

Teaching Fellow:

Jack Isaac Rabinovitch  
[jrabinovitch@g.harvard.edu](mailto:jrabinovitch@g.harvard.edu)

Discussion Section A:

Mondays 10:30–11:30 AM, 2 Arrow St Room 420

Section B:

Tuesdays 12:00–1:00 PM, 2 Arrow St Room 408

Office Hours:

Mondays 2:30–3:30 PM, 2 Arrow St Room 423

# Midterm Study Guide

Ling 105, Spring 2023

## Consonants and the IPA

The pulmonic IPA consonants which you are required to know for the midterm are shown in (1).

(1)

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		t̪ d̪	c ɟ	k g	q ɢ		ʔ
Nasal	m	n̪		n		n̪	n̪	n̪	N		
Trill	B			r					R		
Tap or Flap		v̪		r̪		t̪					
Fricative	ɸ β	f v̪	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ɟ	x ɣ	χ ʁ	h	h̪
Lateral fricative				ɬ ɺ							
Approximant		v̪		ɹ̪			j				
Lateral approximant				l̪		ɺ̪					

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

(2)

Symbol memorization help:

- The retroflex series looks just like the alveolar series except with a rightward facing hook on the bottom.
- c-like = voiceless palatal: c, ç
- j-like = voiced palatal: ɟ, n̪, ɟ̪, j
- Greek letters tend to be fricatives: ɸ, β, θ, ɣ, χ
- Theta [θ] is voiceless, like in the word ‘theta’
- Taps/Flaps all have an ‘r’ hook, even the labiodental one: v̪, r̪, t̪

(3)

Affricates tend to share place of articulation and voicing. When you hear ‘place + voicing + affricate’, think about the plosive and fricative that have that place and voicing:

- ‘Voiceless velar affricate’ = voiceless velar plosive + voiceless velar fricative: [kx]
- ‘Voiced palatal affricate’ = voiced palatal plosive + voiced palatal fricative: [t̪ʃ]

The non-pulmonic IPA consonants which you are required to know for the midterm are shown in (4).

(4) CONSONANTS (NON-PULMONIC)

Clicks	Voiced implosives	Ejectives
○ Bilabial	b Bilabial	,
Dental	d Dental/alveolar	p'
! (Post)alveolar		t'
‡ Palatoalveolar	g Velar	k'
Alveolar lateral	UVular	s'

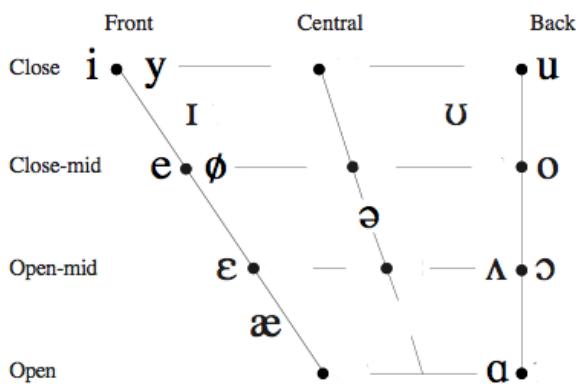
(5) Non-pulmonics:

- a. Ejectives have the ' diacritic
- b. Implosives look like voiced plosives but have a hook facing right on the top
- c. Here is how I've memorized clicks:
  - (i) ○ is round like the lips: bilabial.
  - (ii) | has one 'l' looking mark, like the 'l' at the end of 'dental'
  - (iii) ! instead of Panic! At the Disco, it's (Post)! Alveolar (because it's (post)alveolar)
  - (iv) ‡ looks somewhat as if you took a ſ and turned the top and bottom hooks into horizontal bars; [‡] and [ſ] share a place of articulation: palatoalveolar
  - (v) || has two 'l' looking marks, like the two 'l's in "alveolar" and "lateral"

## Vowels and the IPA

The IPA vowels which you are required to know for the midterm are shown in (6).

(6) VOWELS



(7) 'Neutral' vowels in English:

In unstressed syllable   In stressed syllable

Non-Rhotic	ə	ʌ
Rhotic	ə*	ɔ*

For each tense vowel there is a lax equivalent: these equivalents share the same frontness, high-

ness, and roundedness and only differ in tenseness (8). Generally, the Latin alphabet letters are tense, while their lax equivalents look similar to their tense version but are not traditional Latin letters.

