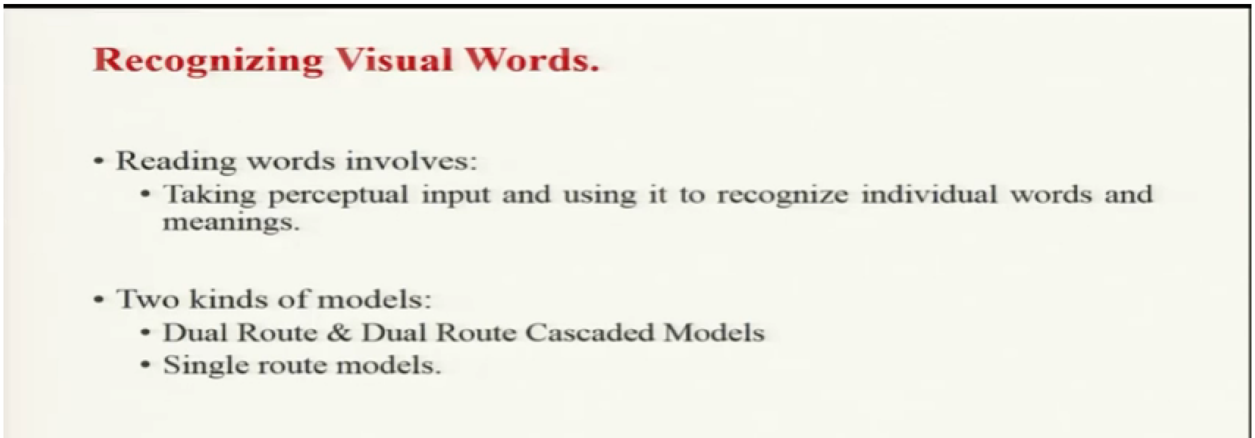
Lecture – 29

Recognising Visual Words

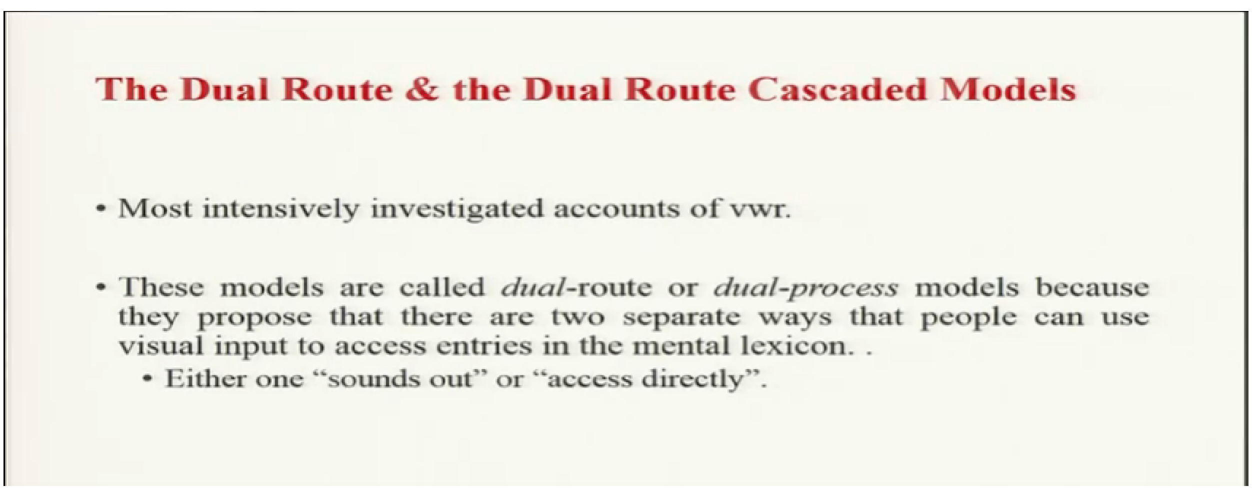
Hello and welcome to the course, introduction to the psychology of language, Dr. Ark Verma, from IIT Kanpur. And we are running in the sixth week of the course. We are talking about reading, we Reece aspects of it in this course. Today we will talk about, a specific aspect of reading that is recognizing visual words. So, when you see words, how do you really make sense of them, do you, say for example, as we were discussing in the last lectures, immediately you know are you immediately able to convert, the grapheme into phoneme and kind of make sense of whatever it is and then, move on to lexical access and making meaning. Or say for example, there are other processes that you might, need to do. How is that really a completion that's basically what we are going to look at, in today's lecture.

Refer Slide Time :(0:59)



