phonemic nasal/non-nasal contrasts have vowels that are *nasalised* allophonically (e.g., when followed by a nasal consonant).
  - Nasal vowels and nasalized vowels are marked using a tilde above the vowel symbol: [ẽ õ]

Nasal Vowels and Nasalization (cont.)

- It is thought that nasal/non-nasal contrasts are common because they increase the number of vowel phonemes without increasing the number of oral articulations a child must learn.
  - In this way, nasalized vowels increase the size of the phonemic vowel inventory without increasing its complexity.
  - However, nasalization is known to obscure some vowel quality contrasts that are more easily heard in oral vowels. Therefore many languages have lost certain contrasts among their nasal vowels that are still present in their oral counterparts (a process called *neutralization*).

Voice Quality

- Because vowels are inherently voiced, voice quality contrasts can be phonemic as well. Just as in the consonants, the three voice qualities are:
  - *Normal voice* (or *modal*): [e o æ]
  - *Creaky voice* (or *laryngealized*) marked using a tilde under the vowel: [ḛ o̰ æ̰]
  - *Breathy voice* (or *murmured*) marked using two dots under the vowel: [e̤ o̤ æ̤]
- All languages have modal vowels. Voice quality differences are relatively rare as phonemic contrasts in the world's languages.

Semivowels (Glides)

- Like the liquids, *semivowels* are approximants (narrow constriction, but not enough to generate turbulence). They are very short versions of the high vowels which are really just formant transitions. The most common glides are:
  - the palatal [j] (counterpart of the vowel [i])
  - the labiovelar [w] (counterpart of the vowel [u])
  - the labiopalatal [ɥ] (counterpart of the high front vowel [y])

Consonantal Effects on Vowel Formants

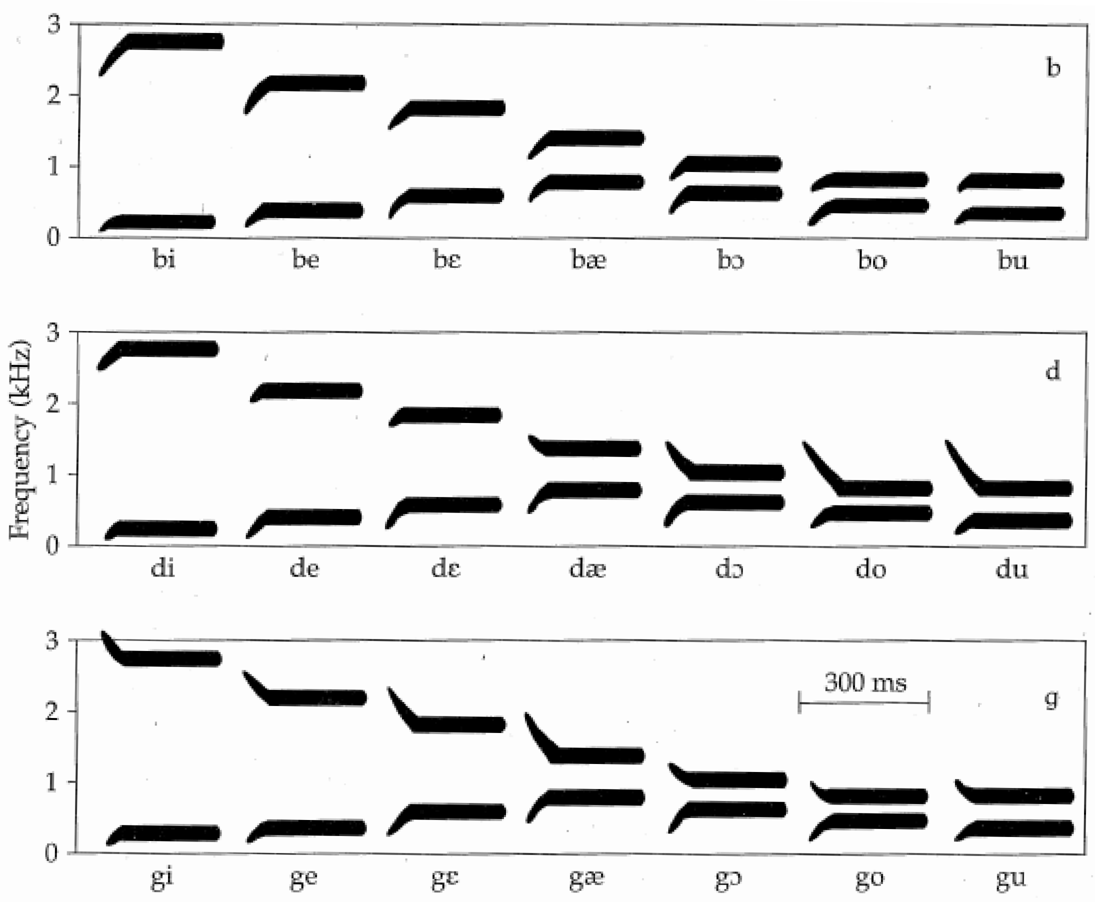


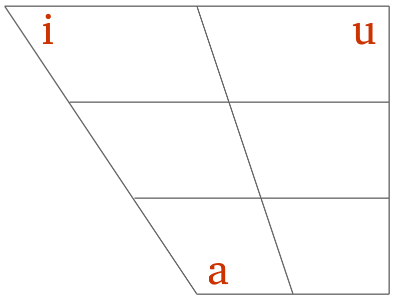
Figure 8.7 F_1 and F_2 transition patterns in stop release used to synthesize [b], [d], and [g] followed by various vowels. Adapted from Delattre et al., 1955, p. 770, and published with permission.

from Johnson, K. (2003). *Acoustic & Auditory Phonetics (2nd ed.)*. Malden, MA: Blackwell.

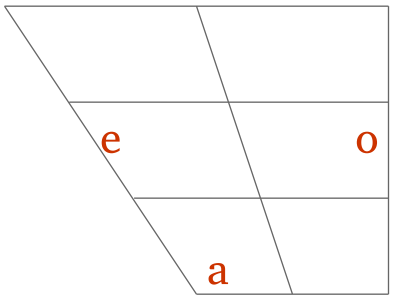
Vowel Systems: Examples

3-vowel Systems

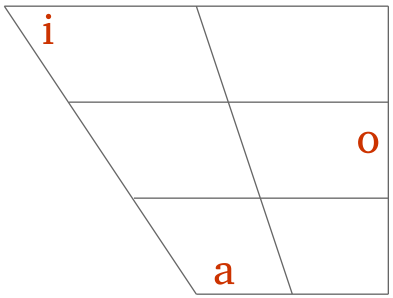
all data from UCLA Phonological Segment Inventory Database (UPSID)



Aleut, Nuxálk, Caddo, Dyirbal, Kalaallisut, Haida, Inuit, Tsimshian

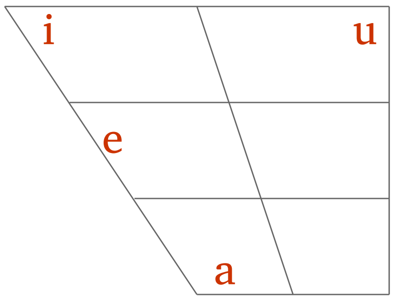


Alabama, Amuesha

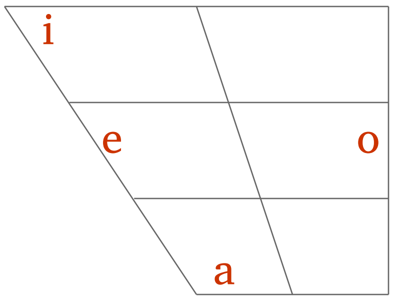


Pirahã

4-vowel Systems

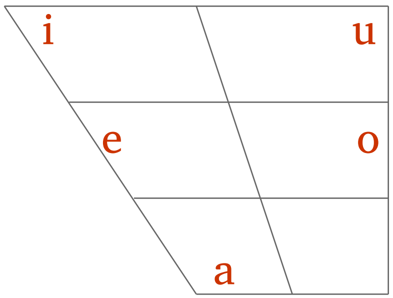


Bandjalang, Moxo, Murihpatha, Shasta



Campa, Klamath, Malagasy, Nahuatl, Tacana

5-vowel System



Ainu, Basque, Hawaiian, Japanese, Luiseño, Nubian, Russian, Spanish, Tagalog, Tlingit, Yaqui, Zulu, Zuni

Reference Chart: Vowels

	Front		Central		Back	
	Unround	Round	Unround	Round	Unround	Round
High Tense	i	y	ɨ	ʉ	ɯ	u
High Lax	ɪ	ʏ	—	—	—	ʊ
Mid Tense	e	ø	ɘ	ɵ	ɤ	o
Mid Lax	ɛ	œ	ə (ɜ)	ɞ	ʌ	ɔ
Low	æ	ɶ	a	ɐ	ɑ	ɒ

Feature values based on Hayes (2009). *Introductory Phonology,* p. 98

Vowel Formants and Duration

Measuring and plotting vowels

See Praat handouts

HW: Vowel Analysis Concepts

Lab: Vowels

Measuring Vowel Duration

- How to measure vowel duration depends on many factors, the most important being:
  - What larger phenomena are you investigating?
  - What sounds are on either side of the vowel?
- The first things to look for are abrupt changes in signal amplitude, changes in periodicity, and the presence of formant structure.
  - When measuring near fricatives, look for abrupt reductions in the intensity of turbulence and the onset of higher formants (F_2, F_3, and sometimes F_4).
  - When excluding the following consonant's closure, look for a marked drop in overall intensity and a loss of energy in the higher formants.
  - When measuring around approximants, it is often extremely difficult to choose a boundary point. Choose a method that minimizes experimenter bias and is suited to the comparisons you plan to make in your analysis.

Measuring Vowel Formants

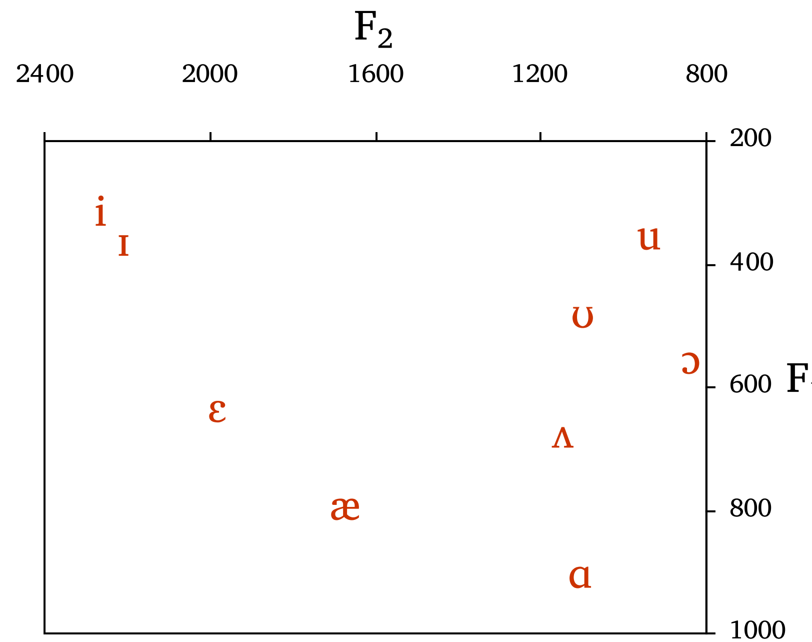
- Since formant frequencies can change through the course of a vowel articulation, it is not always a straightforward task to measure them.
- Monophthongs typically have minimal movement of formant frequencies, usually restricted to the beginning and ending 10% of the vowel's duration.
  - This beginning and ending formant movement is usually the result of coarticulation with the preceding/following consonant.
  - Thus, a measurement anywhere during the *steady state* of the vowel (typically between the 20% and 80% points of the vowel's duration) is usually adequate.

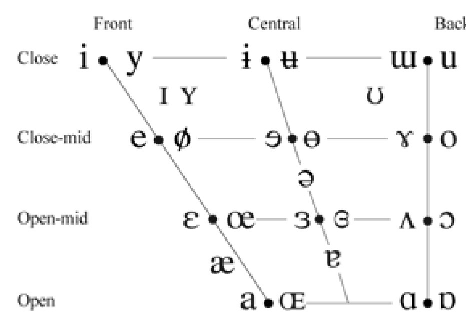
Measuring Vowel Formants - Diphthongs

- Diphthongs involve substantial movement of the formant frequencies, and often do not show an obvious beginning and ending steady state.
  - It is common to consider the 20% point to be the *nucleus* of the diphthong and the 80% mark to be the *offglide*.
  - The 50% point is also sometimes measured to give a sense of how quickly the diphthong moves from nucleus to offglide.
- When comparing monophthongs to diphthongs, they should usually be compared at the 20% mark for both vowels.

Plotting Vowel Formants

- The main features used to describe vowels are *high/low* and *front/back*.
  - Early linguists treated height and backness as physical measures of tongue position during vowel articulation.
  - Today these features represent differences in the F_1 and F_2 values for each vowel:
    - High vowels have low F_1 values.
    - Back vowels have low F_2 values.
- To maintain the historical charting standards of *high = up* and *back = right*, we plot the formant values so they increase downwards and leftwards.





Formant data source: http://alt-usage-english.org/Fontana_formant_numbers.html

Consonants: Place of Articulation

Location of vocal tract constriction

Read: LJ 7 “Articulatory Targets” (up to p. 172)

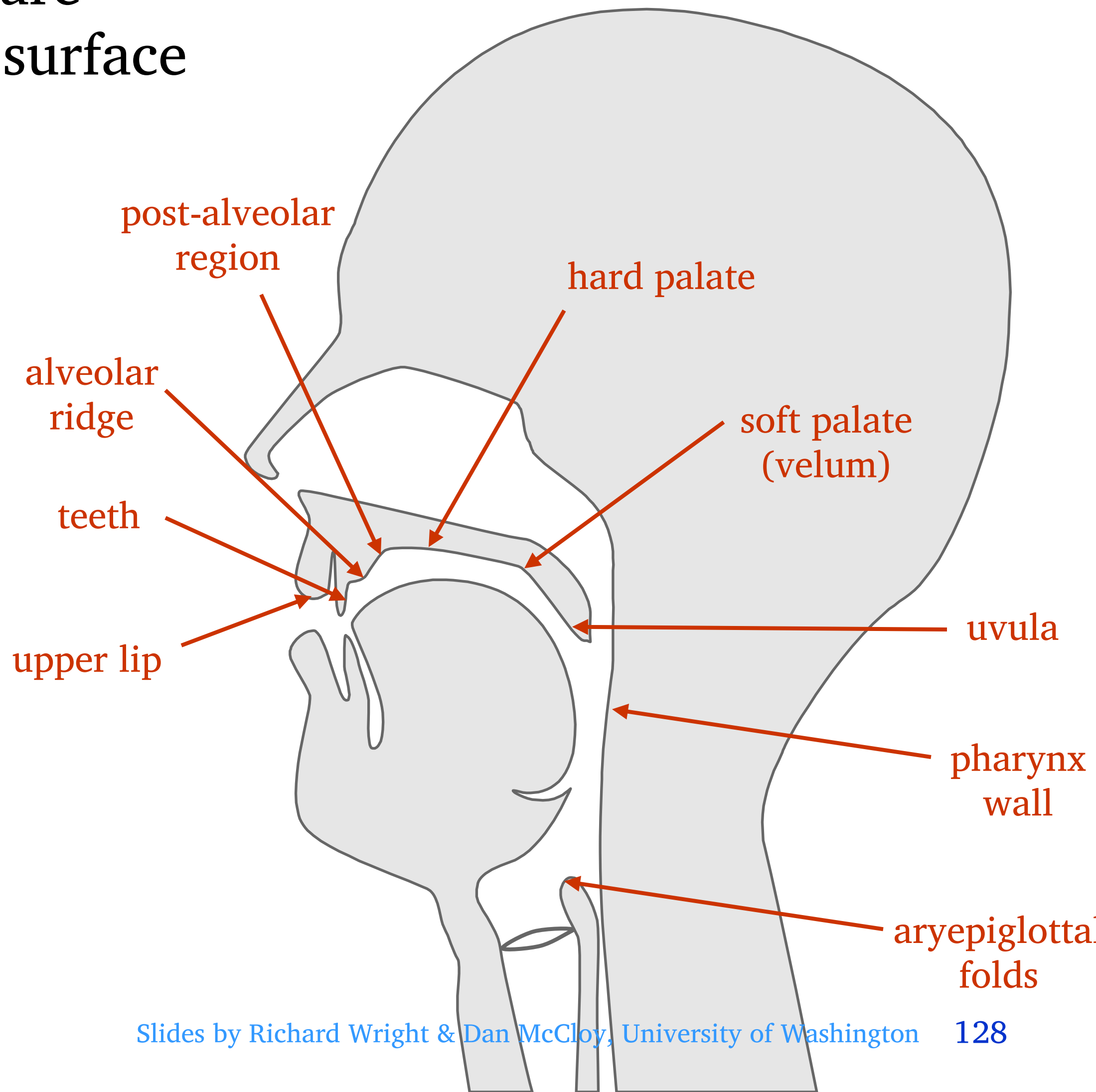
Begin HW: IPA (All consonants - articulation)

Describing Consonants

- All consonants involve a constriction made by an *active articulator* toward a target region (the *passive articulator*).
- Consonants are usually described with two main terms:
  - *Place of articulation* describes which active and passive articulators are making the constriction.
  - *Manner of articulation* describes the way in which the articulators interact (i.e., how close together they are).
  - *Voicing* is often left unstated for approximants, nasals, and trills, since they are (almost) always voiced.
- As discussed in previous classes, there are a variety of airstream mechanisms and states of the glottis that may be contrastive on consonants as well.

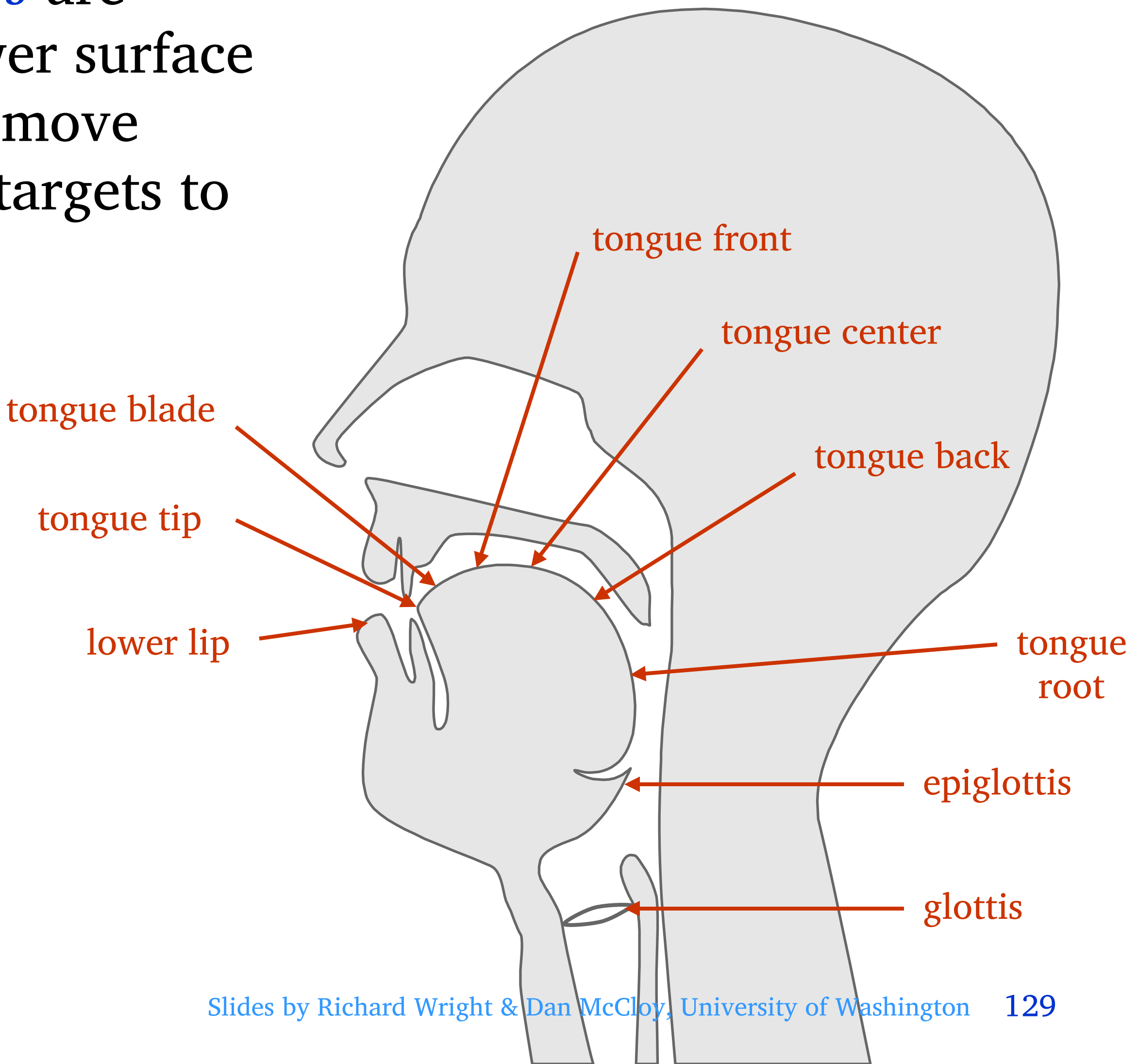
Passive Articulators

- Most *passive articulators* are located along the upper surface of the oral tract.



Active Articulators

- Most *active articulators* are located along the lower surface of the oral tract, and move toward their passive targets to create a constriction.

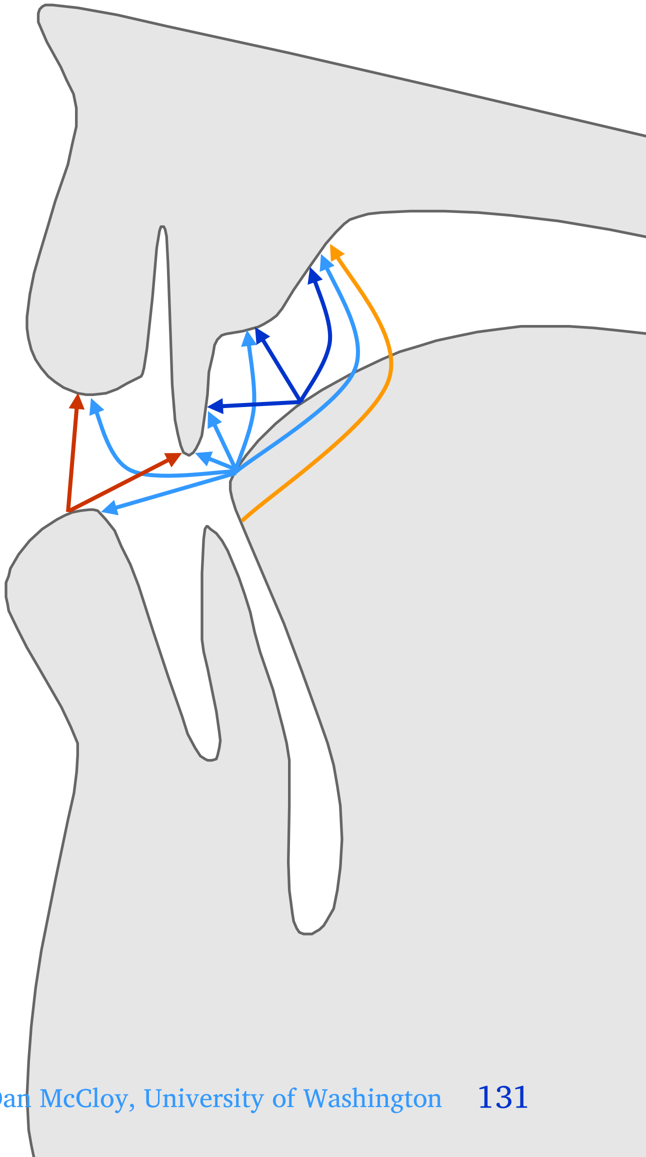


Terminological Conventions

- Descriptions of place of articulation usually refer to the passive articulator (e.g., an *alveolar* stop).
- Other times, both the active and passive articulators are mentioned (e.g., a *labiodental* fricative). This occurs most frequently at the front of the mouth, where a target may be shared by several active articulators.
- *Epiglottal* and *glottal* sounds are the only sounds described with their active articulators only.

Anterior Places of Articulation

- In the anterior (front) region of the mouth, the primary active articulators are: the lower lip (*labial articulations*), tongue tip (*apical articulations*), and tongue blade (*laminal articulations*).
- There are four main anterior targets: the upper lip, upper incisors, alveolar ridge, and the post-alveolar region.
  - Note in particular two possible (though unusual) articulations: tongue tip to lower lip, and underside of tongue to alveolar ridge (a *retroflex* articulation).
- These definitions contain the necessary detail for describing subtle differences between languages, but not all of the definitions are contrastive (or even attested) in any one language. Thus phonological analyses often use a less detailed set of definitions for within-language descriptions.



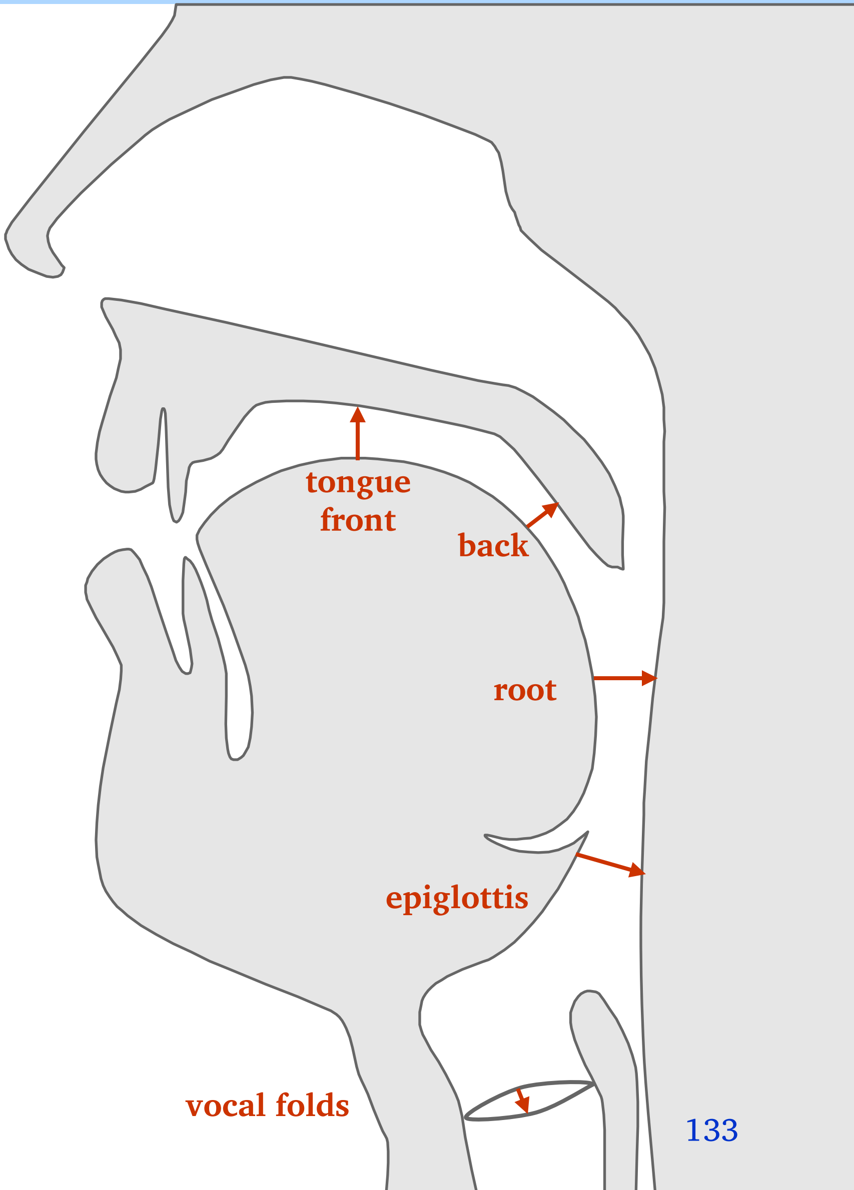
Reference Chart:
Anterior Places of Articulation

	Passive Articulator	Active Articulator	IPA Examples
Bilabial	upper lip	lower lip	[p b m ɸ β]
Linguolabial		tongue tip	[t̼ d̼ n̼]
Labiodental	lower edge of upper incisors	lower lip	[ɱ f v]
Interdental		tongue tip/blade	[t̟ d̟ n̟ θ ð]
Apical Dental	back side of upper incisors	tongue tip	[t̺ d̺ n̺]
Laminal Dental		tongue blade	[t̪ d̪ n̪]
Apical Alveolar	alveolar ridge	tongue tip	[t d n s z]
Laminal Alveolar		tongue blade	[t̻ d̻ n̻]
Apical Retroflex	post-alveolar region	tongue tip	[ṭ ḍ ṇ ṣ ẓ]
Palato-Alveolar		tongue blade	[t̠ d̠ n̠ ʃ ʒ]
Sub-Apical Retroflex		tongue underblade	[ʈ ɖ ɳ ʂ ʐ]

Chart based on Ladefoged and Maddieson (1996) p15.

Posterior Places of Articulation

- In the posterior (back) region of the mouth, the active articulators are the tongue front, back, and root, the epiglottis, and the vocal folds.
- The main posterior targets are the hard palate, velum, uvula, and the pharyngeal wall.
- Note that a glottal stop does not really involve any passive articulators, as the vocal folds close against each other.



Reference Chart:
Posterior Places of Articulation

	Passive Articulator	Active Articulator	IPA Examples
Palatal	hard palate	tongue front	[c ɟ ɲ ç ʝ]
Velar	velum	tongue back	[k g ŋ x ɣ]
Uvular	uvula	tongue back	[q ɢ ɴ χ ʁ ʀ]
Pharyngeal	(upper) pharynx	tongue root	[ħ ʕ]
Epiglottal	(lower) pharynx	epiglottis	[ʡ ʢ ʜ]
Glottal	—	vocal folds	[ʔ h ɦ]

Chart based on Ladefoged and Maddieson (1996) p15.

Reference Chart: Places of Articulation

	Passive Articulator	Active Articulator	IPA Examples
Bilabial	upper lip	lower lip	[p b m ɸ β]
Linguolabial		tongue tip	[t̼ d̼ n̼]
Labiodental	lower edge of upper incisors	lower lip	[ɱ f v]
Interdental		tongue tip/blade	[t̟ d̟ n̟ θ ð]
Apical Dental	back side of upper incisors	tongue tip	[t̪̺ d̪̺ n̪̺]
Laminal Dental		tongue blade	[t̪̻ d̪̻ n̪̻]
Apical Alveolar	alveolar ridge	tongue tip	[t d n s z]
Laminal Alveolar		tongue blade	[t̻ d̻ n̻]
Apical Retroflex	post-alveolar region	tongue tip	[ṭ ḍ ṇ ṣ ẓ]
Palato-Alveolar		tongue blade	[t̠ d̠ n̠ ʃ ʒ]
Sub-Apical Retroflex		tongue underblade	[ʈ ɖ ɳ ʂ ʐ]
Palatal	hard palate	tongue front	[c ɟ ɲ ç ʝ]
Velar	velum	tongue back	[k g ŋ x ɣ]
Uvular	uvula	tongue back	[q ɢ ɴ χ ʁ ʀ]
Pharyngeal	(upper) pharynx	tongue root	[ħ ʕ]
Epiglottal	(lower) pharynx	epiglottis	[ʡ ʢ ʜ]
Glottal	—	vocal folds	[ʔ h ɦ]

Chart based on Ladefoged and Maddieson (1996) p15.

