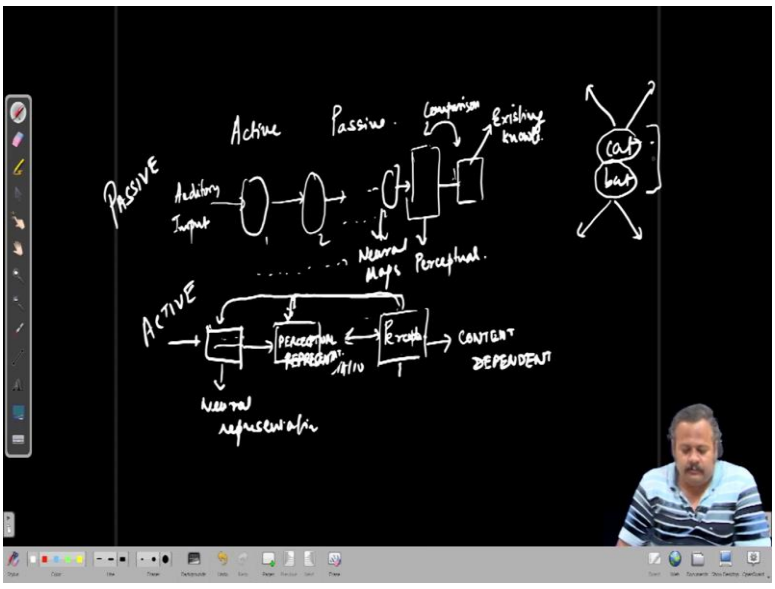
Cognition and its Computation
Prof. Rajlakshmi Guha
Prof. Sharba Bandyopadhyay
Biotechnology and Bioengineering
Indian Institute of Technology, Kharagpur

Lecture - 48
Speech Perception

Welcome in our discussions on language and speech. So, you have learnt about the components of speech and how speech is produced in your earlier lecture. And so today we will discuss about how speech is perceived. And so what do we exactly mean by speech perception? And that is essentially if we take speech the speech sounds which is which consists of a sequence of phonemes which together make up syllables.

And then syllables together make up words and the there is a stream of words which finally, provide the speech sounds. And so, understanding the meaning of each and every word and the sense of the entire speech together is what we mean by speech perception. So, in terms of the ideas of speech perception we usually think of there being established set of meanings from learning and experience and with the auditory inputs, we managed to compare it with the existing vocabulary or existing meaning of words and we match with them to frame the or actually the perceive the speech.

(Refer Slide Time: 02:12)



So, we can have it in this manner, where we have speech perception that is active and that which pass that which is passive. When we mean by passive, we essentially mean that we have an auditory input which is the speech in this case, which then has neural representation at multiple stages 1, 2 as we have know along the auditory pathway into the auditory cortex. And then we are comparing it with, there is a this is neural maps that are getting formed gradually over the different stages.

And then we have a perceptual map, that is created based on comparisons with the existing knowledge. And so this comparison tells us the meaning of the sequence of words. So, this is a unidirectional kind of process, when we talk of the passive way of speech perception, but as you have already known, throughout the our lectures and particularly even for speech attention plays a big role in terms of our understanding of speech, especially when there are many different speakers available.

Like we studied that attention in in the context of multiple speakers and talked about an experiment where the neural responses recreate the original low frequency spectrogram of the attended speech and not the unattended speech. So obviously, there is a act, there are active mechanisms in speech perception. And so, what that does is that they provide, it is essentially again the same where we have the neural representation finally at one stage.

And then from the neural representation, we compare with perceptual knowledge that tells and then based on the error or anything that tells us to modify the neural representations throughout the pathway. And I am sorry, the neural percept this should be perceptual representation after the neural representations.

So, we have just as above these are the neural representations, let us say this is the final stage which then has the perceptual representation. And then from the comparison, the errors or basically some top down input provides the change in the neural representation or even in the perceptual representation so that we make it sound like what or actually perceive it like what is matching.

