

Fundamentals of language Acquisition

Prof. BIDISHA SOM

Dept. of HSS

IIT Guwahati

Week 06

Lecture 026

Lec 26: Major theories: Nativist and constructivist theories

Hello and welcome. We will start with Module 6 today. Module 6 will focus on syntax, which is basically looking at sentence structure and its properties, the rules, and how children acquire sentences; within that, we will cover the smaller aspects of it all today. So, so far we have looked at how we started building up from the time the child is able to understand the sounds of all languages, followed by only their own language, and then we moved on to speech segmentation. We also looked at word acquisition, and so on. So, now, of course, the logical next step will be for them to start speaking in sentences.

So, this is where syntax comes into play. This is a broad roadmap that we will follow. We will start with the problems that children have with respect to speaking in sentences or learning sentences, and then discuss how the major theories in the field help us understand those problems and their solutions. So, primarily we will look at nativist and constructivist theories of language learning here.

Lecture 2 will cover syntactic developments, followed by constraints on productivity, morphosyntactic dependency, movement dependency, and so on. So, this is a broad roadmap. However, sometimes it might happen that we are not entirely able to stick to the lecture. Sometimes, lecture 1 might spill over into lecture 2 and so on. So, these are broadly the topics that we will be covering, even if they are not in a very airtight compartment.

So, when does it all start? We have seen how understanding sounds begins in the very beginning, from the few days after birth; infants have been seen to be able to distinguish sounds. So, when does syntax make an appearance? So, around 2 years of age, children begin to produce multi-word utterances. Before, it was single-word utterances; now we are reaching multi-word utterances. So, they start by putting words together when they have

roughly a vocabulary of around 50 words at that time. Of course, there might be individual differences across languages as well, but this is the estimate that we typically have: that around the time they have 50 to 60 words, they start putting one word after another and try to create longer utterances.

So, this is typically considered the first stage of syntax acquisition in production. Remember, we talked about the difference between comprehension and production. So, in all stages of their language acquisition, it is always comprehension that precedes production. Production happens a little later. Similarly, the same structure applies to syntax as well.

So, when they start putting two words together to create some sort of phrase, that is when the production starts; that is when we say that they are able to produce sentences. However, the comprehension, even in this case, predates the production. So, because they are just beginning to put two words together. So, this sentence structure is, of course, very short as opposed to adult speech. So, typically two words, sometimes three words; this will be the length of the sentences that we see.

So, this is the utterance that will consist largely of open class lexical elements. That is how these are the characteristic features. So, open largely based on open class. So, you will have nouns, verbs, and adjectives, but what will be missing are the closed class elements like articles, auxiliaries, prepositions, and conjunctions. So, all those things will be missed.

Generally, they are not there at the very beginning. So, it is only the words. So, the two nouns could be one noun and one verb, and like this. So, inflections are also missing, like the plural marker, or, depending on the language, various kinds of markers that you have in inflectional morphology; those things are also missing. So, this speech is called telegraphic speech, and then eventually they make sentences; from this stage, they go on to creating further, slightly more difficult sentence structures, longer sentence structures, and so on.

But they are still shorter than adult speech, and this is the stage where they are said to be learning syntax. Now, what is syntax? Syntax is the way sentences are formed in any given language. The rules of forming sentences in a given language are roughly what we call syntax. Now, why is this important? Let us go; I am sure all of you have read Lewis Carroll's work. So this is from his work where the poem Jabberwocky is quoted.

You can see this is complete nonsense; none of the words is actually an existing word, but they have been arranged in such a way that it looks like a poem. So, it was brilliant, and

the silvery, slithy toves, it goes like this. So, none of them are words, but they are nicely organized in a poetic format, and then Alice says the poem "somehow fills my head with ideas; only I do not know what exactly they are". So, it feels like you understand this because it has structure. So, that is the very reason; even if it is entirely complete nonsense, it still makes you feel there is a sense to it.

