Tools and Technologies of Language Documentation

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

Department of Humanities and Social Sciences

IIT Kharagpur

Week-03

Lecture-11

Lecture 11 : Designing Questionnaire, Part 2

Hello! This is Lecture 11 of the course, Tools and Technologies of Language Documentation. Today, I will talk about some more aspects of designing questionnaire. So, I already mentioned about Translation Method. But today, I will talk about some other methods like Picture Book, Object Naming, Video, Role Play, Discussion Method, Interview Method and Observation Method. Of course, you can always add to this list, but these are the basic list, which we use, generally, for language documentation. And when we talk about these methods, there are certain methods, which are more useful for certain type of targets and others are useful for the other type of targets.

Like when I talked about Translation method, then Translation method is a easier way of collecting words in isolation or sentences. So through Translation method. you can easily get those type of data, but when we talk about other types of words, then we can see that some methods like Discussion method, Interview method or Observation method can be useful for Language Attitude and Vitality as well. Because for language attitude and vitality of course, you cannot use methods like Translation method or Picture Book method.

So Translation method, Picture Book method, Object Naming method or creation of video, all these can help in collection of words, sentences, narrations, speech data, various types of actually speech data, but then again, it can also overlap. It is not a very straightforward division. So these are some of the methods, I will talk about today. So narration is a good means of speech data, because through narration you can get fluent speech. You know that there are various morphosyntactic rules, which can change pronunciation of a particular word or it can change morphology of a word.

So, that is why when you get narrated data or narrations, you can analyze the data, which it is actually used in day-to-day conversation. So, narration is a very good way of

collecting natural speech data, but the problem or difficulty with narration is that if your ears are not habituated to the sounds of the particular language, then it might be very difficult for you to understand the words or the sounds of the language. a drawback of it, but once you are habituated to the sounds of the language, once you know the basic patterns of the language, then it can be a good method for collecting more fluent speech data. When you start collecting narrations, then at times, language experts can be little hesitant about it, they might not talk about it. So if you say just "give me a narration", they will not understand what to talk about.

Or if you say that, "tell me a story" then initially, one can be hesitant and say that "I do not remember any story" or "I cannot talk about stories", like that. But to begin with, it is very good if you start with children's stories. So, you can say that "What are the stories that you tell to your children or grandchildren?". Why it is very helpful? Because most of the parents, grandparents or elderly people, they know certain stories which are common in the community. And these stories because they are children stories, they have a very simple   structure.

The grammatical structure of the sentences are simple. So that helps you in analyzing the data quickly, because you are also new to the language and that is why, when you get a simple structured sentence, it is easier to annotate and a transcribe, that is why it is easier to begin with children stories. And again as I mentioned, most of the adults of a community they generally know certain children stories. Another good way of starting is that you can ask the language expert to talk about their daily routine. Because stories at times, they might say that "I cannot recall" or they might not be very willing to say, but daily routine they cannot say that "I have forgotten"; there is nothing to forget.

So, daily routine you can ask "What do you do when you wake up in the morning and after that?". So "What do you do? What is your daily routine?" You can also ask about their siblings or their children. You can ask about their daily activity. So, "What does your child do?" or "What does your sibling do?" So, these things they can narrate; one can easily relate to these things. So if the topic is relatable, then of course, language experts will feel more comfortable and they will talk about it.

So, it is good to ask them something which is relatable. Another very relatable topic is what is happening at present. So, if you are in a field, you are among the speech communities. So, lots of people are there and something is happening, may be a marriage ceremony is going on or some other thing is happening; simple things like a mother is cooking food. So these thing, you can ask someone to narrate, "tell me what is she doing".

So, "tell me everything in details, what are the spices, what are the vegetables, what are the things which are used, what is she doing?" or if any event is going on, there can be sudden events as well. So, once we were in the village of Purulia, collecting data for Kurmali, when we saw that a fire broke in the village. And then the villagers were running around and all and then there was a little girl. We asked her to narrate "what is happening?" because we could not understand. And then this girl narrated the whole thing in Kurmali and we recorded it with her permission and her parents' permission.

So those type of things can be done. So if something is happening then, but of course, we should not disturb what is happening, we can just be there and take a note of it. We can collect the data. And when I say collecting narrations, you can also collect narrations if someone is like delivering a speech or someone is talking about something, you can collect those, but then one has to be very conscious that one gets the permission to collect it. As I mentioned in my previous lectures as well, you can never record anything with a hidden recorder or camera; the permission has to be always there.

May be it can be taken before or after the event or after the recording, but it has to be there. Always adopt an ethical way of collecting data. Picture books are also very interesting. Either you can show picture card prompts, which are also very easily available. You can easily get some card prompts with pictures, various types of pictures and ask them "What is this? What is that?" You can create your own picture book.

