



**Fundamental Concepts in Sociolinguistics**  
**Professor Doctor Om Prakesh**  
**School of Humanities and Social Sciences**  
**Gautam Buddha University, Greater Noida**  
**UP-201312**  
**Lecture 13**  
**Dialect**

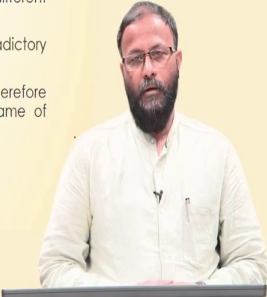
(Refer Slide Time: 00:15)



  
NPTEL

**Dialect vs. Language**

- The debate on language and dialects is very old and continues till date.
- It refers to the linguistic attitude and values that we assign to a variety.
- The idea is socially, culturally, and politically rooted.
- There is no difference between a language and a dialect in terms of linguistic characteristics.
- There is no universally accepted criterion for distinguishing two different languages from two dialects (i.e. varieties) of the same language.
- A number of rough measures exist, sometimes leading to contradictory results.
- The distinction (dichotomy) between dialect and language is therefore subjective (arbitrary) and depends upon the user's preferred frame of reference.



Welcome to class. Today we are going to talk about a very important and debated topic in sociolinguistics called dialects. So we are familiar with the term language with all its properties and characters. But when we talk about dialects, we have certain social, political, and non-linguistic references. And this issue has been debated till date. It continues till date. What is the dialect and is it different from language? Is it inferior to language? How do we define a dialect? Who speaks this variety and how can we identify whether a person is speaking a particular variety is language or dialect? It is a debatable question.

And we do not have a clear cut widely accepted universally accepted definition of dialects because we have different frames of references in which we identify dialects. But again, the results are debatable. So the dialogue or the debate on the difference between dialect and a language is much older and it continues till date.


So while we classify or name or identify x as language and x as dialect, then it refers to a linguistic attitude and values that we assign to that variety. So there is no linguistic argument for naming it as a particular dialect. This whole idea of dialect is socially, culturally, and politically rooted. We need to understand that.

So there is no difference for a linguist; for example, there is no difference between a language and a dialect, because what you call a language has been in some time a variant or a dialect which has got institutional support, which has got a prestige value attached to it, which has got the support of the elites of the society, which is perceived as standard language. It is no longer a dialect.


But all languages that we speak are a favoured dialect. When we say favourite dialect, that means we are talking in terms of socio, cultural, and political favours given or extended to a particular dialect which becomes standard dialect and language. But there is no universally accepted criterion for distinguishing two different languages from two dialects.

However, a number of rough measures exist to identify and classify language and dialects. However, these results are contradicting and always debatable. So we can summarise it as something like the distinction or dichotomy between language and a dialect. It is therefore subjective, arbitrary, and depends on the user's preferred frame of reference. We will talk about this frame of references shortly.

(Refer Slide Time: 04:09)



- "There is no real difference between dialect and a language". [Sapir]
- "In the course of survey, it has sometimes been difficult to decide whether a given form of speech is to be looked upon as an independent language or as a dialect of some other definite form of speech". [Grierson]
- In practice it is sometimes impossible to decide the question in a manner which will gain universal acceptance.
- The two words 'language' and 'dialect' are in this respect like 'cup' and 'mug'. We all understand that these two are different, however, between these two a dividing line cannot be accurately drawn.



So as we stated, there is no real difference between a dialect and a language as Sapir observes. If you look at the statement of Grierson, he too observes in the course of the survey. Grierson (04:25) was doing a survey of Indian languages. He created volumes and conducted a very very systematic survey. Then he came across this dichotomy and he referred to and observed, in the course of the survey “it has sometimes been difficult to decide whether a given form of speech is to be looked upon as an independent language or as a dialect of some other definite form of speech.” This is what Grierson observed and this dichotomy continues.


So the two words, language and dialects, are in this respect are like cup and a mug. So we know a cup and a mug, and a mountain and a hill. Now we all understand the difference between a cup and a mug. We all perceptually understand the difference between cup and mug, of their size and purpose. But can we draw a definite line of distinction between cup and a mug? We cannot. We cannot draw a distinct line or definite line between these two items, cup and mug.

So find a couch for that matter. So it is perceptual how you look at it. So if we increase the height of a cup to a level which is very less than the height of the mug or vice versa, we lower down the height of the mug to the extent that it does not become a cup. But to what extent? It is difficult to point out and understand.

