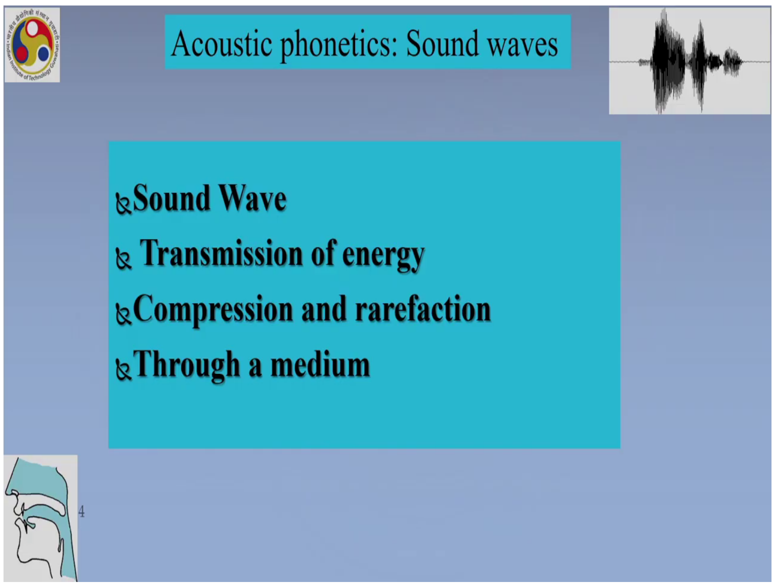
Phonetics and Phonology: A broad overview
Professor Shakuntala Mahanta
Department of Humanities and Social Sciences
Indian Institute of Technology, Guwahati
Lecture 09
Acoustic Phonetics 2

Welcome to the second unit of Acoustic Phonetics. This is the continuation of the NPTEL MOOCs course on Phonetics and Phonology: A broad overview. So again in this part we will talk about acoustic phonetics and we will also look at how the acoustics of vowels can be understood in acoustic phonetics.

(Refer Slide Time: 01:04)

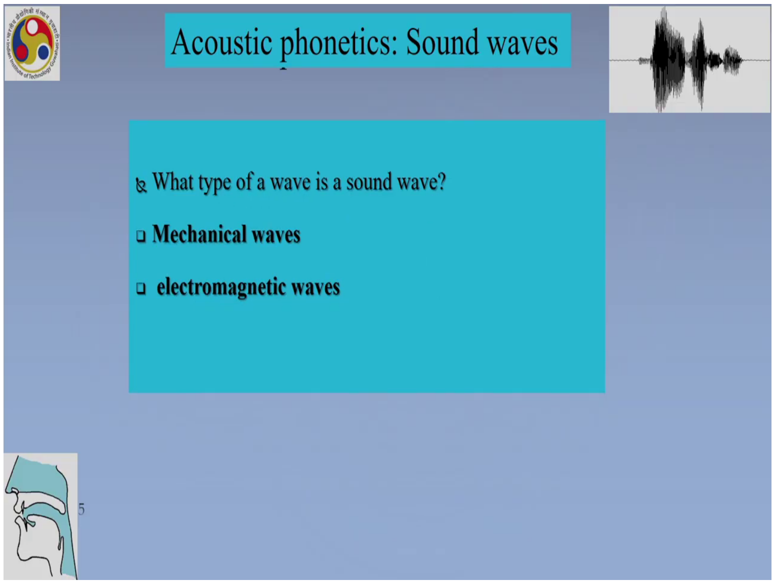
So let us briefly recap what we did in the last class. So we studied how sound is variation in air pressure which is detectable by the human ear.

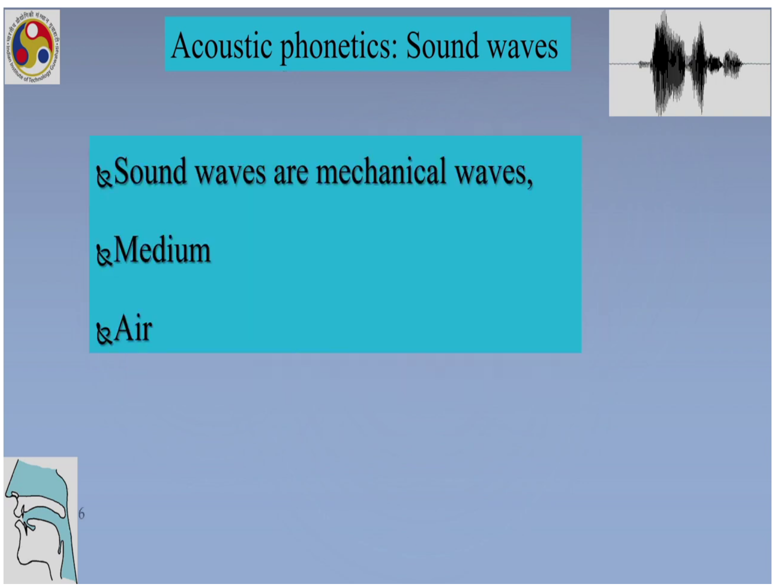
(Refer Slide Time: 1:18)



And we also saw how sound wave is actually transmission of energy, compression and rarefaction through a medium which is air in the case of linguistic sounds.

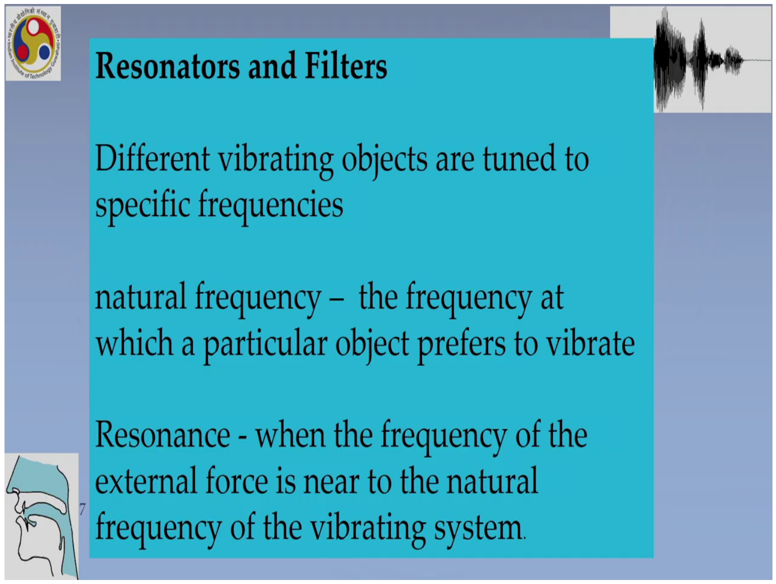
(Refer Slide Time: 1:30)





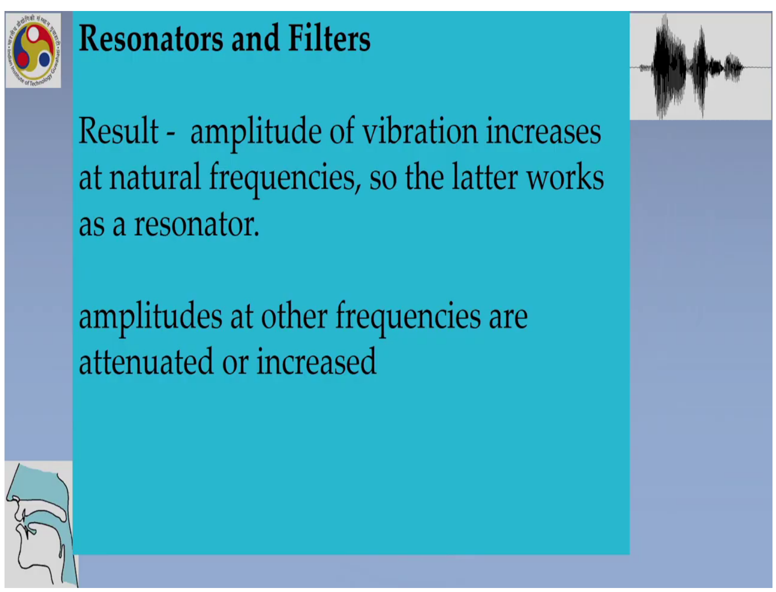
So sound wave is a mechanical wave. We also saw that. And we saw that mechanical waves need medium and that is air for speech sounds. And we also saw how, towards the end of the talk in the last lecture, we saw how resonators and filters play a role in sounds.

(Refer Slide Time: 1:55)



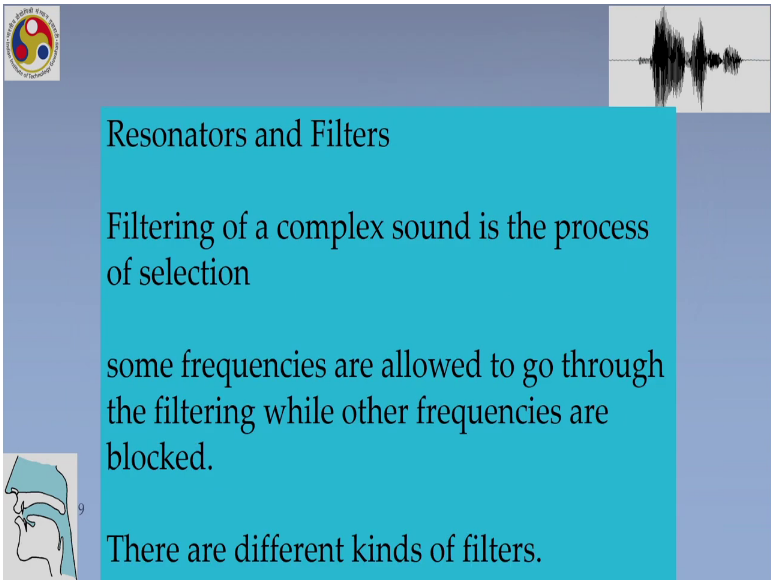
So we saw the different vibrating objects are tuned to different specific frequencies; and each object might have its own natural frequency, the frequency at which a particular object prefers to vibrate. And that may be different from object to object. So when the frequency of an external force is near to the natural frequency of the vibrating system then we have something which is called resonance. We will see how resonators work in the human speech system in this lecture.

(Refer Slide Time: 2:32)



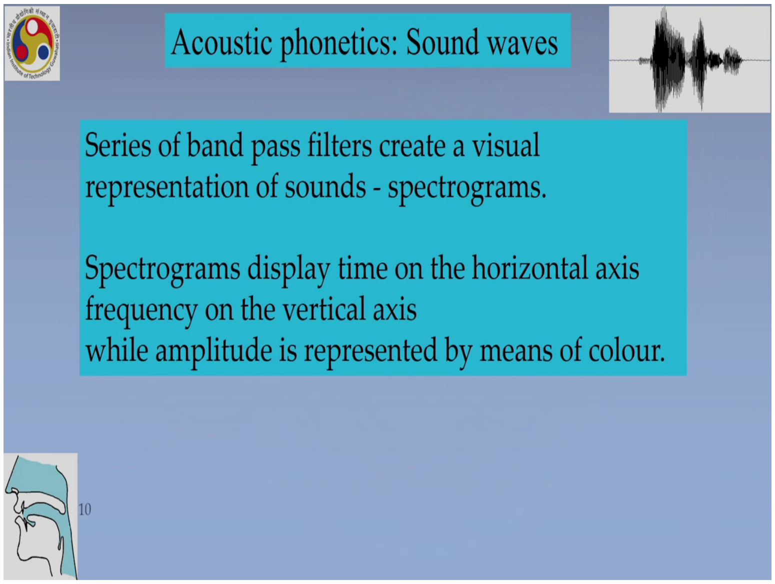
So as a result of this resonating property, amplitude of vibration increases at natural frequencies. So the latter works as a resonator. And depending on the resonating frequencies of the resonator the amplitudes at certain frequencies are attenuated, at certain frequencies they are increased.

