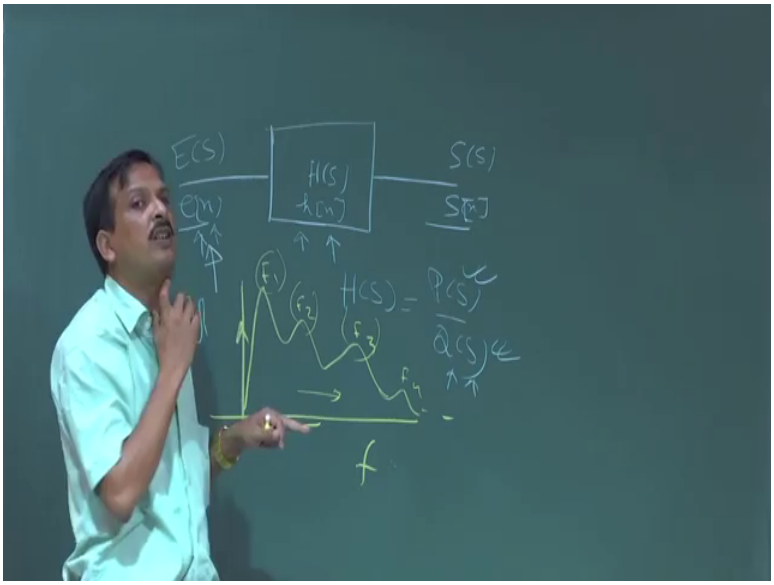
Digital Speech Processing
Prof.S.K.Das Mandal
Centre for Educational Technology
Indian Institute of Technology Kharagpur

Lecture – 06
Place and Mannerat Articulation

So let us start. That last class we have discussed about that engineering models of the speech production systems, we said that this bugle tube modeled as a filters, and glottis vibration can be model as a source. So, we said the source filter model.

(Refer Slide Time: 00:36)



We said let us this is a filter that is h let us H s and there is a source called E S or I can say e n and h n in digital system and the output is the speech either s s or small s n. So now, if I say that this is a filter. So, if the source is exist. So, all voice sound the source is modified by the filter. And source has only property the fundamental frequency contour and it is nothing but a impulse. So, if you see here in the slide. So, it is nothing but a impulse. So, excitation generation nothing but a impulse. And that excitation is coming through a filter which is modeled as a linear filter, why we consider linear filter? And how this model is mathematically described we will discuss later.

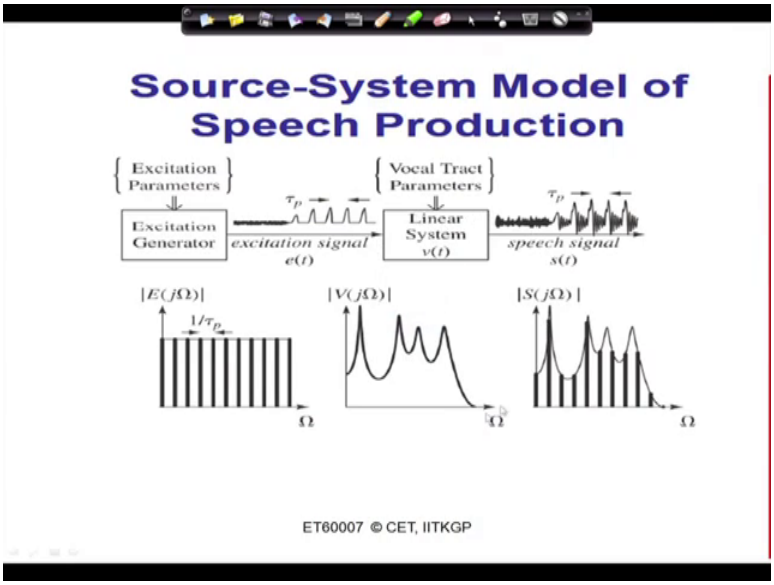
So, this is nothing but a linear filter and convolution of e n and h n produce speech. So, in time domain if it is convolution frequency domain it is nothing but a multiplication. So now, how do I classify the voice sound. So, voice sound is classified based on 2

property one is the fundamental frequency of the excitation which is constant for throughout the all voice sound. Now different model of filter can produce different voice sound. So, if I say as you know H s is nothing but H s is a filter model that is nothing but sum p s by q s as you know in any model is nothing but a pole and 0 kind combination of pole and 0.

So, root of this thing is provide me 0 and root of this thing is provide me the poles. So now, you know as for that control systems if you have the background of electronics, that all pole related to the formant frequency or resonance frequency later on resonance frequency. So, for every pole there is a resonance frequency. So, if q s is a first order system first order then it is only single pole. So, I can the single resonance frequency if it is a multi order systems n th order system say n number of resonance frequency will be there.

So, if I see if I look at the frequency response of this linear system it will be look like this.

(Refer Slide Time: 03:09)

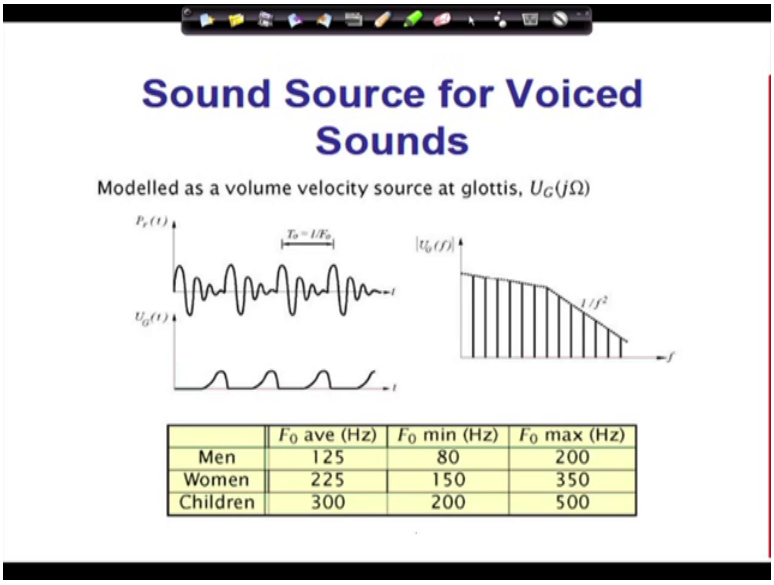


If you see look like this. So, every peak is called formant. So, all voice sound are characterized by those formant. So, this is first formant, second formant, third formant, and forth formant. So, on the formant will be there. So, if it is frequency domain if this axis is frequency and this axis is the amplitude of that frequency and omega is the analog

frequency it is a analog frequency, then it is nothing but a multiplication. So, this frequency response of the unit impulse will be modified by the filter frequency response.

