

## **Tools and Technologies of Language Documentation**

**Prof. Bornini Lahiri and Prof. Dripta Piplai (Mondal)**

**Department of Humanities and Social Sciences**

**IIT Kharagpur**

**Week-07**

**Lecture-33**

### Lecture 33 : Writing Grammar

Welcome to the 33rd lecture of the course, Tools and Technologies of Language Documentation. Today, I will talk about Writing Grammar. First, I will begin with what is grammar and then different types of grammar. There are different types of it, though I will not mention all of them, I will mention the two basic ones. And then, I will talk about those topics which are needed to be included in a sketch grammar. So, when we are talking about grammar it can be quite descriptive and at the same time, it can also be a smaller and just a sketch of a language.

So, we can begin writing a grammar with just a sketch and then as we get more data or we can devote more time into it, we can always make it a overall descriptive grammar. So, when we talk about grammar, grammar is the systematic description of the features of a language. So those, you can say, are the rules which are there in a language. So, syntactic rules, phonemic rules, all the rules and regulations of a language are mentioned in a grammar.

So, grammar rules of a language governing the sounds, words, sentences and other elements as well as their combination and interpretation. So, overall aspect of language is grammar. So, when we say grammar, we say that every language has a grammar. It does not need to be a written language to have a grammar, because when people communicate they always follow certain rules which are part of the language. So, without that it is not possible.

So, depending on the grammarians approach a grammar can be of two types prescriptive or descriptive. So, the grammar books that you might have read or studied in your schools are more prescriptive in nature. That means, they tell that whether a particular

sentence is correct or not or what are the rules which you should always follow while using the language. And mostly when we look at the prescriptive grammar, those are the grammars which are mostly for the written ways. Of course, in some of the grammars, the speaking style is also included or the rules governed in speaking styles are also included, but it is different from the descriptive grammar.

Descriptive grammar is more about how actually language is used in the discourse in day-to-day life. So, it is different from the prescriptive grammar. In prescriptive, it is more like you are learning the rules, it is said that according to the standard of the language this is the norm, this is the prescribed way of speaking or writing a particular language. But in descriptive one, there is no prescription, there is no saying that this is the correct way of saying it. Whatever ways are adapted by the speech community to interact to communicate, all those aspects are actually included in the descriptive grammar.

And there is no notion of so called that this is correct, this is not. So, everything which is included or which is practiced by the speech community can be included in the descriptive grammar. So now, when we are talking about writing grammar and specially so for the endangered languages or smaller or minor languages, then what we want to do is that we aim to write a descriptive grammar of these languages. We do not prescribe that this is right and this is wrong, we just describe what is happening in the speech community, how do they communicate, how do they interact. So, everything which is part of a discourse, everything which is a part of communication can be included in the descriptive grammar.

The way people talk, the way people write, if they do not write their language, they only speak, then the way they speak, the structure of the language because when there is no script, people only speak, then also there is certain rules or forms those are always followed by the community. So, those things are included in the descriptive grammar. Now, if you are planning to write a sketch grammar, what can we include in that. So, these are the important points which we generally include in a grammar. So, I know that just by listening to this lecture, you might not be able to write a grammar, And this lecture is not aimed for that, but this is just to give you a flavor or to know that what are the things which are included in a sketch grammar.

So that, if you want to actually pursue this, you can go deeper into these topics. You can work on them and then plan accordingly. So, this is to just let you know what are the contents, which are generally expected from a sketch grammar. So, introduction, phonology. So, there will be a chapter on introduction, phonology, morphology, syntax, pragmatics and social linguistic aspects and sample text.

Sample text is also important. So, sample texts are those texts which we collect. So, I will talk in details about each of these units. So, in introduction, we include all the information related to the speech community, that is the geography where they stay. So, if they stay in various different places, you can also plot those or you can only focus on that particular place from where you have collected data and you are writing grammar for that particular place.

So, if you are working on a one particular variety, then you can focus on that particular location. You can draw a map, you can show the exact location, you can show the location of the village, the district, also demography, how many people actually use the language. Vitality, so by now you know, how to develop vitality questionnaire, how to know about a vitality of a language. So, those information are also included in the introduction part. And this is so when we are talking about the smaller or endangered languages.

So, these are some of the important aspects which we include. Then comes the variation. So, as I mentioned a language can be spoken in 3, 4 different district or 3, 4 different states and in those cases what do we find? We find lots of variations. So, you should mention which variation you are focusing on or if you can or if it is possible you can include all the variation, that is even good. So, if you can do that.

Then talk about all the variation, you can also give a brief note about the differences lying between these variations. How, for example, a Jharkhandi Mundari is different from say Odisha Mundari. So, giving a little note about that or how West Bengal Mundari is different from these two. So, if possible those examples can be given, the same sentence can be written in all these three varieties to show the difference. So, all those can be marked or noted in the variations of a part.

