Qualitative Research Methods
Assistant Prof. Aradhna Malik
Vinod Gupta School of Management
Indian Institute of Technology - Kharagpur

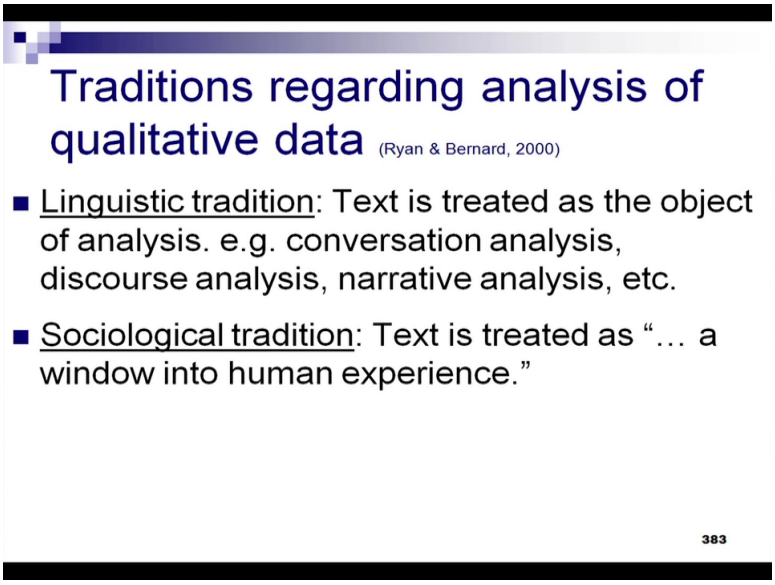
Lecture 36
Data Management and Analysis Methods

Welcome back to the NOC course titled qualitative research methods, my name is Aradhna Malik and I am helping you with this course and we were talking about different ways of collecting and analyzing qualitative data. Now what we will talk about in this class today is, we will summarize whatever we have done till now very, very briefly and we will move on to some more ways of collecting and analyzing textual data, which we briefly mentioned in the last class.

(Refer Slide Time: 00:53)

So we will, the topic of today's classes' data management and analysis methods and this will carry our conversation forward okay and this is from a paper by Ryan and Bernard in the handbook of qualitative research methods, the book that we have been referring to right from the beginning okay.

(Refer Slide Time: 01:11)

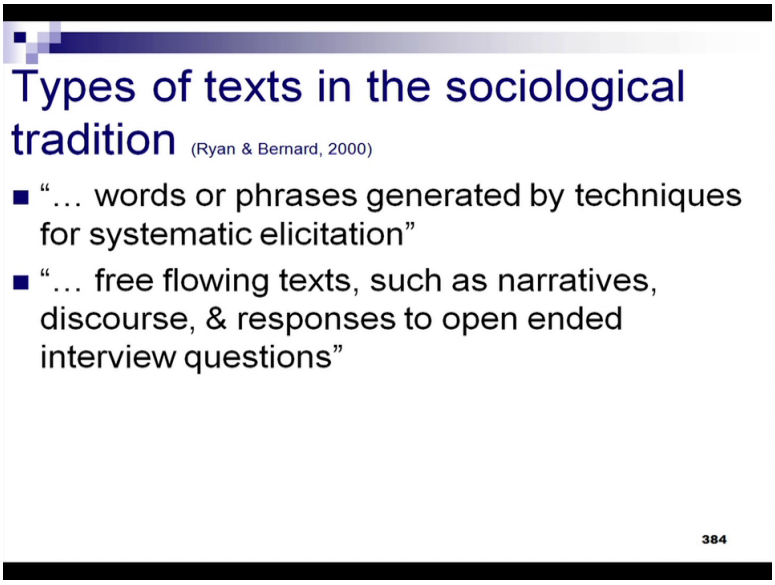


So then in this paper the authors discuss two traditions regarding the analysis of qualitative data and they say that qualitative data is analyzed in primarily two ways one is the linguistic tradition in which the text is treated as the object of analysis, for example the conversation, we have conversation analysis we call the, we talked about discourse analysis, narrative analysis etc.

So we have already discussed this as to how the text or the actual meet of a conversation is treated like the object of analysis, even in discourse analysis the meat of the conversation, the meat of the interactions recorded between or the interactions between two or more people that are recorded, the interactions are taken and then they are analyzed to see how they represent the social situation that they are coming from.

The second tradition here that will talk about is sociological tradition, where the text is treated as "a window into human experience." So this is something that we briefly mentioned in parts overtime but we have not really discussed this and we are going to discuss this much greater detail now. We have already discussed the linguistic tradition, so we will not discuss that anymore okay.

(Refer Slide Time: 02:41)



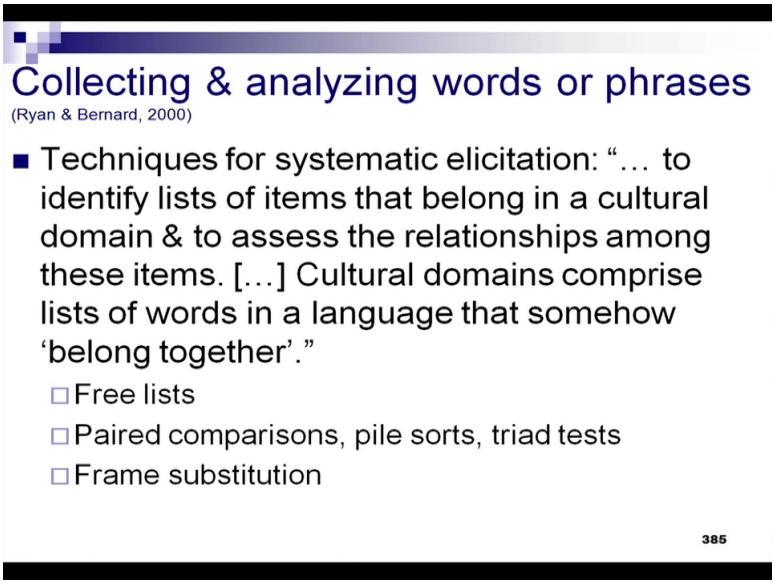
Now as far as sociological tradition is concerned we are going to discuss how text are analyzed there are two types of text in the sociological tradition, the first one is "words or phrases generated by techniques for systematic elicitation", words or phrases that are generated as a result of forcibly, elicitation means we are stimulating the situation to generate words or phrases, to generate representations of itself.