Details of how what should be the formant structures will come on clear the (Refer Time: 03:53) model this speech production system. So, all the voice sounds are characterized by the formant frequency for different voice sound as different kind of formant frequency. That is why they lower they sound like different sound. So, then if just this we have already explained that why that woman had spotted f 0 and for one frequency, then I have go to this directly.

(Refer Slide Time: 04:19)



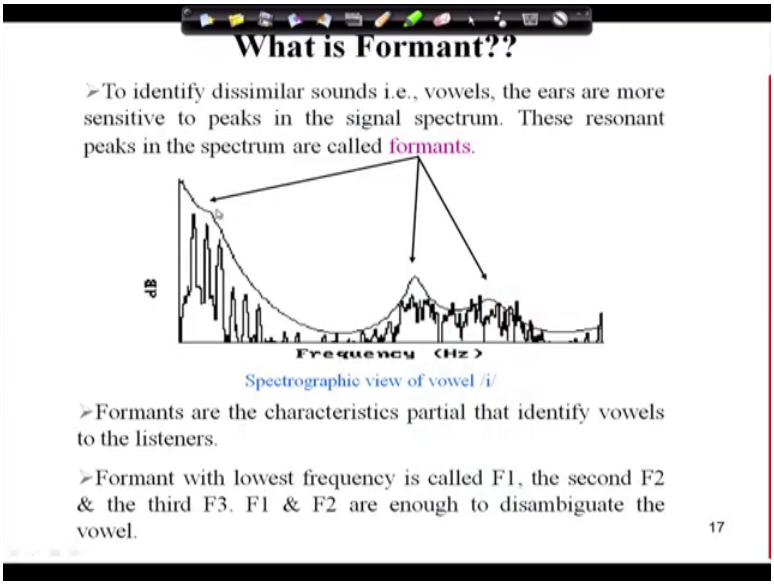
	F_0 ave (Hz)	F_0 min (Hz)	F_0 max (Hz)
Men	125	80	200
Women	225	150	350
Children	300	200	500

So, there is some average f 0 value if you see the average f 0 value for man is f 0 average one 50 hertz and minima is 80 and maxima is 200.

If you say the woman it is average is 225 almost above 200 hertz and children it is above 300 hertz kind of mean frequency. So, if you that you are experience that if you see if you know the singing. So, base frequency that somebody has be sound be flat those kind of scale you are using that is nothing but a match with your average f 0 frequency. Now formant frequency also I will come on what should be the human (Refer Time: 04:59) minimum formant frequency and all those variation I will come on when we discuss about the uniformity modeling of the speech production system.

So, any voice sound I can say that it is characterized By the formant frequency or resonant frequency.

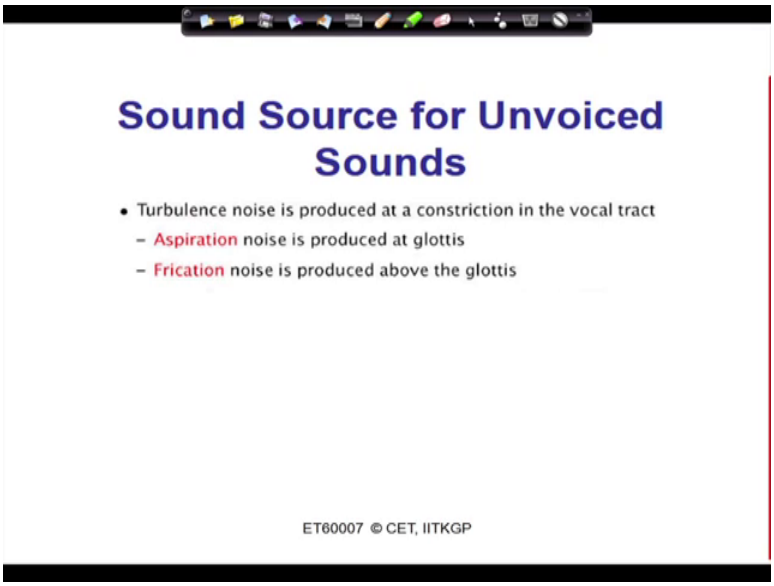
(Refer Slide Time: 05:11)



So, formant frequency is nothing but the resonant frequency. So, if I say any voice sound vocalic sound if this is frequency and this is amplitude it will look like this. So, this is first formant f 1 this is f 2 this is f 3 so on f 4. So, on going on once it is going up the amplitude will be going down that will discuss why this is going down. Now if the voice sound is characterized by the formant frequency. So, formant frequency depends on the structure of the this vocal tract.

How this formant frequency will change depending on the h s. So, H s will be modified based on the structure of the whole tube. So, different structure can produce the different kind of voice sound and this has the different kind of formant frequency. Now if I related to the vowel sound if I spitted to the sound which is unvoiced sound we can say, that there is a difference between the unvoiced and silence.