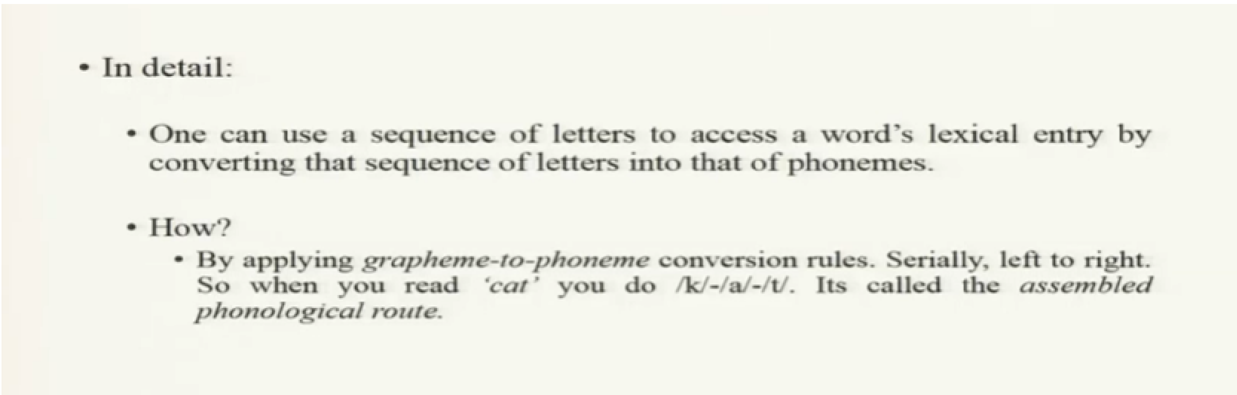
Now, reading a verb, reading a words. Sorry, involves taking the perceptual input that is the visual impression of the word. And using it to recognize individual words and kind of form, from that, move to lexical access and understanding word meaning. There are two kinds of models that have been sort of genetically proposed, with respect to how people accomplish word recognition and basically, these are called the, 'Dual Route' and the dual route cascaded models. And then the single route models, which are more from the machine learning you know, neural networks sort of domain.

Refer Slide Time :(1:34)



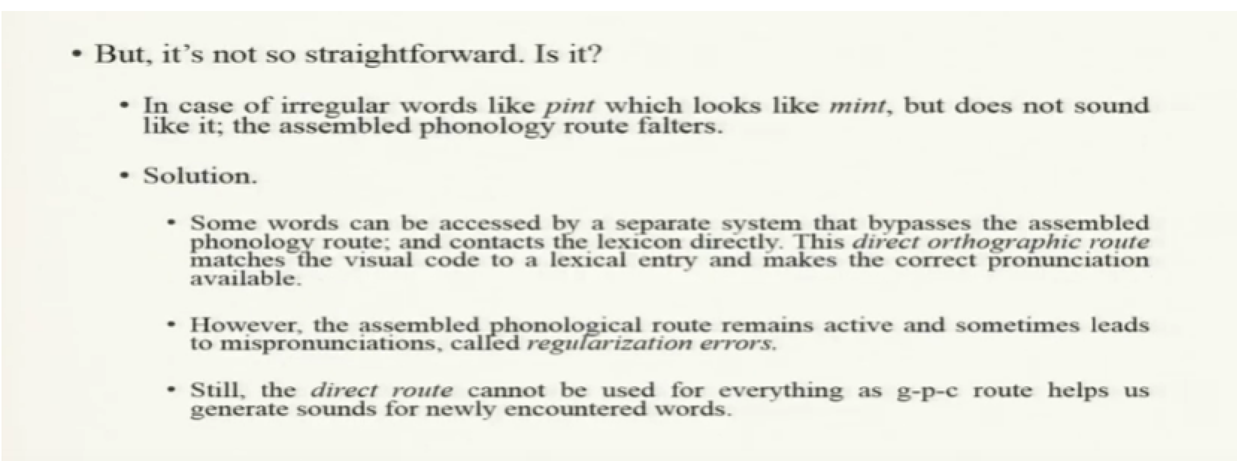
So let us look at, the models. First let us look at the Dual Route and the Dual Route cascaded models. Now, the Dual Route accounts are probably the most, extensively investigated accounts of visual word recognition, these models are called, 'Dual Route or Dual Route you know, access Cascaded Models'. Because, they propose that there are two separate pathways, there are two separate routes, using which people, can use the visual input to, recognize individual words. So say for example, in one of the ways, you can just sound out, the given word that is converting each of the letters, to its sound and then kind of you know, combining all of those sounds and then speaking it. Or say for example, the other way is a more direct way, you look at a word, you directly map it, with something that you already know and you can you know, use that as a way to read the word.

Refer Slide Time :(2:32)



In detail let us look at, say for example, in the first model that in the first route that is the sounding out route. What you will need to do is? you will need to kind of, get to the sound of each and every letter, in a particular word, you need to concatenate those sounds, attach those sounds together and then create a you know, a phonetic phonological version of the word and then use it, to perform lexical access, reach to the form and then, later reach to the meaning. This is basically called the, 'Grapheme'. I mean, this is basically called a, 'Indirect Route' or the 'Phonological Route'. Okay? Basically, using it uses what are called, 'Grapheme to Phoneme Conversion Rules', to reach from letter configurations, to the sound configurations.