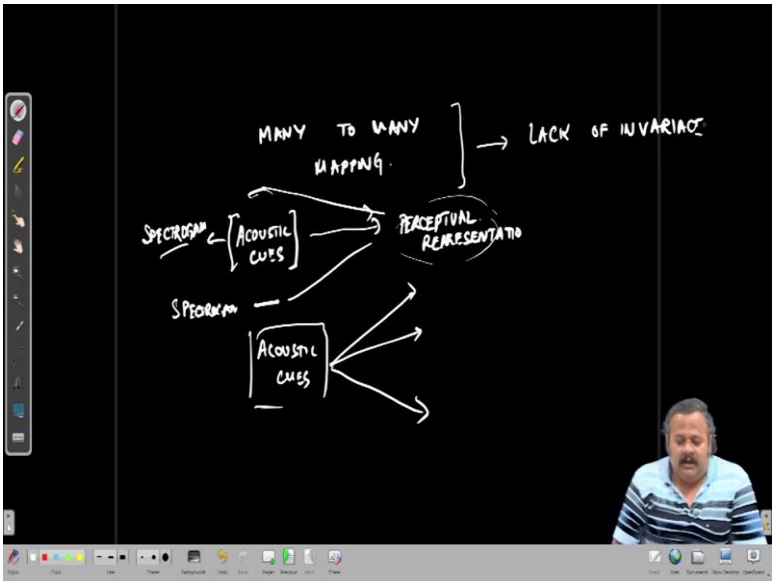
And this perceptual map based on knowledge is very much or the stored maps are very much context dependent. So, it is not just the auditory cues that are important in terms of speech perception, but there are many other things it is so if we take a simple word

depending on let us say a cat or a bat or something. Depending on the context of the sentence or rather the context I mean what the associated words are in the sentence.

The meaning of this bat can be very different and perceived as two different things, and even cat if we have some short form of category here, these also can be perceived as totally different things. Although the accused acoustic cues are exactly the same. So, the contexts, even the context matters and even the acoustic cues for the same sound matters, because of the intonations that we talked about as a component of speech, how long the vowel is or what is being stressed or and so on can determine the actual perceptual meaning of a particular word.

Although the acoustic features are different, then we can still perceive them as the same sound. Although one may be in an interrogatory tone or one maybe an with an emphasis, but still we are perceiving it as the same, in the same way although the acoustic content of that particular word is different.

(Refer Slide Time: 08:22)



So, this brings us to the idea of many to many mapping, that is major problem in speech perception or understanding of speech perception and being able to apply it through computational principles.

So, what we mean by that is what we were just discussing, that we have these acoustic cues a number of them which is essentially given by the spectrogram as we have

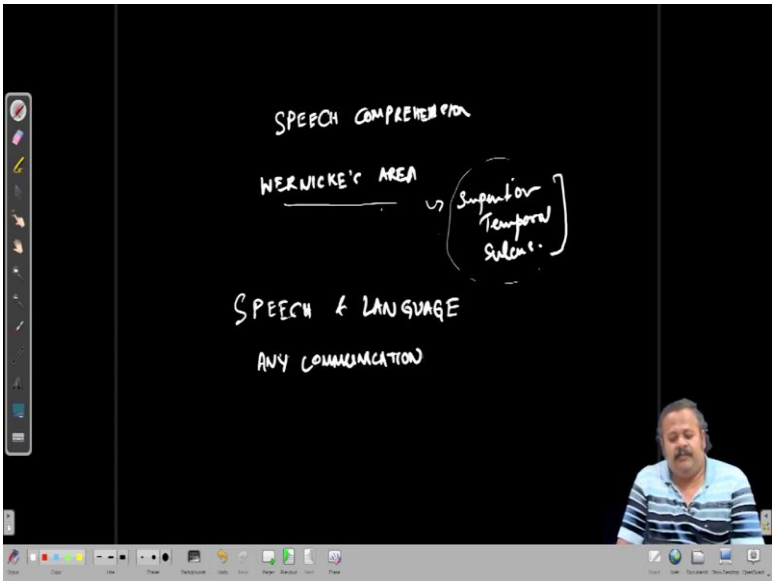
discussed earlier in one of the lectures um. So, these spectrograms of the same meaning so here is the perceptual representation. So, of the same word can have many different acoustic or spectrograms, but they are all perceptually the same, that is they appear as the meaningfully the same word.