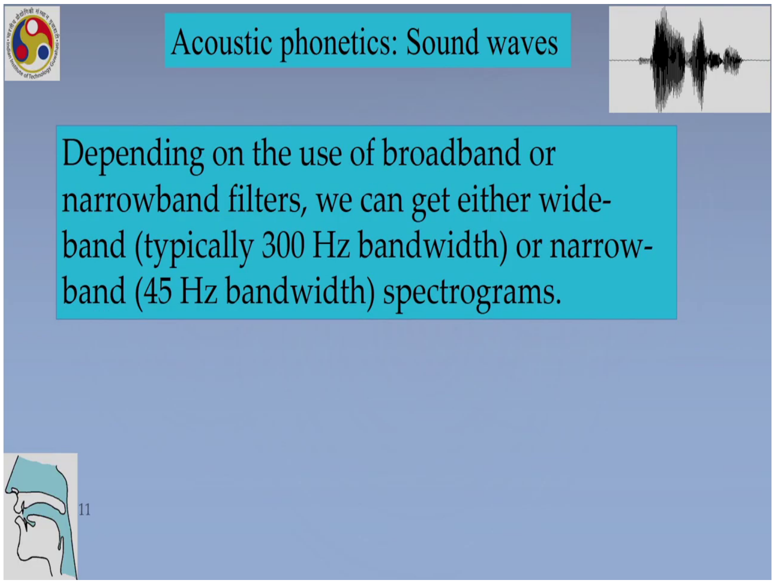
(Refer Slide Time: 3:04)



Complex sound is a process of selection of sound frequencies which are allowed to go through the filtering while other frequencies which are not allowed to, which are dampened or which are blocked.

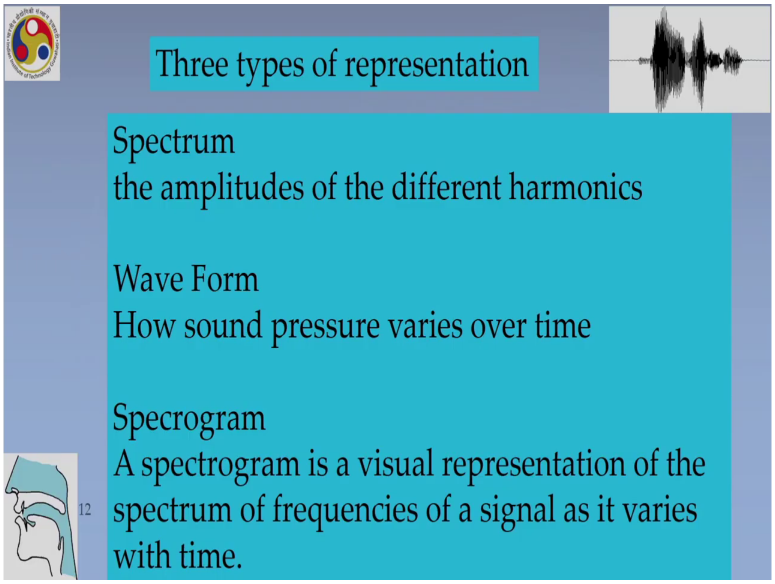
(Refer Slide Time: 3:20)





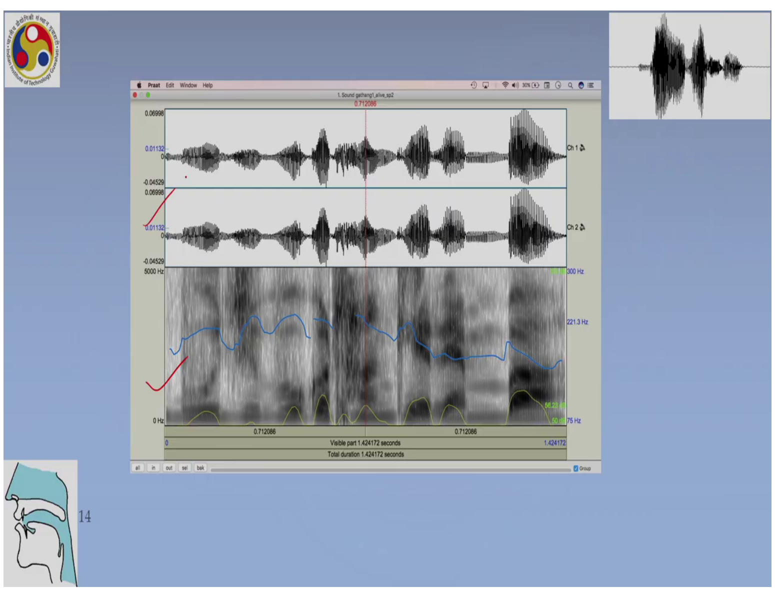
And we also saw how there are couple types of filters. There are band pass filters and we have narrow band filters, wide band filters. And then depending on whether it is a narrow band or a wide band, the certain frequencies are allowed to go through and certain frequencies are not allowed to.

(Refer Slide Time: 03:44)



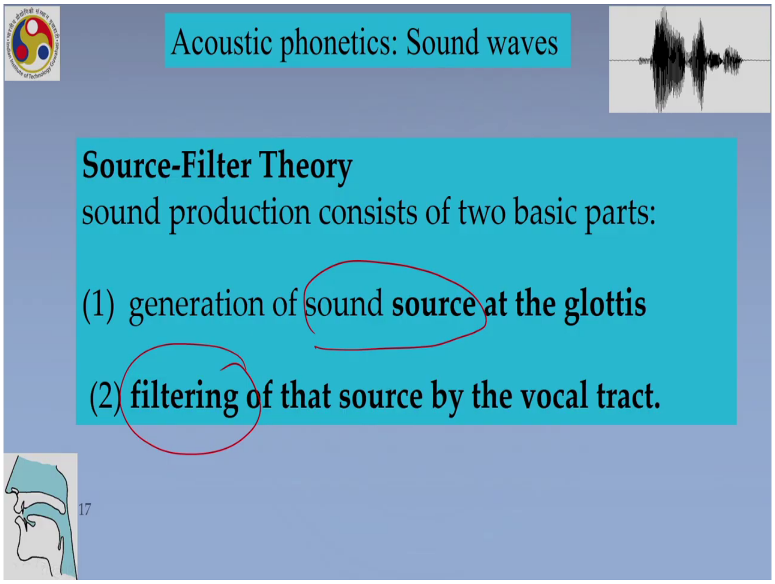
So speech sounds are represented in three different ways. So let us begin there. So the spectrum which shows the amplitudes of the different harmonics, the waveform shows how sound pressure varies over time, and a spectrogram is a visual representation of the spectrum of frequencies of a signal as it varies with time.

(Refer Slide Time: 04:14)



Now this is a wide band spectrogram and a wave form that is a word here and entire sentence here. You see what is below here is a spectrogram and here is the wave form. So how did we get these things? So we will study, we will look at these things in this lecture.

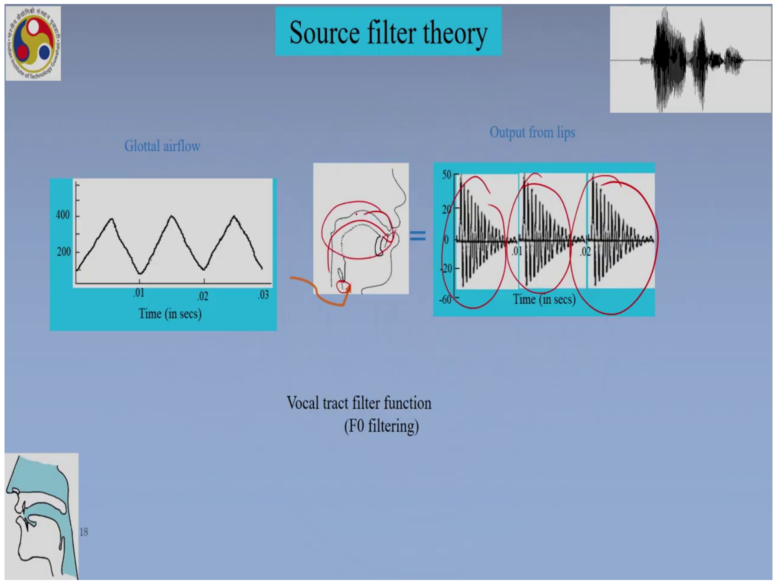
(Refer Slide Time: 04:36)

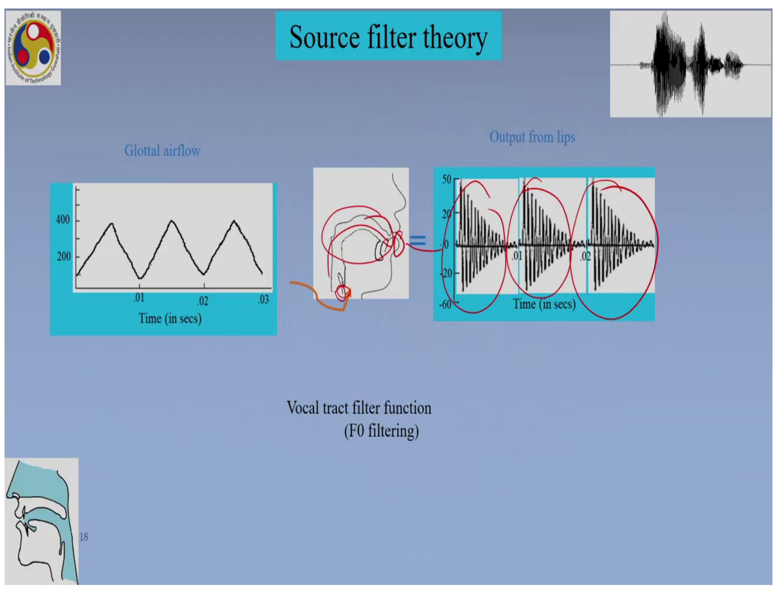


So before we, I will come back to those different types of representation, I want to first talk about something which is very basic and considered very basic to the understanding of how we hear sounds. So there is a particular theory which is called the Source-Filter theory. So sound production, which believes basically that sound production, consists of two basic parts. What are the two basic parts?

One is the generation of source sound at the glottis and the other is the filtering of that sound by the vocal tract. So remember there are two things. There is a sound source. And there is filtering of that source. So these two play a very important role in understanding speech sounds. And the rest of the lecture will be dedicated to understanding this.