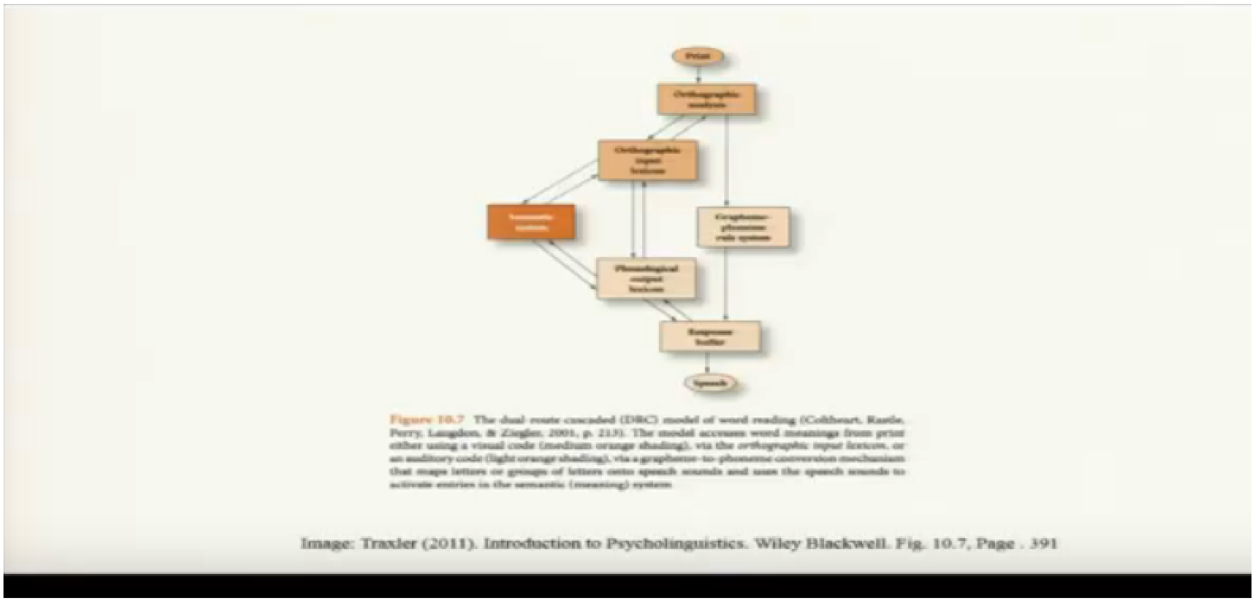
Refer Slide Time :(3:17)



But, it's not really very straightforward and I mean, if you kind of think that I am saying that you know, you will very easily use these letters, to convert them into words, into sound and then speak the name. And following something called the, 'Graphing Phoneme Conversion Rules'. We really need to also talk about, whether the M what these rules are and whether they are very easy to apply. Now, one of the things that is very peculiar to English or some of the other orthographies as well is that there are both kinds of words in English, there are regular words in English and there are irregular words in English. What are the regular words? The regular words are we're in, a similar bill form to phoneme conversion rule can apply and it will lead to a similar kind of a sound, you know generation. The irregular words are where you cannot really apply the same rule, to get the same kind of sound. Then we give an example; say for example, M I N T is mint. But, P I N T is pint. Or say for example, TO is to and GO is go, HAVE is have, where whereas, SAVE is save. So you see that, there are multiple of these instances, where the same configuration of letters, will not lead to the same configuration of sounds. Now, if this is the case, the phonological route will not really be, you know, always effective in creating, reliable and a consistent, sound based representation or a phonological representation, of the incoming visual word, of the visual word that you're trying to read.

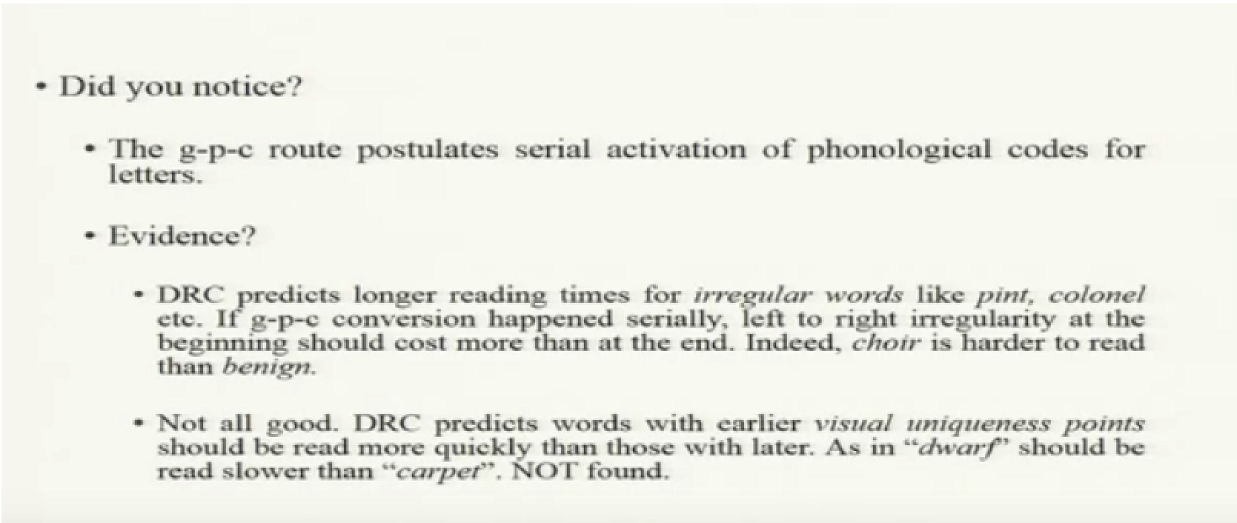
Now, this is basically the something that can lead, to you know, second learn, second language learners of English, into making mistakes, also initially kids when they're learning to read, they would also make these kind of mistakes. Now, the mistakes that people sort of make, in this you know in this route basically are referred to as, regularization errors. What are the regularization errors? When the assembled phonological route, remains active and sometimes leads to miss pronunciations that is called, 'Regularization Errors'. Okay? And what is the other way to do this? The other way is actually when you come across these exceptions, yesterday I was talking to you about, Bou, Qet, pronounced s bouquet and psy, CHO pronounce a psycho and not as, a psycho, something like that. So how do you really read those kind of words? Now for those kind of words the dual route models specify a particular separate route, which is called a, 'Direct orthography Route'. The direct orthographic route basically, matches the visual port directly to a lexical entry and makes the correct pronunciation available, you do not need to assemble, the photo logical route, you know you know assemble the phonological representation, by combining each letter and creating its sound. Because, if you do so you will lead and you know eventually leads to particular kinds of errors. The errors would be of the kind that I just mentioned, the regularization errors, suppose if you are kind of you know, using the assembled phonological route, to read some of these exception words, then you will obviously fall into mistakes and those mistakes will be called as, 'Regularization Errors'. Also, however the direct route cannot be used for everything, as the GPC routes will generate sounds, also for newly encountered words. So for example, if you're reading the FIR, a word, for the first time in your life, supposes say for example, you come across the word like psychology or say for example, they're also interesting words like, GENRE, genre. You know so for example, if you're coming across these new you know, words and you've not heard them, being said earlier, you are more likely to use the similar phonological route and then come up with, incorrect pronunciations. So, the direct route can only kind of make available, correct pronunciations, for words that you already know or for words that you've heard earlier, in time that is basically what you know, this looks like.

