Appreciating Linguistics: A typological approach
Dr. Anindita Sahoo
Department of Humanities and Social Sciences
Indian Institute of Technology, Madras

Lecture – 55
History of the developmental Articles

Hi, hello everyone. Welcome to this session of my NPTEL course Appreciating Linguistics: A typological approach.

(Refer Slide Time: 00:24)

Let us go back to the discussion of the very first set of questions that I had in the previous slides. I said following Moravcsik's book Introducing Language Typology; one of the developmental processes can be traced by using articles as a tool. What do you think, how did articles come into existence in the history of language? A lot of crosslinguistic distribution of definite and indefinite articles have been studied.

Roman Jakobson, one of the finest linguists that we have ever had, he says that languages differ essentially in what they must convey, not in what they may convey. There is a difference between must and may. When we say language A is different from language B, we must account for the solid differences that we observe through the empirical evidence or

through the empirical data. If you assume something, it may convey that does not really help us to find out the differences.

When you are trying to figure out or when you are trying to understand articles, the idea here is that you have to talk about or you have to find out what sort of difference they must convey. So, on this note, let us see how the article story has been talked about or has been discussed in typological literature.

Jakobson's insight is born out by the fact that many languages do not require that this distinction be made; they have no articles at all. For example, Russian for example, Korean, in Hindi we do not have an article. So, considering some languages have articles, some languages do not have, this clearly shows that crosslinguistically the presence of articles is varied. It is clearly demarcated, it is clearly differentiated one type of language type A which has articles, language type B, the languages is in this type which do not have articles and they are not overwhelmingly frequent.

If you look for the numbers, Dryer had this sample of 566 languages and out of this 566, 337 have words or affixes functioning as definite articles and 41 of them have definite, but no indefinite articles. So, 337 plus 41, that is the number of languages which might have articles or which have articles, but the rest do not have; and in the sample size of 473 that we discussed less than half like 204 have indefinite articles and 81 have indefinite articles and no definite one.

There are two different samples we have; one sample 566, and in this 566, 337 have words or affixes that might function as definite articles. 41 have definite, but no indefinite. Then, another sample of 473, 204 plus 81, that is the division. So, these are just some numbers I want you to look for.

This crosslinguistic distribution of articles does not allow either for unconditional statements like all or most, or for implicational; if x then what, most languages have definite articles also have indefinite ones. So, this kind of striking convergence or this kind of crosslinguistic distribution cannot be accounted for when you look at the numbers like this. Then how we are going to draw the crosslinguistic generalization as far as the existence of articles are concerned.

Before that, I just want you to have a look at the data, all the examples given over here. In English, the lack of non-phrasal status of articles, I do not want this I want that, that is ok, but when you say I do not want this, I want the, that is going to be bad. So, here the article though it is a definite one, you cannot have an definite article without a noun. So, that is ok, but the would be bad.

Now, look at 2 a, these are nice apples, I want one of them. These are nice apples, I want an of them that is going to be wrong because though an would mean basically one that is going to be ungrammatical, non-phrasal status. So, when you look for an article it must have a phrase like an apple, a book. It must be considered as a phrase; independently without considering it as a phrase, such kind of articles do not exist in English.

And, in case of suffixed articles, look at the languages like Bulgarian, Albanian and Rumanian, so trup-at that means, the body; konj-at the horse; mik-u in Albanian the friend; djal-i the boy. So, -u, -i, -at these are the suffixes which work like articles.

Did you notice there are two different types of articles crosslinguistically which can be accounted for? One is definite, indefinite; sometimes these definite articles could be treated as or could be manifested as suffixes. English is one type; Bulgarian, Albanian and Rumanian, they would be the other type.

(Refer Slide Time: 06:49)

Then, the third set of data that we have here is Hungarian. In the Hungarian example, the articles lose their inflections of their ancestors, but eventually they have the trace, but then the inflections are lost. Look at the Hungarian data, this apple, and the apple. This apple is ez-t and the apple is az. Apple is alma-t here and ez-ek-et az then alma-k-at whatever I am not going to read this. So, when you say this and the, you see the inflection has been lost over here.

So, ez-t, ez-ek-et. So, this at in the apple accusative, disaccusative and here it is the. So, this; this and that, that accusative marker or they are going they are basically losing the inflection mark over there. When you look at the history or when you look at the division of articles in different languages, what sort of generalization we can draw?

This is tiny bit of data, but I would expect you to go back and check more examples. Now, since we were talking about the initial, final and intermediate levels, what sort of crosslinguistic recurrent pattern we can draw from this initial and final division?

Let us look at the first generalization in connection with the articles. The first question that we are talking about as far as the diachronic perspective is concerned is that how do you think articles came into language which whatever language might have it, let us say indoor in languages like Hindi and Oriya they do not have, but English does have it. I will very briefly talk about what could be the possible pathway or what could be the possible process by which the definite and indefinite articles have been derived?

Let us think about two examples that man versus the man. Semantically, generally they have a similar kind of meaning, but not identical. That man is also definite, the man is also definite, but when I say that man, semantically it is a little different. This gives us some idea probably the demonstratives like this, that, these, those and the definite articles like the, they might have some kind of connection. Similar situations or some similar scenario are also attested for indefinite articles. So, that versus the, we can consider the English data.

