Teaching Fellow: Jack Isaac Rabinovitch
jrabinovitch@g.harvard.edu

Discussion Section: Boylston Hall, Room 335
Section 1: Thursdays 3:00–4:00
Section 2: Thursdays 4:30–5:30

Office Hours: Boylston Hall, Room G03
Mondays 12:15–1:15

Discussion Handout 4

September 30, 2021

(1) Today:
   a. Formants and Vowels
   b. Reading Spectrograms: Voicing and VOT
   c. Listening (phone recognition) Practice
   d. Allophony Practice

Formants and Vowels

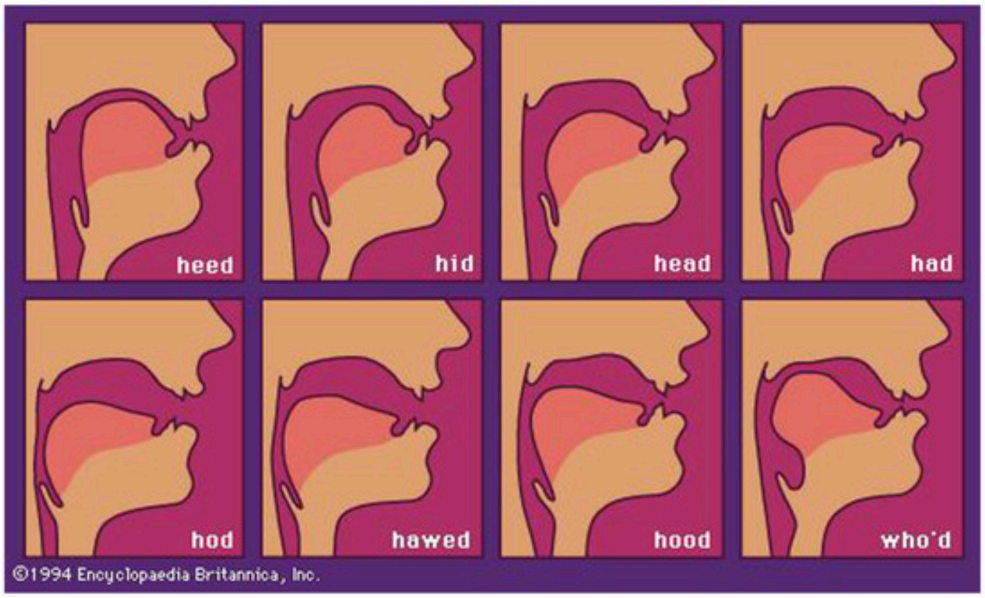
(2) A formant is a broad frequency which has a particularly high amplitude for a given sound: 'local maximum".

(3) F0 is not a formant but the fundamental frequency: it represents the pitch of speech.
   a. Typically between 80–300 Hz
   b. Variation of F0 is formed by the length and tension within the vocal folds of the larynx
   c. Corresponds with tone: high F0 = high tone, low F0 = low tone

(4) 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

(5) 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 places, 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

(6) Note how high back vowels [u] and [ʊ] still leave room in the pharynx!



(7) From Rose et al. (2006) [baɪd]