So, may be picture books for body parts, picture books for flowers, for birds, for various types of things. Picture books can be created. One can actually visit the field and create a picture book out of it or one can have basic picture books like that of body part. So, how does picture books help? So, if you ask someone "What do you call your forehead?", one might not know what "forehead" means or there can be certain other type of division within the head. It might not be for head and head.

So, if you show a particular picture and ask about that particular part, then one might easily tell you about that. So like parting of hair, there can be a word for it or there might not be. So if you use Translation method, one can get confused. One can get you some wrong data, if they do not have a word in their language or if they do not understand what are you saying about, but it is very clear if you show them a picture and ask "What is happening?. So in this picture book, you can see that certain parts, the English words are visible.

So one can actually hide the English part and then show the picture and ask. So that can be done or nowadays, we often carry our tablets or smartphones in the field, there also we can show them a pictures and ask them "What is this? What is that?" But always, we need

to index it properly. So, we should have the picture and the targeted word, matched with each other, otherwise there can be confusion. So, one should take care of that. So, picture books are also useful for collecting data related to locations, because when you are saying "The ball is under the table" or "near the table", there can be again confusion because in different language, the locations are perceived differently.

So, if you are saying "The tree is behind the house", "The tree is beside the house", "The tree is in front of the house", there can be different words for it or there might not be. Like in English, we have words for "above" and "on" right. So, I will say "The fan is above my head", but "..hair is on my head", but if I translate it in my native language, which is Bangla, then I will use the same locative marker for both "above" and "on".

So now, there can be differences depending on the location, depending on the language, what words are used for different types of location. So, it is always good to use picture books in those cases. So these are few examples, you can of course, add more depending on various other type of location. So if "The ball is in the box", there can be a certain type of structure, but if something "is inside a well", there can be a certain different type of word used. So if you have lots of pictures and then you go to the field, then you can get more accurate data and at the same time, it is also very interesting when they look at the picture, think and tell, rather than you give them one sentence and ask them to translate.

And picture books are also good with the children. If you want to collect data from the children, you can also use picture book, you can ask them "What is this? Where is the ball?" tell me. So, things like that. So, picture books are also good with the children. And you can also imagine, when you are speaking like 10, 20, 15 sentences and asking them to translate it, the whole thing can be little boring or tiring, but with different types of colorful pictures, it can be little interesting.

Again, I will mention that when you are creating pictures or picture books, you should be conscious about the speech community. Do not bring pictures which are very alien for the community, always try to make it relatable to the informant. Picture books are also good when you want to get some comparisons, like say "fat and thin", "short and long". So, you can have two figures one is fat, another is thin, one is short, one is long. So you can get those types of words.

So you can ask them to compare or you can ask sentences for those type of objects. But abstract things are again little difficult to be collected, sometimes translation methods can be used, sometimes different types of picture books can be used. For colors also, if you say "what do you call red? what do you call blue?", it can be little difficult because "blue" for my cognitive sense can be different than your, depending on the language we speak.

Like you know, for some communities there are only 7 colors in a rainbow, but for other communities we have seen, there can be 14 colors in a rainbow. Depending on our perception of the colors, we also have different names for the colors, like everything else.

So if we can create color cards and then show them the cards and ask them "What are these colors?", then you can get better names, because there can be names for different shades of yellow, which is actually yellow and which is actually orange. In one language, shades of orange can come under shades of yellow, but for other language, shades of yellow will come with shades of yellow, not with orange. So, things can differ depending upon the speech community. So, what is yellow for me, might be orange for you. So, it is always good to carry certain types of color cards and then use it.

So, these type of some of the processes are adapted by various teams which work on language documentation. Now, these were the picture books I was talking about, which you go well prepared before going to the speech community. So, you prepare you have your own picture cards, you have your own picture books and all, but then when you go to the field and you find new items, what do you do? So, a way of collecting these new items is that you can collect pictures of these items. So as I mentioned, this is a teamwork. So maybe, one or two members of the team can actually move around the speech community, talk to them, get more close to them and click pictures of different items   surrounding   the   area.

And then out of those pictures, a picture book can be created and then after a few days, you can use that picture book for collecting data. So, you can ask like you can see in this figure, Figure number 1, you can ask "what is this figure, what do you do with it, where is it used, what is it made of?" There can be various questions related to it. So, similarly for Figure number 2, you can have 2, 3 more than 2, 3 items and then you can ask "what is it called in your language, what you do with these items, how are they made, where are they used?" So as I said in documentation, it is always necessary to collect as much information as you can and you know that in every domain language is used. So, it is always good to collect more and more information because of course, you are getting more and more linguistic data. And also to get the world view, which I talked about.