- (8) Tense and Lax Vowel Equivalents:

Tense	[i]	[e]	[a]	[u]	[o]	[y]	[ø]
Lax	[ɪ]	[ɛ]	[æ]	[ʊ]	[ɔ]	[ʏ]	[œ]

- (9) Dimensions of vowels which will change the symbol:

- a. Height: High vs. Mid vs. Low
- b. Backness: Front vs. Central vs. Back
- c. Roundness: Rounded vs. Unrounded
- d. Tenseness: Tense vs. Lax

- (10) Dimensions of vowels which require diacritics:

- a. Nasality: Nasal vs. Oral
- b. Rhoticity: Rhotic vs. Non-rhotic
- c. Length: Long vs. Short
- d. Phonation: Modal vs. Breathy vs. Creaky

## Diacritics

- (11)
- a. Devoicing: placed on voiced segment to represent voiceless counterpart: [b] vs. [b̥]
  - b. Syllabicity: placed on consonants to represent that the sound is syllabic: [n] vs. [n̩]
  - c. Rhoticity: placed on a vowel to represent that the vowel is rhotic: [ə] vs. [ə̥]
  - d. Breathy: placed under segment to represent breathy voicing: [b] vs. [b̬]
  - e. Creaky: placed under segment to represent creaky voicing: [b] vs. [b̫]
  - f. Aspiration: placed on consonant to represent aspiration: [p] vs. [pʰ]
  - g. Nasalization: placed on oral segments to represent their nasal equivalent: [a] vs. [ã]
  - h. Length: placed after segment to represent that segment is long: [a] vs. [a:]

## Allophony

In phonology and phonetics, we distinguish sounds as they exist in the air (phones) and sounds as they are represented in the mind (phonemes).

- (12) Phoneme/Phone Definitions:

- a. Phoneme: Sounds as represented in the mind. Represented /as such/
- b. Phone: Sounds as they exist in the real world. Represented [as such]
- c. Allophonic Rules: Rules which take phonemes and the contexts in which they exist and return the phones.
- d. Allophone: Two distinct phones which represent the same phoneme.

We first have a mental representation of a word, and then apply allophonic rules to get phones.

- (13) Phonemes + Context → Allophonic Rules → Phones

For instance, when we think of the word “bid” our brain conjures up the representation /bɪd/, which we pronounce [bɪd], where the allophonic rules determine that the [b] can be devoiced when utterance initial. When we add a suffix -er, as in “bidder” the context of the /d/ changes: it is now between two approximants /bɪdə/. An additional allophonic rule tells us that in between approximants, /t/, /d/, and /n/ become flaps, thus the surface form of “bidder” is [bɪɾə] rather than [bɪdə], even though the [d] shows up in “bid” [bɪd].

- (14)    a. /bɪd/ → [bɪd]  
          b. /bɪd/ + /ə/ = /bɪdə/ → [bɪɾə]
- (15)    A guide to English Allophony:
  - a. Vowels nasalize before nasal codas: ‘win’ /'wɪn/ → ['wɪ̃n]
  - b. /t/ and /d/ become flaps [ɾ] between approximants (including vowels) when the following syllable is unstressed.
    - (i) ‘flatter’ /'flætə/ → ['flæɾə]
    - (ii) but not ‘ideal’ [aɪ'dɪɫ]
    - (iii) /n/ can also flap in the same environment to become a nasalized flap [ɾ̃]:  
      /'wɪnə/ → ['wɪ̃ɾə]
  - c. /t/ can become a glottal stop [?], especially before nasals: ‘button’ /'bʌtn/ → ['bʌʔn]
  - d. a voiceless stop is aspirated when initial in a word or stressed syllable:
    - (i) ‘pick’ /'pɪk/ → ['pʰɪk] and ‘camper’ /'kæmpə/ → ['kʰæmpə]
    - (ii) but not ‘spin’ [spɪn]
  - e. a voiced plosive is optionally devoiced utterance initially: ‘boat’ /'bɒt/ → ['bɒt]
  - f. approximant consonants {ɹ j w l r} devoice after a voiceless consonant. In these instances, the voiceless consonant does not aspirate: ‘play’ /'pleɪ/ → ['pleɪ]
  - g. /t/ is usually affricated (becomes affricate [tʃ] before /i/): ‘tree’ /'tri/ → ['tʃi]
  - h. /l/ becomes velarized ‘dark’ [ɫ] in coda position: ‘bell’ /'bel/ → ['bɛɫ]
  - i. Word finally, voiceless stops are optionally unreleased [t̚] and/or glottalized [ʔt̚]:  
      /kæt/ → ['kʰæt̚'] or ['kʰæʔt̚']