(Refer Slide Time: 06:12)



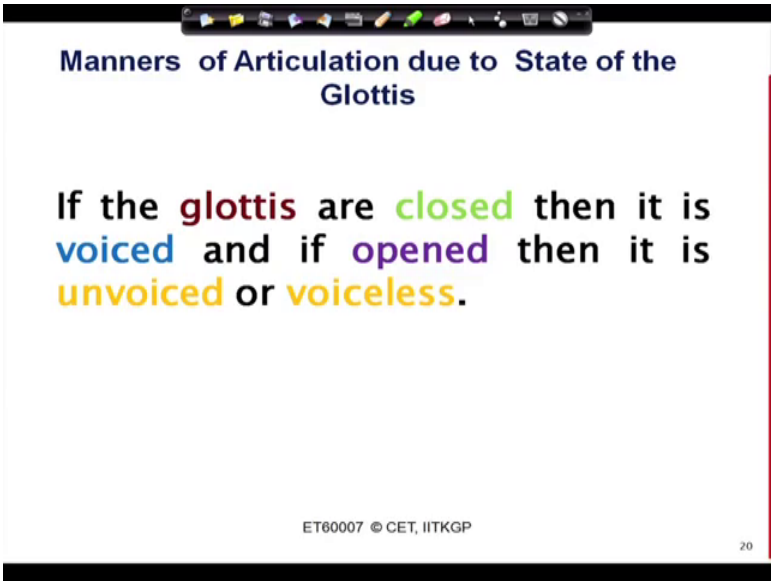
So, aspiration friction are called unvoiced sound unvoiced means vocal call is not vibrate, but still there is a noise kind of sound. The source is aspiration and friction. How the aspiration is produced? Suppose you have a tube. You create a narrow constriction in the tube now what will happen the air pressure lungs produce an air pressure and create a narrow constriction using the tongue, then what will happen the force air try to find out a path and there is a friction kind of noise sound will be generated. So, those kind of sound is called aspiration and friction sound will be generated. So, details I will come in that manner of articulation on that place of articulation discussion.

Now, forget about the signal processing engineering model part. Human vocal tract produce different kind of sound. How do I classify those sound? How do I describe? Forget about the classification, I want to describe the different sound produced by the human being for speech communication. So, I can I can say that whatever the sound is coming out from the mouth has 2 property. One is the tongue position how the tongue is moved and to produce that sound that is and place where that is placed place in the position means where the sound is produced, that is called place of articulation. So, during the articulation the airstream through the vocal tract must be obstructed in some way So that I can produce the sound. The place where the obstruction is takes place is called the place of articulation.

For example suppose I produce want to produce the vowel. So, I can see the vocal cords is vibrating and hip of the tongue is raised. So, it that the place where the obstruction is happened is called place of articulation and manner, manner means how I produce that sound; that means, when I produce the sound the airflow whether the air is fully flow, whether the airflow is stucked, whether the airflow is passes through a narrow cavity. So, all kind of airflow mechanism are produced the manner of articulation.

Now, you can see the manner of articulation has also 2 major difference one is called either this vocal if I say the source is also part of manner of articulation. Then either source can present or source can absent. So, either glottis can vibrate or glottis can absent.

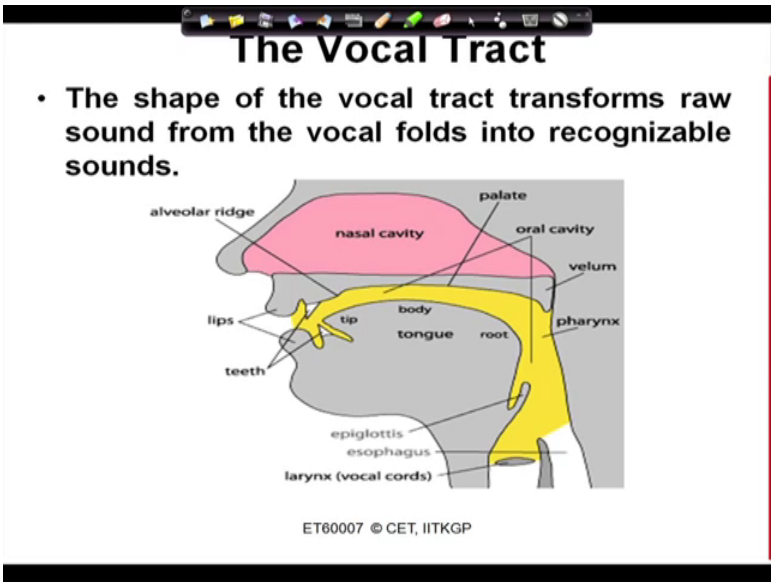
(Refer Slide Time: 09:07)



So, glottis has 2 manners, when the glottis is closed. Then the voice sound is produced while it is totally open then the unvoiced sound is produced. So, anything can be voiced or unvoiced. So, 2 part based on the glottis it has 2 part either it can be voiced or it can be unvoiced.

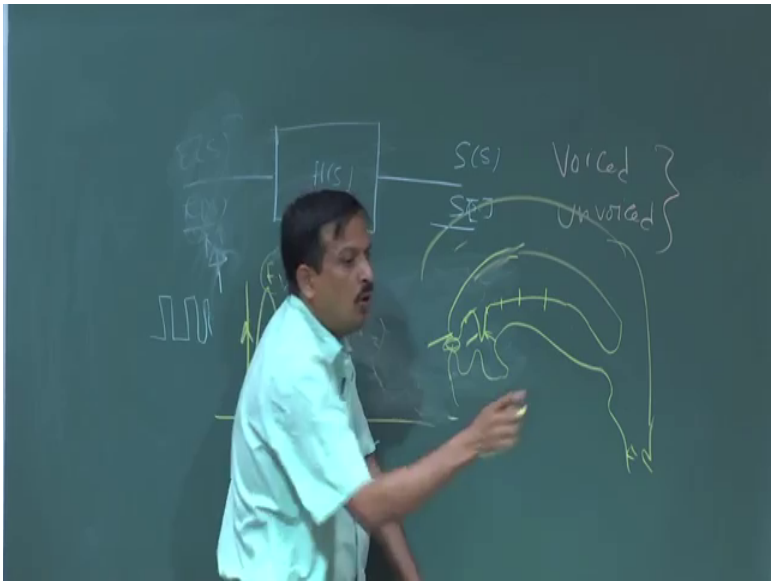
Similarly, place of articulation has produced different sound. How do I describe the place? So, different place are if you see the bilabial, bilabial, labiodental, dental, alveolar, post alveolar, retroflex, palatal, velar uvular pharyngeal and glottal. So, if I see the upper track of the human speech production system which is there in here.