Let us say if we have similar sort of thing for the indefinite ones. For this I would like to add a bit of Turkish data. In Turkish both the definite and the indefinite article and the numeral

one, they have the similar form Let us forget about Turkish, let us talk about English also. The same English data for the indefinite articles, we can say one book and a book. Don't you think they have some kind of connection?

This is just a hint that indefinite articles might have some connection with numerals like one and definite articles might have some connection with demonstratives. We will explore more things in the book, but then that is just one set of data that I want you to focus on. And when I say the demonstrative that and the definite article the, they are similar, but not identical; so, to understand this, let us look at the data given on the slide here.

1 a: I do not want this, I want that. I do not want this, I want the that is ungrammatical. So, that means, that and the might be semantically similar, but they are definitely not identical. Same is the case with 2 a: These are nice apples; I want one of them, but I cannot say they. These are nice apples; I want an of them though the indefinite article an seems to be derived from the numeral one. So, some kind of derivational link we can find it out and that is not the only thing that you need to remember as far as development of articles are concerned.

There are also instances where you see the articles gradually are losing the segments like the phonological erosion is also happening; not only through the segmental levels, but also through the. So, that, the – there is a segmental loss. This is what we would call loss of segments, but besides that, during the process of article development, the phonological erosion is also happening

I would like you to look at the data for Bulgarian and Albanian that I discussed a while ago. In this case also, because of the phonological erosion the articles have become just a suffix; they do not remain an independent morpheme anymore. Look at trup-at. So, trup-at is not trup and at separately; -at is the dependent morpheme. This is also an instance of phonological erosion or reduction of the phonological form.

One thing is that segment erosion from that to the, one to an; the second one is the phonological reduction where the articles have become suffixes. Besides that the article, there are instances where the articles have also lost their morphological inflection. So, not only segment erosion, but also phonological erosion, as well as the morphological erosion. So,

what kind of morphological erosion is happening or and what sort of inflections the new articles are losing you have to check the Hungarian data for that matter.

Look at the Hungarian data in 6. This versus the, see ez-t and versus az. The -t inflection is missing here. Ez-ek-et, so the becomes just az. The plural inflection is missing and then the accusative marker is also missing. So, this ez- from ez-t and ez- from ez-ek-et is also an example of morphological reduction. So, this is the pathway. If you go back, the initial stage and the final stage, what are the different intermediate paths or what are the different intermediate stages that the articles have been through?

On this note, we will go to the generalizations because whatever we are going to discuss now we have to focus on the four stages or the four points that we discussed. So, what could have been the initial stage, what could have been the final stage, what are the intermediate stage and what could be the conditions of this change from the initial to the final? And, remember we are discussing articles and we are trying to find out what could have been the initial stage of the article.

So, what does generalization 1 say? Point A; point A is about initial stages. Let us read what is written here. It is written given a definite article, it is likely to have arisen from a demonstrative that versus the; the is the definite article and the typological generalization is that if a language has articles and demonstratives both, then very lightly the definite article has come from the demonstratives and given an indefinite article, it is likely to have arisen from the numeral one. So, just like the definite has risen from the demonstrative, the indefinite has arisen from the numeral one.

So, what was the initial stage? The initial stage was demonstrative, final stage is definite article. Initial stage is numeral one; final stage is indefinite article an or a and that is about the point A. What does point B say about the final stages? Given a demonstrative, it may change into a definite article. Demonstrative had certain features which resulted in the change or which resulted in the occurrence of a definite article; given a numeral one, it may change to an indefinite one. So, if your language has a demonstrative, it might result in a definite article; if a language has a numeral one, it might change it into the indefinite article.

So the initial stage is demonstrative, final stage is definite; initial stage is one, final stage is an or a. Now, once we decided about the initial and the final stages, what should be the intermediate ones? Because both the changes are instances of gramaticalization, that is the change in the grammatical category, what are the mediary stages? There has been phonological change, semantic change, and morphological change.

Change in the phonological level is an intermediate stage, change of semantic level is an intermediate stage and change of morphological level is also an intermediate change. So, these are the intermediate stages going through that a demonstrative becomes definite and numeral one becomes indefinite. So, this is the pathway.

Now the fourth point is what are the conditions that could be the reasons why such kind of changes happened? There is primarily one reason that is language contact. Because of the language contact this might trigger or accelerate the development of articles. So, the-that is the story of articles that we have in our typological literature.

Let us summarize it. We have to find out four points of crosslinguistic recurrent generalizations. The first point is about the initial stage, second is about the final stage, third is the intermediate stages and fourth is the conditions that we have to talk about.

As far as the articles are concerned, the initial stage are demonstratives, and the definite articles are derived from the demonstratives. Indefinite articles are derived from the numeral one. So, what are the final stages? Demonstratives have some quality which may change into definite article and numeral one has some quality which might change into indefinite article.

So, that is the initial and the final. What are the intermediate stages? Phonological changes, semantic changes and morphological reduction as in Hungarian; phonological changes as in English; semantic changes also as in English the data that or Rumanian data also we had.

So, these kind of changes are there in the intermediate stages and what could be the possible reasons why these changes are happening? This has been triggered or accelerated because of the language contact. Articles are one of the finest examples how language contact leads to language change.