Abhigyan Ghosh - 20171089

Speech Signal Processing [ECE448]

20 August 2020

Assignment 1

1. Question 1

1.1. **Coarticulation**: It is the idea that each speech sound is affected by every other speech sound around it, and each sound slightly changes according to its environment. Example: *feel* [fi:l] →[fi:jəl][i]

1.2. **Phonation**: It is the process by which the vocal folds produce certain sounds through quasi-periodic vibration.[ii]

1.3. **Fundamental Frequency**: It is the lowest frequency in a periodic waveform. Since the fundamental is the lowest frequency and is also perceived as the loudest, the ear identifies it as the specific pitch of the musical tone.[iii]

1.4. **Epochs**: Epoch is the instant of significant excitation of the vocal-tract system during production of speech.[iv]

1.5. **Formants**: A formant is a concentration of acoustic energy around a particular frequency in the speech wave. Each formant corresponds to a resonance in the vocal tract.[v]

1.6. **Pitch**: In speech, it is the relative highness or lowness of a tone as perceived by the ear, which depends on the number of vibrations per second produced by the vocal cords.[vi]

2. Question 2

It is true that female pitch is more when compared to male pitch. Women speak at almost. An adult woman's average range is from 165 to 255 Hz, while a man's is 85 to 155 Hz.[vii]

3. Question 3

Speech is human vocal communication using language. It is the most natural use of language for humans.

Signals can be categorized into two groups: periodic and aperiodic. A signal which repeats itself after a specific interval of time is called periodic signal. A signal which does not repeat itself after a specific interval of time is called aperiodic signal. Speech is unique as it is quasi-periodic, that is, it repeats over short periods of time but over a long time, it is aperiodic.

4. **Question 4**

I am the son of Chhanda

/ai/ /ae/, /m/ /d/, /a/ /s/, /a/, /n/ /au/, /f/ /chh/,/n/, /d/, /A/

4.1. Acoustic-phonetic description of the regions (MOA and POA):

/ai/: It is a diphthong

/ae/, /m/: It is a diphthong followed by a bilabial nasal

/d/, /a/: It is a dental unaspirated stop followed by middle vowel

/s/, /a/, /n/: It is a dental fricative followed by back vowel followed by dental nasal

/au/, /f/: It is a diphthong followed by an unvoiced unaspirated fricative

/chh/, /a/, /n/, /d/, /A/: It is an unvoiced aspirated palatal followed by a middle vowel followed by a dental nasal followed by a voiced unaspirated dental stop followed by long vowel

4.2. Time varying system description:

/ai/: Tongue hump at alveolar ridge, narrow opening at alveolar ridge, VT system is narrowly open.

/ae/: Tongue hump is low and it is in central position of the vocal tract (VT) system, VT system is widely open

/m/: Opening of velum and closure at lips.

/d/: Complete closure at dental

/a/: Tongue hump is medium, and it is in front position of the VT system, VT system is moderately open

/s/: Narrow constriction at dental

/a/: Tongue hump is low, and it is in central position of the vocal tract (VT) system, VT system is widely open

/n/: Complete closure at dental opening of nasal cavity

/au/: Tongue hump is mid and back position of the vocal tract (VT) system, VT system is widely open and cylindrical in shape