(Refer Slide Time: 10:06)



If you see the different upper cavity can be part in different place. Bilabial then dental then can be labiodental means leap and dental can be there. Or I can draw that this will be very easy if I draw it let us I draw a human speech production system I do not find that duster. So, if I see it let schematic diagram of the human speech production system I want to draw it. So, if I see let us I start from labiodental, bilabial, labial, then which can goes to the teeth, teeth Palatal and then goes velum.

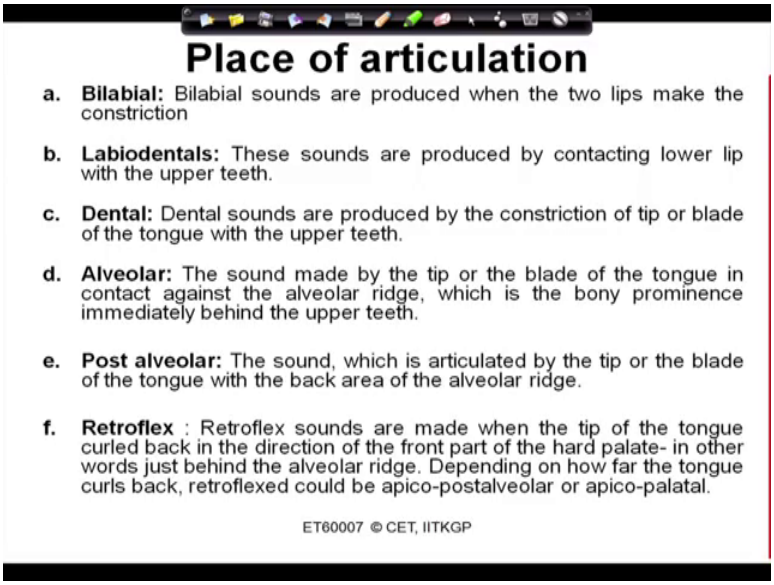
(Refer Slide Time: 10:47)



So, there will be a velum cavity in velum then. So, lower cavity is bilabial teeth and then tip of the tongue and then come in the voice. So, here either glottis can closed or open if it is closed then I call voice sound if it is open then I call unvoiced sound.

Now, when if the tip of the tongue touches the teeth then I call it is a dental sound. Or blocks can be happened in here constriction can be happened in labial position also. So, this is called labial lips dental then you can you can say the alveolar post alveolar alveolar post alveolar baler. So, all kind of place I can show you in different schematic diagram one kind of place for different kind of sound production systems. Now So, depending on the place the sounds are classified in different labial bilabial when the sound are produced 2 lips.

(Refer Slide Time: 12:01)

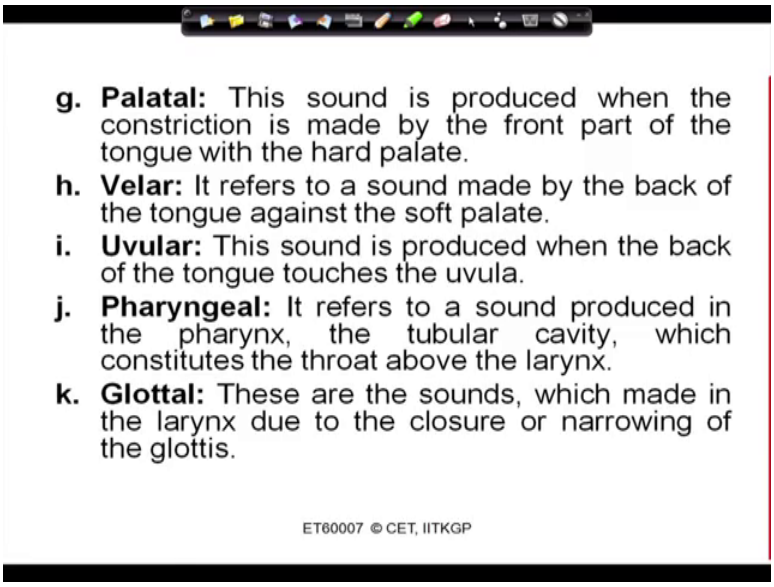


Labiodental when the sound are produced by a constriction of a lower lips with upper teeth, upper teeth. Dental when the sound is produced by the constriction of tip of the blade of the tongue with the upper teeth. [FL], labeal bilabial pah bilabial 2 lips are close together to produce the pah. Bilabial then [FL] if you see the tip of the tongue touches the dental position. So, it is stopped dental then alveolar the sound made by tip or the blade of the tongue within the contract against alveolar region alveolar region after the teeth the alveolar region then we can say alveolar sound.

Then after alveolar there is a post alveolar, then after post alveolar retroflex sometime which is not a place of articulation it kind of manner of articulation find the tip of the

tongue curved back and touches the upper cavity in post alveolar region we call retroflex sound. So, tip of the tongue curve back [FL] the if the tip of the tongue is curve back and touches the upper cavity. That is why it is called retroflex sound. Similarly if I say there is a palatal sound.

