Indian Institute of technology Kanpur

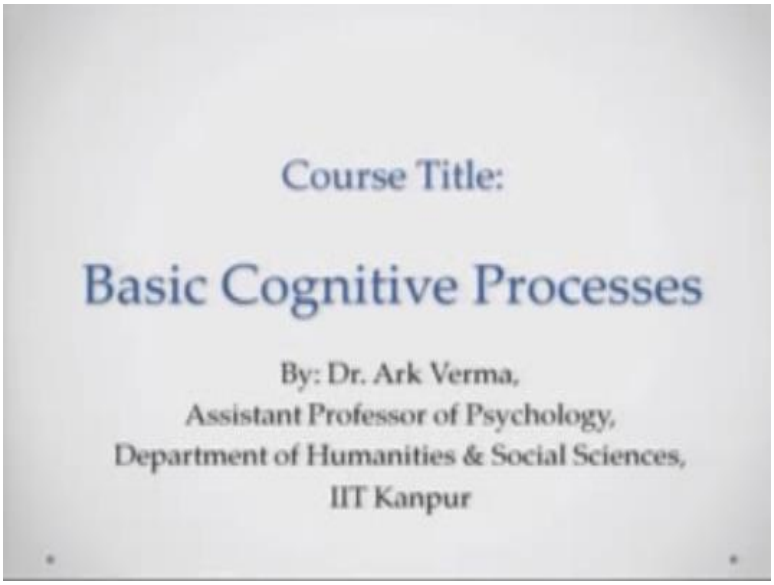
National Programme on Technology Enhanced Learning (NPTEL)

Course Title
Basic Cognitive Processes

Lecture – 24
Auditory Perception - II

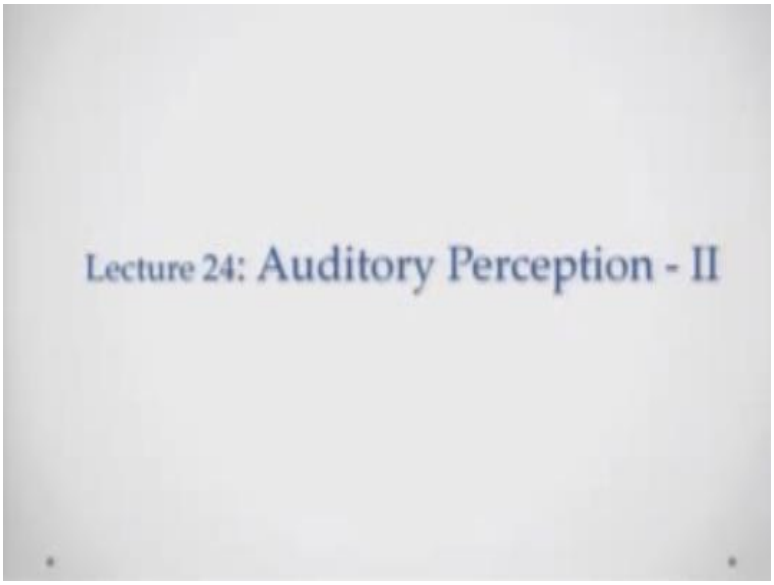
By
Dr. Ark Verma
Department of Humanities and Social Science
Indian Institute of Technology Kanpur

(Refer Slide Time: 00:17)



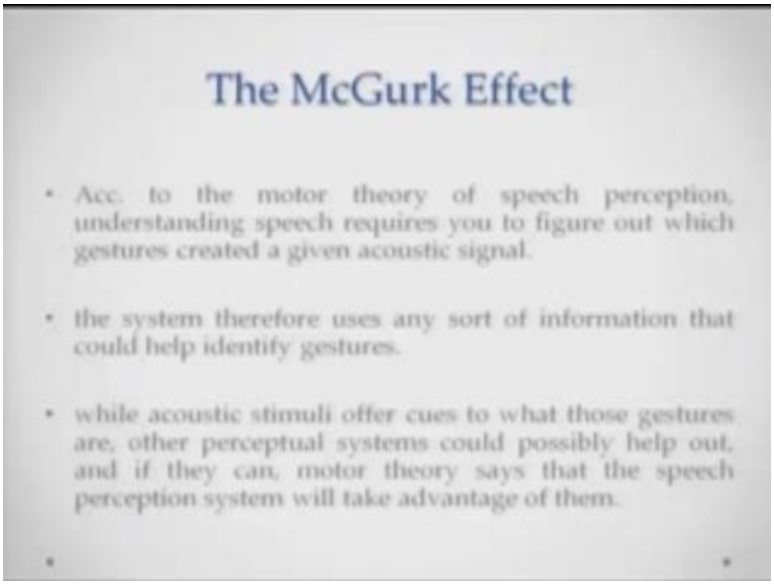
Hello everyone, welcome to the post basic cognitive processes I am doctor Ark Verma from IIT Kanpur if you remember in the last lecture we were talking about auditory perception.

(Refer Slide Time: 00:24)



we were talking we talked a little bit about the physiology of the ear and some physical characteristics of the sound we also talked about we started talking about the theories on motor theories of speech perception the theory we were talking about while I ended the last lecture was the motor theory of speech perception and we will try and continue from there onwards in this lecture as well.

(Refer Slide Time: 00:46)



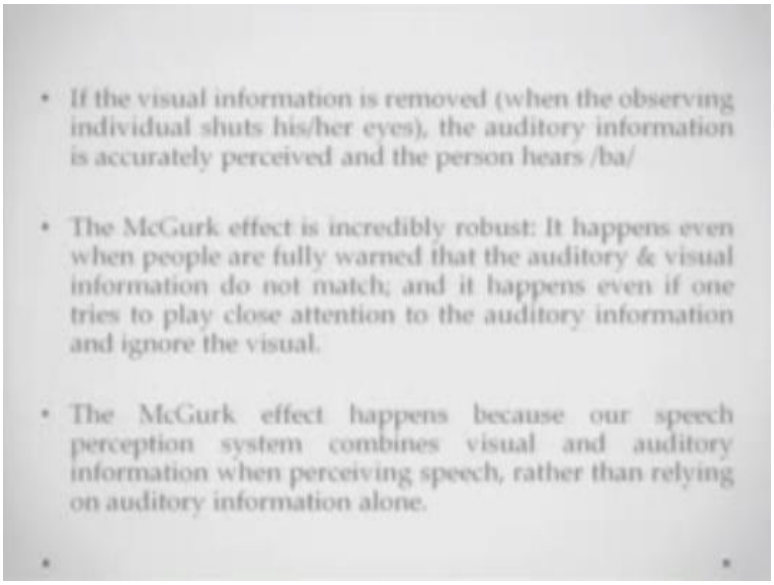
Now an interesting effect of the motor theory of speech perception is basically that it says that understanding speech gestures requires one to you know figure out whatever gestures have created any given acoustic signal the system therefore uses or will require to use any sort of information that can help identify these gestures now while the caustics only offer clue as to what those gestures possibly are the help can be taken from other perceptual systems as well if the perception systems can provide this kind of help they can provide any kind of clue as to what the gestures are.

The motor attorney says that the speech perception system will take up this information use this information and use it in understanding speech in fact two non auditory perceptual systems like vision and touch have already been shown to affect speech perception and the most famous demonstration of this multimodal has effect on speech perception is the McGurk effect which first was reported by mega Ian McDonald in 1976 now this mega effect you can find a lot of videos about on YouTube etc…

But the curse of this McGurk effect is that this happens when people are watching a video of a person talking but the audio portion of the tape has been altered for example the video might be

showing a person speaking ga but the audio signal is of a person speaking but generally what happens is that people perceive neither go or ga perceive a combination which you know comes out as the sound of da so this happening is the person is saying you are hearing us but the video is of a ba and what the system is doing it is combining these two information in some sense and coming up with the entirely new sound that is the why is this happening this is happening.

