The Psychology of Language
Dr. Naveen Kashyap
Department of Humanities of Social Sciences
Indian Institute of Technology-Guwahati

Module No. #01
Lecture No. #08
Speech production – II

Hello friends, Namaste. Welcome back, to this course, the Psychology of Language. And, this is Lecture number-8, in the series. Now, Lecture number-7, focused on, the Production of Speech. And, Lecture number-8, is a continuation of that particular lecture, where we look at, how Speech is understood. What are the Theories of Speech production, which suggest, how speech is produced? And, we will also look at, Speech production, from a social aspect like, Infant Babbling, and also the various type of speech disorders.

Now, before we go, into the details of today's lecture, let us take a journey back, of how we started. So, we started off, the journey, by explaining, what is the need, for this course of psychology of language. So, we started off by explaining, what is language, and defining the meaning of language, in terms of, the very basic system of language, which is called the, Animal communication. So, we looked at, Animal communication systems, and the characteristics of that systems.

We also looked at, why animals communicate. And, later on, we focused on the Human language system. We looked at, how Human language system, is different from the Animal language system. We focused on, how the Human language is built up, right from the speech sounds, to the morphemes, the simple words, or word endings, the complete word formation, the sentence, the rules for sentence, the structure of sentence. This is, syntax and grammar. And, also looked at, how the discourse happened. That is, talking between people.

Now, since we have looked at, the existence of language, we also dealt, a little bit into the history of language, in that section. And so, we looked at, how language evolved, from Neolithic Humans, from Monkeys, Chimpanzees. So, we looked at, the story of how a language evolved, in this present form. We looked at, how pidgins are waystations, which explained how, a language developed slowly, from the Protolanguage which the, Chimpanzees and Gorillas spoke, or the early Humans spoke, to the present form of language.

So, that is what, the story was about, in the first two lectures, of looking at, language, and its history. Now, obviously, the second section focused on, the Research in Language. So, we looked at, the research process of doing language. We started off, by explaining a piece of research, which is very famous in language research, which was the development of the N400, which is event related marker, for the semantic categorization.

So basically, semantic errors happen, we see the N400. And, so by describing that, I describe to you the process of doing, experimentation in language. We went into details of how, hypotheses are made, how problems are constructed, what is the way of doing research in language, what are experiments, how do you do the experimental design, what are experimental groups, and the variations of designs like, the Between subject and Within subject designs.

And, that was the focus on, the say the third and fourth lecture, where we looked at the, research design of doing language, or how do we do research in language. In addition, we also looked at, certain basic experiments in language. And, we carried those experiments, we use those experiments as model, to explain the research process in in language studies. We looked at, how the variables are defined, and what kind of variables are chosen, for doing research in the language.

And lastly, we looked at, the language and the Brains, areas of the Brain, which is related to language. And, in particularly, we detailed the idea of, the Wernicke area, and the Broca area. The Broca area known for, production of speech. And, the Wernicke area, which is, perception of speech, understanding meaning. So, that is what, we did in, Lecture number 3 and 4, we looked at, the details of doing research in language. Obviously, once these two were clear, we started dwelling more into, the idea of language.

So, the third section, we jump into the third section, which is le Lecture number 5 and 6, we looked at how, Speech is perceived, how do you perceive Speech, or how do you listen to Speech. So, we looked at, the Auditory system in detail. We looked at how, sound behaves as a Wave. And, we looked at the, function of the Auditory canal, and the function of the Cochlea, the Basilar membrane, how they are arranged, and un in what way they perceive, or they are able to extract speech, from sound waves.

We also looked at, the speech stream, what is speech composed of, in terms of spectrograph. So, we looked at, speech is continuous. And so, we looked at the way, the speech is manifested, or it is spoken. We looked at, how vibrations how a vibration is expressed, in terms of, fundamental frequencies. Vibration of the vocal fall fold, and what is Phonation, what is Prosody, and those kind of things, we looked at. And, how Prosody and Phonation, they help us in, perceiving or understanding speech.

We looked at, the various characteristics of spoken language, which is, how vowels and consonants are expressed, in the speech stream. What are Formants? And, things like, Sorants, Formants, Fricatives, Plosives, these are all, characteristics of the speech. So, we looked at, those kind of speech streams. We focused on, how speech is perceived, in categorical. Speech is not perceived, in its in in in word by word, but speech is perceived, categorically.

So, that is all the another thing, that we studied. And, we use Warren & Warren experiment, to un understand, how speech is perceived, categorically. And lastly, we looked at, the development of Speech perception, how speech perception developed. And, we looked at, how Infants develop the speech, or develop the speaking ability. So, we looked at evidences to that. And, the various kind of things, related to Infant speech perception. And lastly, we focused on, various theories of Speech perception.

So, we looked at, 3 Basic Theories. We looked at, the Motor Theory, which believes that, speech is perceived through Motor actions, or the various systems involved in speech. We looked at a, directly opposing theory, which is the Auditory Framework Theory, which opposes this idea, that Motor perception, or gestures, or the Motor system, which produces speech, the articulators, which produces speech, is anywhere involved in Speech production.

And so, the General Auditory Framework said, this speech is like any other acoustic signal, it is not special. So, they provided evidences for that. And lastly, we focused on, the idea of Direct realism, which basically talks about, the idea, this speech signal itself, has all the information necessary, and so we do not need, another theory of Speech perception. Now, once you had understood, how speech is perceived, how speech is listened to, we thought about going, into a little bit detail, about those apparatus, which actually produced speech.

And so, that was Lecture number-7, where we looked at, those apparatus, which produced speech. So, we started our journey, by explaining the, vocal tract and Speech perception. So, how vocal tract actually, leads to the, Phonation or the speech stream. So, we looked in detail, on those things of, how the vocal tract is made, and how it produces speech. And, how the Glottis box, they produce speech.

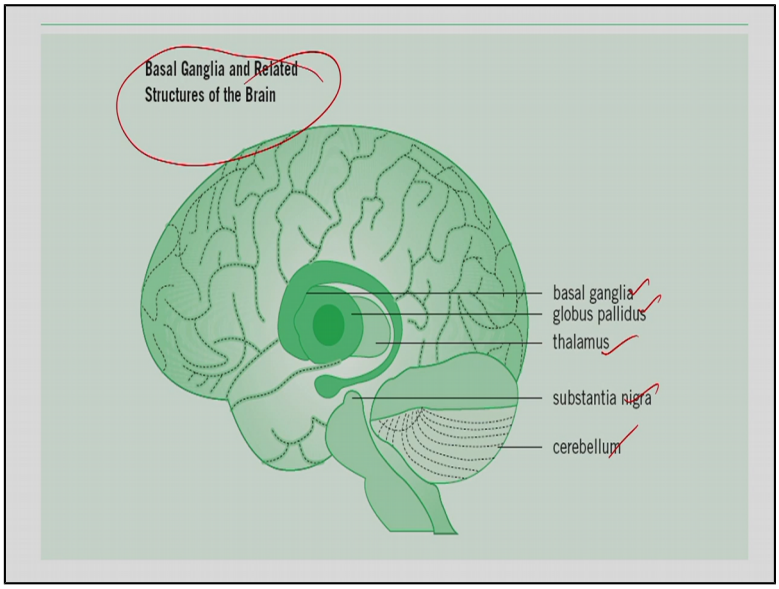
So, we looked at how, consonants and vowels are produced. And, we looked at, the various manners of consonants. So basically, consonant is produced, by restricting the air, which is coming from the vocal fold. So, restricting in, some way. So, there are three ways of restriction, the manner of articulation, the voicing, and the place of articulation. So, we looked into detail, how consonants are produced. And, if we do not restrict the air, which is coming out of the vocal fold, we generally get a vowel.

And so, vowel is basically produced, in coordination, with the Jaw, the Tongue. And mainly, the Jaw and the Tongue. So, these positions. And, we looked at, how different vowels are produced, and the Triangular Theory of Vowel. So, we looked into, details on to, those things. We looked at, how Plosives and Affricates are produced. And, several other interesting facts, related to producing of speech, the vowel, the consonant, and also the idea of, what is a Diphthong, and so on, and so forth.

We additionally, we looked at the, speech areas of the Brain. So, we looked at the Geschwind-Wernicke model, where we talk about, how speech is produced, or what areas of the Brain is, involved in Speech production. So, not only the Wernicke and the Broca's area, we also looked at, several other Cerebral areas, which are involved in, production of speech. So finally, we ended our lecture on the, Minimal Network of Overt Speech production. That is what, we are looking at, the three minimal areas.

So, we have the starting mechanism, which starts with the Supplementary Motor cortex, and enter the Cingulate, we have the Phonetic plan generation. So, this is done by, Broca area, and Primary Motor cortex. And, then you have, the coordinating movement, which coordinates the movement of the Vocal articulators. And, that is done by, Cerebral Basal ganglion, Thalamus. And so, this is where, we ended.

(Refer Slide Time: 11:32)

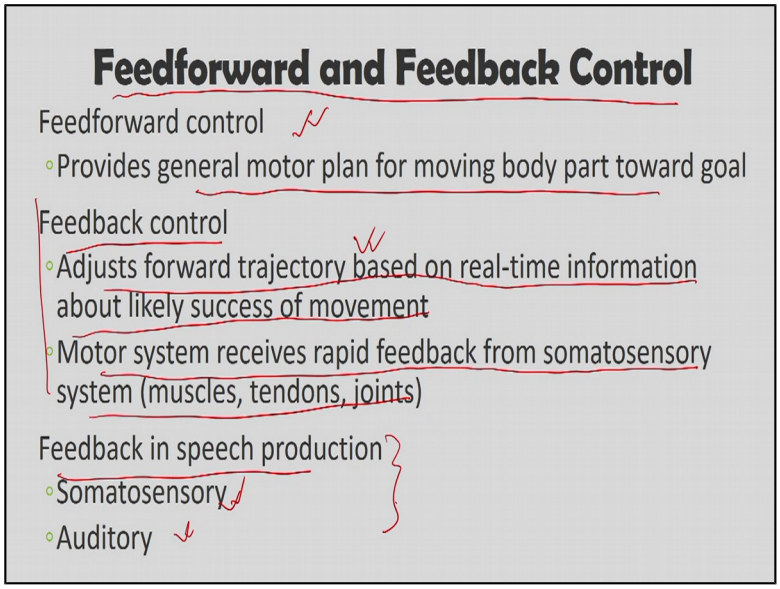


In today's lecture, what we are going to do is, we are looking at, models of Speech perception. So, this is from the last lecture, where we looked at these areas, Basal ganglia, the Globus Pallidus, the Thalamus, the Substantia Nigra, and Cerebrum. And, these are the areas, which are related to, production of speech. So, in today's lecture, what we are going to see, is various models of Speech production, and the social aspect of speech.