(Refer Slide Time: 13:25)

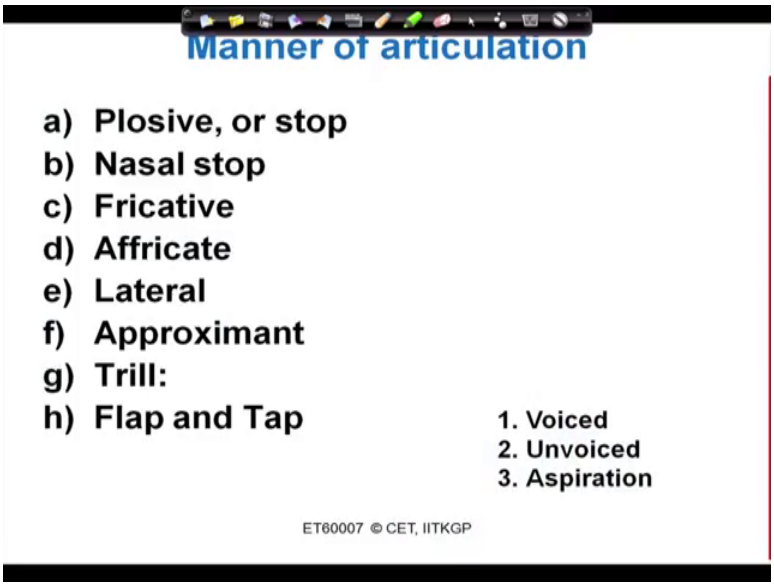


When the constriction made in the front part of the tongue with the hard palate hard palate means middle of the cavity hard palatal region, when the palatal region is come then it is called palatal sound.

Then after the palatal there is a velar sound when the back on that now they see that after here tongue cannot go back. So, the back of the tongue can also move when the back of the tongue raise and touches the velar position to produce the sound we call velar sound. Ka if it is a ka then back of the tongue is raised and touches the velar position then we call velar sound. Similarly uvular pharyngeal and glottal sound when the sound are coming from the glottal region then we call glottal sound ok.

So, depending on the place of the constriction sounds are described and that description is called place of articulation because the constriction is made different position. So, tube structure will be different different kind of frequency will be produced by the tube and that is why the sound is different. Now if I go to the manner of articulation as I described.

(Refer Slide Time: 14:39)



The manner of articulation depending on the source glottis has 2 manner. Either it is voiced or unvoiced or aspiration is there. So, voiced unvoiced and aspiration may be there. So, these 3 are the source and there is a another play manner of articulation which is called plosive or stop nasal stop fricative nasal stops are plosive fricative lateral approximant trill trill flap or tap.

What do what do you mean by the plosive sound? Now if you see if on the tip of the tongue close the upper cavity. If I say pa let us 2 lips are close together. Pa, what is happening? In the back cavity the after before that closure. So, if the closure is here then the sounds the air pressure in here is building up. Because the cavity is closed pa and suddenly this is released and sound burst is coming out, sound burst is coming out. Pa that is why it is called plosion or plosive pa plosion.

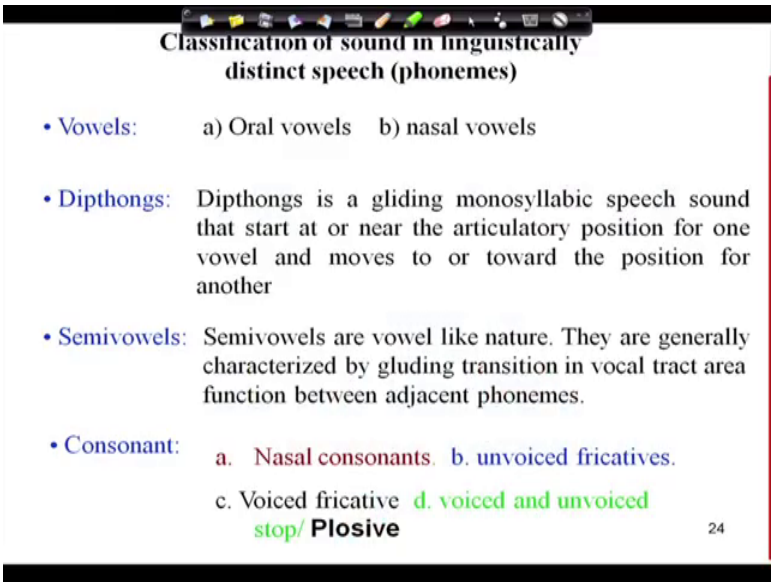
So, if it if it a sound is if the consonant sound this is called consonant sounds if the consonant sound is followed by a vowel pa then what is happening now back of the cavities sound pressure is increases it creates a plosion that is why it is called plosive. Some time suppose if I say [FL], then what is happening there is a vowel sound is producing by the vocal track. And suddenly it wanted to produce a consonant you want to produce a constriction. So, it is want to stop. So, then it is called stop sound. So, plosive or stops it can be either plosive consonant or it can be a stop consonant. So, if it is a followed by a vowel then it is a plosion will be happened.

If the if it is after vowel there is a sound ka then it called this ka is stopped it is not plosive. So, this is plosive and stop then nasal stop, sometime if you say na ma bilabial nasal ma ma. So, if the velum is open that is why it is called nasal cavity is attached. So, this is nasal sound. Then there is a fricative affricate lateral approximant fricative if you say sh sh. So, it is fricative sound affricative sound has a 2 property affricate. First part is plosion second part is friction, that is why it is called affricate sound I will show you. I mean animation for that affricate means first part is plosion second part is friction that is why it is called affricate sound. Then lateral sound if you see the tip of the tongue touches upper cavity and sound is coming the 2 side of the tongue uh la.

So, side of the 2 side are open and the sound is or air is coming out from the 2 side of the tongue then it is called bilateral lateral bilateral sound if it is single side it is coming unilateral sound. So, it is lateral sound then approximant like that we have bengali has 2 approximants approximants are look like a vowel and if I say [FL] wah and approximant or glide you can say then trill rh ra. So, tip of the tongue touches the upper cavity and it vibrate trill rah rah if you say tip of the tongue touches the upper cavity and it vibrate rh ra. So, it is trill.