That is where syntax comes into play because it follows the canonical sentence structure of the English language. So, it is organized this way. As a result, you feel there must be a subject, there must be an object, and there must be some action happening somewhere in the sentence. It is just that I am not exactly able to point it out, but that kind of feeling you get is there. However, if you change that structure from the canonical to a non-canonical way, even if the words do not make any sense, we do not know what is a noun, what is a verb, or whatever else is in this structure; still, if you just play with the structure a little bit and create this kind of output, then you will find that the 'I almost understand' feeling will leave you because the sentences are now completely jabberwocky.

So, it does not make any sense. So, this is where syntax comes in. Syntax is the meaningful way of putting one word after another and making sense. So, following syntactic rules is an essential aspect of learning a language. So, you cannot just get away with learning the sounds and the words, morphology and the words, but you also need to learn the rules of the sentences.

So, like all the other factors that we have seen so far, this is also something typically developing children master, and they do so at exactly the time they are supposed to master it. So, there is a kind of universal trajectory through which children go, whether it is learning the sounds, learning the words, or learning the sentences. So, this does not seem to be a problem. There is a period where they make mistakes, as we will shortly see, the kind of errors that they make as they are getting there. So, there is a wobbly start somewhere, but over a period of time, that is also taken care of.

So, they finally manage to speak in a perfect, perfectly grammatical way in a very short span of time. So, the question we are asking now is how that happens. Like all the previous modules we have seen, we are asking the question of how the child does what it does because the entire mechanism of learning to speak in a language into which you are born comes so naturally, and it is such a rapid kind of, you know, progression. The child manages to speak around the age of 1, and by 2, they are already making small sentences, and you know it goes really quite fast. So, how do they do this? This is the focus of this module.

So, let us start with some of the typical features of telegraphic speech; we have already

seen some of them. For example, we have already discussed that telegraphic speech, which is when they speak in two words, typically tends to use whole words only. So, the content words—nouns, verbs, and adjectives—like this, they will omit most of the time; they will omit the function words and inflections. So, this is how it typically starts. So, it looks like, because it resembles a telegram from old times, there used to be something called a telegram.

So, where there were only the content words and no inflection, none of those auxiliaries would be there. So, that is how it got its name, telegraphic speech. So, then we also see some other kinds of errors that seem to be universal. Mastering negation, for example, takes a bit of time. Negation is a kind of structure where languages, of course, differ.

So, you have the simple negative versus the prohibitive negative, and then, depending on the language, the structure will differ, and so on. So, typically in English utterances, we see this kind of errors: no-go movies I no like it. So, in this kind of sentence, I do not we need a 'i do not like it'. So, there has to be this kind of structure, but they just know how to put a 'not' there.

So, I no like it. This is a very common structure. Similarly, complex sentences are difficult and take some time before one arrives at a complex sentence with a relative clause. The dog that chased the cat was barking. Passives are also another problem area, and this is where they are slightly slow in learning to pick it up and use it as well. Of course, there are language-specific problems.

So, an example will be that the cat was chased by the dog. So, this is a passive construction not in the English language especially, it has been seen that children take a little bit of time to master the passive structure. However, these structures, or let us say the mastery of these structures, are language-specific; not all children from all languages exhibit the same kind of problem. So, as I was just saying, the data I was talking about comes primarily from English, but if we look around at other languages, this might not be the case. For example, this is what we have found in the literature: that English, German, and Hebrew, in the case of children from these languages, even 11-year-old children, rarely use the passive voice because it is not a canonical structure; we do not need to use it just because grammar has that possibility, and it does not really mean we need to use it; it is there, it is a structure that is possible.

So, even 11-year-olds were found to be using them very rarely. However, certain

languages make it more prevalent, make it more common, or, let us say, a more high-frequency kind of structure. For example, one language that literature typically gives an example of is that of Sesotho, which uses a passive structure very often. So, children start using them even at 2 years of age. In Italian, 3-year-olds use relative clauses more often than English ones.

So, it is actually a language-specific property; it may not always be true that all children across all languages will find negation, passive constructions, and relative clauses difficult. Another very well-known case of telegraphic speech and a well-known feature of telegraphic speech is that of the case of root infinitives. Root infinitives seem to be a very common feature across languages. This is where the child's first sentences show the use of root infinitives in the sense that the verb is not inflected because, as we just discussed, telegraphic speech does not have the inflections; often, they lack the inflections. So, as a result, the root verb is used as it is, and that is found in many languages of the world, as some examples here I have quoted from the authors.