Doubly Articulated Consonants

- Many sounds involve simultaneous constriction at more than one place of articulation. For example, the approximant [w] involves a velar constriction as well as a labial constriction. Other examples:
  - voiceless labio-velar fricative [ʍ] (a.k.a. [w̥])
  - voiceless postalveolar-velar fricative [ɧ]
  - voiced labio-palatal approximant [ɥ] (a.k.a. [jʷ])
  - voiced/voiceless alveolo-palatal fricatives [ɕ ʑ]
- If a pre-made symbol does not exist, double articulations can be indicated by a tie bar
  - Example: [ʒ͡ɣ]

Place Terms Review

- *Anterior* – made in front of mouth (tongue blade forward)
  - *Apical* – active articulator = tip of tongue
    - *Retroflex* – active articulator = underside of tongue tip
  - *Laminal* – active articulator = tongue blade
- *Posterior* – made in back of mouth (body of tongue backward)
  - *Dorsal* – active articulator = back of tongue (tongue *dorsum*)

Consonants:
Manner of Articulation

Degrees and types of vocal tract constriction

Read: LJ 7 from “Types of Articulatory Gestures” (p. 172-end)

Finish HW: IPA (All consonants - articulation)

Manner of Articulation Defined

- The main determiner of a consonant's *manner of articulation* is the degree of *constriction* (or *stricture*) between the active articulator and the target region.
  - There are also several manner distinctions that are made based on the style of constriction or accompanying articulations or combinations of manners.
- There are three basic degrees of constriction that are *contrastive* (i.e., are *phonemic*) in the world's languages: *stop*, *fricative*, and *approximant*.

Stops

- A *stop* involves complete closure of the oral tract so that the airstream cannot escape through the mouth.
- Stops often come in voiced/voiceless pairs and may be made with a variety of airstream mechanisms, e.g.:
  - Pulmonic egressive stops (*plosives*)
  - Glottalic egressive stops (*ejectives*)
  - Glottalic ingressive stops (*implosives*)
  - Velaric ingressive stops (*clicks*)
- Stops may be accompanied by the full range of phonation types (voiceless, breathy, creaky, or modal).
- Examples: [p pʰ p’ b bʱ b̰ ɓ kʘ] (bilabial stops)

Fricatives

- *Fricatives* involve nearly complete closure of the oral tract but with a narrow opening. Forcing all the air through a narrow opening accelerates the rate of airflow and creates *turbulence* (also called *frication*). Frication is audible and is what gives fricatives their “hissing” quality.
  - Fricatives often come in voiced-voiceless pairs within a language.
  - Pulmonic fricatives can be accompanied by the full range of phonation types.
  - Many fricatives can be made with the glottalic egressive (ejective) airstream mechanism. Exceptions are the lower pharyngeal sounds [ħ ʕ] and the glottalic sounds [h ɦ].
- Examples of fricatives (given in voiceless/voiced pairs):
  - [ɸ β f v θ ð s z ʃ ʒ ʂ ʐ ç ʝ x ɣ χ ʁ ħ ʕ h ɦ]

Notes: Sibilant Fricatives

- Fricatives can be divided by their auditory properties:
  - *Sibilants* are louder, with more energy in higher frequencies
    - Examples: [s z ʃ ʒ ʂ ʐ] and affricates w/ sibilants: [ts dz tʃ dʒ]
    - English treats sibilants as a group when adding sibilant morphemes (e.g. [s] or [z] for plural, possessive, or 3rd person singular verb) – two sibilants can't be pronounced adjacently in the same syllable

punt	[pʌnt]	→	*punts*	[pʌnts]
punch	[pʌntʃ]	→	*punches*	[pʌntʃəz]

  - *Non-sibilants* (everything else) are quieter, with less energy
    - Examples: [ɸ β f v θ ð ç ʝ x ɣ χ ʁ ħ ʕ h ɦ]
    - English doesn't restrict the combination of a non-sibilant + sibilant

puffs	[pʌfs]	*myths*	[mɪθs]
gives	[gɪvz]	*bathes*	[beɪðz]

Approximants

- *Approximants* involve a narrowing of the oral tract but not enough to create turbulence in the airstream.
- Approximants are always voiced and always made with a pulmonic egressive airstream mechanism. They may be accompanied by the three voiced phonation types (modal, creaky, breathy).
- Examples: [ʋ l ɹ ɻ ʟ ɰ w]

Affricates

- *Affricates* are a combination of two sounds: a stop released into a (usually) *homorganic* (same place of articulation) fricative with the same voicing and phonation type. Like fricatives, affricates can be either pulmonic egressive or glottalic egressive (ejective) sounds.
- Examples: [pf bv ʦ ʣ ʧ ʤ kx gɣ]
  - Note: some fonts include *precomposed digraphs* for the more common affricates (e.g., [ʤ]), but all affricates can also be written as a sequence of the separate stop and fricative symbols (e.g., [dʒ]).
  - If you wish to emphasize that two sounds form an affricate and not separate sequential sounds, you may use a *tiebar* (e.g., [b͜v] or [g͡ɣ]).

Oral Stops vs. Nasal Stops

- As mentioned, stops involve momentary complete stoppage of airflow through the oral tract. However, airflow may pass through the nasal cavity if the velum is lowered. Such articulations are called *nasal stops* (in contrast to *oral stops* in which the velum is raised, closing off the *velo-pharyngeal port*).
  - Generally, oral stops are simply called “stops” and nasal stops are called “nasals.”
- Nasals may be voiced or voiceless and may be accompanied by modal, creaky, or breathy phonation types.
  - Nasal stops: [m̥ m n̥ n ɳ̊ ɳ ɲ̊ ɲ ŋ̊ ŋ ɴ̥ ɴ]
- Stops may also be either pre-nasalized or post nasalized.
  - Prenasalized stops: [mb nd ɳɖ ɲɟ ŋg ɴɢ]
    Postnasalized stops: [bm dn ɖɳ ɟɲ gŋ ɢɴ]

Trills, Taps, and Flaps

- *Trills* are generated by the airstream causing an articulator to vibrate, creating a series of rapid closures separated by brief moments of airflow past the constriction.
  - Trill vibration is caused by the Bernoulli force: airflow over a semi-rigid articulator pulls the articulator toward the target (causing a closure) at which point pressure builds behind the closure (forcing it open again).
  - Examples: [ʙ r ʀ]
- *Taps* and *flaps* are ballistic: the active articulator is thrown towards and bounces off the target, resulting in a very brief stop-like closure. Taps move (more or less) straight toward their target; flaps move back-to-front (usually retroflex).
  - Examples: [ⱱ ɾ ɾ̃ ɽ ɺ]

Notes: Nasal Tap

- *Nasal alveolar tap* [ɾ̃]
  - In American English, tapped /n/ occurs in the same places as where /t, d/ are tapped (after a stressed syllable, before reduced vowel or syllabic consonant).
  - Examples: *sinner, dinner, any, bunny, sunny, pony, kennel, funnel, penny, Jenny*, etc. (compare to how Forrest Gump says "Jenny" with a full or long [n] closure duration).

Notes: Labiodental Flap

- *Voiced labiodental flap* has a new symbol in IPA: [ⱱ]
  - The Ladefoged & Johnson book is out of date when they say there is no IPA symbol for this. They use * in Table 7.6 and describe it at the end of the Trills, Taps, and Flaps section (p. 177-178).
  - Tuck bottom lip behind upper teeth, flap against them.
  - Occurs mainly in African languages, e.g. Kera:

[ⱱīw]	‘see something pass quickly’
[vìw]	‘hearing something pass by’
[wīw]	‘something attached
[pə̄ⱱāw]	‘escape quickly’
[pə̄wāw]	‘noise in forest’

Central vs. Lateral Consonants

- Most speech sounds are *central* (they result from constriction along the center of the vocal tract).
- *Lateral* sounds involve an obstruction at a point along the center of the vocal tract but with incomplete closure along one or both sides of the tongue. Laterals may be approximants, fricatives, affricates, or flaps.
  - Lateral approximants: [l ɭ ʎ ʟ]
  - Lateral fricatives: [ɬ ɮ]
  - Lateral affricates: [tɬ dɮ]
  - Lateral flap: [ɺ]
  - Lateral click: [k‖]

Airstream Mechanisms

Types of airflow and corresponding speech sounds

Read: LJ 6 “Airstream Mechanisms” (up to p. 147)

HW: IPA (implosives ejectives clicks)

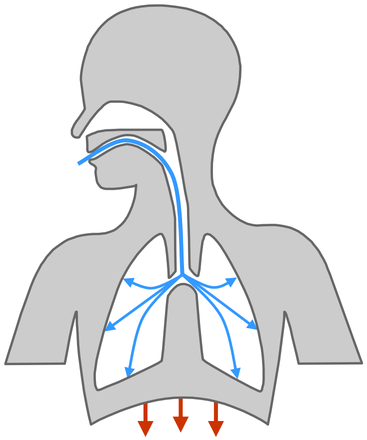
Transcription: Non-English Airstream

Types of Airflow

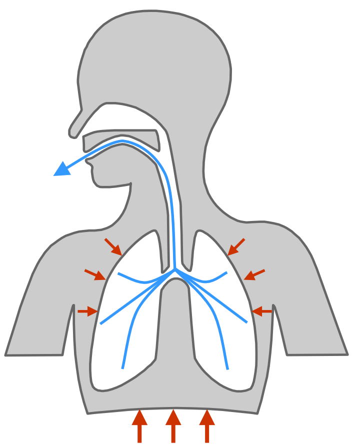
- All speech sounds are vibrations in the air, and require airflow for their production. There are many ways to move air into or out of the vocal tract.
  - The *pulmonic* airstream mechanism uses the lungs to create airflow.
  - The *glottalic* airstream mechanism uses the up-and-down movement of the glottis to compress or rarefy the air trapped in the vocal tract between the closed glottis and an oral closure.
  - The *velaric* airstream mechanism uses the tongue to compress or rarefy air trapped in the vocal tract (usually between the tongue and the hard palate).
  - *Egressive* airflow is air moving out of the vocal tract; *ingressive* airflow is air moving into the vocal tract.

The Pulmonic Airstream Mechanism

- Inhalation
  - The diaphragm flexes down, creating an area of low pressure in the lungs
  - the lungs passively expand as air moves in through the vocal tract
  - The ribcage expands to accommodate the lungs' expansion



- Exhalation
  - The diaphragm relaxes upwards
  - abdominal muscles push the diaphragm from below and the intercostal (ribcage) muscles pull the ribcage in and down, squeezing air out of the lungs through the vocal tract

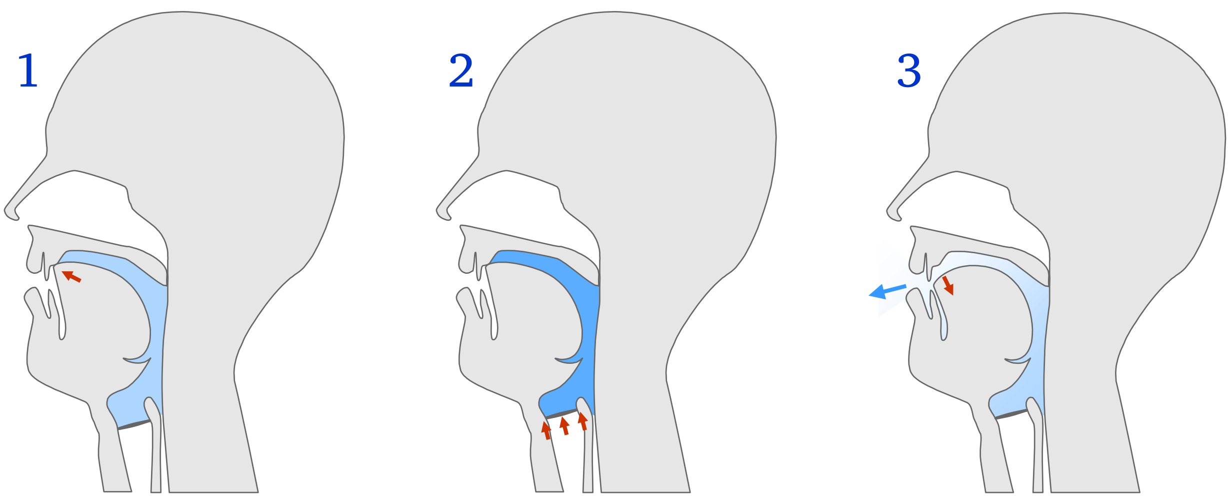


Pulmonic Sounds

- Pulmonic Egressive Sounds
  - All human languages have pulmonic egressive sounds.
  - All vowels, glides, and nasals are made via the pulmonic egressive airstream mechanism.
  - In many languages (e.g., English) all speech sounds are pulmonic egressives.
- Pulmonic Ingressive Sounds
  - These are not phonemic in any known language, though in some cultures ingressive sounds are used for paralinguistic purposes (e.g., back-channeling in French, expressing shock in English)

The Glottalic Airstream Mechanism

- Glottalic Egressive Sounds (Ejectives)
  - Air trapped between the glottis and an oral closure is *compressed* by *raising* the larynx.
  - When the oral closure is released, a burst of compressed air exits the oral cavity, generating the speech sound.
  - All ejectives are voiceless, and are transcribed with a diacritic added to the symbol for the corresponding pulmonic voiceless consonant: / p’ t’ k’ /



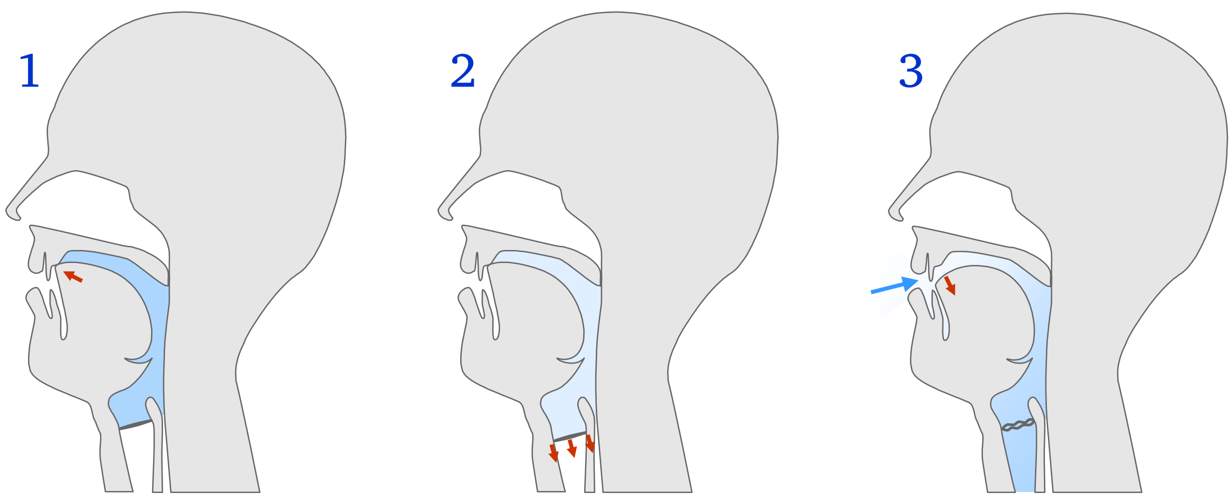
Quechua

Plain	
“tongue”	[gaʎu]
“bridge”	[ʧaka]
Aspirated	
“shawl store”	[q^{h}aʎu]
“large ant”	[ʧ^{h}aka]
Ejective	
“tomato sauce”	[q’aʎu]
“hoarse”	[ʧ’aka]

http://archive.phonetics.ucla.edu/

The Glottalic Airstream Mechanism

- Glottalic Ingressive Sounds (Implosives)
  - Air trapped between the glottis and an oral closure is *rarefied* by *lowering* the larynx.
  - Implosives are almost always voiced; usually voicing happens as the larynx moves down. In such cases rarefaction is slight because air flows through the glottis as the vocal folds vibrate.
  - Implosives are transcribed like their voiced pulmonic stop equivalents but with a right-pointing upper hook: / ɓ ɗ ɠ /



Sindhi

Aspirated "snake hood"	[p^{h}aɳu]	
Voiceless "leaf"	[panu]	
Voiced "forest"	[banu]	
Implosive "field"	[ɓani]	

http://archive.phonetics.ucla.edu/

The Velaric Airstream Mechanism

- Velaric Ingressive Sounds (Clicks)
  - uses rarefaction of trapped air between velar and anterior closures (labial, dental, alveolar, palatal) to create speech sounds.
  - Clicks can be voiced, voiceless, aspirated, unaspirated, nasal, oral, and may be accompanied by a variety of secondary articulations (e.g., lip rounding).
  - Examples: / ʘ ‖ ! /

!Xóõ

Bilabial
"get stuck" [kʘʔôo]

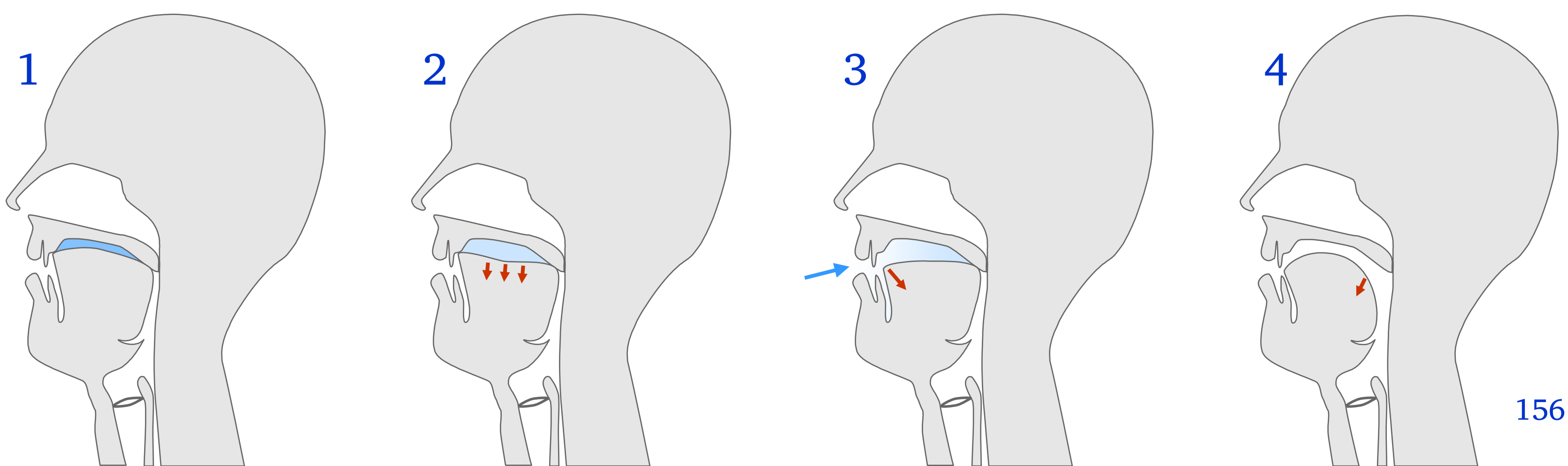
Dental
"die" [k|ʔâa]

Alveolar
"be seated" [k!ʔáã]

Alveolar Lateral
"not to be" [k‖ʔàa]

Palatal
"shoot you" [kǂʔāa]

http://archive.phonetics.ucla.edu/



The Velaric Airstream Mechanism

- Velaric Egressive Sounds
  - Though it is possible to produce these sounds, they are not attested in any known language.

Transcribing Clicks

- Since all clicks involve two places of articulation, they are typically transcribed with two symbols:
  - [k g ŋ] (or [q ɢ ɴ] if the rear articulation is uvular) to represent the voicing, place and manner of the posterior closure, and
  - one of [ʘ | ‖ ! ǂ] to represent the place of the anterior closure. Examples:
  - k| = voiceless dental click
  - g‖ = voiced lateral click
  - ŋʘ = nasal bilabial click
  - Note: The (post)alveolar click can be described as alveolar and the palatoalveolar as palatal.

Clicks	
ʘ	Bilabial
|	Dental
!	(Post)alveolar
ǂ	Palatoalveolar
‖	Alveolar lateral

Transcribing Clicks - Notes

- Some recent research suggests that many clicks are made with a uvular constriction rather than a velar one. Because of this, some phoneticians prefer the term "lingual ingressive" rather than "velaric ingressive" to describe the airstream mechanism.
- Note that, though not an official IPA symbol, occasionally you may see the anterior portion of the dental click transcribed as [ʇ] instead of [|], especially in cases where it helps avoid confusion with the prosodic marker [|]

Name That Click! (Zulu)

• "a conversation"	íŋˈŋ‖oːk‖o	nasal lateral; voiceless lateral
• "to make clear"	ˈk|àːk|á	voiceless dental
• "to pound"	ˈg‖oːbá	voiced lateral
• "to undo"	k!ák!á	voiceless alveolar
• "to be loose"	ˈk‖èːgá	voiceless lateral

Phonation

States of the glottis and Voice Onset Time

Read: LJ 6 from “States of the Glottis” (p. 148)

HW: Obstruent Dynamics (voicing, aspiration, VOT)

HW: IPA (All vowels - articulation, phonation)

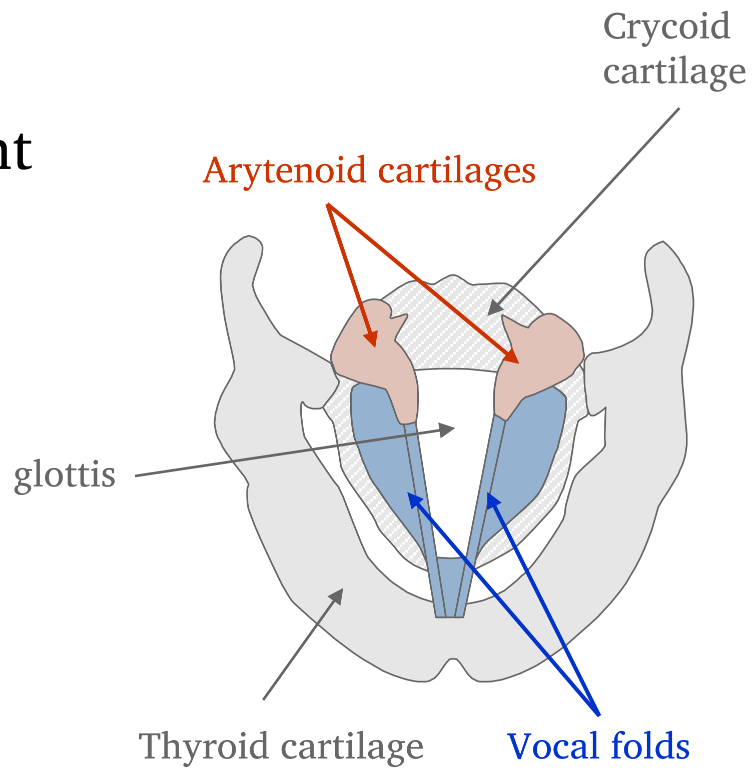
HW: IPA (all consonants - articulation, phonation, airstream)

Transcription: Non-English Phonation

Lab: VOT

Anatomy of the Larynx

- The *vocal folds* are drawn together or apart by the movement of the *arytenoid cartilages*.
- The *glottis* is the space between the vocal folds through which air passes.
- The position of the vocal folds at right corresponds to a *voiceless* sound, e.g., [p t k f θ s ʃ h m̥ l̥]



Three Types of Phonation – Breathy

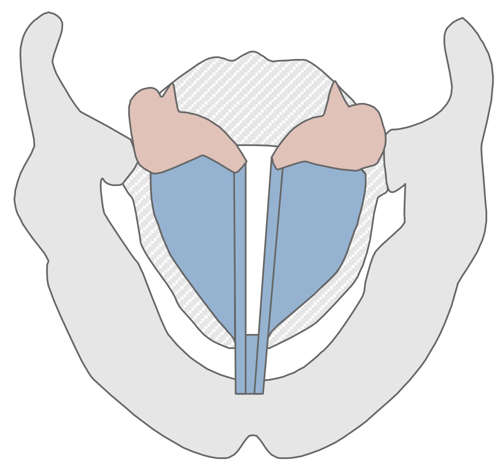
- *Breathy Voice (Murmur)*
  - Vocal folds vibrate, but are spread further than normal voicing.
  - Often the vocal folds never fully close during vibration, which lends an [h]-like quality to sounds.
  - Any sound that can be voiced can have breathy voicing.

Mazatec

Modal
"for a while" [t^hǽ]

Breathy
"horse" [^{n}dǽ̤]

http://archive.phonetics.ucla.edu/



Three Types of Phonation – Breathy

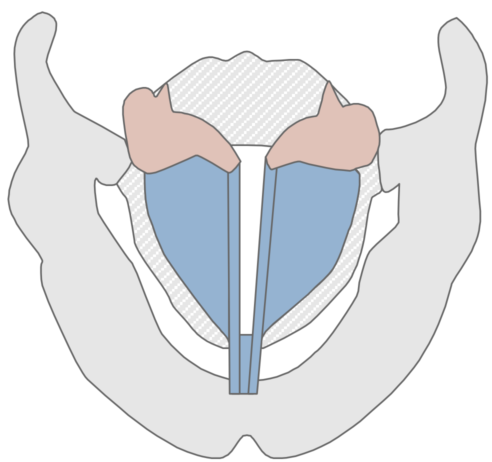
- Transcription of breathy sounds:
  - The "breathy h" has its own symbol: [ɦ]
  - Other breathy sounds are transcribed with a diacritic:
    [e̤ z̤ n̤ d̤ l̤]
  - Breathy stops (but only stops) can be transcribed with a superscript diacritic instead: [$d^{ɦ}$].
  - Some authors use the aspiration symbol with voiced stops to indicate breathy voicing (e.g., [d^h]). This is not compliant with IPA standards and should not be done.

Mazatec

Modal
"for a while" [t^hǽ]

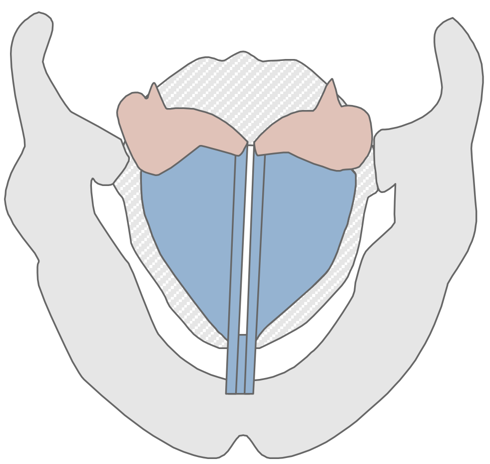
Breathy
"horse" [^{n}dǽ̤]

http://archive.phonetics.ucla.edu/



Three Types of Phonation – Modal

- *Normal Voice (Modal)*
  - Vocal folds are held closely together and vibrate regularly along their full length, creating moments of full closure alternating with moments of airflow
  - Unlike breathy sounds, there is no accompanying [h]-like sound in modal voicing
  - Examples: [b d g m n ŋ v z ʒ w ɹ l i u ɑ]



Three Types of Phonation – Creaky

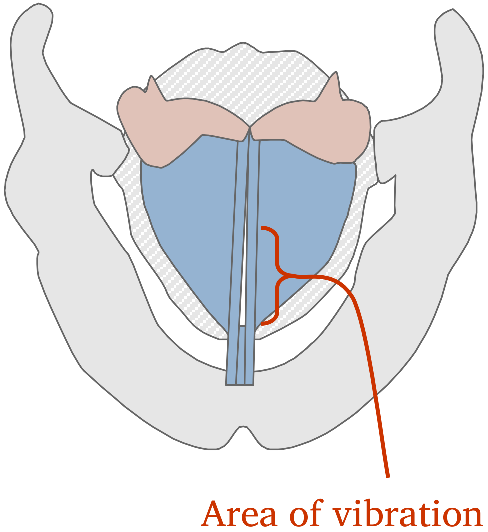
- *Creaky Voice (Laryngealization)*
  - The vocal folds are held tightly together, such that they vibrate irregularly and only at the thyroid end, along roughly half their length.
  - Any sound that can be voiced can have creaky voicing.
- Transcription of creaky sounds:
  - All creaky voiced sounds are transcribed with the same diacritic (an under-tilde). Examples: / d̰ n̰ z̰ l̰ w̰ ḛ /

Mazatec

Modal
"for a while" [tʰǽ]

Creaky
"becomes" [nǣ̰]

http://archive.phonetics.ucla.edu/

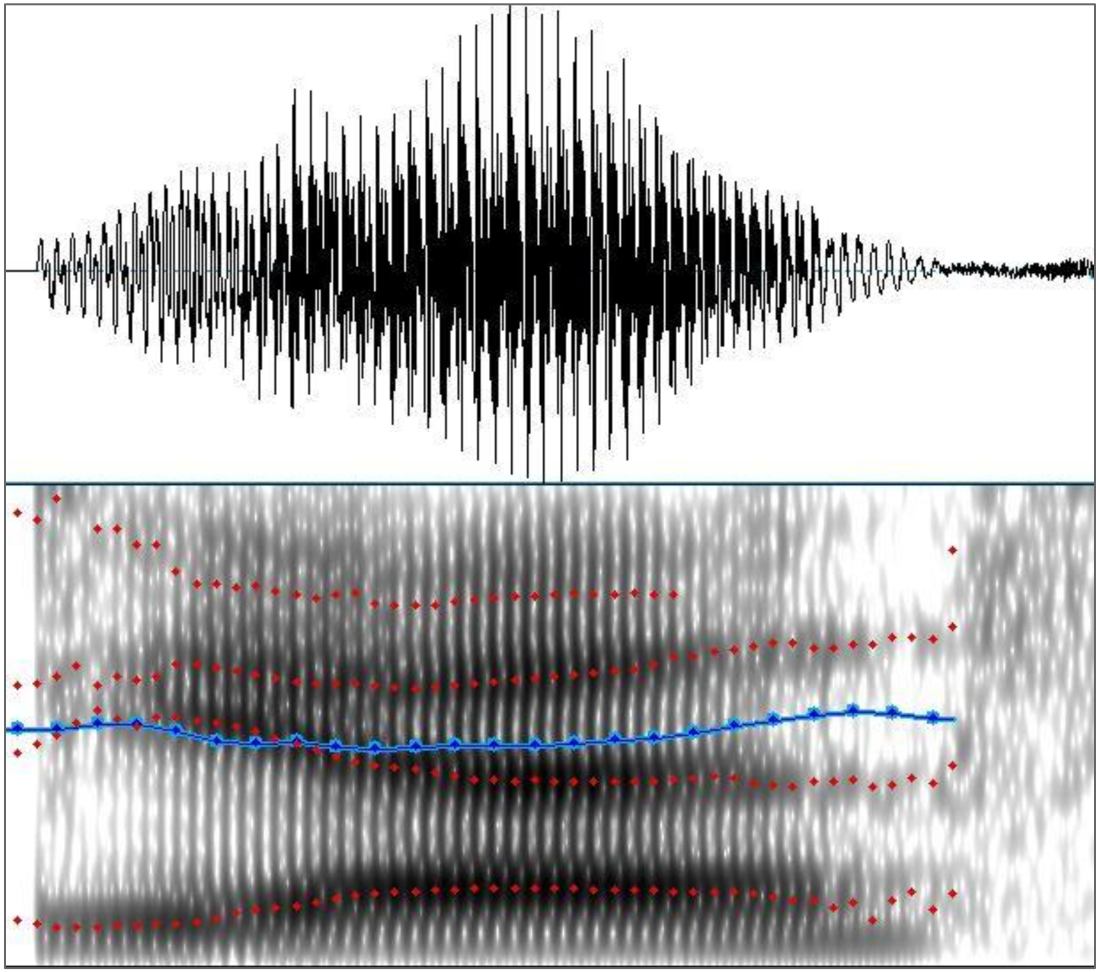


The Open Quotient

- One way to think about the different phonation types is to consider the percentage of time the vocal folds are open during a given vibrational cycle.
  - In voiceless articulations, the vocal folds are open 100% of the time
  - In breathy articulations, the vocal folds are open 66% of the time
  - In modal articulations, the vocal folds are open 50% of the time
  - In creaky articulations, the vocal folds are open 33% of the time

Phonation Examples – Modal

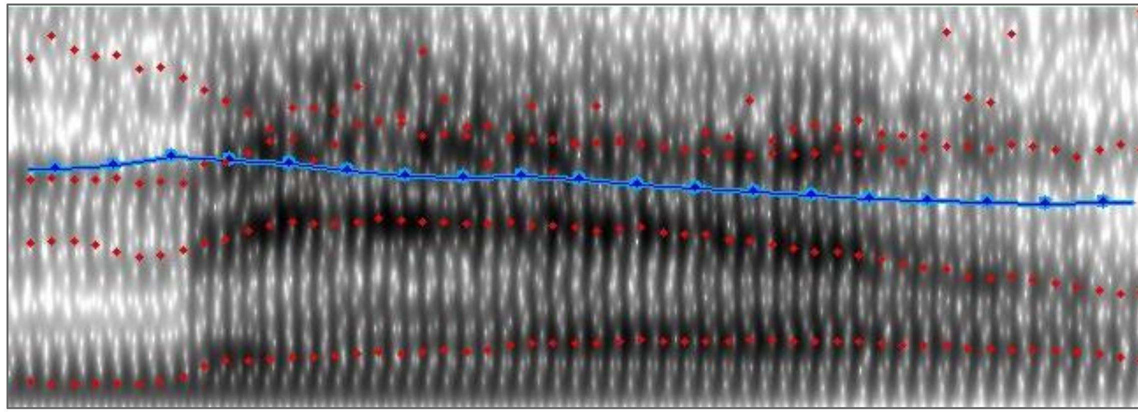
- Modal voicing on a spectrogram:
  - Regular spacing between pulses (space = silence = vocal folds are closed about half the time)
  - Pitch (blue line) and formants (red dots) track well