So we go by extremes, like if I have a cup, in its ideal extreme form. If I have a mug, it is in its ideal extreme form. But the similarity between the two is very close and the area between the two extremes is so blurred. You cannot draw a distinct line. This applies to language and dialect as well.

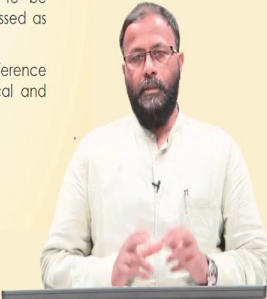
So we cannot point out what are the particular qualities that a particular variety becomes a dialect. So what are the particular characteristics of a dialect or what are the things that does not qualify a particular variety to be a language is difficult to decide. So it is more of the attitude of the speakers. And it is more of the perception or the prestige value or sociolinguistic attitude that we have towards the particular variety.

(Refer Slide Time: 07:37)



### Some 'acceptable' notions about Dialects

- A dialect is a regional, temporal or social variety within a single language.
- It is the product of individual's geographical and class origin.
- It differs in grammar, pronunciation and vocabulary from the standard language, which is in itself a socially favored dialect.
- So a dialect is a variation of language sufficiently different to be considered a separate entity, but not different enough to be classed as separate language.
- There is no clear qualitative linguistic measure to indicate where difference of dialect becomes difference of language. The issue is political and social, not linguistic.



However, there are some acceptable and I have put it in quotes. So we roughly understand a dialect in terms of something like; a dialect is a regional, temporal or social variety within a single language. We can take this as an acceptable notion about dialect, something like it is the product of an individual's geographical and class origin.


So we can take that as a broad acceptable notion. Then it differs in grammar, pronunciation, and vocabulary from the standard language which is in itself a socially favoured language. So a language x which the society or the speech community considers to be a language, this variety will have syntactic, morphological, and phonological variation.

So a dialect is a variation of language sufficiently different to be considered a separate entity, but not different enough to be classed as a separate language. So you have some overlaps in terms of syntactic structures in terms of morphological similarities, process, and similarities in terms of pronunciation and accent.


So it is not that they are not different enough to be regarded as a separate language or not even identical enough to be regarded as the same. So this grey area continues in understanding these two different concepts: dialect and language. As linguists, we do not differentiate because all dialects qualify equally to be called and classified as a language. So what we call dialects is something which represents the bias towards a particular variety. So there are no qualitative

linguistic measures to indicate whether the difference of a dialect becomes a difference of language. The issue is political and social and of course not linguistic.

(Refer Slide Time: 10:04)



- Everybody speaks a dialect, which is not seen as some kind of derivation from the norm of standard language. There is no linguistic justification for saying that one dialect is better than another rather it is a social judgment that leads people to say that a particular dialect is the correct one.
- Dialect is not an important type of language variation for teaching. All language teaching, however, at least implies an assumption about the best dialect to teach.
- Dialects are dialects not because of linguistic reasons but because of the political and cultural reasons. It is customary to describe them as varieties of a language according to users.



Now we will come to the linguistic ecology of India where this question is, and in fact in a multilingual society like India, this question is very dynamic and difficult to answer. Because we have different attitudes and perceptions about different varieties and the issue of language is very emotional. We identify it as emotional so it is a very vigorous marker of our identity. The speakers are very sensitive towards it.

So it is a very sensitive issue and particularly in multilingual societies like India, we will come to the Indian scenario in a while. So everybody speaks a dialect language is a very common notion. We speak a particular variety of it, the language nobody speaks, we speak a variety of dialects, which is not seen as some kind of deviation from the norms of standard language. Take the example of Hindi for that matter.

Now which Hindi do we call the real Hindi language or standard Hindi language? Hindi spoken in Lucknow or Hindi spoken in, let us say Delhi? Hindi spoken in Patna? Hindi spoken in Guwahati? Hindi spoken in Chennai or Hindi spoken in Bangalore? Hindi spoken in Mumbai? Which Hindi is a real Hindi? That is a debatable question.