So basically, if you look, into the world around you, for any movement, let us say, any movement, if I want to grab a cup, now there are several movements, which are there. A program has to be, Ventral program has to be created, a Ventral map has to be created. It signals, to the hand, and the hand moves, from one position to another, the hand moves forward. And, there are two systems, a Feedback system, which tells the hand, where to go, and the Feedback system, which tells the hand, where to stop.

So, if the system says that, move the hand forward, the hand should not overshoot the cup. And so, there is a Feedback system, which actually says that, where it should stop, and it gives you the accuracy of, holding the cup. Now, similar to this program, the speech system, its movement are similar to, limb movements. So, limb movement for grabbing a cup, is very similar to the system, that are available for the holding of, or the speech articulation.

(Refer Slide Time: 12:52)



And so, they involve, the speech articulators, the instruments, or the parts of the body, which produce speech, they are generally, they involve, the Tongue, the Jaw, and the Lips, at various positions. So, Tongue, Jaw, and Lips, are the main articulators, with the vocal fold, for producing speech. Now, the process is, so rapid basically, these three things are the one, which you can see, and other areas, other parts of the body, are also involved, which you do not actually see.

So, we have here, the Feedback, and the Feedforward, control systems of speech. So, how our Feedback, and Feedforward models, control speech, or how speech is produced. Now, the Feedforward control, or the Feedforward processes, they provide the general Motor plan, for moving a body part, towards a goal position. So, the Feedforward systems, they give the general Motor plan, for moving the body part, towards the goal.

And, Feedforward systems are those systems, which actually tell you, how to start the action. Now, similarly, in the Feedback system, they adjust the forward trajectory, based on real-time information about, likely success at the movement. So, if I start moving my limb, or if I start speaking, in the same way, the action or the signal to start speaking, is done by the Feedforward system. But, Feedback system will, adjust the articulators, of what to speak, and what not to speak. And, this is based on, real-time information, which is coming.

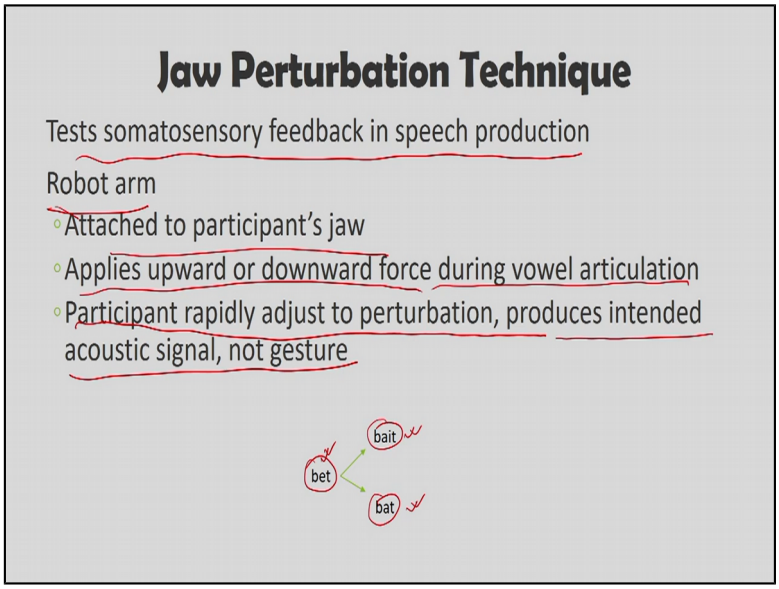
So, Feedback system, they control, and adjust the forward trajectory, based on real-time information, as it has been, said here. Now, currently, the models of Speech production, generally assume, a Feedback, and a Feedforward system. Now, Motor systems receive rapid

Feedback from Somatosensory system, Muscles, Tendons, and Jaws. And, Feedback in Speech production are done by, Somatosensory system, and Auditory system. Now, just as the Feedback and Feedforward system, works in the limb, in the voice also, or in the Speech production also, we have the Feedback, and Feedforward system.

So, Feedback system, generally in terms of speech, are done by the Somatosensory system, and the Auditory system. Now, in the Speech production, the Somatosensory system, provides continuous Feedback, on articulator movement. So basically, this Somatosensory system, keeps on giving you real-time Feedback of, how the articulation is being done, on those parts of the body, which produce speech.

For example, the Tongue, the Jaw, and the Lip, and the vocal folds. The Auditory System Feedback is also important. But, Auditory system Feedback, is a delayed system. So, Somatosensory system, is a real-time system. It tells you, what is the movement, or how your various Jaw or Tongue is moving, for producing speech. The Auditory system, is a late system. Now, for studying the Auditory system, the Somatosensory Feedback system, we use something called the, Jaw Perturbation Technique.

(Refer Slide Time: 16:06)



The Jaw Perturbation Technique. And so, what is the Jaw Perturbation Technique. It can be used, to test the hypothesis, that Motor speech systems, uses Somatosensory Feedback, to make corrections during articulation. So, the idea that Somatosensory system, they make corrections, while making speech, that can be tested through, this the Jaw Perturbation Technique. And so, what it is? It tests the Somatosensory Feedback, in Speech production.

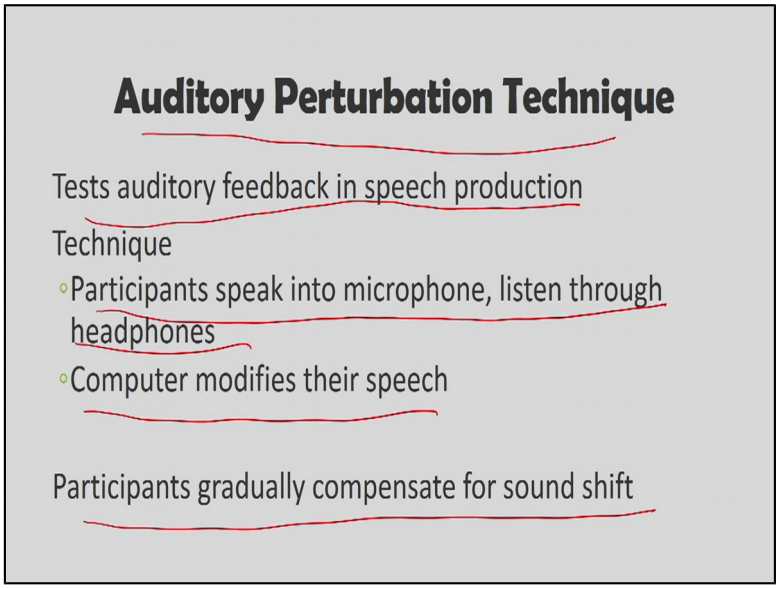
So, what do we do? We attach a robotic arm, to the participants Jaw. So, the participant is asked to produce speech. And, a robotic arm is basically attached, to this part of the Jaw, the lower Jaw. Because, that does the one, which makes the movement. What the robotic arm does. It applies, upward or downward force, during vowel articulation. So, when somebody is producing vowel, the robotic arm is applying a, upward and downward force.

Now, participants rapidly adjust the perturbation, produces intended acoustic signal, not gestures. So, what happens is, when this happens, the robotic arm is producing, some force on the Jaw, but the participants is not able to produce, the right vowel. But, what happens is, when he when he sees that, when the Jaw movement is there, and when the Somatosensory system, it senses that, an external force is creating, an external body is creating the force, it adjusted itself in real-time, to produce another vowel, or another speech sound, which is there.

So, it adjusts itself, in real-time. For example, the people were asked to produce, a bet. And, they were asked to, an upward or downward force of producing, the upward force was produced. The actual sentence, that was produced by people was, bait. And, when a downward force was there, they produced, bat. And so, this is the correction, that was done.

Now, the finding from, this study suggests, that the goal of the speech Motor plan, is not the production, of a particular gesture, but rather, particular acoustic signal. So, this studies say that, the speech Motor system goal, is not for a production of a particular gesture, but for the production of, or the correction of an acoustic signal. The exact role of Auditory Feedback, in Speech production is, actually less clear.

(Refer Slide Time: 18:42)



So, this is for the, Somatosensory system. But, what about the Auditory system. The Auditory system, has also been held responsible, for providing, Feedback. Now, the exact role of this Auditory system is, exactly a little bit not clear. The reason is that, there is a large time lag, with Auditory Feedback, since sound waves need to travel, from the mouth to the ears, and to be produced by the Auditory cortex.

Now, the Somatosensory system, is a real-time system. And so, it provides real-time corrections, or Feedback corrections, on to movements of the articulators, movement of those parts of the body, which produce speech. But, the Auditory signal, or the Auditory Feedback, is a slow Feedback. The reason is that, the word has to be produced by the mouth, it has to be heard by the ear, and then processed by the Auditory cortex, only then, it can produce the Auditory Feedback. The time lag is, too great.

So, the time lag between, producing the speech, and hearing the speech, is too great, to produce useful Feedback, during articulation of a syllable. And so, that is one of the reason. Now, what is the role of Auditory Feedback, that can be tested in or Auditory Feedback in Speech production, that can be tested by something called a, Auditory Perturbation Technique. And so, what happens in the Auditory Perturbation Technique, it closes the role of Auditory Feedback.