Through these items also, you get to know the community better. Some types of objects you can of course, click and create picture books and then there can be certain objects, which you can point and ask in the field itself. If you do not have time or resources, where you can click pictures and create a picture book and again visit after few days, you can actually in the field ask them "what is this?". So, if you see this figure, one of course, are baskets, but they are of different shapes and they have their own names, it is taken from   Mahali   community.

Other picture is a garland of fish head. It has been kept there to get dried in the sun, which is used in different types of curries by Dhimal community. So, these are the pictures which we have clicked. So, and at the same time, we have asked them "what do you call it?". So, they have a name for it. "Where do you use it, why do you use it?" So there are various questions which we could ask, depending on the object.

So those are also a good means of collecting data. So you are in the field, you collect as much information as you can by just pointing at the objects and asking the informant to tell about it. And when these objects are surrounding them, then the chances are that the informants will be able to tell about that because those objects they are using in their day-to-day life. And of course, it is that we should not use or ask about certain things which are not much important, like you cannot point towards a television and ask or point towards   a   smartphone   and   ask.

So, these are very common-sensical. We do not do that; we go for more indigenous type of items. So, now creating videos are again very very useful way of collecting data, which can be used with the adults as well as with the children. Children feel very excited and entertained and at the same time you also get narration data. So generally, when you show video, you ask them what is happening in the video. Videos are very useful for collecting   action   words   and   adverbs.

So if a person is walking in say three different ways in the video, the language expert can explain about the walks, can give different words- adverbs for describing the way the person is walking. So, you can get these type of words from the video, you can also get different types of comparative words from the video, but video should be very short, like 1 minute- 2 minutes, like that; it cannot be very long video. And again, the video items should not be alien to the community. They should be relatable. One can relate to the items or the performances which are happening in the video and they should react accordingly.

So generally, these videos are silent videos, there are no dialogues or anything in these type of videos. There are certain actions being done. So it can be anything, and videos nowadays are very easy to be created. So we see in social media, that people are everyday creating lots of reels and videos and posting it. So what we can do for documenting languages? We can also create these smaller videos, or what you can say, reels, we can create those in our smartphones and then show those to our language experts and ask them   what   is   happening.

So, that can give you more accurate data, and also narration. So, you will get proper

narration and fluent speech through those type of videos. But again while making a video, one has to be very conscious. It should not be something which is insulting to the community and at the same time, it should be very neat and it should be ethical as well.

So, keeping all these considerations in mind. And of course, I am talking about creating videos because you cannot take those videos, which are posted in social medias and all, and keep it muted and get the data, because they come under copyright issues. That is why, if you create your own video and if you can create that in the field, maybe you can ask some of the language experts to help you in creating those type of videos, then that is very good and relatable to the community and then you can get more data. And videos, as I said, are also not boring; when you are collecting the data, language experts also feel little entertained. So, there is another way which is very new, but quite often used is the Role Play method, where you involve the language experts or the community members to play different types of roles. because you cannot actually go to the scenarios, like you cannot go to a bank and then try to document or record.

Maybe, people will not allow you to do that. They have privacy issues; you will hamper their work and all. So, what if you ask someone to play or act like a customer and another person, like a bank manager or maybe the third person like a office manager or something else, like a peon. You can actually involve two three people and ask them to play role play; small script type of things. And generally, a topic is given. So, you can ask, you people can talk about how to give loan, how to open an account or something else.

So, how to relate your Aadhaar card or something like that, with the bank account. There can be various types of scenarios, real life scenarios, not very imaginative, but real life scenarios and then they have to act accordingly. Like two strangers meeting in a market, maybe one is hurt or something falls from one's pocket and the other is actually picking up and saying things. So, how do they talk or when two friends meet in the market, how do they greet and then they talk.

So, all these give you the real life scenario. So, because when they are enacting, we assume that they are doing it, as they do in their real life. So this type of role play can actually give you more realistic data. So those thing can also be used for collecting data and these are also entertaining. So those, who are involved in the part, will also not feel bored; they can actually act, but they also need lots of time from the community. So when you are closer to the community, they know you well, and then you can request them to do these type of activities for collecting or documenting the language.

Discussion and Interview method is used and as I said, it is more useful for language attitude or vitality related topics. So, one or two popular topics are given to the speech

community members and then they are asked to discuss. They basically discuss among themselves and you keep the recorder there. You can also act like a catalyst; in between you can ask few questions or you can guide them towards something, you can do that and collect more material. So you get linguistic data, you get to know about their attitude and opinion.

So what do they think about their own language, whether the attitude is positive or negative. So by now, I assume you know about Language Attitude. So, you can get to know about the attitude towards their own language or attitude towards their neighboring language by giving them these type of topics. Interview method is important to collect language attitude, of course. At the same time, you can also collect words and sentences by   using   Interview   method.