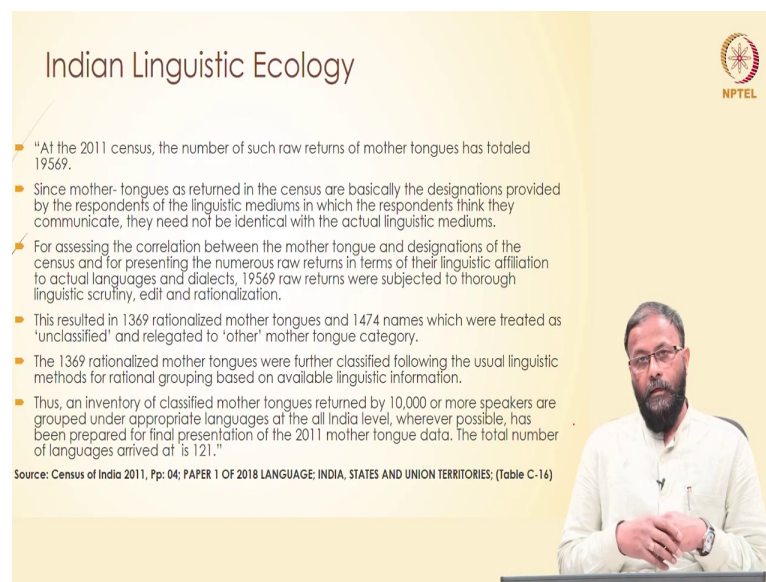
So which Hindi you call a standard Hindi and what about these other varieties? What about the Hindi of Mumbai for that matter? What is the Hindi of Guwahati for that matter? What do you

call it? Is it Hindi and if (11:55) it is Hindi, we cannot find a lot of lexical differences, pronunciation differences, and in extreme cases, like Deccani Hindi and Urdu, you find a lot of syntactic differences. How do you classify them? That is the question. Do you believe that the same Gujarati is spoken all over Gujarat or the same Tamil is spoken all over Tamil Nadu? A debatable question.

So there is no linguistic justification for saying that one dialect is better than another, and it is all about social judgement. It is all about our judgmental values that were assigned to a particular variety. As a linguist, we do not distinguish and we do not put them in a hierarchy, which language is superior and which language is inferior. All languages are equal, equally competent enough to be standard languages.

It is a deliberate attempt to assign institutional support to a particular variety or a particular variant or a dialect which emerges as standard language. Otherwise there are no linguistic reasons why a particular variety should not be treated as proper language. Dialects are dialects because linguistic dialects are not dialects because of linguistic regions, but because of political and cultural reasons. It is customary to describe them as varieties of a language according to the uses.


(Refer Slide Time: 13:44)



**Indian Linguistic Ecology**

- "At the 2011 census, the number of such raw returns of mother tongues has totaled 19569.
- Since mother-tongues as returned in the census are basically the designations provided by the respondents of the linguistic mediums in which the respondents think they communicate, they need not be identical with the actual linguistic mediums.
- For assessing the correlation between the mother tongue and designations of the census and for presenting the numerous raw returns in terms of their linguistic affiliation to actual languages and dialects, 19569 raw returns were subjected to thorough linguistic scrutiny, edit and rationalization.
- This resulted in 1369 rationalized mother tongues and 1474 names which were treated as 'unclassified' and relegated to 'other' mother tongue category.
- The 1369 rationalized mother tongues were further classified following the usual linguistic methods for rational grouping based on available linguistic information.
- Thus, an inventory of classified mother tongues returned by 10,000 or more speakers are grouped under appropriate languages at the all India level, wherever possible, has been prepared for final presentation of the 2011 mother tongue data. The total number of languages arrived at is 121."

Source: Census of India 2011, Pp. 04; PAPER 1 OF 2018 LANGUAGE, INDIA, STATES AND UNION TERRITORIES; (Table C-16)



Now this is what I was talking about. I just wanted to quote from the Census of India 2011 and the data will create another round of debate. Because if you look at the returns, the number of

returns as different languages by the respondents, the number is almost unimaginable. You cannot imagine the number, you cannot guess the number. Can you guess how many varieties people have responded and they all claim it to be their mother tongue, they all claim it to be a separate language (14:26).

So I will quote from the Census of India 2011. The source is the Census of India 2011, page 4, paper 1 of 2018, section language, India states and union territories table C-16. I quote from there, what do they say, I quote from the 2011 Census, then the number of such raw returns of mother tongues has totalled to 19569. Can you believe 19569? A rough estimate says that we have 6000 or approximately 6000 languages in the whole world.

And look at the Census. The respondents return 19569 varieties as their mother tongue. Amazing number that talks about the linguistic identity issues of these speakers and how they look at their own variety and how they identify with that. I quote again, since mother tongues are written in the Census are basically the designations provided by the respondents of the linguistic mediums in which the respondents think they communicate. They were not identical with the actual linguistic mediums.