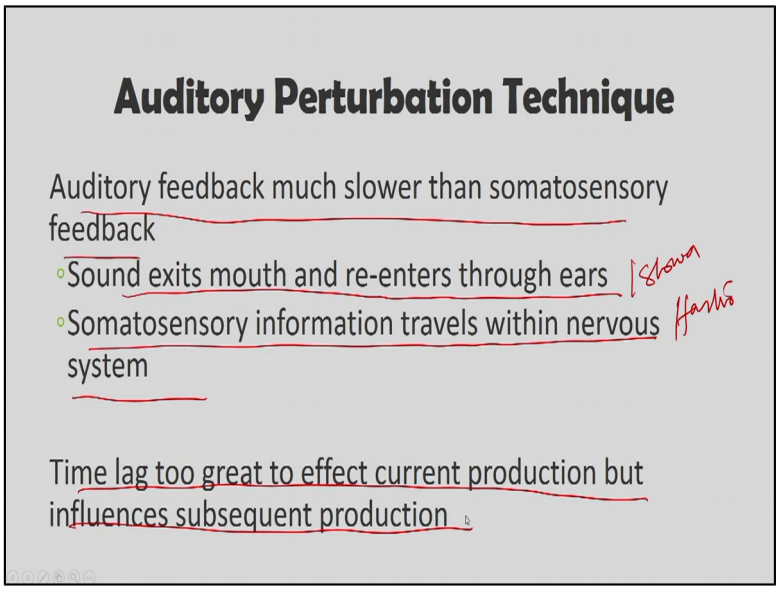
Now, it tests the Auditory Feedback in, Speech production. What is the technique like? Participants speak into a microphone, and listen through a headphone. So, you speak into a microphone, and you hear your own speech, in a headphone. What happens here is, the computer modifies their speech. What it does is, it delays the speech. So, here the participants,

repeat a syllable, into a microphone, and listen through headphones, of their own modified voice, by a computer.

The findings suggest, that we use Auditory Feedback, not to make, course correction of articulatory gestures, in progress, but rather to address, subsequent articular movements. So, we do not use it for course correction, but we use it for adjusting subsequent articulatory movements. Since, this is a delayed signal, so it is used for, the next articulatory movement, which is there. So, computer modifies the speech, and so you use to hear a modified speech.

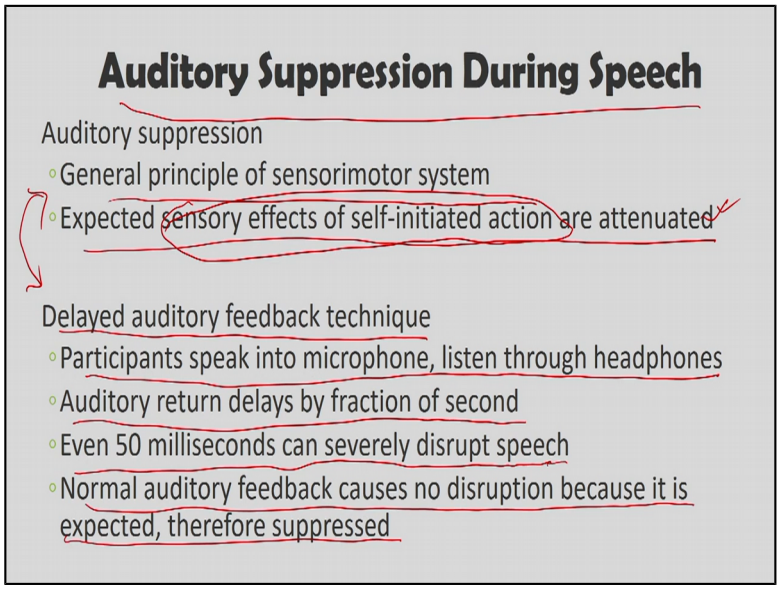
And so, participants gradually compensate for, the sound shift. So, you hear it later, and so what we do is, since something has already been spoken, the Auditory signal, it is not used for gesture correction. It cannot do gesture correction because, it is a late signal. So, what it does is, it actually is used for compensation, or the sound shift. So, if the sound is still being produced, or it has been produced, so you compensate that. And, that is how, the Auditory Perturbation Technique is used.

(Refer Slide Time: 21:30)



Now, then there is something called so Auditory Feedback is since, the Auditory Feedback is much lower than Somatosensory Feedback, sound exits mouth, and re-enters, through the ears. And, Somatosensory system information travels, within the nervous system. And so, this is a slower Feedback, and this is a faster Feedback. Now, time lag, too great to affect current production, but influences subsequent production. So, that is what the, Auditory Perturbation Technique is.

(Refer Slide Time: 21:53)



Now, this Auditory perturbation technique, suggests something called, an Auditory suppression happening, during speech. An Auditory suppression is, it is a technique in which, people, they perceive their own act. And, when they perceive their own act, they do not respond to it. For example, Auditory suppression is similar to the fact that, if you tickle yourself, you will know that, you are tickling, and you will not laugh.

So basically, those acts, that you actually do, on yourself, the goal is also perceived, and so your response to that particular act is, dampened. And, that is what is called, Auditory suppression. So, Auditory suppression, it is a general principle of Sensory Motor system, and expected sensory effects of self-initiated action are, attenuated. And, that is what, I am talking about. Sensory effects of self-initiated action are, attenuated, or it is lower down.

So, if you tried, tickling yourself, or if you try doing something to yourself, your response to that particular action, will be attenuated. So, I have never seen somebody, tickling himself and laughing. And, that is because of, the suppression, Auditory suppression. Now, the delayed Auditory Feedback technique requires, research participants, to speak while listening through headphones, and their own voice, which is delayed, by a fraction of a second.

So, how is this Auditory suppression, actually tested. It is tested through something called a, Delayed Auditory Feedback technique. Right. So, what we are doing right now, we are focusing on, the Feedback systems for, Auditory speech productions. So, we look at the Somatosensory

system, and we are not looking at the Auditory system, and looking at the idea of Auditory suppression. So, we use something called a, Delayed Auditory Feedback technique.

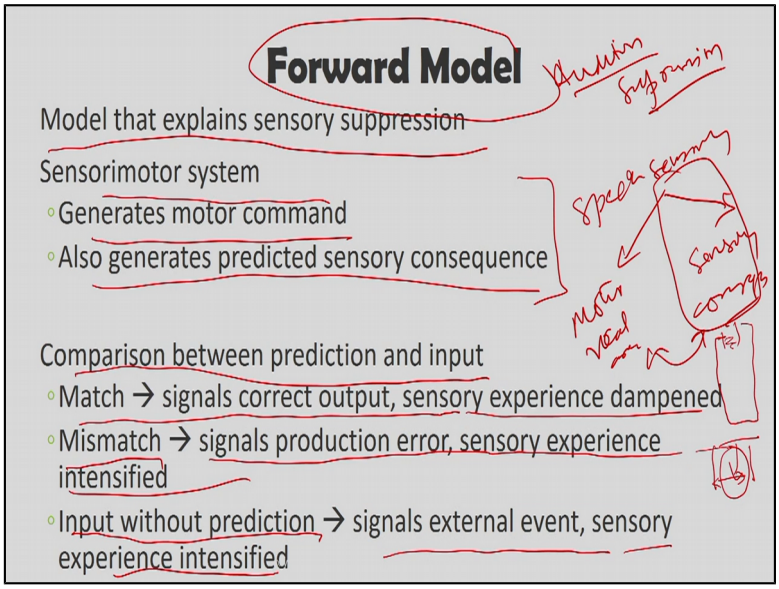
Now, here the participants, speak into a microphone, and listen through a headphone. Now, Auditory returns delays by, fraction of a second. So, when you talk in the headphone, what happens is, your speech is delayed artificially, by a computer, by fraction of a second. Now, when that happens, so even if 50 milliseconds can severely disrupt speech. So, if that happens, if you hear your speech at even 50 seconds later, people keep quiet, and they are not able to complete their speech.

Now, Normal Auditory Feedback, causes no disruption because, it is expected, therefore suppressed. Now, even if there is a time delay, in in in your speech, so normally what happens is, when you speak something, the Auditory system expects a delay, in hearing it. And so, it suppresses itself, it suppresses the sound, which you are hearing in your own ear, through your mouth. But, if there is a delay, in in the production of speech, by the mouth, and the hearing, even by 50 milliseconds, then it will cause an interference.

And, because of this interference, this speech will be affected. Now, when you when you do that, there is a disruption of Auditory suppression. Or, this disruption of Auditory suppression, it disturbs person's ability to actually speak. Now, recent EEG, and Neuroimaging shows, that speech interference, due to delayed Auditory Feedback, is related to, Auditory suppression. So, this kind of thing, where a fraction of a second is added, or a 50 millisecond delay is added, into your speaking, and your hearing, that can interfere with your speech.

That has been proved by, recent EEG, and Neuroimaging techniques. Now, Auditory suppression, is related to the general principle of Somatomotor system, in which, the expected sensory effects, of a self-initiated action is, attenuated. And, that is what, we were talking about. This can be explained, in terms of, something called the, Feedforward model.

(Refer Slide Time: 25:36)



And, model explains, some of the sensory suppression. So, this particular Auditory suppression, that when you hear your own voice, it is suppressed by the ear, or the Speech production system, that can be explained by, the results of that experiment, can be explained by the, Feedforward model. It explains, sensory separation, by proposing that each time, the sensory Motor system, generates a Motor command. It also generates, a predicted sensory consequence.

So, when a sensory system, gives a command for the speech articulators, or the speech organs, to start talking, it also generates, a predictive sensory consequence. It also tells, or it generates, a consequence also, from that particular act of speaking, for comparison against the actually sensory input. So, the sensory system, not only gives a command, to the articulator, for producing speech, it also generates, a consequence, to that speech. And, this consequence is compared to the, in to the actual sensory input.

Now, if the prediction, and the input match, the sensory experience is, dampened. Now, the prediction, and the inputs match, now if what you are hearing from your so when the Somatosensory system, it starts, or it prepares your vocal cord, or the vocal system, for speaking, it gives a command, to the vocal system to speak, at the same time, it also produces, or it also creates a generates a predicted sensory consequence, that you are going to hear.

And so, this this two, it creates. So, when you speak, and when this speech is heard by the ear, and when the ear, what it hears, matches with, what the Somatosensory system has already suggested, it is going to hear, this will be dampened, or what you hear will be dampened. The speech will be dampened, and so, you can speak. If you hear yourself, at the same time, as you

are speaking, what will happen is, or if it does not dampen, what will happen is, there will be a break in your speech.