Then about the script; if the language has a script, then you can talk about that and if it does not have and it borrows script from some other language or uses the script of some major language of the area, you can talk about that you can mention about that. About bilingualism, that is also important, how much is the bilingualism present. So, are the community members fluent in some other languages, are they bilinguals, multilinguals, you can also mention those. Because when we are talking about bilingualism or multilingualism, we know that one language affects the other one. So, a bilingual speaker can actually shift or borrow words, so those features can be included, and for that, it is important to know the profile of the speakers or community members.

And if there are certain important works which needs to be mentioned, if there is already a dictionary present or if there are certain materials which are already present, which are

actually important, then you can mention those as well. So, this is a basic information about the language about the speech community and that goes in the introduction part. So, you talk little bit about the geography, you talk little bit about all these things. You can also talk a little bit about the language, in the sense that which family it belongs to or little bit about the neighboring languages or sister languages, those things also you can include in the introduction part. Then comes your main grammatical analysis.

So, it generally begins with the phonology. So, in phonology, you know that we divide sounds as vowels and consonants and that same pattern, we follow here while describing the speech sounds of a particular language. We talked about different types of vowels those are present in the language diphthongs. We talk about how many vowels are there, vowel inventory and overall phonemic inventory. We talk about the consonants and also various rules governing them.

So by now, you know that in certain distributions, consonants or vowel sounds can change or they might not occur in certain positions. So, all those things, you can mention in the phoneme in the phonology part. And not only about vowels and consonants, we also mention about the supra-segmental features like tones, intonations In many Tibeto-Burman languages, we do find tones. So, depending on the tone, meaning of a word changes.

So, we can mention those. We need to mention those when we are dealing with a tonal language, prosodic phenomena, we need to talk about syllable and syllable structures, we talk about. Consonant clusters, so how many consonants can actually come together without a vowel. In some languages, there can be two consonants, consonant clusters or there can be even more like 3, 4. So, how a syllable is formed in a language or what are the consonant structure pattern in the language. So, of course, you need to go details into phonology to work on all these aspects.

But these are the things which are included. One must try to identify the longest and the shortest possible syllable in a language. So, that also one can identify and give example. Various rules related to distribution and environment of the phonemes can be given. So, all those things, everything, as many information as you can give about the sound system of the language, all those things are included in the phonological part.

In some endangered languages, we do see that the tone is slowly getting endangered or in some languages, we have heard that earlier there was tone, but now we do not find any tone in the language. So, in those cases those things can also be mentioned or maybe in few words, you can find tone, while in other words there is no tone. So, the tone is getting slowly endangered. So, one can mention all those aspects in the phonology.

Then comes morphology. So, morphology is all about words you know. So, various word formation process can be included in morphology, there can be different types of word formation, like adding of suffix, adding of prefix; those can give another meaning of a word or there can be combination of two words and then we get another word. So, different types of word formation process should be included in morphology. There can be compounding and reduplication. So, in certain languages, especially in Indian languages and South Asian languages, we do find reduplication: two words are repeated to get a meaning.

I can give you an example from Hindi, when I say "main parte parte so gaya", then "parte parte", I am repeating to say "so gaya". So, here "parte" in isolation is a verb, but when I repeat it and say "parte parte", that becomes reduplication and it becomes an adverb, which is talking about "sona" which is sleep. Here, I mean to say that I slept while studying. So, in English I cannot reduplicate, but in many Indian language, you can think about this sentence how will you say in your language, "I slept while studying"? In Hindi, I can say it like "main parte parte so gaya". So, "parte parte" is the reduplicated form.

So, these type of structures which are not found in English or other languages, but are there in the language you are studying, you can always talk about those. So, in morphology, we talk about person, number and gender. So, PNG stands here for Person, Number and Gender. We can talk about case, class, classifier, definiteness, adpositions, pronouns, particles, adverb, adjective and word formation process. So, all of these can be talked about in noun morphology and in verb morphology, we need to mention about Tense Aspect Mood.

So, TAM stands for tense aspect mood, voice, nominalization, particles, types of compound verbs, negation, verb classes, all these goes under verb morphology. So, when we are talking about morphology, we classify morphology into two groups, noun morphology and verb morphology. So, all the process those are happening in the noun part comes with the noun morphology. So, case can be marked with the noun, definiteness, at position. So, all of these can be marked with the nouns.

So, this comes under noun morphology. And then as you know tense, aspect and mood, voice, different types of verbs, all this comes under verb morphology as they deal with verb. So, morphology is all about the words and what are the information these words are carrying. So, verb we know there is a verb. How does verb carry different types of information like tense, aspect, mood or person, gender, how is it carrying it if it is carrying. In some Indian languages, we also see honorific being marked in the verb.