So they are generated by techniques for systematic elicitation, techniques that are used to systematically stimulate the generation of words and phrases, for example interviews, closed ended interviews, we ask a question and we know the realm in which the answer will come okay.

The other type of text here is the "free flowing text such as narratives, discourse and responses to open ended interview questions", where we don't know what will get, so there are two types of text one is expected category of texts, the other is this unexpected ocean or you know unexpected wealth of a text that comes to us in the form of narrative, in the form of responses to open ended questions, like what did you do?

Why did you do it? How did you do it? Okay or discourses the conversations rooted in a social situations that are influenced by an in turn influence the environment's at the occurrence, so they cannot be predicted okay.

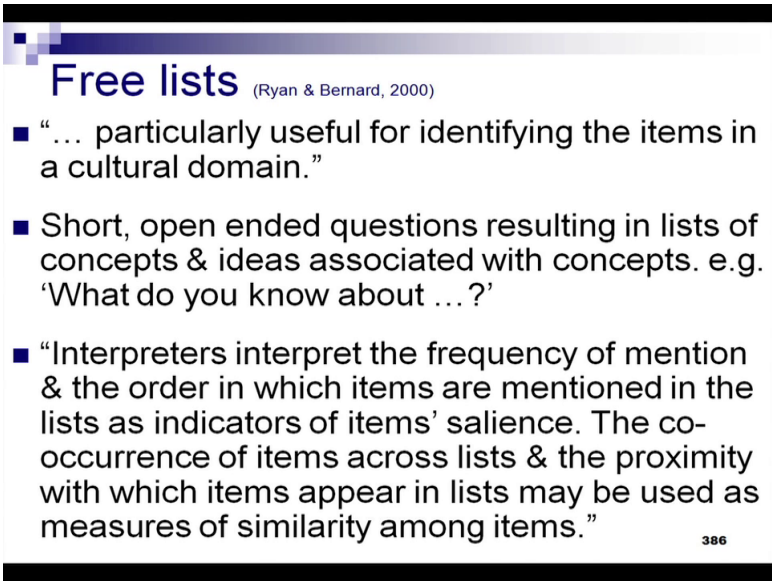
(Refer Slide Time: 04:36)



Collecting and analyzing words or phrases: how do we collect and analyze words or phrases, techniques for systematic elicitation involve, they are these techniques is "to identify list of items that belong in a cultural domain and to assess the relationships among these items. So we bring out or we generate list of items, list of concept that belong in a cultural domain that belong and to assess the relationships between these items.

Now cultural domains comprise list of words in a language that somehow belong together." They are connected to each other by way of the social milieu that their part of. The types of techniques are, there are three broad categories of techniques that we use here, one is called free list, the other is paired comparisons, pile sorts, and triad tests and the third one is frame substitution, we will discuss each of these in detail now.

(Refer Slide Time: 05:46)



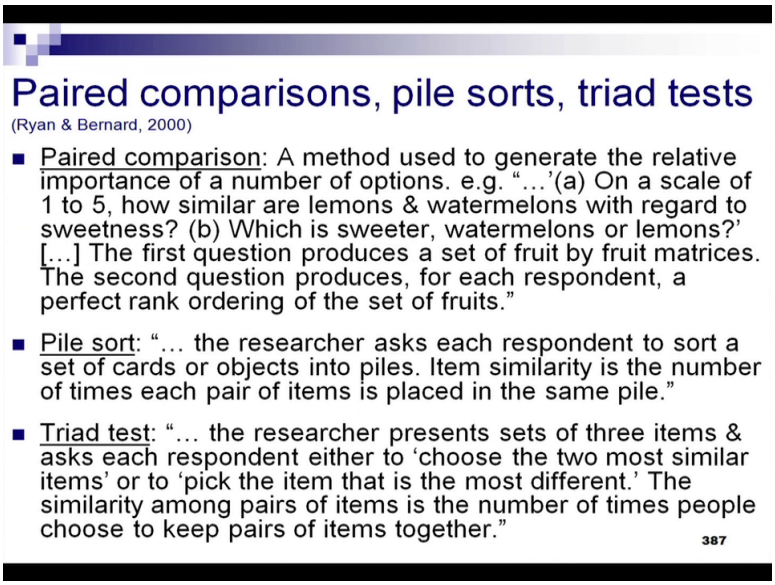
The three lists are “particularly useful for identifying the other items in a cultural domain”, and they consist of short, open ended questions resulting in lists of concept and ideas associated with concepts, for example what do you know about life at IIT, Kharagpur? What do you know about weather in West Bengal? What do you know about the food habits of Indian students living abroad? What do you know about travel in New Delhi or something like that?

So what do you know about or what did you do on this trip? So there are open ended questions, what do you know about is? it can be anything we don't have categories in mind, so these are short questions that result in long list that can be even remotely associated with the concept understudy.

