

Fundamentals of language Acquisition

Prof. BIDISHA SOM

Dept. of HSS

IIT Guwahati

Week 03

Lecture 015

Lec 15: Speech production

Welcome back; today we will talk about speech production in infants, and this is the last lecture of this module. So, in terms of speech production and stages of vocalization, why do we call them vocalizations? Because, in the initial stages, children do not really speak, so to speak. They do not really have words or sentences. So, it starts with some vocalizations going through various kinds of stages, as you can see. So, I quoted this from a particular study. So, from 0 to 1 month, there is this reflexive, uncontrolled vocalization, such as crying, burping, and coughing.

So, these are the initial stages of initial sounds and initial vocalizations, which do not really have any meaning as such. I mean they indicate something; they have indexical meaning, but they are not really sounds or words as such. They are primarily reflexive and uncontrolled in that sense, aren't they? So, in the beginning, for the first 2 to 3 months, there is the cooing stage, which is mostly made up of vowel sounds, like various kinds of vowels in a continuous stream. That is the first and second stage.

The third stage will last 4 to 6 months, during which there will be squealing, growling, and various kinds of exploration of new sounds. So, this is where consonants start to appear, and this is where they will begin to explore new sounds and try to produce them. And then there will be the 7 to 9 months when the canonical babbling stage starts, using the same syllables repeatedly, like "ba ba ba." So, this is the babbling stage. So, like this, there is now a stage where syllables begin to appear.

And then, after 10 to 12 months of variegated babbling, there are sounds that resemble adult speech with varying tones, and so on and so forth, something like this. So, varying sounds and syllables will be put together, and they will almost sound like words, almost

like meaningful utterances. So, these are the stages. While they are trying to speak, imitate, or be speakers of the language in the community, there are various kinds of stages. So, in the beginning stages, there are just some meaningless sounds, but when some semblance of words appears, these are things that are very common.

To quote Barbara Lust, sounds are sometimes omitted. So, this broke will be pronounced as "bok." Sounds are substituted for each other; so, a rabbit will be called a wabbit. Sounds within a word get assimilated; so, doggy becomes goggy. This part comes here: /da/ gets replaced with /ga/, and so on.

Then, repeating syllables within words, "tummy" will be said as "tum tum," "bottle" as "baba," and so on. So, these are some signs of initial speech production in children that are more or less common across various languages around the world. So, why does it happen that children do not start speaking, you know, just like that? The reasons are numerous. So, the primary question here, however, is the reason we will get to in a minute, but we have already seen that children can perceive sound very well in the beginning. Starting from a few days old to a few months old, they can eventually figure out very fine-tuned differences.

So, sounds distinguish between familiar and unfamiliar languages, including a mother's language, another language, and various kinds of differentiation of sounds or voicing aspirations, and so on. So, they have a very sophisticated speech perception ability. Now, this is not matched by their ability to produce the same sounds or different ones. So, there is a mismatch that occurs between speech perception and speech production. That is one of the most interesting things about speech production in children.

Now, the question is, why does it happen? Children, of course, want to speak; they want to speak because that is what we see from the very beginning. They start vocalizing and trying to say something, but it does not really happen. One of the reasons that has been put forward is the primary reason, which is articulatory constraint. What does "articulatory constraint" mean? That reminds us that we talked about the specific kind of vocal tract that humans have, the specific kind of larynx that we have, and the muscles that help us. So, in very small children and infants, even though they understand things as their brains understand them, the articulatory apparatus is not yet fully equipped to produce the same sounds that they perceive; that is the difference.

So, the vocal tract is the wrong shape in newborn babies at the beginning. So, newborn babies have a shorter vocal tract than adults, and the larynx is also placed higher, as in other higher primates. Our larynx starts to descend only sometime in the first few months of a child's life, around 3 to 4 months of age. So, that is when the vocal tract appears

similar to that of an adult. So, that is one problem.

The second problem is that the tongue is larger compared to the child's head and other structures; the tongue is much larger than necessary. So, as a result, the tongue muscles are not well adapted; they are not fully developed or well adapted for the subtle movements that are necessary for creating speech sounds. So, as a result, vocalization happens, but speech sounds are not produced. So, speech sounds need more fine-tuning for adept handling than do vocalizations. So, that is why we see vocalization in the very beginning, but eventually, it gets sharpened into speech sounds.