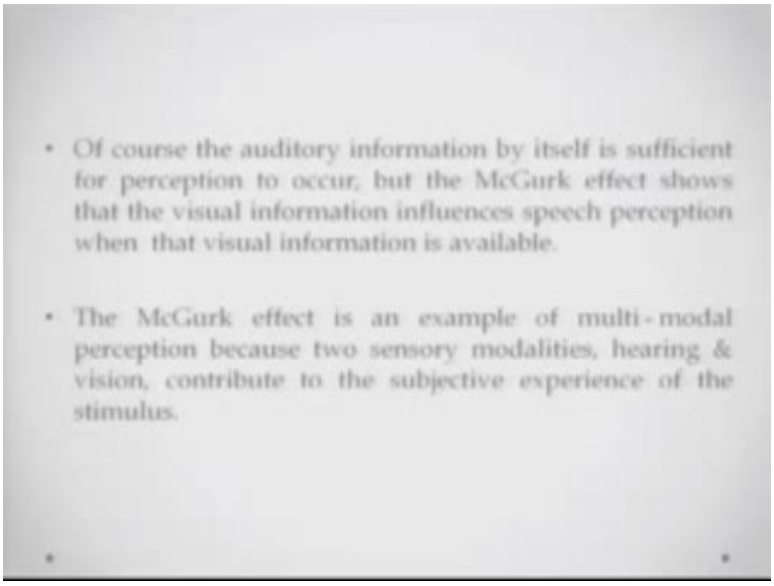
(Refer Slide Time: 02:34)



Because if the visual system information is removed the auditory information is accurately perceived and the person hears ba so if you kind of close your eyes and listen to this information the visual thing is gone and then you really perceive whatever the person was saying ba now this mega effect has been shown to be incredibly robust it happens even when people are fully born even if they are told they see the audio is different video is different you have to still try and understand it still you know happens people still cannot really control integrating of these two information and coming over this third category.

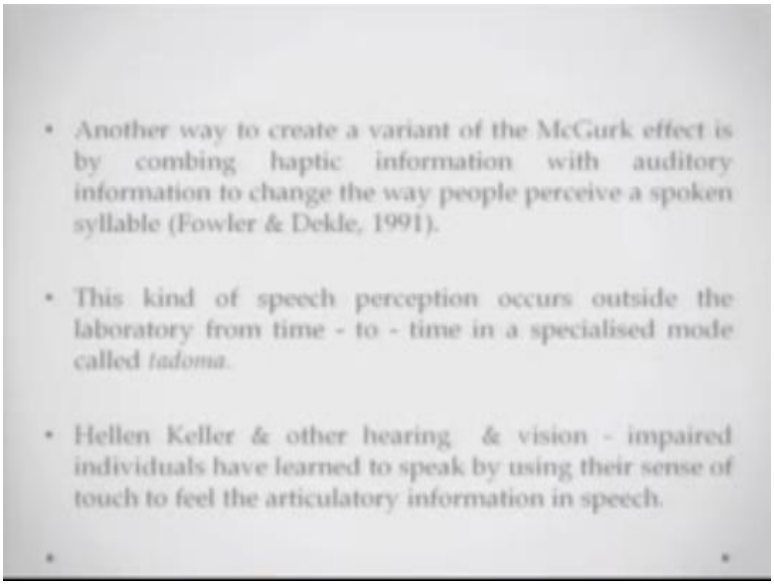
Now the McGurk effect happens because our speech perception system combines both visual and auditory information for perceiving speech rather than relying on the visual or auditory perception alone.

(Refer Slide Time: 03:26)



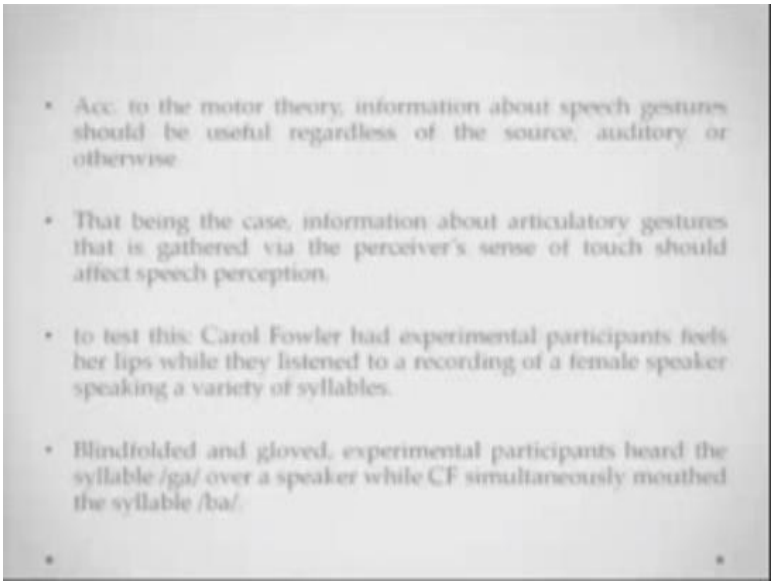
Of course the auditory information by itself is sometimes sufficient for perception worker but the McGurk effect shows that the visual information also influences a speech perception whenever it is available say for example if you are talking on phone these kind of effects are certainly not there the McGurk effect is an example of what is called the multimodal perception because two sensory modalities visual and auditory are actually being used in order to create this subjective experience of the sound.

(Refer Slide Time: 03:50)



Another way to create a variant of the McGurk effect is by combining hap tic information with auditory information to change the way that people perceive a spoken syllable this kind of perception that occurs outside the laboratory from time to time in a specialized you know module is called tadoma now Helen Keller people who are you know who are not who are visually impaired sometimes they learn to speak by using their sense of touch to feel whatever articulatory information is presented remember we are still talking about the motor theorists fish perception. And the goal is still to identify whatever gestures we will use and that is supposed to help us understand the out.

(Refer Slide Time: 04:29)

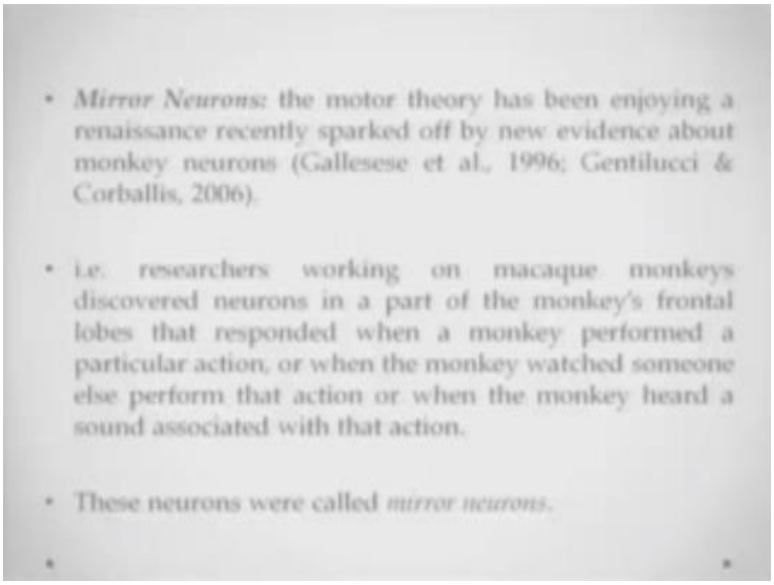


Now Carrie Fowler did this experiment and they actually she had her participant of the experiment feel her lips while they listen to a recording of a female speaker speaking a variety of syllables now blindfolded and gloves these experimental participants heard the syllable girl over the speaker why can we Fowler simultaneously mouth the syllable ba again as in the McGurk effect traditionally they reported hearing the syllable of da the motor theory explains both these versions of the McGurk effect the visual and the haptic one a stemming from the same basic processes.

As the bowl of the feature production system is not respected spectral analysis of the auditory input rather it is to figure out the set of gestures that have been you know producing these sounds the motor theorized to tell us that both of these information the visual and haptic are basically being used in order to make these judgments and that is what is leading to this combined perception that is happening under natural circumstances the visual auditory and touch informations will anyways all line up and they will not be conflicting like in the case of the mega effect video.

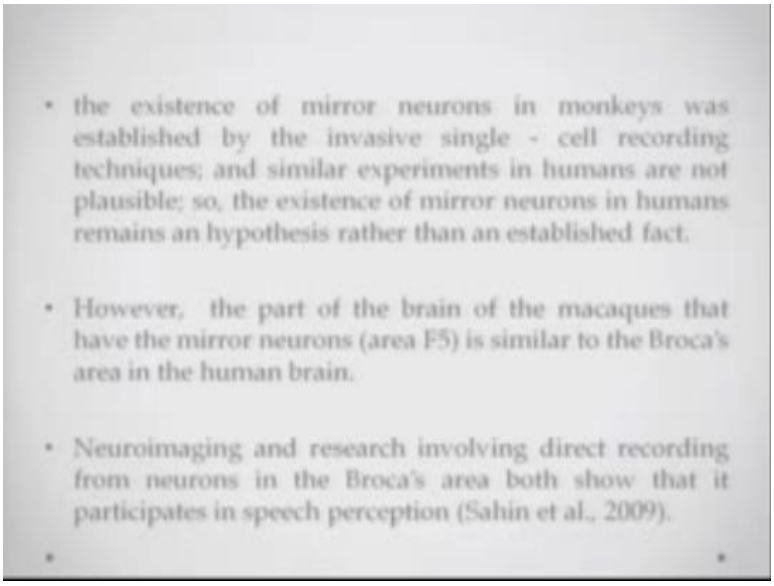
So in that sense it will all work fine but if you are creating an experimental situation like this it might lead to confusion like we saw in the perception of the syllable.