Interpreters interpret the frequency of mention and the order in which items are mentioned in the list as indicators of items salience, the salience of items, and the importance of items. The co-occurrence of items across list and the proximity with which items appear in list may be used as measures of similarity among items.”

So these are free list and the number of times words come together, the number of times they are generated, you know the proximity that they are generated, you know sorry, the way they are generated in relation to other words, their proximity with other words within the same category and the frequency of the similarities and proximity will give us some idea of the kind of context that we are talking about, so this is our free lists, okay.

(Refer Slide Time: 07:41)



The next technique that will talk about is called paired comparisons, pile sorts and triad tests. Now I put all three on one Slide and I'm sorry for the clutter, but it's helpful if you see all three together. Paired comparison: is a method used to generate the relative importance of a number of options, for example and again taking the example from book itself, on a scale of 1 to 10, how similar are lemons and watermelons with regard to sweetness? So you grade them.

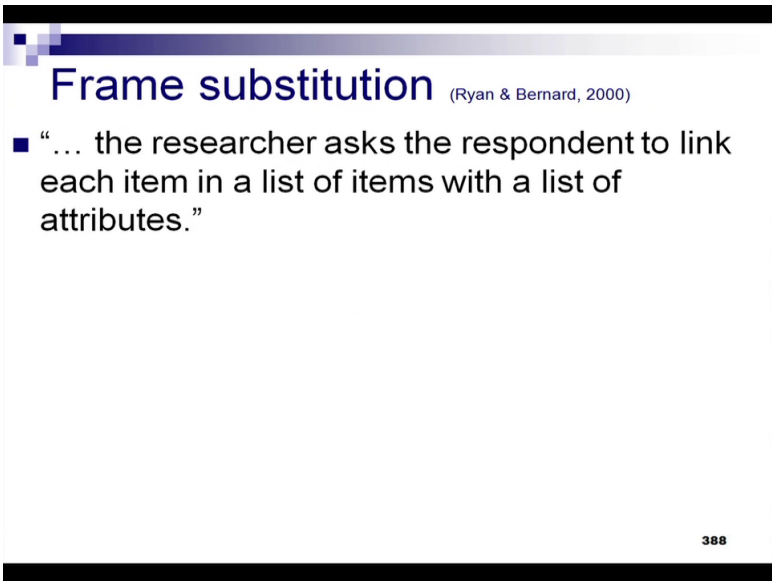
Then the second question you'll ask is which is sweeter, watermelons or lemons? The second question, the first question produces a set of fruit by fruit matrices and the second question produces, for each respondent a perfect rank ordering of the set of fruits." Which is sweeter? Which is next? Which is next? Etc, and how do they relate to each other? This technique is called as the paired comparison technique.

The next technique is pile sore technique in which "the researcher asked each responded to sort a set of cards or objects into piles. Items similarity is a number of times each pair of items is placed in the same pile." So and again you know these cards are made very carefully and then we sort them out, we ask the responded to sort them into different piles, so this is another technique

The third technique is the triad test in which "the researcher presents sets of three items and asks each responded either to choose the two most similar items or to pick the item that is most different. And the similarity among pairs of items is the number of times people choose to keep pairs of items together."

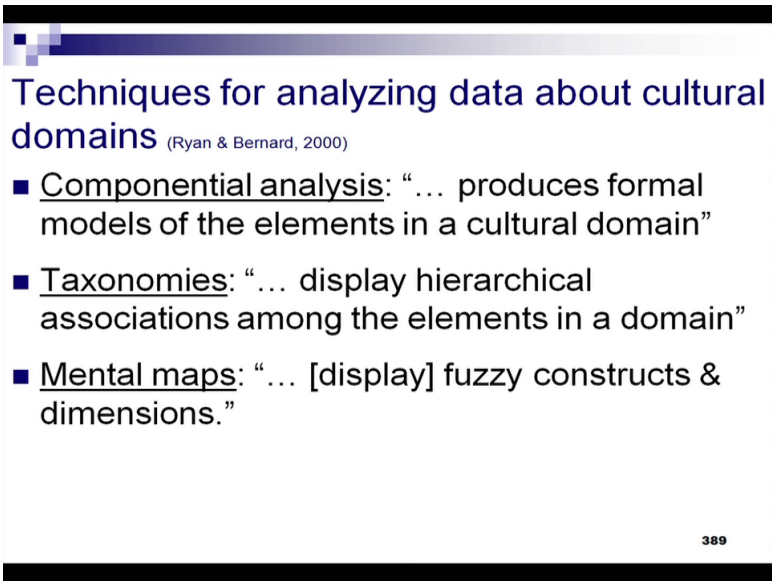
So you give them three things and you say which two seem most similar to each other or which one of these is different and the number of times people put items that they feel or similar together in the same pile is called the triad test and that gives us some idea of the way people categorize the concepts that they are faced with okay.

(Refer Slide Time: 09:52)



Frame substitution is another technique in which "the researcher asked the respondent to link each item in a list of items with a list of attribute." So there is cross connection you find out which items is linked to which attribute, you give people a list of items on one side and list of attributes on the other and these two are connected, okay.