Refer Slide Time :(7:37)



And here I can show to you, the dual route cascaded model, of reading, this was put forward by you know, max cold heart, Cathy Russell and Perry and language and Ziegler in 2001. And this is how the model looks, you can see the model starts at the top, you are looking at, something in print, then basically we'll carry out an orthographic analysis, basically looking at what is this word really composed of? What are the letters and so on? And then once you have that, you'll basically be and carry out, what is called the orthographic analysis and you'll either move to the orthographic input lexicon, wherein you're kind of analyzing each of these letters, you kind of then doing the phonological output lexicon, you're carrying out the GPC conversions and basically, you know you are kind, yeah, so basically you're then reaching a response buffer. So, this is the path on the right, orthographic analysis graphing phoneme rule system and response buffer. So you convert that and you reach the assembled phonology, this is called, 'Assemble Phonological Route'. The one on the Left, where basically is the, system that kind of is the direct route, so you do go from go to the orthographic input lexical, what is, what are the words that I know of and then you kind of directly have the phonological output lexicon, it directly gives you the correct phonological output and then it can go to the response buffer, then you have the speech. You have sort of, also a unit called the 'Semantic System' that is attached to the orthographic analysis. Because, it kind of simultaneously activating entries, as soon as you are reading the word, this is activating entries in the semantic or the meaning system at the same time. This is basically, what is called the dual route cascaded model? Because, the flow of activation in this model is in cascade in nature, as soon as you start doing the orthographic analysis, the phonological and the semantic analysis are already started. Okay? So, that is basically why this model is referred as or dual route cascaded model.

Refer Slide Time :(9:44)



Now, let us go into a little bit more detail, if you see in the assembled phonological route, the GPC route basically or the graphne phoneme conversion pattern the path, potion its serial activation of phonological code of letter. Let us say for example, if I have to read play, plae, PLAY, will kind of you know, be generated in a sequence. Now, the evidence for this, the evidence basically is, say for example, if you see the DRC route predicts longer reaction to, longer reading times for irregular words like pint, colonel, etc. If the GPC routes, you know if the GPC conversion had to happen serially, left to right is regularity at the beginning, should cost more in terms of time, as compared to left and due to the irregularity in the end. Now, what am I actually saying, I am saying is that if you were to read these words, in a serial manner as the GPC thing is saying, the irregularity in the beginning of these words, because you're converting will prove to be more costly, as if the irregularity is the end of the word. So, basically and that is exactly what is found, say for example, if you were read over like choir, CHOIR, you'll find that it is harder to read, than benign ,BENIGN .Because, benign the irregularity is more towards the end and inquire the CH is being pronounced esker is in the beginning. So this is one. Now, this is also not completely, alright. Because, say for example, the DRC predicts that words with, visual uniqueness points should be, read more cave quickly than, those with laugh, don't we'd have, which have later visual uniqueness, say for example, dwarf the visual uniqueness point DW is you know, earlier and this is read, slower than carpet. This is the prediction of the model, which is not really you know, confirmed by the literature. So, not all the assumptions of the dual route models are kind of you know, found to be correct and obviously then that is why there remains scope for more research and more understanding, of the processes involved in visual word recognition. Now, let us move to the single route models, the single route models basically have grown out of, the tradition of parallel distributed processing. And basically, they are based on neural network modeling and so on. They typically contain three kinds of units, they contain orthographic units, which represent the orthography, the written version of the letters, there they have phonological units, which can contain the sound representations and they have semantic units, which represent the meaning based representations. There's also a hidden layer of processing that, between and that kind of connects, between each major component unit.