So, sensory Motor system, it generates Motor command, also generates sensory, but as consequences. So, the speech system, speech sensory system, it generates two commands, when it starts speaking. It generates a command, a Motor command, two vocal areas, for producing speech. At the same time, it also generates, a sensory consequence, that what it is going to say, or what sound will be produced. Now, when the sound is actually, it starts from the mouth, and reaches the ear.

And, in the ear, if the ear is not able to hear, or if what the ear is hearing, and what it has it has been said by the Somatosensory system, that what it should hear, if the match is not there, then there will be heightened reactions. But, if what the Somatosensory system, what consequence it produces, or what it tells the ear, that it is going to hear, and what the ear hears, if it is matching then, a suppression happens to, what you are hearing.

Now, comparison between, prediction and input. Somatosensory system, while producing speech, will send a Motor command to, generate a voice. And, this will be the input, to the ear. And, at the same time, it will also produce, something called the consequence, it could also produce a signal for, what consequence will happen, due to this particular Motor command, or what sound it is going to hear, it will tell the ear.

Now, if there is a match between, what the ear is hearing, and what it has been suggested to the ear, that it will hear, then what will happen is, there will be a dampening of the sound, that you are hearing, from the ear. So, if there is a match, the signal signals correct, output sensory experience is, dampened, and so what you hear will be, dampened. Now, if there is a mismatch, signals production error, sensory experience is intensified.

Now, if what you hear, and through the input, and what consequences have been said to you, if there is a mismatch, signal production error happens, and sensitive experience is are intensified, and also, input without prediction. Now, if there is no prediction, if the sensory system does not predict, what you are going to hear, but you hear something from an input, signals and external event, and sensory experience is, are intensified.

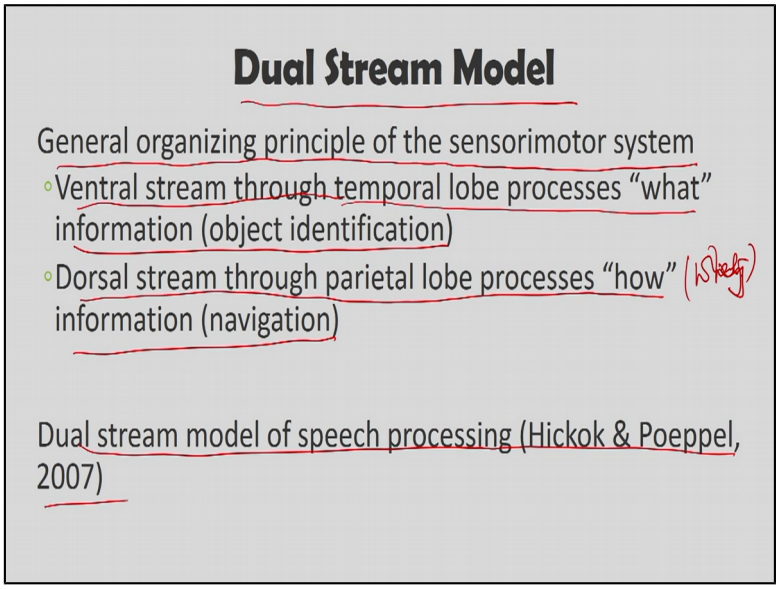
And, so here, what happens is, you start thinking that, it is not something, that you have said. You will start thinking that, it is coming from, somewhere else, and the sensory experience is intensified. It is not your sound. And, that is why, there is no interference. So, when we speak, we also hear, what we are speaking, but that does not dampen our speech system, or that does not dampen us, that does not interfere our speech. Because it is dampened.

But, if somebody else speaks, then the sensory consequences are not, and so sensory experiences are ins intensified, and so we start hearing to them, and responding to them. We do not respond to, what we are speaking. So, Auditory suppression, is strictly a response, to end immediately, prior speech event. Now, Howdin Nagarajan, 2011, they suggested Auditory suppression, they prevent Feedback interference, during Speech production.

Now, by the time, the Auditory Feedback, reaches the speech Motor system, the command for the next articulatory gesture, has already been sent. So, it provides no useful information whatsoever, about the current state of Speech production. Now, using the Auditory Feedback, speakers, modify subsequent articulators, to compensate, when they hear their voice, altered in unexpected way.

So, what generally the Auditory speech signal does is, since it is a late signal, what it does is, it only does a correction to, later speech signal, going to come. So, that is about the, two Feedback system.

(Refer Slide Time: 31:55)



That is therein, in the Auditory Speech production. Now, there is a Dual Stream Model, which has been proposed, for production of speeches. So, what is this model. So, Alan and his colleagues, in 2001, they map the dual processing system for audition. Now, this dual model system, it is actually coming from, I mean, its borrowed from, Weasel perception. Now, in Weasel perception, the general organizing principle, for sensory Motor system, in individual perception, you have two streams.

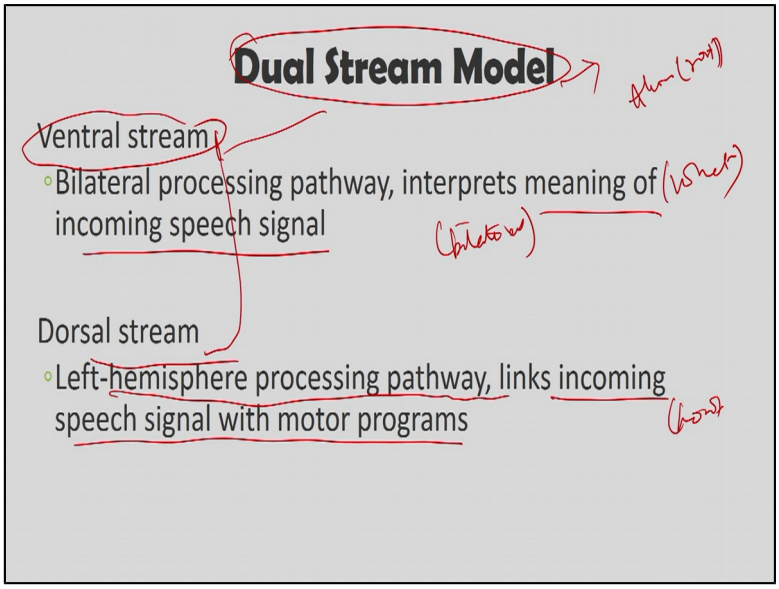
There is a Ventral stream, and there is a Dorsal stream. The vent Ventral stream, which is running from the Auditory cortex to the temporal lobe, is basically recognition, object recognition. And, the Dorsal stream, which is running from, the visual cortex, to the parietal lobe, is basically for navigation, for spatial movement. Now, these two systems, are very popular in in the visual perception. And, they have been borrowed in, Auditory perception also.

Now, the Ventral stream, in in the visual or any sensory Motor system, through temporal lobe, processes, what information, object identification. As I said, in the visual stream, the temporal lobe process, is about, what information is being said, what is an object. So, object recognition. And, the Dorsal stream, parietal lobe from the occipital lobe, it looks at something called, how information. Earlier, it was called the, what information, why information, why am I seeing. Later on, it was identified in to, how.

Since, the Dorsal stream will talk about navigation, or the Dorsal stream is involved in navigation, it is about, what, or how the particular navigation, has to be done. Now, the Dual Stream Model, for Speech production, has been processed by, or has been proposed by, Hickok

& Poeppel, in 2007. Now, the Speech production system, according to Alan and his colleagues, in 2001, the Speech perception system is organized into, the what system, which is the Ventral stream, and how, which is the Dorsal stream.

(Refer Slide Time: 34:01)



So, Alan in 2001, it says that, the Speech production system are processed, or it has been divided into, the Ventral stream, and the Dorsal stream. And, the Ventral stream is in the, what system, and the Dorsal stream is in the, how system. So, Hickok & Poeppel, in 2007, they proposed a Speech perception. And, the Ventral stream is a, bilateral processing pathway. The Ventral stream, that you see, in Speech perception. So, Speech perception, is basically having, two streams of information, or there is two ways of processing.

There is a Ventral way, or a Ventral stream of processing of speech, and there is a Dorsal stream. The Ventral stream, starts from the temporal lobe, to the Auditory cortex. And, the Dorsal stream, it starts from the parietal lobe, to the Auditory information. Now, what is the Ventral stream, according to Hickok & Poeppel. What does it compose of? So, the Ventral stream, is a bilateral processing pathway. So, Ventral stream is bilateral.

Which means that, it is in the both hemispheres, and interprets the meaning of the incoming speech signal. So, it is, first of all, it is a bilateral. Which means that, it is parallel. There are two Ventral streams, first of all. And then, what it does is, it interprets meaning of the incoming speech. In comparison to that, the Dorsal stream is basically, your left hemisphere, processing pathway. So, it is not on the right hemisphere. The Dorsal stream, is the left hemisphere, processing pathway.

And, it links, incoming speech signal, with Motor programs. So, what does it does? It links the, incoming speech signal. The Ventral stream, makes the meaning of the incoming speech. And, the Dorsal stream, takes the incoming speech signal, and it links the speech signal, into the speech Motor programs, so that, what should be said. So, something is said, and your ear hears that, it takes this signal.

And then, the meanings of this speech, is done by the Ventral system, but this speech is heard by the Dorsal system, and it sends, a signal Somatosensory signal, to the Auditory to the voice producing areas, the vocal areas, to articulate the next speech. That is what, the Dorsal system does. It sends the, Motor signal. The Ventral stream, is basically, what does it mean. And, the Dorsal stream, is basically, how do you say that. So, it is a left hemisphere processing pathway. It leaves, incoming speech signals with the, Motor signals.

Now, traditionally, the traditional view, that language is literalized to the, left side of the Brain. Now, Neuroimaging studies, it suggests that, both hemispheres of the cerebral cortex, and a number of subcortical systems, are recruited for, speech processing. So, it says that, not only left hemisphere, but both hemispheres are involved in, speech processing. Now, there are parallel Ventral streams, in both the hemispheres, with the right hemisphere, processing meaning, over large timescales.