(Refer Slide Time: 05:38)

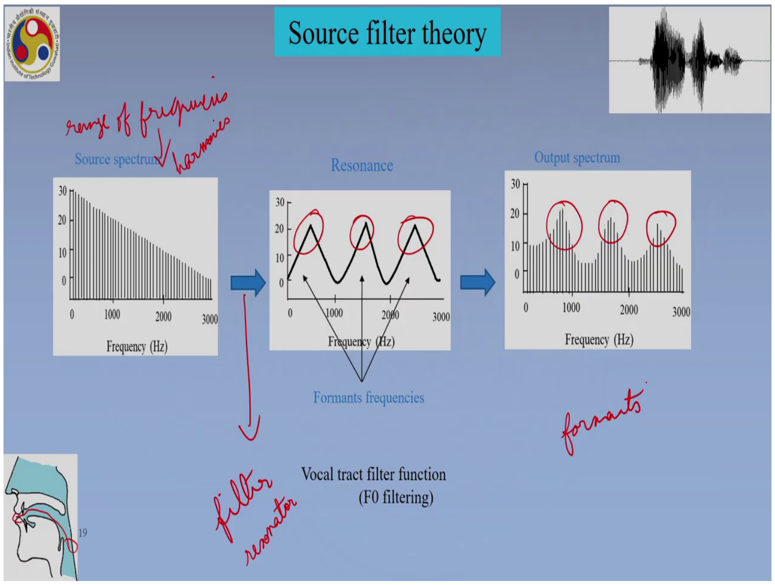




So now what we see here is a wave form. It is a glottal airflow. Now the glottal airflow, this can be thought of as the pulses of the glottal, at the glottis. As it is opening and closing so there is periodicity and this is the basic periodicity at the glottis. Now, so this is the system, supposed this is the glottis region and this is the sound which is produced there, this is the basic sound source.

Now when this passes through the resonating chamber inside a vocal tract we get something like this. As you can see this is a much more complex wave form. So this which just shows the peaks and the valleys; and to this which shows so many frequencies, we have a filtering function inside the vocal tract.

(Refer Slide Time: 6:48)

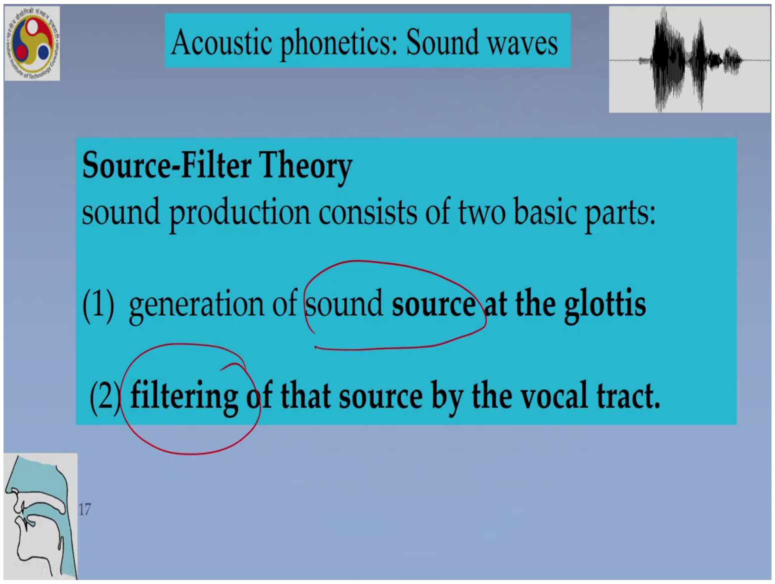


So what we have at the source is a spectrum, a source spectrum. Now what is a source spectrum? This represents a broad range of frequencies, range of frequencies. And this range of frequencies can also be called harmonics. So as we saw in the last class, harmonics are the integer multiples of the fundamental frequency. And now this goes through what we can call a resonating chamber. Now once it goes to a resonating chamber, this now results in your formant frequencies with this peaks.

So basically this is what the filter does. So the filter acts as a resonator producing all the formants. So, as a result we have these peaks. So from here where you can see only the broad range of frequencies with the harmonics, and now you see after going through the filtering you can see frequency at the peak and you can see this peaks which have a lot of energy, so energy here, energy here, energy here, energy here.

So this peak, the spectrum, the source spectrum at the glottis, at the glottis level goes through your supraglottal filtering function and as a result we get these regions of energy which may also be called formants.

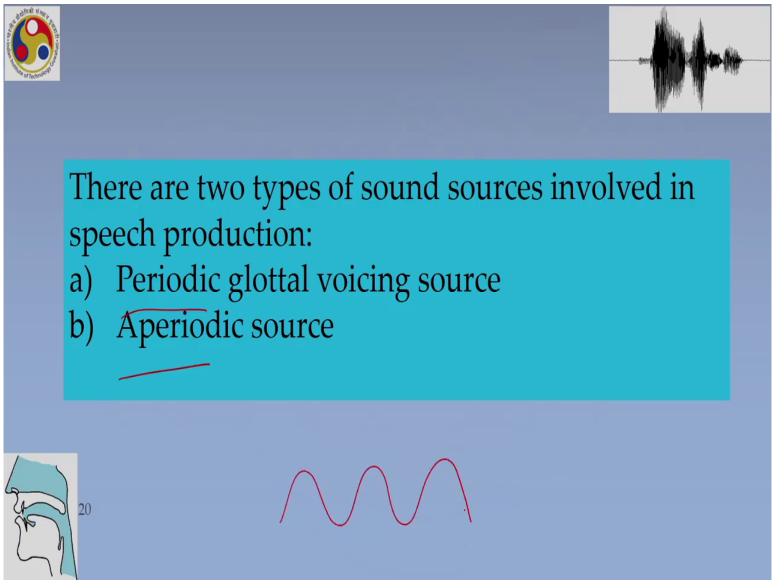
(Refer Slide Time: 8:55)



Now this is what a diagram showed you that this is generation of source sound at the glottis and then there is filtering of that source by the vocal tract. And understanding that glottal airflow which undergoes filtering function and produces the output from the lips. So when we have the output from the lips and what we have at the source are two different things.

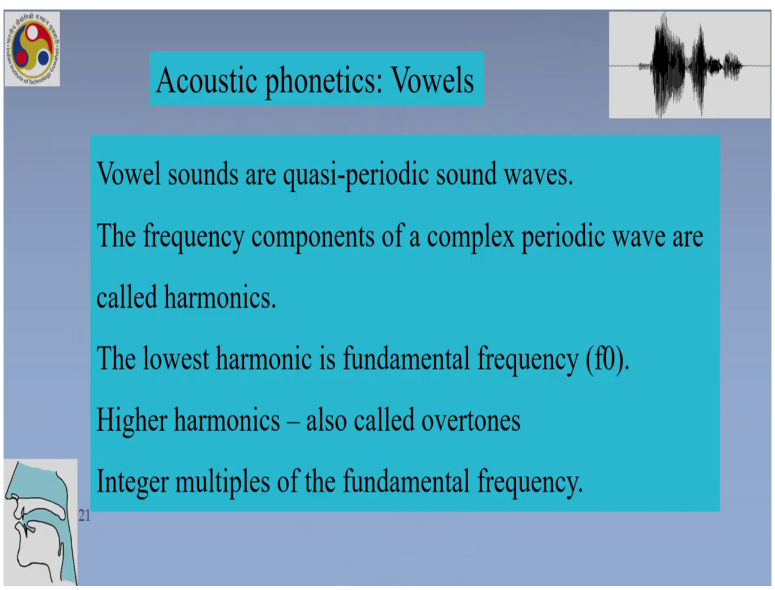
So one of the primary contributions of Source-Filter theory is also that the source is independent. That the source is independent, and what we have as a result of the filtering acquires quite a lot of the attributes of the supraglottal resonant chambers inside our vocal tract.

(Refer Slide Time: 09:50)



So the sound source can be both periodic and aperiodic. There may be couple of types of aperiodic sources. But basically we can think of it as periodicity which we saw in the form of the glottal pulses. And then there could be also aperiodic sounds.

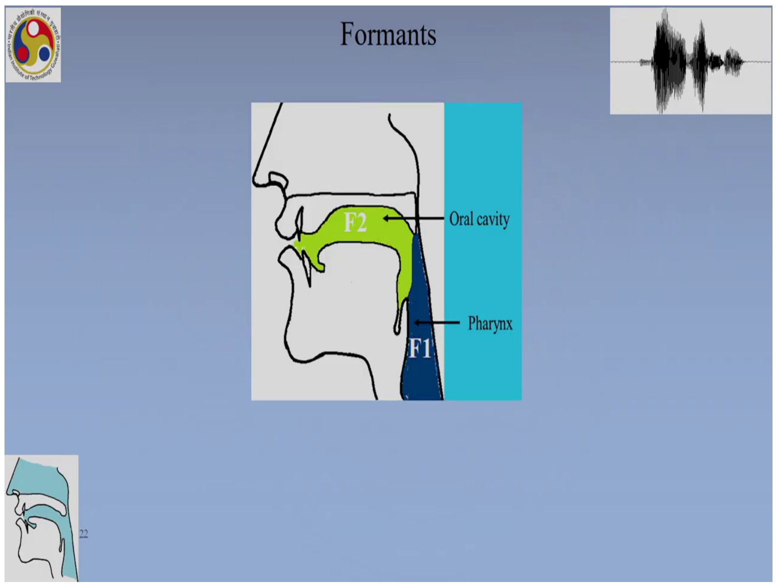
(Refer Slide Time: 10:10)