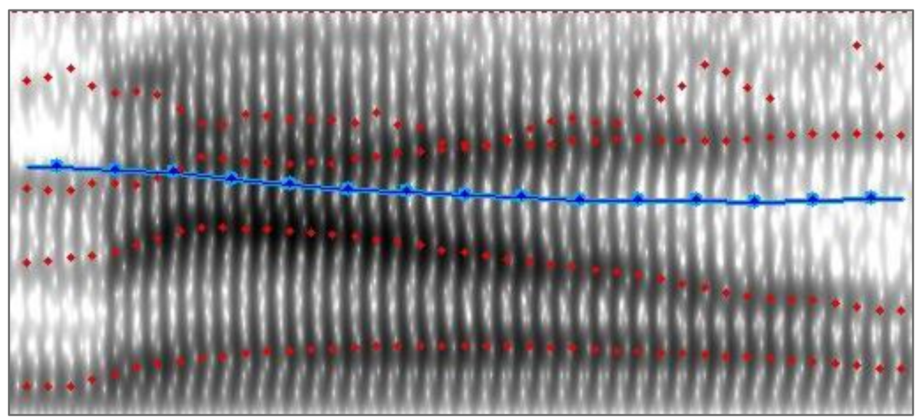
[ɛ] in "yes" said by PNW female

Phonation Examples – Breathy

- Breathy voicing on a spectrogram:
  - Pulses closer together than for modal, spaces less clear (less space = vocal folds are closed less of the time)
  - Pitch (blue line) and formants (red dots) track well for both



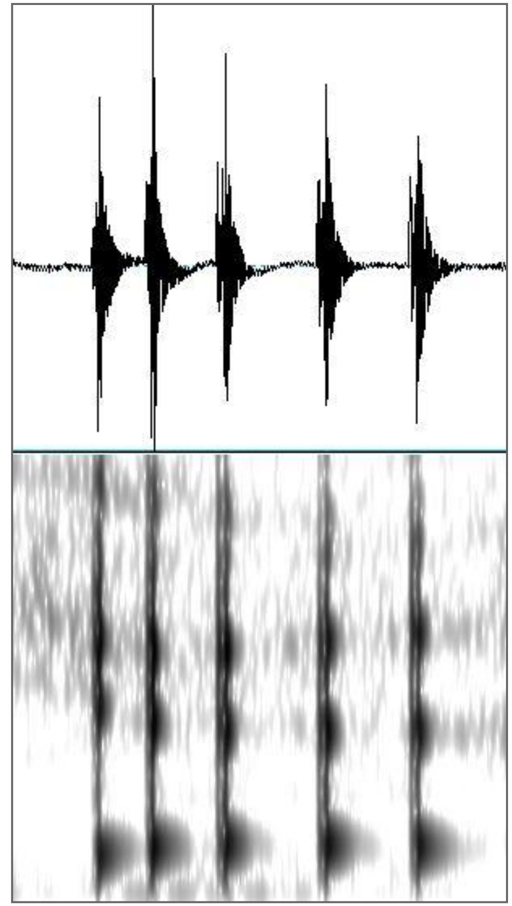
Breathy: [mɛ̤l] ‘palace’ in Gujarati



Modal: [mɛl] ‘dirt’ in Gujarati

Phonation Examples – Creaky

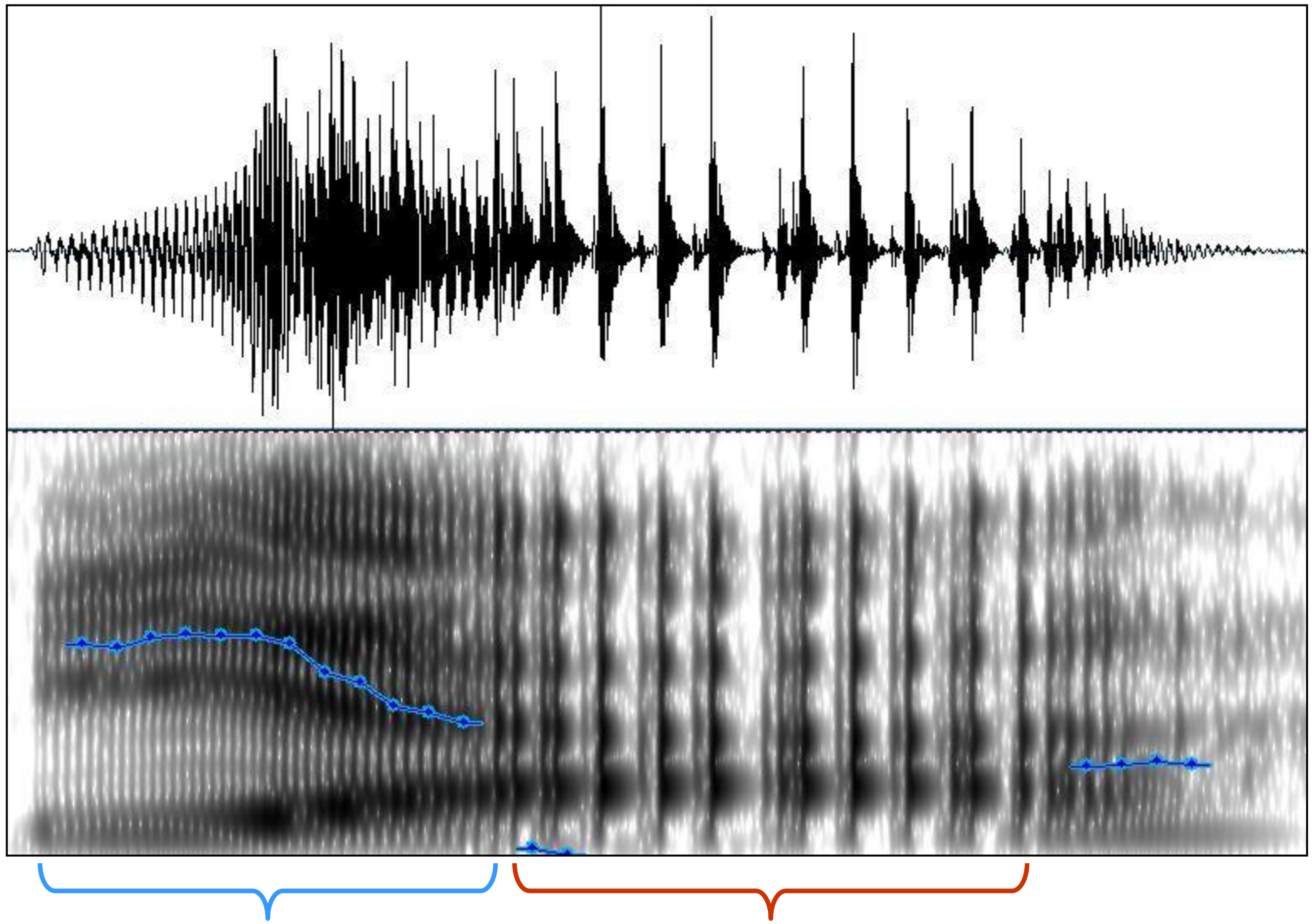
- Creaky voicing on a spectrogram:
  - Sharp spikes (glottal pulses) in waveform
  - Wide, irregular spacing between pulses (space = folds are closed)
  - Often causes errors in pitch tracking, sometimes formant tracking (not shown)



[ɛ̰] in "yes" said by PNW female (extremely creaky)

Phonation Examples

- Spacing between pulses:
  - Even
  - Wide, irregular
- Pitch tracking:
  - Working
  - Lost



modal creaky

[jɛ̰ə] "yeah" said by PNW female

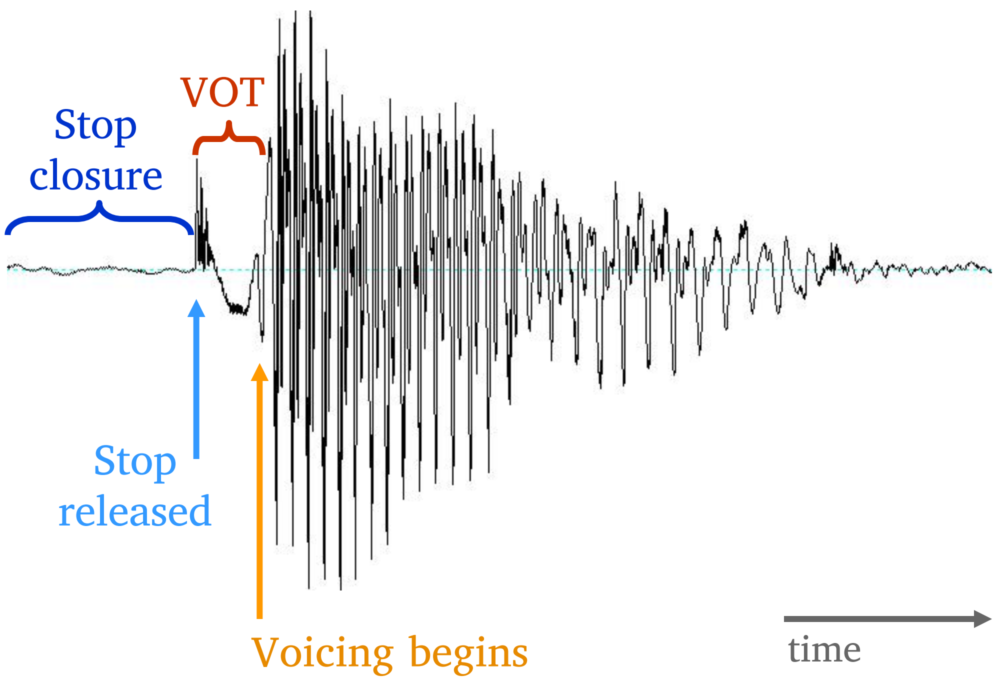
Links: Laryngeal Settings Videos

Watch videos of the larynx at various settings.

- Glottals and Pharyngeals: http://web.uvic.ca/ling/research/phonetics/lands.htm
- Pharyngeals summary: http://web.uvic.ca/ling/research/phonetics/jipa96.htm
- States of the glottis: http://web.uvic.ca/ling/research/phonetics/SOG/index.htm
- Various voice quality distinctions: http://cw.routledge.com/textbooks/9780415498791/chapter11.asp

Voice Onset Time (VOT) Defined

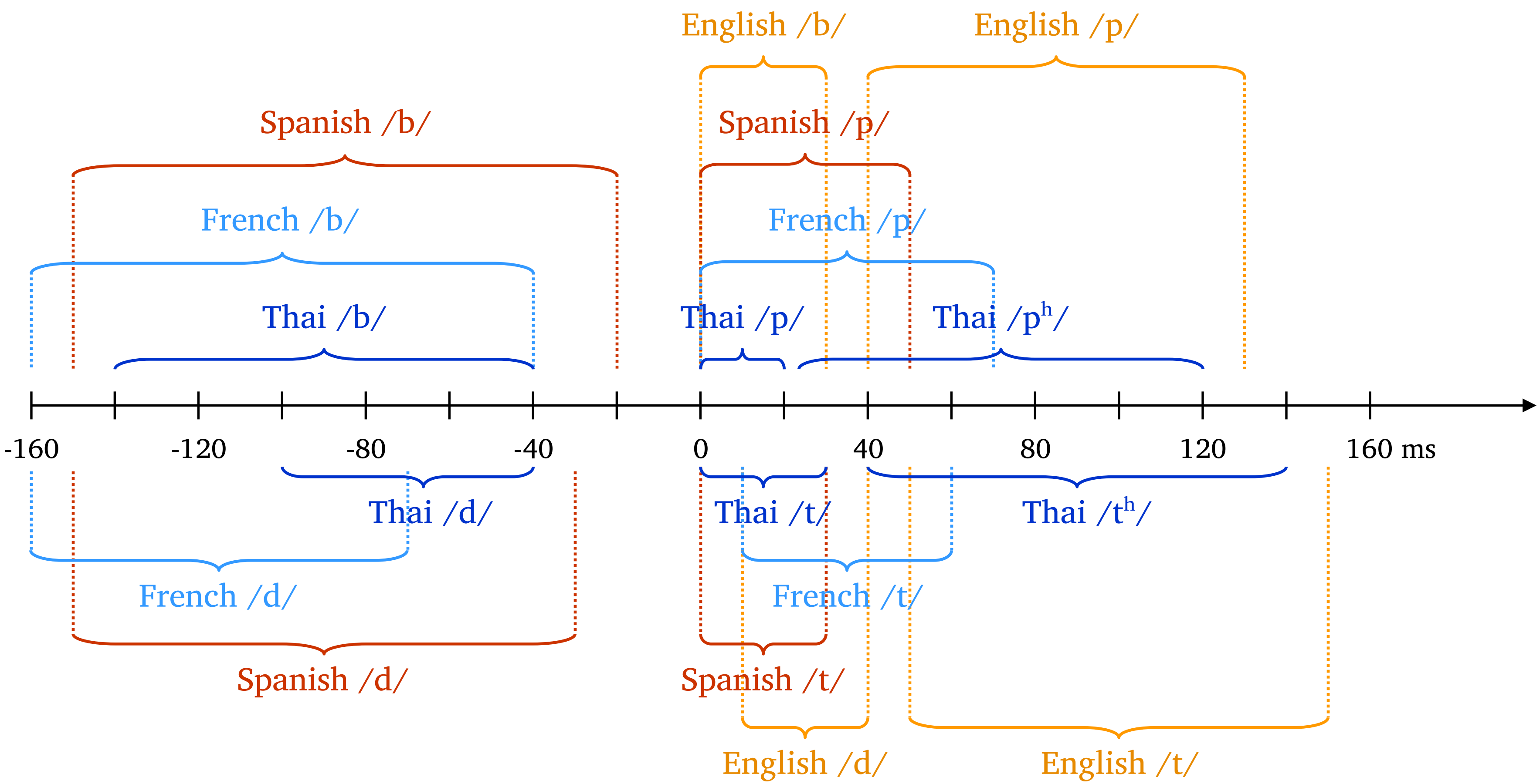
- *Voice onset time (VOT)* is the time between the release of a stop closure and the beginning of vocal fold vibration.
- VOT is the basis for our perception of whether or not a stop is voiced, voiceless, or aspirated.



Variation in Voice Onset Time

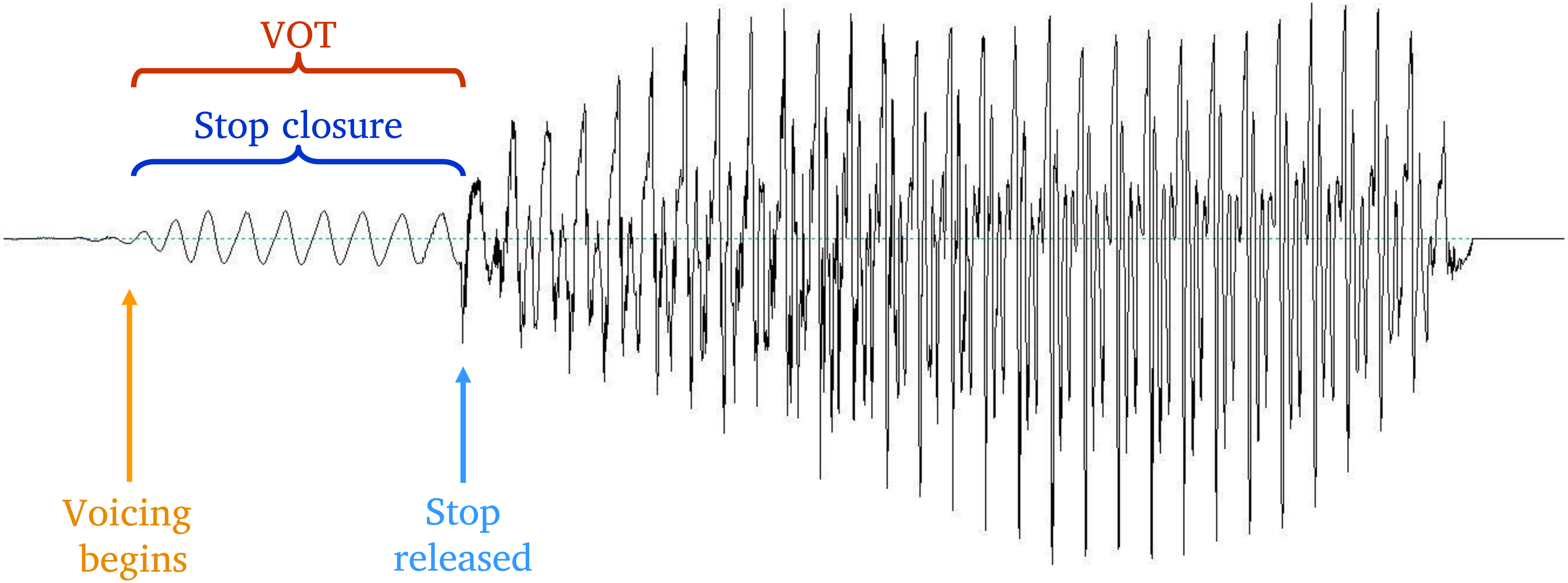
- Negative VOT
  - voicing begins before the stop is released (e.g., voiced stops in French and Thai).
- Short VOT
  - voicing begins almost immediately after the stop is released (e.g., voiced stops in English).
- Intermediate VOT
  - voicing begins shortly after the stop is released (e.g., voiceless unaspirated stops in French, Thai, and English).
- Long VOT
  - voicing begins long after the stop is released (e.g., voiceless aspirated stops in Thai, and the English aspirated allophone).

Cross-Linguistic Comparisons of VOT



Sources:
Rosner, B.S. et al. “Voice-onset times for Castilian Spanish initial stops.” *Journal of Phonetics* (2000) 28, 217–224.
Kessinger, R.H. and Blumstein, S.E. “Effects of speaking rate on voice-onset time in Thai, French, and English.” *Journal of Phonetics* (1997) 25, 143–168.

Estimating VOT from Waveforms

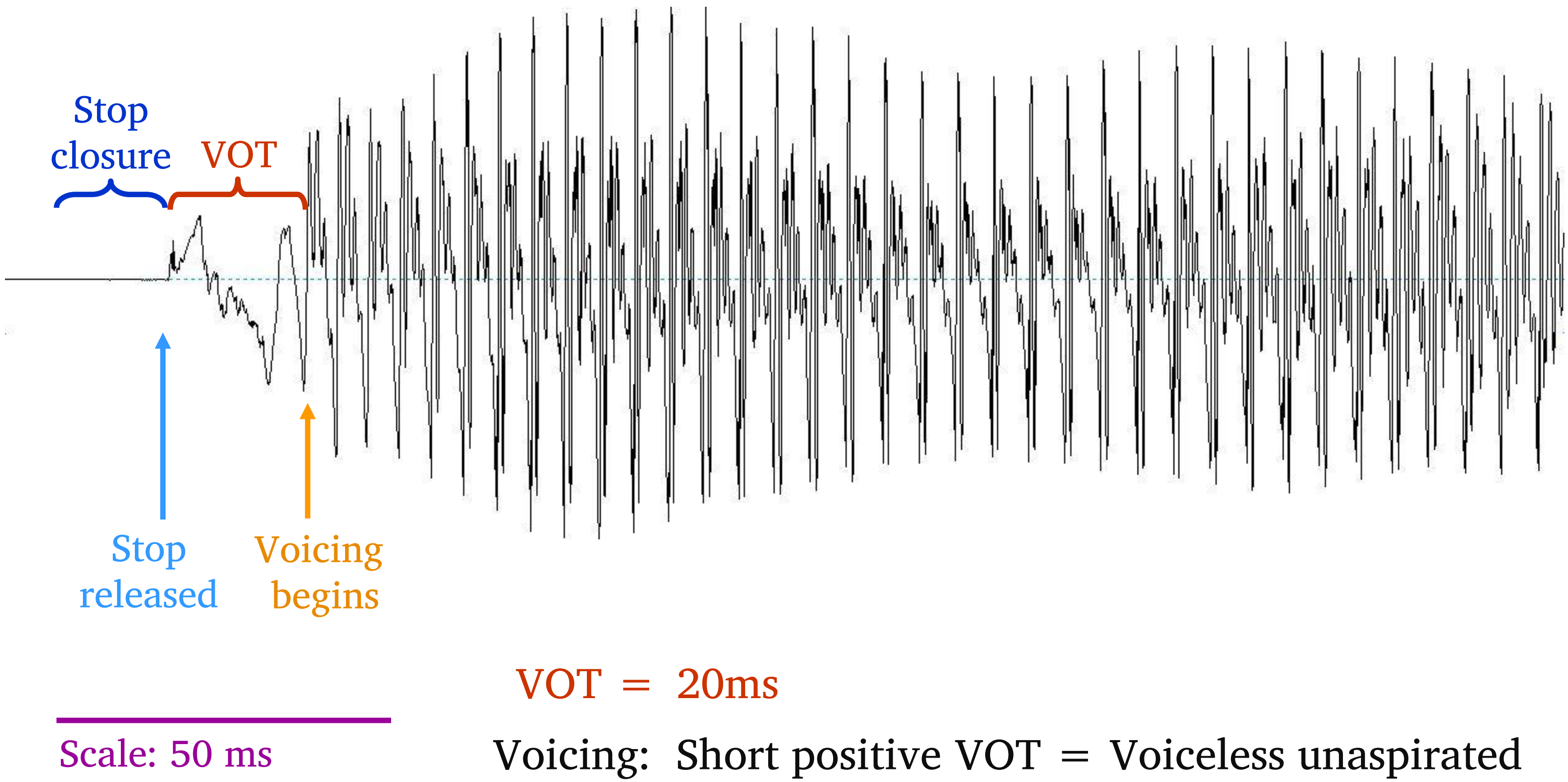


VOT = -50ms

Scale: 50 ms

Voicing: Negative VOT = Voiced

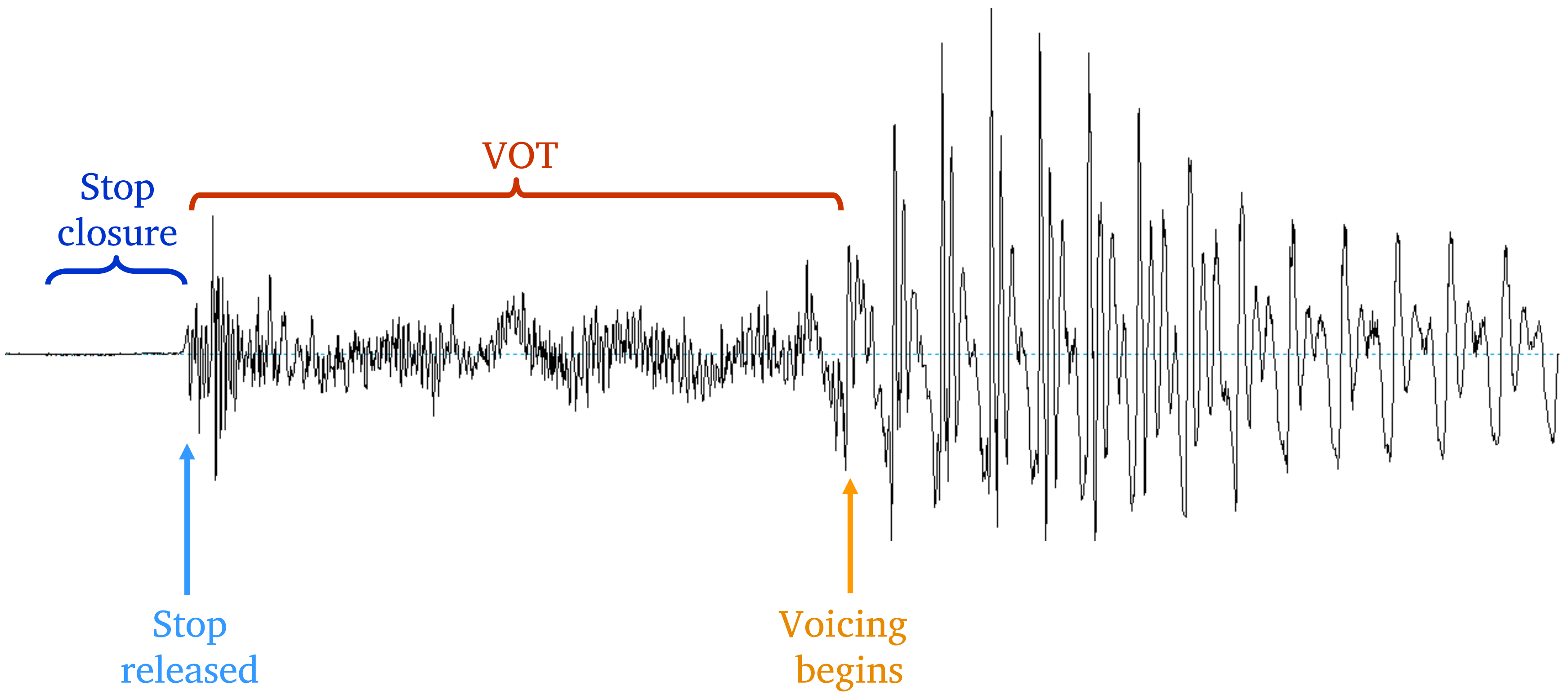
Estimating VOT from Waveforms



VOT = 20ms

Voicing: Short positive VOT = Voiceless unaspirated

Estimating VOT from Waveforms



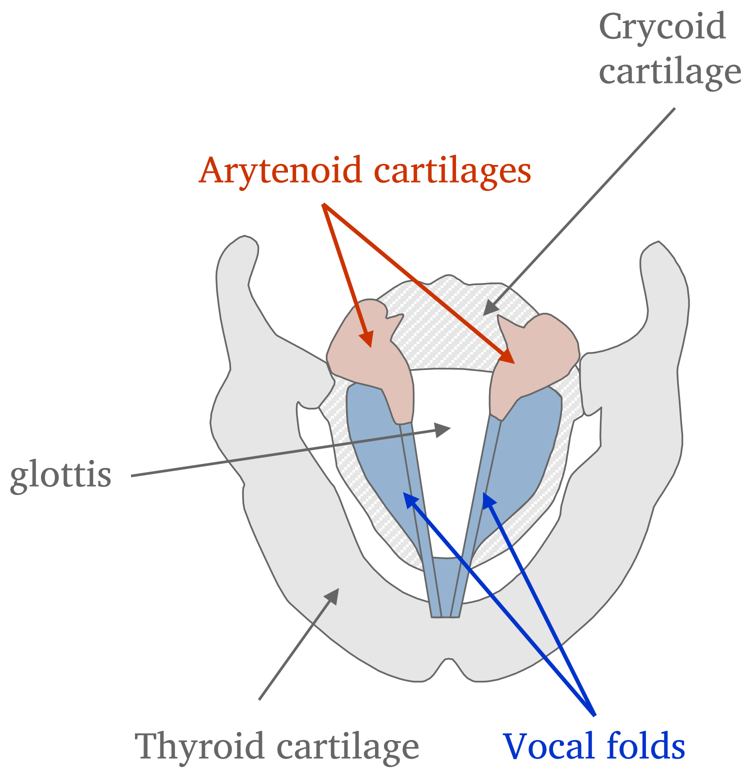
VOT = 95ms

Scale: 50 ms

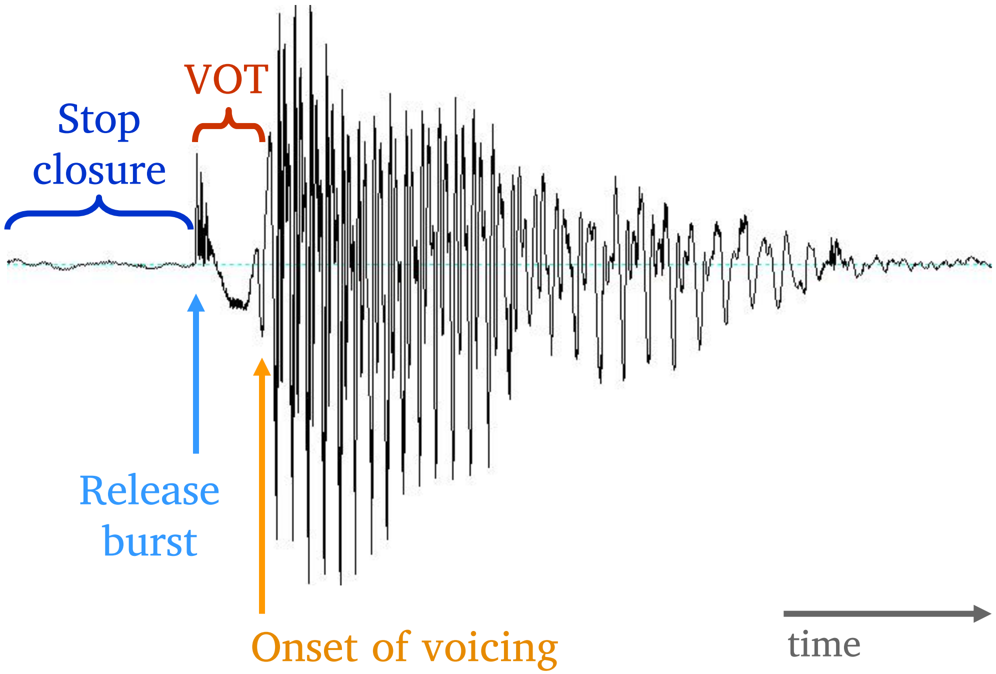
Voicing: Long positive VOT = Voiceless aspirated

Review: Anatomy of the Larynx

- Clockwise from the bottom right corner…



Review: Voice Onset Time (VOT)



Review: Voice Onset Time (VOT)

- Modal voiced stops have __negative__ VOT.
- Voiceless unaspirated stops have _short positive_ VOT.
- Aspirated stops have _long positive_ VOT.
- Breathy voiced stops (sometimes called "voiced aspirates") have _(long) negative_ VOT.

Syllables

Groups of phones

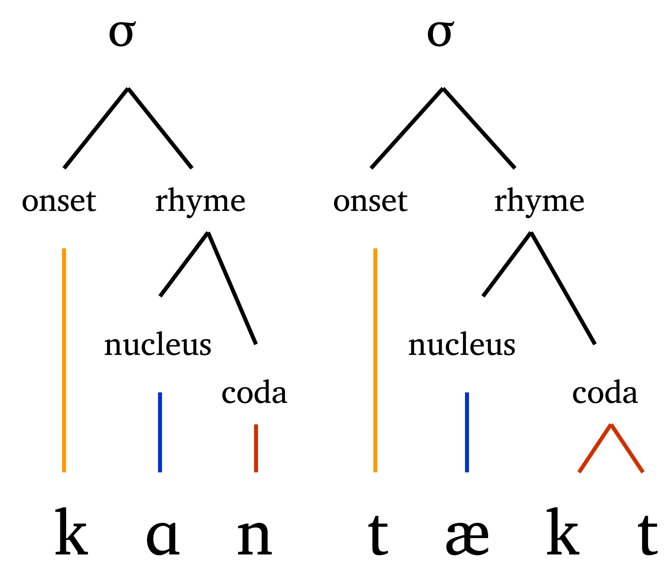
Read: LJ 10 “Syllables” (up to p. 249)

The Phenomenon of the Syllable

- Syllables are remarkable for their near-universal psychological reality:
  - Virtually all speakers can intuitively divide words in their native language into syllables.
  - Speakers of the same language will agree about syllable counts for the vast majority of words in their language.
  - Speakers will usually agree about where the boundaries between syllables lie within a word.
- Nonetheless, there are words that people syllabify differently, and a theory of the syllable must account for this variation.