## Diagram Reading

- (16)    Waveforms:
  - a. The Horizontal Axis represents time.
  - b. The Vertical Axis represents change in pressure (sound is just the change in pressure through a medium).
  - c. The path on the diagram represents the change in air pressure at a given time.
- (17)    Spectrograms:
  - a. The Horizontal Axis represents time.
  - b. The Vertical Axis represents frequency.
  - c. Shading represents amplitude (intensity of sound): dark = high amplitude = louder.

(18) Vowel Diagrams:

- a. The Horizontal Axis represents frontness / F2 (front = high F2)
- b. The Vertical Axis represents openness / F1 (open = high F1)
- c. Dots on the diagram represent vowels
- d. Paths on the diagram represent diphthongs, where the ‘location’ of the vowel changes over time.

## Vowel Space and Formants (F1 and F2)

The vowel space can be seen in two ways: as a space where the dimensions correspond to the placement of the tongue (high, low, front, back), and the other as a space plotting F1 and F2.

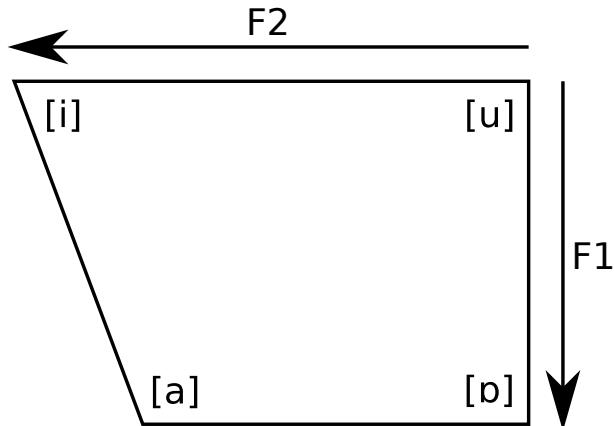
(19) F1 is the lowest formant: it represents the lowest of the loudest (non-fundamental) pitches.

- a. F1: Typically between 300–700 Hz
- b. Variation of F1 is formed by the size of the pharynx
- c. The more open the jaw and lower the tongue placement, the more the body of the tongue is pushed into the pharyngeal area, taking up space
- d. Low/Open vowels = less pharynx space = higher F1
- e. High/Closed vowels = more pharynx space = lower F1

(20) F2 is the next lowest formant.

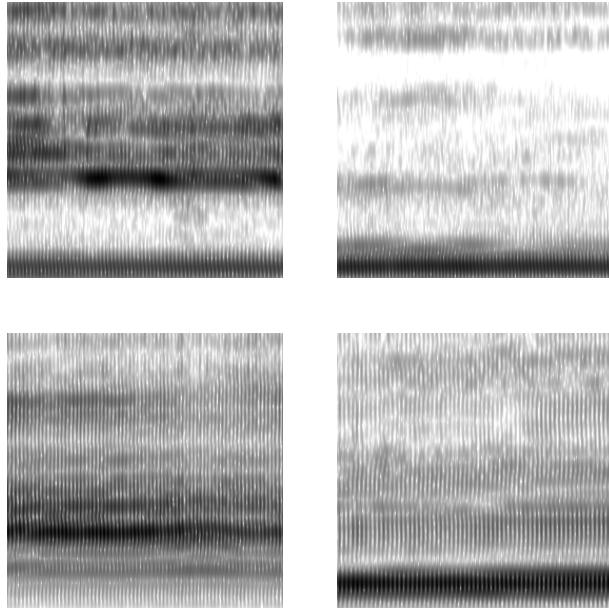
- a. Typically between 700–2000 Hz
- b. Variation of F2 is formed by the size of the oral cavity
- c. The further in front the tongue is placed, the less room in the oral cavity
- d. Front vowels = less oral cavity space = higher F2
- e. Back vowels = more oral cavity space = lower F2
- f. Lip rounding makes the mouth larger, and contributes to F2: round = lower F2