So, how is it being marked all this discussion goes under morphology. Then we come to syntax. In syntax, what we talk about is about different types of structures of a sentence. How are sentence structured? So, we can talk about different types of sentences like simple sentence, complex or compound sentences. So, we can talk about all those, we can talk about word order.

We know that in different languages, words orders are different. So, there can be subject-object-verb, which we call SOV or it can be SVO like subject-verb-object. So, there can be different types of word order. So, what is the word order of that particular language that you are describing? So, when you are talking about for example, if I give you example of Hindi, you know it is subject-object-verb. "Ram am khata hai"; it is the most common example that we hear.

So, we see that the sentence is ending with "hai" which is the auxiliary or helping verb. So, it is SOV, but in English, we do not find the same structure. In English, I will say "Ram is eating mango", that means, SVO. So, subject-verb-object is the pattern. You will have to know what is the word order of the sentence and when we say word order, we look at the simple sentences.

We do not look at other types like question type of sentences or any other sentences, we only look at the simple basic sentence, declarative sentences those are taken for the word order. And then again, in some languages, we can move the words here and there and still the sentence remains grammatical. Like in Hindi we can do that, we can say "Ram aam khata hai". Like for example, "Sita achchi ladki hai", "Achchi larki hai Sita", I can say both.

So, I am actually moving the words. So, you can see that in "Ram aam khata hai", I cannot do it. I can do it. So, these type of descriptions you need to give, but if you compare the same sentence in English, you can see that you cannot do that right. So, you have to see whether the word orders are flexible or not. In most of the Indian languages, what we find is that word orders are flexible they are in free word order.

So, it is free that means, they are not rigid, they can be here and there depending on our focus of the sentence. So, depending on the where we want to focus, that is how we change the words. So, position of the words, that is very very important aspect of a language. In some languages, what you see is that in some type of static sentences, the auxiliary or the helping verb can be dropped. Like for example, in Bangla, if I say like "My name is x", I do not need to say "Mera naam x hai", "hai" I do not need to say, I can just say "Amar naam x".

So, you can see that it is a type of verbless sentence. There is no auxiliary or helping verb required to complete the sentence. So, that can also be a thing. So, those type of combinations of sentences where it can be done, where it cannot be done, how it is done in some languages and not, so depends on the data of your language. The language in which you are working, you need to explore those and know what is the correct form, how is it mostly used, what are the different clauses, how are the clauses combined and then there are different types of sentences. So, all of those comes under sentence types which you talk about, how are different clauses joined or sentences joined together to create complex sentences, complex predicates, we discuss all those in syntactic types.

We also talk about coordination and subordination, how we find in some sentences there is a coordinative clause and in other there is a subordinate clause, how they are combined. So, there are complements. So, there are lots of things and I am sure if you want to work on it, you will have to go into details. So, I am just giving you a brief outline. So, based on that, syntactic description of a language is given.

So, negation is a very important aspect of syntax. Agreement is very important. Agreement you see, like in Hindi, you say "Ladka jaata hai", but in "ladki", you will say "Ladki jaati hai". So, "boy goes", "girl goes", verb does not change in English. But in Hindi, verb changes depending on the masculinity or femininity of the noun. So, there is an agreement between the noun and the verb depending on the gender of the noun.

So now, whether the language has agreement or does not have an agreement. In some cases, we see there are agreement based on gender, number, person. Sometimes also, honorificity in Hindi also, we find depending on the honorificity of the noun, there is an agreement with the verb.

So, if I say "Ap jaenge". So, "Ap jaenge", "Tum jaoge". So, the verb pattern is changing for "go". So, both the sentence if translated in English will mean "you will go". So, but when I am saying with the plus honorific thing, when the pronoun becomes honorific, the pronoun "you" is becoming honorific and it is becoming "ap". So, I will say "Ap jayenge". So, the verb form is "jayenge", but when I am saying non honorific "you", I will say "Tum jaoge".

So, the "jaoge" and "jayenge", there are two forms. And these forms are chosen depending on the pronoun that is used. So, we see that there is an agreement, this is the feature which we call agreement, when the verb has to agree with the noun. So, noun and verb, both have to agree; they both need to be in the same platform. So, if there is second person singular, then there is one form, but if there is second person singular honorific,

then the verb is changing depending on that. So again, if this second person becomes plural, then again the verb form will change.

So, depending on that, how verb forms are changing, that is agreement. Depending on various features, agreement can be determined. So, it can be based on pronominal, gender, number, honorificity. So, different things influence agreement pattern of the verb.