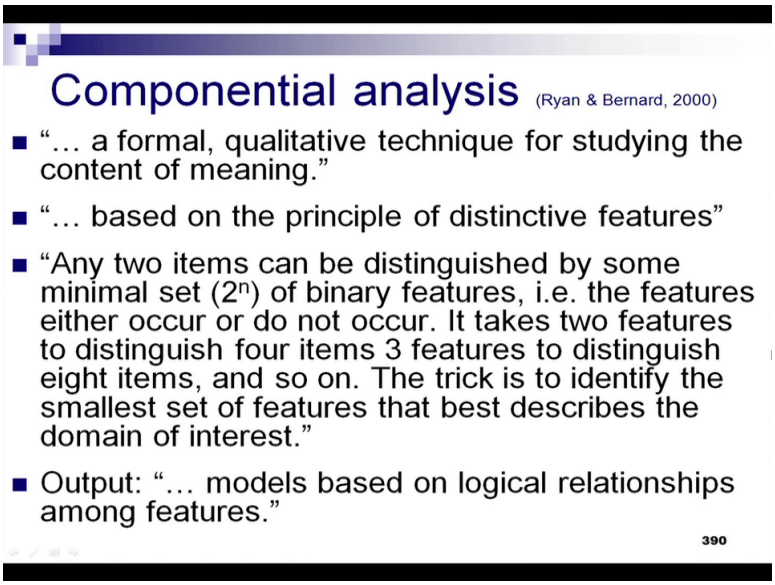
(Refer Slide Time: 10:18)



Now we come to the techniques for analyzing data about cultural domains, we talk cultural domains, so we use some techniques for analyzing data about cultural domains, the three techniques that we will discuss now are called component analysis, taxonomies and mental maps. Component analysis "produces formal models of the elements in cultural domain", we will talk more about this in a minute.

Taxonomies "display hierarchical associations among the elements in a domain" so you categorizing and then you organize them, arrange them, in a hierarchical manner and mental maps "display fuzzy constructs and dimensions."

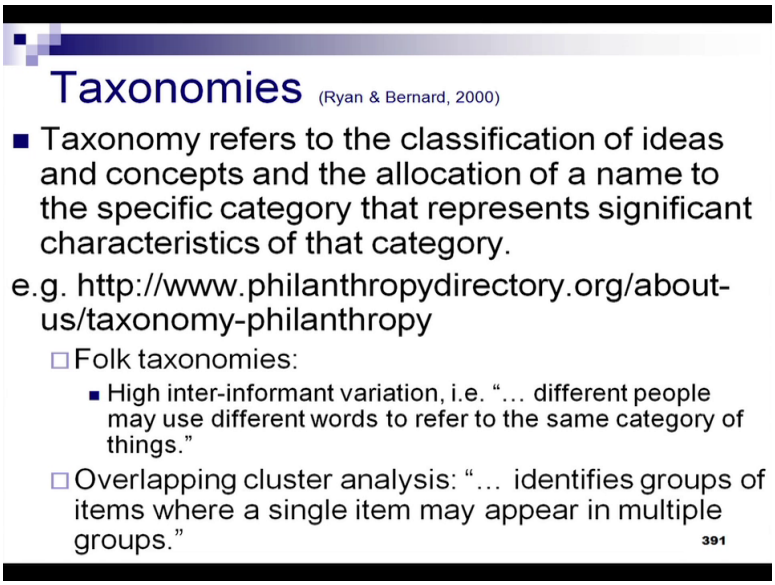
(Refer Slide Time: 11:04)



Component analysis is "a formal, qualitative techniques for studying the content of meaning." And its "based on the principle of distinctive features", "any two items can be distinguished by some minimal set i.e. 2 to the power n of binary features, i.e. the features either occur or do not occur. So it takes two features to distinguish three items.

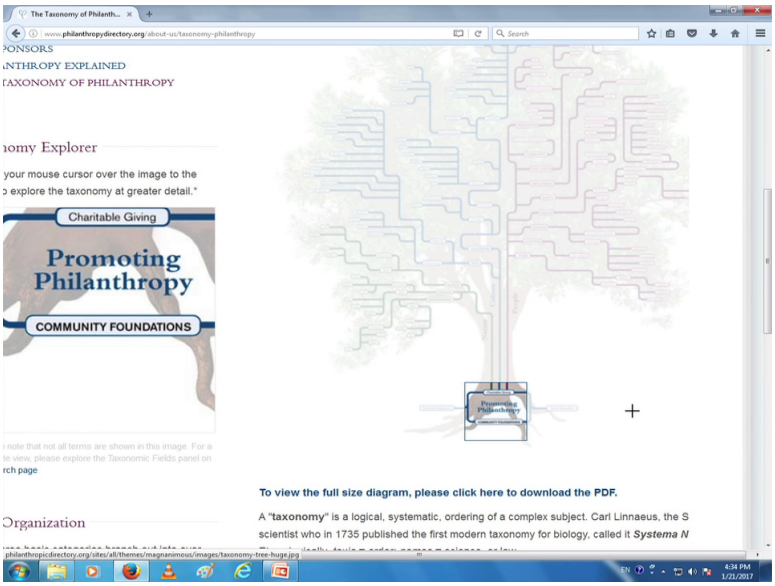
Three features so the 'n' represents the number of features, three features to distinguish eight items, three features and so on. And the trick is to identify the smallest set of features that best describes the domain of interest. And the output is "models based on logical relationships among features", that is called component analysis.

(Refer Slide Time: 11:51)



The next technique that we will talk about is taxonomies. Taxonomy refers to the classification of ideas and concepts and the allocation of a name to the specific category that represents the specific characteristic of that category and the hierarchical placement of items in different, of categories in relationship to each other. The categories that represent a concept in relationship to each other.

(Refer Slide Time: 12:57)



Let me show you a pictorial representation of what taxonomy looks like, let me just copy this and I will just take you to a website this is a very, very nice representation of what a taxonomy looks like. So this is a taxonomy of philanthropy that I found on the Internet, now this I will show you, this is very beautiful, is very nice looks like a tree, if you can see, so as you can see this is the base of the tree, the root says promoting philanthropy.

And over here we have technical assistance, you can see on the left side it says technical assistance, then we have charitable giving so the concept of your talking about is promoting philanthropy, here how do we promote philanthropy by giving people technical assistance, we also promote philanthropy by charitable, charitable giving.