So, for example, in this case, this is a Danish example by a 2-year-old child. So, what this sentence basically means is she sleeps (infinitive). So, this is the infinitive form of the verb that he uses in this case. So, the infinitive form of the verb 'to sleep'. Similarly, in the French example, this is a child named Daniel, 1 year and 11 months old, who uses the same kind of structure.

So, little baby sleep. So, the infinitive form of the verb 'to sleep' is utilized here. We also have similar examples from Dutch. This is a child 2 years and 6 months of age; again, the same kind of thing, and similarly, you go through all the languages. So, German, then Russian, and we have Dutch. So, many languages seem to have the same kind of structure where children, when they are speaking in telegraphic speech, which consists of 2-word or 2 to 3-word sentences, seem to use root infinitives very often.

Root infinitive means the word will be used in its infinitive form, which is the root form. Similarly, yet another very common feature is that of the null subject, which is basically the omission of the subject in the case of a short sentence by kids. So, this happens even in those languages where it is not common in adult speech. So, is it a problem? Is it something that they do not understand? Why do they drop the subject marker? That is another question that syntacticians have been asking. And this again happens around the same time in children's speech when the root infinitive structure appears.

So, around the same age. So, for example, you have a sentence like "fell in the briefcase." This is from a child named Eve, who is 1 year and 10 months old. And this is from Brown (1973). Similarly, this is a Danish example where they are not using the noun or the subject.

Similarly, first little book read. So, who is reading the little book? No, that is not mentioned. Then it goes under the stool and all of that. not drive tractor. So, who is not driving the tractor? The subject is not there. Similarly, like helicopters clean all of that, they are not even using the subject.

So, this is called the case of null subjects in telegraphic speech; this is also a very common occurrence. However, again like the previous case we saw, this is also not entirely uniform, even though we see many cases; it is not entirely uniform that the differences typically occur, or let us say the frequency of occurrences depends on various factors. They can be individual children, they can be, you know, across languages, sometimes even within the same child across time. So, the same child might sometimes use the proper structure with the subject; sometimes, the same child might omit the subject. So, there are variations that have been reported in the literature in the case of null subjects as well.

Again, the sentence length we have started by saying that the sentences children utter at this stage are very short, like 2 words or 3 words. So, that is also not exactly universal. So, in this case, vocabulary might differ, and as a result, some cases' sentence structures might also differ. So, for example, this is a sentence from an Italian girl who is 2 and a half years old; she produces a rather long sentence which, in English, would be translated as "Teeth get washed after eating, right, Mama? Not before." So, it is a rather long and slightly complex sentence for a 2-and-a-half-year-old girl.

But on the other hand, an English boy, a boy of the same age, is creating a sentence like "Daddy, can you help me find it, please?" So, this kind of sentence is not always that children will be using only 2 or 3 words; they are sometimes also capable of creating sentences that are pretty lengthy. So, the main gist of the matter up to this point is that the syntactic development of children has many complex aspects, and it is not always a linear process. So, it is not like 2 words, then 3 words, and then 4 words; this will be the case everywhere. Of course, there is some amount of universality, but there are also a lot of variations across languages, among children within the same language, and even within the same child over different times. So, similarities and differences between the target grammar and the child's grammar might vary at different points.

Children occasionally use grammatical morphemes in the early two-word stage, but they often omit them. So, this is what we mean by even within the same child, there might be differences. So, if they are not using certain grammatical morphemes, it does not mean that they do not know them because they might be using them in some other constructions. So, that is why I said that this is not a linear development; there are various kinds of, you know, complex layers of things that happen and, of course, differences across languages that are

always there. Then, to understand all of these, there are two primary domains that have been looked at and that have been investigated within this.

So, there are two probable approaches. So, the approaches to these questions differ. And they typically will cover the notions of both acquisition mechanisms as well as the syntactic representation. So, what do the children already have in their innate grammar? What do they already know? How is it represented? How is that syntactic knowledge represented in their mind? That is one domain. The other domain, of course, connected, is the acquisition mechanism. So, depending on these two notions, there are different approaches.