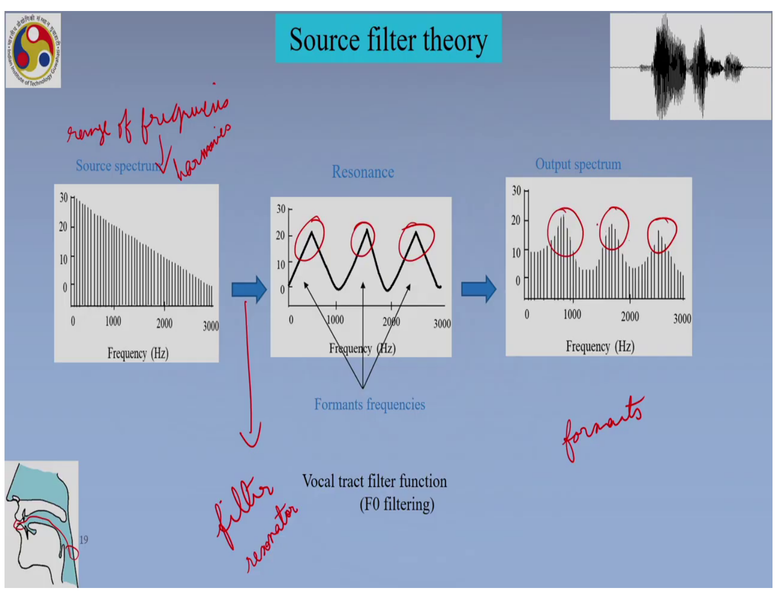


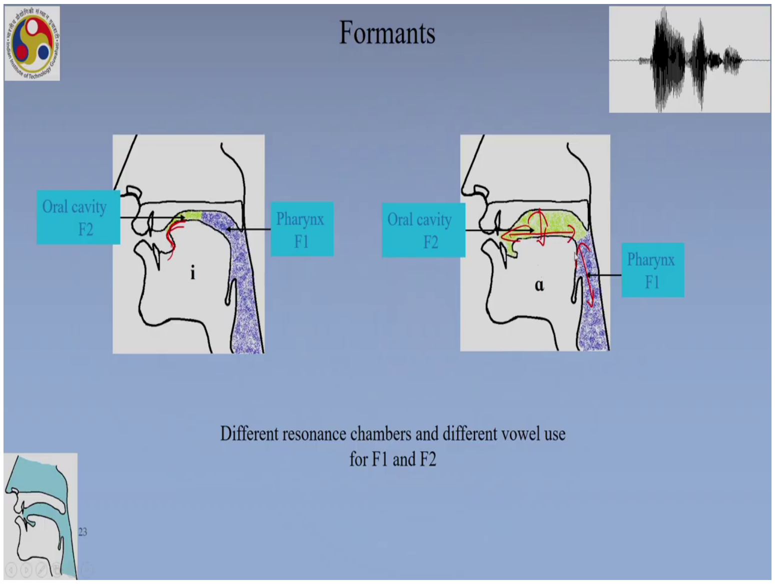
So coming back to vowels. Vowel sounds are periodic but if you want to know more details it could also be called as quasi-periodic but you can think of them as periodic sound waves. And the frequency components of these, of the complex periodic waves are called harmonics. So we already saw the harmonics at the glottal source.

So, you can think of those harmonics at the glottal source and then the lowest harmonic is fundamental frequency, that is, the rate of vibration of the vocal chords. And the higher harmonics are called overtones which are integer multiples of the fundamental frequency.

(Refer Slide Time: 10:59)





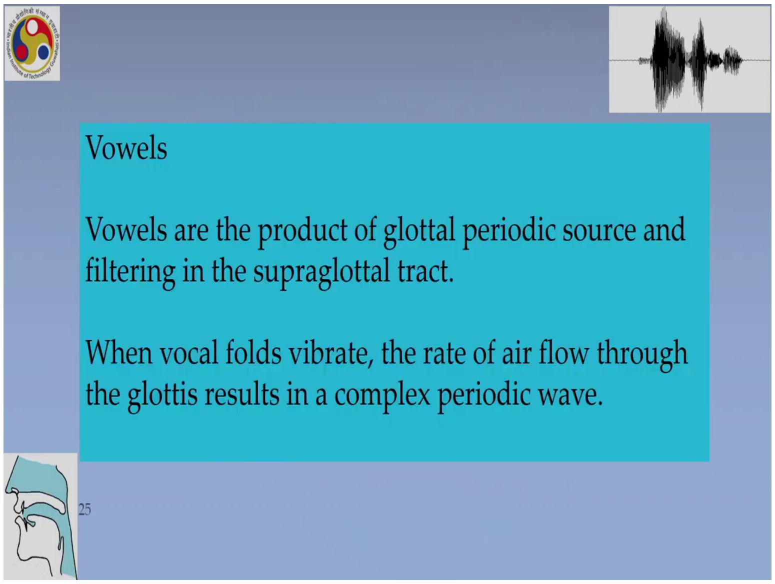


So now formants, if we go back we can see that these are formant frequencies. These peaks of energy which are the output of the, of the resonant chamber, these are the formants. Now formants get their properties mainly from the cavities. So that is the oral cavity and the pharyngeal cavity. So here are, you see that we have two vowels; one is i and one is a.

So in the production of i you will see that we have a region in the front part of the mouth where the tongue is at a higher position unlike ɑ where the production of which the tongue is at a lower position. And both the pharyngeal cavity and the front cavity are both are large unlike the first vowel. So these are the resonant chambers.

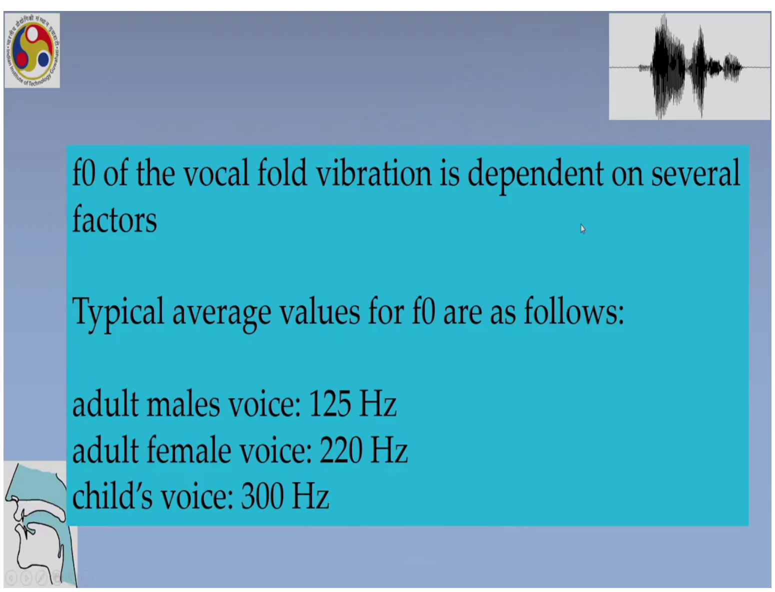
So when we talk about the oral tract now we have to remember that there are many chambers here. So this is, this could be one chamber, this could be one chamber, so this, so these are the various supragrottal chambers which give their vowels their specific characteristics.

(Refer Slide Time: 12:26)



So vowels are the product of glottal periodic source and filtering in the supraglottal tract which we just saw. And when the vocal folds vibrate the rate of air flow through the glottis result in a complex period wave.

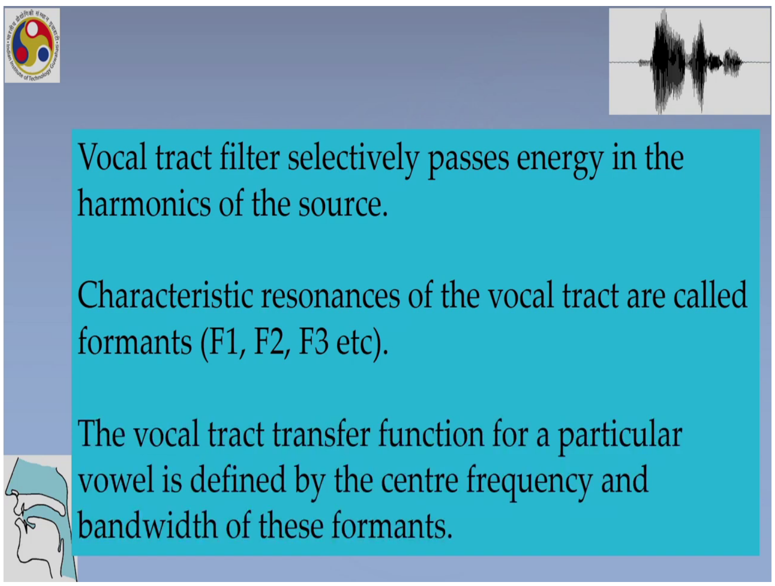
(Refer Slide Time: 12:45)



And also couple of important things about the fundamental frequency. So as we just learnt that fundamental frequency is the rate of vibration of the vocal folds, so, whereas the harmonics are the multiples of that base of that fundamental frequency. There are many other frequencies which are overtones, which are multiples of that fundamental frequency.

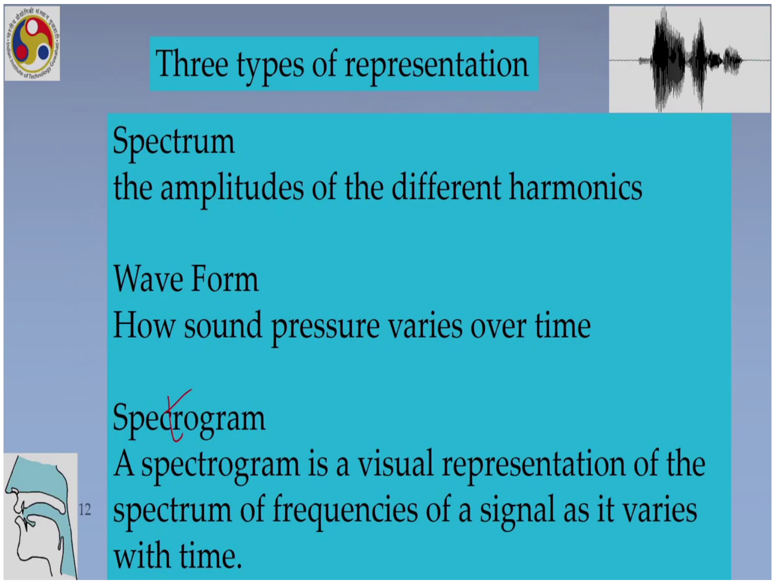
So a vocal fold vibration is dependent on several factors. Typical average values for fundamental frequency are as follows. So, however, these are supposedly the average values for adult males lower than adult female voice and then the highest is the child's voice at 300 Hertz and adult female voice at 220 Hertz and adult male at 125 Hertz.

(Refer Slide Time: 13:47)



So what does the vocal tract filter do? Again it selectively passes energy in the harmonics of the source, and as a result of that we have characteristic resonances of the vocal tract, and these are called formants; F1, F2, F3.

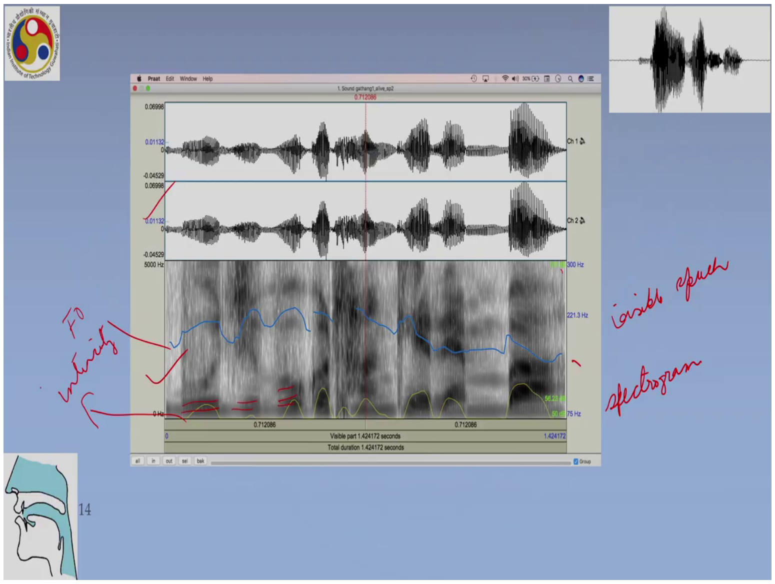
(Refer Slide Time: 14:18)



So I would like to again go back to what we presented in the initial part of this lecture, the three different types of representation of a sound unit. So you can have a spectrum. So you saw what a spectrum is. So we have power spectrum which shows all the different harmonics. And then we have wave form which shows how sound pressure varies over time. And spectrogram is a visual representation of the spectrum of frequencies as it varies over time.