Philanthropy is rooted in community foundations and here if we see this is donor education, so these are the three ways in which we promote philanthropy, now when we talk about charitable giving and we go up different categories in which people give one is nature, the other is culture and the third is people or human resources. When we talk about Nature then there are further classifications here plants, animals and environment.

Now each of these is divided into various things, so the environment is divided into wild and artificial, the wild environment is further divided into land, water, air. The artificial environment is further divided into land, water, air and so on. Similarly culture, arts, education, so these are the two broad divisions that this particular website has come out or is talking about and they're talking about specifically with regard to charitable giving.

So and promoting philanthropy, so they say that you know when we talk about culture, we promote philanthropy through two broadways and that is arts and education. And arts we have performing arts and we have material arts, and then in performing arts, we multimedia, we have theatre, we have dance, we have music, so then we have the multimedia again an artificial form of art. Theatre, Dance, music, involve human beings, similarly material arts involve here.

There are so many different categories here, general, mixed media, painting, architecture, sculpture, culture, literature, etc, etc., all of these things are given here, so this is called taxonomy. We come up with categories and these categories are then organized in the form of a hierarchy, so its a very beautiful way of doing it okay.

Now we also have something called as folk taxonomies: folk taxonomy is referred to taxonomies regarding or taxonomy is related to folk life or rooted in the folk environment. The problem with this taxonomy is that there is high inter-informant variation i.e. “different people may use different words to refer to the same category of things.” When we talk about human beings, when we talk about the natural world, when we talk about the social

environment, we are talking about something very, very dynamic, very, very complex so it may not, so it is not always possible, it's rarely possible with any other side, it is rarely possible to classify something totally into or bound something totally by a set of definitions and parameters.

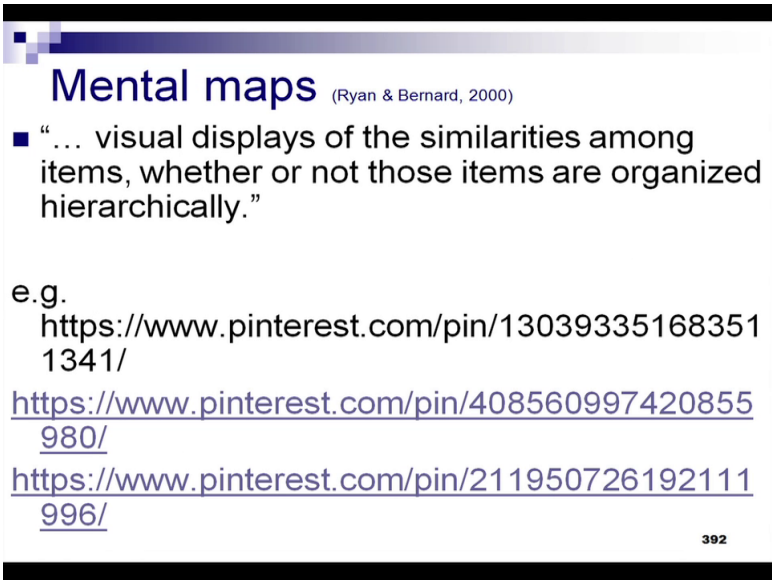
It is not possible to restrict to the categorization of one concept to one category with the result overlapping occurs and we are using language, after all we are using language to describe whatever it is that you want to talk about and language then limits in terms of what we can describe and how we categorize things, so many times how we define a particular constructor concept they may fall into another category also.

Different people are going to refer to the same concept the same experience in different ways, especially in folk life, where the language is limited, where the exposure is limited and where unless the researcher understands folk life, the way the researcher interprets the situation is going to be very, very different. So any unfamiliar, you can say it can apply to any unfamiliar situation, fair enough.

But when we talk about folk life we depend on the information we receive from I respondents and that will vary, so there is high inter informant variation and different people are likely to use different words to refer to the same category of things and to way to the way these categories of things are arranged in the hierarchy, so taxonomy is essentially involves hierarchical arrangement of these categories okay.

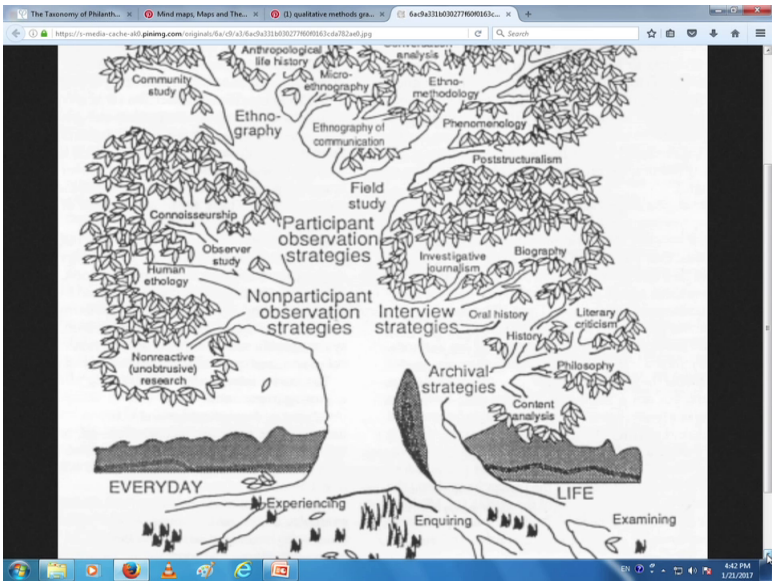
We also we use overlapping cluster analysis which “identifies groups of items where is single item may appear in multiple groups.” and then we use this cluster analysis to analyze items that appear different in more than one group you know, we analyze items that appear more than one group differently, so that is all about taxonomies.