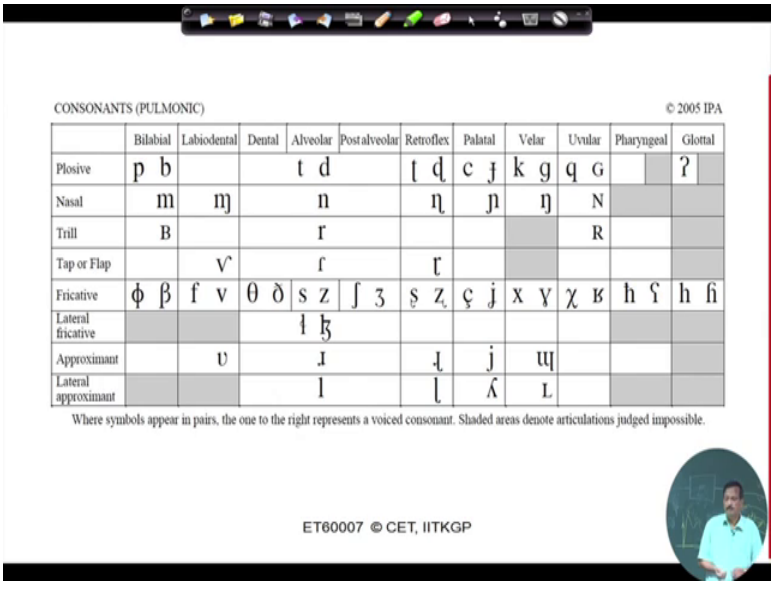
Then tap or flap if your back tip of the tongue back of the tip of the tongue touches the upper cavity it sing like a flap then it is called flap or tap la la la. So, back of the tongue touches the upper cavity and produce that sound that is why it is called tap or flap. Now I go to the details discussion on that.

(Refer Slide Time: 19:14).



Then if you see sound can be classified 2 things one is called oral vowel and nasal vowel vowel sound is described in the tongue movement, I will come later on then dipthong semivowel and consonant. So, those are the you can go through the slides those you know you can go through that.

(Refer Slide Time: 19:33).



If you see how the sound is sound is unique it does not require the language if I say ka for Hindi ka for bengali ka for English all is ka sound ka how the sound sequence is produced then the language rule is come out. So, if I want to describe the each and

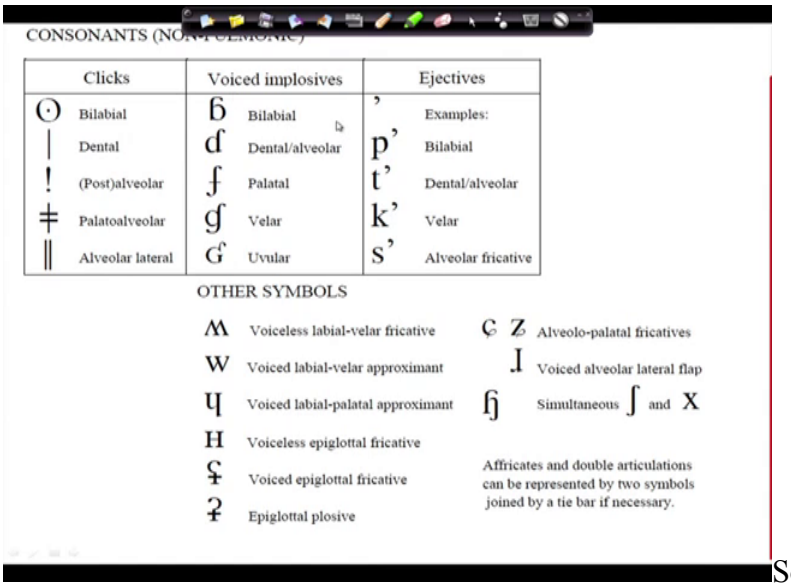
individual sound which is produced by the human being, and try to find out a symbol for each individual sound. So, depending on the place and manner of articulation what I do I try to find out or try to assign a symbol to each particular sound ok.

So, what I want depending on the place and manner of articulation I try to assign a symbol to each and every sound. So, maybe some language has some group of sound some language may not have some group of sound. So, that is, but both sound which is exist in the language may be exist in other language also. So, that is exist.

So, sound which is distinct sound which is produced by a different constriction and different kind of mechanism by the human vocal tract has to be assign a particular symbol. That is called IPA international phonetic alphabet. So, that is why it is called phonetic alphabet I can assign a symbol to each and individual sound depending on the place and manner of articulation now sound. Like a consonant I already discussed we have a consonant we have a vowel. We have a dipthong we have an approximant those all has to be assign some symbol. So, depending on the place and manner of articulation we assign those symbol.

If you see this chart there is another call if you see the consonant which is written pulmonic. It can be pulmonic I think a non pulmonic also I have symbol.

(Refer Slide Time: 21:47)

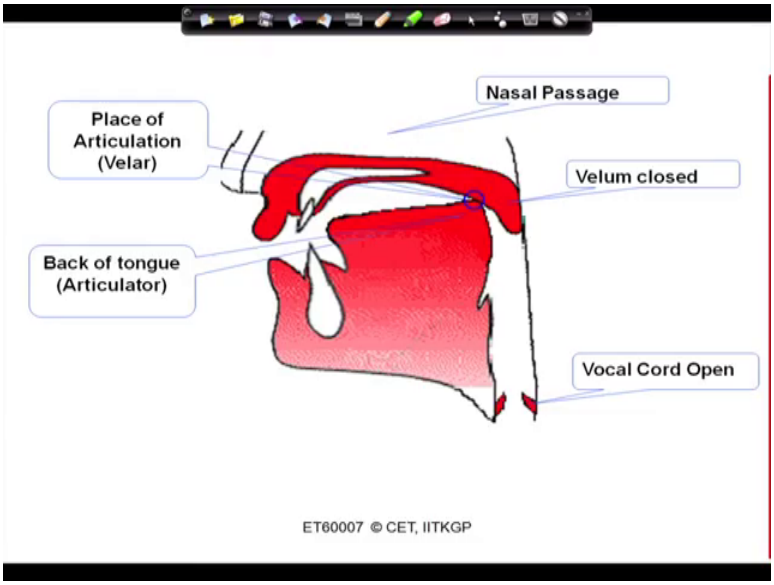


So, most of the cases the sound which is produced during the air flow is from lungs to mouth is called pulmonic sound. If the during the production of the sound if the air flow in this direction, it is from

lungs to mouth then we call this a pulmonic sound. Non pulmonic if it is a air flow is reversed then we call it is a non pulmonic sound. So, any pulmonic consonant in that case the air is coming from lungs to mouth. So, that is the pulmonic sound.