So a spectrogram is very useful for linguistic analysis as it shows the variation over time whereas spectrum with all the information of all the frequencies may be used for other types of analysis. The spectrogram is used for linguistic analysis.

(Refer Slide Time: 15:05)

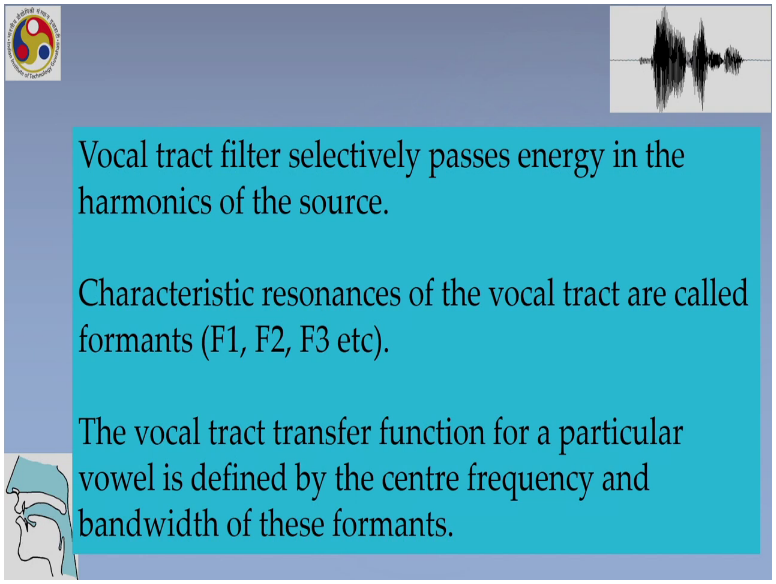


So this is a spectrogram. We will study in the next class, in the next lecture how to read spectrograms. Till that time you can see that, in the bottom part of this picture you can see that this looks like a photograph and this is called, for that reason it is called visible speech. And what we have here is our wave form.

And as you can see, you can see the varying amplitudes over in the wave form there. And here in the spectrogram you can see the dark regions where the peaks, energy peaks which tell you about, about the formants, the main formants F1, F2, F3. And then also you have other information like intensity and fundamental frequency. So here the blue line will show you F0 or the fundamental frequency. The colored lines imposed on this spectrogram show you other information like fundamental frequency and the intensity.

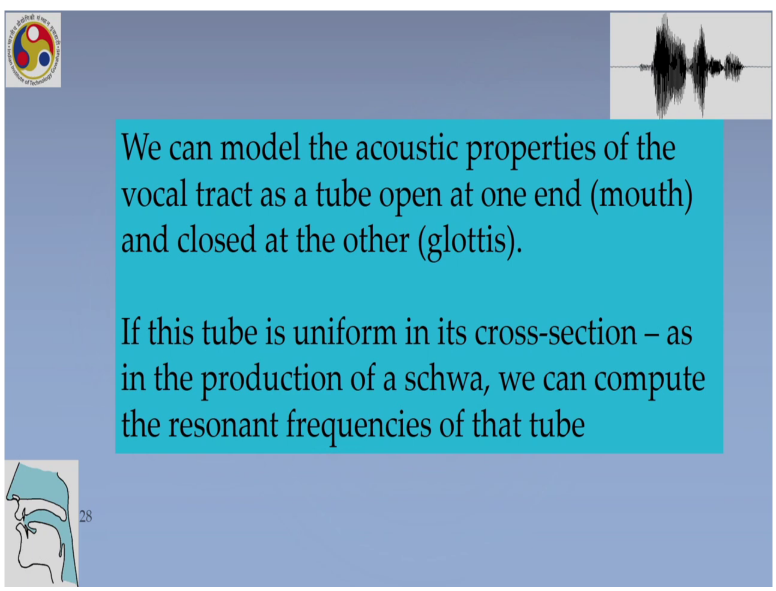
So this is a spectrogram and the wave form of a recording and it is a screenshot of a recording and its spectrogram and wave form and this you can get through various speech softwares like Praat where this was done. In the next class, in the next lecture we will see how we can do these analysis in free softwares which are available such Praat.

(Refer Slide Time: 17:04)



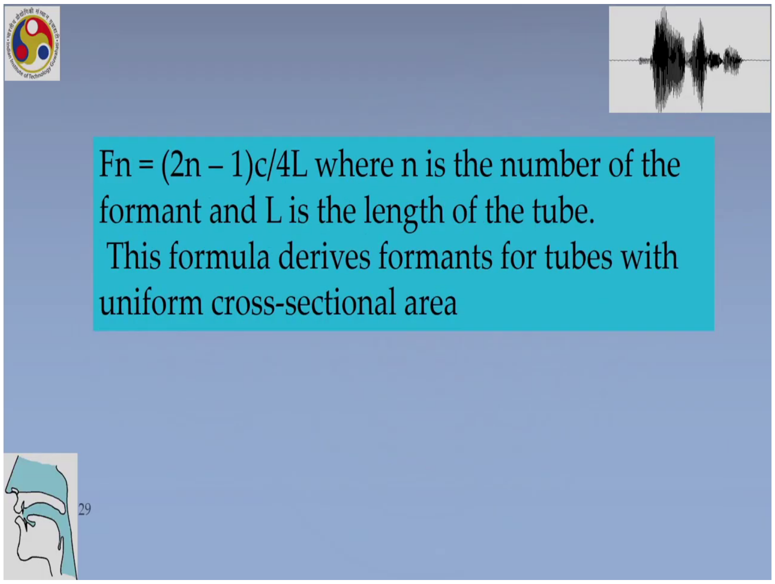
So coming back to vowels. So as we just said the vocal tract filters, selectively passes the energy, that is, the harmonics of the sound source. So characteristic resonance of the vocal tract are called formants. And the vocal tract transfer function for a particular vowel is defined by the centre frequency and bandwidth of these formants.

(Refer Slide Time: 17:29)



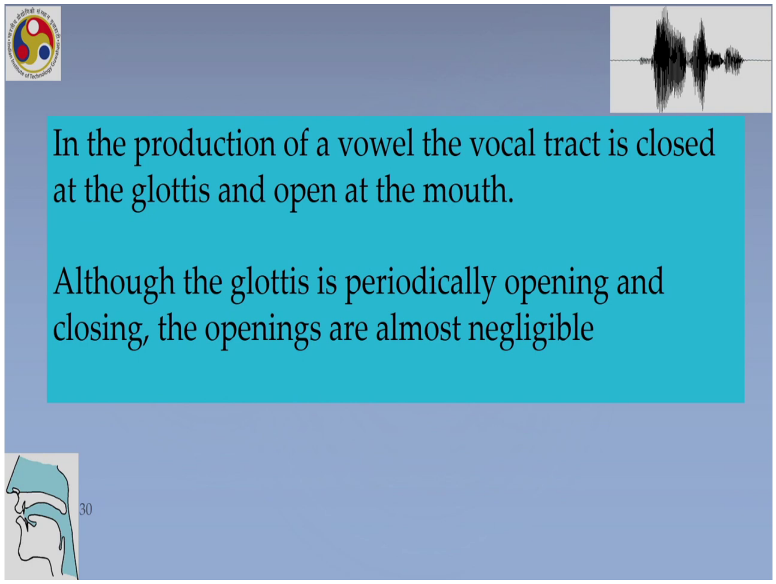
We can also model the acoustic properties of the vocal tract as a tube open at one end and closed at the other. So if this tube is uniform in its cross-section as in the production of a schwa, we can compute resonant frequencies.

(Refer Slide Time: 17:47)



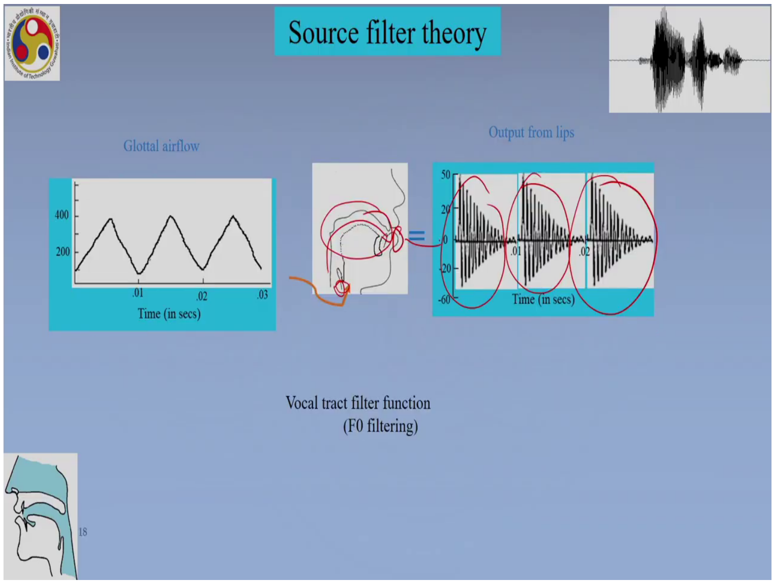
And there is a formula where you have, n is a number of the formant and, L is the length of the tube. And this formula derives formants for tubes with uniform cross-sectional area. So this is how we get the formant values.

(Refer Slide Time: 18:04)



So the reason for this formula is that, the production of a vowel, the vocal tract is closed in the glottis and open in the mouth. So you can think it as closed in one end and open at the another. So although the glottis is periodically opening and closing to produce the periodicity so this is the way that the vocal tract is thought of as the length of the vocal tract and it is, that is why we have 4L. And suppose we have a certain x length of the tube for a certain person then you can multiply that with the length of the tube into 4L. And if it is F1 and this is how you will get the values.

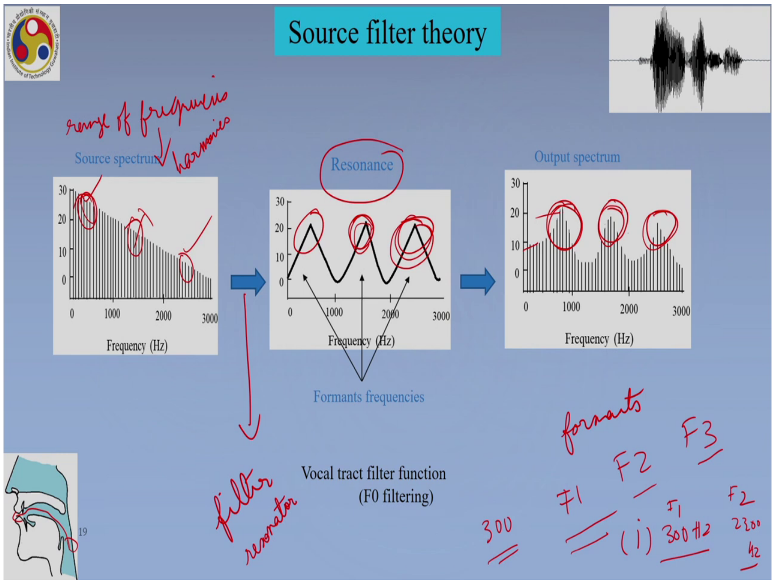
(Refer Slide Time: 19:04)



So before we talk about individual vowels, for another time let us look at the Source-Filter theory before we move on to looking at vowels. So as we have talked about before, the glottal airflow goes to the Vocal tract filter function and gives us the output from lips. So what we see here, the up and down movement, so this is because resonator filters the glottal airflow and results in these kind of movements.