(Refer Slide Time: 18:16)



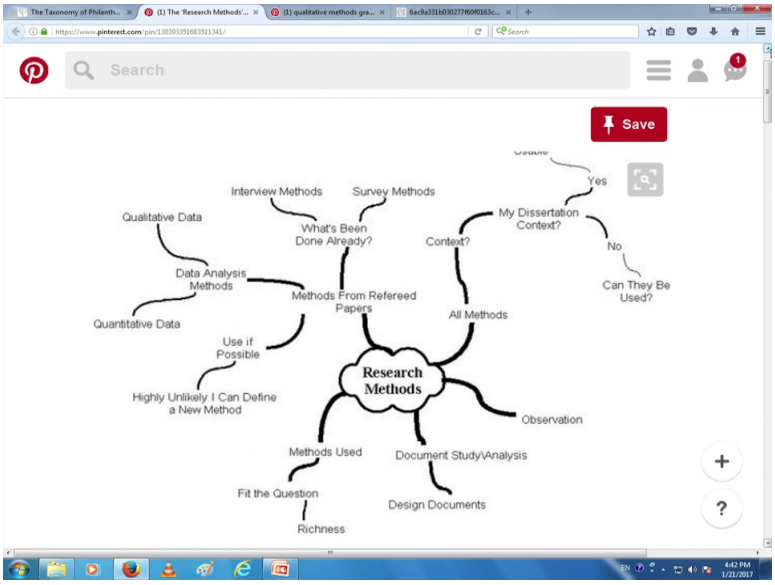
Next technique that we are going to talk about is mental maps. Mental maps are "visual displays of the similarities among items, whether or not those items are organized hierarchically." For example and show you something very nice here, this is something that relates to what you are doing as students.

(Refer Slide Time: 18:50)



So we have a mental map that is given here qualitative strategies in educational research. So this is one mental map.

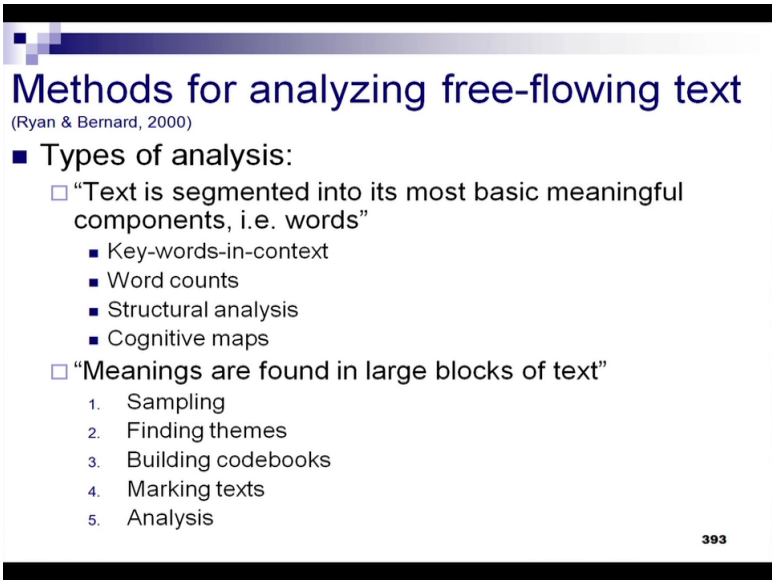
(Refer Slide Time: 18:53)



This is the other one, the research methods mental map is simpler and let me increase the size of the page, looks like we come up with ideas with the concepts that we can think of and we put them, on a piece of paper and then we see how they may be connected to each other and these categories, or these different concepts that are related may be mobile, you know we can take them and we can arrange them according to our, according to the way our understanding develops as we generate ideas, so this is a mental map, this is another technique.

There are some more that you can go through at your own convenience, you may need to create a login and password for the site, I found is quite useful okay.

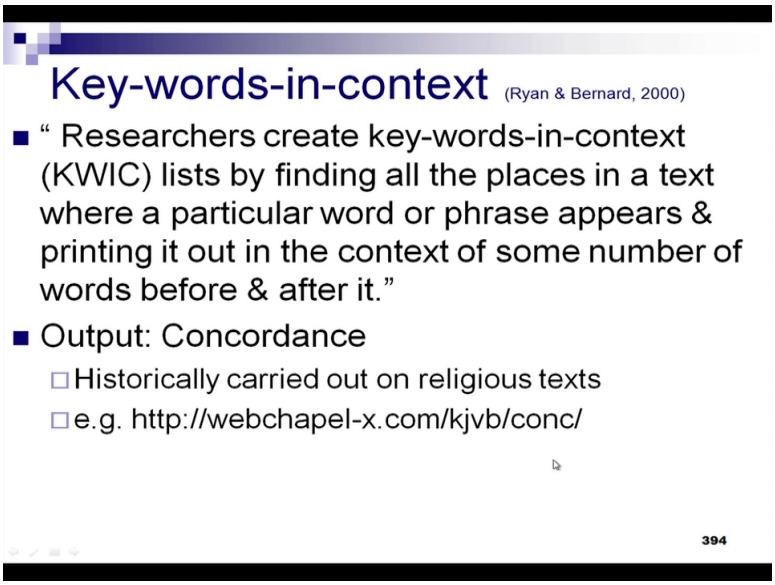
(Refer Slide Time: 19:50)



Now we come to methods for analyzing free flowing text, various types of analysis that are conducted the first one is, you know there are two categories, first : "text to segmented into