/f/: Narrow constriction between lower lip and the upper teeth

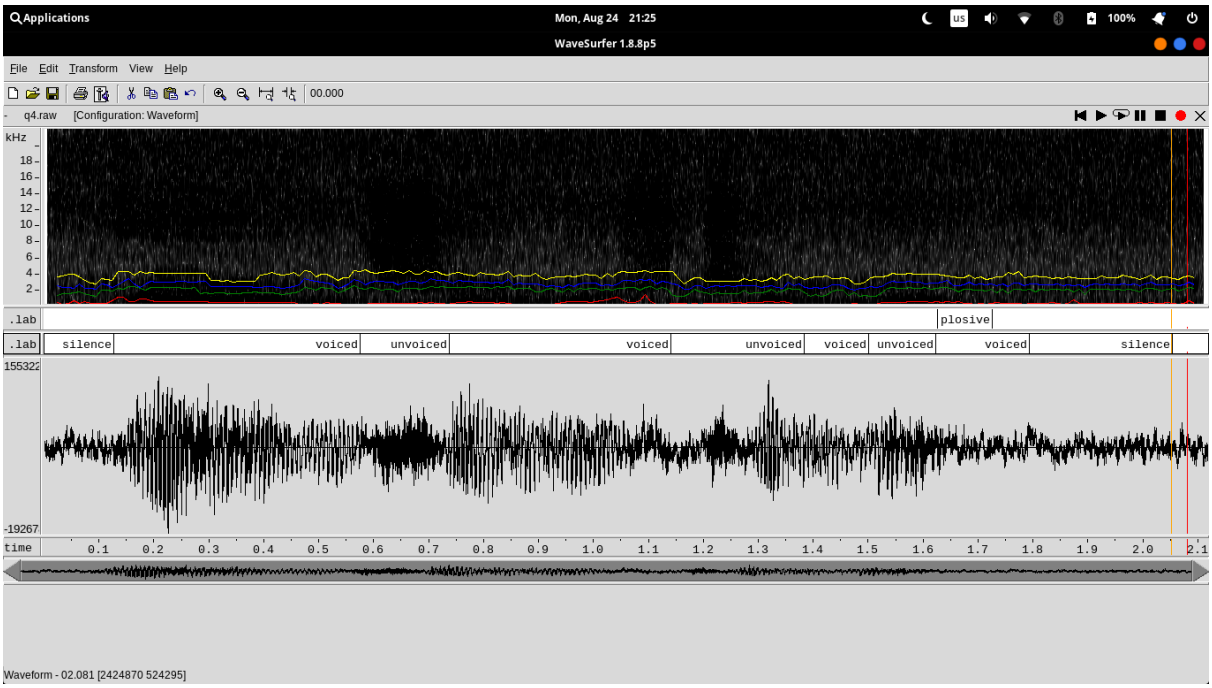
/chh/: Complete closure at palatal

/n/: Complete closure at dental opening of nasal cavity

/d/: Complete closure at dental

/A/: Tongue hump is low, and it is in central position of the vocal tract (VT) system, VT system is widely open

4.3. Spectral details for sounds units present in the waveform:



Waveform and spectogram of utterance

Phoneme	F1	F2	F3	F4
ai	552	1646	2380	3661
ae	646	1618	2320	3446
m	372	1623	2356	3412
s	287	1725	2648	4324
a	404	2358	2967	3977
n	254	2119	2714	4326
a	610	1962	3360	4315
f	370	1616	2495	3099
chh	512	2058	2590	3505
a	413	1944	2695	3462
n	501	2218	2997	4015
d	280	2076	2992	3891
A	320	2124	2920	4267

Table of first four formant frequencies for each phoneme

5. **Question 5**

I am from Jamshedpur

/ai/ /ae/, /m/ /f/, /r/, /o/, /m/, /j/, /a/, /m/, /S/, /e/, /d/, /p/, /u/, /r/

5.1. Acoustic-phonetic description of the regions (MOA and POA):

/ai/: It is a diphthong

/ae/, /m/: It is a diphthong followed by bilabial nasal

/fr/, /o/, /m/: It is an unvoiced unaspirated fricative followed by dental semi-vowels followed by back vowel followed by nasal

/dj/, /a/, /m/, /sh/, /e/, /d/, /p/, /u/, /r/: voiced post-alveolar affricate followed by a middle vowel followed by a bilabial nasal followed by palatal fricative followed by front vowel followed by unaspirated voiced dental stop followed by unvoiced unaspirated bilabial plosive followed by back vowel followed by alveolar semi-vowels

5.2. Time varying system description:

/ai/: Tongue hump at alveolar ridge, narrow opening at alveolar ridge, VT system is narrowly open.

/ae/: Tongue hump is low, and it is in front position of the vocal tract (VT) system, VT system is widely open

/m/: Opening of velum and closure at lips.

/f/: Narrow constriction between lower lip and the upper teeth

/r/: Partial closure of VT with tongue tip at alveolar ridge

/o/: Tongue hump is medium, and it is in back position of the VT system, VT system is moderately open and cylindrical in shape

/m/: Complete closure at lips opening of nasal cavity

/dj/: Complete closure at dental, Complete closure at palatal

/a/: Tongue hump is low, and it is in central position of the vocal tract (VT) system, VT system is widely open

/m/: Complete closure at lips opening of nasal cavity

/sh/: Narrow constriction at palatal

/e/: Tongue hump is medium, and it is in front position of the VT system, VT system is moderately open