Now something that we have to note is that the result of this output from the lip that we are studying now as the result of the output and suppose these are formants of vowels, what is important to note is that, once the source pushes the air through the vocal tract then it takes the shape of the resonant chambers inside the vocal tract. So even though it has some properties of the glottal airflow it also takes the properties of the resonant chambers, because we are going to study F1, F2, F3. So the formant, 1 formant, 2 format, 3, so why do we get formant 1, formant 2, formant 3? Suppose these are the individual harmonics.

(Refer Slide Time: 20:25)



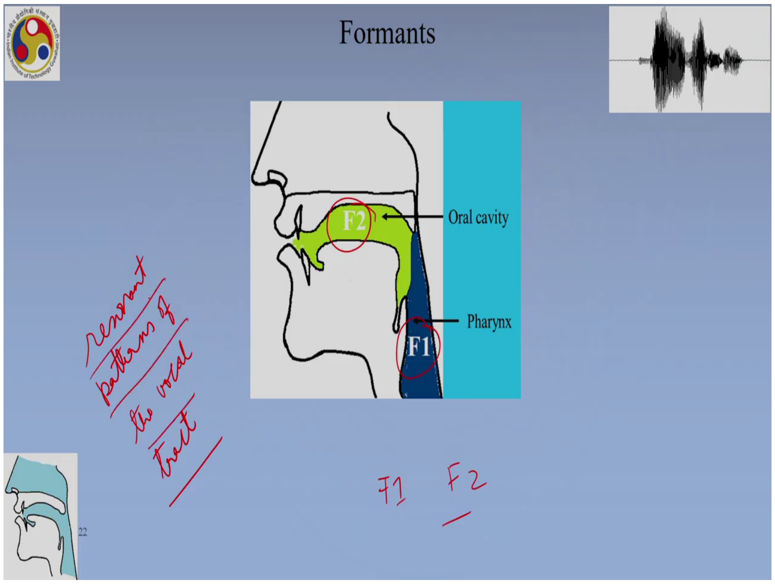
So let us see the other diagram. Suppose these are the individual harmonics. And we have a broad range of frequencies here, the harmonics. And the frequency response, the peaks that we see here. This is the resonant cavity. So the responding parts are the parts which matches. Suppose, suppose the frequency here, there is the frequency here. There is the frequency here. There is the frequency here. So the frequencies here may not receive because of the way the filtering cavities are, the frequency here may not be attenuated.

So it may not receive that kind of a push here in the resonant chamber because it does not match the inherent frequency in the resonance chamber. So, where as, this gets more, this gets highlighted through the resonating chamber. This may not be highlighted from the resonating chamber. This may be highlighted by the resonating chamber.

The resultant frequency is the resultant, so what we get F1, F2, F3 is because of the way the cavity inside, the supraglottal cavities that we have which will give shape, which will give the particular properties of the frequencies. So if F1 is 300 Hertz that is because this frequency matches the frequency that is pushed out which is coming from the glottal source.

Suppose there is a vowel, we have vowel e with frequencies 300 Hertz and 2200 Hertz. So why is F1 300 and why is F2 2200? Suppose this is F1, suppose this is F2. Why did we get these two particular formants? It is because the resonating chamber promoted these frequencies and attenuated or demoted or blocked certain other frequencies. Now that is a role of a resonating chamber, a filtering cavity like the vocal tract.

(Refer Slide Time: 22:51)



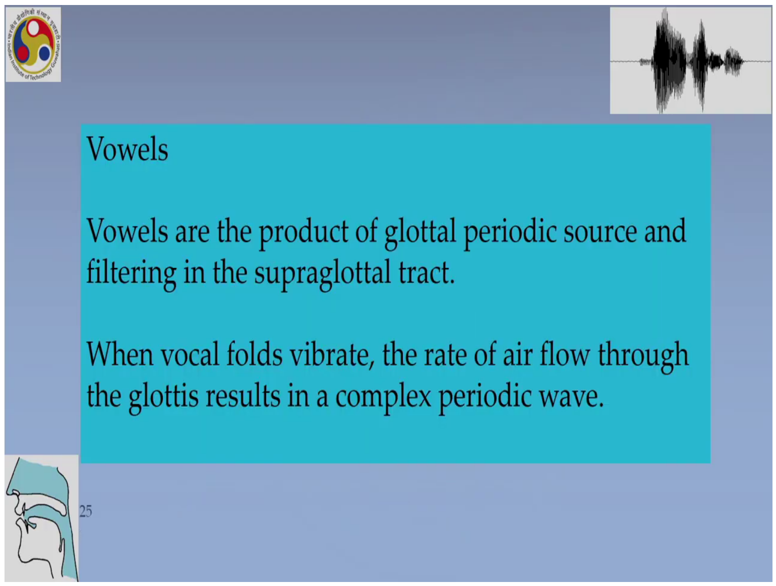
So now again continuing with formants. The human vocal tract has a particular spectral shape and size. And that results in the formant frequencies that we often get as a result of the source, air going through the supraglottal cavity. And we find that particular resonant frequencies are promoted, particular resonant frequencies are not promoted as a result of the particular spectral shape and size of the vocal tract.

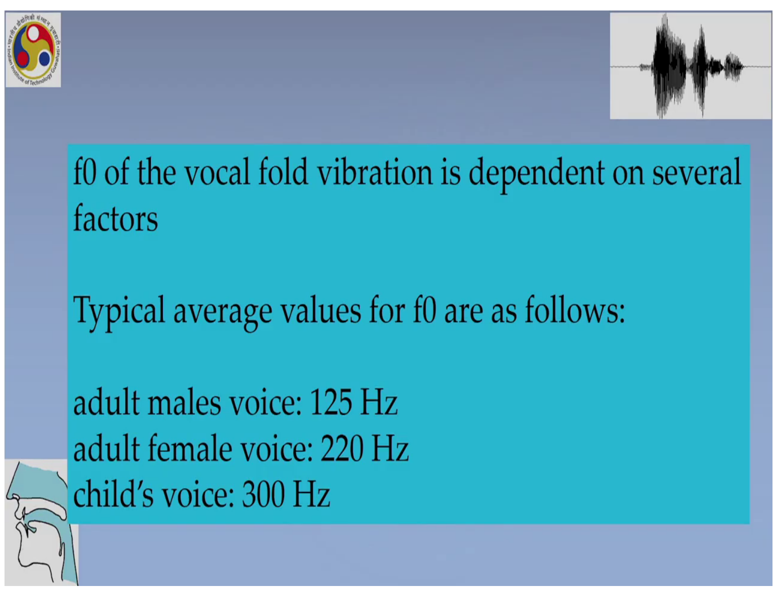
So we have two things. As again we have the source and we have this resonating chamber. Now we are just explicitly showing here which are the two chambers responsible for, suppose F1 and F2. So this pharyngeal cavity gives shape to the first formant. The oral cavity gives shape to the second format.

And that is why we have the values, the particular values for high front vowel like e its particular frequency is the formants because of the way these two chambers are in the production of the sound e. And again if we take the vowel a, it is because of the way these two resonating chambers, the frequencies that the resonating chambers promote will result in the particular formant frequencies.

So what are formants? Formants are resonant patterns of the vocal tract, resonant patterns of the vocal tract. So these patterns of the vocal tract are not just the harmonics which we get from the source, which, the result of that source spectrum going through these resonating chambers results in our particular formants. And these are much more complex than what we have at the source.

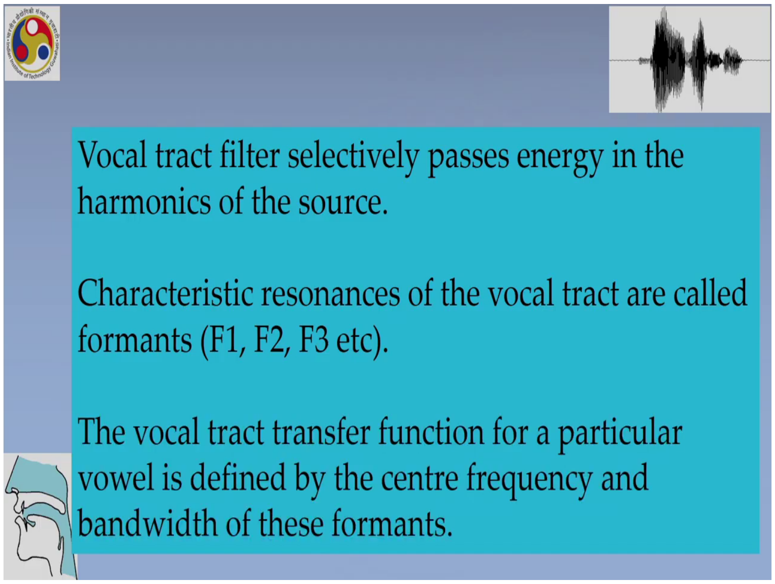
(Refer Slide Time: 25:14)





So we have talked about how vowels are product of the glottal periodic source and filtering in the supraglottal tract. And when vocal folds vibrate the rate of airflow through the glottis results in the complex periodic wave. So as we just said, it is much more complex than the harmonics that we have at the glottal source.

(Refer Slide Time: 25:38)

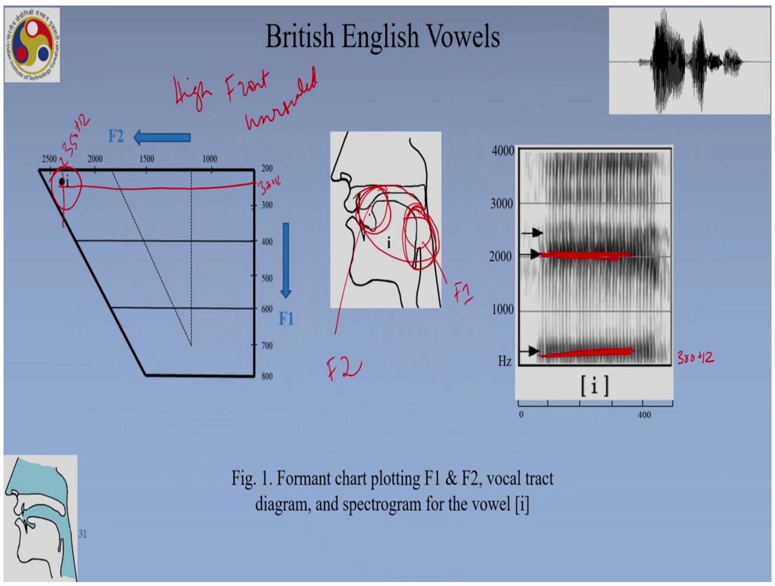


Now we have talked about the particular characteristics of sounds; how there is a fundamental frequency which always depends on the size and shape of the vocal tract; male, female fundamental frequencies may vary but those are generalizations. You can often have outliers. You can have adult males, adult females whose frequencies are not exactly the ones given here. So you can always expect those things.