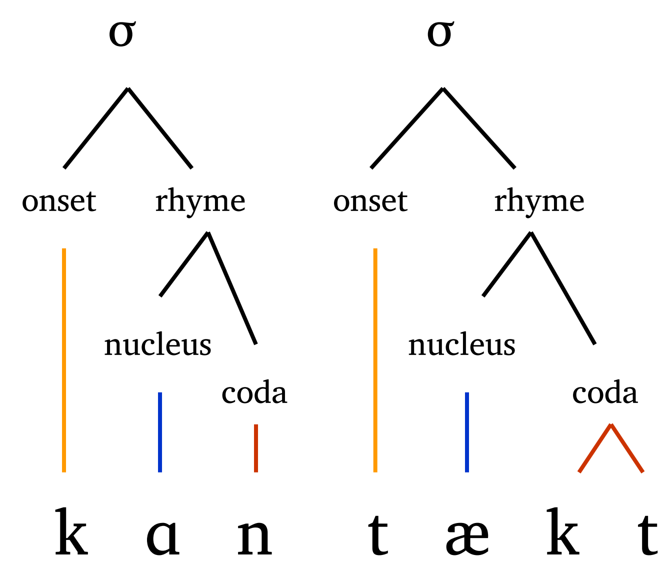
The Tree Model of the Syllable

- Syllables are often subdivided into a hierarchical structure comprising an *onset* and a *rhyme** with the rhyme further subdivided into a *nucleus* and a *coda*.
  - Syllables are commonly represented by the Greek letter σ (lower-case sigma).
  - *Also spelled *rime*



Syllable Composition

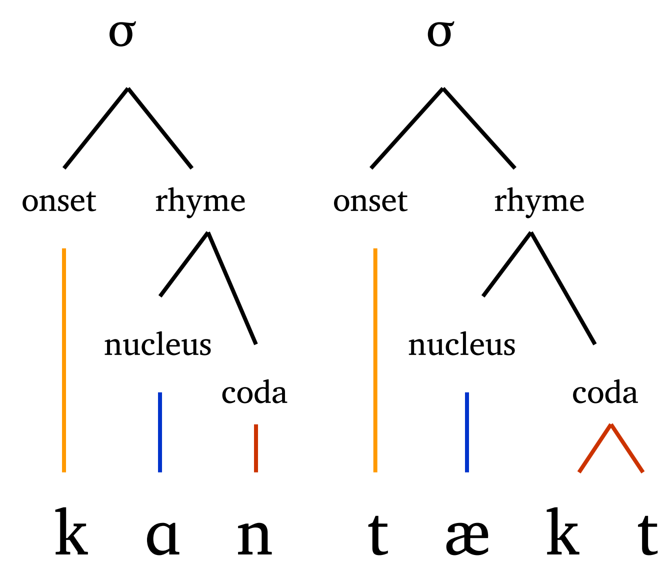
- Every syllable consists of at least a nucleus (and therefore a rhyme), but onsets and codas are optional.
- Languages tend to prefer syllables with onsets and disprefer codas.
- The inventory of allowable coda consonants and *clusters* (groups of adjacent consonants) is usually more restricted than for onsets.
  - e.g., only nasal codas allowed



- Any segment can be a nucleus, but every language has vocalic nuclei. Many also allow *sonorant consonants* (nasals and liquids (laterals and rhotics)), and a few further allow *obstruents* (non-sonorants).

Syllable Weight

- *Heavy syllables* have branching rhymes. *Light syllables* don't. Onsets don't affect weight, even if they are *complex* (branching).
- Phonological rules often apply based on syllable weight (light, heavy, super-heavy).
  - e.g., Stress falls on the first heavy syllable from the end.
  - What counts as a super-heavy syllable varies by language, if the distinction is important to the phonology.
  - Languages may exclude some segments from adding weight.



Theories of the Syllable

- The precise definition of *syllable* remains controversial even today, though the various attempts can be categorized as a few basic types of explanations:
  - Syllables reflect *mental rules about allowable patterns in the speech signal* (e.g., patterns of sonority or prominence).
  - Syllables are *physiological units* (such as the sounds produced by a single chest pulse, or the sounds associated with a single unit of speech planning in the brain of the speaker).
  - Syllables are an *emergent phenomenon* or *epiphenomenon* of articulatory and/or perceptual constraints.

- More about syllables in LING 451: Phonology I

Stress and Intonation

Rhythm of words and sentences

Read: LJ 10 “Stress” and “Length” (p. 249)

Read: LJ 5 up to “Intonation” (p. 118)

HW: Word stress

Suprasegmental Features

- *Suprasegmental features* involve more than single segments. There are four principle suprasegmental features:
  - *Length* – Segments and syllables can differ in length, creating phonetic, and sometimes phonemic, contrasts.
  - *Stress* – Stressed syllables are pronounced with greater energy.
  - *Tone* – Pitch levels or contours on syllables, or sometimes whole words, can change lexical or grammatical meaning.
  - *Intonation* – Variations in pitch and timing over phrases contributes to grammatical (syntactic boundaries, sentence type, etc.) and paralinguistic meaning (discourse functions, sociolinguistic features, speaker properties, etc.).

Phonemic Length

- *Phonemic length*: a segment's length is contrastive (causes a change in meaning)
  - Many languages have contrastive vowel length:
    - Tamil: [pal] 'teeth' vs [paːl] 'milk'
  - Some languages have contrastive consonant length. Long consonants are often called *geminates* and sometimes called "doubled."
    - They can be transcribed using a length marker after the segment (preferred in this class): [kitːa], [paːl]
    - Or by doubling the segment: [kitta], [paal]
      - More narrowly for stops: [kit̚ta] to show that the first stop is unreleased

Phonetic Length

- *Phonetic length*: segment or syllable length varies due to phonetic or intonational factors (coarticulation, phrase boundaries, stress, etc.). Examples:
  - English tense vowels tend to be longer than lax counterparts.
  - English vowels tend to be longer before voiced consonants than voiceless counterparts.
    - The vowel's length can serve as a cue to the consonant identity (especially helpful in noise and/or when stops are not released).
  - Stressed syllables tend to be longer than unstressed.
  - The more syllables in a word/phrase, the shorter each syllable.
  - *Phrase-final lengthening*: Ends of phrases (the last syllable or last few syllables) tend to be longer.

Strong and Weak Forms

- Many phones occur in *strong forms* and *weak forms*, meaning their pronunciation varies depending on context
  - *Citation form*: the way a word (or syllable) is pronounced in isolation. This is the form used to teach new words (often referred to as "careful pronunciation")
  - *Strong form*: the pronunciation of a word when it has emphasis or main sentential stress
  - *Weak form*: the pronunciation of a word in a sentence without emphasis or stress

Example: Different Forms of “and”

- Citation form: [ǽ̃nd]
  - “and” said in isolation has a low front vowel, a nasal stop, and a voiced alveolar stop which is usually released.
- Strong form: “I want this ***and*** that” [ǽ̃nd]
  - The strong form mimics the citation form, but may show more variation. Even in the strong form the nasal may be realized only as nasalization on the vowel ([ǽ̃d]) and the final stop may be absent or assimilate to the following consonant (“…and the…” [ǽ̃nd̪ ðə]).
- Weak form: “a bit of this and that” [ə̃nd], [ə̃n], [ə̃d], [ə̃], [n̩d] , [n̩]
  - This form shows the most variability (depending on the social and communicative context). If the consonants are present they almost always assimilate to the following consonant’s place of articulation.

Assimilation

- *Assimilation*: when one sound changes to become more like another sound (usually as a result of coarticulation between adjacent segments).
  - *Anticipatory assimilation*: one sound is affected by anticipating the gesture for the next sound.
  - *Perseverative assimilation*: the gesture for one sound perseveres into the gesture for the next sound.
- Examples:
  - Voicing of one segment changes to that of nearby segment(s)
  - The place of a nasal changes to that of a following consonant
    - impossible, inconsequential, sangria, tenth, emphasis

Stress

- *Stress* is rhythmic emphasis that (usually) makes a syllable louder than other syllables of its type. More energy is spent on stressed syllables, e.g., by pushing more air out the lungs.
- Segments become more exaggerated in pronunciation, as in citation form or strong form. Example features that increase *perceptual salience* (how easy it is for listeners to identify the sounds):
  - Stressed syllables tend to have higher pitch and longer vowel duration
  - Less assimilation and coarticulation between segments
  - Length contrasts (including phonetic ones) may be exaggerated
  - Aspirated stops have longer aspiration
  - Word-final stops are more likely to be released
  - Vowels have more peripheral formants

Stress and Vowel Quality

- Vowels often exhibit a change of quality between stressed and unstressed occurrences. Vowels that retain their quality even when unstressed are called *full vowels*. In contrast, some unstressed vowels become *centralized*, or *reduced*. Examples:
  - "exploit" [ək ˈsploɪt] (stressed full vowel)
  - "exploitation" [ˌɛk sploɪ ˈtʰeɪ ʃən] (unstressed full vowel)
  - "explain" [ək ˈspleɪn] (stressed full vowel)
  - "explanation" [ˌɛk splə ˈneɪ ʃən] (unstressed reduced vowel)
    - Reduced vowels in English are broadly transcribed with schwa [ə]
    - Note that stress marks are written before the syllable onset (if any).

Intonation

- *Intonation* is the rhythm and melody of a sentence.
  - Involves variations in pitch and timing (length and pauses)
- A group of words that has a cohesive intonational pattern is called a *tone group*.
- The *tonic syllable* is the syllable in the tone group that carries the main pitch change, called the *tonic accent*.
  - Tonic syllables are usually the last stressed syllable in the group unless the speaker wants to draw attention to a different word or morpheme.

Examples: Tonic Syllables and Tone Groups

- "I like pho*netics"
  - The whole sentence is a tone group, and the stressed syllable in "phonetics" [nɛ] is the tonic syllable.
- "I *dislike phonetics"
  - The whole sentence is a tone group but the emphasis is on "dislike," so the stressed syllable [dɪs] is the tonic syllable.
- "*Mikey likes phonetics"
  - The name "Mikey" (contrasting with someone else) is emphasized, so the stressed syllable [mɑɪ] is the tonic syllable.
- Note: by convention, tonic syllables are usually marked with a preceding asterisk in the orthography (as above).

Primary and Secondary Stress

	exploit		exploitation				explain		explanation			
	ək	ˈsplɔɪt	ˌɛk	splɔɪ	ˈtʰeɪ	ʃən	ək	ˈspleɪn	ˌɛk	splə	ˈneɪ	ʃən
tonic accent	–	+	–	–	+	–	–	+	–	–	+	–
stress	–	+	+	–	+	–	–	+	+	–	+	–
full vowel	–	+	+	+	+	–	–	+	+	–	+	–

- The term *primary stress* refers to the syllable that has both stress and the tonic accent, whereas *secondary stress* refers to a syllable with stress but no tonic accent.
  - Primary stress is transcribed with an upline before the syllable, while secondary stress is transcribed with a downline before the syllable. Example: "independent" [ˌĩn də ˈpʰɛ̃n də̃nt]

Intonation and Sentential Prominence

- Recall that *intonation* is a term for the rhythm and melody of a sentence.
  - Languages will often have default intonation patterns for different sentence types: e.g., in English declarative sentences, the last stressed syllable in a tone group will have greater prominence and is said to have *primary stress* and carry the *tonic accent*.
  - Tonic accent is typically marked by a *pitch excursion*, which makes that syllable either higher or lower in pitch than the surrounding syllables. Whether the excursion is up or down for a similar intended meaning can vary by dialect.
  - Tonic accent can manifest on other stressed syllables in the sentence as a way of marking *sentential prominence* or *focus*.

Rhythm: Syllable-timed vs. Stress-timed?

- An old system of dividing languages – there is little quantitative evidence, but it's still referred to, often in ESL or other language-teaching texts:
  - It was thought that in "syllable-timed" languages, each syllable is about the same length. (Turns out, that's not really true.)
  - "Stress-timed" languages were thought to use stress as the major organizing unit of rhythm, e.g., preferring a pattern of stressed-unstressed-stressed-unstressed syllables.
    - Many languages do have a rhythmic preference for keeping stressed syllables apart (and may move stress to maintain the pattern), but there are many interacting factors (i.e., it's just not that simple).
    - Stressed syllables do tend to be longer, but also in "syllable-timed" languages, and not necessarily in every "stress-timed" language.

Rhythm: Word Stress

- A different system with a similar typological goal of dividing languages based on their rhythmic timing tendencies:
  - *Variable word stress*: The location of word stress is not always predictable from the segmental structure; word stress can be lexical (a property of each word) or indicate grammatical info, e.g., syntactic category
    - e.g., English noun/verb pairs: (to) in*sult* vs. (an) *in*sult
  - *Fixed word stress*: Stress lands on a predictable syllable of a word, e.g.:
    - (Near)-universally, e.g., Czech: stress on first syllable of a word; Polish, Swahili: on penultimate (second-to-last)
    - For grammatical function, e.g., tense, part of speech (cf. English)
    - Predictable based on syllable structure, e.g., heaviest syllable
  - *Fixed phrase stress*: Stress falls on predictable locations in the phrase
    - e.g., French: stress comes phrase-finally (at/near end of phrase)

Rhythm: Mora-timed

- *Mora*: a weight/timing unit between segment and syllable in a hierarchical (e.g., tree) structure.
- All languages have syllables, but some use morae as timing units.
  - e.g., Japanese: [nippoŋ] 'Japan' can be analyzed as having two syllables but four total morae: [ni p po ŋ]
  - In Japanese, each mora has about the same length.
- Notes:
  - Onset—simple nucleus (non-branching) is a common mora shape.
  - Branching syllable parts (onset, rhyme, nucleus, coda) often add morae.
  - Some languages require morae to have a minimum number of segments, or syllables to have a minimum number of morae, so they add segments or morae when a morphological process creates a too-light unit.

Tone

The effects of pitch on meaning

Read: LJ 10 from “Intonation and Tone” (p. 254)

Preview HW: Intonation & Tone

Transcription: Non-English Tone

(Lab: Pitch)

Pitch

- Recall that *pitch* is the perceptual correlate of the fundamental frequency of a speech signal.
- Pitch is primarily modulated by changing the tension in the vocal folds: higher tension yields faster vibration and thus higher f_0.
  - Pitch is also affected by the rate of airflow through the glottis: a faster rate of airflow causes the vocal folds to vibrate faster.
  - To some extent the position of the vocal folds affects pitch as well; for example, creaky-voiced sounds tend to have a slower rate of vibration (and thus lower pitch) than their modal-voiced counterparts
- *Pitch affects meaning in all languages, in one way or another.*

Lexical Tone

- In some languages, the pitch of a sound can change the lexical meaning of the word it is part of.
  - The phenomenon is called *lexical tone*; such languages are called *tone languages* (or *tonal languages*).
    - Some examples of tone languages are Mandarin Chinese, Vietnamese, Hausa, and Zulu.
    - About 49% of the world's languages are tone languages.
- Example from Mandarin Chinese:
  - ma 55 (媽) "mother" ma 35 (麻) "hemp"
  - ma 214 (馬) "horse" ma 51 (罵) "scold"

(here tone is transcribed using numbers; more on that shortly)

Grammatical Tone

- Pitch can also act as an inflectional element, so that the basic meaning of the word stays the same but grammatical features like tense or case are affected. This is called *grammatical tone*. Grammatical tone is common in the tonal languages of Africa.
  - Example: Edo
    [ì mà] "I show"
    [ì má] "I showed"
    [í mà] "I am showing"

(here tone is transcribed using diacritics; more on that shortly)

Level Tones and Contour Tones

- Tone is usually considered a characteristic of syllables, with the tonal information most clearly audible on the syllable nucleus (i.e., the vowel or syllabic consonant).
  - *Level tones* (or *register tones*) have a relatively stable pitch for the duration of the syllable.
  - *Contour tones* change pitch noticeably over the course of the syllable.
  - In some languages, tone appears to operate at the word level, so that polysyllabic words will always show the same tone on all syllables.

Transcribing Level Tones: Diacritics

- The simplest tone languages (e.g., Shona, Zulu) have just two register tones (high and low). High tones are transcribed with an acute accent [á] and low tones with a grave accent [à].
  - In languages with three register tones (e.g., Yoruba), the high and low tones are transcribed as above; the middle tone is often left unmarked but can also be marked by a macron [ā].
  - In languages with more than three register tones, the additional tones are transcribed as extra-high [a̋] or extra-low [ȁ].

Transcribing Contour Tones: Diacritics

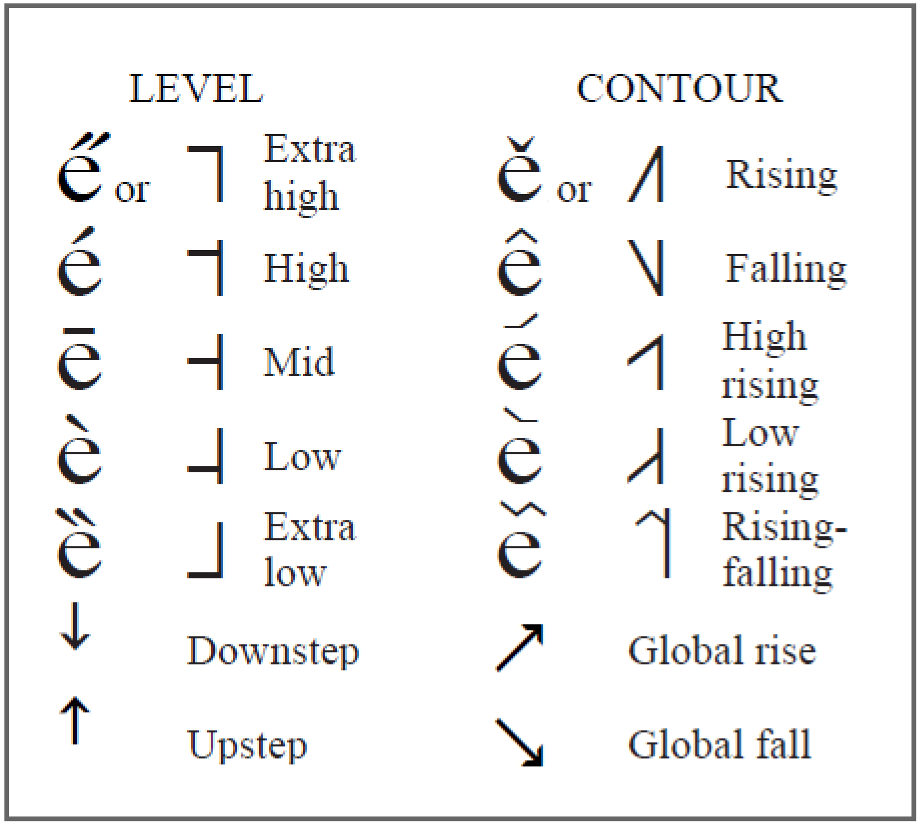
- Contour tones can be transcribed by combining register tone accents into new diacritics.
  - a rising tone combines low + high: ` + ´ → ǎ
  - a falling tone combines high + low: ´ + ` → â
  - a low-rising tone combines low + mid: ` + ¯ → a᷅
  - a high-rising tone combines mid + high: ¯ + ´ → a᷄

Transcribing Tone: Lines and Numbers

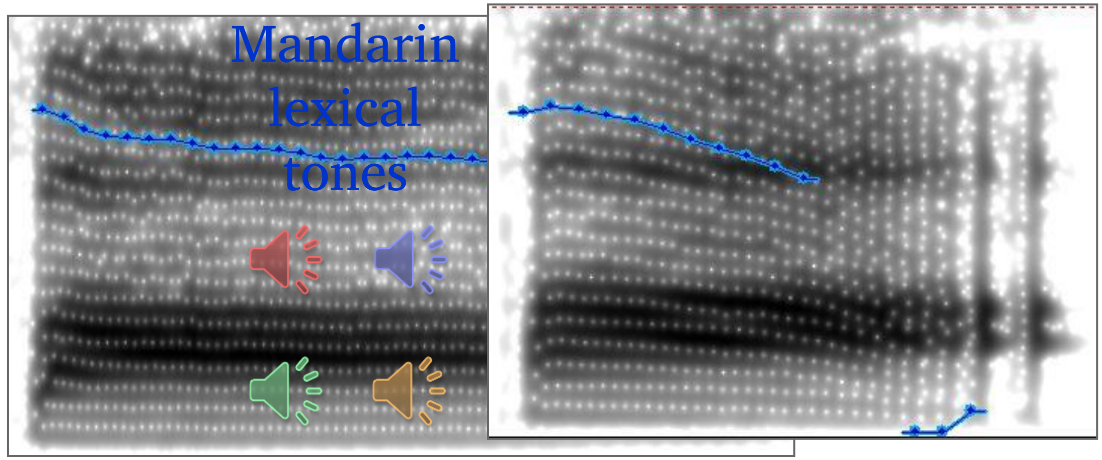
- Tones can also be transcribed with line graphics following the syllable, or sequences of numbers showing the beginning and ending points of the tone (1 = extra-low, 2 = low, 3 = mid, 4 = high, 5 = extra-high).
- Examples (imaginary words):
  - Extra-low tone: da˩ (da^{11})
  - Mid tone: da˧ (da^{33})
  - Falling tone: da˥˩ (da^{51})
  - High rising tone: da˦˥ (da^{45})
  - Low dipping tone: da˨˩˦ (da^{214})

Transcribing Tone: IPA Chart

- The IPA uses two of the systems: diacritics and lines (*tone letters*)
  - Hint: To create contour tone letters in Office, use Insert > Symbol. Enter the beginning tone letter, then the next tone letter (and then the third, if applicable), and they will overlay each other to create one contour tone letter.
  - Example: mid-low-mid: click on the symbols ˧ ˩ ˧ in that order; the result: ˧˩˧
  - Do the same using the online character pickers, then copy and paste the result: ˧˩˧

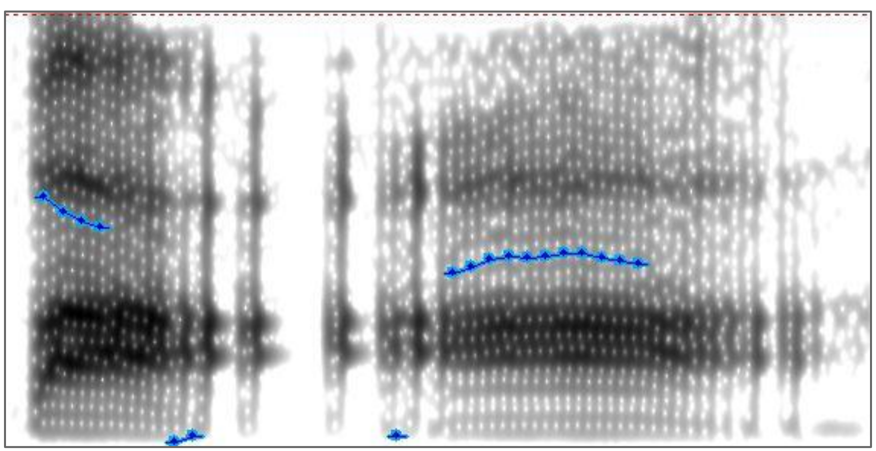


Contour Tones on Spectrograms

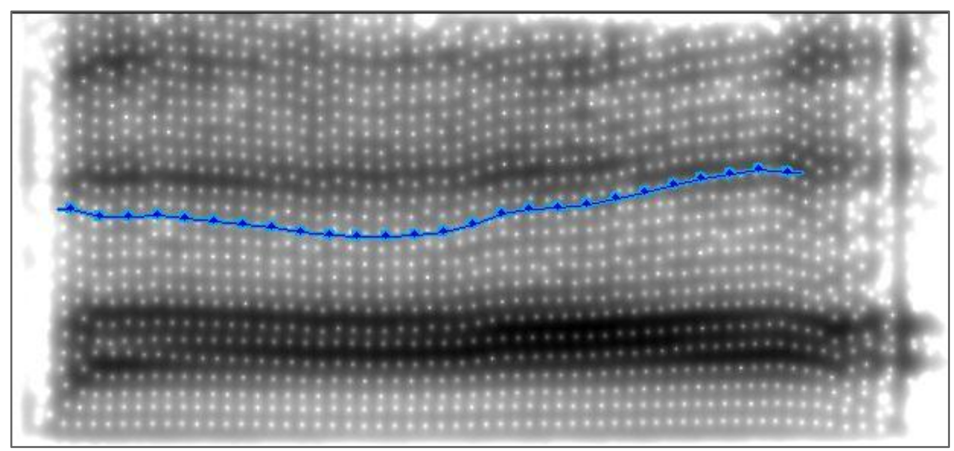


‘eight’ [pa˥] high level

‘father’ [pa˥˩] high-to-low falling



‘to hold’ [pa˧˩˧] mid-low-mid dipping



‘to pull out’ [pa˧˥] mid-to-high rising

Tone Overlap and Tone Sandhi

- Just as neighboring segments influence each other through coarticulatory processes, neighboring tones can also influence each other.
  - Pitch changes during speech tend to be continuous; we don't abruptly jump from one pitch to another, but glide smoothly between pitches. So for example, a low tone that follows a high tone will often begin with a downward movement of pitch.
  - Systematic changes of neighboring tones are called *tone sandhi*. Common types of tone sandhi:
    - Rules that make neighboring identical tones different. For example, Mandarin 好冷 "very cold" [xɑʊ^{214}] + [ləŋ^{214}] → [xɑʊ^{45} ləŋ^{214}].
    - Rules that spread the first tone of a polysyllabic word across the entire word: e.g., Shanghainese "universe" [tʰi^{51}] + [di^{15}] → [tʰidi^{51}].

Pitch Accent

- In some languages, tone is only pronounced on certain types of words or certain words in the sentence.
  - Such languages are often called *pitch accent* languages, though this term is not always used consistently in the literature.
  - It is thought that tone in pitch accent languages is not truly phonemic, but gives some information about the form of inflected words (e.g., whether its root form is mono- or disyllabic).
  - Examples: Japanese, Serbo-Croatian, some dialects of Swedish.

Example of Pitch Accent: Japanese

- Pitch-accent languages allow only one syllable within a word to carry non-neutral tone:
  - [háʃi] “chopsticks”
  - [haʃí] “bridge”
  - [haʃi] “end, edge”
- The rest of the tone group will have a neutral tone, which can depend on the pitch pattern of the accented word:
  - [háʃionɯɹɯ] “paint chopsticks”
  - [haʃíonɯɹɯ] “paint a bridge”
  - [haʃionɯɹɯ] “paint the end”

Other types of Tone

- The use of tone in language is not an “all or nothing” phenomenon.
  - Some languages have a small number of words that differ only in tone, but which preserve historical information about the language (e.g., low tone indicating the loss of what was historically a syllable-final consonant).
  - Example: Scottish Gaelic.

Transcribing and Measuring Pitch

ToBI and pitch tracking

Read: LJ 5 from “Intonation” (p. 118)

Finish HW: Intonation & Tone

(Lab: Pitch)

Pitch tracking demo/exercise

- Open Ibibio transcription practice sound files in Praat.
  - Use standard pitch settings and turn on pitch tracking.
- Compare minimal pairs with different tones.
  - Can you see the difference between High and Low tone?
  - What is a rough estimate of this speaker's average pitch for High tone and for Low tone? (The speaker is saying clear citation forms of the words; in connected speech, tones will vary more, being higher or lower relative to preceding tones.)
  - What patterns do you notice in disyllabic (two-syllable) words? For HL patterns? For LH patterns?

Transcribing Intonation: ToBI

- A more fine-grained way of marking intonation is the *Tone and Break Indices (ToBI)* system.
  - In the ToBI system, intonation is described as a series of *target tones* on the stressed syllables.
    - Each intonational phrase has at least one target tone (also called *pitch accents*). The last pitch accent in a phrase is called the *nuclear pitch accent*.
    - Nuclear pitch accents are always followed by a *phrase accent* and a *boundary tone*.
  - *Break indices* are numbers representing the degree of separation between words. They range from 0 (no break) to 4 (a long pause, as between sentences).

Marking Transcriptions with ToBI

- ToBI markup is typically added to lines (or *tiers*) above the IPA transcription (one tier for tones and one for break indices).
  - Phrase accents are marked as high or low (either H– or L–).
  - Boundary tones are marked as high or low (either H% or L%).
  - Nuclear and pre-nuclear pitch accents can have six forms:
    - H* (high)
    - L* (low)
    - L* + H (rise beginning on stressed syllable)
    - L + H* (rise ending on stressed syllable)
    - H + !H* (high with slight downstep)
    - !H* (downstepped high; occurs after H* in sentences with downdrift)

Common ToBI Patterns

- Certain types of utterances tend to have similar ToBI representations on the tonal tier. Examples from English:
  - H* L– L% (declarative statement)
  - L* L– H% (direct address)
  - L* H– H% (question)
  - L+H* L– H% (expression of surprise)
  - L+H* L– L% (reprimand)

British (Peter Ladefoged)

American (Bruce Hayes)

Example: ToBI Transcription

Break Index	1	1	1	4
Tonal Tier	H*			H* L– L%
Segmental Tier	meɹɨ ˈænə	meɪd̪	ðə	ˈmaɹ mə leɪd
Orthography	*Marianna*	*made*	*the*	*marmalade*

data adapted from http://www.ling.ohio-state.edu/research/phonetics/E_ToBI/ToBI/ToBI.1.html

Phonetic Transcription

Written symbols for speech sounds

Read: LJ 2

Preview HW: English word transcriptions

Transcription

- *Transcription* is the use of a set of agreed upon written symbols to accurately describe sounds of a spoken language.
  - *Phonemic transcription* describes spoken language using only contrastive sounds (phonemes). Phonemic transcription is often called *broad transcription*.
  - *Narrow transcription* notes the allophonic nuances. There are many levels of detail in narrow transcription.
  - A *phonetic transcription* is the narrowest, noting all of the detail in an individual's pronunciation of speech.
- Depending on the language being transcribed, any alternation (e.g., aspiration) may or may not mark a phonemic contrast and thus may or may not be included in a broad transcription.

Transcribing Variation

- Variation in speech may be noted in a transcription by using different phonetic symbols, or by using diacritics to modify symbols.
- In practice, it is very difficult to perform phonetic transcription without the use of recordings and software tools. Most transcription is *impressionistic*, meaning it involves listening to speech and noting as much of the detail as possible.

Choosing a Transcription Method

- The goal of the research determines the appropriate level of transcription. The most widely used level is what Ladefoged refers to as “fairly broad” which notes systematic allophonic variation, but not speaker-dependent variation or suballophonic variation (such as that caused by fatigue, random variation, or even detailed intonational effects).
- In this class we will concentrate on “fairly broad” transcription:
  - we will note things like aspiration, nasalization, flapping, voicing/devoicing…
  - but not some of the harder to hear effects like vowel shortening or unreleased stops.

Limitations of Transcription

- Transcription uses discrete units (segments) to describe a dynamic, continuous phenomenon (speech).
  - As such, even the narrowest transcription is an imperfect record of the speech signal it represents. In reality, the boundaries between segments are not always clear, and the relative durations of segments can vary quite a bit.
  - Nonetheless, transcription is a powerful tool for representing certain aspects of the speech signal (e.g., the sequence of sounds, dialectal variation, which sounds are phonemic in a given language, etc.)

