영어음성학 _ Class Summary (Wk3)

2018130913 영어영문학과 김동현

-English consonants & vowels : 영어의 자음과 모음은 여러가지 기준(조음 위치, 유성음/무성음, 이중모음/단모음 등)으로 분류될 수 있음. 이러한 기준들이 향후에 음성학에서 다루게 될 기본적인 요소들과 관련이 있음.

-Phonetics : 음성학. 음운론(Phonology)이 소리가 어떠한 시스템 하에 형성되는가를 인지적인 수준에서 다루는 것이라면, 음성학은 그보다 더 깊이 들어가 물리학적인 측면 에서 소리를 바라봄. 음성학을 이루는 3가지 주요한 요소로는 Articulatory phonetics(인 간 발화가 시작되는 메커니즘), Acoustic phonetics(발화가 공기를 타고가는 원리), 그리고 Auditory phonetics(귀를 통해 발화를 듣는 것)을 꼽을 수 있음.

-Articulation : 인간은 vocal tract의 각기 다른 부분을 움직여 다양한 소리를 만들어냄. (Vocal tract는 larynx, pharynx, oral cavity, nasal cavity.로 이루어짐) 조음 기관으로는 크게 다섯 가지를 꼽는데, lips, tongue tip, tongue body, velum, larynx가 이에 해당함. 그리고 lips, tongue tip, tongue body를 묶어 articulatory process, velum을 nasal과 non-nasal을 구분하는 oro-nasal process, 글고 larynx를 유성음과 무성음을 구분하는 phonation process라 일컫는다. 우선, phonation process는 성대의 진동에 따라 진동이 있을 경우 유성음, 없을 경우 무성음으로 본다. 한 편, oro-nasal process는 velum이 lower되어 nasal tract가 열릴 경우 nasal, 그렇지 않을 경우 non-nasal sound로 여긴다. 이외에도 articulatory process를 통해 많은 소리들을 구분한다. 추가적으로, 이러한 조음 기관들을 constriction이 일어나는 위치와 정도를 통해 보다 세부적으로 구분한다.

-Phoneme : 음소. 의미를 갖는 소리의 최소 단위. Acoustic phonetics는 praat을 통해 이후에 더 자세히.

영어음성학 _ Class Summary (Wk4)

2018130913 영어영문학과 김동현

-Vowel acoustics : 소리가 만들어지는 mechanism. sine wave에는 소리의 특성을 구분 짓는 두가지 성분이 있는데, frequency(Hz)라 불리는 1초동안에 생긴 sine wave의 수와, sine wave의 크기가 이에 해당함. 이를 이용해 Praat을 통해 소리를 직접 생성해볼 것.

-Simplex/complex tone in spectrum: Frequency와 magnitude에 따라 sine wave의 형태가 결정됨. 참고로 이 세상 모든 신호는 여러 다르게 생긴 sine wave의 결합으로 표현됨. 이는 여러 simplex tone의 결합(synthesis)으로 complex tone이 보여지는 것을 이야기하는데, 각 tone들을 x축이 frequency, y축이 amplitude인 형태로 나타내는 것을 spectrum이라 함. Complex tone을 이루는 여러 다른 simplex tone 중 가장 느린 simplex tone이 우리 목소리의 pitch를 보여줌. 한 편, praat에서 sine wave의 아랫 부분에 나오는 그림은 spectrogram으로, spectrum을 시간축으로 늘여놓은 것에 해당함.

-Source & Filter: 우리 인간이 내는 소리는 입모양에 따라 달라짐. 본래 목구멍 아래에서 성대가 떨리며 나오는 소리 자체를 source라 한다면, 성대로부터 입술까지의 tube가 filter 역할을 함으로써 각각의 소리를 다르게 하는 것임. Source의 spectrum을 보면, F0가 fundamental frequency, 즉 pitch(number of vibration in a second)를 의미하며, 그 이후에나오는 것들을 harmonics라 하는데 이 harmonics들의 합으로 voice source가 정해지는 것임. 여기서 각 sound의 spectrogram에서 어두운 산맥 부분을 분석하면 formant들을 얻어내는데, 이 formant의 특성만으로 우리는 인간의 vowel sound들을 구현해낼 수 있다. 또한, 여기서 formant 1을 모음의 혀의 높낮이, formant 2를 혀의 front/backness로 나타내면 마치 IPA에 나오는 vowel chart와 같은 결과물을 보게 된다.