(Refer Slide Time: 05:48)



The motor to do speech perception repeatedly talks about the importance of understanding the motor aspects of speech and it has been basically a very popular theory but say for example another way was found which could help and support the meticulous fish perception this other way was a chance discovery of by researchers were working on macaque monkeys and they discovered that particular neurons in the monkeys frontal cortex responded when the monkey performed the action but they also responded when the monkey observed a particular action these neurons were referred to as the mirror neurons.

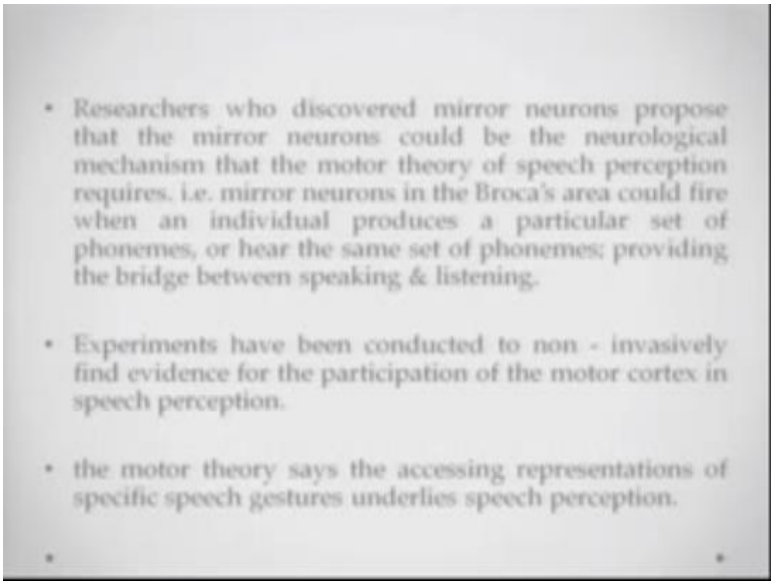
(Refer Slide Time: 06:26)



Now the existence of mirror neurons in monkeys was established by invasive single cell recordings and in that sense they have not these kind of recordings because they are not possible to do with humans there is a hypothesis that the human brain which is very similar to the monkey brain also contains this similar kind of neurons however the part of the brain of the macaques that have the mirror neurons is similar to the Broca's area of the brain which also is involved incidentally in production of speech.

So it also does something motor which is related to speech you can remember the Broca's area from the lecture I gave on the brain in behavior thing neuro imaging in research involving direct recording from neurons in the Broca's areas show that Broca's are a participates in speech perception.

(Refer Slide Time: 07:12)

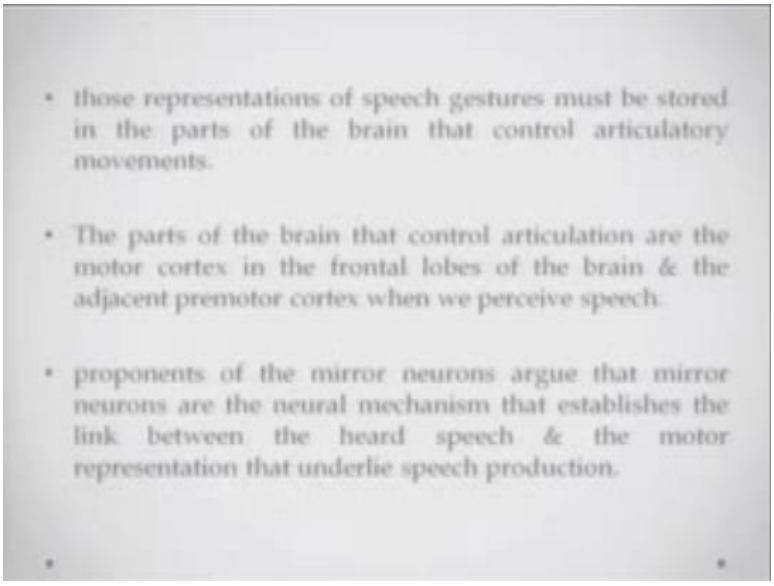


Now researchers who discovered mirror neurons proposed that the mirror neurons could be the neurological mechanisms that the motor theory of speech perception requires that is these mirror neurons in the Broca's area could fire when an individual produces a particular set of phonemes or here the particular set of phonemes and providing the bridge between speaking and listening.

So if you are speaking a set of warnings the mirror neurons in the Broca's area are firing and if you are listening to the particular sound the same neurons are firing again so that there is the same neuro logical structure that is involved both in speaking and listening there have been experiments conducted to non-invasively find evidence for the participation of mirror neurons or the participation of the motor cortex in speech perception obviously the motor cortex has been known to you know participate by speech production is there but remember we are talking about speech perception.

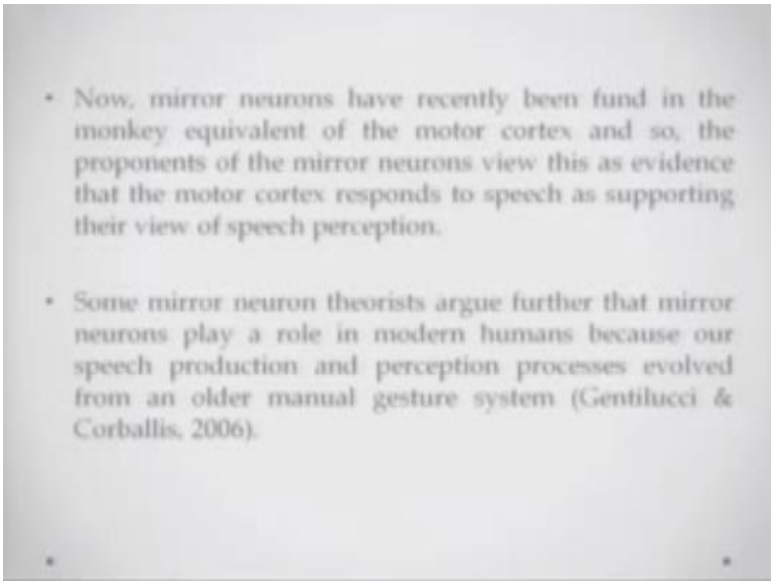
The motor theory says that accessing the representations of specific speech gestures must underlies speech perception.

(Refer Slide Time: 08:10)



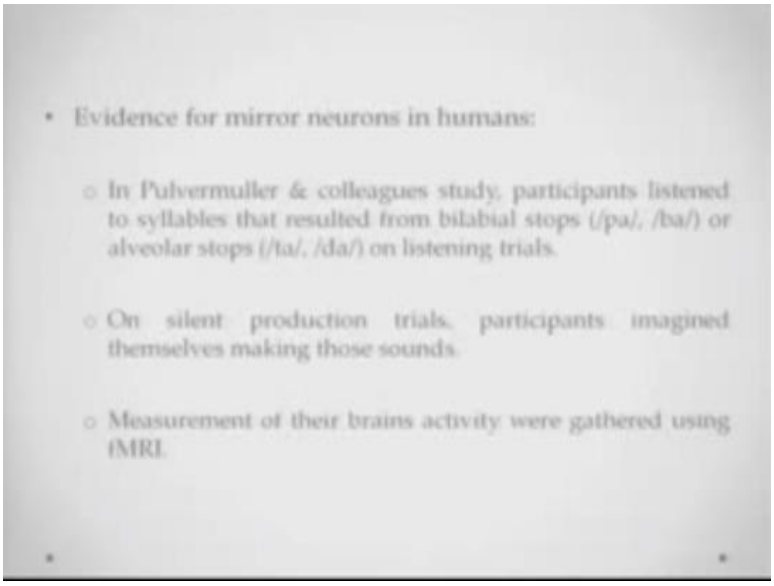
The representations of speech gestures must be stored in the parts of the brain that control articulatory movement wherever those fascism rings which are involved in making these movements must also be the parts of the brain that store this information about what gestures have been used the parts of the brain that control articulation are the motor cortex and the frontal lobes of the brain and the adjacent premotor cortex that are used when we perceive speech now the proponent of the mirror neurons are cued that mirror neurons are able are basically the neural mechanism that will establish the link between heard speech and the motor representations that underlies speech production.