The two main approaches we will be discussing. So, one is called the maturational aspect; the other is called the continuity approach. The maturational approach and the continuity approach are kind of opposed to each other. So, the maturational account looks at the development of syntactic knowledge over a period of time. So, they start because it is focused mostly on the maturational stages. So, as a result, the child starts simple and over a period of time builds up their knowledge and then eventually acquires the target language structure.

So, the child's grammar and the target grammar are not the same in the very beginning. Now, why does it happen? It happens because the child's language development, specifically in syntactic development, in this case will be based on, or let us say dependent on, neuronal maturation as well. So, there are two theoretical positions that adopt this view: one is, of course, the usage-based approach, and the other is the generative structure-building approach. So, the main argument here is that children build the representations of their target language incrementally. So, if that is the case, if they are not born with the entire structure intact and they build up incrementally as they go after structure A, they learn structure B and so on.

So, this then solves the problem of the disparity between early representation and the target structure. Why, in the beginning, do they speak the way they speak, and why, ultimately, over a period of time, do they only start to have the complex structure in place? That is a given; it already happens. So, this maturational account addresses that problem pretty easily. On the other hand, we have something called the continuity approach. The continuity approach does not assume any maturationally induced changes; all the changes that happen, starting at 1 year and 10 months to 2 and a half years and 3 years, are there for us to see, but they do not believe that those changes are due to some kind of maturationally induced qualitative change.

Rather, it is due to some other reasons. They believe in the theoretical position that supports full competence or a strong continuity approach, and they assume that adult-like

syntactic representations are already present even in the early stages, including the stage where they are only speaking in two words. So, the entire structure is already there, even though they are not able to produce it as such. At an underlying level, at the deeper level, the representation is already there. So, why do they speak the way they do, and why do they make those mistakes? As per this approach, the initial deviations from the target language occur due to other factors, not because they need to mature or learn, but because there are other problems. Which are the problems? Problems with morpho-phonological realization, lexical gaps, the interface between morphology and syntax, the discourse conditions for the realization of morphosyntactic properties, and so on.

So, it is not like the structure is not there; the structure is already there, the representation is already there, but in order for them to realize it, they need to take into account all these other problems, and that is why we see that in the beginning they do not approximate the target structure the way they should. So, these are the problems. Now, overall, if we look at the child's language acquisition, there are a few very well-known major problems that, depending on whichever theory you are talking about, all the theories consider these problems to be some of the most important issues to solve. So, we will look at them one by one; the first problem is called the logical problem of language acquisition. It is also called the famous poverty of stimulus, and there are other names given to it; this is a well-known problem: how do children learn to speak perfectly grammatical sentences even though the stimulus is not so rich, not always so rich? So, children receive a very small arbitrary sample of the target language as input because you see no child gets a very systematic kind of an input.

So, no parent, unless they are linguists—I doubt even linguist parents—speaks like that to their children. So, there is no methodical way of speaking to a child when the child is beginning to speak in words and two-word stages or whatever. There is no structured input that they receive; what they get is very, you know, in a very scattered way, in an arbitrary way, and a very small size of sample that they actually receive as input. But they learn to produce, understand, and judge an infinite number of sentences and sentence structures.

So, this is the famous poverty-of-stimulus problem. So, if children are able to arrive at the complex structure of their language in spite of such a small number of inputs, that means they have innate knowledge of the structure, and they are also able to generalize based on that innate knowledge. So, this was the most important point of argument that Chomsky had forwarded in his criticism of verbal behavior. So, this is the same problem we are discussing now: they are not able to learn sentence structure and language as a whole because their learning cannot be solely dependent on the environment, as the input is just not sufficient; that is the problem. Now, if there is innate knowledge already existing there and all they need to do is generalize or apply the same rule to everything, then we have a

different problem. What is the problem? That problem is about generalization, and generalization might actually take at least two different kinds of trajectories.