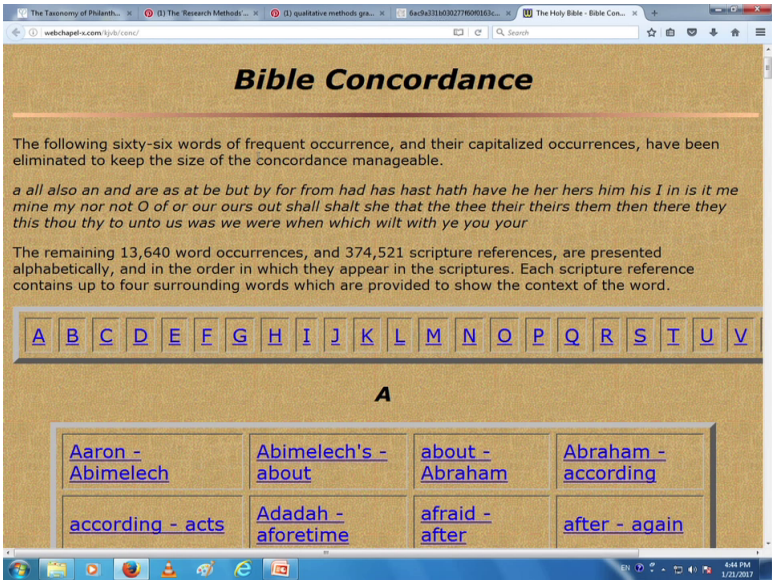
its most basic meaningful components i.e. words" and the second method is "meanings are found in large blocks of text" which means the text has to be coded and will do coding in the class. We will discuss the first set of texts and first set of techniques okay.

(Refer Slide Time: 20:21)



Key-words-in-context: "researchers create key-words-in-context (KWIC) lists by finding all the places in a text where a particular word or phrase appears and printing it out in the context of some number of words before and after it." The output here is called concordance and this was historically carried out on religious texts and I found a very nice example of concordance on the Holy Bible I will show it to you.

(Refer Slide Time: 20:50)

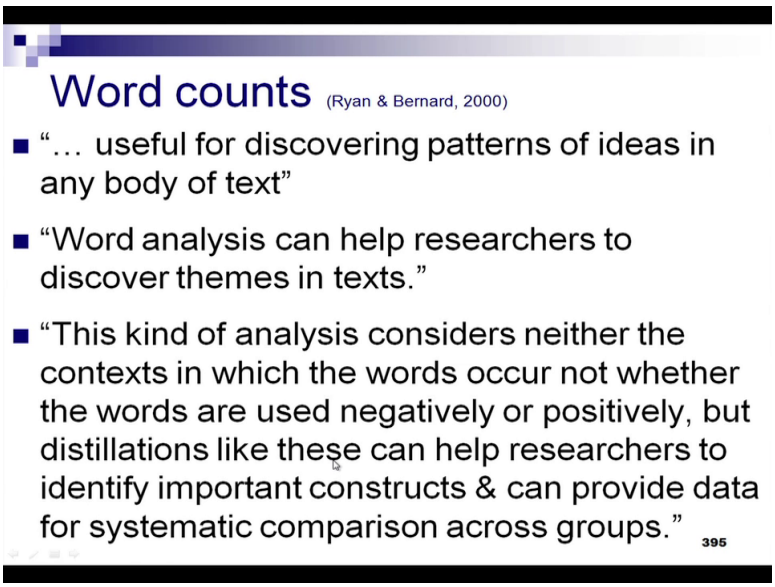


So this is what a output looks like, Bible concordance, all in blue that's a big problem, the Holy Bible concordance and they have come up with these words that appear in context here,

The following sixty six words of frequent occurrence and their capitalized occurrences have been eliminated to keep the size of a concordance manageable. So these words have been put in an alphabetical order and when you click on one, so the references come up with it or the contacts is highlighted with the word okay.

So this is called a concordance. here Aaron to Abimelech, I am sorry I hope I'm pronouncing this right when he say Aaron, all of these words that come before and after it come along with it and this context is also mentioned here, if you see a full screen of what you are seeing, you will realize what I am talking about, you have the website address you can copy it and see it for yourself this is called it concordance, alright. And in short the most common abbreviation for this type of technique is called is KWIC key-words-in-context.

(Refer Slide Time: 22:52)



The next technique is word counts and it's simply a physical counting of words that appear in a particular text. The number of times the same word appears in a particular text, it's "useful for discovering patterns of ideas in any body of text." And it can help researchers to discover themes in text because you know how many times that word appears.

Now one thing, that one is to be cautious of here is, that "this kind of analysis considers neither the context in which the words occur or not whether the words are use negatively or positively, but distillations like these can help researchers to identify important constructs and can provide data for systematic comparison across groups."

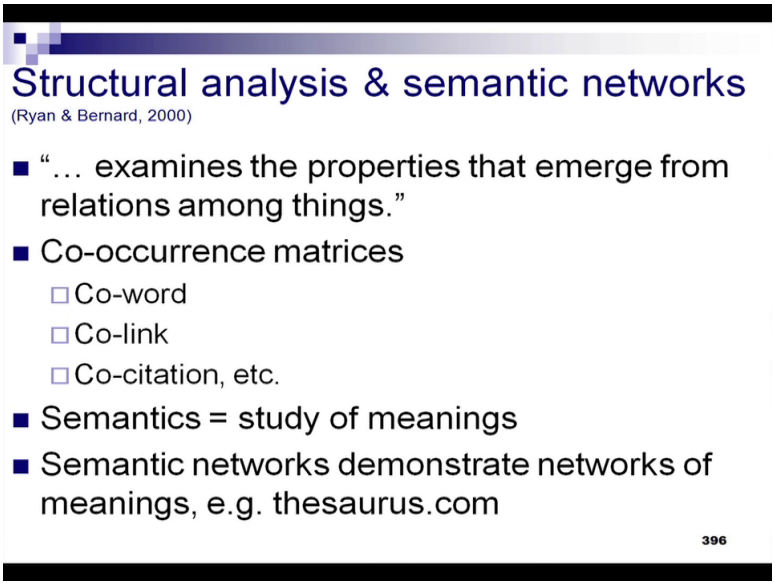
The contexts is not taken into account, the usage is not taken into account, whether the word is used positively or negatively is not taken into account, all that is taken into account is how many times a word appears within that text.

