

Tools and Technologies of Language Documentation
Prof. Bornini Lahiri and Prof. Dripta Piplai (Mondal)

Department of Humanities and Social Sciences

IIT Kharagpur

Week-05

Lecture-23

Lecture 23 : Fieldwork Conversations: Challenges, Precautions and Solutions

Welcome to the 23rd lecture of the course, Tools and Technologies of Language Documentation. Today we will have a discussion session. So it is not a lecture. It is more about talking and sharing experiences. So today, I have Vani and Anindita with us and Anindita is a native speaker of Sylheti. We know that Sylheti is not endangered.

But then there are various aspects of the language, which are endangered and there are various issues related to Sylheti, which she will talk about. And Vani is working in Mahasui folklore. She is also a community member and we know that Mahasui is an endangered language, as marked by UNESCO. So, UNESCO marks it as Definitely Endangered.

That is why, we know it is an endangered language. But Vani, you being a community member, how do you think the language is endangered or why do you think the language is endangered? So, thank you for introducing the language, ma'am. So Mahasui, it is spoken in two districts of Himachal Pradesh, Shimla and Solan, and we have a very large population of Mahasui speakers, and the number ranges from 1 lakh to 5 lakh speakers. But what is interesting is that, most of these speakers, they belong to the older generation. and they'd speak with themselves in Mahasui, with the parent generation Mahasui, but with the grandchildren, there's a shift that is evident.

So for example, my parents and grandparents, they will speak in Mahasui amongst themselves, but they switch to Hindi with me. And for reference, I'm 30 years old, so our generation, we do not speak Mahasui fluently. And the generations that are coming after us, they have a problem understanding Mahasui as well. So we see that the number of speakers, it is reducing with time. And we see there's a change in community attitude as

well.

So with education, what the community members started feeling was that, in order to learn Hindi and English, they give preference to speaking these languages in the home domain. So slowly, Mahasui is losing its domains as well. So that is why, we are losing our fluency in the language. And with this, I think there's a loss of identity that the community is facing.

Right. So, when we are talking about identity Anindita, in Sylheti, we know that Sylheti is not an endangered language, but what we see is that often, it is clubbed with Bangla as a variant of Bangla and that is why there are lots of things related to identity which is related to Sylheti. So being a Bangla speaker, I can understand little bit of Sylheti, but not whole Sylheti. If someone is talking in fast Sylheti, I will not be able to understand. But little bit what I understand, I will not say it is a variety of Bangla, because that much of mutual intelligibility is there in most of the Indo-Aryan languages, as they are part of a continuum and in continuum, neighboring languages they share mutual intelligibility. But then we see issue of prestige related with Sileti.

So, will you say something about it? Ma'am, as you mentioned, Sylheti is definitely not an endangered language. However, neither in India nor in Bangladesh, with a total population of about 11 million people, Sylheti is still not a recognized official language. In both the countries, it is clubbed together with Bangla. Having said that, since it is not officially recognized and thus, there are different problems that come out of it. For example, you might find Sylheti being spoken in marketplaces, in streets, in cricket fields and stuff, but you will not find Sylheti being used for formal purposes, for education and things like that.

So, since Sylheti is not an instrument for formal education, what often happens is certain aspects, if not all of Sylheti, but then certain aspects of the language do seem to be endangered or going towards it. It definitely is not in its safest form. Yes and when you are saying that Sylheti is clubbed under Bangla, that is what we also see in Census right? In census also, in Indian Census, language census, we do see that Sylheti comes under Bangla, it is not given a status of individual language right? So in this scenario, what can be the challenges related to data collection when we talk about Sylheti? Ma'am, the first challenge that I can think of, that I have faced in the field is that the very awareness, sometimes the informant does not even want to be recognized as a Sylheti speaker. Sometimes the informant would want to recognize himself or herself as a Bangla speaker. In other places, the informant would not even recognize themselves as a Sylheti or a Bangla speaker, but rather would just respond in English or Hindi.

So, I think first to even find a speaker who is ready to speak Sylheti, basically it all depends on the place that you collect the data from. Around the Brahmaputra valley, where it is the native place where Sylheti is spoken, there is a lot of Sylheti that is found and people are ready to give you data and are ready to show different aspects of their language. However, when you come out of that native land and you go to any other place in mainland India, I think the first problem or obstacle that we find is that there is nobody or there is very few people who would even recognize or identify themselves as Sylheti speakers, but would rather go along with the tag of being a Bangla speaker and then move on to speaking Standard Bangla. So, it depends largely on the location.

Yes. So, it is very important to know the exact location from where data can be collected. So, Vani how did you locate your places, from where you are going to collect data? So, when I started my field work, so Mahasui is spoken over large area. So, it is the entire district Shimla almost, where the population is scattered. It was very difficult to plot where to get data from, because we cannot cover the entire area per se and go to every village. So what we did, we did a small pilot survey first and this helped us locate where the language vitality is more.