(21)



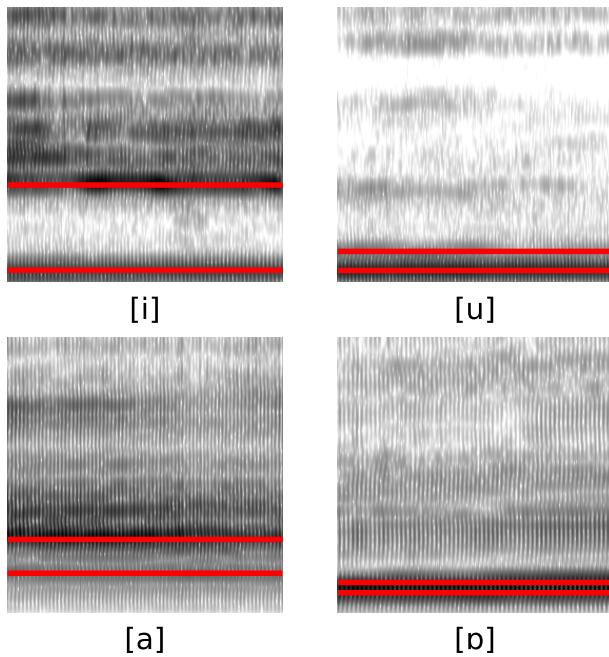
In spectrograms, the formants appear as dark lines at varying heights. F1 is always lower than F2, by definition.

(22)



It can be difficult to see in isolation what counts as the two lowest bands. The red lines in (23) will help.

(23)

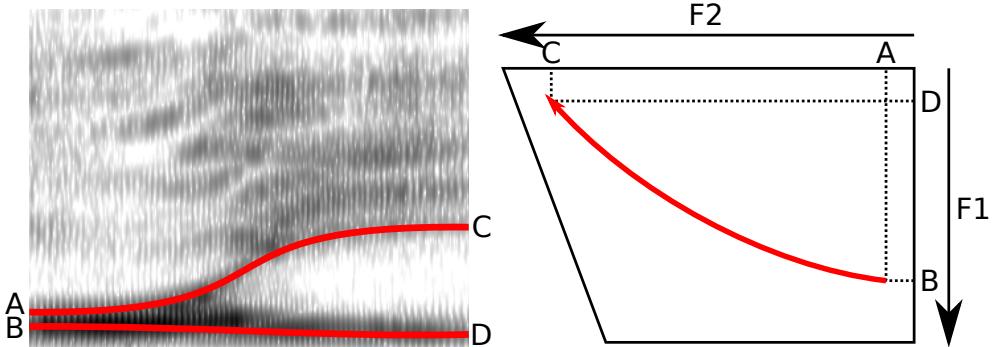


As you can see, [i] has the largest gap between the two: F1 is low while F2 is high. For [u], the F2 is lowered, and the F1 remains low (it has not moved). [a] has both F1 and F2 at high positions, while [ɔ] has a low F2, but a high F1, resulting in two bands which almost blend completely together.

Diphthongs show change in the vowel space. In (24) you can see how I plotted a recording of myself waying [ɔɪ], F2 rises from a relatively low position to a high position, while F1 lowers (the lowering is much more subtle because the range of F1 is smaller!). If I plot over time the position of F1 and F2 as coordinates in the vowel space, the result is a curved path, denoting the

change in vowel quality over time.

