1. a) Corticulation refers to changes in speech articulation of current speech segment due to neighboring speech. In other words, before some characteristics close, the second characteristics start. For example, when you pronounce a world liked 'kitAb', before 'k' finishes, 'i' starts.

b) Phonation is the physical process behind sound production. Energy or air is given from lungs to glottis to where the vocal folds are. When air is pushed through glottis, the pressure drop in larynx causes vocal folds to vibrate, thus leading to voicing.

c) Fundamental frequency refers to frequency at which vocal folds vibrate. It refers to the frequency of (quasi) periodic structure of voiced speech signals.

d) Epochs are instants of significant excitation or glottal closure instant. They tell us how many times vocal folds vibrate and when does air enter the glottis, i.e., location of impulse trains

e) Formants are pitch harmonics. They are concentration of acoustic energy around a particular frequency. They are broad spectral maximum that comes from acoustic resonance of human vocal tract.

f) Pitch is rate of vocal folds vibration. In general children have more pitch than females, who have more pitch than males.

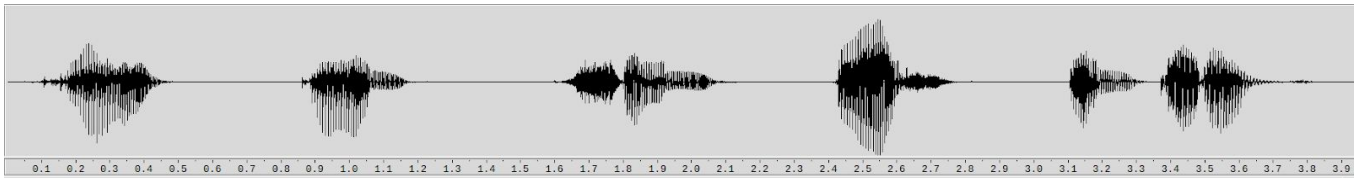
2. Female pitch is more than Male pitch as the rate at which vocal folds vibrate in females is more than that of males. Males are more likely to increase volume for emphasis while females are more likely to increase pitch for emphasis. Females have higher pitch range.

3. Speech is the most natural, intuitive and preferred means for humans to communicate. They are signals that have combinations of vowels, consonants, semi-vowels, etc. and variability of speech is in the form of languages, accents, etc. They exist in real-time and have varying duration, are quasi-stationary, have varying amplitude that makes it difficult to analyze as compared to other signals. Speech signals have meaning. Speech is legal sequence of legal sounds produced by humans. It is natural mode of communication

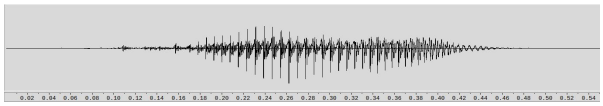
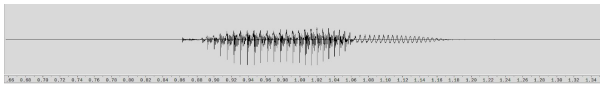
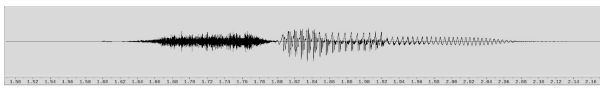


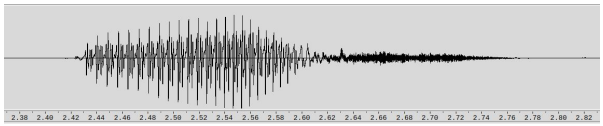
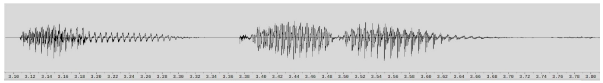
# Question 4

## (a) Waveform



I have added the audio (.wav) and annotated transcriptions (.lab) files in the zip for a better look at the waveform. We clearly see High-frequency regions with periodic nature are Voiced regions, high-frequency regions with non-periodic nature are unvoiced regions.

