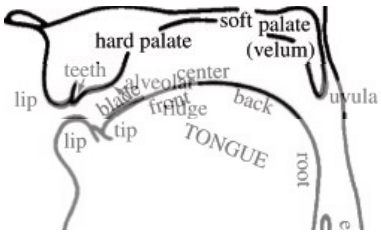
[2019-2 영어음성학]

이름(학번): 민지우(2016130575)

Week1 Summary

Spelling in English is inconsistent with its pronunciation, which call for the study of **phonetics**, all about the speech sound. We look at the actual sound of a word, and use phonetic symbols to describe it. There are about 44 distinguishable sounds that form words: **phonemes**, much diversified than 26 letters in the English language. Each phoneme represents vowels(monophthongs, diphthongs) or consonants. The study area of phonetics can be divided into three categories; articulatory, acoustics, and auditory, each accounting for production, transmission, and hearing of speech sound.

✱ Articulation



Making a speech sound is related to physical movements of human organs. Depending on subtle changes in the vocal tract (pharynx, epiglottis, uvula, soft palate(velum), hard palate, alveolar ridge, teeth, tongue body, tongue tip, and lip) while air flows, different sounds are articulated. We can divide five speech organs which control constrictions: lip, tongue tip, tongue body, larynx(vocal cords), and velum.

(1) Articulatory process in lips/ tongue tip/ tongue body

- Constriction Location(CL)
  - lips(bilabial/labiodental): p, b, m, w, f, v
  - tongue tip(dental/alveolar/palate-alveolar/retroflex): θ, ð, t, d, s, z, n, l, ʃ, ʒ, ʧ, ʤ, r
  - tongue body(palatal/velar): j, k, g, ŋ, h (+ change in oral cavity makes different vowels)
- Constriction Degree(CD)
  - upper part(nearly closed)> stop(p, b, t, d, k, g) - fricative(f, v, θ, ð, s, z, ʃ, ʒ, h) - approximant(w, l, r, j) - vowels <lower part(opened)

(2) Phonation process in larynx(voice box)

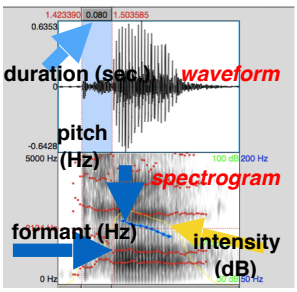
- Vocal cords(=vocal folds) vibrate in a varied degree when air flows from the lung.
- Voiced(vocal cords closed): b, d, g, v, ð, z, ʒ, ʤ m, n, ŋ, l, r, w, j
- Voiceless(vocal cords opened): p, t, k, f, θ, s, ʃ, h, ʧ

(3) Oro-nasal process in velum

- When the velum is lowered, the nasal tract is opened, so we can breathe through nose and produce nasal sounds m, n, ŋ.
  - Velum: m, n, ŋ
- *When the velum is raised, the nasal tract is closed, and we can produce every vowel and every consonant except for nasal sounds m, n, ŋ.*

We can predict any phonemes with description of constricting environment, for example, “t,” is a sound produced when velum is raised(not nasal sounds), larynx is opened(voiceless), constriction location is at tongue tip(alveolar) and constriction degree is upper part(stop).

✱ Acoustics



With Praat, we can examine the duration, intensity(loudness), pitch(65-200Hz for male speech, and 145-275Hz for female speech), spectrogram, formant(ed dots signify different vowels) of sounds. **Pitch** is defined as the number of occurrences of repeating event per second(Hz). We can see the repeated event as the vocal cords vibrate repeatedly. For example, 236.5Hz means there was 236 times of vibration in vocal cords per second, single event took 0.004228 seconds to be concluded. A pure tone of certain frequency shows sine wave and every natural sound can be represented with simplified sine wave of same frequency.

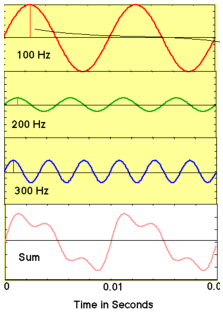
[2019-2 영어음성학]

이름(학번): 민지우(2016130575)

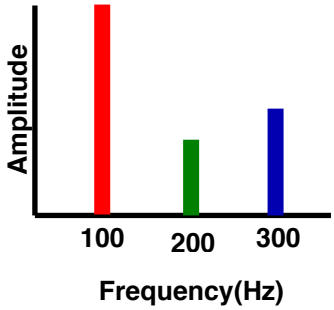
Articulation	CL	CD	Velum	Larynx
/p/	bilabial	stop	raised	open
/d/	alveolar	stop	raised	closed
/z/	alveolar	fricative	raised	closed
/n/	alveolar	stop	lowered	closed

Week2 Summary

Every signal(including sounds) can be expressed with a synthesis of various sine waves. We call sounds of a sine wave as a simplex(pure) tone and a sum of sine waves as a complex tone. A sine wave is shaped by frequency and amplitude of a sound over time. A lower frequency means fewer vibrations in vocal cords, which produces sparse graph and vice versa. 100Hz signifies 100 times of vibrations per second and, if doubled, it becomes 200Hz. Amplitude; air pressure, is not an intrinsic value of a sound source but varies over a wide range depending on the distance, closer the bigger, thus, producing bigger graph and vice versa.



In the graph left-side where sound waveforms are depicted, x-axis means time in seconds and y-axis means value. In spectrum right-side(spectral slice of spectrogram), x-axis corresponds to frequency(Hz) and y-axis corresponds to amplitude over time. Spectrogram is a temporal concatenation of spectrum, its x-axis being time and y-axis being frequency. Light-Dark gradation represents the amplitude of each frequency, lower in frequency being darker in the case of a pure tone spectrogram.



Human voice source is a sound at larynx, which can be measured by ElectroGlottoGraph(EGG). When we focus on the sound pitch(magnitude and frequency, source can be sorted as a complex tone, consisting of harmonics. A complex tone is a sum of pure tones at integer multiples of the lowest pure tone(fundamental frequency = F0 = pitch). Source-filter theory represents speech production process where sound sources of distinct spectrum are filtered by the resonant properties of the vocal tract. While source shows gradually decreasing graph in the frequency spectrum, filtered by vocal tract(audio) shows zigzagging(carved) graph with peaks and valleys. Amplitude Peaks in the frequency spectrum of the sound are analyzed as formants, the spectral shaping by acoustic resonances of the vocal tract. The formant with the lowest frequency is called F1, the second F2. In a graph where x-axis is F2 and y-axis is F1, corresponding to the geometry of mouth, we can locate every monophthong as spots and diphthong as lines, as x-axis determines front/back position and y-axis determines height of vowels.