So, the right hemisphere, process meaning, and it takes a long time, for processing meaning. And, the left hemisphere, processing is, for a short period of time. And, it is analytical processing, that the left hemisphere, actually does. Now, the Dorsal stream, originates in the Speech perception area, in the Lateral fissure, which is the Wernicke area, and extends to the Posterior inferior frontal gyrus, which is the Broca area, as there is a primary Motor cortex of the left hemisphere. So, that is what it is.

The Dorsal stream, it starts from the Wernicke area, and goes to the Broca area. Now, in the model, the Ventral stream provides an interface between, pram processing below the level of word, which is Phoneme, Prosody and Syllable. Now, this model of Dual Stream System, it says that, the Ventral streams provides, an interface, for processing, those sounds, or processing those aspects, which are below the level of the word. For example, at the level of Phonemes,

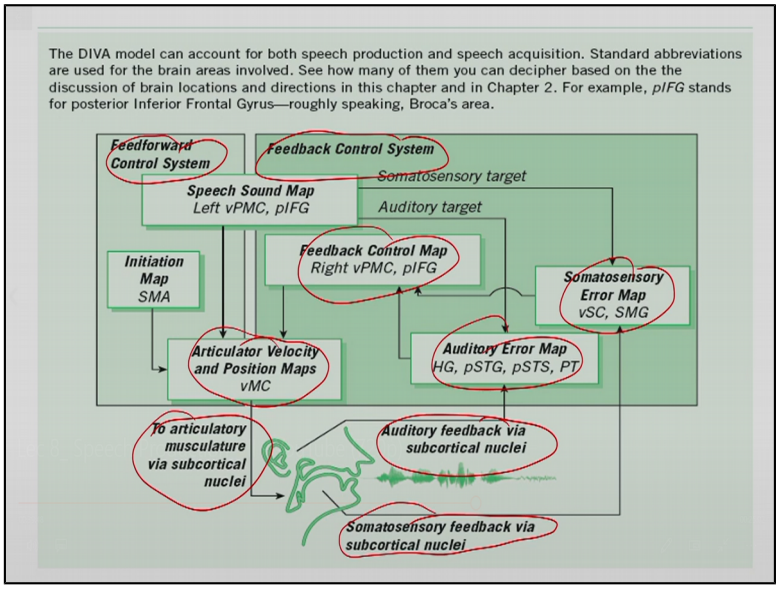
Prosody, and syllables, and processing at the word level, are also processing, at above word level. For example, Sentence and Discourse.

So, this kind of processing, below word level, and above word level, the interface of processing these, are through the Ventral stream. Now, the Ventral stream, ac accounts for Clinical and Neuroimaging data, required for retrieval of, words from the meaning. That is what, the specification of Ventral stream is. Now, what is the Dorsal stream. Now, the Dual stream, looks a lot like, the Wernicke.

So basically, how does the Dual Stream Model, it compensates, or it matches with the older theory of Wernicke and Geschwind, that we have looked at. What happens is, what is the similarities between them on, how does the Dual stream, accommodate those theories. First, the Dual stream looks a lot like, Wernicke-Geschwind model. So, it is basically, the same kind of interpretation, which is there. Also, the Dorsal stream, it provides a mechanism for Motor Theory.

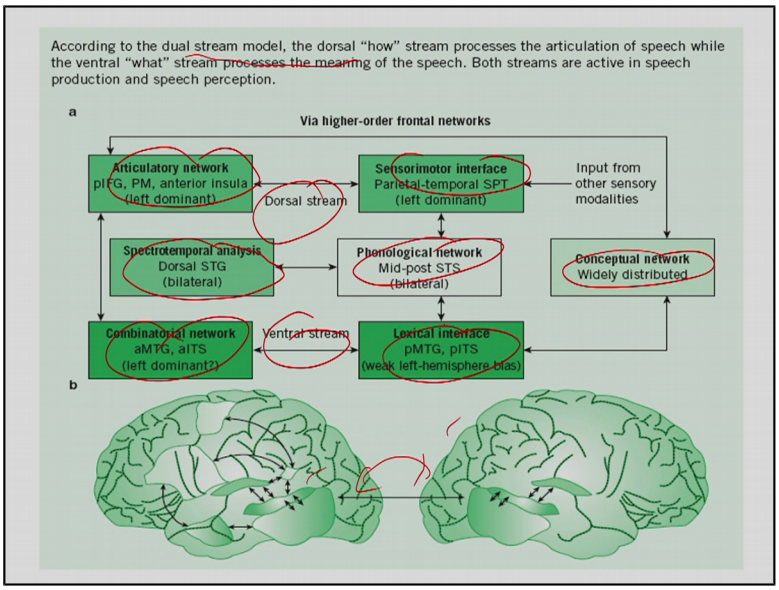
The Dorsal stream provides a reason, that Motor Theory of Speech perception is, somewhat correct. Also, the Dorsal stream also provides, a mechanism to confer, Phonological short-term memory. So, this is the Dual Model. And, this is the details of the, Dual Stream Model of Speech production. Now, there is another model of Speech production, which is called, so this is, what it is.

(Refer Slide Time: 39:21)



You have the, Auditory Feedback, via Subcortical nuclei. The Somatosensory Feedback, via Subcortical nuclei. To, Articulatory musculature for, Subcortical nuclei. You have the, Articulator Velocity, and Position Maps. Feedback Control Maps. Auditory error Maps. Somatosensory Error Maps. Feedback Control Systems, are there. And, this is, Feedforward Control System. And, this fid for Feedback fid control Feedforward control system, generally are used for, production of speech. And, this is how, the DIVA model, actually works.

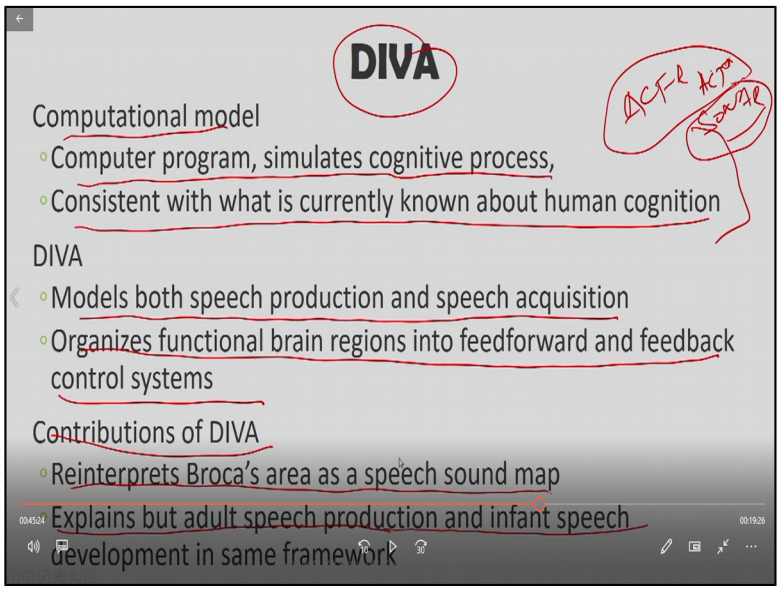
(Refer Slide Time: 39:48)



So, this is basically, that was the DIVA Model, that I explained. So, it should have come, later on. This is the, how the Dual Stream model, really works. This is the, Dorsal stream. This is the, Ventral stream. So, you have the Articulatory network, Somatosensory interface, Input from other sensory modalities. You have, Conceptual network here, Phonological network here, Lexical interface here, temporal analysis here, and Combinatorial networks here. And so, these

are the, two Wernicke area, and how the information, in the left and right hemisphere, they interact for, producing the speech stream.

(Refer Slide Time: 40:24)



And, besides that, we have another theory, which is called the, DIVA Theory, the D I V A Theory. Now, DIVA is basically a, computational model of Speech perception. So, what does it say. So, DIVA incorporates, all the functional Brain region, organizing them into, Feedforward and Feedback Theory. So, it is basically a, computational model. What is a computational model? A computer program, simulates cognitive processes, that is what, a computational model does.

And, this consistent with, what is currently known about, Human cognition. So, example for example, you have the ACTR and ACT*, or the Sonar, and these are computational models for cognition, which are provided by Anderson. Similarly, in in linguistics, in Psycholinguistics, you have the DIVA program, which is a computational model. What does the model do? It incorporates, all functional Brain region, into Feedforward and Feedback Models.

So, model both, Speech production, and speech acquisition, it organizes functions Brain regions, into Feedforward, and Feedback control systems. So basically, the Feedforward system, programs the Motor command for speech, and which includes the Somatosensory Motor area, the Broca area, and the Motor cortex area. So, the Feedforward system, it programs the Motor command for speech.

According to DIVA model, the Feedforward system, it programs those Motor area, or Motor sequences, for Speech production. And, Somatosensory in Auditory targets, the expected sensory consequences, are sent to Feedback control systems. Somatosensory system, and Auditory system, which are Feedback system, they are sent to Feedback control systems, where they are compared, against the actually sensory input.

So, Auditory Feedback, and Somatosensory Feedback, they are sent back to control systems, which are there for Feedback, and they compare the actual sensory input with, what the input it is receiving, the ear is receiving. Now, any detected Auditory, or Somatosensory errors, if the control system, it gets any kind of error, which may be a Somatosensory error, or an Auditory error, they are sent to Feedback control map.

Now, this again, is sent to a Feedback control map, which sends Feedback to Motor cortex, to modify its Motor command, and articulators. So, if the Somatosensory system and Auditory system, it detects an error, in to what it has, what the speech has been produced, and what the ear is hearing, this error, is sent to something called the, control map, where the control map processes this information, and sends it back, to the Motor cortex, to modify the Motor command for articulator.

So, if it sees an error, if what you are speaking, is not what you are hearing, or what is expected, what will ha happen is, the articulators, the speech articulators, the Jaw, or the Tongue, or the Lips, or the vocal fold, it will be adjusted. Now, DIVA is basically, called the directions into, velocities of articulators. Now, what are the contributions of this DIVA system, what does it propose. First of all, it proposes a new way of thinking about the Broca’s area. Traditionally, Broca's area was about, Speech production, or production of speech areas.