(Refer Slide Time: 08:48)



Mirror neurons have recently been found in the monkey equivalent of the motor cortex as well and so proponents of their mirror neurons view this as evidence of the fact that all motor neurons respond to you know speech perception as well so mirror neuron theorists argue further that mirror neurons also play a role in modern humans because our speech perception in production processes are evolving from a manual gestures a theory about language evolution is that because we are using manual gestures initially. That is why our mirror neuron or motor cortex is involved in perception on production speech as well.

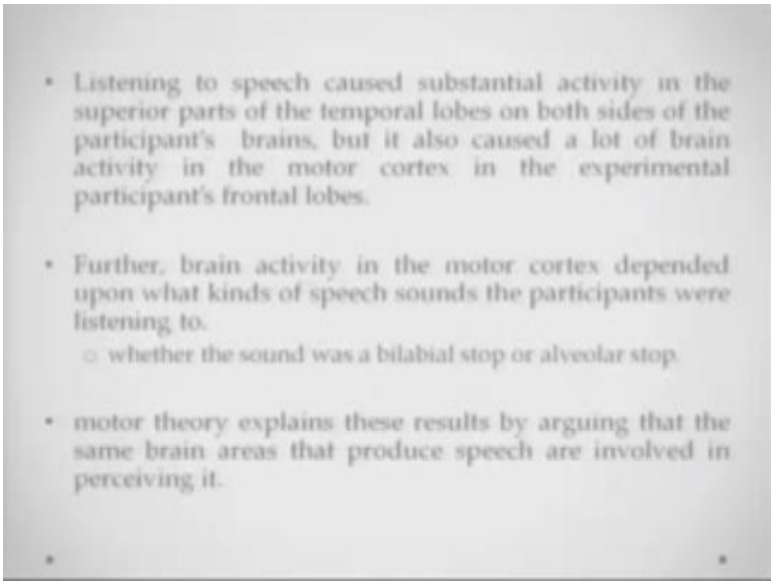
(Refer Slide Time: 09:23)



There is a story for another day but let us examine some evidence about this involvement of mirror neurons and the motor cortex in human speech perception so Palmer Muller and caliz they conducted a study where in participants were to listen syllabus that resulted from either by pulver stops like particular or ba or alveolar stops like to or da on listening trials on silent prediction trials these participles imagine themselves making these sounds, so there is production and rejection both happening measurements of the brain activity were gathered using fMRI if you remember fMRI basically measures the amount of flow of oxygenated blood to the areas that are involved in particular cognitive tasks.

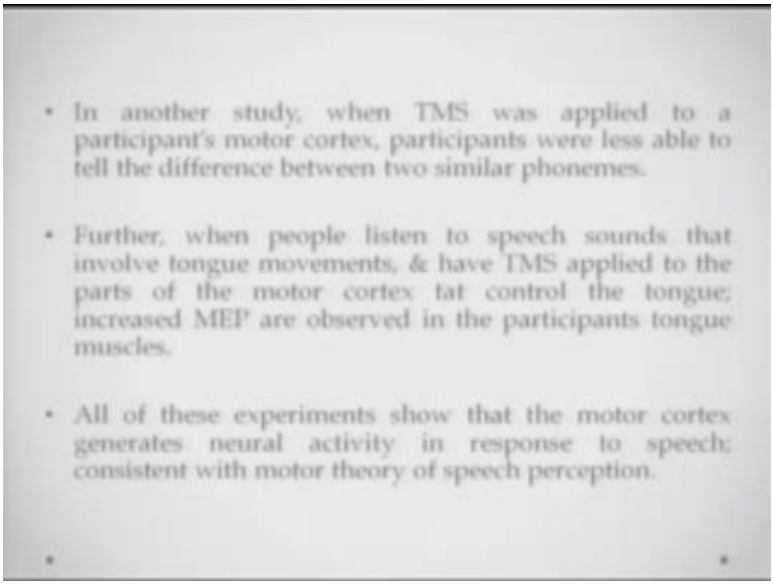
Now listening to the speech caused substantial activity in the superior parts of the temporal lobes on both sides of the partisans brain but it also caused a lot of activity in the motor cortex in the experimental part sense frontal lobes further brain activity in the motor cortex depended on what kind of speech sounds the partisans were listening to, so there were different activations depending on whether the sound was a by label stop or an alveolar stop this result is explained by the motor theory.

(Refer Slide Time: 10:34)



And they say that you are the same areas that produce the speech are involved in perceiving it.

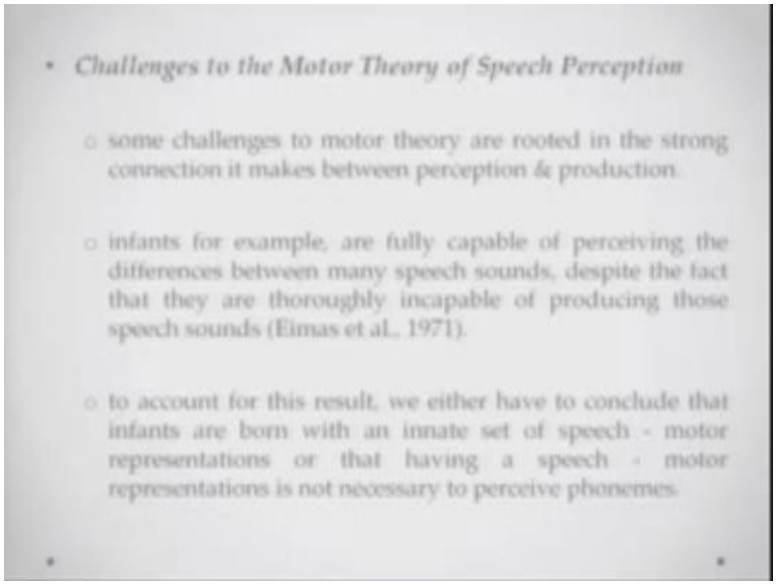
(Refer Slide Time: 10:40)



So it is one kind of a confirmatory evidence in another study when TMS was applied to a participant motor cortex participants are less able to tell the difference between two similar for you so if those areas are not working your perception and understanding of these two phonemes might also be attenuated further when people listen to speech sounds that involve tongue movements and have TMS applied to parts of the motor cortex that control the tongue increased motor evoked potentials are observed in the person's the.

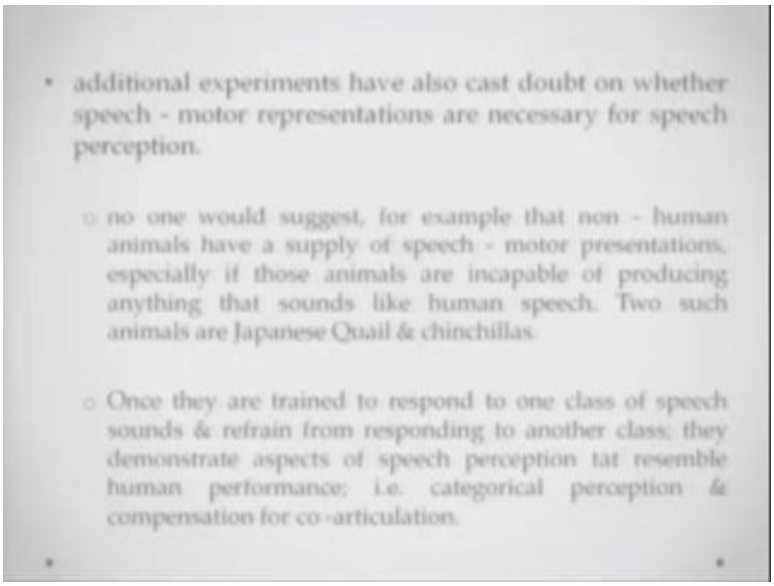
So there is some processing happening there as well all of these experiments put together show that the motor cortex indeed generates neural activity in response to listening speech consistent with what the motor theory has been seen.

(Refer Slide Time: 11:21)



But there have been some challenges to the motor theory of speech perception as well some of the challenges are rooted in the link that they make between perception and production you can say for example infants are fully capable of perceiving and understanding speech despite the fact that they are thoroughly incapable of producing these speech sounds to account for this we will either have to conclude that the infants are born with this innate set of speech motor representations or that having a speech motor representations is not necessary for perceiving phonemes. If we accept the latter we are kind of violating what the motor tell you are saying.