And that gives us an idea of the theme emerging in relation to these words now, one thing that every research on must be careful of is choosing the words that one wants to count, so we need to pick the right markers of that text, after all why are we looking for these words, these words will give us an idea of what the text is all about, so the world that appear most commonly.

For example, is you say ‘the’, ‘THE’ doesn't make sense, but more significant matters may be useful, so or the word travelled, if it's something a spiritual text or if it’s a philosophical text or if it’s a text documenting, say a research in tribal areas the word travelled may not be as important as the world Culture or living or family, maybe the word family would be more important, family or field or fruit maybe more important when we are doing research in tribal area.

So let's see how many times the word family occurs in this narrative and then we get an idea of why this word occurs and of course you know it’s an iterative process, we do it, we go back to the text and then we do it and we go back to text, but some themes emerge.

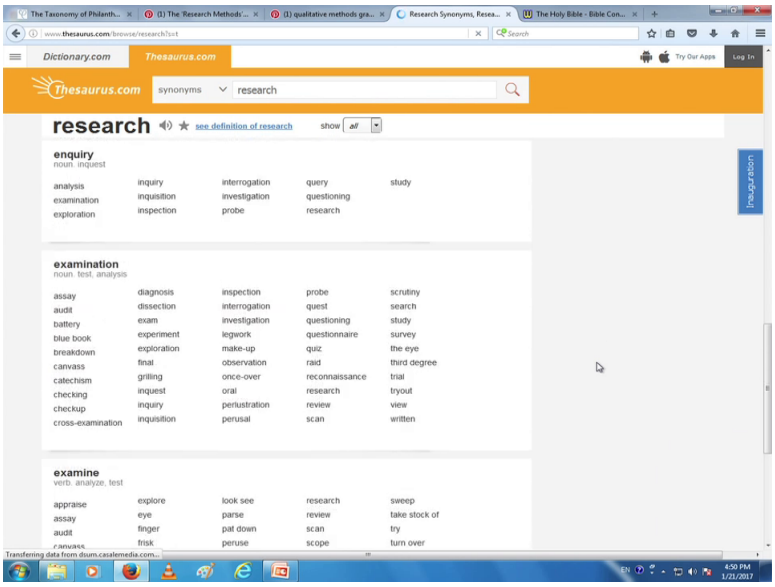
(Refer Slide Time: 25:17)



Structural analysis and semantic networks, so this technique “examines the properties that emerge from relations among things.” Co-occurrence matrices involve co-word, co-link, co-

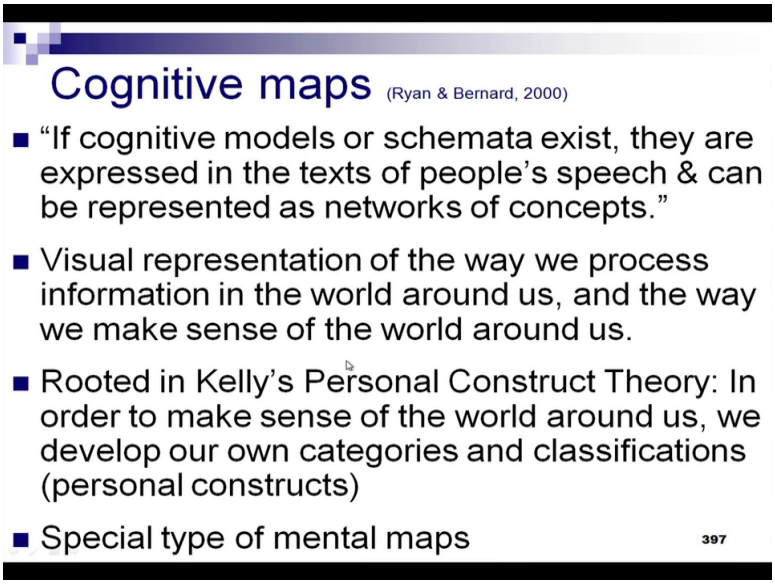
citation, etc., Semantics is the study of meaning. Semantic networks demonstrate networks of meanings, for example thesaurus.com. These are all public website, thesaurus.com.

(Refer Slide Time: 25:48)



This particular website gives you a network of meanings and visual thesaurus is another one, maybe I'm doing the wring, I will put a link to this. There something called as a visual thesaurus and it actually shows you how these words are connected, so these networks demonstrate networks of meanings and these networks can be used to see, how different meanings of different words related to each other.

(Refer Slide Time: 26:21)



The last topic in today's lecture is cognitive maps this is another technique and in this cognitive maps are special types of mental maps in which the way, it is a visual representation of the way we process information in the world around us, and the way we

make sense of worlds. And the premise here is that “if cognitive models of schemata exist, they are expressed in the text to people speech and can be represented as networks of concepts.”

They are rooted in Kelly’s Personal construct Theory: so which says that in order to make sense of the world around us, we develop our own categories and classification and these are special types of mental maps that show, that demonstrate visually what you know how our thinking process involved in relation to a particular concept or environment or phenomenon etc., so this is all we have time for in this lecture, we will discuss some more around, along the lines of, we will talk about coding, we will talk about analyzing chunks of textual information in the next class, thank you very much for listening.