English Consonants

Features and Alternations

Read: LJ 3

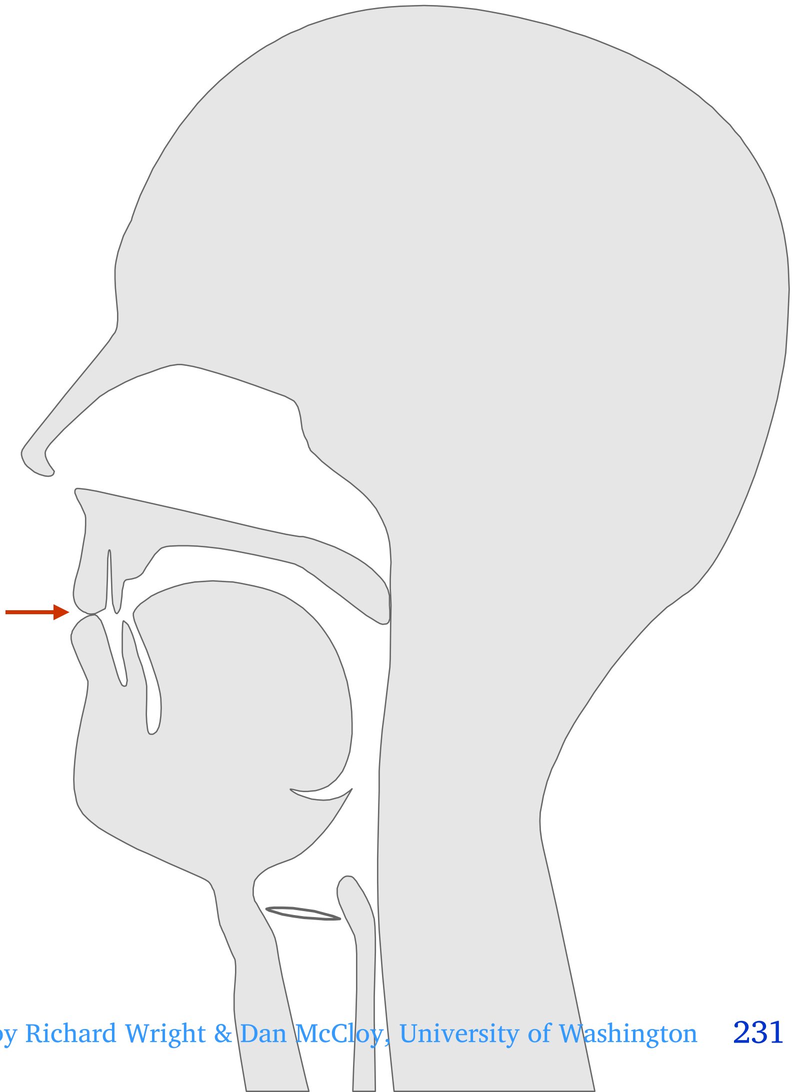
Begin HW: English word transcriptions

The Consonants of English

	Bilabial	Labio-dental	Dental	Alveolar	Palato-alveolar	Palatal	Velar	Glottal
Stop	p b			t d			k g	(ʔ)
Nasal	m			n			ŋ	
Tap				(ɾ)				
Fricative		f v	θ ð	s z	ʃ ʒ			h
Affricate					ʧ ʤ			
Approximant	w			ɹ		j		
Lateral				l				

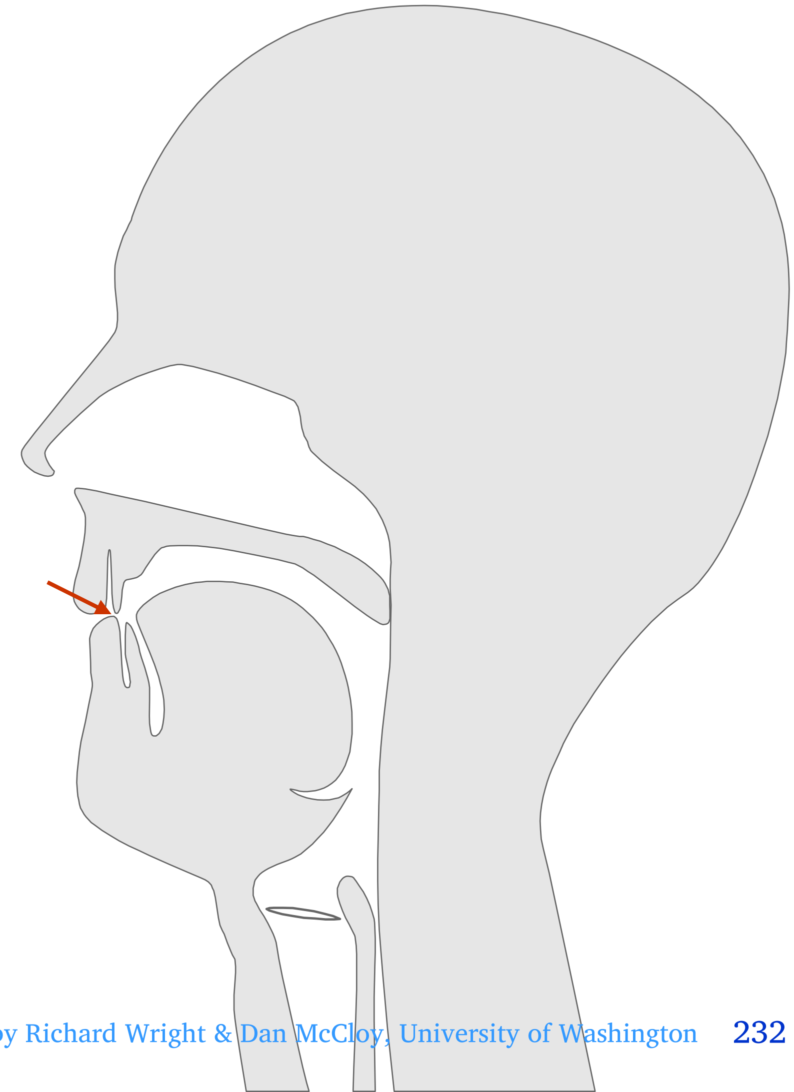
English Consonant Places of Articulation

- *Bilabial* sounds involve a constriction made by the lips.
  - Examples: / p b m /
  - The /w/ sound in English involves simultaneous bilabial and velar constrictions.



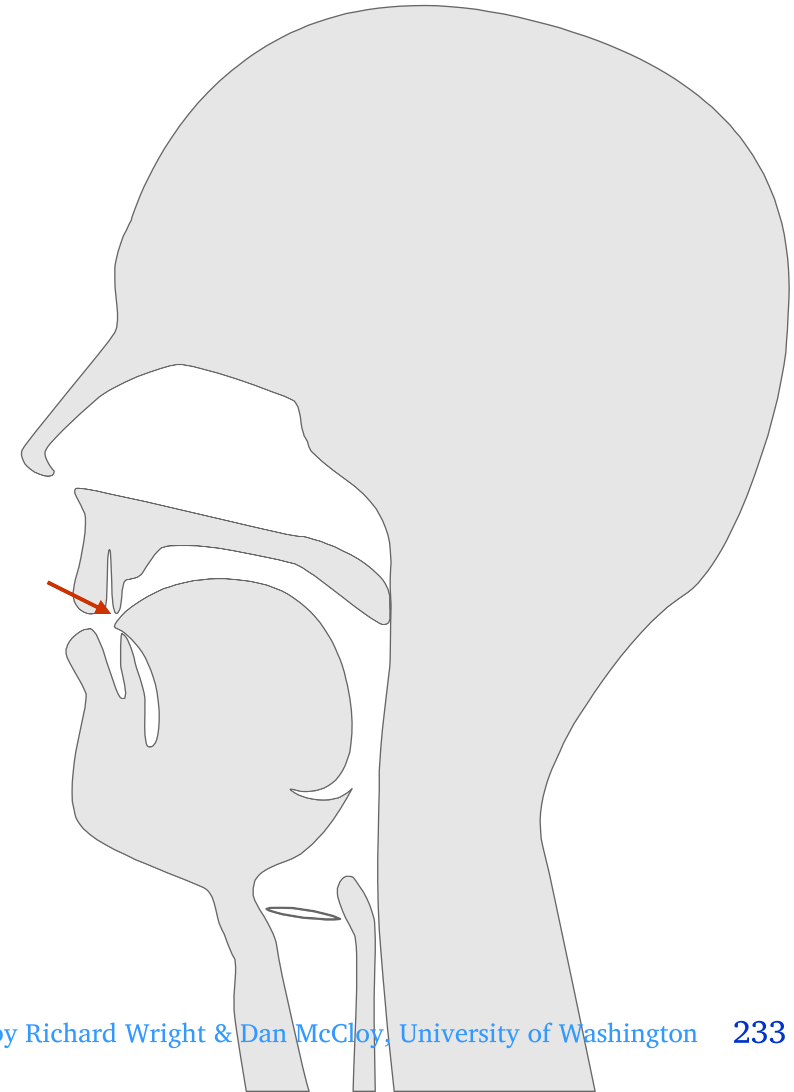
English Consonant Places of Articulation

- *Labiodental* sounds involve a constriction made by the lower lip and upper teeth.
  - Examples: / f v /



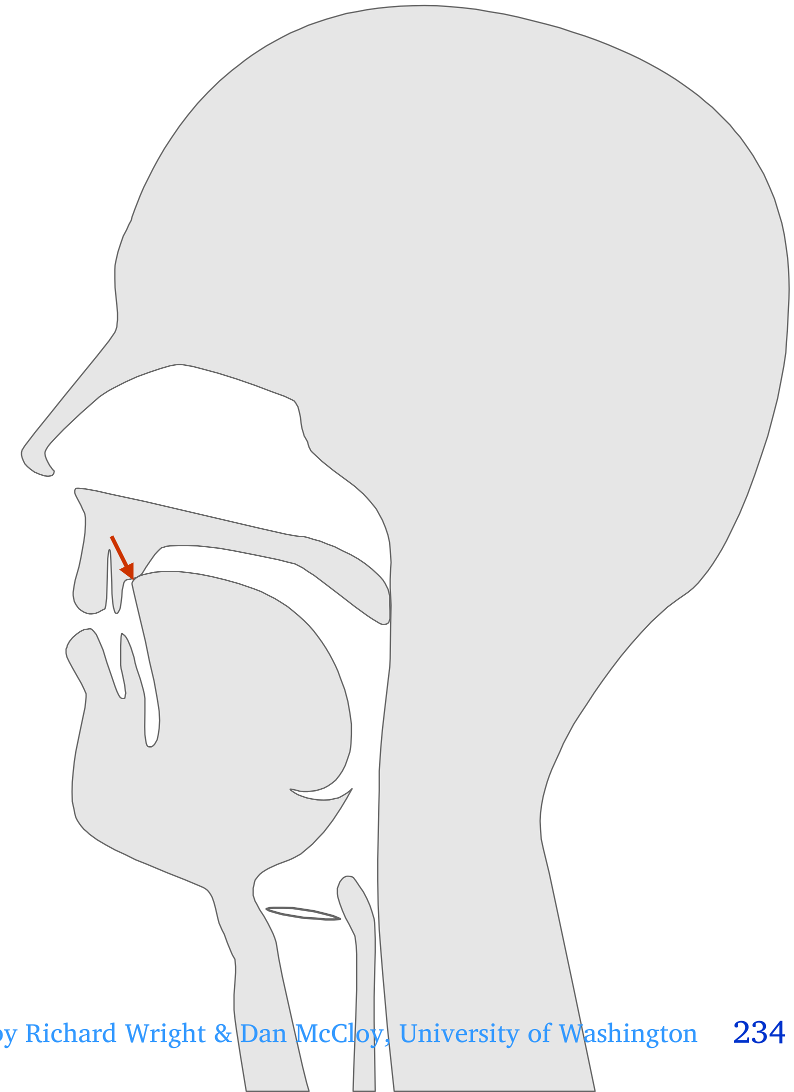
English Consonant Places of Articulation

- *Interdental* sounds involve a constriction made by placing the tongue tip between the front teeth.
  - Examples: / θ ð /
  - many British English speakers pronounce / θ ð / as *dental* rather than interdental



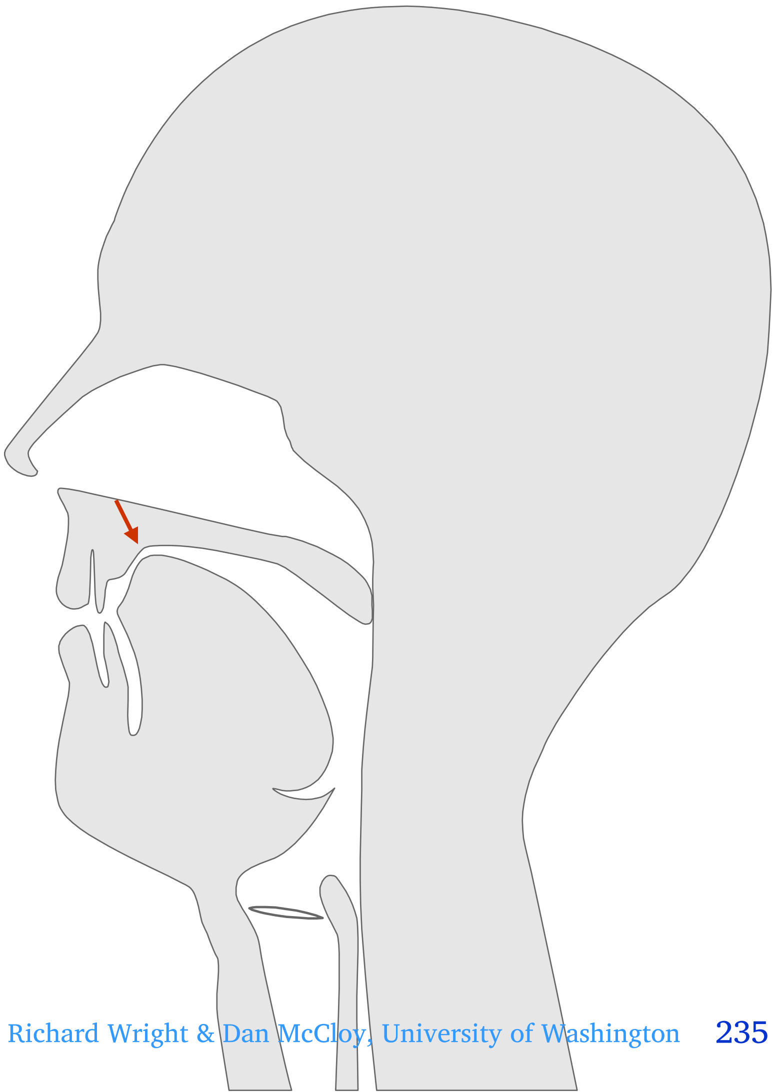
English Consonant Places of Articulation

- *Alveolar* sounds involve a constriction between the tongue tip (or blade) and the alveolar ridge.
  - Examples: / t d n s z l / [ɾ]
  - /ɹ/ is produced by some English speakers as an alveolar sound.



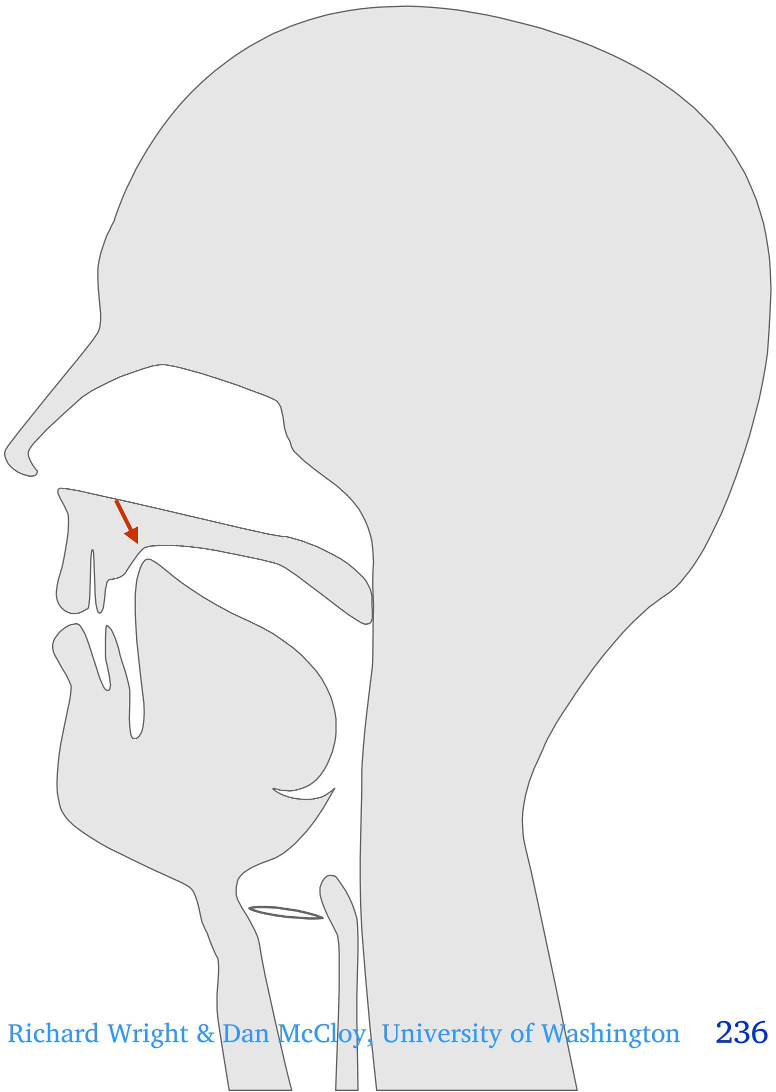
English Consonant Places of Articulation

- *Palatoalveolar* sounds involve a constriction between the tongue blade and the postalveolar region.
  - Examples: / ʃ ʒ ʧ ʤ /



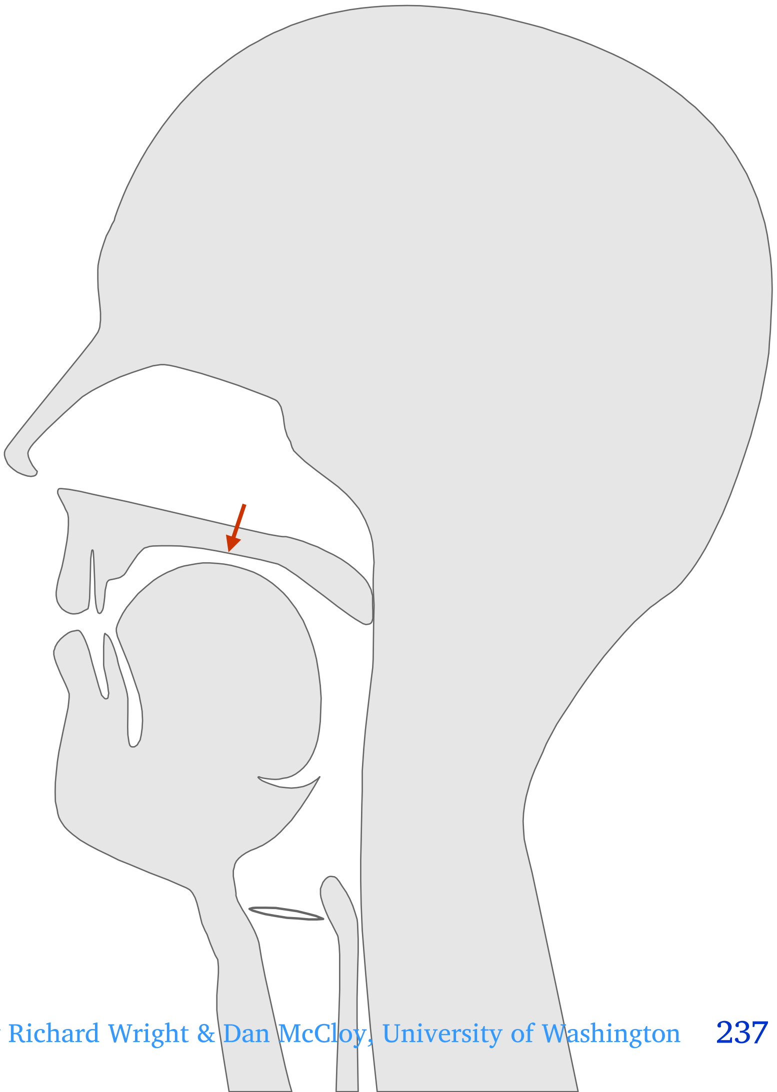
English Consonant Places of Articulation

- *Retroflex* sounds involve a constriction between the tongue tip or underblade and the postalveolar region.
  - For many English speakers, the /ɹ/ sound is pronounced with a retroflex articulation. If you want to specify a retroflex /ɹ/ in a narrow transcription, use the IPA symbol [ɻ].



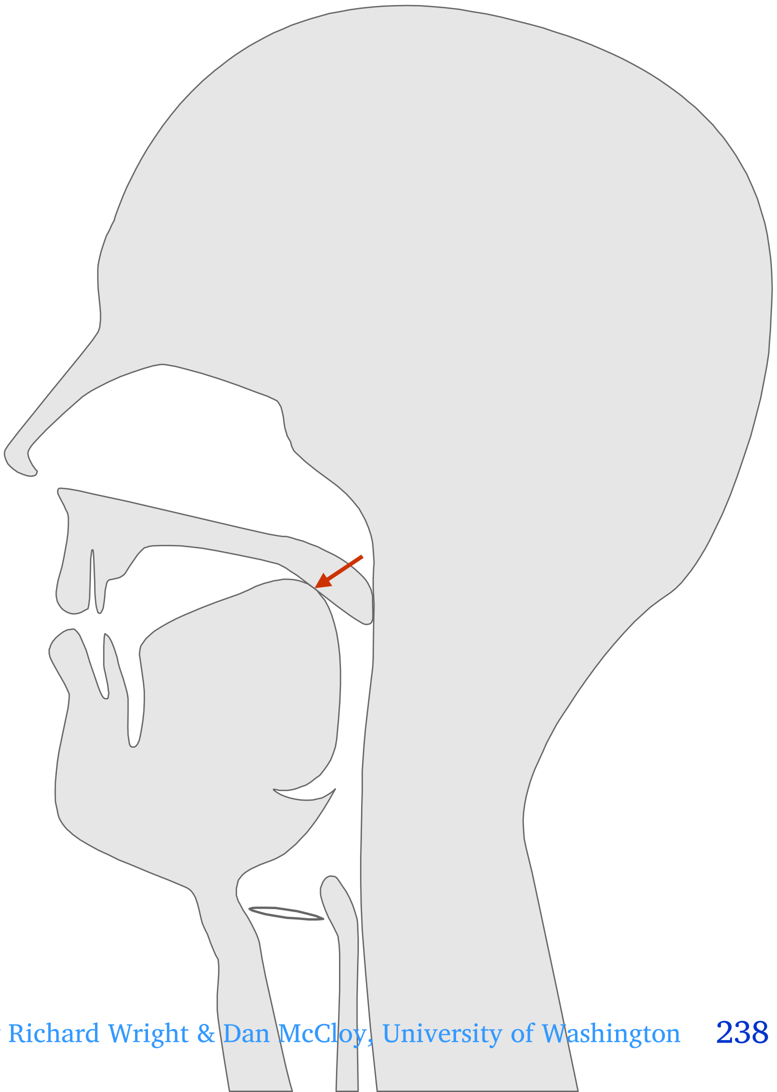
English Consonant Places of Articulation

- *Palatal* sounds involve a constriction between the tongue tongue front and the hard palate.
  - Example: / j /



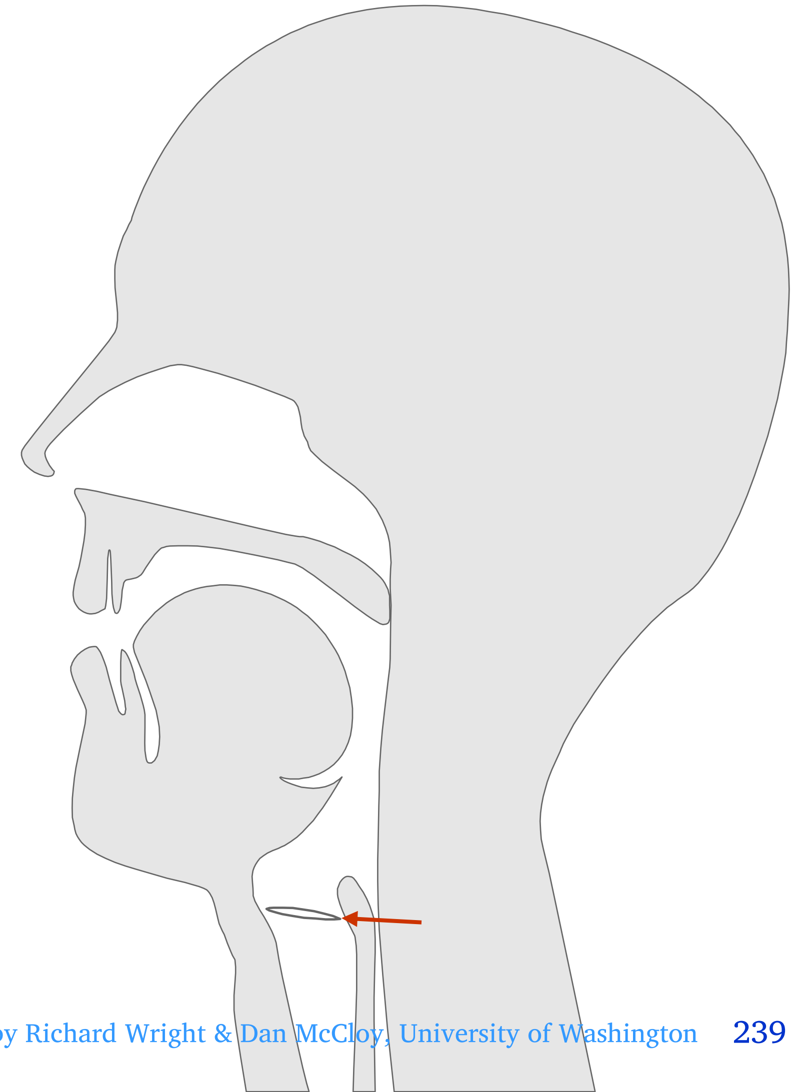
English Consonant Places of Articulation

- *Velar* sounds involve a constriction between the tongue back and the velum.
  - Examples: / k g ŋ /
  - Some English speakers make the /ɹ/ sound with a velar constriction instead of an alveolar or post-alveolar one. This is sometimes called the “bunched r” and can be represented in narrow transcriptions as the velar approximant [ɰ].

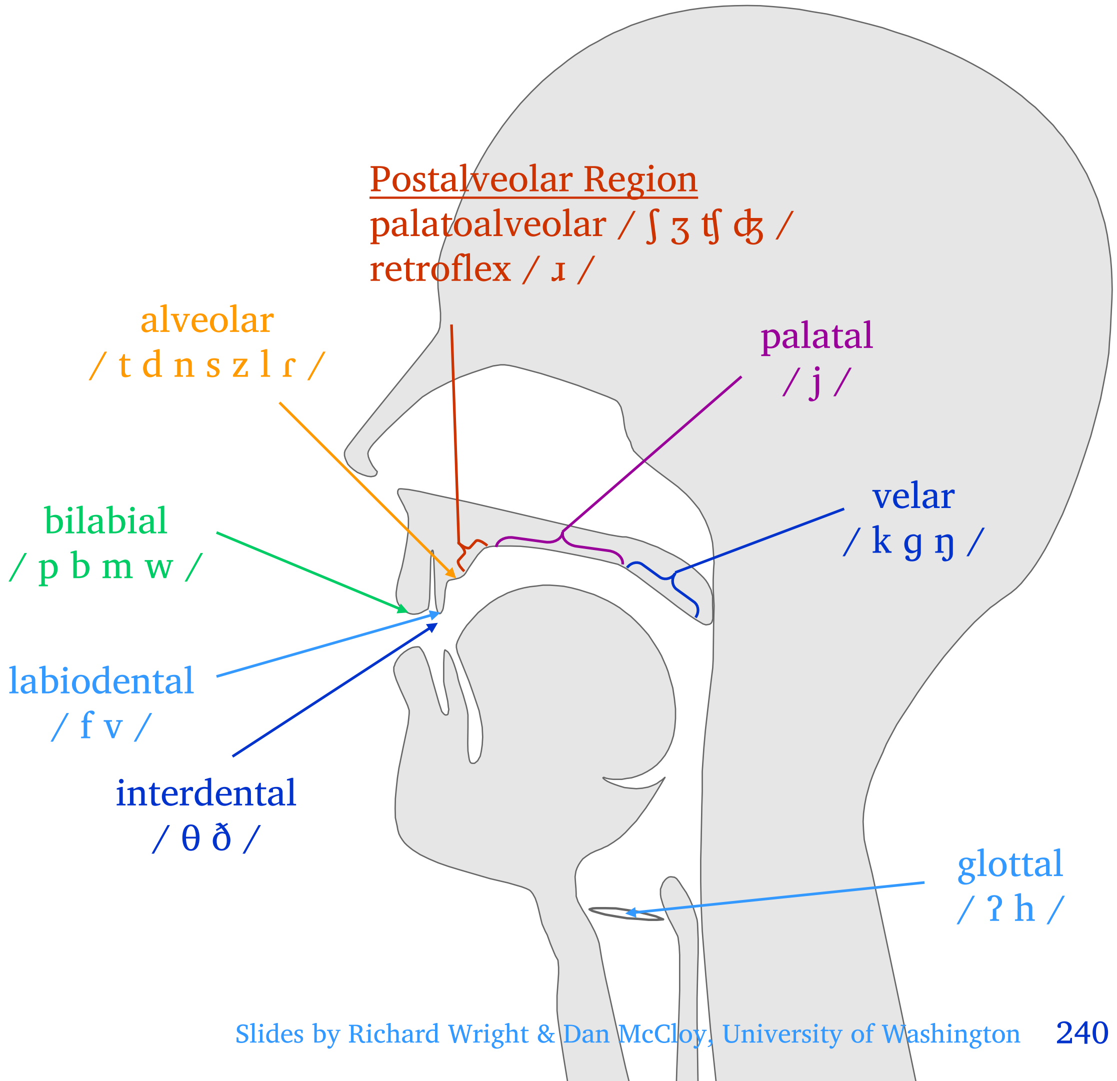


English Consonant Places of Articulation

- *Glottal* sounds involve a constriction made by the vocal folds.
  - Examples: [ʔ] / h /



English Consonant Places of Articulation



Coarticulation and Alternations

- *Coarticulation* occurs when the articulatory gestures for neighboring segments overlap.
  - Example gestures that spread to nearby segments: voicing, nasality (velum lowered), lip rounding, place of articulation
- *Alternations* are phonetic variations that occur as the result of phonological rules, which are stated in terms of phonetic or phonological environments. Alternations may or may not be related to coarticulation.
  - Example environments: Word/syllable-initial or –final, in a stressed syllable, between voiceless segments, preceding a nasal
  - Allophones are a type of alternation.

Common English Alternations

- *Devoicing*: approximants that follow an aspirated voiceless oral stop become partially devoiced.
  - Example: the /l/ in “please” and “play” is voiceless [l̥]
- *Nasal Plosion*: stop consonants that follow a stressed syllable and are *homorganic* (have the same place of articulation) with a following nasal are released nasally when the velum lowers for the nasal.
  - Example: in “hidden” the /d/ is nasally released [dⁿ]
- *Lateral Plosion*: alveolar oral and nasal stops are released laterally when they occur before a homorganic lateral.
  - Example: in “middle” the /d/ is laterally released [dˡ]

Common English Alternations

- *Palatalization*: consonants move to/toward a palatal place of articulation.
  - Dorsal consonants are fronted or anterior consonants are backed (*palatalized*) before high front vowels and palatal consonants.
    - Examples: [ka, ku] “caw, coo” but [ci, cju] “key, cue”
    - Examples: Japanese /si/ → [ʃi], English [n^juz, ɲuz] “news”
  - Palatalization can also refer to the addition of a *secondary articulation* to a non-palatal consonant: [t^j, n^j, k^j, s^j] etc.
- Anticipatory *labialization*: consonants have secondary lip-rounding when followed by a rounded segment.
  - Examples: [t^wwaɪs, t^wɹi] but [taɪt, ti]

Common English Alternations

- *Aspiration*: the voicelessness of voiceless unaspirated stops overlaps into the following vowel (or approximant) creating an /h/-like sound. The aspirated allophones of voiceless stops [p^h, t^h, k^h] occur in word initial position and at the beginning of stressed syllables.
- *Flapping*: alveolar stops (oral /t, d/ and nasal /n/) become flaps (oral [ɾ] or nasal [ɾ̃]) when they follow a stressed syllable and precede an unstressed syllable.
  - Example: “catty” and “caddy” both have oral flaps [ɾ]; “canny” has a nasal flap [ɾ̃].
- *Dark /l/*: In many English dialects the alveolar lateral /l/ at the end of a word has a velar constriction instead of or in addition to the alveolar one. This type of /l/ is called “dark” (or *velarized*) and is transcribed with the symbol [ɫ].