And vocal tract filter selects the energy that passes through the harmonics of the source. And the characteristic resonances are called, therefore, formants as we have already discussed.

So let us now go to individual British English vowels. Let us see how each of these vowels can be produced. Now before that I would again like to recap what we had just studied about the Source-Filter theory of sound production. So there is a two-stage process. That is what it says, and the generation of sound source has its own, the source has its own spectral shape and size and structure. And this is then filtered through a resonating chamber. And this resonating chamber has its own properties of the vocal tract. And most of the filtering property of the sound source is carried out by the cavity which is anterior to the glottal source.

(Refer Slide Time: 27:26)

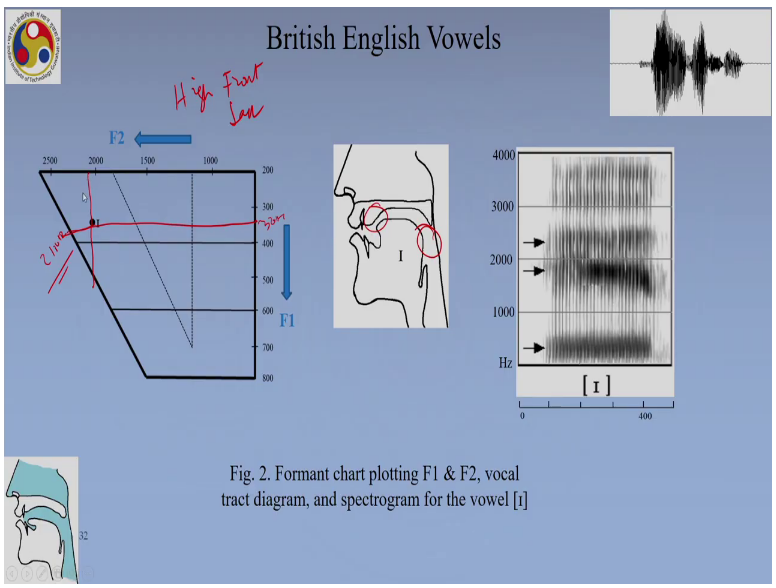


Now let us look at these British English vowels and see how we can explain what we have just studied with regard to source filter and the formants that are produced as a result of that filtering. So let us take the vowel i. So this is the British English vowel i and we can see that as the tongue position is raised towards the front part of the mouth. So we see a little bit of less space there while we see the wide cavity, the pharyngeal cavity is wide, the front cavity is narrow. And what do we get here?

So what happens when we get this particular combination of our resonating chambers here? So these two resonating chambers; one here for the F1 and one here for the F2. Now as we already studied, F1 is the pharyngeal cavity, is responsible for the F1 and the front cavity is responsible for F2. So this particular combination of the pharyngeal cavity and the front cavity results in these, these F1 and F2 values. So this is roughly around 300 Hertz and so as it is plotted here, as you can see, it is roughly around 300 Hertz. So it is around, say let us say 370 Hertz. And here it is around say 2300-2400 Hertz. So we get these two formant frequencies as a result of these two cavities.

So if you recall from our study on articulatory phonetics, so what do we call the vowel i, what are the 3 labels that we will give for i? It is a high, front, unrounded vowel. So it is high and it is front. So, and the back cavity is particularly wide because it is a high vowel and it is a front vowel that is why the front cavity is particularly narrow.

(Refer Slide Time: 30:04)



Secondly now let us take the lax vowel; the English lax vowel, high, front but lax vowel. So the production of this British English vowel, what do we get? We get that, that the F2, suppose the F2 and the F1 values are, so here the F1 value is between 300 and 400 Hertz. It is similar to i but slightly, slightly lower than i.

And that is why we see that the front cavity, there is little bit more space than I and the back cavity is similar to I has, is pretty wide. And that is why the F1 value there is still around 300 to 400 Hertz, so around 320 Hertz or so. And this is around 2100 Hertz or so. Because it is lax it is slightly low and the F2 value is also similar to that of i.

(Refer Slide Time: 31:26)

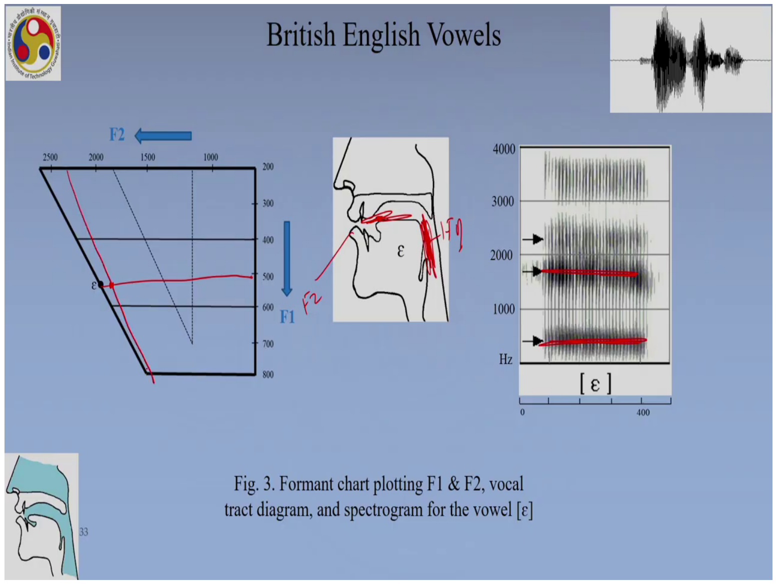


Fig. 3. Formant chart plotting F1 & F2, vocal tract diagram, and spectrogram for the vowel [ɛ]

Now let us take a vowel which is lower than i and \ɛ\, it is a for instance. Remember it is also a front vowel. But the only difference is that this is not as high an i as I and \ɛ\ but it is slightly lower and it is \ɛ\. So what happens when we have a vowel like that? We have again, because, so remember front cavity and back cavity, so this is, this produces F1 and this is F2. So again F1 is, is around 500 Hertz and F2 similarly is between 1000 to 2000, around 1800 Hertz. So it should be somewhere here and it is around 1800 Hertz.

(Refer Slide Time: 32:28)

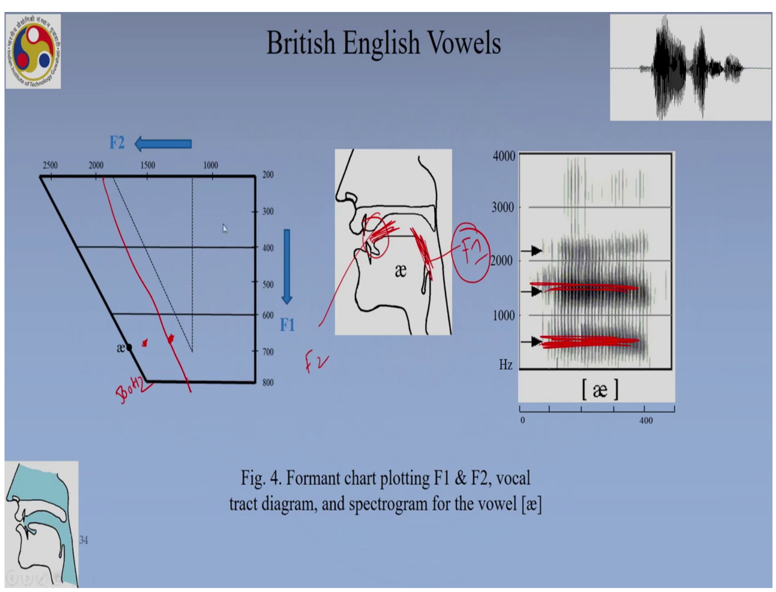
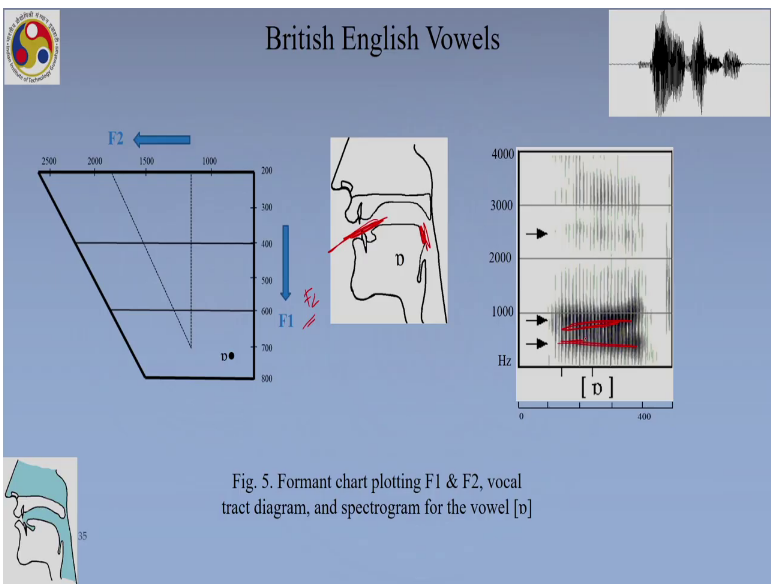


Fig. 4. Formant chart plotting F1 & F2, vocal tract diagram, and spectrogram for the vowel [æ]

Now let us take, still going down this, this F1 F2 levels; so going down in terms of height. So we had the highest vowels i, e and then slightly lower E. And then the lowest here ae. So when we have a low vowel like ae which is also front and which is low, so which means, its front means there is slight narrowing in the front cavity which produces, as we know, is responsible for F2.

And we have a wide pharyngeal cavity which is responsible for F1. And then, so what do we get here? So we get, so it is a front vowel. That is why the F1 value, the pharyngeal cavity which is wide, the values will be lower. It is again around 500 Hertz. And this is between 1000 and 2000 Hertz. So the F1 for ae, it is higher than 500. It is around 700 to 800. And F2 is around 1800. So it is, should be like this.