영어음성학 _ Class Summary (Wk5)

2018130913 영어영문학과 김동현

-코딩: 기술이 발달해감에 따라 우리 사회에서 대두되는 자동화 시스템에 의해 코딩이 필요해짐. 코딩은 computer language로, 일종의 language에 해당함. 참고로, Language는 정보를 담는 그릇에 해당하는 '단어'와 해당 단어에 담겨지는 '정보의 종류', 이렇게 2가지로 구성되는데 코딩의 경우에는 단어가 variable, 정보의 종류가 숫자, 글자 및 list로 이루어짐.

-코딩의 문법: Language에 문법이 있듯, computer language에도 사람과 기계 사이의 communication 방식이 되는 문법이 있는데, 이는 크게 4가지로 구성되어 있음. 1) 변수라는 그릇에 특정 정보를 assign하는 행위 2) 조건에 대한 문법. If conditioning 3) 여러번 반복하는 문법. For loop 4) 함수 형태(입력값에 해당하는 출력값이 존재).

-코딩 활용 : Anaconda prompt 접속 후 jupyter notebook 검색 후 작업할 directory 선택 / Computer language에서 equal sign(=)은 같다는 뜻이 아니라, 오른쪽의 정보를 왼쪽의 variable에 assign한다는 뜻임. *Equal은 variable assignment. 따라서, a=1이라 하면 a라는 variable이 1이라는 정보를 갖게 하는 것임. / Print(a) - print는 일종의 함수로 소괄호 안의 입력값으로 들어가는 variable이 가진 정보를 출력값으로 나타나게 해줌. / 특정 cell 자체를 선택한 후 b 누를 시 below에 cell을 만들어줌. a를 치면 above에 동일한 작용. Cell을 지울 때에는 x를 누를 것. / 똑같은 variable에 서로 다른 두 값을 넣으면 이후에 입력한 값으로 덮어 씌워짐. / 알바벳은 변수, 숫자는 정보로 고정. 따라서 문자열을 때에는 " " 필수. / Run 단축키 : shift + enter / Cell의 마지막에 그냥 변수명을 치면 output 보여줌. / 대괄호할 경우 한꺼번에 정보 넣어줌. 이를 list라고 함. / type은 변수의 특성(int, float, string, list)을 알려주는 함수. / 소괄호는 tuple이라고 나오는데 list와 동일.

영어음성학 _ Class Summary (Wk6)

2018130913 영어영문학과 김동현

-정보 retrieve : a=1, b=1, c=a+b, c=2. 여기서 c를 실행하면 a와 b 각각의 정보가 retrieve되어 결과가 나오는 것임. 이런식으로 우리는 코딩을 하는데 있어 정보를 retrieve함으로써 다양한 활용을 하는데, 특정 variable의 정보를 부분적으로 retrieve할 때에는 대괄호를 이용함. 대괄호 안에 들어가는 것은 Index임. 예를 들어, a[0]는 a list에서 0번째 정보를 가져오라고 하는 것임.

-기타 코딩 활용 : semicolon은 여러 줄에 나눠적을 것을 한 줄에 연속하여 적고 싶을 때 사용함. float함 수는 어떤 float() 했을 경우 괄호 안의 variable을 float타입으로 바꿔줌. Int()도 동일.

-Dict: dict는 중괄호를 이용하여 pair로 표현함. 이는 이름 그대로 사전과 동일한 원리를 가지고 있음. 왼쪽이 표제어, 오른쪽이 그에 해당하는 정보인데, 사전이 표제어를 통해 정보에 접근하듯 이 또한 표제어를 index로 씀. 결국 Dict가 정보를 retrieve하는 방식은 '몇번째 정보를 가져와라'가 아닌, '이것에 해당하는 정보를 가져와라'가 됨. 참고로, 표제어는 string뿐만 아니라 숫자도 가능함.