| **Word** | **Waveform for that word** |
|---|---|
| I |  |
| am |  |
| son |  |

| | |
|---|---|
| of |  |
| Antara |  |

## (b) Voiced, Unvoiced, Plosive and Silent regions

I have added the annotated transcriptions file in the zip. Voiced regions have high frequency and periodicity, unvoiced regions have high frequency and non-periodic nature, Silent regions have 0 amplitude, and there is only one Unvoiced-Plosive (3.3736142 to 3.3876560) region in this whole audio, which comes before /t/ in ‘antara’ as a burst which is clearly visible in the audio waveform and spectrogram. Plosive sounds happen because we completely stop air flow.

## (c) Acoustic-Phonetic description

I have used Hindi based IPA to find out the sound units.

| **Word** | **Description according to POA and MOA** |
|---|---|
| I (/A/, /e/) | Middle vowel followed by Front vowel |
| am (/E/, /m/) | Front vowel followed by Nasal |
| son (/s/, /a/, /n/) | Fricative followed by Middle vowel followed by Nasal |
| of (/O/, /ph/) | Back vowel followed by Unvoiced Aspirated Bilabial stop |
| Antara (/a/, /n/, /t/, /a/ , /r/, /A/) | Middle vowel followed by Nasal followed by Unvoiced Unaspirated Dental stop followed by Middle vowel followed by Semivowel followed by Middle Vowel |

## (d) Time-varying System Description

I have used Hindi based IPA to find out the sound units.

| **Sound unit** | **Description** |
|---|---|
| /A/ | Tongue hump is low and it is in central position of the vocal tract (VT) system, VT system is widely open |
| /e/ | Tongue hump is medium and it is in front position of the VT system, VT system is moderately open |
| /E/ | Tongue hump is medium and it is in front position of the VT system, VT system is moderately open |
| /m/ | Opening of Nasal cavity and closure at lips |
| /s/ | Narrow constriction at Dental position |
| /a/ | Tongue hump is low and it is in central position of the vocal tract (VT) system, VT system is widely open |
| /n/ | Opening of Nasal cavity and closure at Dental |
| /O/ | Tongue hump is medium and it is in back position of the VT system, VT system is moderately open and cylindrical in shape |
| /ph/ | Complete Closure at lips |
| /a/ | Tongue hump is low and it is in central position of the vocal tract (VT) system, VT system is widely open |
| /n/ | Opening of Nasal cavity and closure at Dental |
| /t/ | Complete closure at Dental position |
| /a/ | Tongue hump is low and it is in central position of the vocal tract (VT) system, VT system is widely open |
| /r/ | Partial closure of VT with tongue tip at alveolar ridge / Narrow opening at alveolar ridge |

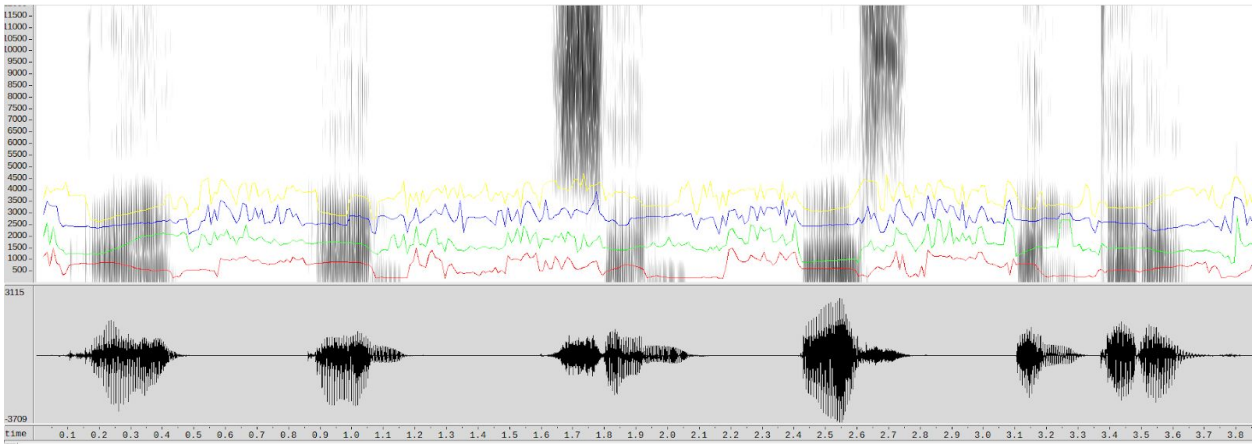
| | |
|---|---|
| /A/ | Tongue hump is low and it is in central position of the vocal tract (VT) system, VT system is widely open |