Refer Slide Time :(12:44)

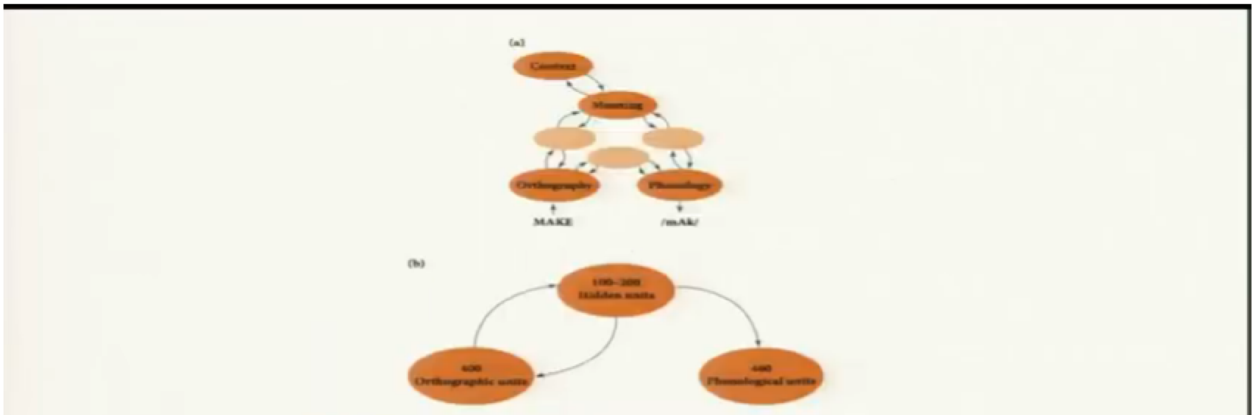
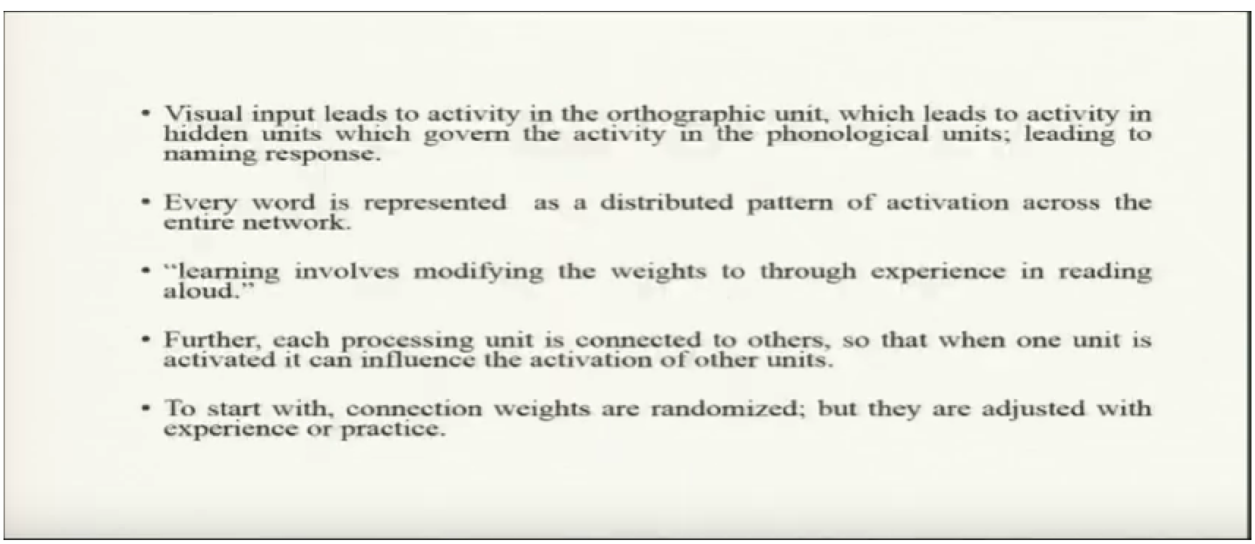


Figure 10.8 A general framework for lexical access (top) and a more specific mechanism for generating pronunciations from visual input (bottom) (from Seidenberg & McClelland, 1989, pp. 526–527)

And that is what basically is the, processing part of the model. So, you can see the model here, in this figure, you'll see that there are, there's content meaning and orthography and phonology. So context is kind of at the top. But, then you have meaning that is, the semantics in its orthography and phonology, you will see each of these units have a layer of hidden, you know, I have a hidden processing units in the middle, which basically will carry out, all the processing. And if you look at the hidden units, there are 100 to 200 hidden units, with respect to between context and meaning and then there are 400 orthography units and a 464 a logical unit. So, basically here in, in B you have a more specific mechanism for generating pronunciations, so if you look at the entire model, Part B is the model is the part of the model that is used to create pronunciation. So how many units will be required for analyzing orthography, to phonology?

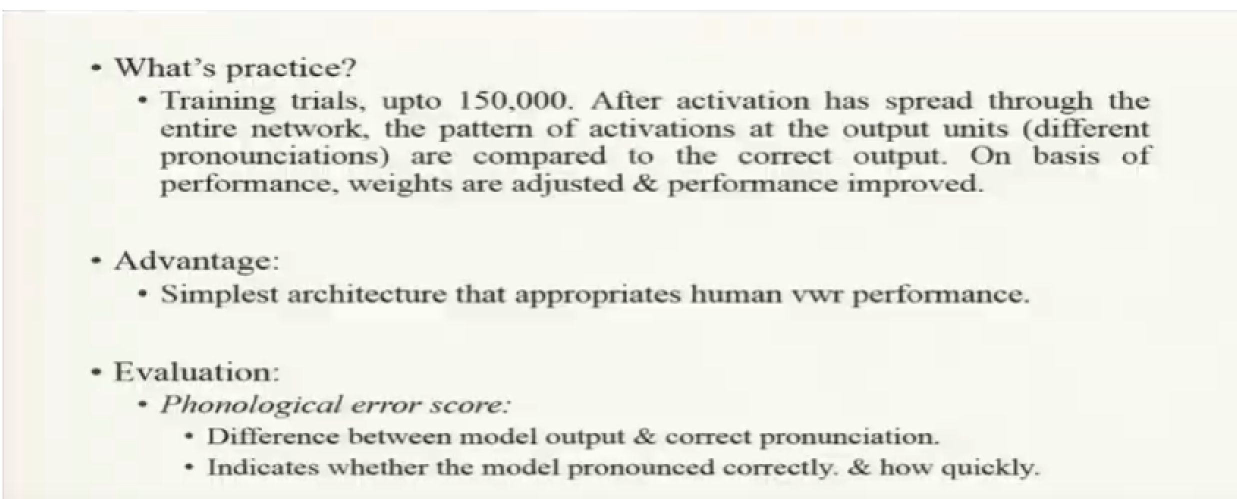
Refer Slide Time :(13:47)