So, as the larynx gets lowered, the tongue muscles get, you know, fine-tuned, and so on. So, the child will slowly overcome the articulatory constraints. So, physiological changes throughout the first year of life finally make the child start to speak. Also, they go through pre-linguistic vocal development during this time. The stages are largely similar across the languages.

So, prevocalic changes, as well as physiological changes, will take care of the child's abilities. So, the development of speech sounds is also dependent on the universal sequence of maturation. So, on the one hand, you have the articulatory constraint; on the other hand, there is something called the universal constraint, which was proposed by Jakobson way back in 1941 and followed up on in 1968. So, the idea here is that there is a kind of constraint that is universal across the world's languages, depending on the stages of maturation. So, there is a universal inventory of distinctive features that is said to underlie all possible phonemes of all languages, a set of distinctive features that is the basis of all phonological differences, and there we have a sequence that is universal.

So, according to Jakobson, these features are organized in an implicational hierarchy. This hierarchy arranges features in a stratified manner, starting with the most common ones and moving to the rarer ones. So, the world's languages will have some set of contrasts, let us say phonological contrasts, that are most commonly found in all languages. For example, let us say the /ba/-/ma/ contrast; the difference between the sounds /ba/ and /ma/ is more common compared to some other contrasts that are specific to certain languages. So, if that is the case, the most common features are at the highest level, and then, hierarchically, the lower ones will be those that are less common.

So, this is how he puts it: the hierarchy looks somewhat like this. So, the most common sounds are called unmarked sounds, and these are the sounds that are learned first. And on the lower end of the spectrum, you have the less common sounds, which he calls marked sounds; those marked sounds will be learned later. Technically, the unmarked sounds are universal, and the marked ones are language-specific sounds. So, some sounds

are, for example, the click sounds that are present in some South African languages, if you remember them.

So, now that could be a contrast that is not very common. So, it could be a marked contrast as opposed to the /pa/-/ba/ type of contrast. So, another important feature of this theory is that universal constraint theory states that infants do not really acquire speech sounds; they acquire only sound. So, they acquire a set of contrasting sounds rather than contrasting speech sounds. So, they need to have a set to learn from.

To begin with, they need a contrasting set, such as the /pa/-/ba/ contrast; for example, they need that contrast. So, they are able to contrast the two sounds, whether they are part of the language or not; that will come later. In the beginning, the stage that we are looking at, infants simply distinguish between two sounds based on their properties: if they are universal, they will be learned first; if they are language-specific, they will be learned later; that is the gist. So, the contrasts that are common will be learned first, and the others will be learned later. Now, there is also a genetic predisposition that is part of this theory.

The idea here is that the developmental stages for infants are genetically programmed and universal. So, the stages of development we have seen with Piaget and others show that the developmental stages are genetically coded for us to develop in a particular pattern. So, this is a result of which it is common to see infants learning or acquiring sounds in the same sequence because we are genetically predisposed to have particular stages of development in terms of mental and language development. So, it is understandable that acquiring sounds will also follow a sequence. For example, bilabial stops are a set found in early speech production across English, Japanese, and Swedish, according to a study carried out in 1991.

Bilabial stops like /pa/ and /ba/ contrast, as I mentioned. Not only that, but the errors are universal too according to this theory. So, this theory predicts a similar kind of error across languages. And this is also the case, for example, in Serbian, Swedish, German, French, and English, where children typically replace /ka/ and /ga/ with /ta/. So, these velar sounds are more often than not replaced by dental sounds in many languages.

So, this is probably easier to pronounce than the velar sounds. Similarly, there is an example given by Lust (2006): the sound "kaka" in Serbian becomes "tata." In the case of German, the same thing is also seen. So, because these kinds of sounds are probably difficult, So, the errors occur across languages in the same patterns. That is what the theory predicted, and this is what has been supported by some findings.

However, we also have some contradictory findings: some languages do not seem to follow the same kind of maturational stages that this theory proposes. In some cases, certain languages follow specific language patterns. For example, labials are less common in Japanese and Swedish than in English or French. Bilabial stops are less common in Japanese and Swedish, but they are more common in English and French.