## (e) Spectral details

I have used Hindi based IPA to find out the sound units.

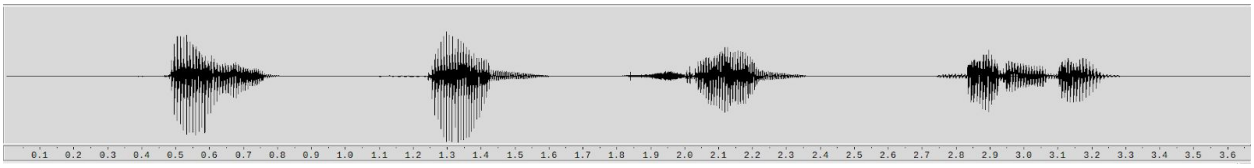
| **Sound unit** | **Spectrogram Details** |
|---|---|
| /A/ | Regular Formant structure (520, 1190, 2390), pitch harmonics |
| /e/ | Regular Formant structure (530, 1840, 2480), pitch harmonics |
| /E/ | Regular Formant structure (660, 1720, 2410), pitch harmonics |
| /m/ | Concentration of low frequency energy and mid range frequencies with no prominent peaks, pitch harmonics |
| /s/ | Concentration of high frequency energy |
| /a/ | Regular Formant structure (730, 1090, 2440), pitch harmonics |
| /n/ | Concentration of low frequency energy and mid range frequencies with no prominent peaks, pitch harmonics |
| /O/ | Regular Formant structure (600, 900, 2300), pitch harmonics |
| /ph/ | Concentration of high frequency energy |
| /a/ | Regular Formant structure (730, 1090, 2440), pitch harmonics |
| /n/ | Concentration of low frequency energy and mid range frequencies with no prominent peaks, pitch harmonics |

| | |
|---|---|
| /t/ | Concentration of high frequency energy |
| /a/ | Regular Formant structure (730, 1090, 2440), pitch harmonics |
| /r/ | Regular Formant structure (490, 1400, 2510), pitch harmonics |
| /A/ | Regular Formant structure (520, 1190, 2390), pitch harmonics |

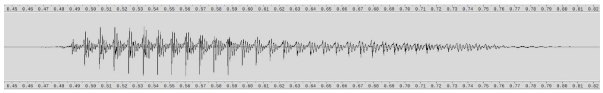
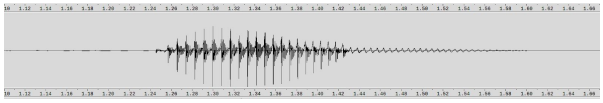
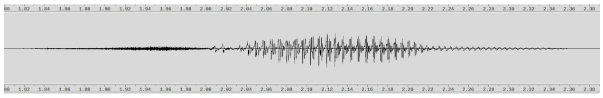
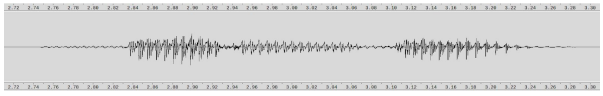