So, what do you do? You ask the language experts to say, like "Can you tell me names of 10 animals, or 10 flowers or names of the months in your language?". So, you just ask them few questions and they have to reply. So they can name any 10 flowers, any 10 fruits or anything. So these are some of the methods through which you can collect data.

Observation method is a very very important method. Of course, you do not actually get prepared for that before visiting the field, but in the field, you have to be always alert, for this method and this has been actually mentioned as a very important method of language documentation. So, Himmelmann says, "of major importance in this regard are specimens of observable linguistic behavior..". So basically, you are observing how they are behaving in their own language, that is, examples of how the people actually communicate   with   each   other.

This includes all kinds of communicative activities in a speech community, from everyday small talk to elaborate rituals, from parents baby talking to their newborn infants to political disputes between village elders. So, everything is involved. How they are using the language and where they are using the language? Where they are shifting to maybe some other language? Everything can be observed. Why I am saying observation method, or why does Himmelmann says Observation method is very important? Because sometimes, some answers you can get from the Interview method, which may not match with   your   observations.

Something is informed, but something else is actually practiced. Because when they are answering you, may be the language expert wants to be more moral or wants to be more bookish. So, they are giving you those things which should happen, but they are not talking about actually things which are happening. So, they are becoming more predictive, rather than saying what should be there, rather than being more descriptive. So

for my own experience, what we observed in a community is that, I am not going to name the community, but there was this let us tell x language. X language's activist, who was talking a lot about saving his language and how actually he promotes his own language and all those things.

But then, when we went to this activist's house, and we ask his children to sing song, any song in their native language, of course, the children were little hesitant earlier, but then when they sang, they said that "We can only sing Bollywood songs. So, is it ok?" We said "ok, fine". But through this whole situation, what we came to know is that, though the activist was talking a lot about his language preservation or using his own language, in his home, maybe he was not practicing it. There can be of course, valid reasons for not doing so, but he could not do so and that is why his children cannot sing a song in their own native language, because they do not know. So, and of course, the children said that they do not know, they only know Bollywood songs.

So, there can be things like that. I will also like to share another experience, where I was told by a community member that we only marry within our communities, we are not allowed to marry outside the community. But then when I was talking to this person, I suddenly said that "Ok, can I also talk to your wife? Will you please call her?". And then he says, "No, I cannot call her because she cannot help you much. She is from another community, she cannot talk in this language or she cannot tell things about this language".

So, this was just a paradox what I witnessed, but this was fact. So, what he was saying was by the rules, this should not happen; might be the community should not marry outside a community, but in practice, they are actually marrying outside the community. And this observation was important because this also tells that within the community, outsiders are coming and that means, they are also bringing their own language. So, there can be mixture of languages. So, you can think about the children of this man, who has father from a linguistic community and a mother from another linguistic community. So, this can tell me a lot about the linguistic scenario of that particular community.

That is why observation is very very important. Another importance of observation is that in observation, we can actually observe various types of rituals and new objects. So sometimes, you are not allowed to record everything; certain rituals are very restricted, restricted only for the community members. You might be allowed to see it, but you are not allowed to record it. So in those situations, what you can do? You can actually keep a note of it, you can observe as many things as you can and then keep a note of it. So when recording is not permitted, when certain rituals do not allow you, in those cases also, observation is very important.

But again, observation is also important to see how they are behaving in their language, where they are using it, where they are not using it. So we should never ignore observation method, we should always be very alert when we are in the speech community and try to collect as much information as we can, also through our observation. So creating a relationship with the speech community is very very important. Of course, you should always go through ethical methods as I said, and at the same time, you should also try to be friend to the community members.

It is not like give-and-take thing, it is also like working together. And it has actually become a method of collecting data, which I will talk about in my next class, where the language experts and the researchers, they together work for the documentation of a language. It is not like the researcher only collecting data from the language experts, together they are trying to save a language or document a under-documented language. So when you go to get closer to the community, you might get informations which are actually not much visible from the outside. So, this is a process of creating local alcohol and this I got from a community where I got very close and then initially of course, they were little hesitant to tell me about it because of various law and orders and other issues, but then slowly when I become close to them, they actually took me to that area where it is created and talked about the whole process, when it is consumed, how it is consumed and everything. So, when you get close to the community members, you can actually get lots of informations about the speech community and then it is your responsibility to actually what are you doing with the data.

You cannot of course, make everything very public and you can of course, take care of the data. So that, the community members are never harmed due to the data that they have given you. So these are some of the references which you can consult and I hope you enjoyed today's talk. Thank you!