So, all of these again, comes on the syntactic structure. And now what we see is that, due to syntactic rules there can be various changes in the morphology of a word which we call morphosyntactic aspects. So, morphology is influencing the syntax of this language; due to syntax, there are some morphological changes in a word. So, these morphosyntactic aspects are also included in syntactic aspect. How they are because in a language, it is very difficult to cut things like we cannot say, here morphology ends and here syntax starts. When we are talking we are using all of these, we are using phonology, morphology, syntax, everything and also, the pragmatic features.

So, when we are talking, we need to describe all of these in our grammar and for the convenience of writing and also for the convenience of readers to read, we divide them into different sections. So now, coming to the pragmatics and sociolinguistic aspects, what we do is that we look about the different aspects which are beyond syntax. So now, I gave you the example of honorific "you" and non-honorific "you". In pragmatics section, you will discuss where honorific "you" is used, where non-honorific "you" is used. So, both the aspects, both the sentences are correct when I say "Ap jaenge" and "Tum jaoge" both the sentences are grammatical, but then one needs to know where to use "ap" and where to use "tum".

So, that is important. So, that comes under pragmatics, where to use what. So, these domain analyses, who speaks what with whom in which context, that is important to know. Sometimes, in some areas, Hindi speaking people also use honorific terms with the children. So, while addressing a child, they might say "ap", what will you eat. So, they are using the honorific word rather than using the non honorific term and if in those contexts they use, that might be taken as offensive.

So, these are the things which are domain analysis in which domain what can be used, ways to greet, how do people greet each other, what are the tabooed words, when are they used, in which context they will not be considered as tabooed. So, there are all those terms and also various sociolinguistic aspects. So, pragmatics and sociolinguistic aspects are also very closely connected. There are various kinship terms and address terms. So, kinship terms you know like various types of relation terms mother, father all those things, but address terms are those ways we address the person.



Like in Hindi, I can say I will say my mother-in-law "Saas". So, "Saas" is mother-in-law in Hindi, but when I address her, I will not call her like "saas, come here", I will say mother, something like that. So, there are various address terms and relation or kinship terms and they are not same, they are, at times, different. The way you call someone is different than what is the relation called. So, there are various differences like that.

So, you can talk about those kinship terms and address terms. Politeness strategies, what are the strategies used to be polite in the language. And then of course, you can also talk about language shift and retention, language contact and borrowing. In some languages, we do see that the languages follow a different rule, a different pattern for the borrowed words. Like in some languages, there can be difference depending on animacy of a particular noun, for the borrowed words, animacy is not marked in the verb or in the agreement.

So, there can be different types of features based on the borrowed words. So, borrowed words can work differently in the language in the grammar of the language. So, you should also incorporate those in the language and specially so, because we are talking about smaller or endangered languages where we see lots of borrowing. So, that is also important to mark how borrowing is actually affecting the language. So, sample text is important because sample text gives you the idea about the real language. So, if someone is reading your grammar, he or she will know about the exact language.

So, it is always good to give narrations because those there one can find words, one can find sentences everything is covered within narrations. Dialogues can be given, folktales can be given, if there are some archaic use, if you are giving a sample of a folktale which has archaic use, the pattern in which the folktale is told is not the pattern people talk in. So, you can also mention that that this is a different pattern which is not used in day to day discourse. Of course, when you are providing these narrations or the sample text, you should provide transcription, glossing and translation, so that the readers can understand. It is ideal to have a couple of narrations, both from male and female consultant; that is important.

So, nowadays what do we find? We find that there is a hard copy of grammar. If there is a hard copy and then we also find a pen drive or some link from where we can also hear the data not only about reading or looking at the narrations, but also, we can click on the sounds and know the sounds of the language. So, contextual knows about the narration, social or cultural significance of the narration should also be provided. So, when was this text said or what are the background about it or the context of it. All those should be incorporated in the text.

Grammars are very important output of language documentation. So, grammars can be very useful in language maintenance or communication. Language revitalization and at the same time, grammar can help in building language technologies as well and grammars can also help in creating various types of pedagogical materials. So, grammar is a very important aspect of a language written grammar. However, it takes proper expertise time and data to write a grammar even a sketch grammar.

So, if you are thinking about writing a sketch grammar, you should give proper time. and you should know about it. So, it is an work of expertise. So, you need expertise time and data. Grammar of many smaller languages are sometimes produced just by copying the features of the major or dominant languages. So, maybe smaller languages of West Bengal; in some cases, we have seen, have produced grammars of their language, but they have just copied the grammar of Bangla, which is not actually an accurate way of writing grammar.

You cannot copy grammar of a language and just put in the features or the example sentences of your language, because that spreads misconception about the language and every language has its own grammar. So, it is not a good idea to copy grammatical features of one language and paste it in some other language. So, it is always better to look into the data, analyze the data and work properly for writing grammar of a language. So, I hope, by now you know what is all about writing grammar and how we can proceed to write a grammar. Please go through these references. Thank you!