(24)



## Aspiration, Voicing, and VOT

(25) Voice Onset Time (VOT):

The amount of time that passes between the release of a plosive and the onset of voicing. Assuming that the release is at “ $t = 0$ ”, the VOT is the relative time of the voicing.

a. Release of a Stop Consonant

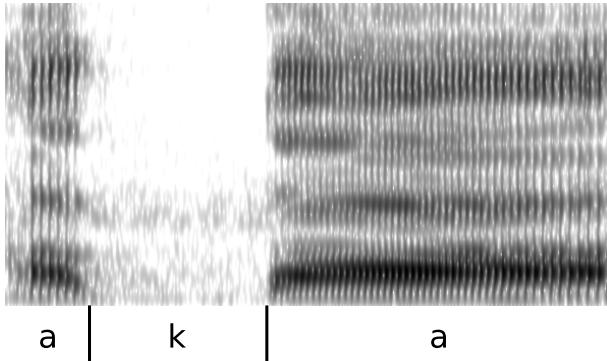
The time when the closure of a stop consonant is released and air can begin to flow from the mouth.

b. Onset of Voicing

The time when the larynx begins to vibrate, causing voicing.

In (26), I pronounce [aka]. The full closure of the mouth caused by the plosive [k] means that air cannot flow out of the mouth. As a result, I produce very little sound, as shown by the lightness in the spectrogram.

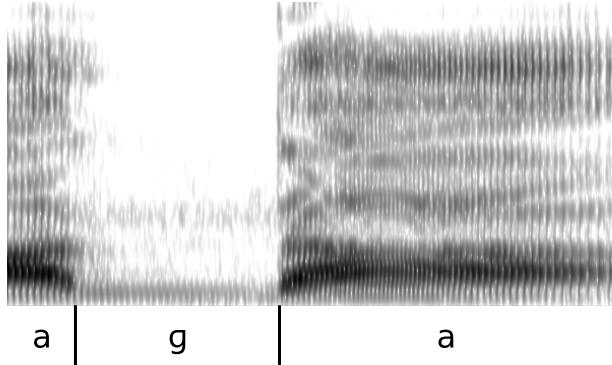
(26)



The [k] is released and immediately is voiced, becoming the vowel [a]. Because release (the mouth opening) and voicing (the vibration of the larynx) are simultaneous, the Voice Onset Time (VOT) is zero.

In (27), I pronounce [aga]. Just like the plosive [k], [g] creates full closure and does not allow airflow out of the mouth. However, as [g] is voiced, air still flows from the lungs into the mouth, and the larynx vibrates, creating a F0, as denoted by the dark bar on the bottom of the spectrogram.

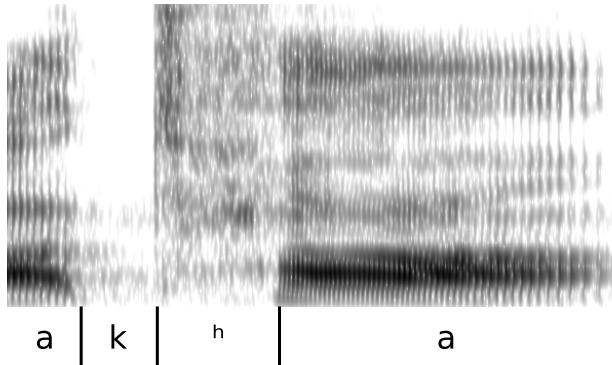
(27)



The [g] is consistently voiced, and release happens well after voicing starts. Assuming that the release is at “ $t = 0$ ”, because the time of voicing is earlier, the VOT is negative.

In (28), I pronounce [ak<sup>h</sup>a], the aspirated [k<sup>h</sup>] is a voiceless plosive like its unaspirated counterpart [k], and thus shows no dark band of voicing. For [k<sup>h</sup>], there is a period of time where the stop has released but there is no voicing. This causes the darker patterning above F0 to be present but with no dark band on the bottom. This is aspiration of [k<sup>h</sup>], where air flows out of the mouth but is not voiced.

(28)



If we set the time of the release of [k<sup>h</sup>] as “ $t = 0$ ”, when we find that voicing happens afterwards; corresponding to a positive VOT.