-String: String은 list와 정보에 access하는 방식이 거의 동일함. 순서를 논하는 index에서 0이 첫번째를 의미한다면 음수는 맨뒤에서부터 거꾸로 세는 순서에 해당함. 또한, range표시는 :을 이용, 1:3과 같이 표현하는데 이는 첫번째에서 세번째 직전까지의 정보들을 retrieve하는 것임. :왼쪽을 표시하지 않으면 0번째부터, 오른쪽을 표시하지 않으면 끝까지, 그리고 둘 중 아무것도 표시하지 않으면 전체를 가리킴. 추가적으로, 함수 len()은 variable내에 있는 정보의 길이를 나타냄. .upper()은 문자들을 대문자로 바꿔주는 함수이고, s.find()는 괄호안의 문자가 몇번째부터 시작하는지 알려줌. 또한, .rindex()는 last index를 찾아주되 count하는건 왼쪽부터에 해당함. .strip()은 잡스러운 것을 지워주는 함수이며, .split()은 string을 괄호안의 것을 도구로 이용하여부분부분으로 나눠주는 함수에 해당함. 한 편, split된 것을 다시 문장으로 복구하고 싶을 때에는 ''.join()을 이용함. .replace('a','b')는 string 속의 모든 a를 b로 바꿔주는 함수임.

-syntax: for loop는 시행에 있어 여러 번을 여러 개 해야할 때 사용함. 즉, 반복되는 행위가 focus되는 것인데, 예를 들어 a안의 정보들을 하나하나 print out 하고 싶을 때 print out이 반복되는 행위에 해당함. For ~ in ~: 가 기본 형태이며, in 뒤에 있는 것을 하나씩 돌려서 i에 각각의 것들을 할당한 후 for 밑에 적힌 행위를 하라고 명령하는 것임. 이 때 명령하는 것은 indent되어있어야함. range()함수는 괄호 속의 숫자 개수만큼의 index list를 만들어줌. 예로, 숫자가 4라면 0,1,2,3이라는 index list를 만들어줌.

영어음성학 _ Class Summary (Wk9)

2018130913 영어영문학과 김동현

-numpy : 일종의 library이자 package. 하나의 numpy 안에 연속해서 또 하나의 package를 만들 수 있음. 그리고 각각의 package들 안에 함수들이 있는 것. 참고로 이 numpy를 이용하려면 import해서 numpy를 불러와야함. import numpy 라고 하면 numpy 속에 있는 것은 전부 쓸 수 있게 되는 것임. 한 편, import를 표현하는 다양한 법이 있는데, 예로 import numpy / numpy.A.D.f 라는 표현을 다르게 표현하는 한 예로 from을 이용한 from numpy import A / A.D.f가 있음. import 한 부분이 기준점이 되는 것임. 또한, 내가 import한 것을 다시 사용할 때 특정 용어로 설정할 수 있는데, 이는 as를 이용한다. (import matplotlib.pyplot as plt)

-기타 함수들: 참고로 행렬은 [2,3]과 같이 표현함. 이 때 앞이 행, 뒤가 열에 해당함. np.zeros는 특정 matrix의 숫자들을 모두 다 0으로 채워줌. 마찬가지로, np.ones는 0대신 1이 들어감. 다만, 주의할 점은 zeros에서 default값이 float로 되어있다는 것. 때문에 1.과 같이 표현된다면 dtype을 int로 해줌으로써 소수점을 사라지게 할 수 있음. np.array는 list를 계산이 가능한 array로 바꿔주는 함수임. np.arrange는 range와 비슷한 기능을 하는 함수에 해당함. 다만 이를 계산 가능한 array형태로 만들어주는 것임. 예로, np.arrange(5)라 하면 array([0,1,2,3,4])가 나옴. np.arrange(0,10)은 0부터 9까지의 수가 나열되고, np.arrange(0,10,2)는 0부터 10미만의 수들을 증가분을 2로 한 상태에서 나열해줌. np.linspace는 등차수열을 얘기하는데, linspace(0,10,6)은 0부터 10까지의 수를 6개로 나누어주는 함수임. X.ndim은 X의 차원을 알려주는 함수이고, X.shape은 차원의 detail을 알려주는 함수인데, 가장 큰 괄호에 대한 정보부터 표시됨. X.dtype은 X속의 숫자들이 어떤 type인지 알려주는 함수이며, X.astype은 type을 바꾸고 싶을 때 쓰는 함수임.

영어음성학 _ Class Summary (Wk10)

2018130913 영어영문학과 김동현

Sound에 관해서 직접 어떻게 만드는지.