Now, basically how does processing happen in this model? Whence the model encounters the visual input, it leads to activity in the orthographic units, which then leads to activity in the hidden units that govern the activity in the phonological units. So, as soon as the, activity is registered at orthography units, by other hidden units, it will start creating some activity in the phonological units part, leading to the

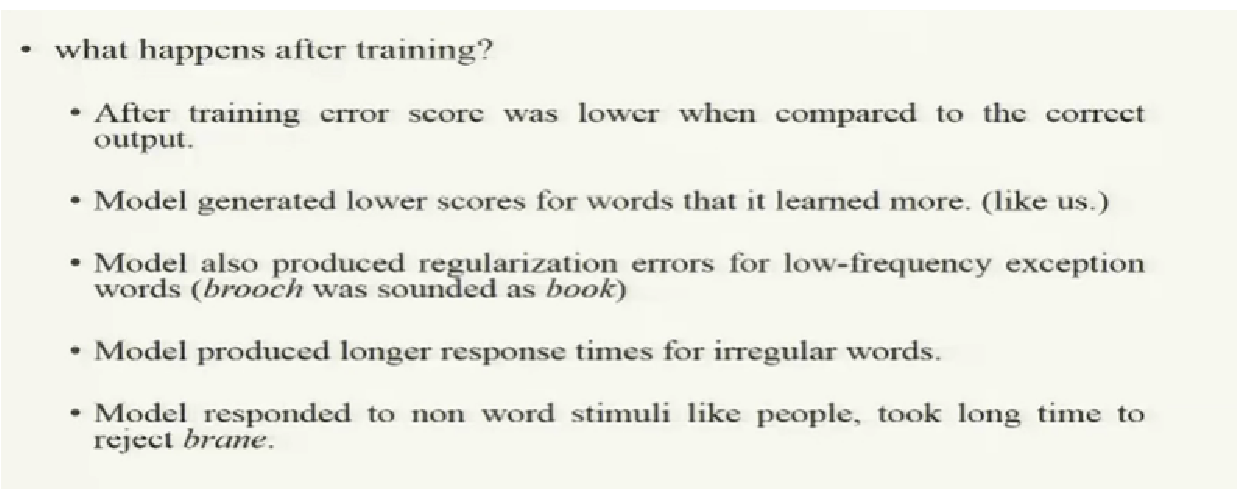
naming response, leading to generation of the sound code. Every word is represented as a distributed pattern of activation, across this entire network, learning basically in this model, involves modifying the weights through experience, in reading aloud. Say for example, the more and more experience that the model gains, the more basically you know, it modifies the weights it basically learns. Okay? How many units I have to put in the orthography part? How many words I have to put in the phonology part? How many do I have to put in the meaning part? And these number of units will probably be, adjusted with respect to the models performance.
So, say for example, the model is performing purely, you can adjust the weights, to you know lead to gradually better performance and once the model is, giving near perfect performance that is sort of the resting count of. Okay? This is the actual number, of units that is used for this particular representation, moving further, each processing unit is connected to others, so that when one unit is activated, it can influence the activation of other units. So for example, if you're hearing words or if you're reading words, in this case, reading words as soon as you, there is some activity you know, registered in the reading part, in the visual part that kind of will create activity in the orthographic part and as soon as, some activity is registered in the orthography part, both the meaning and the phonology part, will start receiving some activation. Okay? So and to begin with say for example, the connection weights between all of these units are orthographic and the phonology the meaning units are sort of randomized. But, they're adjusted with practice or experience say for example, in the first iteration, in the first hundred after first hundred trials the model gives up, forty percent correct meanings, but 40 percent corrects the official recognition, but 60 percent incorrect .So, you kind of again adjust the weights, you try and steer it to getting, the maximum correct performance.

Refer Slide Time :(16:14)