Now this actual linguistic medium and perceived linguistic medium is debatable. I continue quoting again to assess the correlation between the mother tongues and designations of the census. And for presenting the numerous raw returns in terms of their linguistic affiliations to actual languages and dialects, 19569 raw returns were subjected to thorough linguistic scrutiny, edit and rationalisation.

This resulted in 1369 rationalised mother tongues and 1474 names which were treated as unclassified and relegated to other mother tongue categories. The 1369 rationalised mother tongues were further classified following the usual linguistic methods for rational groupings based on available linguistic information.

Thus an inventory of classified mother tongues returned by 10000 or more speakers are grouped under appropriate languages at all India level, wherever possible, has been prepared for final presentations of 2011 mother tongues data, the total number of languages arrived at is 121.

So let us see the data from 19569 to the final tally of 121 languages. What do they do with the rest? How do we treat them? What do you understand by that and why you have so many returns



beyond your imagination? The government of India in Census 2011 claims 121 languages in the country spoken over all states Indian territories. And here you have a raw data that gives you 19569 languages as mother tongue varieties as mother tongue. Now do we call them languages? So do we go by the users and how they perceive it and the value they assign? So this is a debatable question.

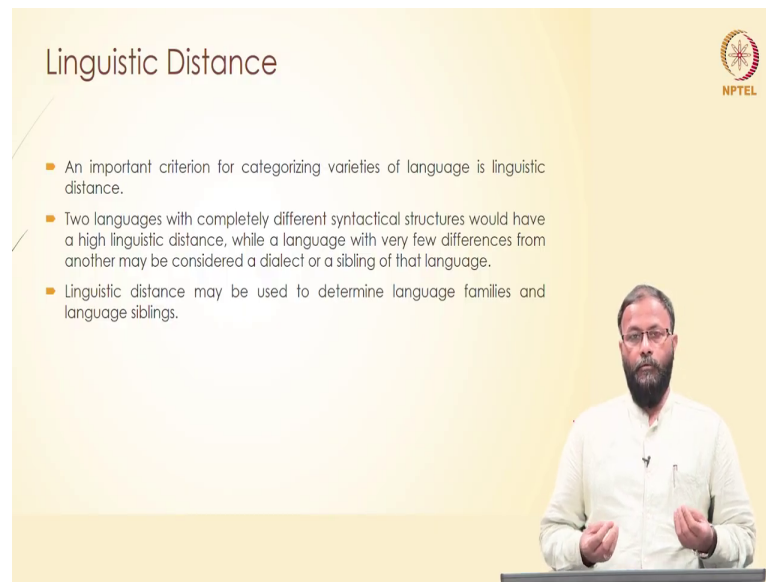
(Refer Slide Time: 18:49)



Now as I told you, deciding upon dialect and language, whether a particular variety  $x$  is a language or a dialect, we have certain frames of references. So we do not give a definite definition and category assigned to it. But we have certain frames of references. For example, linguistic distance between these two varieties, then mutual intelligibility. Another frame of reference and socio-political factors. So these are three main frames of references, where we can decide whether  $x$  is a dialect or a language. So let us go to each of them one after the other.



(Refer Slide Time: 19:45)



The slide is titled "Linguistic Distance" and features the NPTEL logo in the top right corner. It contains three bullet points:

- An important criterion for categorizing varieties of language is linguistic distance.
- Two languages with completely different syntactical structures would have a high linguistic distance, while a language with very few differences from another may be considered a dialect or a sibling of that language.
- Linguistic distance may be used to determine language families and language siblings.

In the bottom right corner of the slide, there is a small video inset showing a man with a beard and glasses, wearing a white shirt, speaking at a podium.

So when we say linguistic distance, what does it mean? An important criterion for categorization of varieties of languages is determining the linguistic distance between two varieties. So two languages with completely different syntactic structures would have a high linguistic distance. While a language with very few differences from another may be considered a dialect or a sibling of that language. What do we call Hindi and Urdu? It is a political question, but linguistically, we hardly find any syntactic difference.

So if you look at the linguistic distance between Hindi and Urdu, they are the two varieties of the same language. Are they two dialects, what do you call them? This is a very sensitive question. Because if you go by linguistic distance, we hardly find any linguistic distance between these two languages. We know them as two different languages not because we find linguistic distance. We know them as two separate languages because of socio, cultural, political and religious reasons.