(Refer Slide Time: 34:07)

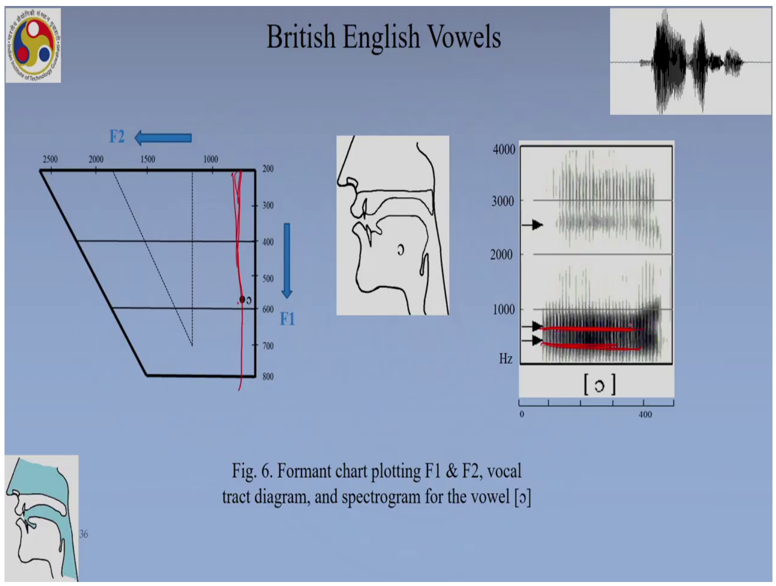


Now we move to the back vowels. Now we saw that when we have high front vowel the F1 is low. There is a wide pharyngeal cavity which is responsible for F1. So whenever we have the front vowels, so we saw that the front and the, there is a, especially with i which is very high vowel and there is a wide pharyngeal cavity, similarly the lax vowel a, and as we went down in tongue height there was an inverse relationship. So the highest vowel has the lowest F1. And it goes up as you, as the height of vowel goes down. So there is an inverse relationship.

What happens to back vowels? So now, depending on whether the vowel is high or low so there will be again changes in the cavities. So the production of o, the F1 and F2 values will be very close. Remember that in the production, when we had i, the F1 and F2 values were very far apart it is a front high vowel.

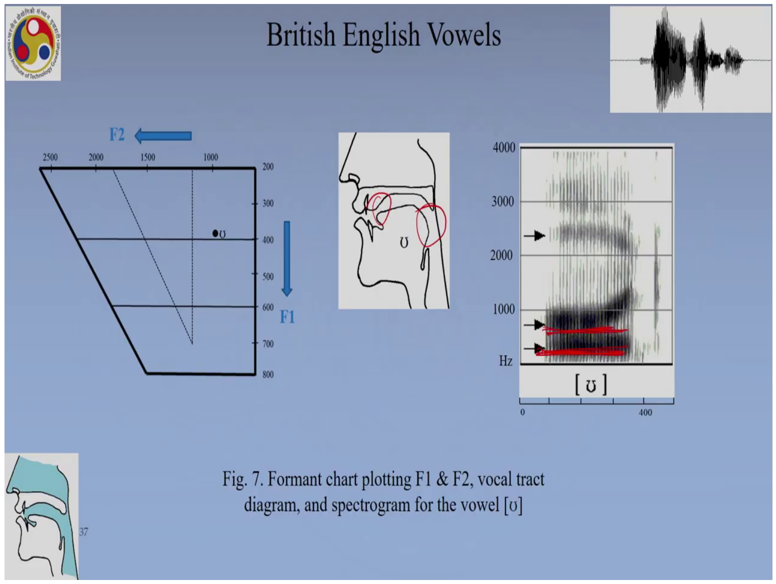
Now back vowels have a low F2. And that is why the F2 values go down. Back vowels, you have narrowing in the pharyngeal cavity whereas you have wider front cavity. So we know that the front cavity is responsible for F2. And now we can see that the F2 value here is much lower than what we had seen for all the front vowels.

(Refer Slide Time: 35:46)



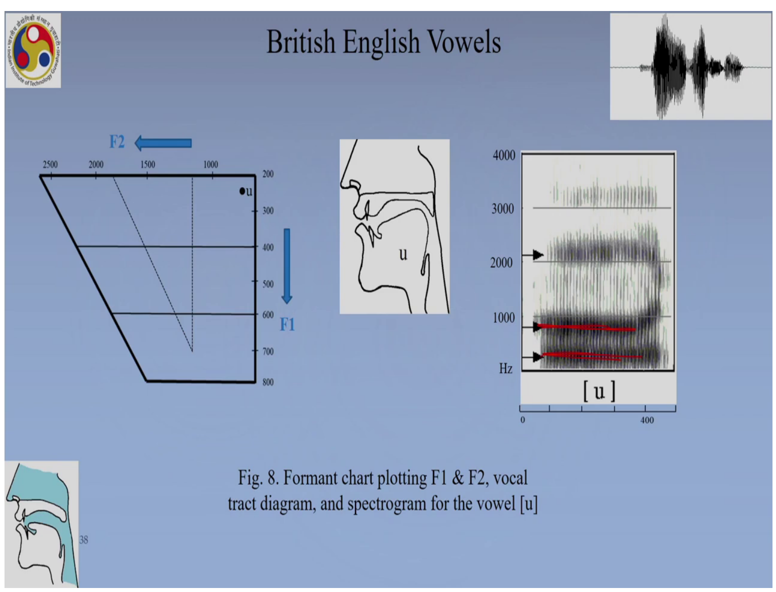
Similarly for o, again we see that F1 and F2 values are pretty close, and again we have an F1 value which is between 500 and 600, and F2 value which is also low. It is around, maybe, let us say, the F1 is around 580 and this could be around 700, 600-700.

(Refer Slide Time: 36:21)



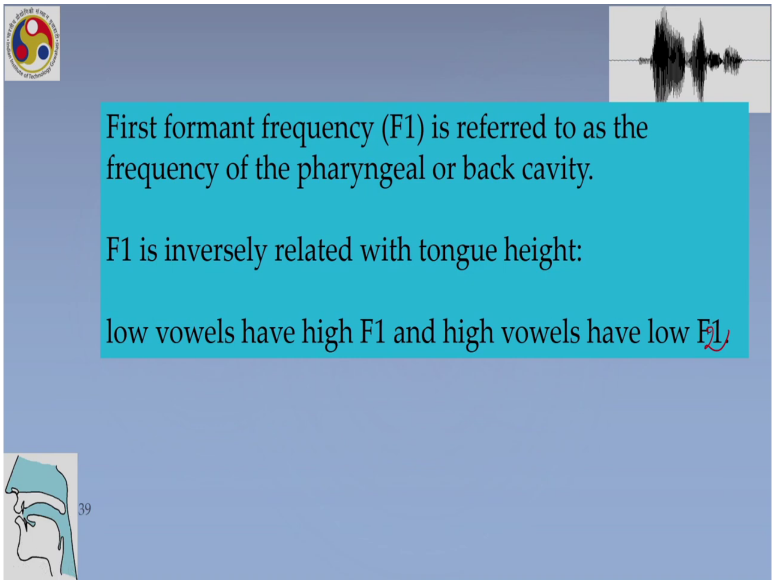
So similarly again as you get higher vowels. Now we have the lax vowel o, so again for the lax vowel o we see that because it is a high vowel, now for the production of a high vowel, we know that there must be narrowness in the front cavity and again because it is a back vowel there is also narrowness in the pharyngeal cavity. Now what happens when that happens? When that happens you will see that the two formants are very close to each other.

(Refer Slide Time: 37:01)



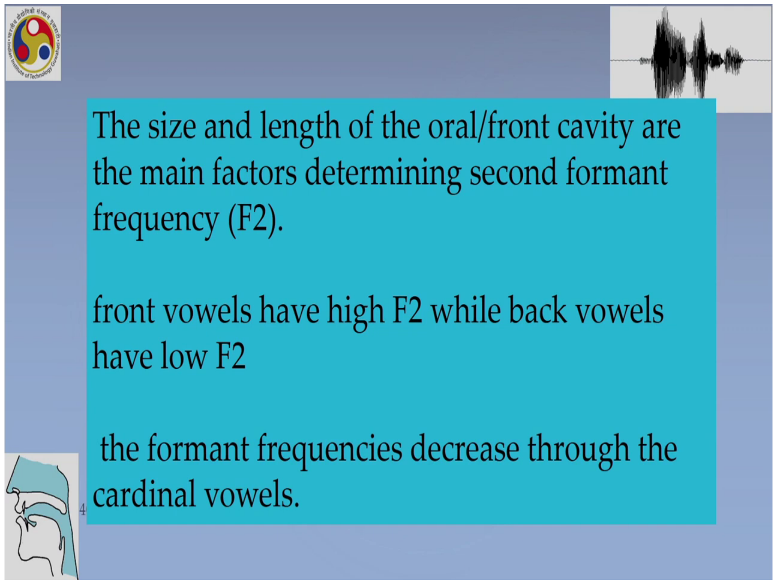
And now this is the highest back vowel, so for the production of u, again we see that the two, F1 and F2 values, the F1 is low because it is a high vowel and F2 is around 1000 Hertz because it is a back vowel.

(Refer Slide Time: 37:27)



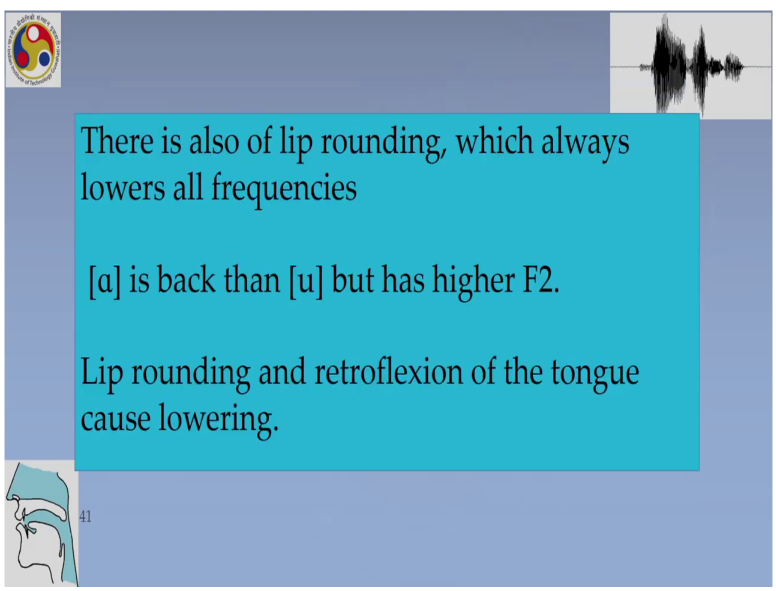
So, summarizing this, first formant frequency is referred to as the frequency of the pharyngeal or back cavity. It is always inversely related to tongue height. Low vowels have high F1 and high vowels have low F1. So the first formant frequency is referred to as the frequency of the pharyngeal back cavity. And F1 is inversely related to height we have seen that.

(Refer Slide Time: 37:57)



And the size and length of the oral front cavity are the main factors determining second formant. So front vowels have high F2 while back vowels have low F2. The size and length of the oral front cavity are main factors for F2.

(Refer Slide Time: 38:17)



And again there is also the factor of lip rounding which always lowers all frequency. a is backed in u but has higher F2 because of lip rounding. Lip rounding and retroflexion of the tongue cause lowering.

So with this we have come to the end of this lecture on acoustic phonetics, and we have studied in this lecture the Source-Filter theory of sound production, and how the formants are responsible for giving the particular properties to individual vowels and how we get formants and how we calculate them and how Source and Filter theory helps us to understand the differences between vowels. Thank you for listening and I will see you again in the next class. Thank you.