

Human voice source

- 성대에서 나는 소리를 직접 캡처한 것. larynx에서 나는 소리=EKG
- spectral analysis 후 첫번째 주파수=발화자의 피치= F_0 =1초 동안의 성대 떨림수=lowest sine wave
- Harmonics**: F_0 에 정수배하는 여러 sine wave들의 합=대음
- Complex tone: 가장 낮은 pure tone을 정수배한 tone들의 합
- *사람 목소리는 다 sine wave들의 합으로 이루어져 있지만 모든 소리가 그런 것은 아님
- *여자 목소리가 남자 목소리보다 spectral analysis를 했을 때 F_0 이 더 크고 간격이 더 넓다

Filtered by vocal tract

- 성대에서 나온 소리가 입모양 거쳐서 최종적으로 발음이 되는 소리
- spectral analysis 했을 때 대음 구조가 바뀌는 것 X, amplitude의 구조가 바뀌는 것
- 스펙트럼의 formants 사이의 계곡을 만드는 것: 사람의 입모양

Source&Filter

- 모든 사람이 /a/할 때 스펙트럼의 패턴이 다 똑같은 것

Formant: spectral analysis했을 때 위로 솟은 부분들=peaks/mountains

- e.g. F_1 , F_2 ...
- 모음마다 F_1 F_2 가 달라서 이 두개만 안다면 스펙트럼을 보고 뭐가 무슨 모음인지 알 수 있음
- 세로축 F_1 가로축 F_2 로 이루어진 표에 각 모음들을 배치하면 사람 입안모양, 모음사각도랑 같음
- F_1 : 혀의 높낮이 결정
- F_2 : 혀의 앞뒤 결정

Spectrogram: spectrum을 시간 축으로 늘어놓은 것

모음 만드는 방법: pitch에 해당하는 주파수를 안다면 그 pitch에 해당하는 sine wave를 만들고 정수배 하기 ex) 100HZ-200HZ-300HZ...

기타 소리가 사람 목소리 source하고 가장 비슷함

Synthesizing source

- harmonics 만들기
- 100HZ 부터 1000HZ까지 amplitude가 일정하게 감소하는 pure tone들을 합성하기
- 무한대로 합친다면 앞부분에 산맥 하나=pulse train
- spectrum of pulse train gradually decreases