Psychology of Bilingualism and Multilingualism
Professor Ark Verma
Dept. Of Cognitive Sciences
IIT Kanpur
Week – 03
Lecture – 13

Hello and welcome to the course introduction to the psychology of bilingualism and multilingualism. I am Dr. Ark Verma from the department of cognitive sciences IIT Kanpur. We are in the third week of the course and from today onwards I am going to talk to you about speech production in bilinguals and multilinguals. What is speech? You know speech is in some sense it is one of the most fundamental skills that we you know take for granted but basically that is the language that we sort of you know that is how we understand our language or communication abilities. Writing came much later, reading came much later but one of the primary manifestations of human communication ability is language which is expressed through spoken words, spoken sentences and hence is basically referred to as speech.

The question that we are going to talk about in the coming two, three lectures is mainly concentrated about how is how do bilinguals speak or how do say for example they manage speaking in two languages without a lot of interference from either one, how do they switch across languages, how do they code mix a lot of times, how do they basically able to, how are they basically able to communicate effectively having known or you know having known two or more than two languages. Now the thing is if we need to understand how you know speech production works in bilinguals or multilinguals it makes sense for us to also you know pay some attention to how is speech production accomplished by monolingual speakers as well because that would give us a baseline in some sense to understand how would speech production actually take place in case of bi or multilinguals, okay. So as I already said majority of human communication occurs via speech, it includes a range of steps from conceptualizing an idea that a speaker wishes to convey and to basically a set of behaviors that would change the physical environment. Say for example if I give you a particular topic to speak or if I intend you know stand here and intend to speak about bilingualism the first part of the process is for me to conceptualize what do I want to talk about, the second would be formulating how do I want to talk about it and the final would be articulation where I will create you know some kind of behavior which is basically moving my vocal apparatus creating sound that carries meaning which will be understood by my listeners.

So it is this tripartite process of conceptualization formulation and articulation that basically explains the process of speech production and this would be true for

monolinguals as well as bilinguals and multilinguals as well. Let us take an example you know for example a compliment conversation how does this really start say for example you could you meet somebody you like let us say their shirt and you go on and you tell them oh I like your shirt this is a very good color and the person would listen it and would react appropriately saying oh thanks and then you can you know take the conversation forward by saying oh where did you get it oh I got it from this particular market in this city and so on and so forth. So this is just an example of a typical conversation we have all kinds of conversations say for example a lot of times you will see that conversations typically revolve around generic formulas for example if you want to have a weather conversation is typically you know a decent conversation starter when you are meeting somebody for the first time so you know typically people would be like oh today is really hot or cold or rainy or anything outside and the person would say yes yes I am also feeling cold or I am also feeling rather warm and then you would sort of say okay I hope the weather becomes warmer or colder as it goes ahead and so on and so forth. So typically if you see this process of having a conversation basically has these multiple steps where you are talking about a particular topic you are conceptualizing how you will put yourself across and then you formulate the right words the right sentences the right form of how you want to convey the message and going forward you can basically you know take the message forward to other topics and so on and so forth. So there are you know it could be a lunch time conversation as well there could be any number of conversation that we have in the same sense but if you look at these conversations closely you could ask yourself certain questions for example how do we retrieve these linguistic representations say for example when we want to talk about let us say somebody's shirt so I know what a shirt is I have the conceptual representation of a shirt or let us say if I am going to talk about weather I know what weather is what I am talking about I am talking about the general climatic condition of that place in that moment.

So how do we sort of from these thoughts how do I pick up linguistic representations that you know which are best suited to convey this idea and then how do we organize these representations we do not speak you know randomly we do not speak just a constellation of words which are not organized we typically speak in a particular order we typically speak when we speak our words are grammatically correct say for example I will not say oh the weather how it is oh the weather how is it we will basically oh how the weather is today something like that. So we also organize these representations in a particular order in sense of grammar as well as semantics for example if you want to talk about a particular idea you will basically you know organize your you know ideas in a particular sense that it seems coherent and it seems like a logical flow of you know conversation. So finally the question that one can ask is that how are these representations once you organize them retrieve them how are they translated into a form that the motor system

can use to generate the physical actual gestures which are utilized to articulate them. See the first two parts conceptualization and formulation are still abstractions of the process of speaking but the final part which is articulation is actually a physical activity because you need to move your tongue you need to move your mouth you need to stop the flow of air in your vocal apparatus in order to speak in order to create speech. So these are the three very important questions that we can ponder about speech production in this lecture I am going to talk mainly about monolingual speech production but I will encourage you to you know take a pause at every moment and say oh how would this happen when a person has two languages and we will basically see how we sort of go ahead with this.