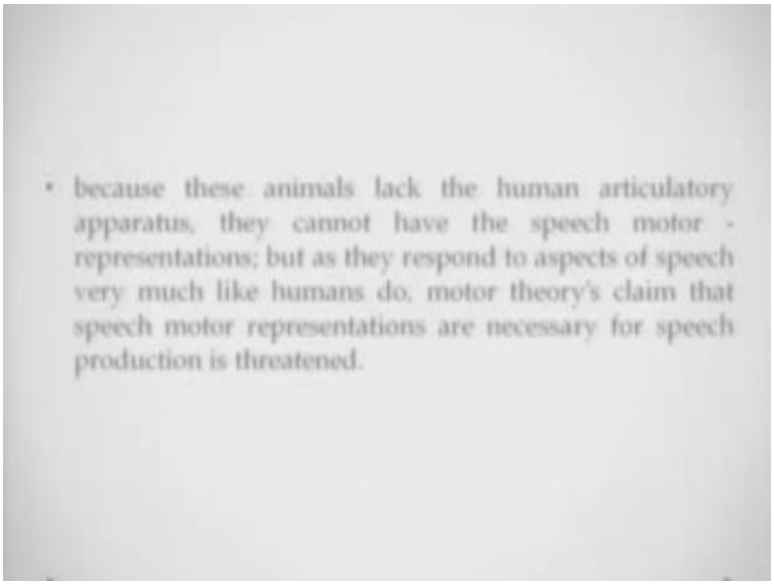
(Refer Slide Time: 11:57)



Additional experiments have also cast doubt on whether speech motor representations are necessary for speech perception now no one would suggest that non-human animals have a supply of speech motor representations you know which are pertain to human speech sounds but it has been found that animals like the Japanese quail and chinchillas they also respond to you know particular class of screech sounds and the frame from responding to other class of speech sounds telling us that they have this aspect of speech perception and they can differentiate between these different sounds.

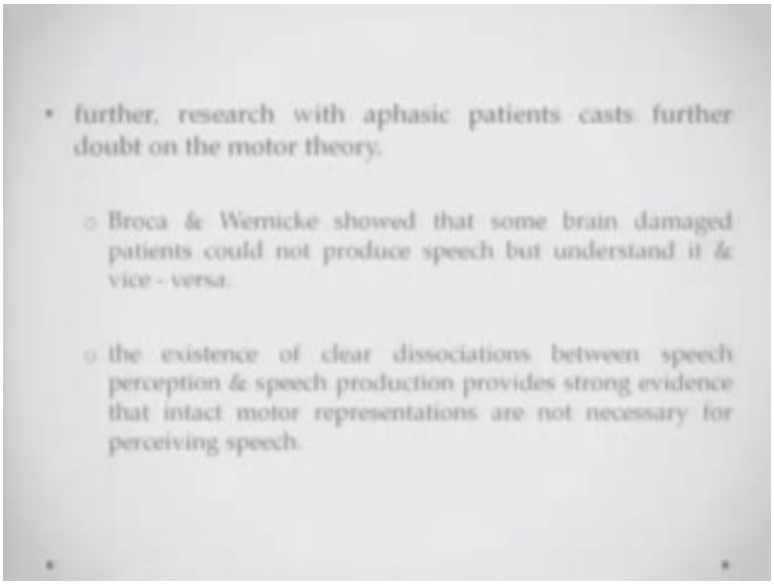
Now the motor theory would say that they also know which gestures are involved in producing these which kind of is a non-starter.

(Refer Slide Time: 12:35)



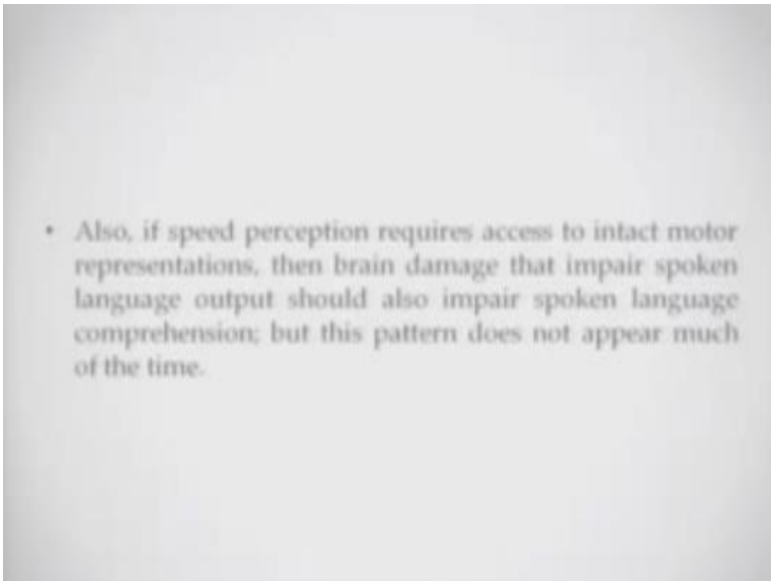
Because these are all slag the human article articulate the operators they cannot have the speech motor representations as I was saying but as they respond to these different aspects of speech very much like humans who do the motor theories claim that speech motor representations are a necessary part of speech perception is kind of weakened.

(Refer Slide Time: 12:54)



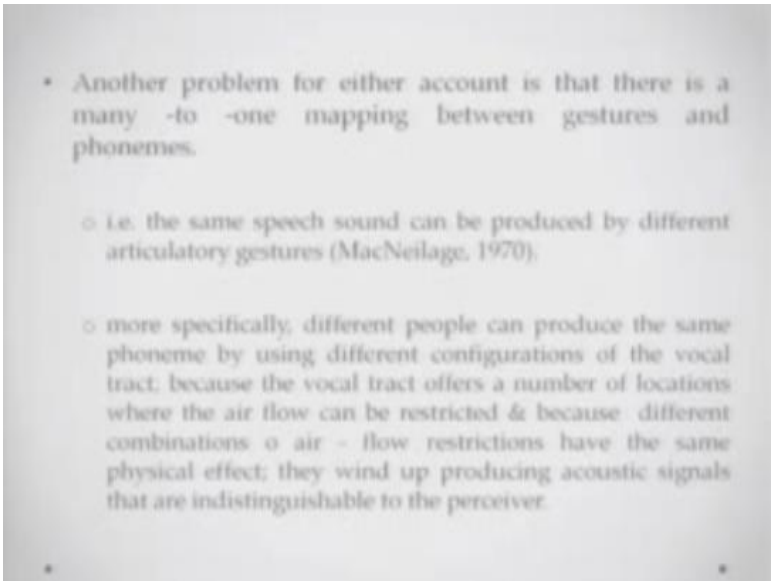
Further research with a physic patients also casts doubt on the motor theory say for example both Broca and Wenicke they showed that some brain-damaged patients could not produce speech but understand it while others could understand speech but not produce speech if you were to listen to the motor theory schemes then you will say that you know this is not really possible the existence of these cleared associations between speech perception and production systems provides strong evidence against the account of the motor theory.

(Refer Slide Time: 13:22)



Also if speed perception requires access to intact modern representations then brain damage that impair spoken language output should also impair spoken language comprehension as I was saying.

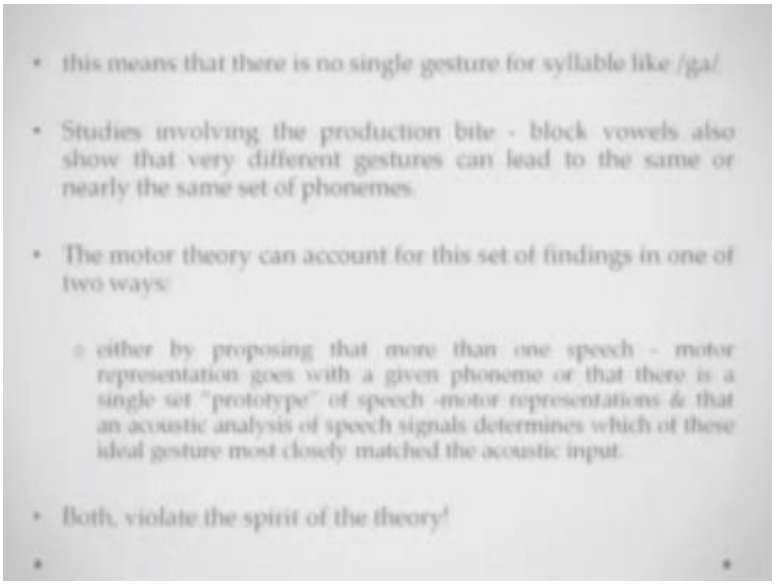
(Refer Slide Time: 13:32)



Now the final problem or another problem about this account is basically that one has to say the same speech sounds can be produced by different articulatory gestures it has been shown in a study by MacNeilage in 1970 more specifically different people can produce the same phoneme by altering configurations of the vocal tract because the vocal tract offers you know a number of locations where the air flow can be restricted and because different combinations of airflow restrictions have the same physical effect they wind up producing similar requested signals and which are indistinguishable to the perceiver.