So, similarly, it sounds like this. And these are common among young Quiché Maya speakers in Guatemala, though English learners acquire these quite late. Now, obviously, because this contrast or these kinds of sounds is certainly not universal, these children do learn it in the beginning. English learners do not learn it until quite late. So, this is not, hence it does not hold universally; that is the idea here. So, similarly, Bulgarian, Estonian, and Swedish children learn to produce the /va/ sound earlier than English children.

So, what these findings primarily point to is that there is probably not a universal pattern that holds for all languages in terms of how children learn to articulate certain kinds of sounds. Universal theory proposes that there will be a universal pattern, but there are contradictions. There are various hypotheses and theories about why these kinds of things happen. These theories try to take into account the various kinds of differences and patterns that we see in infant speech production. One of the more well-known ones is called the mushy mouth, mushy ear hypothesis by Lust.

Again, the idea is that the mispronunciation occurs because of a misperception. Mushy ear as in they do not hear it properly. So, if they do not hear it properly, the pronunciation is also affected. So, if you hear a mispronunciation, you will pronounce it in the wrong way. So, the idea here is that infants often pronounce words incorrectly because they do not perceive them accurately.

Now, this is contradictory because we have seen that children actually perceive sounds with a great deal of sophistication, even at a very early age. The point that this theory discusses is that identifying sounds in isolation is one thing, while identifying them in continuous speech and producing them is a different matter. Children, typically developing infants, hear those sounds not in isolation but embedded in continuous speech. So, as a result, they may not actually perceive it differently.

This is an example of such a study. They created minimal pairs using nonwords and nonsense words. So, this /Bih/-/Dih/ minimal pair. The infants were younger than 17 months old. This is how they did it. So, there was a nonsensical word paired with a picture that was supposed to represent its meaning.

So, /bih/, this looks like a crown, but for now, just ignore that it is a crown. This is, these were put together. So, when they heard the sound, the word "bih," they saw it. When they heard /dih/, they saw this. So, this is, of course, because the child does not know either the crown or the chemical composition.

So, for them, this is something new. So, this is how new words, nonsense words, and new objects are paired. In the familiarization phase, they hear the sounds and see the corresponding pictures. In the testing phase, the combinations were reversed. So, when they saw this in the familiarization phase, it would differ in the test phase. So, in the test phase, this will be shown in relation to this sound, and vice versa.

So, this is in the testing phase. But in the familiarization phase, they will see it as it is. What the dependent variable here was is a longer time spent looking at the pictures. So, the idea in the test phase was the combination, because it was reversed to check if they could identify the change. Now what they found out was that at 14 months, infants could distinguish between sounds that were totally different, like /lif/ and /nim/, but they could not distinguish between /bih/ and /dih/. There are things that are completely different, meaning they are not minimal pairs.

These two are minimal pairs because there is only one difference. First, sound changes; that is why it is a minimal pair. So children could distinguish between the non-minimal pairs, but they could not distinguish between the minimal pairs. After 17 months of age, they could also differentiate between the minimal pairs. So this suggests that misperception could be one reason for mispronunciation because, until then, they have not been able to perceive it correctly.

There is another very well-known study on the fish. This is called the FISS phenomenon. This is a conversation between an adult and a child. Child says, "There are fish in the pail." So an adult says, "Is there a fish in the pail?" An adult is imitating the child's pronunciation of the word "fish." The child is pronouncing "fish" as "fiss," and the father or mother, whoever is there, is also pronouncing it the way the child is pronouncing it.

So, the child does not agree, "No, not fis, FIS," she stresses it but pronounces it the same way. Similarly, again, an adult said, 'Is this your fish?' No, my fish. She continued to reject the adult's pronunciation until he said the following: This is your fish. Yes, my fish. So, what is happening here is that mispronunciation does not lead her to misperceive it.

She is pronouncing it as "fis," but she refuses to acknowledge the adult's pronunciation as a proper one. The adults should pronounce it as "fish." So, only when the adult pronounces it as "fish" does the child agree, but then he or she cannot agree. So, the gap

between the understanding and pronunciation that this theory proposes does not hold here because the child perfectly understands that it should be this; this is how it is perceived as "fish," but the child is not able to pronounce it correctly. There are many other examples of such cases; for example, the word "jump.