So as I said the tripartite process of speech production has three parts conceptualization, formulation and articulation very basically conceptualization means thinking of something to say what do I want to talk about formulation how do I figure how do I you know express this idea in the best possible way and articulation is the final part of moving the muscles to produce the sounds that can be perceived by the listener. A very interesting model a very interesting theory on this account was presented by William Levelt in a model which he named as the Weaver plus plus model and we are going to sort of go through very briefly with some of the stages of this model you can see that at the top here is the conceptual preparation in terms of lexical concepts what is a lexical concept? A lexical concept is basically concept which can be expressed in the words of our language in case of bilinguals it could be you know it could be a concept that could be expressed in two languages or three languages you know whatever the number of languages we know typically you will come across more than one lexical concept at any point in time so you need to do lexical selection you will need you will reach a level after that which is called the lemma level which has both semantics and syntax information you go to morphological encoding where you are basically wondering about the type of word that is going to be used you create a morpheme then you go to the phonological encoding part where you are sort of you finalize the morpheme that you have to communicate but you have to assemble sounds for it create the phonological word go to how the word has to be you know actually produce in phonetic encoding and then is the gestural score which leads us to sort of you know produce speech. Let us go in a little bit of detail of these models because at this point in the slide I have just sort of given you an overview but we can obviously visit it in a bit more detail and try and appreciate how this would change or what will be the special tweaks in this model that we need to do if we want to explain bilingual or multilingual word production in this. So as we have seen speech production can be viewed as a process involving a sequence of mental steps each mental process accomplishes a smaller goal a sub goal and the output of one mental process works as the input of the next one. Now this is something which is peculiar to this model it basically follows what is called a feed forward kind of a process.

So let us say we start with something like okay you know conceptual preparation which is at the top here and during conceptual preparation we are choosing about the ideas that we want to speak about we are choosing the idea that we wanting to speak about or do I want to have a lunch time conversation do I want to talk about whether do I want to talk about the nice dress that a person is wearing what is it that I want to really talk about. The output of this process once you have a sense of what you really want to talk about is a lexical concept and a lexical concept is in basically an idea for which your language has a label. So the idea is say for example if you want to talk about you know let us say beauty or you want to talk about a particular person's shirt or whether your language should have words to be able to express that idea. To take a more concrete example if you want to talk about a female horse there is a word in English language called mare and if you are seeing and talking about a female horse you can use the word mare and it basically encompasses all the meaning that you want to convey in the word mare. However if one were to talk about a female elephant there is not a single word that encapsulates that idea so you will be forced to use both the words female and elephant and that will be your lexical concept.

So this is basically how what we are doing is we are converting a pre-verbal message into actual verbal thing by mapping the thoughts onto these words which are these lexical concepts. Now one could ask that can all ideas be very neatly expressed with individual words probably not because a lot of times you will see people telling you that oh I am so happy that I cannot express in words or I cannot describe in words what I am feeling at this point in time. So basically what is happening is this process of mapping the thoughts to these lexical concepts is referred to as a process called lexicalization and lexicalization basically serves as an intermediary between non-linguistic thought processes and the linguistic items or lexical items that can produce verbal expressions to convey these thoughts. So this is precisely what we are trying to do. Now a lot of times it would happen that the number of lexical concepts that you have for a given idea may be more than one.

So for example if you want to talk about let us say a particular emotion let us say you want to talk about being happy so you can be talked you can talk about oh I am very happy I am very elated I am overjoyed I am you know in a blissful state you can basically come up with any number of words or a large number of words to sometimes express a single idea. Now what would happen in that kind of a scenario? In that kind of a scenario what would happen is that you will need to select you will need to perform lexical selection you will need to select the appropriate word for that point in time. Do you think that overjoyed is the best expression to convey what you are feeling at the moment or just happy is the best one or elated is one so you select from any number of

words that you are made available with you can basically say okay this is the word that I want to go with and this at this process the output that comes out is called the lemma. Now a lemma is an abstract mental representation which reflects a slightly intermediate stage between activating an idea and activating the exact speech sound. So it is something which is very similar to a word but it is an abstraction it is not an exact word itself it is an abstraction of that word you can say it is a root idea for a given word which carries with itself both semantic information what is it that I want to convey within the word that I am choosing as well as syntactic information how do I want that thing to you know play out.

