

< 1 주차 9/5 >

- Praat: 음성 분석 프로그램
 - 소리 = 숫자. x 축은 시간, y 축은 분석값. 그래프는 그 숫자 값들이 연결된 것
 - 그 값들은 끊임없이, 연속적으로 이루어져 있음. 쪼개는 기준에 따라 나누어지는 거지 실제 소리는 continuous 하다. 즉, 목적에 따라 나누어 값을 뽑기 나름이다. (ex. 1 초를 얼마간격으로 나눌 것인가? Sampling frequency 44100Hz = 점 사이당 44100/1 초)
 - cut, pitch(높낮이), intensity(세기. 폭이 넓으면 intensity 가 강하고 좁으면 약하다), 분석본에서 위는 wave, 아래는 소리의 스펙트럼(빛을 프리즘으로 본 스펙트럼처럼)
 - (+ 프리즘에서 빨간색은 더 온도가 낮고 보라색은 더 높음. 빨간색이 더 slow 하고 저주파, 보라색이 더 빠르고 고주파. 즉, 위로 갈수록 고주파, 아래쪽으로 갈수록 저주파)

< 3 주차 9/17, 9/19 >

- Phonology(음운론)_cognitive, abstract
 - English consonants
 - English vowels: 단모음(monophthongs), 장모음(diphthongs)
- Phonetics(음성학)_physical
 - Study on sound system
 - Articulation: 소리를 만들어내는 사람의 원리
 - Acoustic: 공기에서 소리가 어떻게 나는가. 물리적인 것.
 - Auditory: 어떻게 듣는가. (Ex. 귓바퀴는 소리를 증폭시켜 듣게 해줌)
 - The vocal tract: 크게 nose, ear, pharynx, larynx 로 구성
 - Upper part: Palate, soft palate(velum), uvula(목젖), larynx(후두), alveolar.
 - Lower part: lip, epiglottis(후두개), tongue → tip, blade, body.
- Articulation
 - Phonation process(with larynx)
 - Larynx(voice box) open → voiceless (no vibration on vocal cord)
 - closed → voiced
 - Oro-nasal process(in velum)
 - Velum lowered → nasal (nasal tract open), 코로 숨쉴 때
 - Raised → 비음 뺀 모든 자음 + 모음
 - Articulatory process(in lips, tongue tip, tongue body → 3 constrictors)
 - CL(좌우) → Lip(bilabial, labiodental), tongue body(Palatal, velar), tongue tip(dental, alveolar, retroflex, palate-alveolar)
 - CD(상하) → Stops > Fricatives > Approximants(r,l,w,j) > Vowels
 - 즉, By specifying velum, larynx, constrictors, CD, CL → 소리가 결정된다.
 - + 모든 모음은 constrictor 중 tongue body 만 사용한다.

< 4 주차 9/24, 9/26 >

- Phoneme

- individual sounds that form words. ex. 'psycho'=orthography, /s aɪ k oʊ/=phonemes

- Acoustics in Praat(DSP: Digital Signal Processing)

- Pitch = 높낮이, 내가 1 초에 몇 번 larynx 를 움직였는가에 대한 계산. 즉 vocal folds 가 vibration 하는 것이 repeating event 이며, 이때 sine wave 가 repeating 하는 것이 성대의 떨림과 일치

- Intensity = 크기 → Pitch 와 intensity 는 서로 독립적.

- 스펙트로그램에서 나타나는 4 개의 빨간띠를 formant 라 하며, 제일 아래부터 F1~F4 라고 부른다. F1, F2 가 무엇이냐에 따라 모음이 뭔지 결정하며, 즉 모음을 구분하는 수치로서 formant 가 쓰인다.

- Complex tone in spectrum

- Sine wave = 가장 기본적인 시그널의 형태. X 축=시간, Y 축=value. Frequency 와 magnitude 에 따라서 웨이브의 형태 결정.

- Time_value 그래프인 Sine wave 는 frequency_amplitude 그래프로 더 단순하게 표현할 수 있으며, 이때 frequency_amplitude 의 막대 그래프를 Spectrum 이라고 한다.

- 스펙트로그램은 스펙트럼을 시간 축을 따라 길게 visualize 한 것이며, 스펙트로그램의 한 슬라이스가 스펙트럼이다.(Spectrum=Spectral slice) 즉, 스펙트럼에는 시간 개념이 없다. 스펙트로그램에서 X 축=시간, Y 축=frequency.

- 이 세상에 존재하는 모든 시그널은 서로 다른 사인 웨이브의 결합으로 표현된다.(Simplex tone→Complex tone) 즉, 모든 신호는 조금씩 다른 사인웨이브의 합으로 표현될 수 있다. 복잡한 형태의 어떤 것을 단순한 형태의 것들로 표현할 수 있다!

- Human voice source

- 소리가 어떤 소리로 만들어 지는가는 성대가 아니라 입에서 어떻게 하느냐에 따라 결정됨

- 즉, 성대에서 나는 소리를 source 라고 하고, 튜브(입)에서 어떻게 소리가 달라지는가를 filter 라고 한다. Filter 가 어떻게 달라지느냐에 따라 /아/ /이/가 달라지는 것.

- Source 만 가지고 spectral analysis 를 하면 gradually decreasing 하는 형태를 보인다. 맨 처음 값이 F0(fundamental frequency)로서 pitch 값과 똑같고, 그 뒤의 값들은 F0 의 X2 값, X3 값으로 쭉 이루어지는데, 이를 harmonics 라고 한다.

- Human voice source 는 harmonics 로 이루어져 있고, harmonic 는 sine wave 의 배음으로 이루어져 있다. 이때 나타나는 amplitude of pure tone 의 패턴은 gradually decreasing 한다.

- Filtered by vocal tract(filtered=carved)

- Filtered 된 소리를 spectral analysis 하면 Jignagging with peaks and valleys 하는 형태를 보인다. X2, X3 의 배음의 구조는 깨지지 않았으나, amplitude 의 패턴이 깨져 더이상 gradually decreasing 하지 않는 것이다.

- EGG(성대녹음) 소리는 gradually decreasing 한데, Audio(실제소리)는 peaks and valleys 의 형태

- Vocal tracts 에 의해 Filtered 되었기 때문

- 첫번째 산맥인 F1 은 formants=VT 가 좋아하는 주파수, Valley=VT 가 싫어하는 주파수라고 보면 됨.

- 각 소리에 따라 어디에 산맥이 나타나느냐가 다르다.

- Source(from larynx)-filter(by vocal tract) theory

- Voice source(source spectrum) –Vocal tract filter function→ Output from lips(output spectrum)

- Spectrogram = Airplane view of temporal concatenation of spectrum

- 진하게 나타나는 부분 Dark band = mountains = formants

- F1, F2 만 보면 충분하며, 그 둘이 달라지는 형태가 모음 도장이라고 생각하면 됨

- F1= 높낮이 결정, F2=Front & back 여부 결정