So perceiver might be listening to two different know the same sound being produced by different articulate as and different kinds of gestural scores then it becomes very difficult you know to say that a single gesture is responsible for a sound like you can say then there are multiple gestures now studies involving and there is these interesting experiments done by a you know called by the block bubbles when people are keeping something in their mouth and then producing the sound and the parcels can still understand it says that you know the motor theory is kind of weakening here.

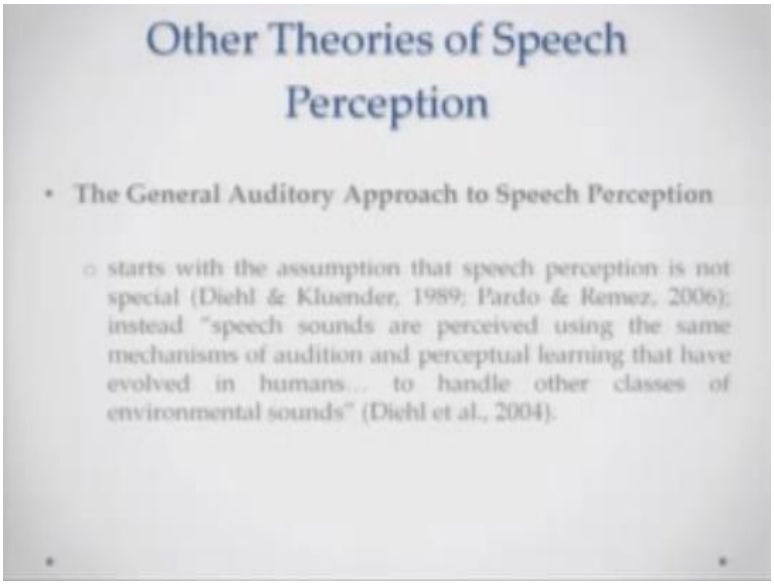
(Refer Slide Time: 14:41)



The motor theory if given a chance will try to account for this set of findings in one of two ways they will say either one more than one speech motor representation goes with a given phoneme or that there is a single set prototype of speech motor representations and that an acoustic analysis of these signals determine which of these are ideal gestures most closely master acoustic output now if you see this closely that both of these things will violate the spirit of what the motor theory is saying I just repeat this once more.

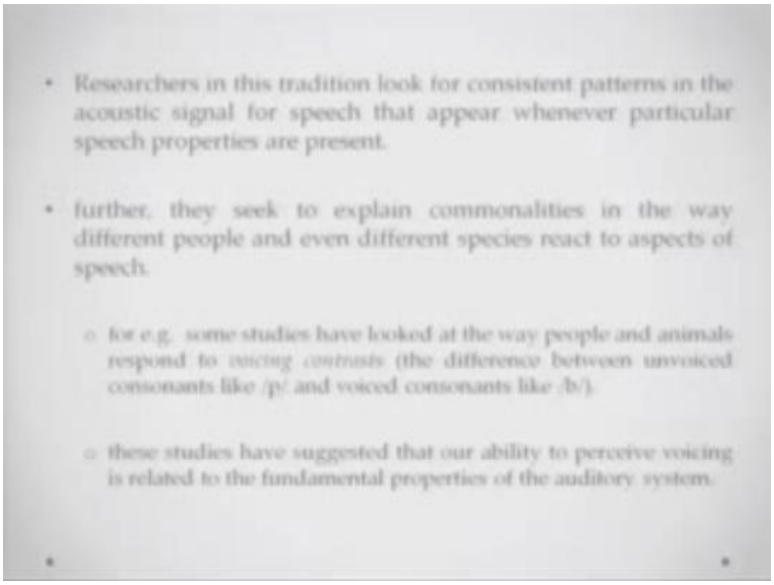
So two reasons for they can give is they can say that the more than one speech modern representation can go with a given phoneme or they can say that there is a single prototype which can be matched to any given gesture both of these things are contradictory to what the motor theory of speech perception has originally claimed in that sense it is kind of weakened for M0 is playable to explain all the findings.

(Refer Slide Time: 15:33)



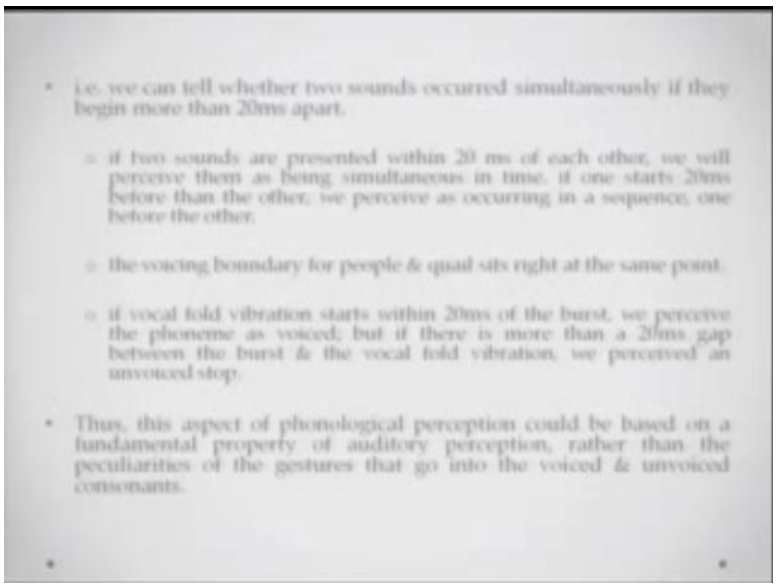
So when there are you know flaws with a particular big theory there are other theories that jump in the other important theory of switch perception is the general auditory approach to speech perception now the general auditory approach basically says it starts with the assumption that speech sounds must be perceived or are perceived using the same mechanisms of audition and perceptual learning that have evolved in humans to handle all other classes of sounds so it says that speech perception is not really special you understand speech as you understand all the other sounds research is in the general auditory tradition look for consistent patterns in acoustic signals for speech.

(Refer Slide Time: 16:11)



That appear of enable particular speech properties are present further they seek to explain commonalities in the way different people and even different species react to aspects of speech for example some studies have looked you know at the way in which people and animal respond to what are called voicing contrast say for example the example I was talking about/ n ba these studies have suggested that our ability to perceive voicing is related to the fundamental properties of the auditory system and not really a special module that was proposed by this moderate theory.

(Refer Slide Time: 16:48)



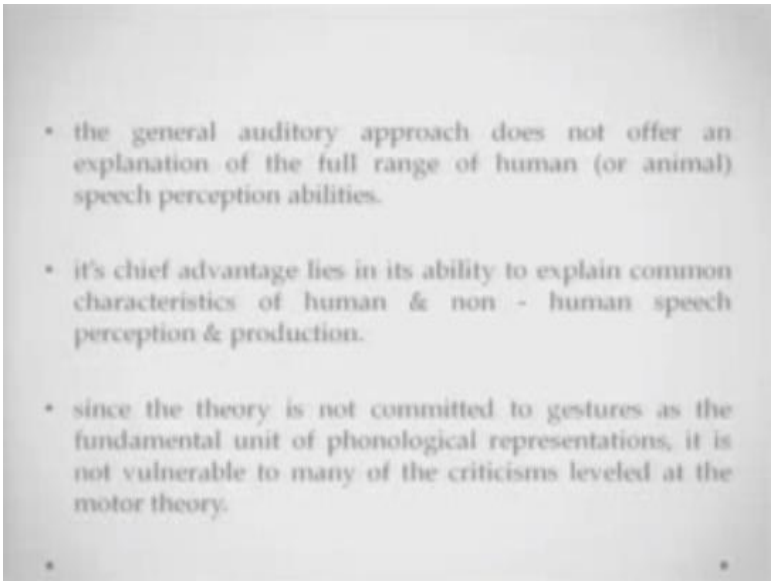
We can tell whether two sounds occurred simultaneously if they begin more than 20 milliseconds apart so it is just a matter of time and not really that it is a special speech signal so if two sounds are presented within 20 milliseconds of each other we will perceive them as being simultaneous in time if one starts 20 missing before or after another we start perceiving them in one before the other in differentials the voicing boundary for people and quails by the way sits right at the same difference of 20milliseconds.

Now you can see that this is this generality between human auditory perception system and the queries auditory perception system by the way quail is a particular bird if the vocal fold vibration starts within 20 milliseconds of the birds we will perceive the phoneme as voice if it starts after 20 milliseconds we will proceed as unwise and bird the same example you have been talking about.