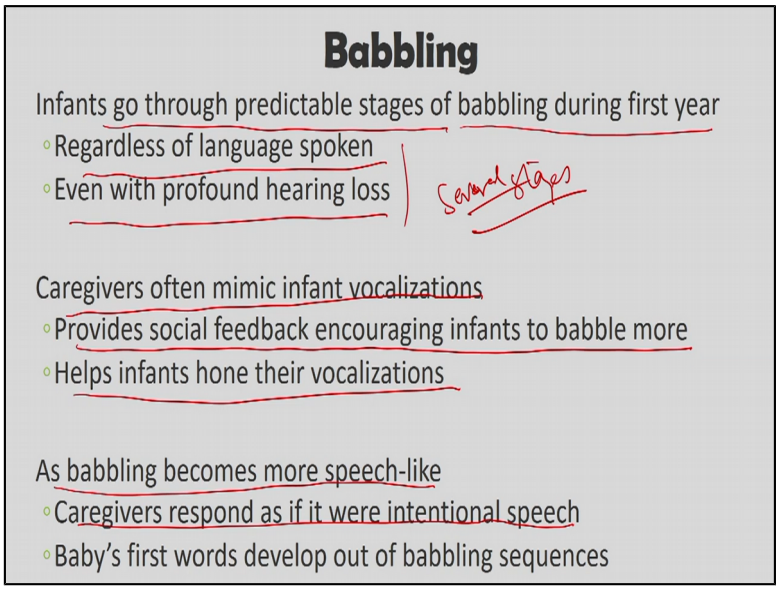
Now, according to the DIVA system, Broca is a spatial area, which is used for Speech production, Speech perception, and gesture production. So, both, perception and production, and gesture production, is also done by, Broca area. Also, Broca area, serves as a speech sound map. Also, Broca area serves as a speech sound map, where any error, that the current Feedback control system, it detects in the Somatosensory, or the Auditory Feedback, and it sends an information to the speech sound map, which is in the Broca's area, where Mirror neurons representing speech sounds, whether perceive or produce, are located.

So basically, the control map, that we have been talking about, in the DIVA, is in the Broca area. Also, DIVA explains, both adult Speech production, and Infant speech development, in the same framework. What it says is that, Infants early Babbling, which is random and repetitive, is driven by the Feedback control system. Now, as Infants babble, it receives sensory Feedback, that it that it can compare against the speech samples, it has already experienced.

So, when Infant Starts Babbling, what happens is that, the Infant generally, first, when he is Babbling, it just babbles. But, when the ear hears, this Infant hears his own Babbling, and compares this with the actual sound of, what an adult is saying, it modifies its articulators, it modifies this, the way it is producing speech. And, so basically, DIVA provides a reasoning for, how this happens, so, reinterpret the Broca's area, as a speech sound map, and also explains, but adult Speech production, and Infant Speech production development, in the same framework.

So, that is what the, DIVA system is. And, this is how, the DIVA system works. So, you have this, feed Feedforward control system, you have a Feedback control system, and then, you have something called the, Somatosensory error map, Feedback control map. And, this is how, the articulators are again, send back information to controller. So, this is the Dorsal stream.

(Refer Slide Time: 45:44)



Now, how does development of Speech production happen? Now, how does Infants develop speech, or how does speech development happen, or the development of Speech production happen. Now, Infant coming to the world, screaming. And, for the next few months, crying is main form of communication. So basically, they start with something called, screaming. The Infants come in the world, with screaming. And then, later on, they cry.

But, after that, the whole process is there of Speech production, or development of Speech production in them. And, that starts with a, Babbling. So, Infants go through, predictable stages of Babbling, during the first year. Babbling is the same, mu, gu, sound, that the Infant is producing. And, how this Babbling, actually leads to, the production of speech, is what we are going to learn here. So, regardless of language spoken, even with profound hearing loss, so you have something called, stages of Babbling.

So, from the first 6 months, Infants progress through, predictable stages of vocalizations. The first 6 months are, predictable areas of, or predictable stages of vocalization. First is the, Phonation stage. In this, from birth to 2 months, Infant produce vowels, like sounds, by vibrating the vocal cords. So, the first 2 months, the Infant's produce vowels. And, these are vowel like sound, which is like, which is not Babbling, exactly. So, it is vowel like sound, and they are produced by, vibrating the vocal cord.

Now, they have little control over the articulators, and they rarely produce a consonant. So, they produce a vowel, but not a consonant, as they say it. At the going stage, which is 2 to 4 months, Infant produce syllables, like sounds, in back of their vocal tract. For example, coo or gu, both consonant and vowels are, poorly formed, and highly variable. But, they start producing consonant and vowel, by 2 to 4 months, by producing sound in the back of the vocal tract. But, these sounds are like, very poorly formed, vowels and consonants.

Now, there is an expansion stage, which is 4 to 6 months, Infant produce a variety of different sounds. Well, from vowels, such as, e e and a h, and consonants like sounds are produced. And, this is called, Marginal Babbling. So, from 4 to 6 months, Infants are able to produce, full vowels like, eee-vowel, like the aah-vowel, so back of the Tongue, front of the Tongue, by Jaw lower position, Jaw upper position, and so, they are able to produce these vowels, but some consonants, they are able to produce. And, this is called the, Marginal Babbling sound.

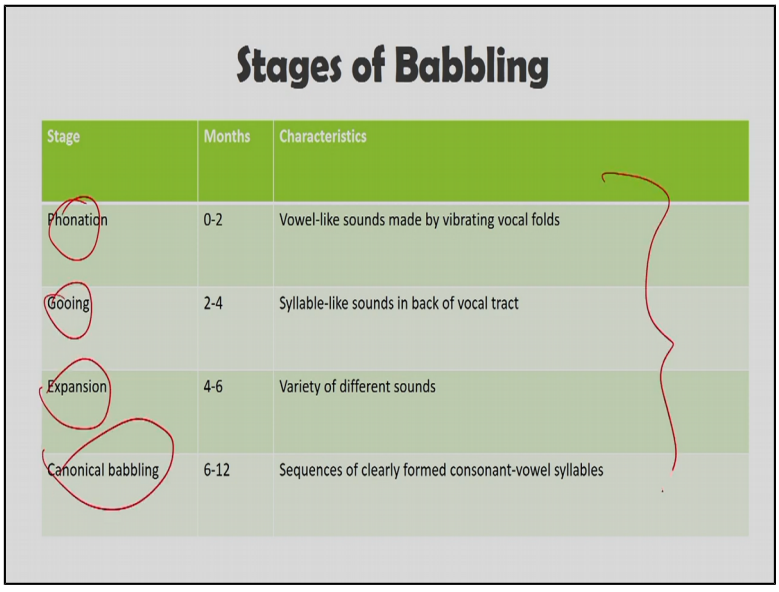
Now, there is a Canonical Babbling stage also, which is from 6 to 10 months. And, this is pre-linguistic vocalization. And, it is characterized by sequence of clearly formed consonant, vowel, syllables, first produced in isolation. For example, ba. Then, re-duplicated. So, in first stage of the Infants are, able to produce clearly, consonants, in isolation. For example, ba. Then, they are able to re-duplicate the syllables. For example, not one just, ba, but they can produce, ma-ma.

So, in the first stage, they are able to just produce, ma-ma, ma-ma, like that. And, later on, they are able to re-duplicate this vowel, and produce two, which is, mama. And then, later, they can produce a variegated syllable like, daddy, which is more than replication of a single syllable. Now, consonant consisting of mainly, oral and nasal stop, produced with Lips or Tongue. So, at the Canonical Babbling stage, they are able to produce consonants, with oral or nasal stop. For example, the b sound, the d sound, the m sound, and the n sound.

And, Caregivers, they respond to Canonical Babbling, as they are intentional speech. So basically, they are able to produce, this kind of speech, so early. There are several stages of Babbling, which we talked about. And, the Infant goes through, these stages of Babbling. So, that is how, the Babbling starts. Now, Caregivers, often mimic Infant vocalization, and they provide social Feedback, encouraging Infants to babble more, help Infants in there, on their vocalization.

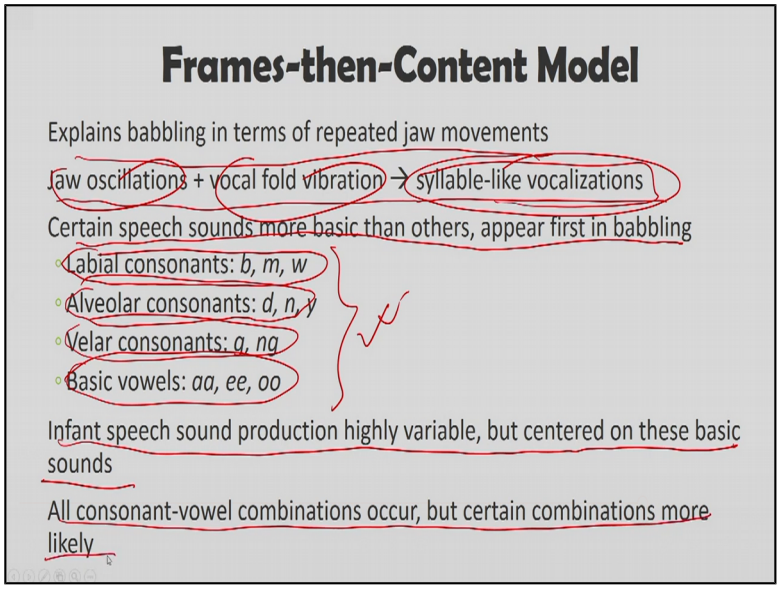
As in, babbling becomes more speech like, Caregivers respond, as if they were intentional speech. And, baby's first word develops out of, Babbling sequences. So basically, that is how, it is. So, Infants, when they produce this Babbling, the Caregivers, they respond to the Canonical Babbling, and as if they were, intentional speech sounds.

(Refer Slide Time: 50:14)



Now, these are the stages of Babbling. The Phonation stage, the Gooing stage, the Expansion stage, and the Canonical Babbling. And, these are the things, that we were talking about.

(Refer Slide Time: 50:22)



Now, there is something called a, Frames-then-Content-Model, which is there, for Speech production. So, what is that? It says it says that, early vocalizations, they are strictly driven by, Motor systems, and not Auditory Feedback. So, early vocalizations of an Infant, they are driven by Motor systems, and not by Auditory Feedback. And so, they are basically then, the idea is that, they are universal in nature. And, Infants are, early vocalization is, universal in nature. Also, deaf children, go through, the same stages.