Now, pulmonic sound depending on the place and manner of articulation they have assigned different symbol. If you see this side is the place of articulation, and this side is the manner of articulation if you see the chart this row wise. Row wise is the place of articulation column wise is the manner of articulation. Now if you see plosive or stop. So, if we see the plosive or stop bilabial plosive. So, if I want to produce a bilabial plosive; that means, I can draw it or I can go to here. So, you can get that animation let us like this, let us go to this pictures.

(Refer Slide Time: 23:27).



So, if I want to produce a bilabial sound bilabial sound is here I yes, the bilabial sound I want to produce a bilabial sound. What is that 2 tip lip lips are closed lips are closed here. So, place of articulation is bilabial. Now lips may be closed or the state of the glottis either glottis can open or glottis can closed if glottis is vocal cords or vocal cords or glottis a vocal cord is open then I call unvoiced sound if it is vocal cord is closed then I call voice sound. So, any bilabial plosive has 2 varieties. 2 manner of articulation depending on the source vocal cords either it can be voice or it can be unvoice.

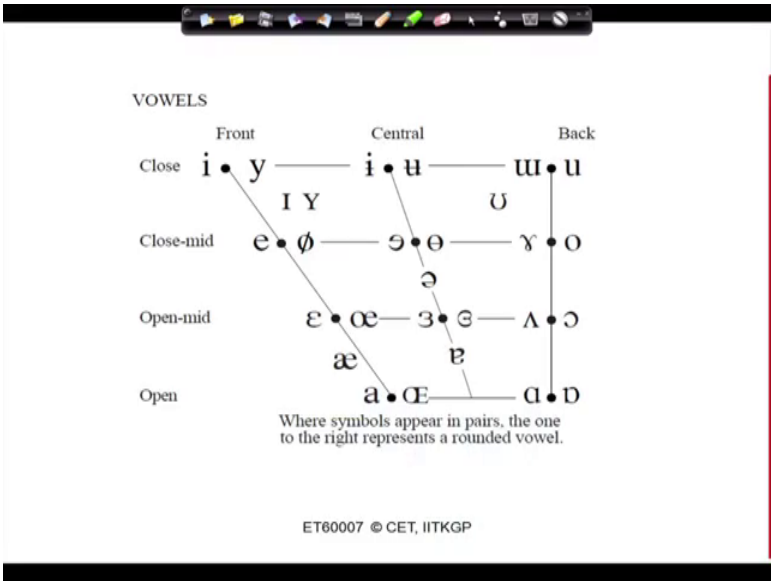
So, if you see the chart it is written p for unvoiced b for next one is for voiced. So, it is 2 place unvoiced and voiced. Similarly if you see toe or ta voiced unvoiced voiced

unvoiced voiced unvoiced. So, maybe some language does not have this sound, but other language may have this sound. So, in if you see that consonant symbols in the IPA chart it depending on the voiced unvoiced symbols are different and depending on the manner. If you see the different symbol bilabial nasal it is a smaller m bilabial trill if it is I will be produce the capital b. Similarly we have produced bilabial till ra ra it is on the alveolar position.

Then tap tap or flap fricative if I say I will produce dental sound sa sa what we next produce palatal sound sa or post alveolar palatal or postal alveolar sound sha sha. Lateral fricative approximant lateral approximant. If it is see that only voiced part unvoiced part is not there that is why it is a single symbol, if you see single symbol. So, depending on the voice that is voice or unvoiced 2 sound and depending on the manner in here different sounds are there. We will describe later on we will describe details.

Similarly, non pulmonic I am not discussing non pulmonic and there is a other kind of symbols are there if you see the IPA chart download it from IPA site there are other kinds of symbols are also there.

(Refer Slide Time: 26:16).

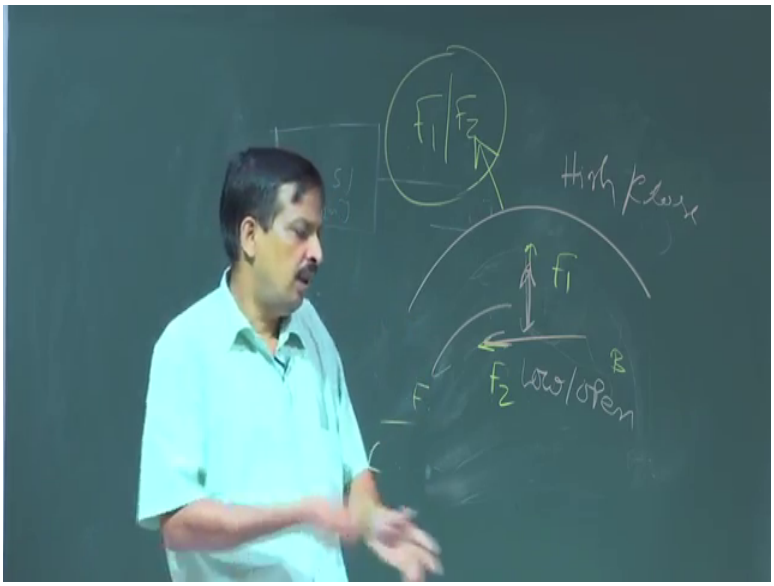


Now I go for the vowels if you see the vowel. How the vowel sound is specified in case of production of the vowel the constriction is not take place. The constriction is not there. What is there the tongue movement. If you see tongue can move front and back, if you

see your tongue can move either back or front back of the tongue can raise or front of the tongue can come out front and front front of the tongue can raise ok.

So, depending on the tongue tongue front and backward movement and tongue height; that means, how the tongue is moving your upper palate is fixed.

(Refer Slide Time: 27:01).