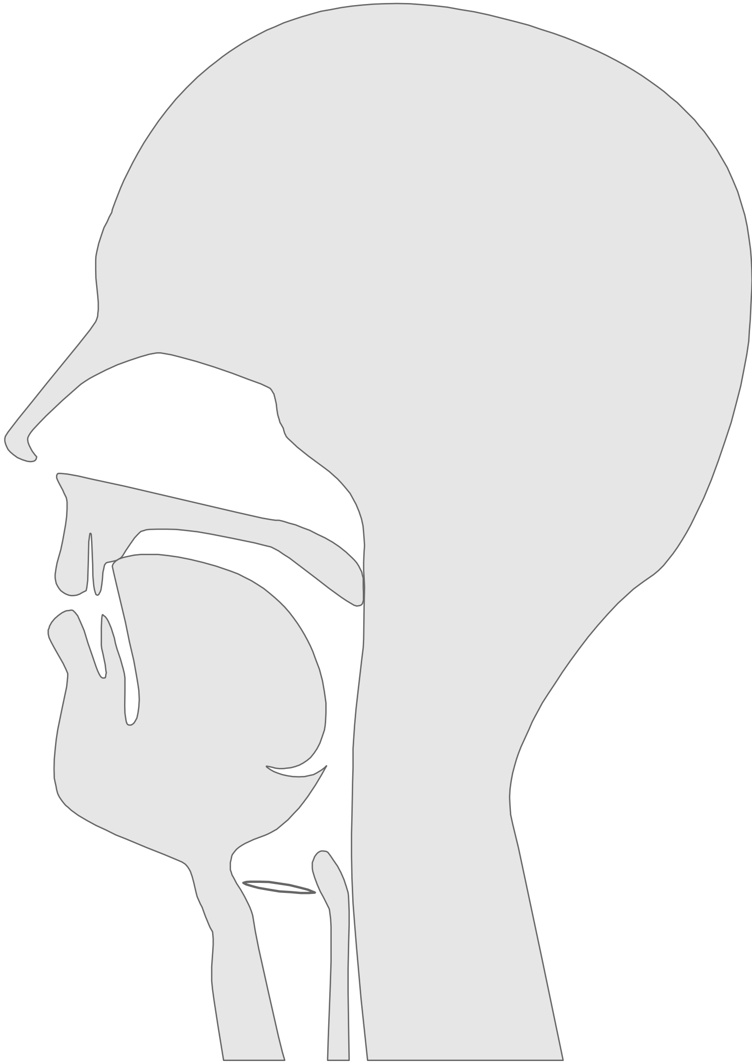
Common English Alternations

- See “Rules for English Consonant Allophones” on p. 72-77 of the Ladefoged & Johnson text.
  - Repeat the example words in each rule (or think of your own). Can you observe the rules in your own speech?
  - To help you perceive the differences, try to say example words with a different alternation. Is it difficult?
  - Highlights in class:
    - Compare your voiceless stops in the environments described in #2 & 5.
    - Try #8-10, 14-16.