So, the Frames-then-Content-Model, actually explains the Babbling, in terms of, repeated Jaw movements. So, we are coming to that. So, Davis & McNeilage, 1994, proposed the theory, that explains Babbling, in terms of Jaw movements. So, this theory was given by, someone called, Davis & McNeilage. So, what did they say? According to the model, Babbling begins around, 6 to 8 months. They say that, Babbling begins around, 6 to 8 months. Because, this is where, the Infant gains control, over the Jaw.

Infant, when he is born, he is not able to control the Jaw, and so he cannot produce speech. But, by the time, from 6 to 8 months of his birth, the Infant is able to control the Jaw, and able to produce, Babbling. Now, Infants will make Jaw movements, without vocalization, but sometimes, Phonation results. So, initially, the Infant may produce, the Jaw movement. And so, the movement will be, alone there, and nothing will be there. But, later on, there will be, some phonation also.

So, Jaw oscillations plus vocal fold vibration will, syllable like vocalizations. Jaw oscillation, a complete by vocal fold vibration, is a sequence of vocalization, then sound like, simple

consonant vowel sound. So, certain speech sounds, will be produced, which are most basic than others, appear first in Babbling. So, Jaw oscillation, which is accompanied by vocal folds. So, Jaw oscillation, accompanied by vocal fold vibration, they sound like simple vocalizations, and that happens.

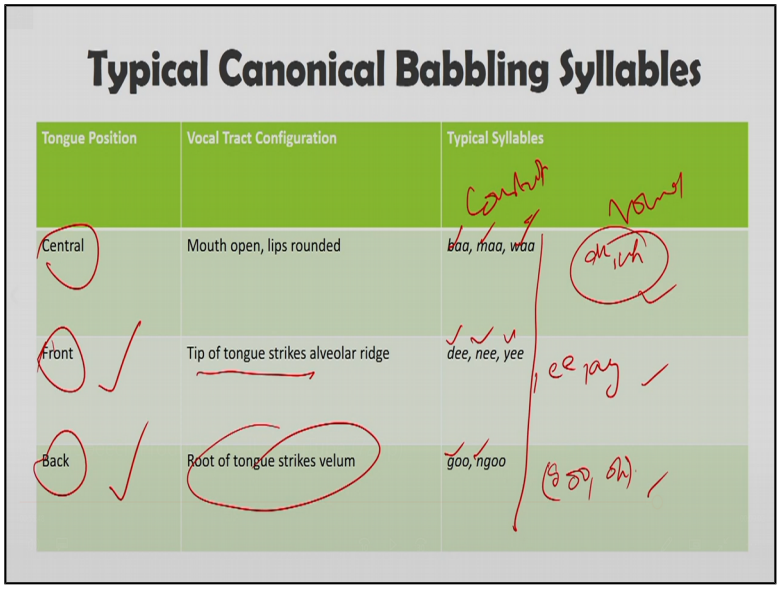
So, initially, what happens is, they control the Jaw, the Infant is able to control the Jaw, and move the Jaw, but anything comes out of it. Later, there is some Phonation. And, after that, there is Jaw oscillation. And, this Jaw oscillation, in a session with the vocal fold vibration, the Infants are able to produce, simple syllable like vocalization. The model also stipulates that, certain speech sounds, are more basic than others, appearing first in in in Canonical Babbling.

For example, the Labial consonants, b m w, the Alveolar consonant, d n y, the Velar consonants, which is n and g, and the basic vowels of, a e o, they are being able, to produce by the Infants. So, nearly all languages, make use of consonants, produces the Lip, which is b and m, the tip of the Tongue, which is d and n, and back of the Tongue, which is g and ng. Now, since this is common to all languages, all Infants goes through, or else, all Infants are able to produce this, speech sounds.

Now, all languages distinguish vowels, based on Tongue position, front e e, central a h, and back o o, as we have, sh from that, basic vowels also produced. Now, early Babbling features, all speech sounds narrowed, later to, phoneme inventory of language. This is the Old Theory. So, the Old Theory says that, babbling features on, when the Infant is producing early Babbling, now they are producing all kind of sounds. And, later on, they by listening to the Babble, they constrict these sounds, and are able to fit their Babbling of producing sound, into the Phonation of the language, mother language, which they are going to speak.

Now Infant but, New Theory says that, Infant speech sound, is highly variable, but cent centred on most basic speech sounds, common to all world languages. It basically says that, most Infants produce, the same kind of speech, when they start. Now, these should be, predictable patterns of consonant, vowel pairs, in Babbling, as opposed to, random mixtures of consonants and vowels. Model suggests, three patterns of freewill. So, Infant speech sound production, highly variable, but centred on, these basic speech sounds. All consonant vowel combination occurs, but certain combinations are more likely, than other combinations.

(Refer Slide Time: 54:30)

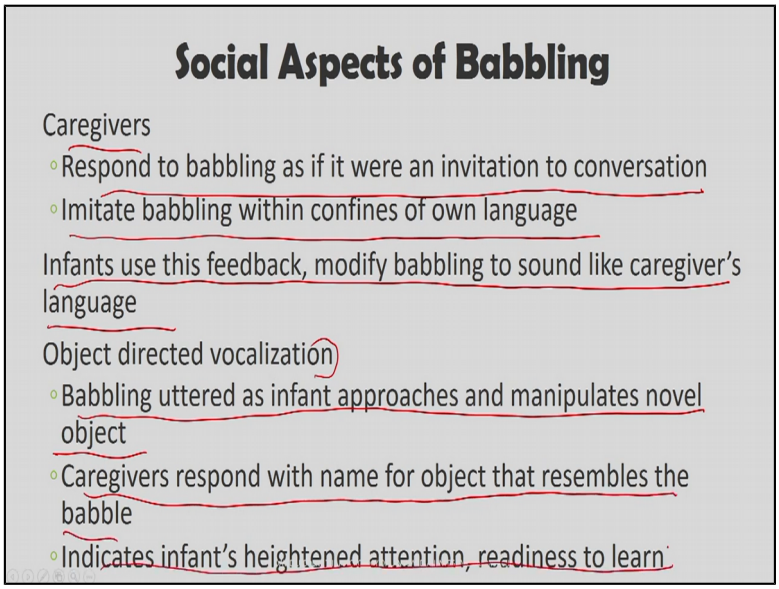


So, as we said, typical Babbling syllables, the model suggests that, three patterns will prevail. You have a Central pattern, you have a Frontal pattern, and a Back pattern. So, if the Tongue happens to be in the central position, during Babbling, the Lips will constrict, and when the Jaw closes, will produce consonants like, b m w. So, b m and w, and vowels like, aah and ooh, so they are able to produce this.

Now, if the Tongue is in the front position, Lips will constrict, the vocal tract by meeting the t-producing consonants like, d and y, and also producing vowels like, e a and a y. So, when the Tongue is in the, front position. But, the Tongue is in the back position, the back of the Tongue will meet, the root of the Tongue will strike the Velum, and they are able to produce, consonants like, g and ng, and vowels like, o and o h.

So, these are these are the vowels, these are the consonants, the Infant is able to produce. And, these are the vowels, that the Infant will be able to produce. Now, the reasonable conclusion, from this model is that, these Babbling patterns, are driven by Motor, and not by Auditory processes.

(Refer Slide Time: 55:50)



Now, let us look at, the social aspects of Babbling. First, the parents interpret, the Infant Babbling, as a attempt to speech. So, Caregivers, they respond to Babbling, as if it were the invitation to conversation. Now, imitate Babbling within confines of their own language. So, what they do is, they imitate the Babbling, within confines of their own language, so they produce the right sound, und or they produce the sound, which is similar to the Babbling sound, and so the Infant listens to that, and they will to respond.

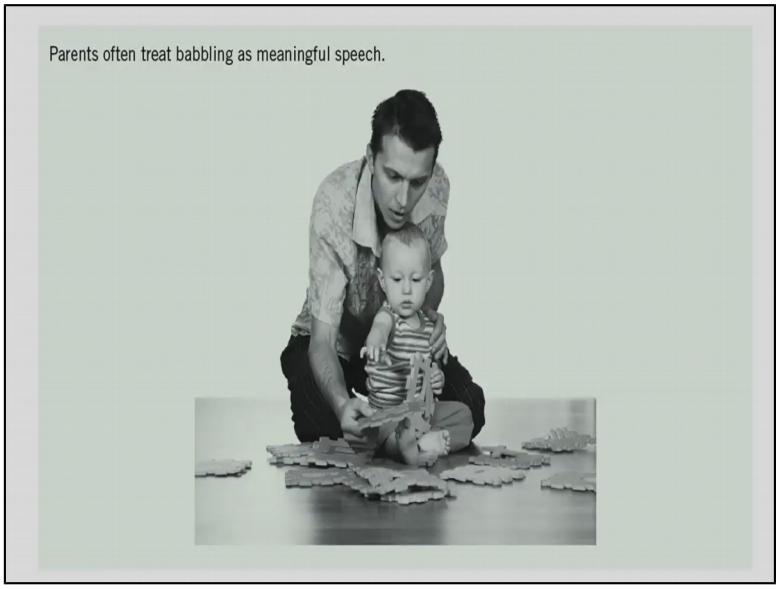
So, when babies babble, parents imitate them. And, the Auditory Feedback, helps the Infant, to home-in on the sound categories of the language, it is learning. Infants were Babbling, to elicit responses from Caregivers. So, the Infants use this Feedback modified Babbling sound, to like Caregiver's language. Now, there is something called, Object directed vocalization. So, when the in Caregiver, or the parent is speaking to the Infant, they use something called, Object directed vocalization.

They show the Infant, particular objects, and speak the name of the object, in the same Babbling language, of the Infant. So, Object directed vocalization, babbling uttered, as Infant approaches and manipulates novel objects. So, babbling uttered as Infant approaches, and manipulates a novel object. Parents respond, by naming the objects. So, when an Infant touches an object, the parent responds, by naming the object, with a word, that matches the Babble.