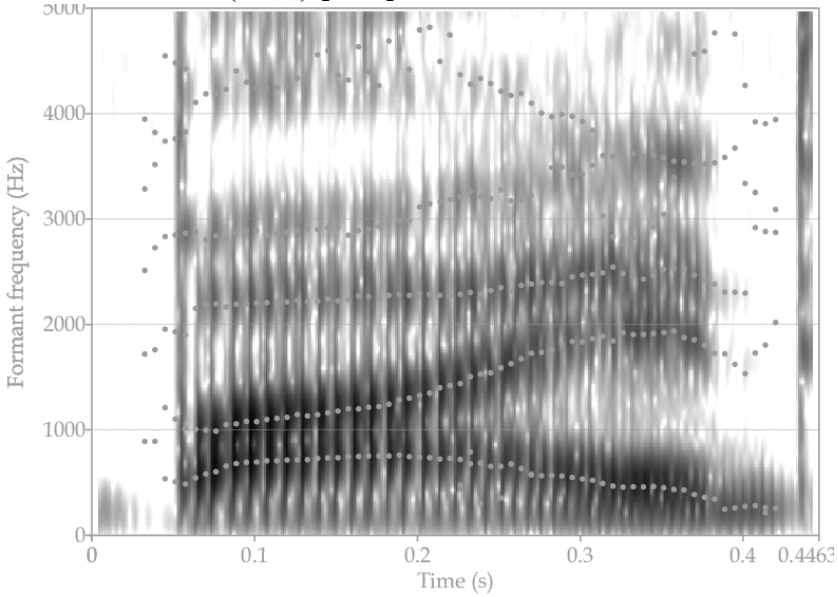


Figure 2: Spectrogram of *bide* (Jonathon 1.2) with superimposed formants showing lack of clear stable F-pattern for first target.

Voicing and VOT

(8) F0 (Fundamental Frequency) is a byproduct of voicing.

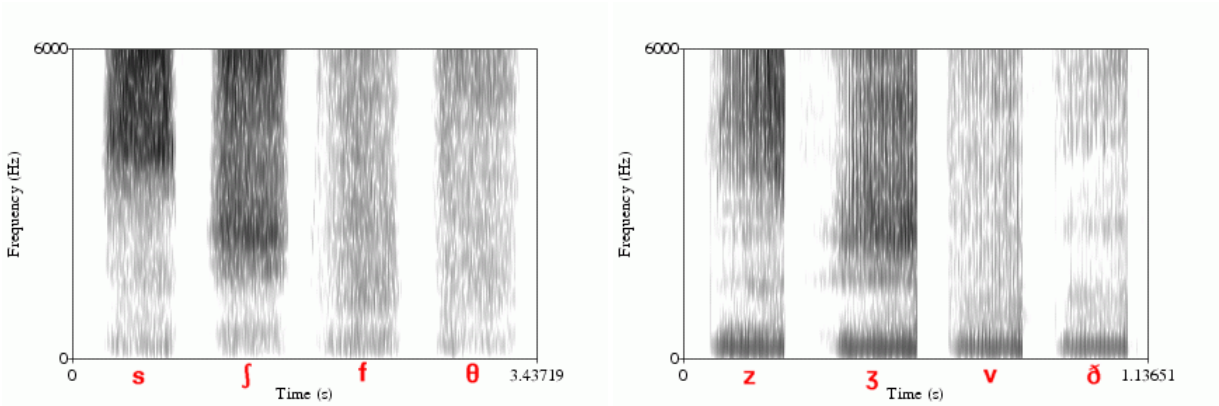
a. If a sound is not voiced, then it will not have an F0
b. F0 is typically too low and specific to a certain tone to see
c. Voicing not only adds F0, but increases amplitude of formants: if something in the low region is particularly dark, it is likely because of voicing.

(9) Frication affects high frequencies

a. Sibilants affect particularly high frequencies: {s, z, ʃ, ʒ} are all very dark on a spectrogram in the high frequency region.

b. These sounds affect a large range, are 'noiselike'

(10) Various Fricatives, voiceless and voiced:

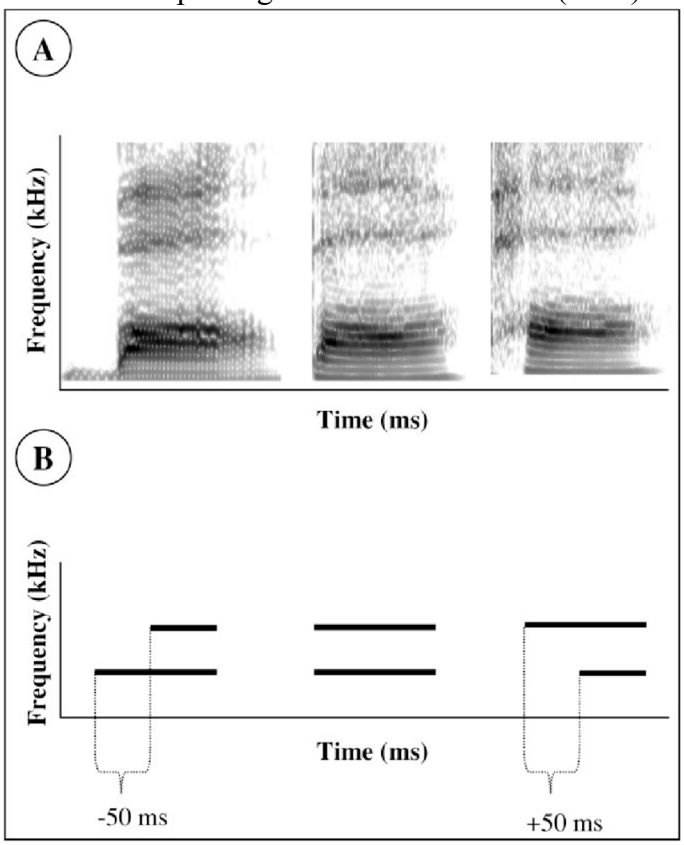


(11) VOT = Voicing Time − Release Time

a. Positive VOT: there is time with release, without voicing: aspirated

b. Negative VOT: there is time with voicing, without release: voiced

(12) VOT in the Spectrogram from Diehl et al. (2004):



More Practice

(13) For the following:

a. Provide a broad transcription.
b. List which allophonic rules (at the bottom of the page) apply to each word.
c. Apply the allophonic rules to the broad transcription to provide a narrow transcription.

(14)
a. dimming
b. flutter
c. redeem
d. skinny
e. mutton
f. temple
g. princess
h. spill
i. gannet
j. twin
k. trespass

(15) English Allophony (all of these should be denoted in narrow transcriptions):

a. Vowels nasalize before nasal codas.
b. Coronal stops {t, d, n} become flaps {ɾ, ɾ̃} between approximants (including vowels) when the following syllable is unstressed.
c. /t/ can become a glottal stop [ʔ], especially before nasals.
d. Voiceless stops are aspirated when they are the first segment in a word or stressed syllable and not followed by a approximant consonant.
e. Voiced plosives are optionally devoiced utterance initially.
f. Approximant consonants {ɹ j w l ɾ} devoice after a voiceless consonant. In these instances, the voiceless consonant does not aspirate.
g. /t/ is usually affricated (becomes affricate [tʃ]) before /ɹ/.
h. /l/ becomes velarized 'dark' [ɬ] in coda position.
i. Word finally, voiceless stops are optionally unreleased [t˺] and/or glottalized [ʔ͡t].

References

Diehl, Randy L., Andrew J. Lotto, and Lori L. Holt. 2004. Speech perception. Annu. Rev. Psychol. 55:149–179.

Rose, Phil, Yuko Kinoshita, and Tony Alderman. 2006. Realistic Extrinsic Forensic Speaker Discrimination with the Diphthong/a□. In Proc. Llth Austr. Int. Conf. on Speech Sci. and Tech, 329–334. Citeseer.

Practice Answers

(16)
a. dimming /ˈdɪmɪŋ/ > (Rules: a, e) > [ˈd̥ɪmĩŋ]
b. flutter /ˈflʌtɚ/ > (Rules: b, f) > [ˈfl̥ʌɾɚ˞]
c. redeem /ɹɪˈdim/ > (Rules: a) > [ɹɪˈdĩm] (Rule b does not apply because of stress)
d. skinny /ˈskɪni/ > (Rules: b) > [ˈskɪr̃i] (Rule d does not apply because of /s/)
e. mutton /ˈmʌtn̩/ > (Rules: c) > [ˈmʌʔn̩]
f. temple /ˈtɛmpl̩/ > (Rules: a, d, h) > [ˈtʰɛ̃mpɬ̩] (technically /l/ is in nucleus position, but it is still the last sound of its syllable.
g. princess /ˈpɹɪnsɛs/ > (Rules: a, f) > [ˈpɻ̊ĩnsɛs]
h. spill /ˈspɪl/ > (Rules: h) > [ˈspɪɬ]
i. gannet /ˈgænɛt/ > (Rules: b, e, i) > [ˈg̊ær̃ɛt˞]
j. twin /twɪn/ > (Rules: a, f) > [ˈtw̥ĩn]
k. trespass /ˈtɹɛspæs/ > (Rules: f, g) > [ˈtʃɻ̊ɛspæs]