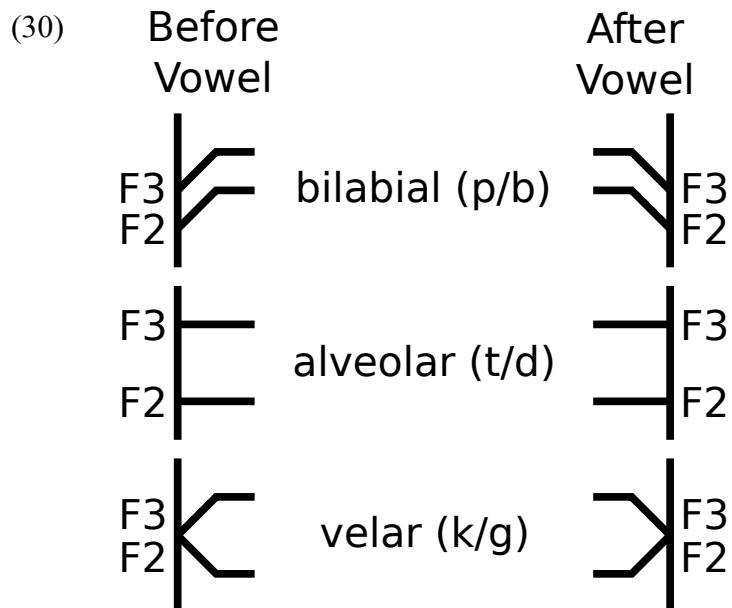
## Dispersion

(29) Perception Dispersion:

- a. In order to make two sounds more distinct, they may differ in more than one ‘dimension’ of perceptual space.
- b. Front/Back distinctions are often parallel to Rounded/Unrounded distinctions.
- c. [s] is unrounded in English, while [ʃ] has slight rounding: this creates an extra dimension in which [s] and [ʃ] can be distinguished.

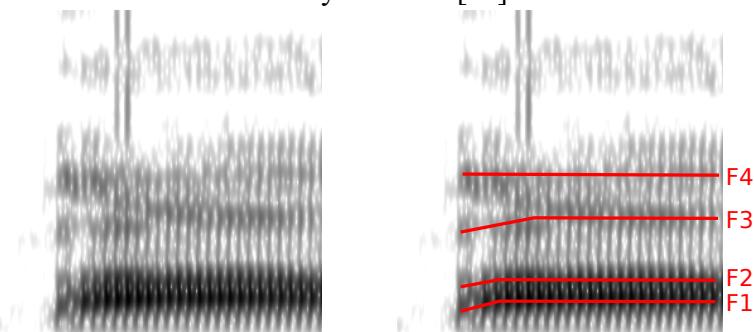
## Consonants in Spectrograms

The effects of plosives on formants are seen in (30). Note that these are F2 and F3! Bilabials have lower relative F2 and F3s to the vowels that follow or precede them, while alveolars show little difference. Velars have a “pinch” where the F2 is relatively high and the F3 is relatively low.

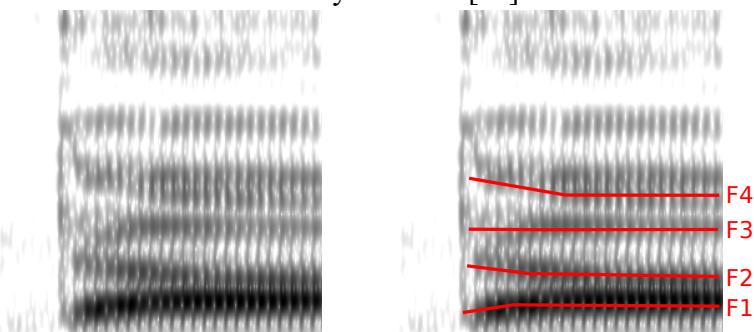


In actuality these are hard to pick out individually in formants, but differences can be seen more clearly through comparison.

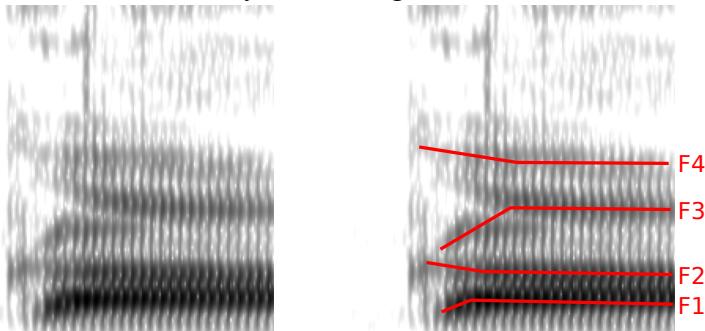
(31) Bilabial onset followed by a vowel [ba]