So it is all about the frame of reference that you take to classify and distinguish a language and a dialect. But so it is that these two varieties of siblings, Hindi and Urdu are siblings, in that case, in that frame of reference, linguistic distance may be used to determine language, families and language siblings.

So can we consider Hindi and Urdu to be siblings because they do not have linguistic distance? Some differences that people refer to is the way we used scripts. Hindi uses Devanagari script,

Urdu uses Nastaliq Perso-Arabic script. Hindi in certain cases draws heavily from Sanskrit origin and Urdu draws heavily from Perso-Arabic origin. But still, if you look at the structures, if you look at the grammatical patterns, if you look at the agreement patterns, we find no difference except for the orthographic visual representation. We hardly find any difference or possibly, you may find a difference.


Let us say in case of genitive phrases, like Sher e Punjab and in Hindi we say Punjab Da Sher, and a few sounds that are different, otherwise, the linguistic distance between Hindi and Urdu remains very minimal. So this is one frame of reference that if two varieties have high linguistic distance, in terms of syntactic differences, morphological differences, phonological differences, then they are two separate languages. If it is low, there may be a variety of any x language or dialect.

(Refer Slide Time: 23:11)



### Mutual Intelligibility

- Mutual intelligibility is the one criterion, which is purely linguistic.
- Two varieties are said to be dialects of the same language if being a speaker of one variety confers sufficient knowledge to understand and be understood by a speaker of the other; otherwise, they are said to be different languages.
- Further problems with this criterion are that mutual intelligibility occurs in varying degrees, and that it is difficult to distinguish from prior familiarity with the other variety.



Then another frame of reference that we refer to in understanding the concept called dialect is mutual intelligibility. So to what extent speakers of two different varieties are able to understand each other? To what extent do they share the linguistic norms? To what extent are they able to communicate effortlessly? That is another frame of reference in determining whether it is a dialect.

So mutual intelligibility is one criterion, which is purely linguistic. And two varieties are said to be dialects of the same language if being a speaker of one variety confers sufficient knowledge to


understand and be understood by speakers of the other, otherwise they are said to be different languages.

So if they are not able to understand each other easily without any effort and if the two speakers or two varieties are able to communicate effectively without certain reservations, they understand each other they can predict. So this mutual intelligibility leads to the conclusion that these two varieties are dialects of the same language. So this is another frame of reference.

So further problems with this criterion are that mutual intelligibility occurs in varying degrees. Now the question is, to what extent I am able to understand the other variety or to what extent the speaker of another variety is able to understand me. It is a degree. It is a varying degree. So what is that exact degree of intelligibility which determines it to be a dialect, again debatable. And it is difficult to distinguish between the prior family with the other variety.

So this question remains confusing and we cannot have a clear-cut answer, but broadly if they are able to understand we say that these two varieties are dialects. If they are not able to understand each other, we say that there are two separate languages, but there are certain cases where it may be possible that these two varieties, because of some historical reasons, may be separated for a longer period of time. So even if they are derived from the same language, they may not be intelligible. So as I mentioned, such frames of references are problematic and they do not give us a very clear-cut idea, however, they are useful to determine a broad dialect and a language.

(Refer Slide Time: 26:08)



The slide is titled "Socio-political Factors" and features the NPTEL logo in the top right corner. It contains a list of bullet points discussing the relationship between dialect, sociolect, and language status. A video inset in the bottom right shows a man with a beard and glasses, wearing a white shirt, sitting at a desk with his hands clasped.

### Socio-political Factors

- In many societies, however, a particular dialect, often the sociolect of the elite class, comes to be identified as the "standard" or "proper" version of a language by those seeking to make a social distinction and is contrasted with other varieties.
- As a result of this, in some contexts, the term "dialect" refers specifically to varieties with low social status. In this secondary sense of "dialect", language varieties are often called *dialects* rather than *languages*:
  - if they have no standard/codified form,
  - if they are rarely or never used in writing (outside reported speech),
  - if the speakers of the given language do not have a state of their own,
  - if they lack prestige with respect to some other, often standardised, variety.
- The status of "language" is not solely determined by linguistic criteria, but it is also the result of a historical and political development.

Now what do you mean by socio-political factors? It is purely the attitude of these speakers that is the attitude of the speakers towards other languages and attitude of other languages speakers towards this particular variety. It is social, political, and has got exclusivity. So in many societies, a particular dialect, often the sociolect of the elite, comes to be identified as the standard or “proper version of a language.”