Simple tone. Sine wave = sinusoidal. Sinusoidal function을 만들어내는 것을 phasor.

Sin().()안에 뭐가 들어갈까? // degrees를 파이를 이용해 표현하는 것이 radian? What is radian?

Sin, cos 함수 입력값에는 degrees가 아닌 radians 이용. 파이가 180도? Sin, cos함수 모두 2파이를 기준으로 형태가 반복됨. Ex) cos(theta) = 0, theta=3/2파이일때, (오일러 공식 찾아봐라)

Vector = 숫자 열.

Projection – flashlight로 비출 때 어떻게 나오는가? 특정 방향(축)에 projection하고 그 방향에서 관측. 이 행위는 우리가 특정 수만을 관측하겠다하는 의도가 담겨있음. Ex) a+bi, x축 a, y축 b, x축 에 projection하는 것은 실수만 보겠다.

실체의 소리를 만들어내기 위해서는 시간 개념 필요. 시간 개념이 들어가지 않은 단순 sin, cos만 으로는 소리 만들 수 없음.

0부터 2파이까지 각도의 벡터를 만드는 법 - (0, 2파이)

default니까 1만큼 increase, 2*np.pi는 6.28... 이므로 6까지 생성됨. Theta가 radian으로 정의됨.

그것을 sin에 넣어 s로 받음. 이를 plot하면 sin곡선 생김.

화면전체가 figure. Ax는 뭐하는 거냐. Subplot은 전체 figure 속 각각의 square. 화면 분리.

괄호 속 숫자 : (221); 줄을 2by2, 열과 행으로 나누는데 각각에 번호 1, 2,3, 4 매김

Ax.plot(theta, s, '.') theta가 x값, sin이 결과값이므로 y축.

영어음성학 _ Class Summary (Wk11)

2018130913 영어영문학과 김동현

- Sampling rate이 100Hz라고 생각. 우리가 표현할 수 있는 숫자의 개수가 1초에 100개. 이 숫자들을 통해 1Hz 짜리 frequency를 표현할 수 있을까? - 있다. 그냥 한번만 주기 그리면 됨. 2Hz도가능 / 그 Hz의 숫자를 계속 올려서 10000Hz의 frequency를 표현할 수 있을까? (10000번을 1초동안 왔다갔다하게) 안됨. 우리가 가진 숫자가 너무 적음. 따라서, sr이 1초에 적당히 있어야 그만큼의 주파수를 표현할 수 있는 것임.

Ex) sr=10Hz, Fr=100Hz 표현가능? 주어진 sr로 표현할 수 있는 반복되는 주파수(Fr)의 개수는 Max. 1/2*sr. (그래프상 한점은 위, 한점은 아래로 하면 점 두개를 가지고 주기를 만들어 낼 수 있음.) – Nyquist frequency = 1/2*sr

CD의 Sr=44100Hz, (nyquist는 22050) 사람의 가청주파수가 20000Hz까지이기 때문.

유선전화의 경우 sr=8000. 통신사에서는 최대한 아껴야하므로. 따라서 표현 소리의 폭이 제한적이라 목소리 구분 힘들기도.

FO (제일 낮은 것). 그 후부터 f1, f2 ... 2배, 3배... - 그 끝은? Sampling rate의해 결정(sr의 절반). 제일 마지막은 Fend (end). Fend = int(sr/2) = nyquist frequency. 소수점 나오면 지저분하므로 int. (여기서 time은 만들었다고 assume) / 우리가 phasor만들 때 time 만들고, 거기로부터 theta만들고, sin에 입력. 그 과정을 여러번 반복하는데 frequency만 바꿔주면 반복하는게 여기 코드의 핵심. for loop 통해 여러번 반복. s라는 변수에 더하고 또 더하고. 쌓아올리기. s=s+ ~~ 식으로 표현.

다만, 여기서 s를 미리 정의 해줘야함. 첫 루프에서 맨처음 s가 무엇인가. s = np.zeros(len(t)) --time vector의 개수만큼 0을 만들어주면 됨. 이후 plotting.

제일 처음은 0에서 시작. 거기서 sin wave 더하고 더하고..

+왜 pulse train? 우리가 sin wave에서 부드러웠던 곡선부분이 거의 없어짐. 극단적으로 갈수록 선하나가 남아있고 00000... 선 하나 0000...