So let us say for example if I am talking about let us say a particular adactivity let us say play am I talking about play am I talking about playfulness am I talking about playing what is it that I am talking about if I am talking about eat am I talking about he ate a sandwich day ago he is going to eat the sandwich now he is eating the sandwich now so you can see the root form of this word is the word ate which is the lemma but it can carry information both about the meaning that I want to convey something about the act of eating but it is also carrying the information about the syntax that is what is the situation about eating that I want to convey. So this is what a lemma is so therefore we have said that lemma incorporates information both about the meaning and the syntax of a given idea. Once you have an activated set of lemmas then you move slightly further you activate the process or you initiate the process of activating the sound codes that we sort of need for speech to begin so then what we are doing is then we are sort of finalizing upon that exact type or that exact variation of a given word that we want to go ahead with and we move to what is called morphological encoding. Before I talk to you about morphological encoding I am sure you would know that morphologic morphemes are you know the basic units of language representation basically in any language they are the basic units in which you know the language can be cut down into or reduced to and morphological and morphemes in that sense are basically two kinds of unit which carry meaning. First is lexical morphemes which can stand alone say for example any number of words you can say play, eat, sleep, you know run, cry etc these are all free morphemes because they can stand alone on themselves but if I change them say for example if I do a playing then the ing part is called a grammatical morphemes because per say by itself it does not have a meaning but once it is attached to the morphemes play it modifies the word plays meaning.

So this one ing, er, ess etc are basically called grammatical morphemes. Now in morphological encoding basically what one needs to do is one needs to sort of finalize it okay this is the word that I am going to go with this is that specific form of the word so am I going with eat or am I going with eating or am I going with ate. So this exact form of word is basically specified at the level of morphological encoding. Now levelt

basically says that each for each word that we need to speak we would basically need to specify the morphological form and hence we would go to the step called morphological specification telling us how the word is going to behave when we place it in a larger utterance. So for example as I was saying the morphological specification for the word eat includes that it is a root form and its past tense is ate or its continuous tense is eating.

So we sort of figure out that okay which form of the word eat am I going to use and then we can move on further. Now remember here whichever form of the word that you finalize you will really need to start assembling sounds of that type. So if you finalized ate then the sounds that you start assembling will be different if you finalized eating then the sounds that you going to assemble will be slightly different. So having selected a set of morphemes to produce morphological encoding starts activating the speech sounds we need to plan the auditory movements that will eventually create the speech signal. The speech sound one produces depends on the morphemes that one has activated and also they have to be organized in the right sequence.

Now again when we speak we are not speaking phonemes in a jumbled order while we are speaking phonemes our words are obviously concatenation of phonemes but we do not jumble them up they have to be executed in the same order for example if we decided that we want to speak the word eating we want to talk about the act of eating then what we will do is we have to not only activate the sounds for the word eat and for the part ing we have to also ensure that it comes out in the right order we cannot say ing eat we have to always ensure that whenever we speak it comes out in that same order eat ing. This is basically also very very important so the right sounds need to be activated and they need to be activated in the right order. So just summing up what have we done so far we started with conceptualization we eventually selected lemmas we went to morphological specification then we figured out which morphemes we are going to go with and then we had information about okay which sounds we need to activate and in what order so that we can go further with the process of speech production. Once you have everything with you once you have the morphemes slotted into the right positions once you have activated their individual speech sounds then what you are ending up with is that lexical entry called a lexeme it is basically how a particular word exists in your mental lexicon which is your mental dictionary and for this we need to activate now start activating now the basic sounds that will create this lexeme and we need to be organized in a particular manner for production. So one of the organizational schemes that is preferred by human speakers is called syllabification we do not speak in individual sounds rather we speak in syllables let us look at a example of that.

Now syllabification has two parts it basically requires us activating a metrical structure say for example if I am talking about the word pat it is a single syllable or if I am talking

about the word patter it is two syllables so you say pat and then there is a break and then you say er so basically we need to understand the metrical structure where the stress would be say for example the word banana has the stress on the second syllable na na and the word panama has a stress on the first syllable panama. So basically what we are doing is we are organizing things according to their syllable structure and also at that same point emphasizing which parts of the syllable will be stressed which parts will not be stressed and this information is very very important because it tells us how exactly we have to produce these things say for example banana will basically involve stress on the second syllable and will speak the second syllable slightly louder than we have spoken the first or the third syllables. Similarly in panama we are going to be slightly louder on the first syllable as opposed to the second and the third syllable. Now this is basically while it seems you know as I was saying in the beginning while it seems that speech production is a very easy job it basically requires us paying attention to these very very minute details and you know managing and coordinating between these different levels of representations so that speech production happens seamlessly and it also happens you know rather rapidly. Typically if you see people do not make a lot of speech errors we speak rather fluently we speak not only with the content that we want to speak but we also have these intonations   these you know different emphasis points etc.

so that we can also communicate part of our emotions part of our emphasis non-verbally as well. So in that sense you can see that all of these activities are rather sophisticated and require a lot of attention. Again as I said banana the stress is on the second syllable you can see the stress marker is on the second syllable here in panama the stress is on the first syllable and you can say the stress marker is on the first syllable. Now people would wonder that oh is syllabification an actual process do we really go through syllabification and for that you could look at how we speak you can just sample any amount of speech that somebody has spoken and you could try and see whether that speech segment is organized according to actual word forms morphology or basically according to a stress scheme. For example if you are talking about the word escorting, escorting typically has two morphemes it has a free morphemes escort it has a bound morphemes ing and typically what we should expect is that if somebody speaking escorting they should basically   do   escort   and   ing.