Review:
English Consonant Articulations

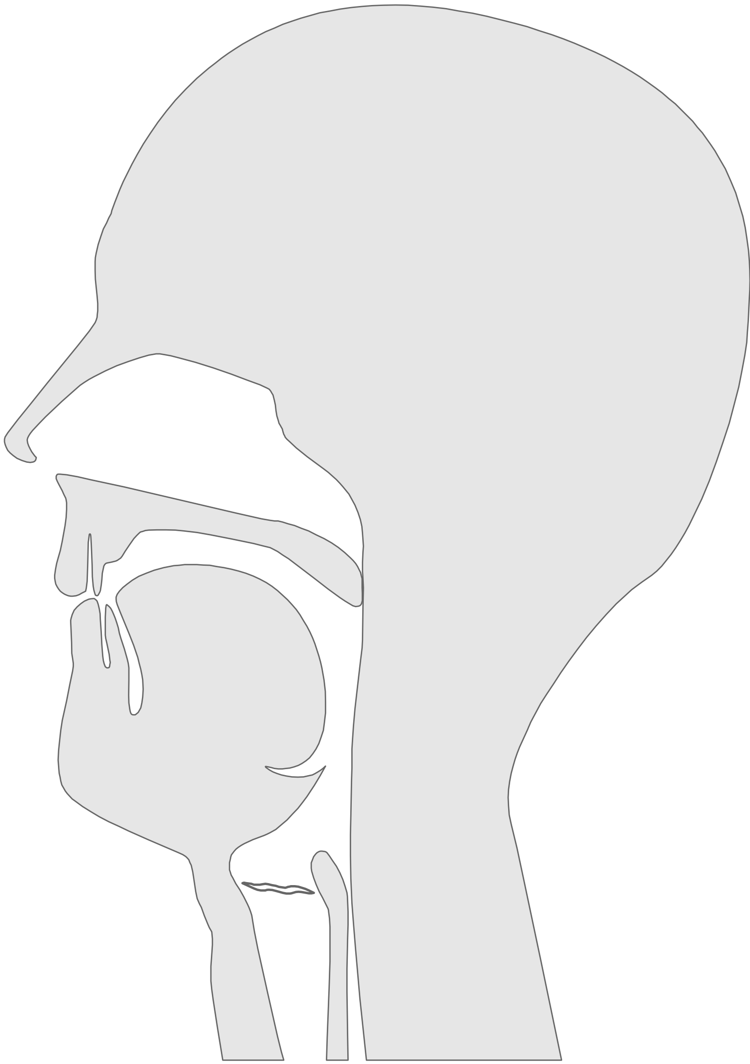
Consonant Articulations

t



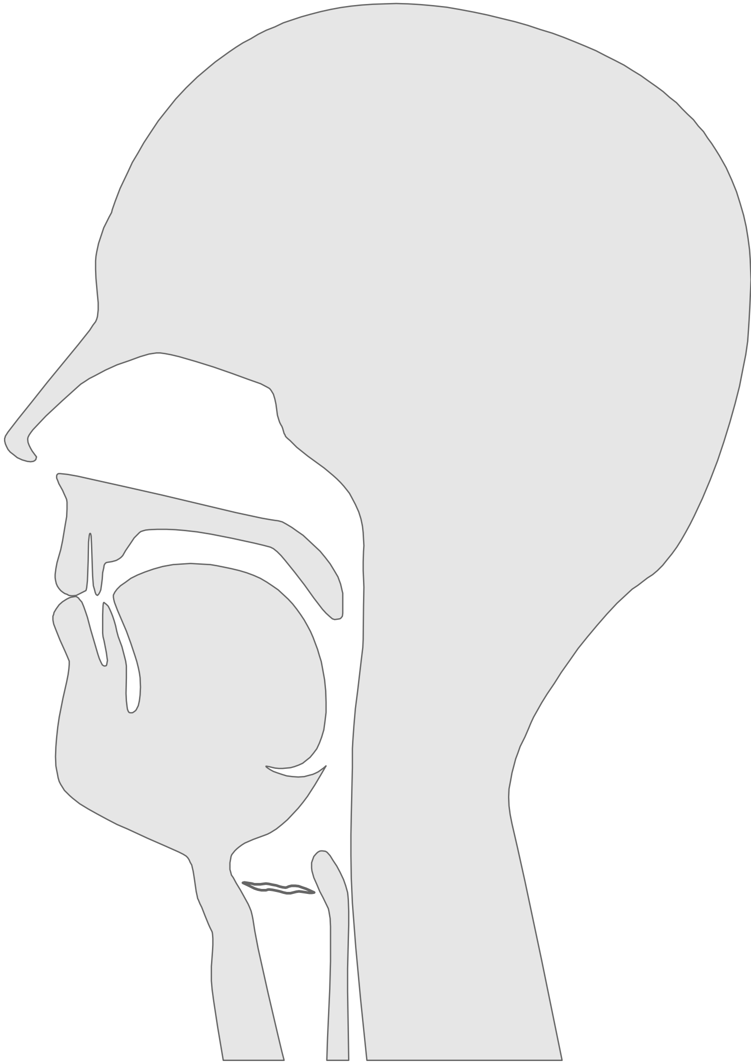
Consonant Articulations

v



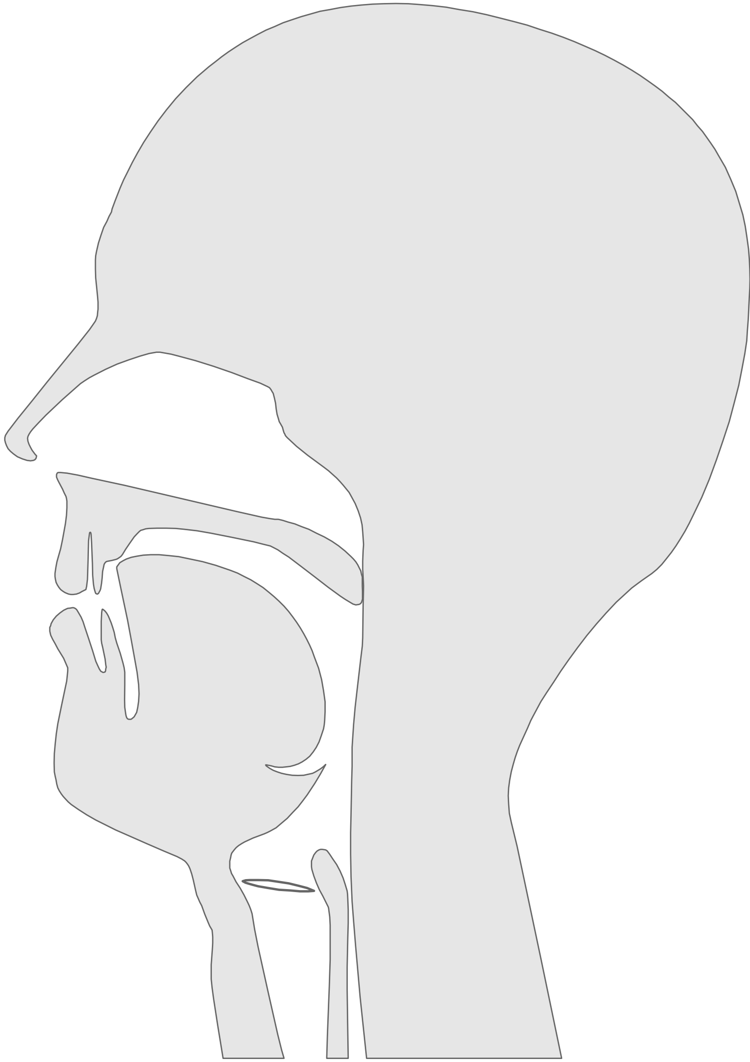
Consonant Articulations

m



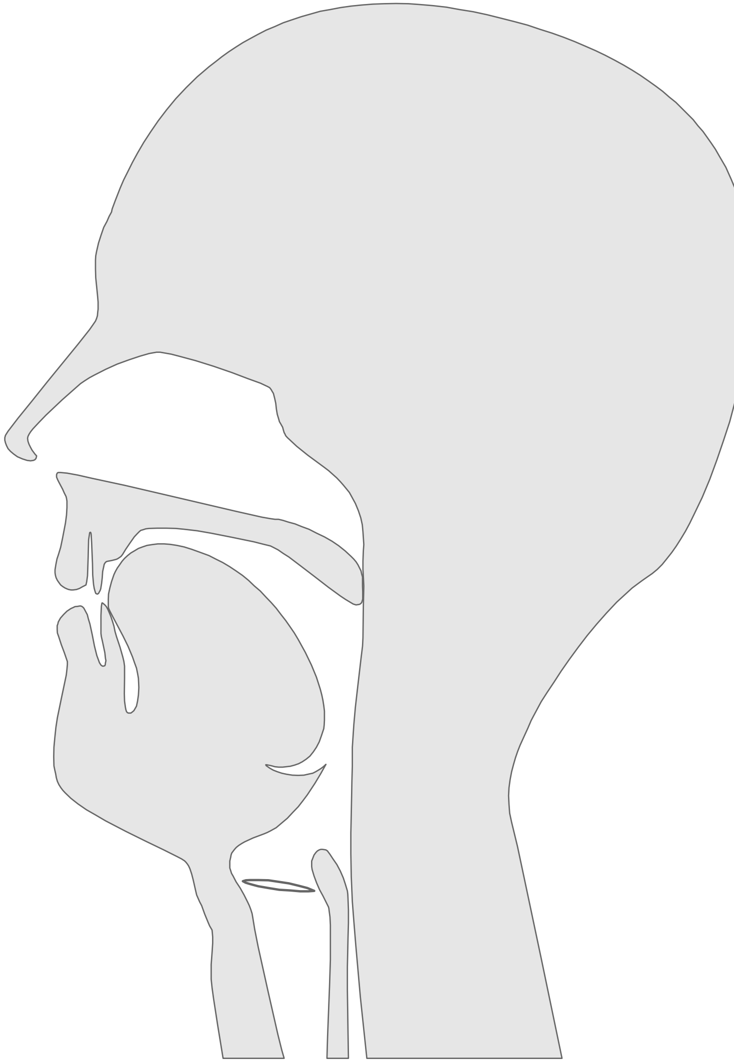
Consonant Articulations

s



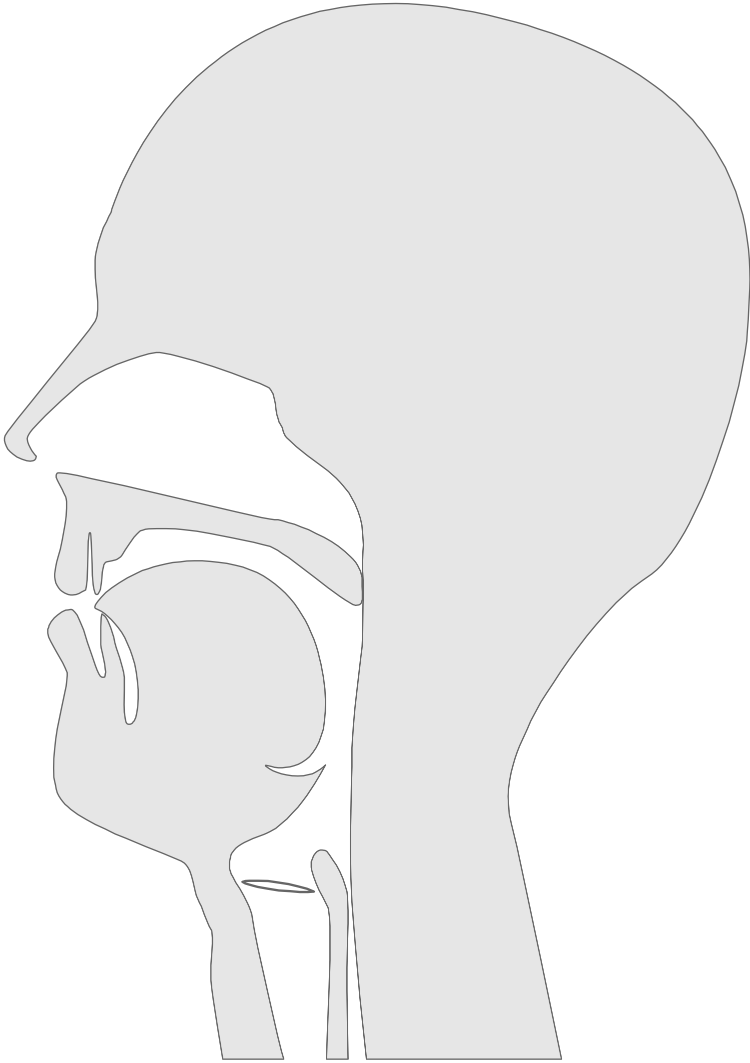
Consonant Articulations

s̻



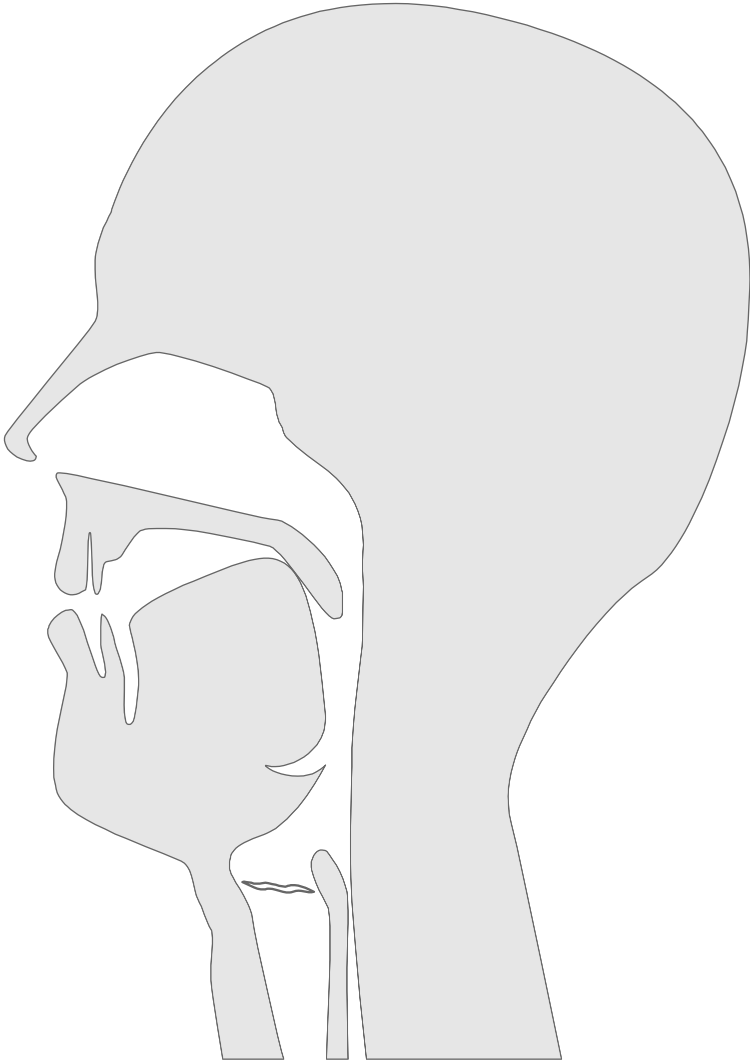
Consonant Articulations

θ



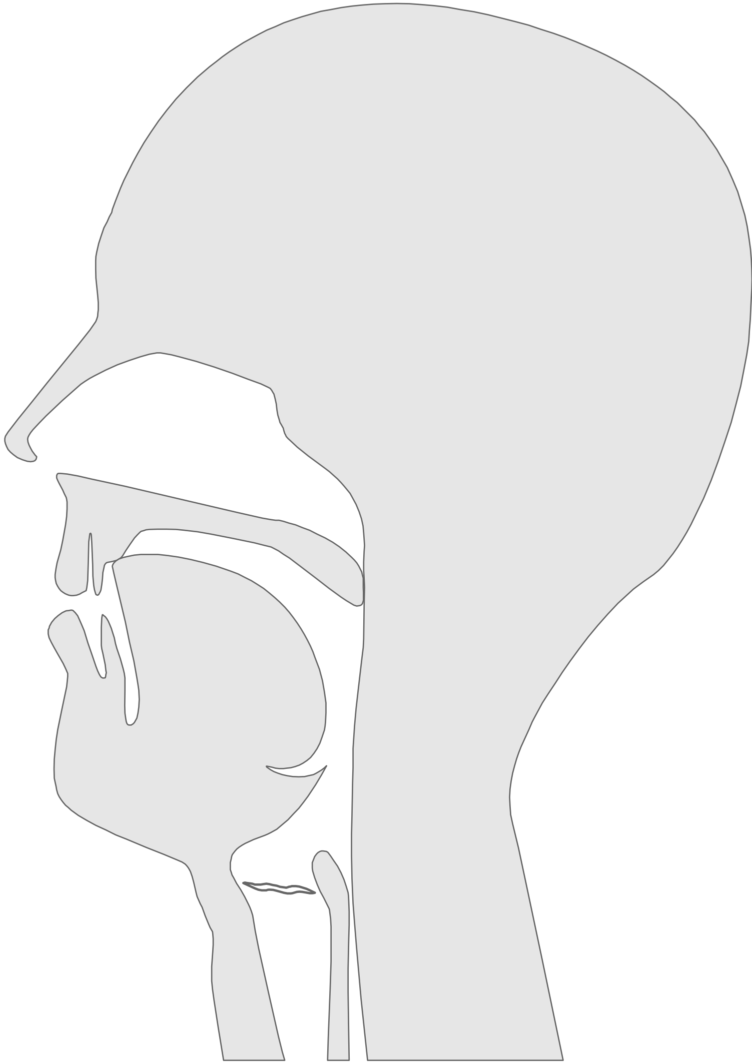
Consonant Articulations

ŋ



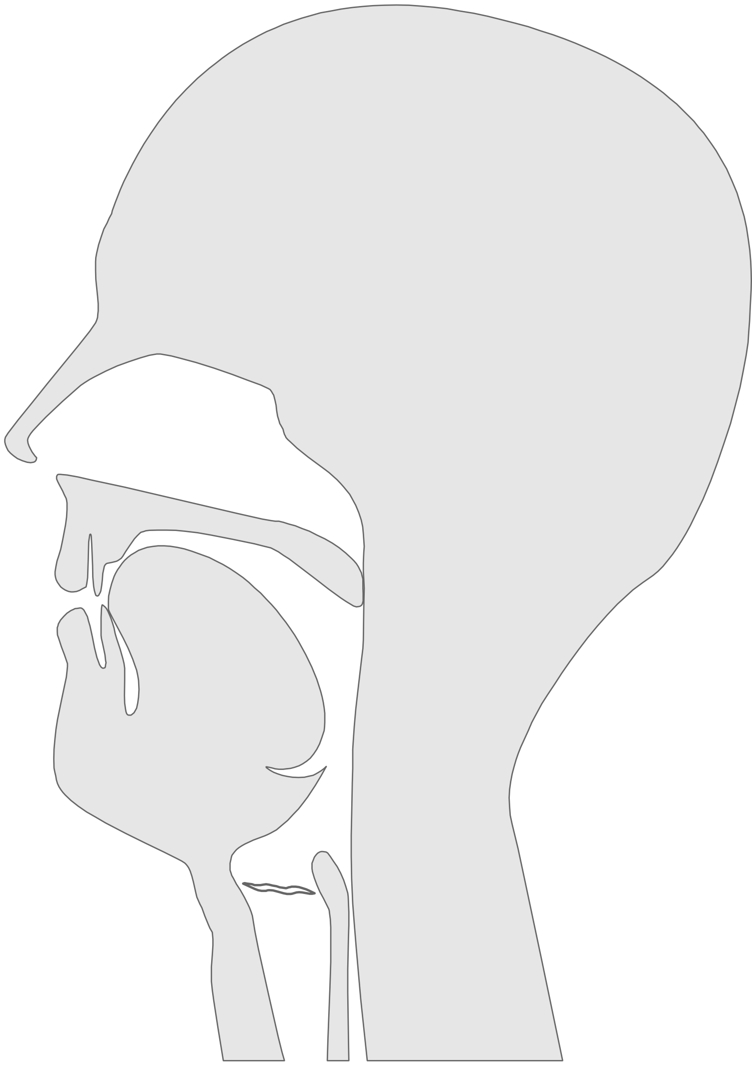
Consonant Articulations

d



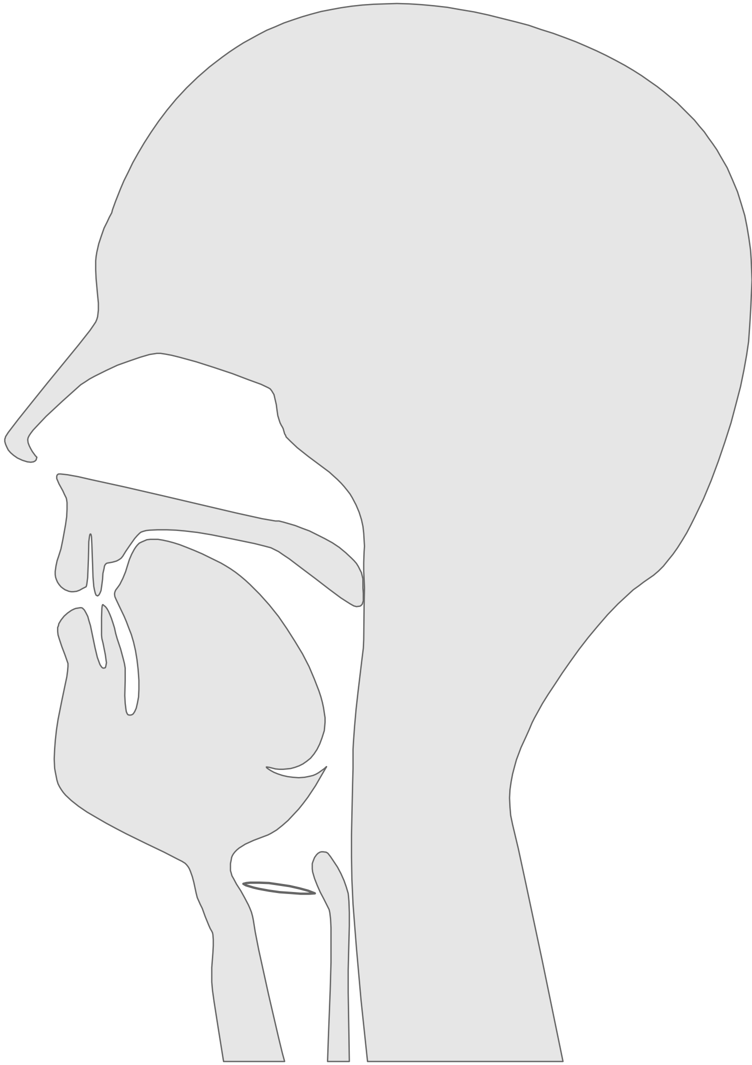
Consonant Articulations

z̥



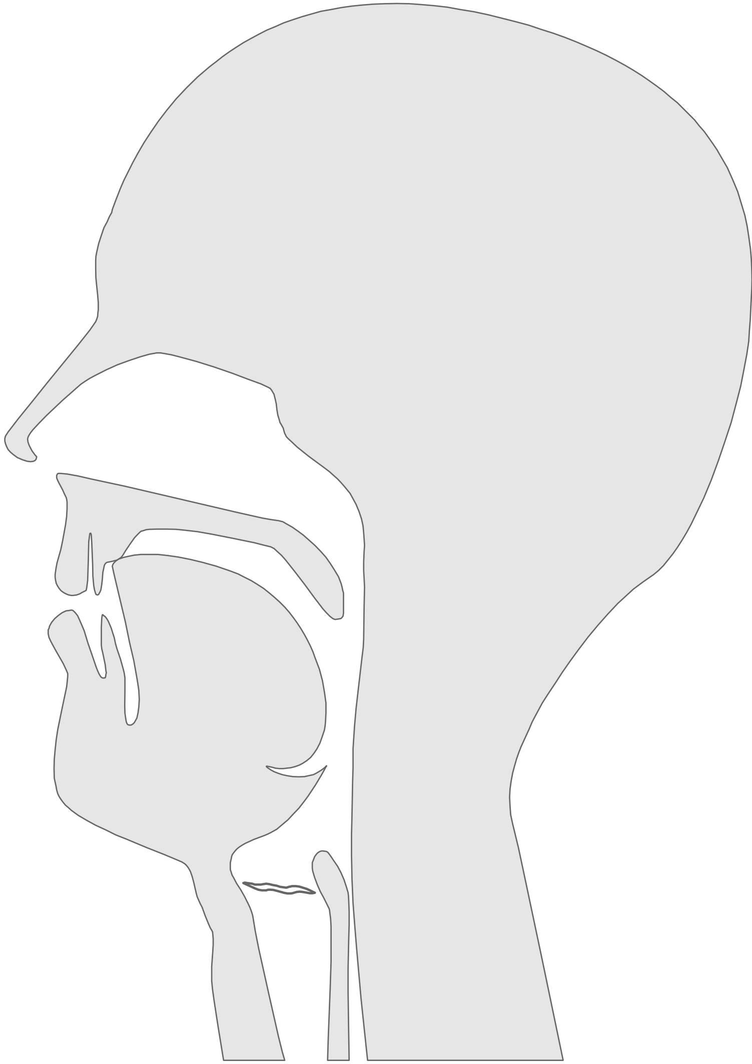
Consonant Articulations

ʃ



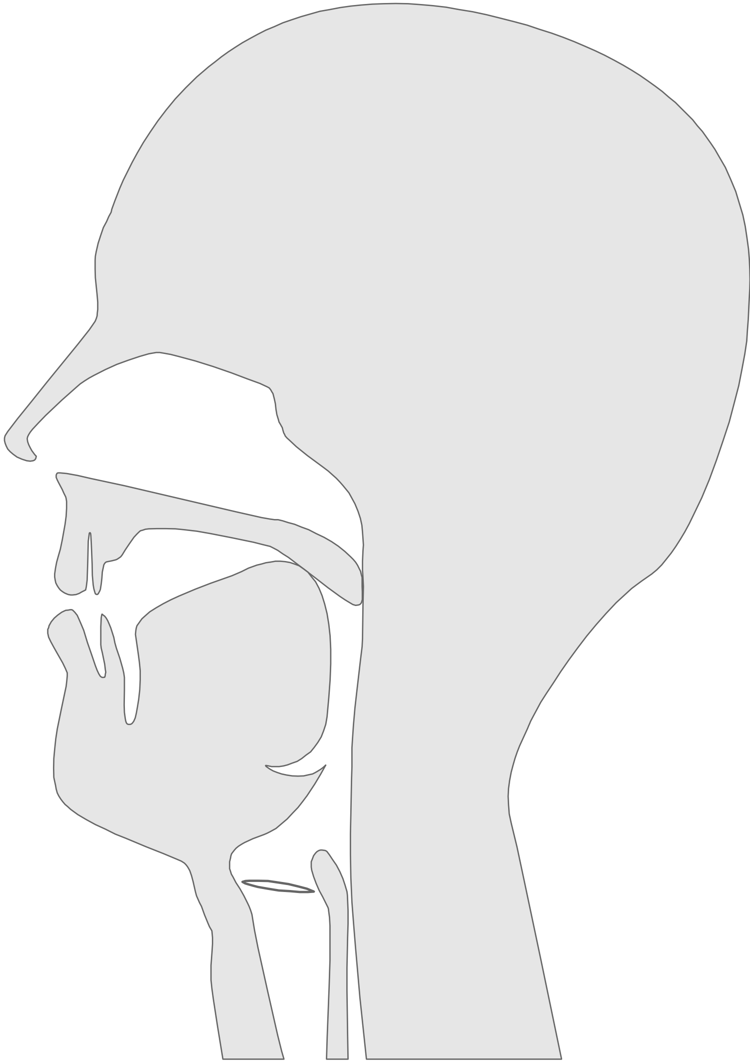
Consonant Articulations

n



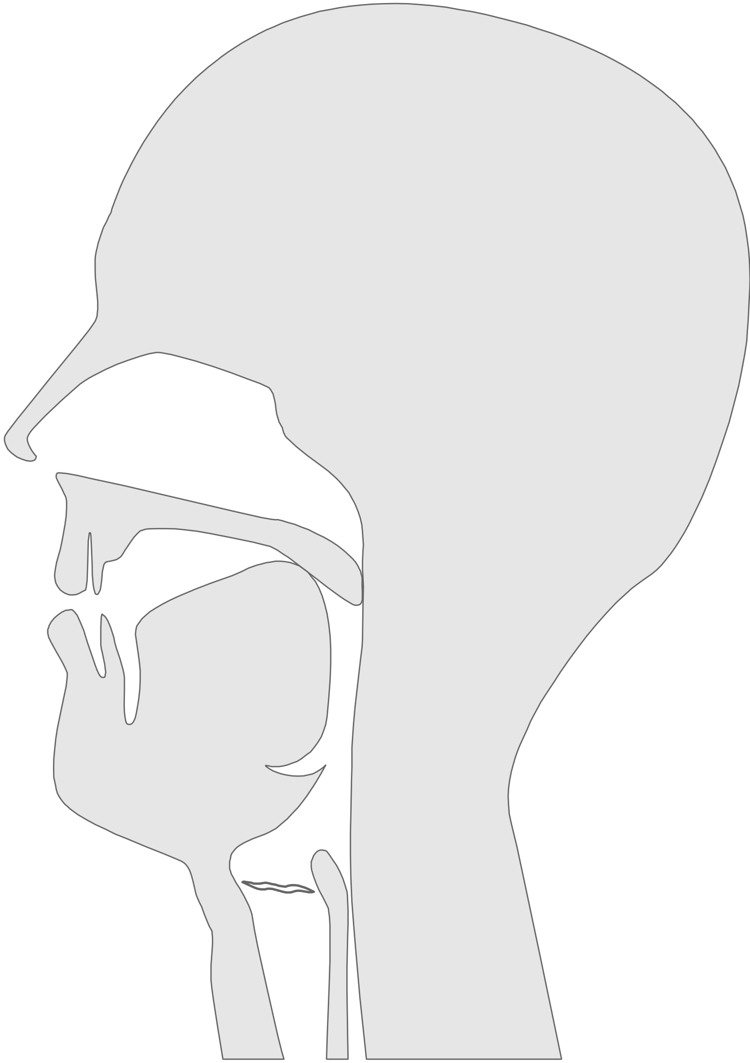
Consonant Articulations

k



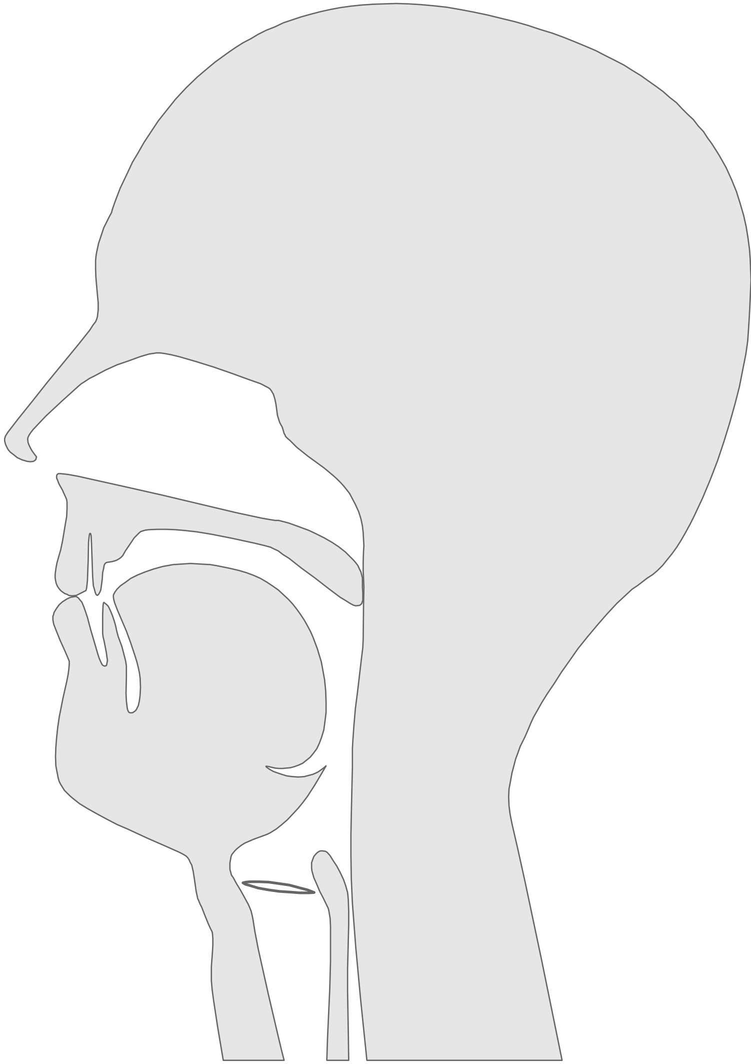
Consonant Articulations

g



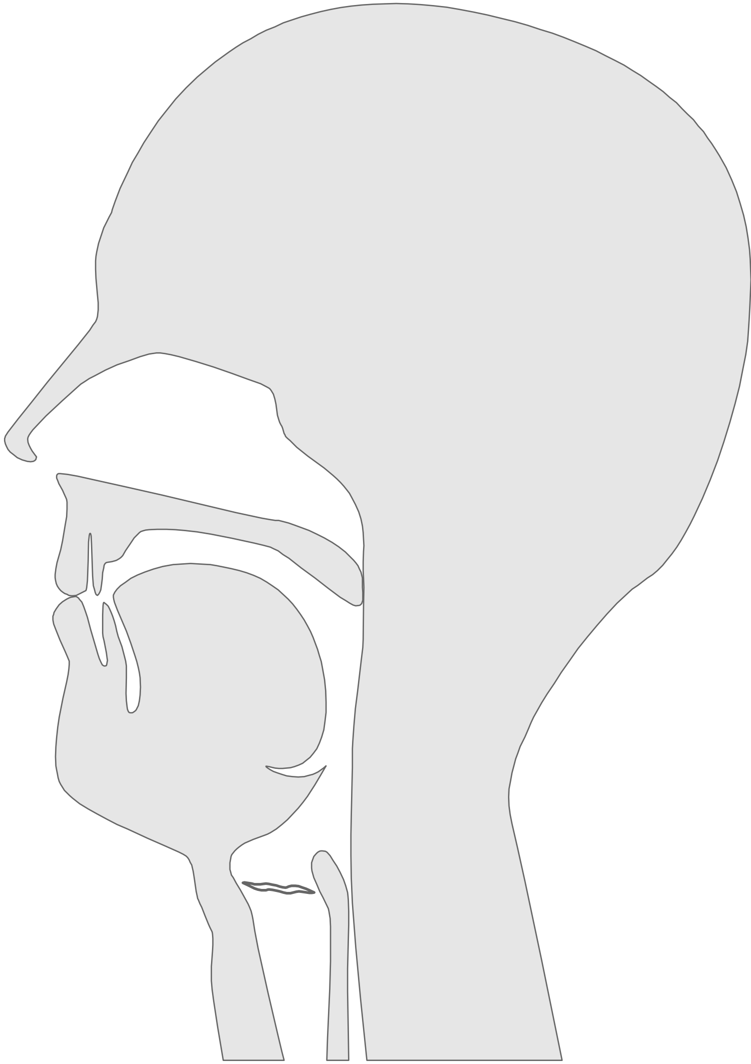
Consonant Articulations

p



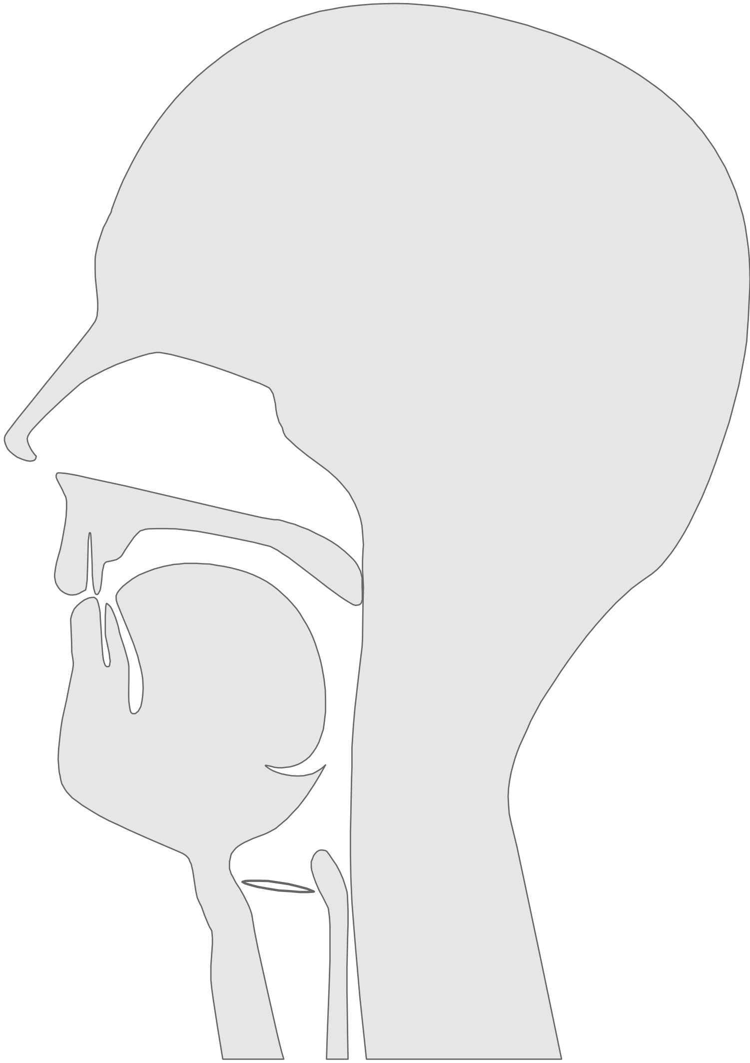
Consonant Articulations

b



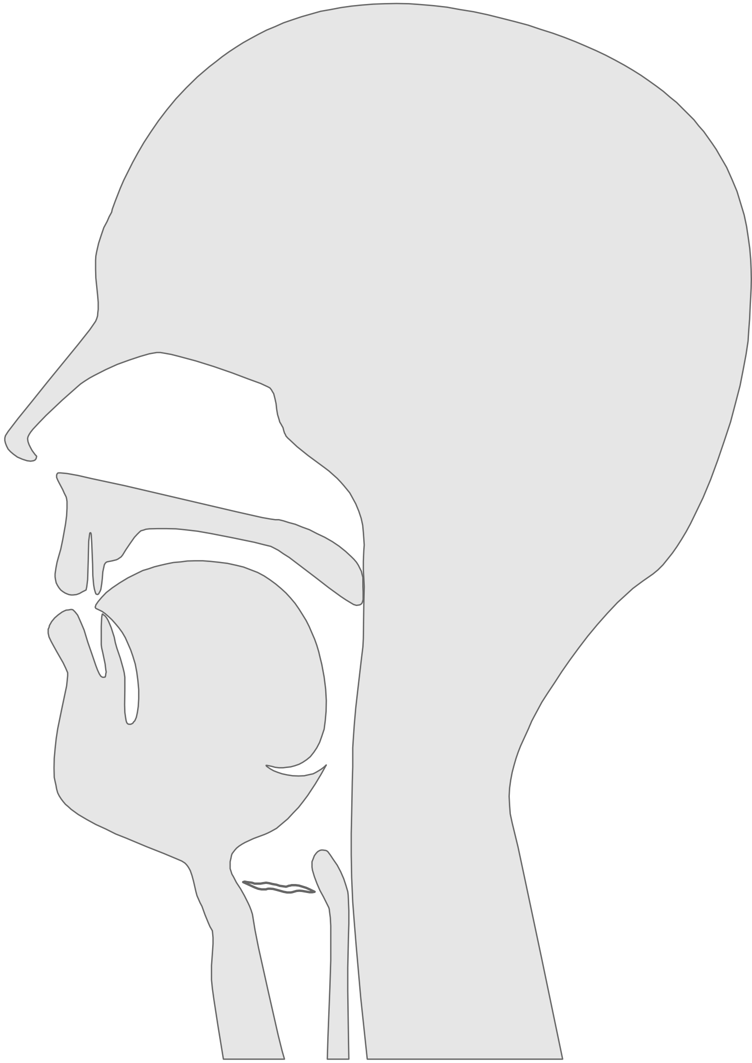
Consonant Articulations

f



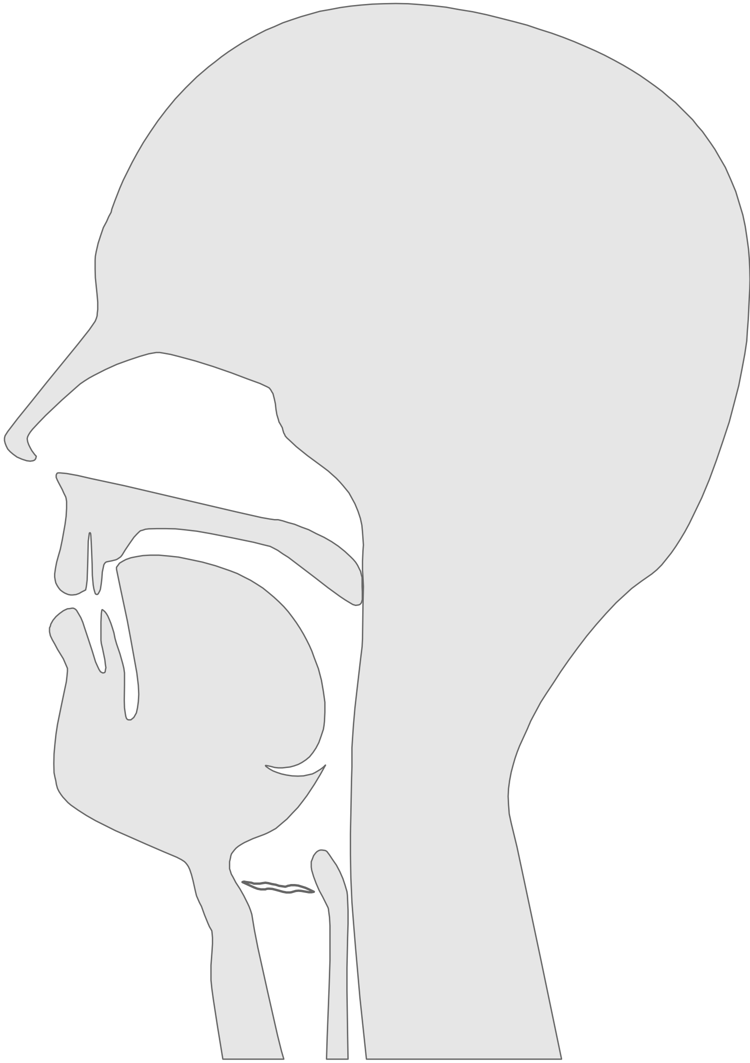
Consonant Articulations

ð



Consonant Articulations

z



English Vowels

Features and Alternations

Read: LJ 4

Finish HW: English word transcriptions

The Vowels of English

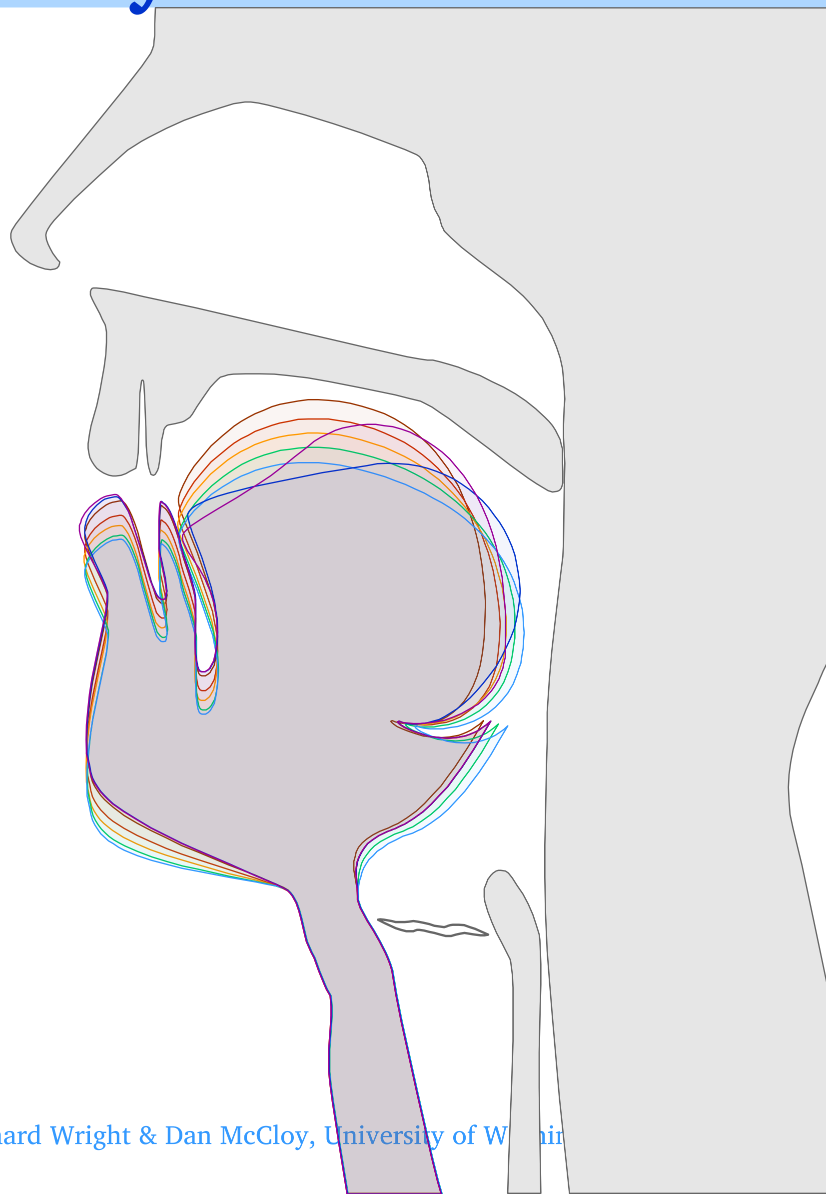
Chart based on Ladefoged *Vowels and Consonants* (2001) p. 28.

General American	Received Pronunciation (British)	b_d/t	k_d/t	h_d	h_t	h_	_r
i		bead	keyed	heed	heat	he	beer
ɪ		bid	kid	hid	hit		
eɪ		bayed	Cade	hayed	hate	hay	
ɛ		bed		head			bare
æ		bad	cad	had	hat		
ɑɹ	ɑ	bard	card	hard	heart		bar
ɑ	ɒ	body	cod	hod	hot		
ɔ		bawd	cawed	hawed		haw	bore
ʊ		Buddhist	could	hood			
oʊ	əʊ	bode	code	hoed		hoe	(boar)
u		booed	cooed	who'd	hoot	who	tour, poor
ʌ		bud	cud	Hudd	hut		
ɝ	ɜ	bird	curd	herd	hurt	her	burr, purr
aɪ		bide	kite	hide	height	high	fire
aʊ		bowed	cowed	how'd		how	hour
ɔɪ		Boyd			Hoyt	(a)hoy	(coir)
iu/ju		butte	cute	hued		hue	pure

Vowel Articulations Overlay

- See also: mid-sagittal x-ray video of five point vowels: http://www.phonetics.ucla.edu/vowels/chapter11/tongue.html

i
ɪ
ɛ
æ
a
ʊ
u



Vowel Quality and Feature Classification

- *Vowel quality* is a term for the auditory basis on which we distinguish one vowel from another.
  - Recall that we distinguish consonants based on *articulatory* features like *voicing*, *place*, and *manner* of articulation. Similarly, we distinguish vowels based on a combination of features, but vowels are commonly classified based on *auditory features* (characteristics of how the vowel sounds), not articulatory ones.
  - The primary classification of vowels occurs along two dimensions described by two pairs of features: Front–Back and High–Low.
    - A secondary (articulatory) feature that can be added to a vowel at any height or backness is lip rounding: Rounded–Unrounded.
  - An additional auditory feature is needed to fully distinguish some pairs of vowels: Tense–Lax.

Examples of Vowel Features

- Examples by class:
  - High vowels: “key” /ki/, “coo” /ku/
  - Low vowels: “cat” /kæt/, “cot” /kat/
  - Tense vowels: “heed” /hid/, “aid” /eɪd/, “who’d” /hud/
  - Lax vowels: “hid” /hɪd/, “Ed” /ɛd/, “hood” /hʊd/
- Examples by contrast:
  - High front vowel “bead” /bid/ vs. low front vowel “bad” /bæd/
  - High front vowel “key” /ki/ vs. high back vowel “coo” /ku/
  - High front tense vowel “heed” /hid/ vs. high front lax vowel “hid” /hɪd/

Monophthongs, Diphthongs and Triphthongs

- Vowels which have a single quality are called *monophthongs*, vowels which move between two or three qualities within one syllable are called *diphthongs* or *triphthongs*.
- Diphthongs are comprised of a *nucleus* and a *glide*. *On-glides* come before the nucleus, *off-glides* after. Triphthongs have two glides (two on-glides, two off-glides, or one of each).
  - Diphthong end points are often different from the monophthong qualities in the system.
  - Off-glides often fall short of their phonemic targets.
  - In English diphthongs, the starting quality is tense and the ending quality is lax. Examples:

    "pie" /paɪ/, "cow" /kaʊ/, "toy" /tɔɪ/, "pay" /peɪ/, "toe" /toʊ/

Transcribing Diphthongs

- Monophthongs are written with a single IPA symbol, diphthongs are written with two, triphthongs with three.
  - Glides are transcribed with normal vowel symbols: [aʊ], optionally with the 'non-syllabic' diacritic: [aʊ̯], or sometimes as a raised vowel: [aᶷ].
- Some linguists consider "cute" to contain a diphthong, usually transcribed as /kjut/.
  - Strictly speaking this is incorrect because transcribing /j/ instead of /i/ implies that the glide is behaving phonologically like a *consonant*, and only *vowels* can form diphthongs.
  - This "on-glide" /j/-sound is more accurately transcribed as a true diphthong: /kiʊt/, or as a palatalization of the preceding consonant: [kʲut]

/ɹ/-sounds and Rotacization

- A *rhotic vowel* is a vowel produced with an "r-like" quality; this feature is called *rhoticization* or *r-coloring*.
  - Rhotic vowels are often transcribed by a small "wing" attached to the vowel symbol: "purr" [pʰɝ], "heard" [hɝd], "murder" [ˈmɝɾɚ]. Another way to transcribe rhotic vowels is to consider them syllabic consonants: "murder" [ˈmɹ̩ɾɹ̩]
  - Some linguists consider vowels followed by /ɹ/ to be diphthongs or triphthongs rather than rhoticized vowels: "are" /ɑɹ/, "err"/ɛɹ/, "or"/ɔɹ/, "ear" /iɹ/, "ire" /ɑɪɹ/
- Note that rhoticization is not always the result of a vowel being coarticulated with a following /ɹ/ sound; in some languages certain vowels are always rhotic regardless of the following sound.

Stressed and Unstressed Vowels

- In many languages some syllables are more prominent than others. They are often louder and have a different pitch than neighboring syllables, and are usually longer in duration. These are called *stressed syllables*.
- In *unstressed syllables* vowel quality contrasts may become diminished or may merge into a single quality. This is called *vowel reduction*. The quality of reduced vowels is usually similar to the mid vowel /ʌ/ found in the word "cup."
- By convention, reduced vowels are transcribed with the symbol [ə] (called *schwa*). Depending on the language or dialect, there may be a variety of reduced vowel qualities that actually occur in unstressed syllables: the lax high front vowel [ɪ] and the high central vowel [ɨ] are common in unstressed syllables for many speakers of English.

Tense vs. Lax

- Tenseness is primarily a way to group vowels based on their behavior in different phonological environments.
  - For describing American English, the tense-lax distinction is only needed to separate pairs of non-low vowels.
    - Lax vowels generally occur in *closed syllables* (those with codas), while tense vowels can also occur in *open syllables* (those without codas). Following this, /a/ (or [ɑ]) groups with tense vowels and /æ/ with lax.
    - Lax vowels are more central and shorter than tense counterparts.
  - See Table 4.2, p. 99 and listen to the words on CD 4.2.
    - Words in parentheses indicate vowels that don't exist in the indicated environment in many dialects. That is, most dialects have no or very few /ɔɪɹ/ and /iʃ/ words, and most pronounce /oʊɹ/ words as /ɔɹ/.

Some Common Vowel Alternations

- *Vowel shortening*: In English, vowels occurring before voiceless *obstruents* (stops, fricatives, affricates) are measurably shorter than before voiced consonants (example: "cab" [kæb] vs "cap" [kæ̆p]).
  - Vowels are longer in open syllables, stressed syllables, and monosyllabic (one-syllable) words.
- *Vowel devoicing*: unstressed vowels that occur after voiceless consonants are often devoiced (example: "potato" [pə̥ˈtʰeɪɾoʊ]).
- *Vowel nasalization*: vowels that occur before nasal consonants are often nasalized (example: "can" [kæ̃n]).
- *Canadian raising*: a characteristic of some dialects of North American English in which the nucleus of diphthongs that occur before voiceless consonants are raised. Example: "fight" [fʌɪt] (raised) vs. "fire" [faɪɹ] (non-raised).

Reference Chart: (Standard American) English Vowels

	Front		Central		Back	
	Unround	Round	Unround	Round	Unround	Round
High Tense	i		ɨ			u
High Lax	ɪ		—	—	—	ʊ
Mid Tense	e		ɜ **			o
Mid Lax	ɛ		ə		ʌ	ɔ
Low	æ		a		ɑ *	ɒ **

* some American dialects

** some British dialects

Feature values based on Hayes (2009) p98.

Basic Phonology

Alternations, Rules, Features, Natural Classes

(A refresher)

Suggested: Review: *Language Files* ch. 3

Phonology Terms (Review)

- Sounds are *contrastive* if interchanging them changes meaning
  - *Minimal pair/set*: words that differ by a single contrastive sound (phoneme)
- *Phoneme*: mental representation of a set of non-contrastive sounds with an unpredictable distribution
  - *Contrastive distribution*: when sounds (phonemes) occur in the same phonetic environment (i.e. to form minimal pairs/sets), they are in an unpredictable distribution and change meaning
- *Allophones*: non-contrastive realizations of a phoneme with predictable distributions based on phonetic/phonological environment (context)
  - *Complementary distribution*: when sounds (allophones) never occur in the same environment, they have a predictable, non-overlapping distribution determined by context and describable by phonological rules

Analogy of allophones and complementary distribution

/Kal-El/

[Superman] (emergencies)

[Clark Kent] (elsewhere)

Alternations

- A single phoneme may be *realized* (produced) as a variety of different allophones. Allophonic variation is called *alternation*.
- Alternations are systematic variations in the pronunciation of a phoneme and can be described by sets of rules that change a phoneme into a particular allophone depending on its *environment* (what sounds are around it, whether it receives stress or not, etc).
- Example:
  - /t/ in "team" is pronounced with aspiration but /t/ in "lit" is not; the /t/ in "steam" also lacks aspiration. In fact, /t/ is *aspirated* (produced with aspiration) in only two environments in English: as the first sound in a word (e.g., the /t/ in "team"), or as the first sound of a syllable with emphasis (stress) (e.g., the first /t/ in "interrogate").

English

"team" "steam" "(s)team"

http://archive.phonetics.ucla.edu/

Phonological Rules

- Alternations can be described by sets of phonological rules that describe the environments or conditions under which a given allophone surfaces.
  - Example:
    /t/ → [tʰ] at the beginning of words or stressed syllables
    /t/ → [ɾ] between (stressed & unstressed) vowels
    /t/ → [ʔ] before /n/
    /t/ → [t] elsewhere
- Sometimes several phonemes undergo similar alternations:

  /p t k/ → [pʰ tʰ kʰ] at the beginning of words or stressed syllables
- Describing sounds in terms of *features* can make it easier to describe such patterns.

Phonological Features

- *Phonological features* (sometimes called *distinctive features*) allow us to describe phonemes or phones as combinations of characteristics.
- The way sounds are described in the IPA chart is a feature system; e.g., for consonants:
  - [place] has 12 values
  - [voicing] has 4 values (voiceless, breathy, modal, creaky)
  - [manner] has 6 values (stop, fric, affr, approx, flap, trill)
    - To distinguish /ɹ/ from /l/ you must either split [manner: approx] into "central" and "lateral" or add a fourth feature [± lateral].
    - Likewise, either split [manner: stop] into "oral" and "nasal" or add a binary feature [± nasal].
    - What about aspiration? or unreleased stops? or the distinction between "dental" and "interdental"?

Phonological Features (cont.)

- Distinctive features can be *auditory* (describe how a segment sounds) or *articulatory* (describe how a segment is produced), though current phonological theory tends to emphasize articulatory features.
- The feature system used in this class (from Hayes 2009) uses pseudo-binary features (either +, – or “0” meaning “unvalued”).
- Example:
  - The voiced bilabial stop /b/ can be described as the conjunction of features [+consonantal, +voice, +labial].
  - /b/ is [0tense] (unvalued for the [tense] feature; the [tense] feature only applies to the consonants [j w ʍ ɥ] and non-low vowels).

Phonological Features (cont.)

- Phonological features also provide a way of organizing the phonemes of a language into hierarchical groups.
  - Example: the words “pie” “tie” “kite” “fight” “thigh” “sigh” “shy” “high” all begin with *voiceless consonants*: /p t k f θ s ʃ h/. Those phonemes can be grouped together as all having the same *voicing* feature: [–voice].
- In this way, phonological features serve a dual purpose. In one sense, they are *smaller conceptual units* than phonemes because they represent only *part* of a phoneme’s articulation.
- In another sense, features describe *higher-level categories* of sounds than phonemes do, because they can be used to pick out classes of phonemes.

Natural Classes

- Phonemes that share many feature values tend to undergo similar alternations in similar environments. Groups of phonemes which share sufficient feature values to have similar alternations are referred to as *natural classes*.
- Example:
  - the consonants /p t k/ are *aspirated* at the beginnings of the words “pie” “tie” “chi”…
  - but are *unaspirated* in “spy” “sty” “sky”…
  - and are (often) *unreleased* word-finally, as in “lip” “lit” “lick”.
  - Thus /p t k/ undergo similar alternations in similar environments, and therefore they form a natural class of *voiceless stops* described by the feature values [+consonantal –continuant –voice].

Coarticulation

- In human speech, adjacent sounds are always slightly overlapped because we physically can't separate the articulations completely. Because of this, some articulations are altered to become more like adjacent articulations. This process is called *coarticulation*.
- Coarticulation is most notable for causing changes in *place of articulation*. Example:
  - in "tenth" the alveolar nasal /n/ becomes dental [n̪]
  - in "inconsiderate" the alveolar nasal /n/ becomes a velar nasal [ŋ]

Types of Coarticulation

- *Anticipatory coarticulation*: a consonant takes on one or more features of a sound that follows it. Example:
  - in “wealth,” alveolar /l/ becomes dental [l̪] before the dental /θ/
- *Preservative coarticulation*: a consonant takes on one or more features of a sound that preceded it. Example:
  - in “play,” voiced /l/ becomes voiceless [l̥] after the voiceless /p/
- Coarticulation underlies many of the common alternations we will see in this course.

Phonemic Analysis

Allophonic Variations and Phonological Rules

Read: Hayes 2

Suggested (optional): Read: Hayes 3

HW: Phonemic Analysis

Phonology Defined

- Phonology is the study of *sound systems* in language. This includes:
  - describing what sets of phones native speakers think of as "the same sound" (*phoneme*);
  - describing the ways neighboring sounds influence each other, and the rules that determine which version (*allophone*) of an underlying sound is realized in a certain environment;
  - describing the permissible sequences of sounds in a language (*phonotactics*).

Phonemes and Minimal Pairs

- The first stage in a phonological analysis of a language is accurately describing the set of *contrasts*. Contrastive sounds serve to differentiate words in a language.
- Sounds that contrast belong to different *phonemes*. Two words that differ by a single contrastive sound (phoneme) are called a *minimal pair*. Larger groups of words that all differ by a single phoneme are called *minimal sets*.
- A single phoneme may be pronounced in a variety of ways that don't change the meanings of words. These non-contrastive variant sounds are called *allophones*.

Alternations

- A single phoneme may be *realized* (produced) as a variety of different allophones. Allophonic variation is called *alternation*.
- Alternations are systematic variations in the pronunciation of a phoneme, and can be described by sets of rules that change a phoneme into a particular allophone depending on its *environment* (what sounds are around it, whether it receives stress or not, etc).
- Example:
  - /p/ in “pie” is pronounced with aspiration but /p/ in “lip” is not; the /p/ in “spy” also lacks aspiration. In fact, /p/ is *aspirated* (produced with aspiration) in only two environments in English: as the first sound in a word (e.g., the /p/ in “pie”), or as the first sound of a syllable with emphasis (stress) (e.g., the /p/ in “impressed”).

Phonological Rule Notation

- To describe allophonic alternations, phonologists write *phonological rules* that capture the important features of the linguistic environment that trigger the alternation.

- The standard syntax for phonological rules is:

/phoneme/ → [phone] / (environment1) ___ (environment2)

/phoneme/	→	[phone]	(environment1)	___	(environment2)
underlying form	spoken as	pronun-ciation	when preceded by environment 1	location of the segment that is changing	and/or followed by environment 2

Example: "/p/ *becomes aspirated at the beginning of words*"

$$/p/ \rightarrow [p^h] \;/\; [_{word} \,\_\_$$

Note that if only one environment triggers the alternation, then only one environment is specified in the rule. "$[_{word}$" indicates the beginning of a word; "#" is another common way to indicate this.

