

# 영어음성학필기

2015410098 컴퓨터학과문승연

음성 분석 프로그램 Praat와 Python(Jupyter)를 이용해서 직접 음성분석을 할 예정.

음성학이란?

★ A study on speech 말소리 자체에 관한 연구

크게 3가지로 분류할 수 있는데

1. 어떻게 소리가 나오는가? (from mouth)
2. 어떻게 소리가 전달되는가? (through air)-> 공기를 타고 가는 과정
3. 어떻게 소리가 들리는가? (to ear)

이중에서 1번이 음성학의 가장 주된 부분이다.

Articulation (소리를 입에서 만들어내는 과정)

Vocal tract는 코 쪽 통로인 nasal tract와 입 쪽 통로인 oral tract로 나눌 수 있다. 그리고 목 안쪽의 larynx라는 부위가 존재한다.

먼저 구강통로의 위 쪽 부분은 lip, teeth, alveolar, hard palate, soft palate(velum), uvula,

larynx가 존재한다 (순서대로 입술, 이, 치경, 경구개, 연구개, 목젖, 후두)

반대로 아래 부분은 lip, tip, blade, front, center, back, root, epiglottis가 존재한다 (순서대로 입술, 혀끝, 혀날, 혀의 앞부분, 몸통, 뒷부분, 혀뿌리, 후두 덮개)

또 연구개(velum)는 비음이 아닌 소리를 낼 때 비강으로 가는 통로를 막는 역할을 한다. 코로 숨을 쉴 때는 비강통로가 열려야하므로 연구개가 내려가고 비음이 아닌 소리를 낼 경우에는 연구개가 올라가서 비강통로를 막는 구조이다. 비음의 예시로는 m, n, ng 등의 소리가 해당된다. 이 소리들은 비강통로를 통해서 소리가 나므로 소리를 내는 도중 코를 막으면 제대로 발음할 수 없다. 또한 우리가 일상생활에서 숨을 쉴 때 코를 통해서 숨을 쉬므로 평상시에는 연구개가 내려와서 비강통로가 열려있는 상태라는 것을 어렵지 않게 유추할 수 있다.

목 안쪽에는 voice box라고도 불리는 후두(larynx)가 존재하는데 소리의 종류가 유성음이나 무성음이나에 따라서 후두가 진동하거나 진동하지 않는다. V, z, l, m 등의 유성음을 낼 경우 후두가

진동하는 것을 느낄 수 있다.

소리를 낼 때 각 constrictor가 사용된다 하더라도 어디에 닿을 것 인지? 얼마나 막을 것 인지?  
이러한 요소들에 따라서 소리가 달라진다.

Constriction Location(CL)의 차이에 의한 소리 변화의 예시로, 아랫입술이 윗입술에 닿느냐 또는 윗니에 닿느냐에 따라 소리가 달라지고 마찬가지로 혀끝이 이빨에 닿느냐 치경에 닿느냐 경구개에 닿느냐에 따라 소리가 변한다.

Constriction Degree(CD)의 예시로 파열음, 마찰음, 파찰음 등의 차이를 들 수 있다.

소리를 포함한 이 세상의 모든 신호들은 단순한 형태의 sine wave가 여러개 합쳐져서 만든 형태이다. 따라서 아무리 복잡한 소리라도 분석해보면 여러가지 frequency의 sine wave들의 합으로 표현할 수 있다. praat를 이용해서 이를 직접 눈으로 확인해 볼 수 있다.

먼저 스스로의 목소리를 녹음한 다음 그 포맷트를 분석해보자. 일정한 주기로 파동이 반복되고 있음을 관찰할 수 있는데 그 주기를 주파수로 바꾸면 바로 그 주파수가 자신의 목소리의 주파수임을 알 수 있다. 필자의 경우 딱 100Hz로 나타났다.

그 다음 자신의 목소리의 주파수와 같은 주파수를 가지는 pure tone을 만든다음 그 소리를 비교해보자. 기계음이기 때문에 사람의 목소리와는 다르지만 같은 음을 내는 것을 알 수 있다. 이때 목소리가 다르게 느껴지는 이유는 그 소리의 source와 소리를 내는 과정에서의 filter의 차이가 있기 때문이다. 여기서 filter란 목소리를 내는 과정에서 발생하는 변형, 즉 입모양으로 인한 변형을 뜻한다.

말소리의 source는 larynx의 소리이다. Siswati어를 하나는 직접 말소리를 녹음하고 또 다른 하나는 사람의 목, larynx부분에 대고 녹음한 소리를 비교해보면 톤이나 음의 높이등은 같지만 발음에서 차이가 느껴지는 것을 알 수 있다. 이는 larynx의 소리는 입모양으로 인해 filtering 되지 않은 소리이기 때문이다.

사람 목소리처럼 여러개의 pure tone(sine wave)의 합으로 이루어진 tone을 complex tone이라고 한다. 사람의 목소리의 주파수가 100Hz라고 가정하면 그 주파수가 Fundamental frequency가 되고 이 배수만큼의 주파수들이 합해져서 complex tone이 만들어진다.

예를 들자면 100Hz가 Fundamental frequency(F0)이고 그 다음 2배수인 200Hz의 sine wave가 F0의 크기보다 낮은 크기로 더해지고 그다음 300Hz의 sine wave도 더해지고... 이런 식으로 계속 반복되어 complex tone이 만들어진다.

실제로 praat를 이용해서 녹음한 목소리의 spectrum을 분석해보면 파동의 peak부분이 자신의 f0인 주파수의 배수마다 나타나는 것을 확인할 수 있다.

이는 실제로 100Hz부터 1000Hz까지의 pure tone을 amplitude를 0.05씩 낮춰가면서 합성해 complex tone 을 만든 뒤 그 spectrum을 분석해봐도 확인할 수 있다.

이때 praat상에서 spectrum분석을 통해 볼 수 있는 spectrum을 위에서 본 모양으로 진한 부분과 연한 부분으로 나타낸 그래프를 spectrogram이라고 한다. 진한 부분이 peak에 해당하는 부분으로 formants라고도 한다.

일반적으로 모음을 비교할 때에는 Formant 1과 Formant 2를 비교해보면 충분하다. 이를 바탕으로 해서 F1과 F2의 위치에 따라 알 수 있는 모음의 종류를 나타낸 그래프를 vowel space라고 한다.