So we took some suburban areas and some villages and over there, we found that it was in the villages that were near the forest covers, where we could find that language is more vital, it's being spoken in every day conversations and they are the ones who remember the folklore. In suburbs, the folklores are almost gone. So, this way we could identify that we had to find out places which were near the forest covers. And then this led us to go and contact the forest department of Himachal Pradesh and over there we have a lot of scholars like us working on forests and the vitality of the forest covers, which is in direct relation to a language's vitality as well. So we clubbed with them, we found places where there were forest departments and forest covers near the areas and we plotted those places and then, we visited those places and found people through random sampling, who would give us data on folklore.

So this is one method and then, another method, on online platforms to look for people who could help us find language experts on folklore. So, I think there are also advantages of collecting data online. So, when you were physically collecting it, maybe you had some advantage because you are from that area, you belong to that area, you know the area geographically, but for outsiders, it is more difficult. But then, when we are talking about online platforms, I also think that in online platform, the language experts or those who are responding to the survey, can keep themselves anonymous. So when they have this option of hiding their name and identity, maybe they can reveal something, which they can't in front of a camera or in front of you, right? So, what do you think? What were your observations? I definitely agree with this.

So, when we were on field, there were some differences we saw, when we were on field and we also collected some data from online surveys and we could see over there, that people gave more free answers. They were more open about the problems the language was facing. When we go on field, the entire family, most of the time, is sitting there. So the younger generations, who are losing the language, they feel some sort of pressure on them, so they give positive answers, where they would not want to. So online platforms, they have a sense of security, where they can express themselves anonymously and over there, we can also find that there are certain trends that are taking place which can be noted.

So, for example, on field they would report that they can speak the language, but on online surveys, when they are given some certain tests, so they would reveal that they are not fluent speakers, they cannot translate and they find it difficult to have a conversation in Mahasui. So, it is actually related to prestige we can say, right. So, when they want, when there is not much prestige related to the language, they can also hide their identity at times. So, what will be your input about Sylheti and prestige? Very interesting question. Now historically, it is important to understand how Sylheti is historically evolving, in itself.

For example, if you look at the 1960s, where the Assam Government decided that Assamese has to be the official language of the state, there was this uprising against that idea wherein they said that even Bangla has to be or Sylheti has to be a part of one of the official languages. But we have noticed that, what happened in 1961, after the uprising was that, Standard Bangla is also an official language in the Bramhaputra Valley in Assam. So what happened is although the Sylheti people were fighting for their mother tongue to be an official language, officially recognized, however it was Standard Bangla that became recognized and not Sylheti. Now, from that perspective if you understand, there are affluent families. So, there are educated families and there are not so educated families.

In educated families, there was seen or it has been observed that the parents teach Standard Bangla to the kids growing up. So, the mother is mostly educated in Standard Bangla. Now, there has to be some motivation for it and I would probably ascribe the motivation to educational needs, because it's Standard Bangla is what you need to get a job, to first pass your school level then, get a job and things. However, when it comes to not very affluent families, they would use Sylheti to even talk to their children. The second interesting thing that I found in the field was that Sylheti as a language, it is very widely used in that area.

However, what happens is that in other areas, Sylheti is not. Like you go to Delhi and ask a person, what language do you speak? They would ascribe themselves as Bengali speakers. They are not Sylheti speakers anymore. Vani mentioned that location is a huge thing. If you are talking to somebody in the native space and you ask them what is your language, they might identify themselves as a Sylheti speaker.

But if you are moving away, they would identify themselves as a Bangla speaker. So, that also talks a lot about the prestige of a language, that am I actually identifying myself as a Sylheti speaker or am I talking in public about something called Sylheti or am I talking in public about Bengali? So that gives in quite a lot. That also depends upon the neighboring communities, how are they perceiving the particular community. Yes i'd like to give you a small anecdote Ma'am. So I remember growing up, when I shifted base, my Bangla was laughed at.

So even as a native speaker, when i am talking in Standard Bangla, of course, there would be certain phonological or even syntactic differences in my Standard Bangla. So, I would be laughed at as a small kid and then slowly, I almost became ashamed of using Bangla in public spaces. So, I shifted to English or Hindi where the identifiability would be less. So, I think that also, everything together also adds to the prestige or the lack of it.

And vitality as well. Yes. So slowly, when a kid is actually laughed at, may be that kid will have the tendency to shift to another language. So, prestige is not only about the speech communities, what do they think about their language, but also what the others think, how do they look at the community or the language. So, prestige is also which we also get from the neighboring communities. So, how it is affected and whatever you said actually, I think these are things which we get through observation, but we should also keep a note of those things. So those observations, there are various types of observations, which we can talk about only if we carry a field note.