So this aspect of this general aspect to phonological perception then could be said to be based on a fundamental property of auditory perception rather than the peculiarities of gesture that go into voice and invoice, so if I if I were a auditory perception system I do not really need to keep tab of whether it is voice or not voice I will just keep track of time and if the time is sufficiently

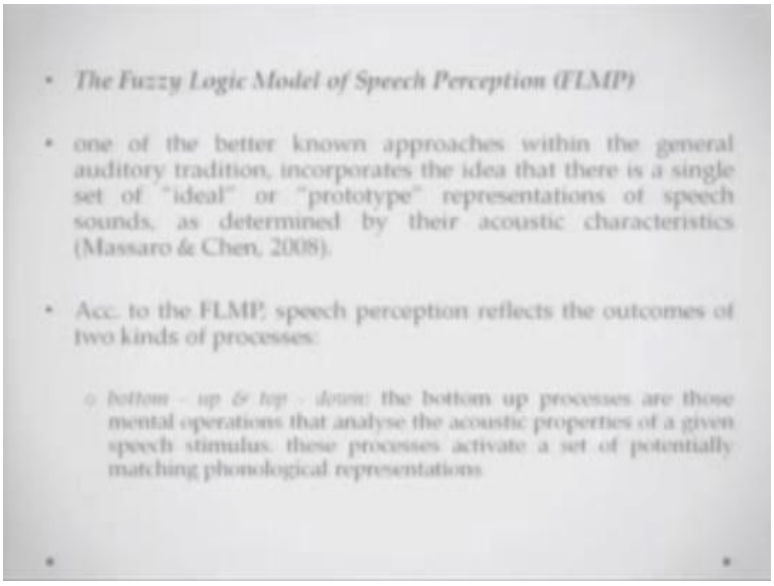
apart I will treat them as different I can be a quail or any other animal and still do this task perfectly.

(Refer Slide Time: 18:09)



The general auditory approach also does not offer an explanation of the full range of human or perception abilities but its chief advantage lies in it is ability to explain common characteristics of human and non-human speech perception and production since the theory is not really committed to gestures per say it as a fundamental unit of phonological representations it is also not vulnerable to the kind of flaws that were associated with motor theory which had said that speech and you know perception and production link is necessary.

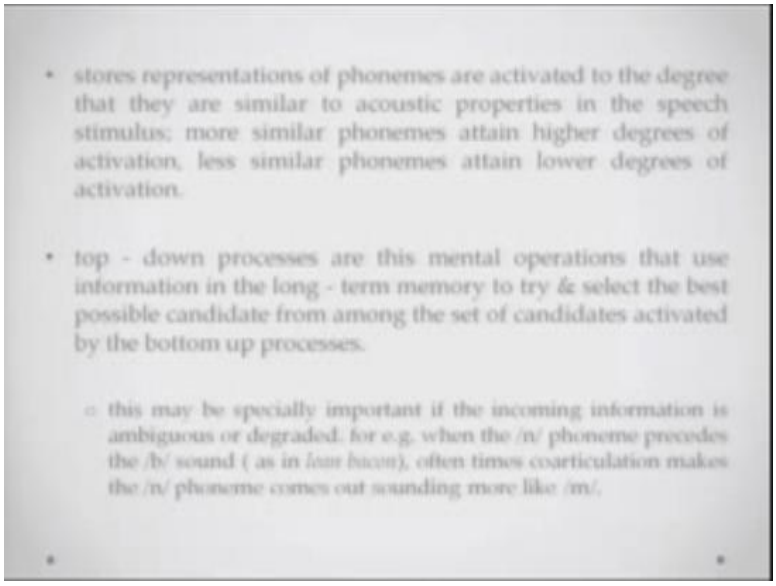
(Refer Slide Time: 18:39)



Another kind of model a different kind of model of speech perception that is a more popular and more recent is the fuzzy logic model of speech perception that is different from both the general auditory and motor theory of speech perception in that it says that you know a better approach or a general auditory tradition is that there is a single set of ideal or prototype representations of speech sounds as determined by their acoustic characteristics.

Now according to this fLMP model speech perception reflects the outcomes of two kinds of processes there are bottom-up processes which are the mental operations that analyze the acoustic properties of the incoming speech stimulus and there are top-down processes which activate a set of potentially matching phonological representations so imagine if you are listening to a sound if somebody is speaking something to you one set is already analyzing this incoming signal in terms of very basic physical characteristics and the other set of operation is trying to look into your memory as to whatever information about this particular sound you have and they meet somewhere in the middle.

(Refer Slide Time: 19:30)



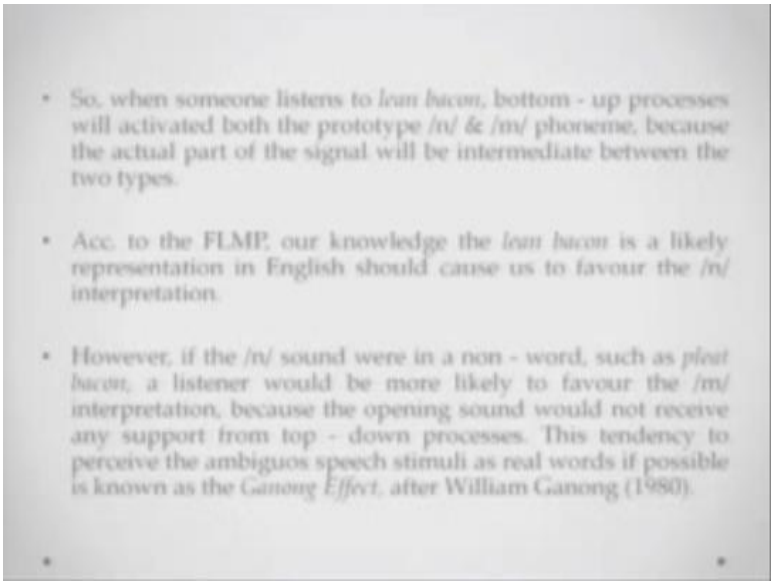
And then you can understand whatever sound is produced now it needs to specify that there are a lot of stored representations of phonemes and they are activated to different degrees and they are similar to acoustic properties in the speech stimulus more similar phonemes attain higher degrees of activation less similar for phonemes achieve lower days of activation.

So if you listen to a particular sound all those sounds similar to this incoming sound that you have heard of we all get activated and you know they will potentially be matched against this incoming sound the top-down processes are these mental operations that use the information in long-term memory to try and select the best possible candidate from among the set of candidates activated by the bottom-up processes.

So the bottom-up analysis has activated so many candidates which you can potentially match to any of this incoming information and the top-down mental operations are actually doing this matching and they are trying to select the best possible candidate which will match this incoming stimulus once that match is made you understand that particular stimulus if you have heard of it earlier.

This may be specially important if the incoming information is ambiguous or degraded let me take an example say for example when the N phoneme precedes the B sound say for example in the case of lean bacon and if I am saying it very fasty in bacon oftentimes the co articulation because N and B are coming so close together a lot of people might perceive this as lean bacon if I am saying and in B so close together but lean will be perceived as lean and a lot of people will report saying a hearing a lean bacon. If I if you know just want to do it you can say this very fast again and again to yourselves and then you can see that what is happening.

(Refer Slide Time: 21:31)



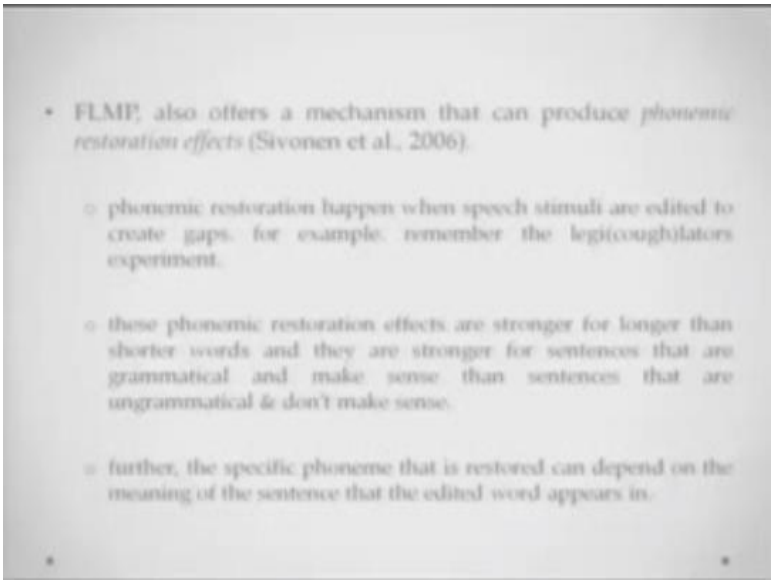
Now see when somebody listens to lean bacon bottom-up processes will activate both the prototypes it will be an activate end and it will activate m so they could be lean bacon and lean bacon both will be activated according to the fuzzy logic model of speech perception our knowledge that lean bacon is actually a meaningful word and is a likely representation in English will cause us to favor lean bacon over lean bacon and that is how we understand whatever has been said.