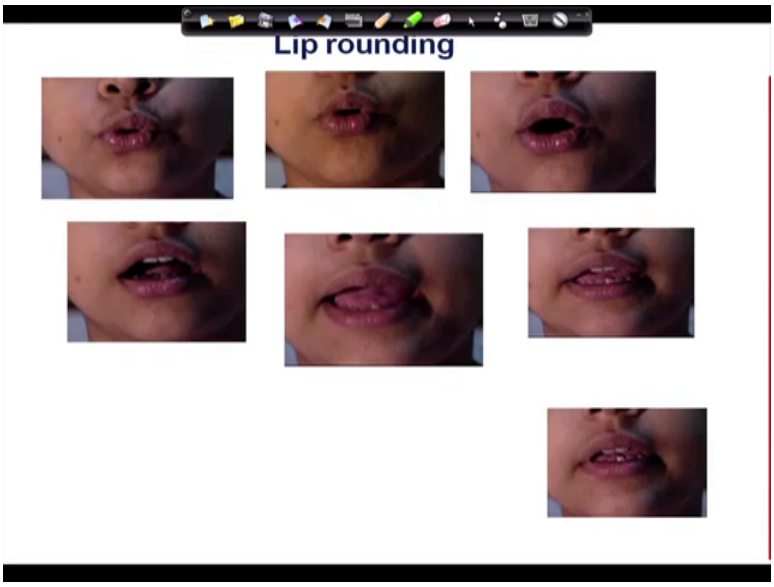


Lower palate is moving tongue is moving. So, tongue height and tongue movement. Frontness and backness and tongue height. So, depending on the front or back where the vowel is produced and depending on the tongue heights all the vowels are described. So, tongue height means if it is high if toungue is raised high then it is closed almost closed that is why it is also called, close, close tongue closer.

Or if it is low tongue height is low then it is also called open mouth. Sometimes we find some literature has closed or high open or low. And depending on the tongue height and tongue forwardness different symbols are there for different vowel sound. So, vowels are classified or vowels are symbolified based under their vowel position. Production position instead of constriction here you call power it will be in the articulator position vowels are described. Either it can be high low or it can be front and back. And there is a another things also it is not consider, but yes that is there.

That if you see some vowels are required lip movement also. If I either say you can say the lip rounding you can it is called lip rounding, sometimes we round the lip and sometimes we do not Round the lips.

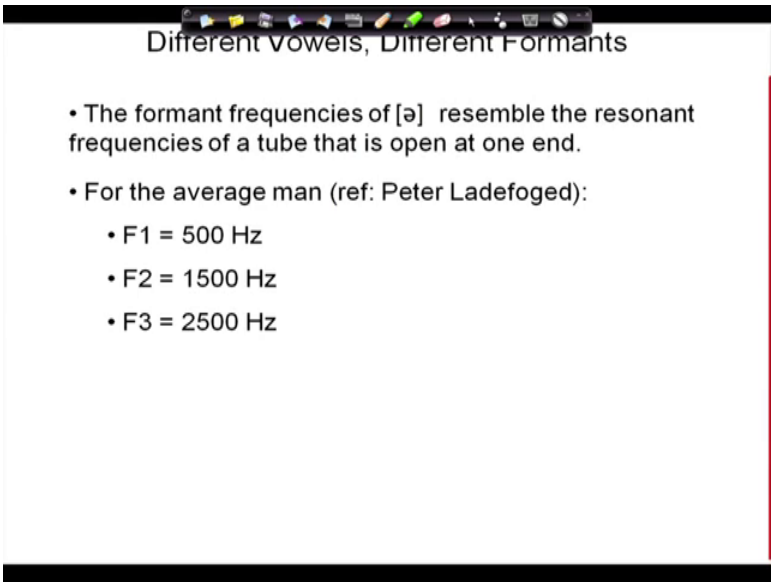
(Refer Slide Time: 28:42).



Here I am showing you the picture for different vowel production. If you say woo lip are rounded. So, lip rounding and the tongue position are describe the vowel sound. I am not playing the vowel there is a different kind of vowel sound you can if it is here you can play the sound. Now as a engineer we do not know what is the vowel tongue height tongue kind of things how do you describe. So, can I find out some parameter which will related to the tongue height and which is related to the tongue forwardness and backwardness?

Yes it is possible. If you see the formant vowels are described based on the formants, and the formants it is f 1 is nearly 500, 500 hertz.

(Refer Slide Time: 29:29).



F 2 is 1.5 kilo hertz f 3 is 2.5 kilo hertz, I will come, how it is 500 hertz like that. Now there is a relation that depending on the tongue height is related to the first formant. Tongue front and backness is related to the second formant and lip rounding is related to the third formant. So, depending a seeing the, So, this is a correlation there is a close relation between the f 1 and tongue height and it is correlation between the f 2 n tongue lowerness the tongue movement is the front and back. So, this is front and back ok.

So, either I can see the f 1 and f 2 I can classify whether this is a high vowel or low vowel. I can classify whether is a front vowel and back vowel. So, the same place same articulatory place can be simulated or can be model or can be placed in a formant measurement plate plain. So, I can back that (Refer Time: 30:52) I can find out the frequency and I can plot them f 1 and f 2 plane which generally call f 1 and f 2 plane. Where the vowel sounds are classified. So now, if I say the f 1 and f 2 are classify the vowel sound different kind of vowel sound. So, I can say the vowel if I the content or whether I said or oh it depends on f 1 and f 2. And f 3 some part of f 3 where the lip rounding is required or not that is f 3

So, f 1 f 2 and f 3 is describing the vowel which vowel it is. So, you can say the content now if you say the formant frequency f 1 f 2 describing the vowel different vowels then why we require the higher formants. Higher formants maybe communicate the speaker style or other information also. So, f 1 and f 2 mainly responsible for which vowel I have

produced. If I change the f 1 and f 2 the sound the may be a becomes o or u becomes o, if I change the vowel if I change the formant. So, f 1 and f 2 describe the vowel quality or you say the vowel which vowel I have produced. And f 1 and f 2 related to the tongue height and tongue backness or front and back is f 2 and tongue height and whether it is closed or open or low or high that depends on the f 1 and f 3 is the lip rounding. So, next class I will take away to that animation and find out that different place of articulation.

Thank you.