Because everything is not recorded. These things might not be said by someone, but when you are visiting different locations, you might observe that in one particular location, people are very easily identifying with the community, but in the other place, they might not be identifying with the community as Anindita said. So these things, I think can come under our field notes. So, when we are collecting or working in a language, we should always have a field note. So, Anindita you used to have certain field notes.

So, will you talk about it? Yes. The first thing that I would like to talk about field.. So I, as a person, would not like to make notes, but then after I visited the first field, I realized that that is one of the most essential things that you will be carrying in the field. Because,

there are certain data that you will get, for example, you will get linguistic data in your recording. But there are certain other data that you can only observe and find out.

So, like for example, if I am asking a person to reply to me in Sylheti and the person hesitates, so, that is not something that you can actually record it, but that can only be put into a field data. I think that is very important. And some others, I would like to pitch in here. So, related to all that you have talked about, on field that what I have observed is that, once I was recording and there were two elderly ladies. So, Mahasui has a lot of language varieties within it.

So, there is a lot of intermarriage within these language varieties and it was observed that these two ladies, who had just grown up together because they were old now, the marriages, they took place when they were 16 years old. So till now, they were mocking each other on the different language variety they spoke. So this, we can't ask them but this is something that we noted on the field and later on, we saw that this was a recurring event that women, who married into a different variety, they were mocked at and this is also one of the reasons why they stopped using the language altogether, and they've shifted to speaking in Hindi with their children. So field notes here are very important. And talking about women, what do you think, how do women participate as a respondent in the field? Are they allowed to speak in the field or because, I have seen at times, women are not allowed to come in front of the camera or the recorder or so what is your experience about it? Yes ma'am, definitely.

So as we know that women often times, they are not given the first preference to speak up. So when we go visit a family, we are introduced to the head of the family, which is a man and they will become the language experts as we go forward. So in this process, we often times, miss on the opportunity to take data from women, who at times know a lot more about the social practices of a community. So, we need to keep into mind that we need to collect data from the women as well. So, here I will again narrate an anecdote from field.

So, I had once gone to a remote village and I was supposed to collect data from an elderly man, but he had to go out due to some emergency. So, I was left alone there with his children and his wife. So, there was no way I could go somewhere else. It was a very remote location.

So, I thought I will collect data from his wife. So I asked the children if I could approach her. And they said, no, she won't be able to give you any data. She's illiterate. She's never gone to school.

And the woman, she was sitting there. And she said, "Yes, yes, I don't know anything. You wait for my husband. He'll give you data". And then I just asked her, had a friendly conversation. And then I said, I'll not ask you anything about your education, just tell me the sentences, whatever you do in your daily life in your own language.

And with this, we proceeded. And she gave a lot of data on Mahasui. She translated sentences for us. And she gave us a lot of data on folklore as well as we proceeded. So we need to involve more women, I think. I think we should be more inclusive, we should actually include everyone.

We should not see that whether the people, they are educated or not. So everyone, educated, not educated, women, men, older people, younger people. So, when we are talking about language documentation, we should actually include as many diversity as we can in the terms of collecting data from person, different types of person and also, as many people as we can approach. So, that we can get an overall idea of the language and the community. So, what will be your take aways for language for documenting specially folklore? For folklore, I will connect it with what I just said.

So as we see that, they think that, if a language does not have a script and it is an oral language, it is not a language. They say that this language does not have a grammar. So, it does not hold any value and it is termed a 'boli'. So this is one take away. While we are collecting data, we have to somehow navigate through the prestige issue they have and we have to find proverbs, folk songs, folk tales and in order to gain this data, we have to get the trust of the community.

So, will you tell us any proverb in Mahasui? Do you remember? So, this is a proverb that my grandmother used to say when I used to really be naughty. And now also, when I'm on field work, so I travel a lot. And my father also, he never sits at home, he's always working.

So then she says .. so that means, "Under an oak tree, only an oak tree can rise". So they take this that, I take after my family, my ancestors. So this is a knowledge system that, a metaphor that they've taken from the ecology around them, the trees that grow there. So, I think this is one proverb.

Yeah, this is very interesting actually. And yes something else, do you remember? I know you are not a fluent speaker of the language but still, if you have And can you also tell us that, while documenting the language, has it affected your fluency in Mahasui? Yes definitely. So before i started working on Mahasui, I could not speak any sentence in Mahasui. I could understand Mahasui, but with fieldwork, I have grown to speak in

Mahasui, although it is broken, but I have learned Mahasui. And during fieldwork, we also observed that when I tried to speak in Mahasui, even though it was broken, the respondents, the language experts, they were more open to giving data.