However if the n verse say for example in a non words such as pleat bacon and if I was a clean bacon or something like that a listener will be more likely to favor the M interpretation because

the opening sound would not receive any sub from top-down processes because there is nothing I more that is called clean bacon so I will go by the bottom-up analysis and I will pick up something that is coming from there.

Now this tendency to perceive ambiguous speech stimuli as real words is possible you know a real word is actually called the gang effect of named after William Gong in 1980s.

(Refer Slide Time: 22:39)

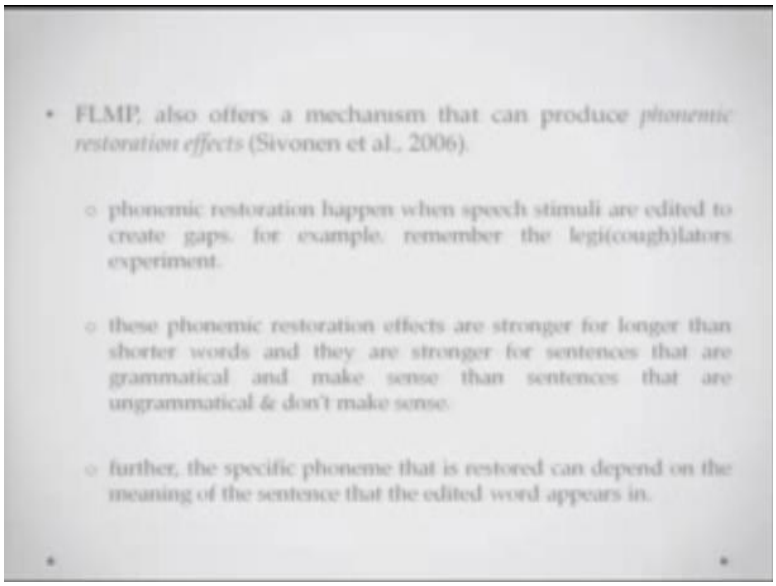


FLMP also offers a mechanism that can produce what is called phonemic restoration effects so for me the restoration basically is when sweets only are edited to create gaps say for example if you remember I was talking in an earlier class about legislators and there was this experiment and people who are healing the word legislators on the head phones and there was a coughing sound where the S is their so leggy cause and nature's and experiments what people do is people do report hearing.

This even if though there is no s presented in the signal itself what is happening here is that you are using your previous knowledge of the word legislators to fill in that s and you do it so well that you are convinced that they are both ns in that signal now these phenomena serration effects

are stronger for longer than for longer worse than shorter words because these make much more sense and are more grammatical than the shorter words which are ungrammatical and they might not make sense further the specific phoneme that is restored can also depend on the meaning of the sentence that is edited okay.

(Refer Slide Time: 23:45)

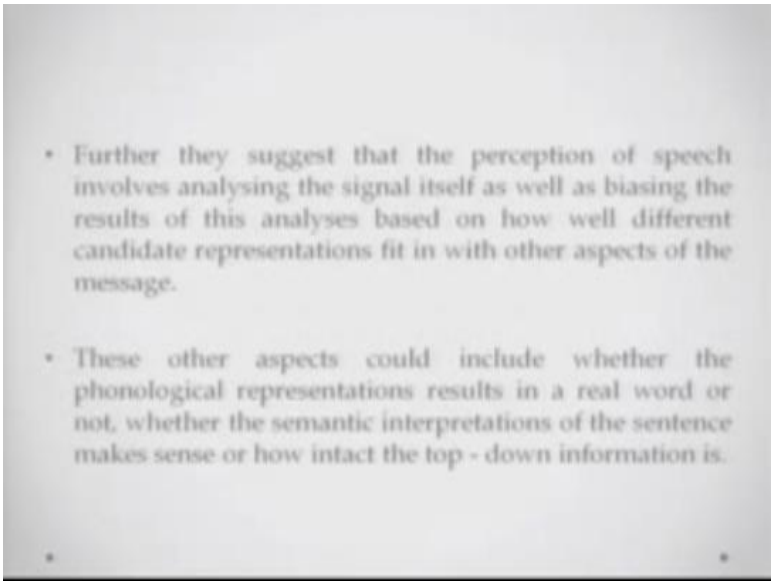


Similarly if there is a different thing if you are hearing the coughing sound let us say the Wagon losses and there is a coughing sound e you will most likely hear the word w so because you will hear the wagon lost it lost it is w which is the more probable thing but if you hear the circus train has lost and the circus has lost a trained eel and there is a coughing sound before e you will probably think that you know you are talking about s your hearing and s there because the circus generally has animals like seals and you say that maybe you know the sound that was there was s.

So how you will do the phonemic restoration actually depends on the context that is built in the sentence previously research involving ERPs have shown that nervous system does register the presence of the coughing noise very soon after it appears in this similar so what is happening is you are registering the cuffing sounds very early but you are doing all the mental processes

possible to fill up that space that is created by the coughing sound all of these cities that there are a variety of possible you know sources of top-down information and these in various source of information and affect the way and a caustic signal is perceived.

(Refer Slide Time: 24:56)

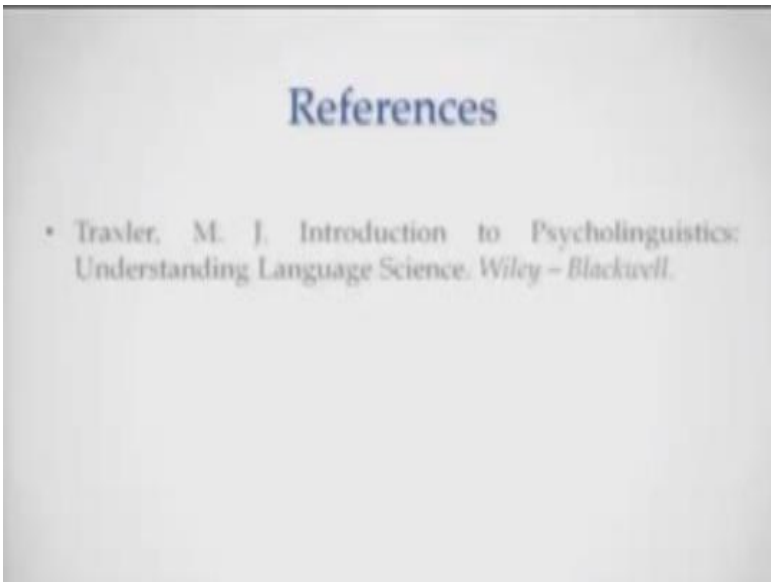


Further they said is that the perception of speech involves anything that involves analyzing the signal as well as biasing the you know results of this analysis based on how will different candidate representations fit in with other aspects of the message so you might have you know be biased to hear wheel or seal or anything else but because I was talking about the wagon you would fill it with veal because that fits in better and you will not fill s there because you know and you will fills in the circus example because veal does not fit their better.

So you are doing this different kind of calculations and your online correcting your perception or speech there these other aspects could include whether the phonological representations result in a real word or not whether the semantic interpretation of the sentence make sense or not and how intact the top-down information is if you do not remember the word exactly if you do not remember what animal the circus have if you do not have any language any knowledge about the animals or the circus.

Then you might not be able to fill the example there with s as you see if you are living in a particular swear in circuses do typically have seals then you will create the sentence that the circus has lost between C this is all about speech perception that we will be talking about.

(Refer Slide Time: 26:09)



In the next class where you begin talking about attention as a common effect thank you.

Acknowledgement

Ministry of Human Resources & Development

Prof. Satyaki Roy
Co – ordinator, NPTEL IIT Kanpur

NPTEL Team
Sanjay Pal
Ashish Singh

Badal Pradhan

Tapobrata Das

Ram Chandra

Dilip Tripathi

Manoj Shrivastava

Padam Shukla

Sanjay Mishra

Shubham Rawat

Shikha Gupta

K.K Mishra

Aradhana Singh

Sweta

Ashutosh Gairola

Dilip Katiyar

Sharwan

Hari Ram

Bhadra Rao

Puneet Kumar Bajpai

Lalty Dutta

Ajay Kanaujia

Shivendra Kumar Tiwari

an IIT Kanpur Production

@copyright reserved