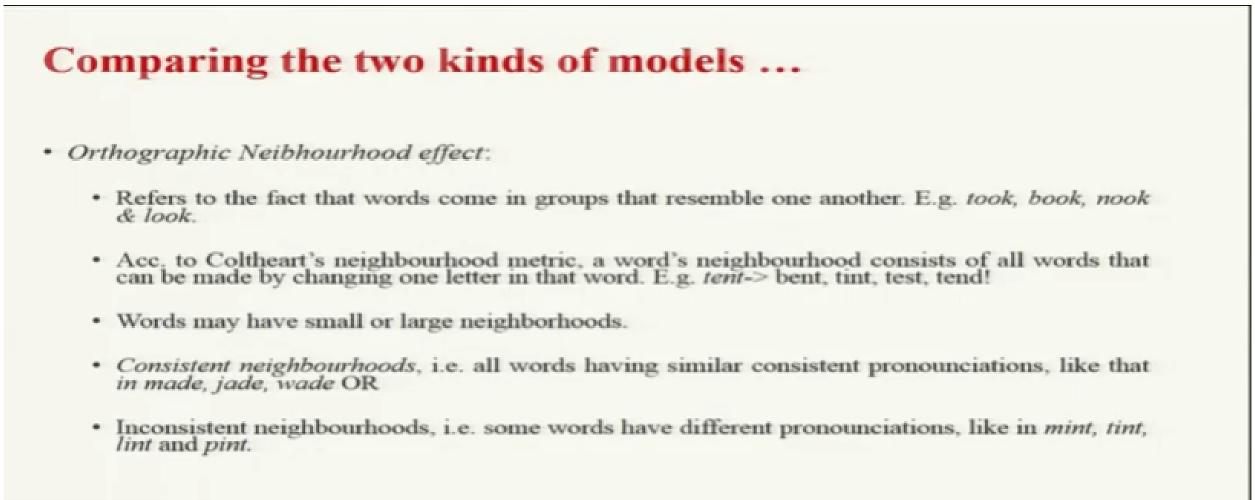
Now, what is what do you actually mean by practice? Practice is basically the training trials that you expose the model to, you kind of put some randomized weights and you give the model, the input and you kind of expect correct, recognition from the model. Usually in the beginning the model will not really be, able to give you correct performance. But, you kind of do the process iteratively, gradually the model trying to figures out, what are the correct weights, using which I can reach the sort of a correct pronunciation. So, these training trials, usually around up to you know 10,1,50,000. Basically are used for the model, to learn you know, what are the processes, what are the things that I need to do, in order to get to the correct recognition .Okay? So, after activation has spread through the entire network, the pattern of activations of the output units are compared to the correct output ,this is the pattern of activation that the output unit ,this is what my result is and you correct, you know compared with the actual result and you try and match this. And this kind of the weights in the model keep get jesting, keep getting at just and, and you know the, output units keep getting different kinds of outputs and it keeps getting compared with the correct output. As soon as this comparison is closer and closes that is how the, you know number of units here will start getting, the sort of fixed. Okay? So these weights are therefore adjusted to get to the best performance of the model. Now, one of the advantages of this kind of simple, architecture of the of the single route networks model is that it kind of resembles human performance, very well, it in the data and output and the response patterns of the single network, a single route network models is, found to be closely matching the performance output of the actual models. Now, we can sort of you know, try and evaluate the performance of this model as well, by various ways, one of the ways is to look at the photo logical error score. Say for example, how well or how correct the phonological representation that the model has come up with, matches the actual correct program you know, pronunciation. Now, basically how do you do it is you differentiate or you compare, the model output with respect to pronunciation, with the actual pronunciation and you indicate basically, whether the model has pronounced a particular word correctly or incorrectly and basically, the difference in the models pronunciation and actual pronunciation, gives you an idea of the phonological error that the model has made. If the phonological error score is minimal, it means that the model is performing all right, the phonological error score is slightly larger, you probably need to adjust the weights, in order to get the model closer to the actual pronunciation.

Refer Slide Time :(18:58)



Now, what happens after the model is trained for a lot of time, you know it's been, say for example, five hundred thousand trials or something like that, the model has started giving, better output. So basically, what happens is, after training the scores, the error scores would be found lower, as compared to you know the correct, winner when compared to the correct output. Also the model was found to generate, lower scores for words that it learned more, so the more practice it has with particular words, the better it gets with reading those particular words, also the model kind of interestingly produced regularization errors, like the humans, would do in the low frequency words, which it would have got, less time to learn with, also it was found that the model produced, longer response times for irregular words. So, it seems that the model is, kind of using the phonological, you know, activations in a slightly more primary sense and that is why it is leading to these kind of errors. Also a model responded to non word like simply, it's say for example, brain, BRAIN, it found it slightly harder to reject those words, much like the human studio, it's following a process which is very similar, to what the humans kind of do.

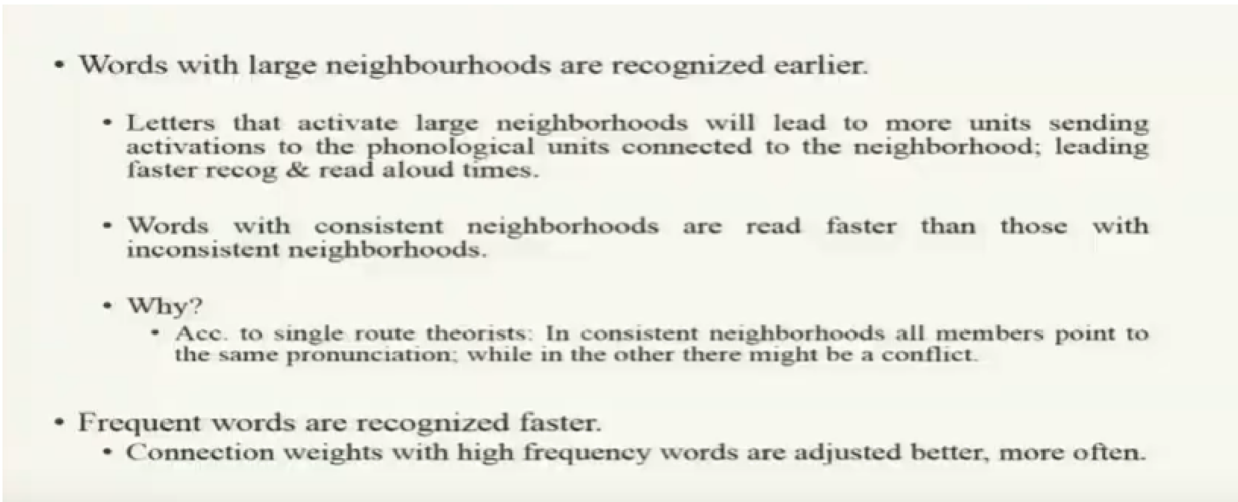
Refer Slide Time :(20:15)