Like, even if the pitch is different by the same person or it is being stressed differently, the spectrograms will change and so the acoustic cues are different, but the perceptual representation is the same, that we see we hear it as the same word.

Similarly, now the acoustic cues can be the same this is the same, but the perceptual representations are possibly many of them, like we talked about the cat with different meanings. And even the sound like depending on what kind of vowel is associated with it, can have different kinds of meaning, the same phoneme can have different kinds of meaning.

So, this many to many problem, which is also can be looked up as or the looked upon as a lack of invariance is the major important point that underlies speech perception and the difficulties of use of developing algorithms for speech perception inherently suffer from this lack of invariance, that is although the acoustic cues are same the meanings are different. Whereas, the acoustic queues are different the meanings is same, that is what we just talked about.

(Refer Slide Time: 11:33)



Secondly, in terms of the brain areas that are involved in terms of speech comprehension, it is believed that the Wernicke's area is a major region in the brain that is involved in speech comprehension. And similarly, this area is also equally involved or very critical in speech production. So, sort of sits in between the it is a motor area that is there are motor neurons that actually are involved in speech production and they are also equally important in speech comprehension.

So, throughout the temporal regions in the superior temporal gyrus, the main regions where speech perception takes place is the superior temporal sulcus and around there in the auditory cortical regions. And surrounding them very nearby is the Wernicke's area, very near by the superior temporal sulcus. So, STS which is again the superior temporal gyrus is also there which is involved in actually auditory processing and auditory perception. So, this region is close by to this Wernicke's area where which is very much involved in speech comprehension.

So, most of the studies about speech comprehension or speech perception from and the corresponding regions involved in the brain, have been through lesion studies in humans. And we say humans because it is a sort of a very very specialized ability among all animals, which is speech and language as you know already know.

So, in terms of expressive language or express our speech as an expressive part of it and there is a receptive part of it. And so that receptive part is basically which has to do with the perception and the expressive part of it has to do with the production of speech.

Both of these are unique in terms of human ability, in terms of abilities or skills in animals, it is unique in humans. But in terms of any kind of communication between animals of the same species or between different species, there are forms of vocal communication that is present between a within a group of animals, within let us say like the bird song, a song bird even rodents and non-human primates and so on.

The big difference here is that the speech and language have an infinite capacity in the sense of creativity, that the number of ideas that we can represent through speech and communicate with others is basically infinite. Which is not at all true for the other animal species that use vocal communication that is communication through sounds produced by the animal, which have a very specific set of purposes that are that are involved.

Like for the honeybees, it there is a kind of communication not verbal or vocal it is through their gestures or their dance, which indicates the entire herd of bees about location of food. So, that is also another kind of animal communication, which is nonverbal and then other animal communications which are verbal are for example, in many species it has to do with mate selection and courtship rituals and that is the specific purpose of that kind of communication.

Then there are communications by this by pups of mice or rodents in distress, which they emit which elicits a return behavior or rather a retrieval behavior of the mother, where the mother goes out in search of the pup and finds it and brings it back to the nest.

So, those are very specific and limited kind of communications that are present in the other species. Whereas, speech has infinite capacity, which makes it I mean intriguing and all the more interesting to study. However, if we consider speech as a tool of language, which is you know as a as something that is involved in expressing ideas or thoughts humans are the only ones that actually are capable of doing that.

And however, the again the although the animals are very different as you will see, when we talk of the different kinds of disorders of speech and language, we have a lot of things to learn or lot of mechanisms that are involve the neural mechanisms, that are involved in speech production and speech perception are involved.

The mechanism that are involved in speech production and perception can be studied in animal models. And that will be the topic of our next lecture, where we understand what we can learn about speech communication or language from animal communications and they are neurobiology or their computational aspects at the neural level in the next lecture.

Thank you.