# Question 5

## (a) Waveform



I have added the audio (.wav) and annotated transcriptions (.lab) files in the zip for a better look at the waveform. We clearly see High-frequency regions with periodic nature are Voiced regions, high-frequency regions with non-periodic nature are unvoiced regions.

| Word | Waveform for that word |
|---|---|
| I |  |
| am |  |
| from |  |
| baroda |  |

## (b) Voiced, Unvoiced, Plosive and Silent regions

I have added the annotated transcriptions file in the zip. Voiced regions have high frequency and periodicity, unvoiced regions have high frequency and non-periodic nature, Silent regions have 0 amplitude, and there are only two Voiced-Plosive regions in this whole audio, which come before /b/ (2.7475140 to 2.8342976) in ‘baroda’ and /d/ (3.0875278 to 3.1096086) in ‘baroda’ as a burst which is clearly visible in the audio waveform and spectrogram. Plosive sounds happen because we completely stop air flow.

## (c) Acoustic-Phonetic description

I have used Hindi based IPA to find out the sound units.

| Word | Description according to POA and MOA |
|---|---|
| I (/A/, /e/) | Middle vowel followed by Front vowel |
| am (/E/, /m/) | Front vowel followed by Nasal |
| from (/ph/, /r/, /O/, /m/) | Unvoiced Aspirated Bilabial stop followed by Semivowel followed by Back vowel followed by Nasal |
| baroda (/b/, /a/, /r/, /o/, /D/, /A/) | Voiced Unaspirated Bilabial stop followed by |

| | |
|---|---|
| | Middle vowel followed by Semivowel followed by Back vowel followed by Voiced Unaspirated Alveolar stop followed by Middle vowel |

## (d) Time-varying description

I have used Hindi based IPA to find out the sound units.

| **Sound unit** | **Description** |
|---|---|
| /A/ | Tongue hump is low and it is in central position of the vocal tract (VT) system, VT system is widely open |
| /e/ | Tongue hump is medium and it is in front position of the VT system, VT system is moderately open |
| /E/ | Tongue hump is medium and it is in front position of the VT system, VT system is moderately open |
| /m/ | Opening of Nasal cavity and closure at lips |
| /ph/ | Complete closure at lips |
| /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/ | Opening of Nasal cavity and closure at lips |
| /b/ | Complete closure at lips |
| /a/ | Tongue hump is low and it is in central position of the vocal tract (VT) system, VT system is widely open |
| /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 |
| /D/ | Complete closure at Alveolar |
| /A/ | Tongue hump is low and it is in central position of the vocal tract (VT) system, VT system is widely open |

## (e) Spectral Details

I have used Hindi based IPA to find out the sound units.

| **Sound unit** | **Spectrogram Details** |
|---|---|
| /A/ | Regular Formant structure (520, 1190, 2390), pitch harmonics |
| /e/ | Regular Formant structure (530, 1840, 2480), pitch harmonics |
| /E/ | Regular Formant structure (660, 1720, 2410), pitch harmonics |
| /m/ | Concentration of low frequency energy and mid range frequencies with no prominent peaks, pitch harmonics |
| /ph/ | Concentration of high frequency energy |
| /r/ | Regular Formant structure (490, 1400, 2510), pitch harmonics |
| /O/ | Regular Formant structure (600, 900, 2300), pitch harmonics |
| /m/ | Concentration of low frequency energy and mid range frequencies with no prominent peaks, pitch harmonics |
| /b/ | Concentration of low frequency energy |
| /a/ | Regular Formant structure (730, 1090, 2440), pitch harmonics |

| | |
|---|---|
| /r/ | Regular Formant structure (490, 1400, 2510), pitch harmonics |
| /o/ | Regular Formant structure (570, 840, 2410), pitch harmonics |
| /D/ | Concentration of low frequency energy |
| /A/ | Regular Formant structure (520, 1190, 2390), pitch harmonics |