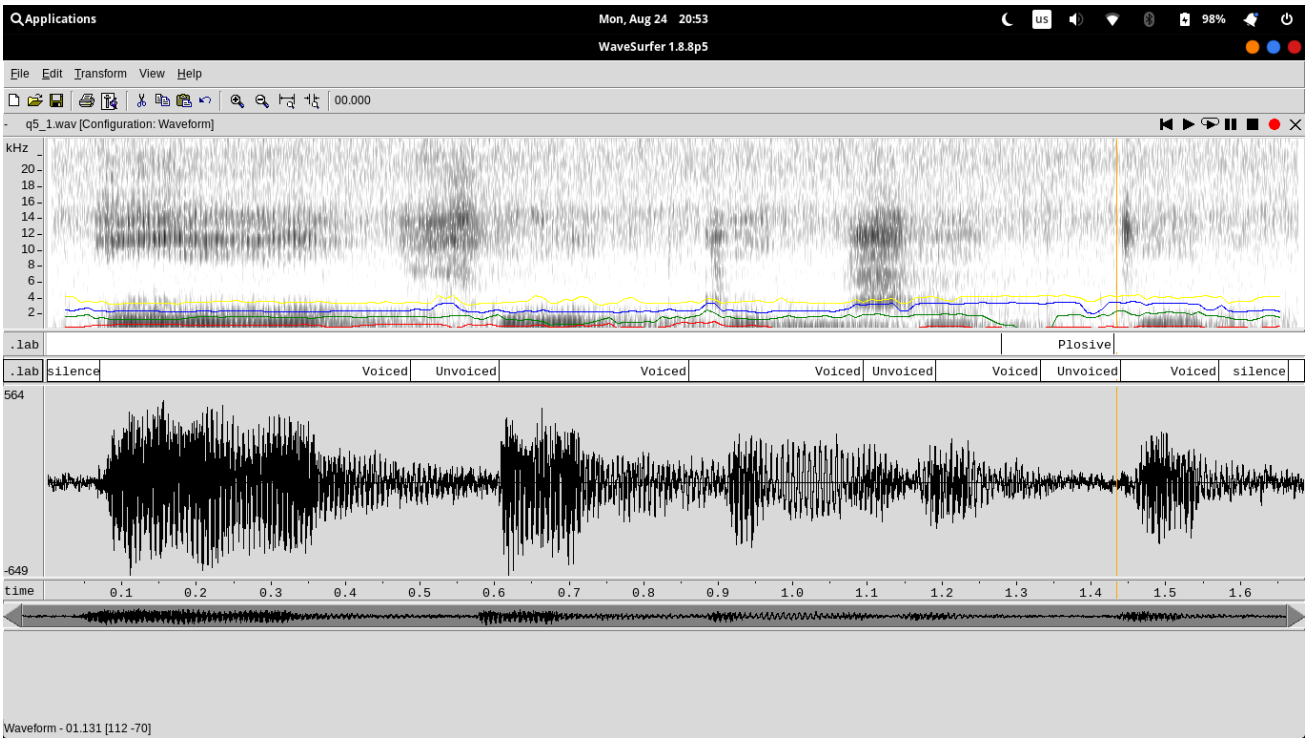
/d/: Complete closure at dental

/p/: Complete closure at lips

/u/: Tongue hump is high, and it is in back position of the VT system, VT system is narrowly open and cylindrical in shape

/r/: Partial closure of VT with tongue tip at alveolar ridge

5.3. Spectral details for sounds units present in the waveform



Waveform and spectogram of utterance

Phoneme	F1	F2	F3	F4

ai	552	1646	2380	3661
ae	646	1618	2320	3446
m	372	1623	2356	3412
f	629	1325	2203	3330
r	534	1272	2213	3435
o	320	646	2259	3318
m	282	1031	2345	3405
j	885	2524	3382	3991
a	539	1311	2201	3684
m	155	1677	2416	3411
S	250	1981	3197	3754
e	422	1924	3289	3644
d	264	120	3414	4150
p	403	2382	3404	4271
u	449	2146	3358	4117
r	292	1779	2150	4064

Table of first four formant frequencies of phonemes

[i] https://www.slideshare.net/Andriyanieka12/9-phonotactics-coarticulation

[ii] https://en.wikipedia.org/wiki/Phonation

[iii]

https://www.hchsmusic.com/uploads/2/0/4/7/20479636/music_theory_and_practice_textbook.pdf

[iv] https://ieeexplore.ieee.org/document/4648930

[v] https://person2.sol.lu.se/SidneyWood/praate/whatform.html

[vi] https://www.britannica.com/topic/pitch-speech

[vii] https://leader.pubs.asha.org/doi/10.1044/leader.FTR1.24022019.44