(32) Alveolar onset followed by a vowel [da]



- (33) Velar onset followed by a vowel [ga]



## Phonation and Airstream Mechanisms

- (34) Phonation Types:

- a. Modal Voice (Voiced): Regular vibrations of the vocal cords
- b. Voiceless: Lack of vibration of the vocal cords; arytenoid cartilages usually apart
- c. Breathy: Vocal cords vibrating but without appreciable contact; arytenoid cartilages further apart than in modal voice
- d. Creaky: Vocal cords vibrating anteriorly, but with the arytenoid cartilages pressed together; lower airflow than in modal voice

- (35) Flow Direction

- a. Egressive: compression causes high pressure; exhalation (if pulmonic)
- b. Ingressive: rarefaction causes low pressure; inhalation (if pulmonic)

- (36)

	Pulmonic	Glottalic	Velaric
Pressure Source	Lungs	Larynx	Velum and Tongue
Egressive Version	'Typical' sounds	Ejectives	(Velaric egressive)
Ingressive Version	(Pulmonic ingressive)	Implosives	Clicks

## Midterm Reminders

- (37) Midterms can be stressful! Here are my tips.

- a. Studying in groups can make studying effective; I find it helpful to partner up and 'quiz' one another (e.g.: tell your partner the name of a sound, like 'voiceless uvular fricative', and your partner write it down, check, repeat).
- b. If you are reviewing material and you notice that you are spacing out/not absorbing it/not processing it, take a break: I find going for short walks (just around my apartment) or regularly getting a glass of water helps to break up the study and give my brain some space to relax.
- c. Don't be afraid to make some noise! Making a sound and feeling the position of your tongue and other articulators is the mouth equivalent of a 'hands-on' experience.
- d. Get a good night's sleep two nights before the test: it'll be much easier to sleep the following day, or help you stay awake when studying the night before the test.

## Practice

(Q1) Transcribe the following words using narrow transcription.

- a. hollow
- b. rider
- c. monstrous
- d. plight
- e. fill
- f. miner
- g. stomp
- h. tomb
- i. carbonation
- j. trickle
- k. cartography

(Q2) For each of the following descriptions, give a corresponding IPA symbol (or symbols, with diacritics if necessary)

- a. voiceless bilabial fricative
- b. voiceless lateral click with uvular rear closure
- c. voiceless palatal affricate
- d. voiced alveolar lateral fricative
- e. voiceless velar plosive
- f. voiceless retroflex nasal
- g. high front tense unrounded vowel
- h. nasal high back lax rounded vowel
- i. schwa
- j. long front mid lax unrounded vowel
- k. breathy high front tense rounded vowel

(Q3) Labiovelars such as [w] use what two broad places of articulation?

- a. Labial and Velar
- b. Labial and Dorsal
- c. Bilabial and Velar
- d. Coronal and Dorsal

(Q4) At which following place is the front closure of a click impossible?

- a. Bilabial
- b. Dental
- c. Velar
- d. Palato-alveolar

(Q5) Which of the following ejectives is the least common?

- a. p'
- b. t'
- c. k'
- d. s'

- (Q6) If a sound has a frequency of 125 Hz, what is its corresponding wavelength (in ms)?
- a. 50 ms
  - b. 8 ms
  - c. 0.008 ms
  - d. 0.05 ms
- (Q7) The following spectrogram represents me saying “shopify”. Please divide the spectrogram into sections representing each of the phones in the word “shopify”, and label each section with its respective phone. Consider diphthongs as two different segments.



- (Q8) Fijian is a language which has both voiced and voiceless plosives. In Fijian, voiced plosives tend to be pre-nasalized, meaning that they are nasalized at the beginning of the stop sound (represented like [<sup>n</sup>d]).
- a. Why might Fijian make the additional distinction of nasalizing voiced plosives? What principle is at play here?
  - b. Fijian is one of many languages that employ pre-nasalization in voiced plosives. There are no languages, however, for which voiceless plosives are pre-nasalized and voiced plosives are not. Why might that be?