For loop부분 해석 중요. F0만큼의 증가폭. s정의 해줘야하는 것 유의.

일단은 뚜껑없는 목구멍에서 나는 소리만 만들었고, 이 이후로는 a,i등의 여러 소리 어떻게 만드느냐 볼 것.

영어음성학 _ Class Summary (Wk12)

2018130913 영어영문학과 김동현

행렬은 직사각형. vector는 행렬 중에서도 한 줄 짜리로, 가로 혹은 세로로 한 줄. (숫자열)

데이터를 입력값으로 중간의 어떤 기계, 혹은 인공지능 (함수 역할)을 통해 변형이 되어 새로운데이터 생성.(출력값). 각 데이터의 형태는 vector. ex) vector형태 음성이 들어가 vector형태 text가나오면 음성인식이라 함. Text가 들어가 소리가 나올 수도. (어떤 vector가 들어가 어떤 vector가나오는 것)

그럼 중간의 기계, 혹 인공지능 (함수)은 어떤 형태?

-행렬.

Vector와 중간 행렬이 곱해지는 원리 = 행렬의 각 열을 가로로 나열 후 vector의 각 수(자리가 일치하는 수끼리) 와 곱하고 각각의 수들을 모두 더한 값을 출력값의 vector로 순서대로 배치. 참고로 중간의 행렬은 많은 기계 학습을 통해 얻어진 수들.

인공지능은 행렬의 곱. 입력벡터를 출력벡터로 바꿔주는 함수의 역할. 입력구, 출력구 모두 음성, 텍스트, 이미지, 등등 뭐든 될 수 있음. Vector의 형태로 표현되어야하는 이유는 행렬의 곱을 가능 케 하기 위해서임.

/ Space 관련 이야기

3*2matrice에서 column vector는 3쪽과 관련. component가 3개기 때문. 점이 표현되는 dimension이 3차원이기에 column. whole space는 3차원 / column space는? 원점과 두 점 연결해서 삼각형 칠해보면 그저 2차원의 plain. 이를 밀가루 밀듯이 다 밀어버리면 전체가 됌. 이를 spanning이라 함. 그럼 나머지 비는 차원은? null space (정확히는 left null. 왜? 좀이따 알려줌)

- +orthogonal(직각,직교의) : 위의 2차원 plain과 orthognal한 것은 원점 관통하는 한 선. 이 선은 1 차원. 이 선이 null space. column vector가 만들어낸 space에 orthognal한 것이 null space.
- +Spanning = linear combination으로 표현가능한 모든 것 (기하적). 결국 위의 예에서 column vector는 각 column vector에 수를 곱하고 둘을 더한 모든 값들의 집합.
- -row space : row vector 가지고 이야기. (강의상 예 참고) row wise whole space = 숫자를 2개 쓰니까 2차원상의 점들. 결국 2차점. row vector들이 만들어내는 space는 2차원을 넘어가지 않음. row space도 2차원.

결국 column space든 row space든 independent한게 몇 개 있느냐가 중요.

그럼 2차원 상에서 independent한게 3개인 건 어떻게 설명? Independent한 것은 한 space에서 나머지 둘의 linear combi.에 의해 나오지 않음. 3개가 아니라 2개인 것. 결국 column관점이든, row 관점이든 independent한 vector들의 숫자(rank)는 늘 똑같음.

Rank는 independent한 것의 개수.

영어음성학 _ Class Summary (Wk13)

2018130913 영어영문학과 김동현

null space 정의 : Ax=0 : AY 뭐가 오든 0이 되게 하는 x. 그 x가 사실은 null space상의 모든 구성부분들. 어떤 입력이 들어오든 output에는 영향 미치지 않는 것. 그 효과가 0이다.

인간은 점점 어떤 일을 할 때, skill이 늘어갈수록 쓸데없는, 딱 필요한 것 이외의 skillful action들을 함. 이러한 것은 출력 쪽에 영향을 미치지 않는 입력값의 영역들에 해당함. 우리가 살아가는 과정은 null space를 늘려가는 과정으로 볼 수 있음.

graph상에서 null space와 평행한 방향성을 가지고 움직이면 결과에 영향을 미치지 않는 입력값 들 but 다른 방향성 : 결과에 영향

eigenvector site

- Av = b. v를 basis vector로 놓고, a1과 a2 1,2와 2,1이라는 2개의 vector.
- -- 전에 봤던 평행사변형 꼭짓점으로의 이동. 기본적 linear transformation

eigenvector: 원점과 입출력이 평행하는 순간이 생김.