So, you have to be very sure. Or you have the point here is that, the object, that the Infant is touching, or manipulating, the parent responds to that Babbling, by producing a name of the object, which matches the babble, that the child is doing. And, that is very important.

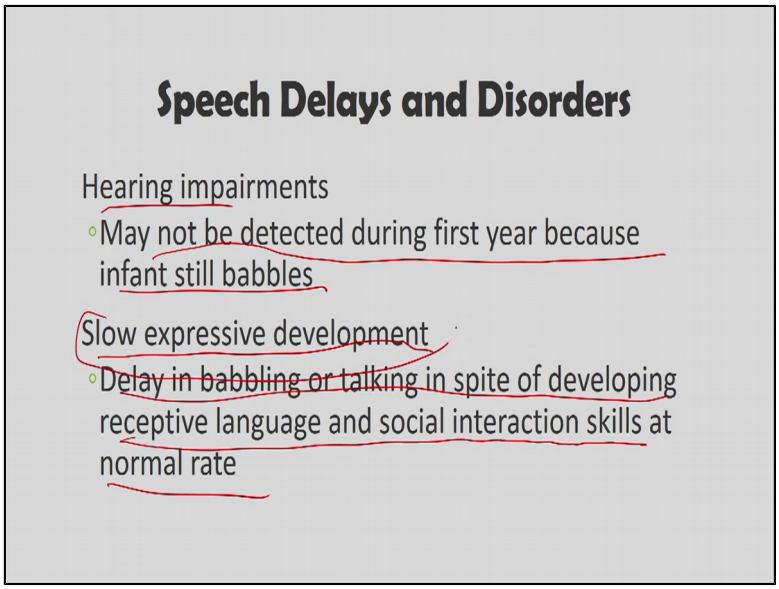
Now, Infant who has received this kind of Feedback, have large vocabulary. So, Infants, who get this kind of Feedback, from the Caregivers, or parents, they have large vocabularies, then Infants, who do not get this kind of Feedback. So, Caregivers respond with name for an object, that resemble the babble, and indicate Infants heightened attention, and readiness to learn.

(Refer Slide Time: 57:52)



So, parents often treat babble, as a meaningful speech.

(Refer Slide Time: 57:56)

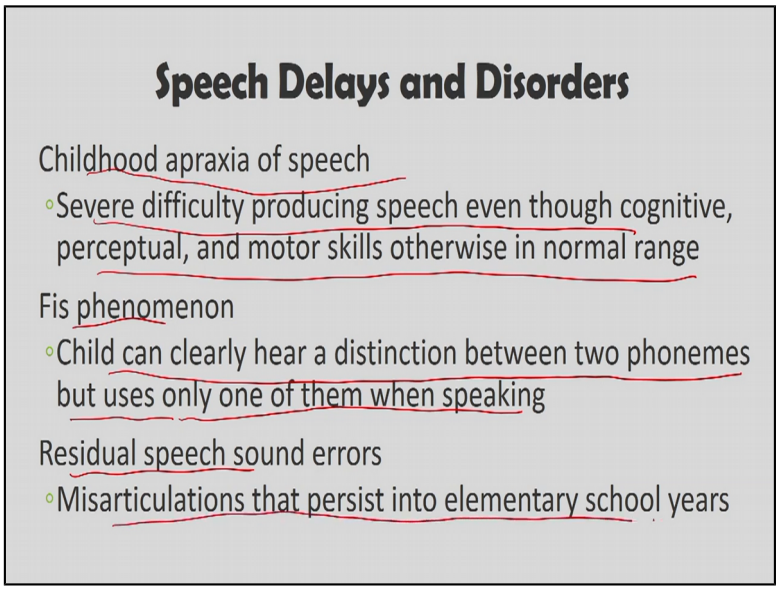


Now, there are something called, Speech Delays and Disorders. Now, hearing impairment, they go undetected, till second year of life. Control vibrations of vocal cord in Jaw movement, same in, it has been found that, for the deaf and the normal people, the Jaw movement, and vocal cord vibration, are same. However, coupling of Phonation, and articulation is impaired in deaf, suggesting that decision, or detrive feedback. So, since they do not hear, the coupling does not happen.

The Jaw movement, and the vibration of the vocal cord, happened, but this is not able to, hear it. They are not able to couple, the Phonation, and articulation, of the speech sound. So, hearing impairment may, not be detected during the first year, because Infants still babble. Now, the Slow expressive development, delay in Babbling, or talking, in spite of developing receptive natural language interactions, can set normal rate. So, delay in Babbling, is not always caused by, hearing impairment.

The delay in Babbling is sometimes caused by, Slow expressive development. For example, a delay in Babbling, in in spite of developing receptive language, and social in in interaction skill, at a normal rate. So, it could happen due to, Slow expressive development. It can also happen, due to childhood apraxia of speech, in which children experience, severe difficulty in producing speech, in spite of, normal cognitive perceptual and Motor skills. It could happen due to, a Slow expressive development.

(Refer Slide Time: 59:22)



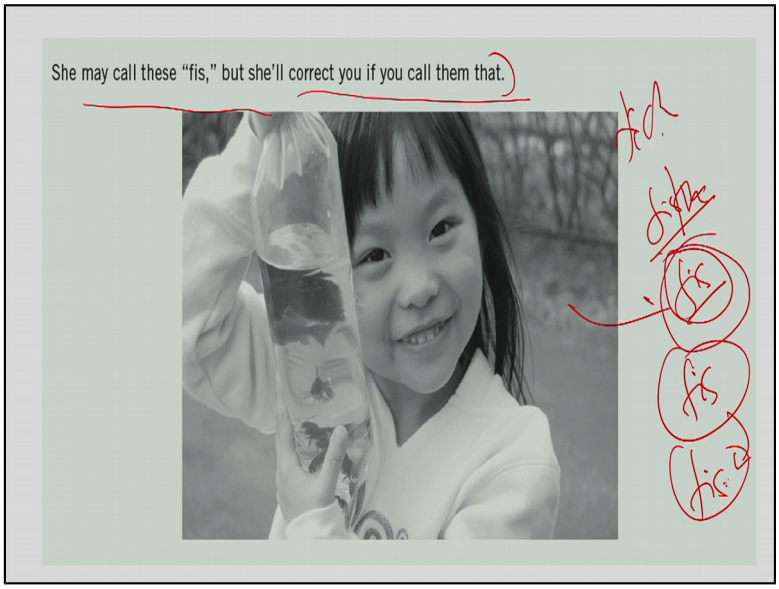
Or, it could also happen due to, childhood apraxia of speech, which is a severe difficulty to producing speech, even though, the cognitive perceptual and Motor skills of the children, are

normal. Now, they also mispronounce certain phonemes, and can hear two phonemes, but use only one. So, example, this is called the, Fis phenomenon. Child can clearly hear, a distinction between the two phonemes, which has been said to them. Fish, for example.

But, can only use, one phoneme. And, this is called the, Fis phenomenon. Children, who have late speech development, they also produce something called, Phoneme substitution. And, there is also something called the, Residual speech error, which is misarticulations, that persist in elementary school years. So, Residual speech errors is misarticulations, that persist into the elementary school years.

For example, Sometimes, Speech production is basically affected by, Residual speech error. For example, the pronunciation of the, l and the r, sound. And so, that happens. And also, Fricative and Affricates made, why? The Tongue for, s and s h, so they make error in, this also.

(Refer Slide Time: 1:00:30)



So, she may call, fis, but she will correct but she will correct you, if you call them that. So, when the adult says, fish, this child is not able to say, or the fish, the two phoneme, that it is hearing, the fi and sh, that that it is hearing, so she will be saying, the child will be saying, fis. But, if the adult says, fish, she will be able to correct you, saying that, this is wrong, it is fish. She will write the fish sound, but she is, the child is not able to produce, the two phonemes together.

And so, that is another error, which is out there. And so, that brings us to the, end of this section, end of this lecture, where we are looking at, how the voice is, or the speech is,

produced. So, what we did here is, we looked at the, Feedforward and Feedback control systems. The Feedforward system, controls the articulator, or the articulator is producing feeds.

And, the Feedback system, it provides Feedback, on to what speech is produced, and provides this error correction, does the error correction. So, the two ways. One is the, Somatosensory system based Feedback, and then the Auditory system based Feedback. Then, we looked at the two stream, the Dual Stream Model, where we looked at, the Ventral stream, and the Dorsal stream.

And, how the Ventral stream and the Dorsal stream, they combined together, to lead to Speech production. Further to it, we looked at, the DIVA model, which is another model, and a more complicated computational model, which explain how, speech is produced. And so, they talked about the, Feedback, and Feedforward systems, and the integration through system, of producing speech, and collection of speeches.

And, they integrate the whole idea of speech production, which has been, captured up till, now. Further to that, we looked at, how speech is developed in children. We looked at, the idea of Babbling, the several stages of Babbling, which are there. And, how this Babbling, lead to the actual Speech production, or development of Speech production, in small children, or Infants. We looked at, several stages of it, and how it is developed.

Further to that, we also pointed out, or we also looked at, the social consequences, of the Speech production, how Caregivers respond to the children, and how Object identification task is used, for developing of Babbling, and how this Babbling leads to, for the Speech production in children. And lastly, we looked at, some kind of speech errors in children. The speech errors like, Fis-phenomena, the Residual speech error, the Childhood apraxia, and Slow expressive development.

And, how these are, what kind of disorders are they, and how these disorders can lead to, either later Babbling, or the slow production of speech, in children. So, all in all, what we need is, we in the last lecture, we looked at, how speech is produced, and in this lecture, we looked at, the several theories of Speech production, or how speech is produced. Now, when we meet next, we will start another, down or more complex things, in Speech production. Now, we were focusing, on the phone level, the phonetic level, of the basic speech sounds.

Now, we will start with, looking at, what are Words, and how Words are produced, and what are the nature of Word production, and what is the way, the Words are produced, and how they handle, and the things like that, we will look in to Sentences, we look into Discourse, and so on, and so forth. So, next lecture, we will deal with, what are Words, what is Word actually, and what is the construction of this, what is the nature of this. But, until we do that, which is in the next lecture, which is Lecture number-9, we will meet again, and it is Goodbye, and Namaste from here. Thank you