" The child is not able to say "jump," but when the father says "jump," the child only says, "Papa can say mama, papa; I forgot I can say 'dump.'" So, the child is not able to say "jump" and says "dump" instead, claiming that only parents can say it. So, they understand that the perception is in place, but the pronunciation is not. So, there is a mismatch. So, examples like this show that they perceive the difference but are not yet able to produce the sounds correctly.

So, it could be a matter of articulatory constraints, not exactly a constraint of perception, but a constraint of articulation. But then we have this: is it only an articulation problem? To counter this, we can discuss the following examples. Now, this is a very interesting and rather old study, but it is interesting nonetheless. So, in this case, the child deforms the word "puddle" but produces "puzzle" instead of "puddle." So, when they have to, this is the target word: the child produces "puddle"; the child says "puddle.

" But when the target word is "puzzle," the child says "puddle." So, it is not that the child is unable to produce the puddle sound. They can, but somehow it does not happen when it is supposed to, although they produce the same sound in the case of a puzzle. So, the idea here is that the articulatory constraint is probably not the issue. What is the problem, then? There are many theories that have been proposed about the problem, but due to a lack of time, we will not go into detail. Now, there are other kinds of things as well; there is something called regression.

Regression talks about the child starting with the proper pronunciation, then regressing into, you know, not pronouncing it properly; there is a regression in their capacity, and then they eventually get back. So, this is one particular study that is well known as Leopold's "Hildegard." So, they started at 10 to 16 months; she could pronounce the word "pretty." At 18 months, it became biddy, and then again at 18 months plus, it went back to pretty. So, this kind of thing also happens; there are all kinds of possibilities regarding pronunciation.

So, there are many theories; another theory, the mushy mouth, mushy ear theory, does not always hold, and now we come to another theory called the template theory. This theory tries to explain infant speech production in terms of tasks. So, there is a template that the child develops, and then whatever they hear, they try to figure it out within the template, and whenever the template does not match or they try to extend the template,

that is when we see the errors. This is what the template theory primarily states. So, the sound-meaning mapping needs articulatory finesse that is not available to the child.

So, what does the child do? This goes against the earlier theory where we say that the word and the meaning are mapped together, and that is what the child figures out. Here the theory says that the word meaning mapping and the sound and meaning mapping probably are not available. So, what they do is resort to creating a Word template. Word templates are well-practiced word patterns that children have already learned. In the beginning stage of their lives, they have learned a few words, practiced them well enough, and heard them an adequate number of times; as a result, a pattern has been created based on sounds.

While learning new words, this template is extended, and that is when the errors occur. So, for example, if the trochaic stress pattern is the template that the child has achieved, For example, take the case of English children: if they have heard a lot of words with a trochaic pattern, then they will take that pattern as the template for words. So, all words should follow this particular template. Now, if they have a new word that does not follow that pattern, they will make an error in their production, and that is one of the possibilities. Similarly, if a vowel followed by the /sh/ fricative pattern is learned, it can be used to produce multiple words like fish, dish, fetch, and vest.

So, the child probably pronounces them as "ves," "fesh," and so on. So, this was, this is, this comes from a study carried out by, I think, Berman, in 1977. So, they have talked about this particular child who seems to have figured out that a vowel followed by the /ʃa/ sound is a template. So, whenever they see anything that is similar, they pronounce it that way. So, fetch becomes fesh, and vice versa. Evidence for this has come from various languages, such as Hebrew, Spanish, French, English, Estonian, Swedish, and so on, where we see that they try to fit the same template for various new words that they learn, which may lead to certain types of errors.

The initial speech production we have looked at so far in this segment has focused only on the very first stages of speech production. So, in speech production, there are two stages: the first stage involves producing the sounds, immediately followed by the production of words. So, these are the studies that we just looked at, which are the beginning of word production, but we will discuss more of them when we talk about word learning among infants in a different segment. So, this is where we have talked about a few theories regarding how they probably segment. In the previous lecture, we talked about how infants probably segment a continuous stream of speech into words through various kinds of theories, and now we are looking at how speech production happens and what probably helps them.

But this is a very rudimentary, very basic idea that we start with, and we will take it forward; we will carry this on in the case of word learning in another module. So, this is where we complete Module 3. Thank you very much. I will add all the references at the end, and you can look up all the more recent findings that either follow or do not fit the previous theories so that you have an updated understanding of how these theories hold today. Thank you.