But if you record different number of people speaking the word you will see that when we are speaking or just observe yourself speaking the word escorting you will see that at s I took a pause cot I took a pause and ing. So what we are basically doing is I am speaking escorting which is basically the syllabic structure in which this word is constructed.   So speech production you can see here is not basically following morphemic boundaries but rather it is following syllabic boundaries and this is a important property of our speech production system because this is how or this is how

smoothly the vocal apparatus actually moves around to create these words. So as should be clear with this example we do not simply activate morphemes we activate the phonemes that go with each morpheme and we produce them in sequence. Moreover after these morphemes are activated we calculate the best way to organize the sequence of these phonemes into syllables and it is the syllables that actually form the basis of how we are going to produce them along with stress information and so on.

So yeah while we do not while we need morphemes and words to plan what to speak speech does not simply involve activating the speech sounds in individual words it involves much more detailed speech planning and it basically involves activation morphemes, words, it figures out the stress marking and so on and it in that sense becomes a slightly sophisticated slightly coordinated activity that eventually ends up in production of speech. Now once you have done syllabification what we are ending up with is called a phonological word how are we going to speak something. So for example if you are going to speak escorting we have three parts we have escorienting this is the set of phonological words and this is how we are going to execute this. According to the Weaver plus plus model of William Levelt one can begin to speak as soon as you have activated all the syllables of a given phonological word. Further we can plan each utterance by activating a number of lemmas and morphemes simultaneously it is a fast process you can see and basically the actual speech movements are planned one phonological word at a time.

So escort, escortentous or escorrenting and these are basically you know you can say many executable programs that are executed one after the other in you know specific sequence so that speech production seems very very you know seamless. So another part that we could sort of you know talk a little bit about is that how do we ensure that people are actually executing this process in an ordered manner and this can be seen and this has been demonstrated through a bunch of experiments some of them done by Wheeldon & Levelt where if you ask people to monitor how they are speaking and if you basically ask them to you know let us say press a button if you know particular phoneme comes you will see that they will press faster if the target phoneme comes from the beginning of the word than if it comes in the middle or the end of the word. So let us just summarize how this Weaver plus plus model or how speech production in monolinguals work or generally how speech production works we begin with a set of ideas that the speaker wishes to express which is the conceptualization phase in the next phase we basically tied the overall idea to lexical concepts basically words which are capable of expressing our ideas because then each language will have its own specific words that it will use to convey a particular idea but also may require combination of words say for example we have compound words we have phrases and so on and so forth. Now after we have finalized on the set of lexical concepts after these have been activated we will activate the

lemmas that correspond to these specific lexical concepts activating lemmas would provide us about information about the morphological properties of the words including information about how words can be combined after a set of morphemes has been activated and organized into a sequence the speech sounds that are required can be activated and again placed in the syllable sequence as we discussed. Phonological encoding involves the activation of metrical structures and syllabification so that we know how each syllable is to be produced in what order and with what emphasis.

The outcome of this is a phonological word consisting of syllable size frames. Finally we come closer to articulation so what we are doing here is during phonetic encoding the speech production system basically consults a set of stored representations for specific syllables. So it is basically again as I said many executable programs so how do I say S how do I say Pat how do I say B and basically these programs are concatenated they are sort of brought together and what the system does is it activates these appropriate syllable representations and places them in the correct order so that they can be executed very quickly one after the other as the demands of the speech production system are. This final representation is called a gestural score which basically creates executable commands of how these syllables are going to be articulated. So that is pretty much a generic understanding or a generic description of how speech production really happens you can see that in this model we started at the top with conceptual preparation we had we tied those ideas to lexical concepts we did some selection selected lemmas did some morphological encoding over which form of the word we are going to use then we have fixed out which morphemes we are going to finally select when to phonological encoding deciding syllables creating a phonological word and then we did phonetic encoding phonetic gestural score which is an executing command and then articulation happens which basically ends up in a sound wave which is a physical stimulus that is received by the   ears   of   the   listener   so   that   comprehension   can   take   place.

So that is all that I wanted to talk about a generic and give you a generic understanding of speech production monolinguals in monolinguals and in the next lectures we will look at some of the processing assumptions of this model and how these processing assumptions may need to be sort of tweaked a little bit in order to account for bilingual or multilingual speech production as well. Thank you. .