I keep on repeating the sentence why we put in quotes, because we cannot subscribe to these ideas, but there are certain conventional practices which exist. So it has something to do with the power, it has something to do with the status, it has something to do with the socio-political share of these speakers of that particular variety.

So the elite class of a society. So the variety of the elite class spoken by a particular elite class may be considered standard language. And there is no linguistic reason for that, it is purely political. So as a result of this, in some contexts, the term dialect refers specifically to varieties with low social status.

So this is again problematic because it is governed by our perceptions and power equations, status assigned to a particular variety and prestige assigned to a particular variety. Dialects are used as derogatory terms, so dialect may be a derogatory term for language. So it is relegated with low perceived value.

And there is another very widely held concept that dialects do not have written forms. So languages which are written forms are standard languages and dialects do not have written forms. Dialects are not taught in schools. So they are not a good variety to teach in schools. Dialects do not have written literature.

Now we need to understand a particular variety which does not have a written system or we do not have a sizable body of literature. We do not have institutional presence of these varieties, in our schools, in our media, and in administrative offices. It has nothing to do with the linguistic properties of the variety. It is all about favouring a particular dialect, which is being used, taught, and practised by the elite and powerful. It is taught in the school grammar that this whole process of rationalisation which elevates that particular dialect to be standard or “proper language.”

But while they hold notions about dialect, something like, if they have no standard or codified form, again, the standard and codified form, who makes it a standard is a deliberate attempt by the system or institutions, which favour a particular dialect, standardise it and use it for various purposes.

They are codified, dictionaries are created, grammars are created, books are created, encyclopaedia is created. So they get institutional support and a particular dialect is elevated as language. Second, such a notion is if they are rarely or never used in writing, which is another notion about dialects.

If the speakers of a given language do not have a state of their own. It refers to political power, if the lack proceeds with respect to some other often standardised variety. So a sense of inferiority is assigned, perception of inferiority is assigned to a particular variety, that is a dialect and the status of language is not solidly determined by linguistic criteria, but it is also the result of historical and political development and this is true to all such languages.

For example, the constitution of India has listed 22 languages in its schedule eight which are considered as official languages of India. So Hindi is an official language of India. English is the associate official language of India and 22 languages listed in schedule eight are considered as official languages of India. What about the 121 languages that the government of India in census 2011 claims. So 99 languages are languages which are not scheduled languages, but again go back to the approximately 1400 varieties which are classified as others, what are these?

So the question of language and dialect is always debatable, and it is the linguistic attitude that determines and it is the frame of reference that determines whether it is a dialect or a language. But in linguistics, in sociolinguistics or for us, as an enlightened informed people, there is no difference between a language and a dialect.

Language is a favourite dialect. Language is a socially, politically and institutionally supported variety. We all speak a dialect. When I am speaking English, I am speaking a dialect of English. But because of socio-political reasons, we consider a particular dialect to be the standard variety and which has written form or it has dictionaries, it has encyclopaedia, it is used in media, it is used in administration, and this becomes standard language.

Otherwise all so called dialects or all varieties equally qualified to be a language. It is just an accident that a particular dialect is favoured and emerges as language. The rest of them are called dialects of the same language. So I hope the idea of dialect is clear now and in sociolinguistics, we prefer to use the term variety, speech variety. In the Labovian tradition, we will talk about in some other lecture, where he has used a variety made.

So dialect somehow has some negative connotation and a derogatory term assigned to a particular speech variety, which is considered a low prestige variety. But that is not an appropriate explanation and there is no linguistic merit in such a belief and in such a linguistic attitude. There is absolutely no linguistic grounding. But this is how we perceive and understand dialects.

Based on the geographical location of a particular variety, we refer to terms like regional dialects. So regional dialects are spoken by people of a particular geographical area, within a speech community. For example, Cockney in London, that is a regional dialect and so many others in Indian context.

Social act refers to a variety spoken by members of a particular group or a stratum of speech community. While a variety of languages used at particular stages in its historical development, may be termed as temporal dialects, such as Prakrit and Pali in ancient India, where Sanskrit was the standard language.

So with these closing notes you can understand dialect as a speech variety which is classified because of socio political factors and mind you not, linguistic factors as a low prestige variety

spoken by a particular class in a particular language in a particular land or geographical location or possibly in a particular point of history of that speech community.

So we will talk about the speech varieties and variations more in next videos and we will see why sociolinguists use the term variety not a dialect because dialect is a controversial terminology ascribed to assign to a particular variety. So thank you very much and we will meet in the next class and continue our debate and discussion.