So, if children need to learn from limited input based on some innate rules, they run the risk of generalization. This generalization can be either restrictive or a matter of overgeneralization. Restrictive generalization is something like the target might contain structures that are not covered by the child's generalizations. For example, in the case of Hindi, let us say, Hindi-speaking children hearing sentences with an overt subject noun phrase may incorrectly generalize that in all cases Hindi will have overt subject marking. This is the case with many other languages as well, but Hindi is closer to home.

So, if the child is hearing sentences in the input where the subject is always overtly marked or the child is producing the generalization thinking that this is the only way of using Hindi sentences. But this may or may not be the case in Hindi; at least it is possible to drop, and you can simply say "khana kha liya." So, of course, "maine khana kha liya" is the canonical structure, but it is equally possible and equally common to say "khana kha liya," the subject is understood in the context of who has eaten. So, this is a case of what is called the restrictive case of generalization. The other side is also possible; the opposite is also possible, which is a case of overgeneralization.

Overgeneralization happens when children have learned to create sentences on their own by learning the pattern but still do not know how to pick the correct verbs or nouns, etc. So, they tend to overgeneralize the rules. So, they have learned the rules, and the same rules they apply everywhere, but there are some cases where they will not apply. There will be exceptions because they have not yet learned the exceptions, so they tend to overgeneralize.

So, I spilled the carpet with juice So, I have painted the box with color. if they have learned that kind of structure. So, they know that this is possible; there will be an agent, and they will be doing something with a structure like this, they will create a sentence like: "I spilled the carpet with juice" This is an example of overgeneralization because it is not possible; English does not allow this kind of sentence structure. Similarly, do not giggle at me; I said her no. So, this kind of structure has some cases where it will be fine if you change the verb in "do not giggle at me," "do not trouble me," and "do not irritate me." However, "do not giggle at me" is not correct, but this is an example of overgeneralization.

So, this is one possibility, and this is one problem that the theories need to take into account. Another problem is what is famously called the developmental problem. Now, this problem is related to the time course of syntactic development. Children tend to go through similar stages of development across the world's languages. Of course, we can see

some amount of inter-participant, inter-individual, and inter-language differences, but there is an overarching structure that seems to hold for almost all children speaking almost all languages.

So, that is another problem that we need to address that the theories need to tackle. Now, why is it a problem? Because languages differ; different kinds of languages have different kinds of structures. So, at the very beginning, at the very root, you can talk about the difference in word order. So, SOV, SVO, and VSO— all kinds of word orders are possible. But even in spite of that, some languages are noun class languages, some are classifier languages; you know there are various kinds of differences that hold across the world's languages, but we still see some sort of a pattern there.

So, that is why it is important to keep in mind the process of how the children master syntax while theorizing. Then we have another well-known problem called the bootstrapping problem. The bootstrapping problem is a kind of circular problem. The syntactic rules operate on syntactic categories. So, in syntactic categories, such as the category of subject, object, or, you know, the verb, these are the syntactic categories.

So, rules apply only to the category of verbs or nouns, but not to the words per se. So, children need to first figure out which is a noun and which is a verb to start with in a sentence. However, this knowledge is contingent upon the understanding of sentence structure itself, unless and until you learn that there is something called a sentence, and a sentence includes a noun and a verb and whatever else depending on the language. Until and unless you know that you will not, the child will not be able to arrive at the syntactic categories. But again, as you can see, without the understanding of the noun and verb, they cannot learn the sentence structure as well. So, this is a circular problem, and this is why Pinker put it this way: it is called a bootstrapping problem because the child must somehow lift himself up by his bootstraps to even get started.

So, this is a very interesting problem in syntax; they have to know the sentence to understand the sentence structure. So, this is a problem that typically all theories will need to talk about; they need to discuss it. And other problems at hand will, let us say, be that verbs do not always denote action. We have seen that when we talk about learning words, learning nouns is comparatively easier than learning verbs, and how learning verbs automatically leads you to the learning of sentences to some extent, if you understand that x does what is done to y. So, there is a verb that needs an object, there is an agent, there is an object on which the action happens, and so on.