Phonological Rule Notation: Disjunctive Environments

- Recall however that /p/ becomes aspirated not only at the beginnings of words, but also at the onset of stressed syllables (e.g., "impress" [ɪmˈpʰɹɛs]).
- One way to handle this would be to write a rule describing multiple disjunctive environments that trigger the same change. This can be done with a "curly bracket":

Example: *"/p/ becomes aspirated at the beginning of words or at the onset of stressed syllables"*

$$/p/ \rightarrow [p^h] \;\; / \left\{ \begin{array}{l} [_{word} \;\_\_ \\ [_{syllable} \;\_\_\; [+stress] \end{array} \right.$$

Phonological Rule Notation: Segments vs. Features

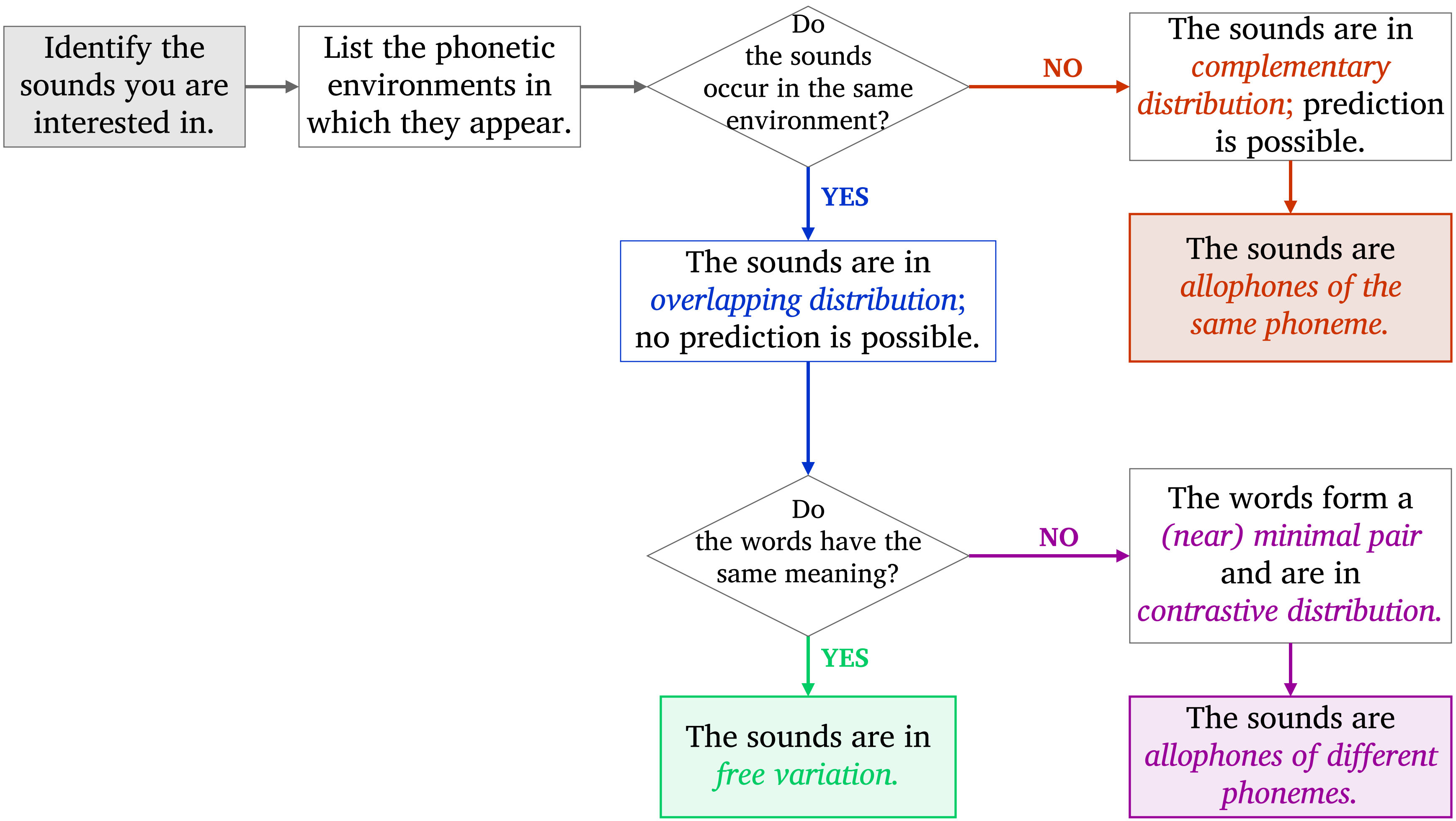
- A further refinement comes when we recall that aspiration happens to other voiceless stops besides /p/. Rather than writing separate rules for /p/, /t/, and /k/, we can write a single rule to describe all three alternations. This is done using phonological features.

Example: *"voiceless stops are aspirated when they are syllable initial and precede a stressed vowel."*

$$\begin{bmatrix} -\text{continuant} \\ -\text{voice} \end{bmatrix} \rightarrow \begin{bmatrix} +\text{spread glottis} \end{bmatrix} \;/\; [_{\text{syllable}} \;\_\_\_ \begin{bmatrix} +\text{syllabic} \\ +\text{stress} \end{bmatrix}$$

Hayes (2009) p122.

Solving Phonology Problems: Flowchart



Adapted from *Language Files 10* (2007), p. 133

Sample Phonology Problem: Spanish

1. What are we interested in? [d] vs [ð]

   Write out the environments in which the sounds occur.

Data

dɾama	"drama"
doloɾ	"pain"
dime	"tell me"
kaða	"each"
laðo	"side"
oðio	"hatred"
komiða	"food"
anda	"scram"
sweldo	"salary"
duɾaɾ	"to last"
toldo	"curtain"
falda	"skirt"

2. Look at the environments. The sounds are in ______________ distribution:

3. Since the sounds are in this distribution, they are likely:

   a) allophones of the same underlying phoneme, or

   b) two different phonemes.

4. Write the rule:

Sample Phonology Problem: Spanish

1. What are we interested in? [d] vs [ð]

   Write out the environments in which the sounds occur.

Data

dɾama	"drama"
doloɾ	"pain"
dime	"tell me"
kaða	"each"
laðo	"side"
oðio	"hatred"
komiða	"food"
anda	"scram"
sweldo	"salary"
duɾaɾ	"to last"
toldo	"curtain"
falda	"skirt"

Environments

[d]	[ð]
#_ɾ	a_a
#_o	a_o
#_i	o_i
#_u	i_a
n_a	
l_o	
l_a	

2. Look at the environments. The sounds are in ____________ distribution:

   [d] never occurs after a vowel
   [ð] always occurs after a vowel

   Note that although [ð] *always occurs* *<u>between</u> vowels in this data set, the following vowel is not conditioning the change, since* [d] *can be followed by vowels too (e.g.,* [anda]*).*

3. Since the sounds are in this distribution, they are likely:

   a) allophones of the same underlying phoneme, or

   b) two different phonemes.

4. Write the rule: /d/ → [ð] / V__

Phonological Features

Grouping phones

Read: Hayes 4

HW: Natural Classes

HW: Phonological Features

What's Wrong with the IPA Features?

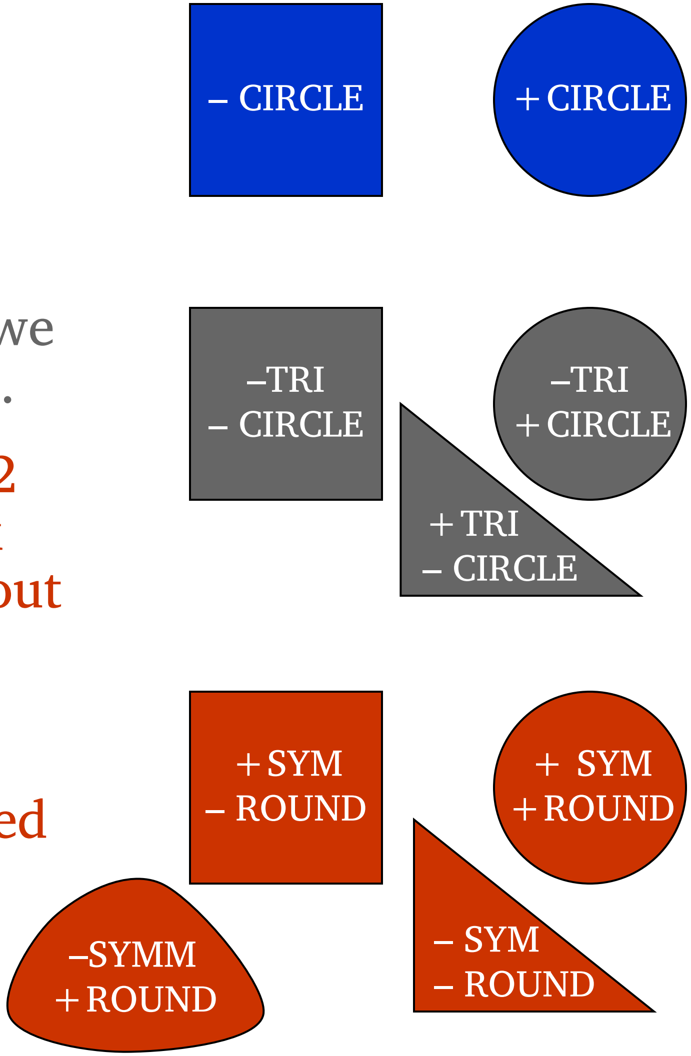
- The feature set used by the IPA is designed to clearly and simply pick out each individual sound from all other sounds.
- It also does a reasonable job of picking out basic *classes of sounds* like "fricatives" or "alveolar consonants," but there are some classes of sounds that are not easily described using IPA features.
  - Example: the class of stops, affricates, and fricatives make up a natural class called *obstruents*. In many languages, all obstruents (and only obstruents) undergo similar phonological changes like *word-final devoicing* or *voicing assimilation to a following consonant in a cluster*.

Which Features to Choose?

- In order to easily describe patterns of phonological rules that apply only to obstruents, we need some feature or set of features that uniquely describes *all obstruents and only obstruents*.
  - The feature that does this is [sonorant]: all obstruents are [–sonorant]
  - All *sonorants* (everything that isn't an obstruent) are [+sonorant]

How Many Features Do We Need?

- If we restrict ourselves to binary features:
  - In a universe with 2 kinds objects we need 1 feature to describe everything.
  - In a universe with 3 kinds of objects, we need 2 features to describe everything.
  - In a universe with 4 kinds of objects, 2 features will still be enough if we pick the features carefully (i.e., if we pick out more abstract features that several objects share). In reality, objects are rarely similar in ways that are most efficiently describable, so we often need more features than the minimum required mathematically.



Consonant Manner Features

Vowels	Glides	Liquids	Nasals	Stops	Affricates	Fricatives
+ syllabic	– syllabic					
– consonantal		+ consonantal				
+ approximant			– approximant			
+ sonorant				– sonorant		
+ continuant			– continuant			+ cont.
0 delayed release				– d.r.	+ delayed release	

Consonant Manner Features: Describing Natural Classes of Manner

- Use the minimum number of features necessary to uniquely pick out the sounds you're interested in:

– Vowels:	[+syllabic]
– Glides:	[–syllabic][–consonantal]
– Liquids:	[+consonantal][+approximant]
– Nasals:	[+sonorant][–approximant]
– Fricatives:	[–sonorant][+continuant]
– Affricates:	[–continuant][+delayed release]
– Stops:	[–delayed release]
– Stops & Affricates:	[–sonorant] [–continuant]
– Liquids & Glides:	[–syllabic][+approximant]
– Liquids, Glides, & Nasals:	[–syllabic][+sonorant]

Consonant Place Features

- Major place features:
  - [labial]: active articulator is lower lip.
    - [labiodental] passive articulator is the upper teeth.
  - [coronal]: active articulator is tongue blade or tip.
    - [distributed]: active articulator is tongue blade (vs. tongue tip): [+distributed] = *laminal* (blade), [-distributed] = *apical* (tip)
    - [anterior]: passive articulator is alveolar ridge or teeth.
    - [strident]: "the airstream is channeled through a groove in the tongue blade and blown at the teeth." [s z ɕ ʑ ʃ ʒ ʂ ʐ], a.k.a. *sibilants*
  - [dorsal]: active articulator is tongue body.
    - Subdivisions use the vowel features [high], [low], [front], [back].

Consonant Place Features (cont.)

- Other place features:
  - [lateral]: air flows around the side(s) of the tongue: [l ɬ ɮ ɺ ɭ ʎ ɫ ʟ].
    All other consonants are [– lateral] (a.k.a. *central*).
  - [round]: includes lip rounding: [w ʍ ɥ] and anything with the [ʷ] diacritic. All other consonants are [– round].
  - [tense]: applies to the semivowels [j ɥ w ʍ].
    All other consonants are [0 tense].

Consonant Place Features
(Using fricatives as examples)

<table>
<tr><th colspan="2">bilabial</th><th colspan="2">labio-dental</th><th colspan="2">inter-dental</th><th colspan="2">alveolar</th><th colspan="2">alveolo-palatal</th><th colspan="2">palato-alveolar</th><th colspan="2">retroflex</th><th colspan="2">palatal</th><th colspan="2">velar</th><th colspan="2">uvular</th><th colspan="2">pharyn-geal</th><th colspan="2">glottal</th></tr>
<tr><td>ɸ</td><td>β</td><td>f</td><td>v</td><td>θ</td><td>ð</td><td>s</td><td>z</td><td>ɕ</td><td>ʑ</td><td>ʃ</td><td>ʒ</td><td>ʂ</td><td>ʐ</td><td>ç</td><td>ʝ</td><td>x</td><td>ɣ</td><td>χ</td><td>ʁ</td><td>ħ</td><td>ʕ</td><td>h</td><td>ɦ</td></tr>
<tr><td colspan="4">+ labial</td><td colspan="20">– labial</td></tr>
<tr><td colspan="2">– labio-dental</td><td colspan="2">+ labio-dental</td><td colspan="20">– labiodental</td></tr>
<tr><td colspan="4">– coronal</td><td colspan="12">+ coronal</td><td colspan="8">– coronal</td></tr>
<tr><td colspan="4">0 anterior</td><td colspan="6">+ anterior</td><td colspan="6">– anterior</td><td colspan="8">0 anterior</td></tr>
<tr><td colspan="4">0 strident</td><td colspan="2">– strid.</td><td colspan="8">+ strident</td><td colspan="2">– strid.</td><td colspan="8">0 strident</td></tr>
<tr><td colspan="4">0 distributed</td><td colspan="2">+ dist.</td><td colspan="2">– dist.</td><td colspan="4">+ dist.</td><td colspan="2">– dist.</td><td colspan="2">+ dist.</td><td colspan="8">0 distributed</td></tr>
<tr><td colspan="14">– dorsal</td><td colspan="8">+ dorsal</td><td colspan="2">– dorsal</td></tr>
<tr><td colspan="14">0 high</td><td colspan="4">+ high</td><td colspan="4">– high</td><td colspan="2">0 high</td></tr>
<tr><td colspan="14">0 low</td><td colspan="6">– low</td><td colspan="2">+ low</td><td colspan="2">0 low</td></tr>
<tr><td colspan="14">0 front</td><td colspan="2">+ front</td><td colspan="6">– front</td><td colspan="2">0 front</td></tr>
<tr><td colspan="14">0 back</td><td colspan="4">– back</td><td colspan="4">+ back</td><td colspan="2">0 back</td></tr>
</table>

Laryngeal Features

- [+voice] applies to voiced sounds.
- [+implosive] applies to implosive sounds.
- [+spread glottis] applies to:
  - breathy-voiced sounds
  - aspirated consonants
  - [h]
- [+constricted glottis] applies to:
  - creaky-voiced sounds
  - ejectives
  - fortis stops in Korean
  - preglottalized consonants
  - [ʔ]

Laryngeal Features

voiceless (except [ʔ h ʍ])	breathy	modal	creaky
– voice	+ voice		
– spread glottis	+ spread glottis	– spread glottis	
– constricted glottis	– constricted glottis		+ constricted glottis

Vowel Features

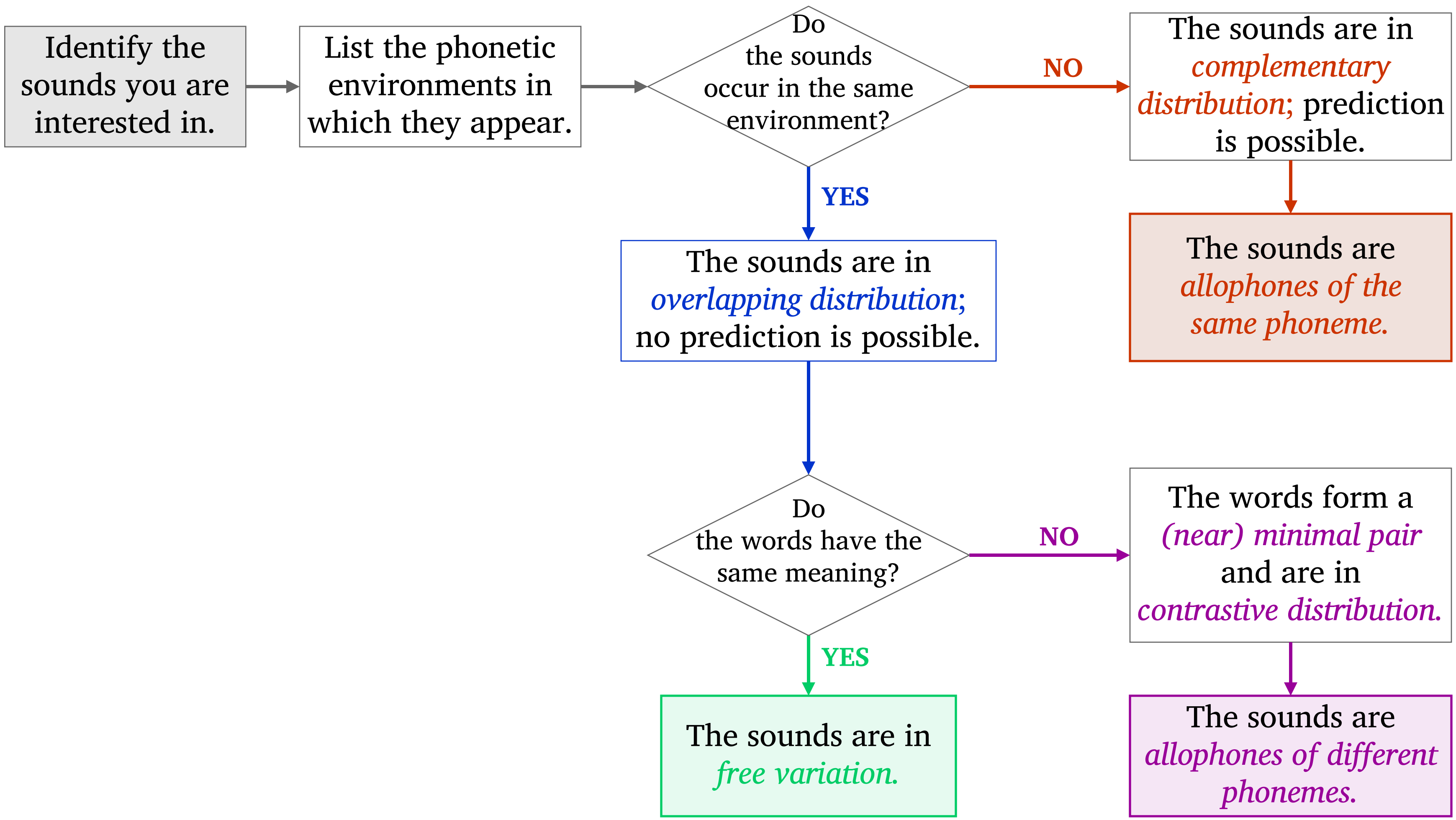
high front				high central		high back			mid front				mid central				mid back				low front		lo cen	low back	
i	y	ɪ	ʏ	ɨ	ʉ	ɯ	u	ʊ	e	ø	ɛ	œ	ɘ	ɵ	ɜ ə	ɞ	ɤ	o	ʌ	ɔ	æ	ɶ	a	ɑ	ɒ
+ high									– high																
– low																					+ low				
+ front				– front					+ front				– front								+ front		– front		
– back						+ back			– back								+ back				– back			+ back	
+ tense		– tense		+ tense				– t	+ tense		– tense		+ tense		– tense		+ tense		– tense		0 tense				
–	+	–	+	–	+	–	+	+	–	+	–	+	–	+	–	+	–	+	–	+	–	+	–	–	+

Other Vowel Features

- [ATR]: used for languages that have a vowel contrast between *advanced tongue root* [+ATR] and *retracted tongue root* [-ATR] articulations.
- [long]: used for languages with a vowel length contrast
- [nasal]: used for nasal or nasalized vowels
- Note that none of these distinctions need to be phonemic in order to use the feature. For example, English does not have any phonemically nasal vowels, but the feature [nasal] can be used to describe allophonic nasalization of English vowels:
- [+syllabic] → [+nasal] / ___ [+sonorant, – approximant]

  vowels *get nasalized* *before* *nasal consonants*

Solving Phonology Problems: Flowchart



Adapted from *Language Files 10* (2007), p. 133

Sample Phonology Problem: Spanish

1. What are we interested in? [d] vs [ð]

   Write out the environments in which the sounds occur.

Data

dɾama	"drama"
doloɾ	"pain"
dime	"tell me"
kaða	"each"
laðo	"side"
oðio	"hatred"
komiða	"food"
anda	"scram"
sweldo	"salary"
duɾaɾ	"to last"
toldo	"curtain"
falda	"skirt"

Environments

[d]	[ð]
#_ɾ	a_a
#_o	a_o
#_i	o_i
#_u	i_a
n_a	
l_o	
l_a	

2. Look at the environments. The sounds are in ____________ distribution:

   [d] never occurs after a vowel
   [ð] always occurs after a vowel

   Note that although [ð] *always occurs* *between vowels in this data set, the following vowel is not conditioning the change, since* [d] *can be followed by vowels too (e.g.,* [anda]*).*

3. Since the sounds are in this distribution, they are likely:

   a) allophones of the same underlying phoneme, or

   b) two different phonemes.

4. Now to write the rule…

Sample Phonology Problem: Spanish

- The rule stated with segments is: /d/ → [ð] / V__
- From this problem we know Spanish has the consonants [d k m n ɾ f ð s l w] and the vowels [i e a u o].
  - In reality there are many other sounds in Spanish (e.g., [β x r]), but these need not concern us here: *use what you are given in the problem.*
  - [–sonorant, +voice] is sufficient to pick out just [d] and [ð] from our given set of sounds.
  - [+continuant] is sufficient to distinguish them from each other; [+distributed] or [+delayed release] would also work.
  - As far as our data goes there are no syllabic consonants, so [+syllabic] suffices to pick out all vowels.
- The final rule, using features: $\begin{bmatrix} -\text{sonorant} \\ +\text{voice} \end{bmatrix} \rightarrow [+\text{continuant}] \;/\; [+\text{syllabic}]\;\_\_\_$

Sample Problem: Chamorro Vowel Harmony

Chamorro Data

gumə	"house"	i gimə	"the house"
tomʊ	"knee"	i temʊ	"the knee"
lɑhɪ	"male"	i læhɪ	"the male"
gwihən	"fish"	i gwihən	"the fish"
pecʊ	"chest"	i pecʊ	"the chest"
tunuʔ	"to know"	en tinuʔ	"you know"
huluʔ	"up"	sæn hiluʔ	"upward"
otdut	"ant"	mi etdut	"lots of ants"
lɑgu	"north"	næn lægu	"toward north"

Chamorro Vowels

	front	back
high	i	u
mid	e	o
low	æ	ɑ

- Identify the changes taking place:
- Note what doesn't change:
- Find the natural class:
- Describe the change in words:
- Write a rule with formal notation:

Sample Problem: Chamorro Vowel Harmony

Chamorro Data

gumə	"house"	i gimə	"the house"
tomʊ	"knee"	i temʊ	"the knee"
lɑhɪ	"male"	i læhɪ	"the male"
gwihən	"fish"	i gwihən	"the fish"
pecʊ	"chest"	i pecʊ	"the chest"
tunuʔ	"to know"	en tinuʔ	"you know"
huluʔ	"up"	sæn hiluʔ	"upward"
otdut	"ant"	mi etdut	"lots of ants"
lɑgu	"north"	næn lægu	"toward north"

Chamorro Vowels

	front	back
high	i	u
mid	e	o
low	æ	ɑ

- Identify the changes taking place:
  u → i, o → e, ɑ → æ
- Note what doesn't change: e, i
- Find the natural class:
  [+syllabic, +back]
  (i.e., back vowels)
- Describe the change in words:
  back vowels become front when preceded by a front vowel in the preceding word (possibly with intervening consonants).
- Write a rule with formal notation:

$$\begin{bmatrix}+\text{syllabic}\end{bmatrix} \rightarrow \begin{bmatrix}-\text{back}\end{bmatrix} \;/\; \begin{bmatrix}+\text{syllabic}\\ -\text{back}\end{bmatrix}(\text{C})\#(\text{C})\;\_\_$$

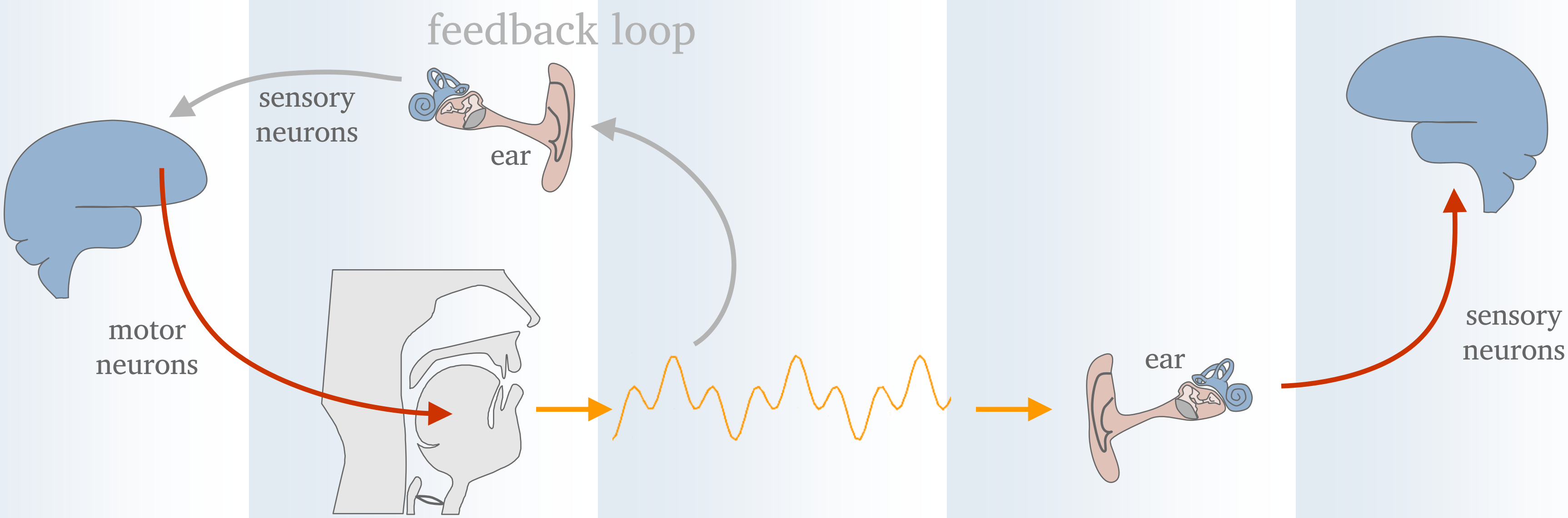
focus set — structural change — conditioning environment

Summing up

Summary of topics covered in this class

The Speech Cycle

Linguistic Level	Physiological Level	Acoustic Level	Physiological Level	Linguistic Level



Signal is initiated in speaker's brain	Signal articulated by the speech organs	Sound waves travel through the air	Listener detects the sound waves	Signal is interpreted by listener's brain

Articulatory Phonetics

- Our focus has been on learning a three-way correspondence between gestures, sounds, and symbols.
  - Gestures can affect each other though processes of coarticulation.
  - This results in “the same sound” being pronounced with different gestures on different occasions, depending on the gestures before and after it.
  - Whether and how we symbolize this variation depends on our interest (what level of detail matters to us).

Acoustic Phonetics

- Our focus has been on understanding the physical phenomena of speech sounds, and learning how to use tools to investigate the nature and properties of those sounds.
  - Speech sounds are complex pressure waves moving through a fluid.
    - Like all waves, speech sounds have *frequency*, *amplitude*, and *phase*.
    - Like all *complex periodic waves*, we can decompose speech into *component waves*, each with its own frequency, amplitude, and phase.
  - Spectrograms help us see and systematically measure variation in speech sounds (variation that we may or may not be able to hear ourselves, either because of the limits of our auditory system, or the particular types of variation we have learned to ignore based on the contrasts that matter in our native language).

Auditory Phonetics

- We have focused on a few key principles:
  - The auditory system is *nonlinear*: it handles some frequency ranges better than others, and some amplitude ranges better than others.
    - Presumably this is why most of the distinctive information in speech is carried in the range of 200 – 2000 Hz (roughly the F1 – F2 range): that is where our auditory systems are most sensitive to small changes.
  - Differences in articulation don’t matter if the resulting waveform sounds the same (recall the several different ways of making an /ɹ/-sound).
  - Our ability to recognize a sound can depend on a variety of different *cues* (recall that word-final obstruents that are fully devoiced can still be recognized as “voiced” phonemes, based on the length of the preceding vowel).

Phonology

- Phonology is about patterns in speech sounds, and about how speech sounds are represented mentally.
  - In that sense, phonology is more like a cognitive science than phonetics (which is more like a physical, physiological, or behavioral science).
  - Even though *phonemes* are represented with the same symbols we use to represent speech sounds, they are *abstract mental representations* of sound rather than measurable physical phenomena.

Where Do Features Fit?

- In this view of phonology, *features* are a way of modeling what kinds of changes cross boundaries from one phoneme to another.
  - For example, many languages don't have a voiced/voiceless distinction for stop consonants. So voicing is not a feature of stop consonants in such languages, since it doesn't cross a boundary.
- Features can also pick out larger groupings (e.g., between *obstruents* and *sonorants*).
- The value of a feature system is determined by whether it allows a full description of the patterns in a language.
  - Several different feature systems might all work; choosing among them depends on what else matters to you (e.g., do the features map to common ideas about articulatory similarity? Is one system more efficient than another?).

Final Exam Review

Exam Breakdown

- Part 1 (28 questions, 35 pts)
  - 3 anatomy/articulator terms
  - 1 VOT
  - 18 IPA symbols
  - 1 phrase written in IPA, you write in the English
  - 5 phonation questions
- Part 2 (27 questions, 40 pts)
  - 4 airstream mechanisms
  - 11 acoustics concepts/terms
  - 5 spectrogram reading
  - 3 stress/tone/intonation
  - 2 features
  - 2 natural class

Basic Terminology: Acoustics

- Frequency: the rate of repetition of a periodic signal
- Fundamental frequency (f_0): the frequency of the lowest component of a periodic signal
- Pitch: the human percept of frequency (also often used in reference to the f_0 of speech).
- Amplitude: a measure of pressure or displacement
- Intensity: a measure of how much energy is in the signal
- Loudness: the human percept of intensity
- Quality: the human percept of several intensities at several frequencies simultaneously

Basic Terminology: Phonology

- Segment/phone: a single speech sound, e.g., [s]
- Phoneme: an abstract mental representation of speech sounds. Phonemes are never directly measured or heard; they are inferred based on native speaker intuitions about what sets of sounds are "the same" and what sets of sounds are sufficiently different to change the identity of the word that the sound is a part of.
- Allophone: a segment as it relates to other segments in the language; allophones are always described in relation to the phoneme they belong to. Never say "[ɾ] is an allophone" but instead say "[ɾ] is an allophone of /t/".

Airstream, Phonation, VOT

- Airstream mechanisms
  - what are they?
  - which one(s) are used to make speech sounds in all languages?
  - which one(s) are used to make speech sounds in no languages?
- Phonation types
  - we usually describe phonation as one of four basic types. What are they? How are they transcribed using IPA?
- VOT
  - what is it?
  - for what kinds of speech sounds does VOT matter?

Stress, Tone, Intonation

- Tone
  - Grammatical vs Lexical Tone
  - Register vs Contour Tone
- ToBI
  - What is it used for?
  - What do the basic symbols represent?
- Stress, Tonic Accent, and Vowel Reduction
  - What are the main acoustic correlates of stress?
  - What is Vowel Reduction? What is Tonic Accent? What's the relationship between stress and tonic accent?

Features

- Features
  - Articulatory vs Acoustic Features
  - What is the purpose of describing speech sounds with features instead of segments?