So, that is one observation. So, you think that the actually language experts' language also affect the respondents language, right? Yes. So, I am sure Anindita, you must have seen this because Bangla and Sylheti are very closer. Yes, I remember during that same field study, I was not sure as to what language to use because whatever language I was using was just being repeated in the answers. So at that point, I didn't want to use Sylheti because I really didn't want the data to be in Sylheti and then, I wanted to use Standard Bangla but even my diction, the choice of words that I was using, the kind of intonation that I was bringing into the language, was also affecting the informants' language. So that is something that is a little tricky and while doing field work, it is also very important to be.

. I think if the researcher, if they are from the community, it becomes easier because then, you can also relate with everything around. But if you are an outsider researcher, you will have to put in a little effort to understand the small nitty gritty of the language to get better efficient data. Yes, that is true, but I will also say that being a community member has its own challenges. So, Vani once you were narrating that how people used to ask you from which village you are and all the details about it and depending on that they used to react.

Yes. As I mentioned, that the language has a lot of varieties. So, if you are a speaker of one, so for example, I speak a language spoken in Kotgarh. So, my data will also be influenced by their perspective of people belonging to that region. So, regional differences within the language speakers, that is also a big factor in getting data. So for example, I once visited a village and I randomly approached an old lady and she gave me data and after that, we started talking and apparently my ancestors had a big feud with that village's ancestors and after they came to know, they were not as open to me collecting data from that area. You were also saying that though you were a part of the community, still you had problems while collecting data related to rituals and festivals. Yes, so as every community has some taboos, some restrictions that we need to be careful of, so as a community member, I think we have a benefit that we are more receptive to getting to know where to stop and what not to do.

So, for example, I was collecting folk songs for Mahashivratri. So, over there, I was allowed to document the event with the audio device. But after some time, they asked me to just observe. They said, please don't record and just observe if you want to. Otherwise, you can leave and that also was a very big thing because that is an event in which mostly

the

men

participate.

Women are just observers there. So for them, to let me record that that I could do only because I was a member of the community. If somebody outside the community would want to do that, they would not have allowed to do it. So, of course, there are some benefits of being the community member and you also know the norms and all those things. So, these are the some of the points which we learn from here, right. So, will you like to add something to it? About some of the attires and maybe, your jacket, which is a traditional one, you are wearing.

So over here, this is traditionally in my region, it's called a saluka and this is also something, that is traditionally worn by all women of Mahasui community. And, there are different changes that are there. So, we can get to know different variety speakers from this. So, for example, if I am, I have just buttoned up, I have two buttons here.

So, this is a sign that I am from Kotgarh. From someone, who will button up three, it will be from a different region. If you wear a cardigan over it, you are from a different region. Under it, there is a different identity. So, all these things I think. These are very very interesting I think and these things are so interesting and they should be actually noted and kept a track of.

So, may be people are shifting to other modern dresses, but this knowledge and this tradition should be at least documented and kept. So, do you think people also use this jacket, the modern generation? I'm sure some people do like, you also have it. Yes, I think we started identifying ourselves with the community more, after COVID when we went back. So, I think there is an increase in the prestige after COVID.

Okay, that's actually a good sign then. Yeah, maybe because community members, they were together and they were using online platforms for using their language during this covid time and maybe that is why, that is also added prestige to the language. Basically, during these discussions, the points which we talked about were importance of field notes, which is very very important. We should have because this type of information, like two buttons in one village, another three in another place. So, these type of observations can be actually kept in the field notes. So, these type of information and then, we also came to know that in languages like Sylheti and there are lots of languages like that, you can see there are different varieties named under Hindi in Census, language census of our country.

So, all these languages might be very different from the cover term that they are kept under. But how do you collect data for those languages because there is prestige related

to it. So, sometimes speakers will identify with that language sometimes they might not they might shift to the other more dominant language. So, it really becomes a challenging task to document those languages. When we are not documenting those languages or when we are not providing materials in that language, then of course, there can be a shift.

There can be a tendency of shift in the language and as Anindita said, certain parts can get endangered like kinship related terms or terms related to food items, those can get endangered. So these are very important points Another important point is that we should be more inclusive in our approach. We should try to include as many people as we can in collecting data. So, we should not be only approaching men or only women. We should be more inclusive, everyone educated, not educated, that does not matter actually when we are talking about language.

So, that is very important and another important aspect is also about collecting data; from where you are collecting. You should locate the exact location, the more authentic location and also, you can use online sources of data collection because online service has its own advantage. Of course, whole of documentation work cannot be done through online studies, but then online studies can also be a part of it. So, that can also be very important. So, these are some of the aspects which we actually talked about today and I hope I have covered all the points and so that is it and I hope you enjoyed our discussion today. Thank you!