Now, you've seen the do Route cascaded models. And you seen the single Route models. Now, what we can do is? We can compare the two models, on different kinds of word recognition phenomenon. Let us look at, say for example, the orthographic neighbourhood effect. Now orthographic, orthographic neighbourhood is basically, specifying say for example, if there is a word, how many say, how many letters you can change, to create another word. Say for example, if there is a word called, 'Book'. You can change just one letter, to change to you know, a form another word look or another letter took and you know, two letters to lead it to show. So for example, the idea is the number of changes you can do to a particular word, to create another word, basically defines how many neighbors, sort of it has, so if you can change one and how many letters, how many different words you can come up, by changing just one letter, how many different words you can come up with changing two letters, how many different words you can come up with changing three letters.

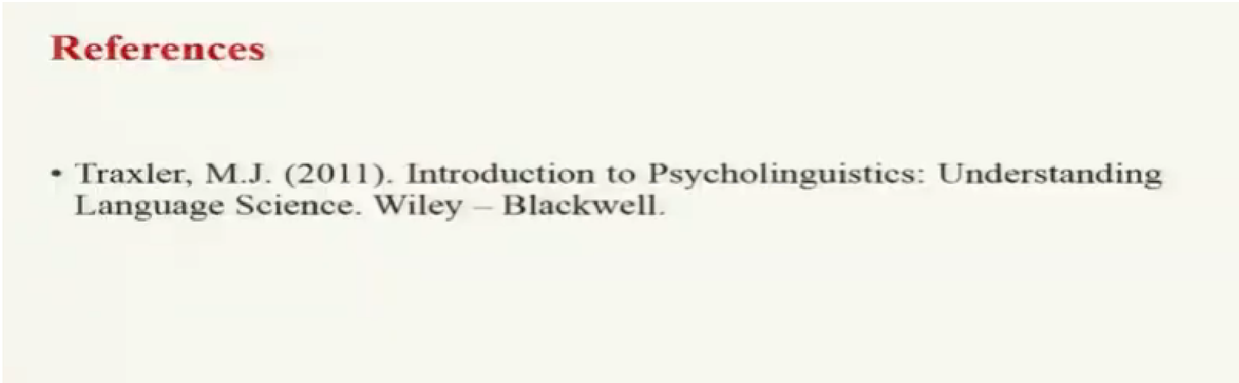
And this is the progressive; this is the orthographic neighborhood, of this word. The lowest number of changes needed and the more number of words you generate, it sort of it gives you an idea about, how many neighbors that this word have, according to cult hurts neighborhood metric, so max cult had pointed out, had devised a way to measure, the neighborhood of particular words and he found out that a words neighborhood basically, consists of all words that can be made by changing just one letter in that word. So, if you can just change one letter and create say for example, 10 words, then 10 is the neighborhood size of that word, you can create one letter and create 1000 words and 1000 is the neighborhood size of that word. And words, different words may have small, versus large neighborhoods, so that is also there and in word recognition experiments it has been found that words with large neighborhoods are slightly faster to recognize. Now, there are also things like consistent or inconsistent neighborhoods. Say for example, all words having consistent pronunciation, so book, took, look, shook, have consistent pronunciations, were say, say for example, there could be inconsistent neighborhoods also, say for example, words like mint, tint, but then you have another neighbor like pint. Okay? So, then they have in, you know the neighbors can have different pronunciations and that would be referred to as, in having, word having, an inconsistent neighborhood.

Refer Slide Time :(22:50)



Interestingly it has been found here, for example, words with large neighborhoods are recognized earlier. And why would that happen, it's basically because letters that activate large neighborhoods, will lead to more units ending, activations from the orthography to the phonological units, connected to that neighborhood, leading to faster, reading a recognition and naming times. Also, it has been founded words, with consistent neighborhoods are also read faster, than those with inconsistent neighborhoods ,you might wonder why, it's basically because inconsistent neighborhood all members, would point to similar pronunciations, versus in inconsistent neighborhoods all members, will start pointing out to different kinds of pronunciations that will lead to difficulty and slowing down. So, words with inconsistent name with, consistent neighborhoods therefore are read faster, as compared to words with, inconsistent neighborhoods. Also, one of the more important effects that you kind of, need to explain is the frequency effects, in the single route models, basically what you can do is, you can adjust the connection weights, with high frequency words, slightly better as compared to the you know, in the low frequency words. So because the model, comes across the higher frequent what's more number of times, it knows or it learns to read those words better ,it learns to adjust the connection weights for those words better, as compared to less frequent words. And that is why the high frequent words will also be recognized faster and you know, with less error, in single route network models.

Refer Slide Time :(24:28)



This is all from me about, the new kinds of models of word recognition that we have read. And I will talk to you about, a different process in the next class. Thank you.