여기서 필요한 사고. vector는 방향. 평행한 선을 eigenvector라고 하는데, 그것과 같은 방향으로 가는 것에 대해 이야기. 여러 가능한 eigenvector의 집합 : eigenspace . eigenvector는 어떤 한 값이라기보다는 방향.

+2by2에서 eigenvector는 2개. 3by3는 3개.. -

-eigenvalue : 각각의 eigenvector에 대해 eigenvalue.존재. 입력값과 출력값 비교. 얼마나 더 늘어난 것인가. 그 비율. 원점으로부터 v의 거리와 Av까지의 거리 비율. eigenvector 2개있다면 각각의 2개 eigenvalue존재.

null space - 사람은 진화해오며 null space쪽 확장시켜감. Ax=b에서 b가 하고자 하는 task라 할 때, 우리가 그 task를 이루고자 하는 과정에서 장애물들이 발생할 수 있는데 이를 피해 결국엔 같

은 결과값을 도출하도록 해주는 것 필요.

-eigen 왜 배우는 것?

- 2by2를 또다른 2by2로 바꿔준 셈. 2by2 column vector를 통해 2by2 eigenvector

eigen = 독일어 고유. 훨씬 더 고유한 것. 더 unique한 것. 주어진 섞여있는 matrix를 통해 이를 쪼개서 훨씬 더 unique한 결과로.

상관관계 - correlation --> r로 표현.

여러 사람들별로 영어, 국어, 수학, 과학 점수값들 쭉 나와있음.

graph상에서 한 점이 한 사람이 됨. ex) y축 영어 x축 수학 이런식. 어떤 graph는 random해보이고, 어떤 graph는 상관성이 있어보임. (영어 수학 graph보다 영어 국어 graph에서 더 상관적인게보임. '두개가 같이간다'.)

-1 < r < 1 / 0에 가까워질수록 상관관계 낮아짐.

+1, -1이 정확하게 나오는 경우 : 완전한 비례 반비례 선상에 있으면. 얼마나 선상에 가깝냐에 따라 1 or -1에 가까워짐. 점 찍힌 형태가 원에 가까워지면(패턴x) r 0에 가까워짐.

여기서 선형대수가 어떻게 쓰일 수 있나.

사람 85명 있을 때, 85개의 차원가진 국어 vector, 영어 vector ... 각 차원이 한 사람이 됨.

원점으로부터 국어 한점, 영어 한점, 수학 한점. 삼각형 식으로 점 두개 잡고 연결해서 그 사이각도 theta, cos theta = r. theta 90일 때 0 (서로 orthogonal), 0일때 1 (방향성 생각)

그럼 그 각도 재는 방법?

inner product / dot product

원점이 있고, 두 vector(a,b)가 있을 때, 그것의 inner product는 두 vector 서로 곱하고 다 더한 것. 안에서 일어나는 product. inner product. dot product 라고 표현하기도함. a (dot) b.

기하학적 해석도 중요. 원점, a, b 있을 때 한 점에서 다른 한점과 원 이은 선에 대해 수직으로 내린 후, 원점으로부터 그 점(수직으로 내려서 만난) 길이 곱하기 원점으로부터 다른점(a or b)의 길이. -- 두 개 vector 있을 때 수직으로 내리고 두 길이 곱해버리면 됨.

-수식으로 정리. (a의 길이 곱하기 cos theta) 곱하기 b의 길이

a, b 길이 구하는 법 : 루트 1제곱 플러스 2제곱 플러스 3제곱 (a가 1,2,3이라 할 때.) (2차원에서 생각해보자. - 피타고라스 공식)

inner product 필요한 이유 : 어떤 signal있을 때 어떤 frequency주파수 대가 많은지를 spectrogram이 보여줌. 그 spectrogram을 직접 만들기 위해 여태 위의 내용들을 다룬 것임.

signal 분석할 때 여러 Hz의 sine wave들 다양하게 만든 후, inner product해버림. 그러면 특정 값이 나올텐데, 서로서로 유사한 성분이 있다면 inner product가 높게 나올 것임.