But there are cases where verbs do not actually denote action; for example, we might have verbs of thought, verbs of feeling, and so on. So, to know is also to think, to love, and all

these are also verbs. So, they do not always occupy the same position in the sentence either. Especially in languages like Hindi, which has a free word order, we can say the same thing in three or four different ways.

maine khana khaya; khana maine khaya, khaya maine khana; all kinds of possibilities. Of course, there are different kinds of semantic overtones. There is some extra information that you are giving, but at a very basic level, at the structural level, there are possibilities of changing the position of the components in the sentence. So, you cannot always depend on the position of the words in the sentence in order to understand which is a verb. Similarly, and at the same time, verbs also carry a lot of extra information, like tense, aspect, mood, marking, etc.

So, this is a complex problem that we have at hand. Now, these are some of the problems, of course, because we are devoting only one module to syntax here. So, I have put together only a few of the most well-known problems, and we will now look at how theories try to tackle these problems. So, we are now going to discuss theories, syntactic theories. So, we are again asking the same question that we have been asking so far: in all cases, how do children learn language, how do children learn words, how do children learn sounds, and now we are at how do children learn syntax? But here, the primary motivating factor, the primary question, and the primary point of departure is the main question of whether knowledge is innate or not.

Based on that, of course, we have the three kinds of theories again. We have the idea that children learn sentence structure also on the basis of imitation. So, whatever they hear, they repeat; they imitate the adult speech, and thereby they learn. Then, we have the nativist theory of innateness, which is generativism, and then the constructivist theory. Children build up their understanding from the bottom up as they learn, starting with small two-word sentences and then three-word sentences, and as they progress, they learn that these are the three theories.

In the beginning, we also talked about three theories. So, we are revisiting them, and let us see now how things go. So, taking behaviorism as the first one, this theory says that children learn by whatever they hear around them. So, imitation is the most important fundamental point of departure. Now, if this has to be true, we will not be able to understand a novel construction as being grammatical.

We have already talked about the poverty of stimulus. So, we will not be discussing it again, but let us say that often we hear new sentences, new kinds of construction, novel constructions, and very young children are able to judge the grammaticality of novel sentences. Now, how does it happen? Because they have never heard the sentence before.

So, if the input, if imitation is the key, then they would not be able to either produce novel sentences or comprehend novel sentences, but they do. So, this idea then it will not be very tenable.

Now, nor we will be able to generate novel structure as we just say. So, for example, the famous example by Chomsky is "colorless green ideas sleep furiously." sleep green colorless furious ideas. So, this kind of sentence is, as you all know, an example he gave to say that if you ask a small child whether the sentences are grammatical or not, even though the sentence means nothing, it is completely novel because such things do not exist; however, the child will be able to judge the sentences correctly. In the first case, this sentence is grammatically correct even though it has no meaning; however, the second sentence is ungrammatical.

This is how we denote ungrammatical sentences with an asterisk. So, they will be able to say that the second sentence is grammatically incorrect. This is what the idea here is against behaviorism. So, children learn a system of syntactic structures that helps them generate rather than imitate. So, the idea that behaviorism was decided to be not very tenable was already kind of trashed by generative theories.

So, we are just revisiting this because we will not be carrying it forward. So, how do children learn? Children learn because there is some kind of innate mechanism that helps them generate rather than imitate, as imitation does not take you very far. So, they learn, for example, that the cat can be combined with a determiner like a or the. So, the cat can combine this with a verb phrase that is running.

So the cat is running. The verb phrase, for example, "is running." So, now the sentence becomes, "The cat is running." The noun phrase can be replaced with a pronoun. So, it becomes that he is running and so on. So, this is not that the child does not do any imitation here, but that the child is merely following the rule of the game.

So, that and that is how they learn the language. So, that is the idea. So, this, and because of this, behaviorism is not tenable. So, we will not discuss behaviorism anymore. In the next lecture, we will primarily look at nativism followed by constructivism. How each of these theories tackles the major problem that we have already discussed before, and thereby we will try to see what evidence the theories have received and what criticisms each of them has received. And then, of course, we will also look at a comparison between these

two theories and discuss the way ahead and where we are today. So, we will stop here today. Thank you.