Lecture – 23
Parsing Sentences - 2

Hello! And welcome to the course, Introduction to the, 'Psychology of Language'. I am Ark Varma, from IIT Kanpur. And, we are in the fifth week of the course, talking about sentence processing. So, in the last lecture, we talked about the garden path theory or sentence processing. We talked a little bit about, what is actually the premise of the garden path theory. Which is basically to try and understand, how do people choose one or more representations in a scenario, where more than one grammatical representation is possible. Also, those representations are all grammatically possible, plausible, acceptable. So, that is what we did, we talked about the three heuristics of lead closure, main assertion, and minimal attachment lead closure and main assertion.

Refer Slide Time: (01:06)

Constraint Based Models of Parsing

- Constraint based models constitute the most prominent alternatives to two-stage models (MacDonald et al., 1994).

- Critical differences:
  - Constraint based parsers are capable of pursuing multiple structural possibilities, simultaneously.
  - adopt a parallel distributed processing architecture.
  - Represent different aspects of the sentences, including their syntactic structures, as patterns of activations spread across large number of interconnected units, resembling neurons in the brain.

Today, we will talk also a little bit more about parsing. We will talk about, a different theory of parsing, a different group of models. So, basically we are going to talk about, constraint based models of parsing. And, the constraint based models of parsing are slightly different, with respect to the garden path theory, the two step theory that we have talked about earlier. Now, we will discuss a little bit about what these constraint based models of parsing are. And, then we kind of talked about, different constraints that operate inorder to kind of inform the parsing procedure. This is almost a group of theories, not really just one theory, but this is kind of a group of theories, each of which kind of plays. You know, it gives a little bit of importance to one of the factors that we will talk about. So, without much ado, let's talk about constraint based models of parsing.

Now, constraint based models of parsing, constitutes the most prominent alternatives to the two-stage models, the kind that lean phrase yet had propounded. And, there are some critical differences, with respect to these constraint based models of parsing, versus the other two-stage model serial models of parsing that were given earlier. Some of these differences are, say for example, the constraint based passes are capable of pursuing multiple structural possibilities simultaneously.

If you remember I was telling you in the last class. That, in the garden path theory or lean phrases two-stage parsing model, each structure can, so, the model is basically being able to evaluate only come up with, only one structure at a time. So, it comes up with one structure evaluates, say for example, the structure with respect to, whether it is thematically integratable or not. And, if it finds it integratable semantically plausible, it accepts it. If not it goes back to the drawing board, creates another structure, checks it again goes back to the metal board, comes again and does it iteratively. One of the major in, one of the most important difference is, therefore, between constraint based models and garden path model is

that these constraint based parsers are, able to pursue multiple structural possibilities one at the same time. So, these can kind of come up with possibility ABC evaluate all of them in a parallel fashion, and then choose the correct one out of there. So, that kind of obviously you know, contributes to speed and efficiency of the parsing process.

Second is, yes they adopt a parallel processing architecture, borrowing from the MacDonald and Rumelhart school of parallel distributed processing. Also, these constraint based parcels represent different aspects of the sentences. Including their syntactic structures, as patterns of activations across a large number of interconnected units. So, the distribution of the sentences of or the information contained in the sentences is, happening much like, say for example, if you remember we were talking about distributed processing, where as far as words were concerned. So, this is in some sense distributed representation of the knowledge that is contained in a sentence in the brain.

Refer Slide Time: (04:14)

- Moreover,

    - Partial and incomplete information can lead to partial activation of multiple mental representations, so at any given point, multiple structural representations may be active in the network.

    - The system, effectively ranks these structural hypotheses, with more activation to the more likely, and less activation to the less likely. So, there is competition.

    - An implicit assumption in most constraint-based accounts is that syntactic structures compete for activation, similar to what happens at the level of word processing in accounts of lexical access like TRACE (the competition assumption is made explicit in some versions of constraint-based processing accounts).

Further on, partially and incomplete information can leads to partial activation of multiple mental representations. So, that at any given point incoming information. You know, basically can help you choose even from the other one, or so at any given point in the model, there are multiple fully active, partially active representations, that are kind of you know going around in the network. Also, the system effectively ranks these structural hypotheses. So, it's not like each of these multiple hypotheses are all activated at the same level, there is a system of ranking as well. So, each of these multiple structural hypotheses are, sort of effectively ranked with more activation to the more likely ones, and less activation with the less likely ones. This kind of kind, you know helps us organize the number of possible representations in a hierarchical order. So, as to we know that the most possible representation is the most easily accessible one, as well. Another implicit assumption in this is that most in most constraint based

accounts. The syntactic structures do compete with each other for activation. Similar to, what happens in the level of word processing in accounts of lexical access like trace, and so on.

So, this is almost like trace. There are multiple interpretations of a, of a sentence possible. All of these multiple interpretations will be differentiatively plausible. So, this is high as possible, this is low as possible, this is so on and so forth. And, they are kind of, you know organized in such a way, not that they're not competing, they're still competing, but then they hierarchically, hieratically organized in the order of plausibility, and it kind of keeps the process efficient and fast. At the same time, giving you all of that kind of bandwidth and you know, evaluating these multiple structures at the same time.

Refer Slide Time: (04:14)

- The second critical difference between the garden path and constraint-based parsers is that the garden path parser relies solely on word category information for its inputs, but constraint-based parsers can draw on a much wider variety of cues to decide what structures to build and the relative emphasis to place on each alternative structure.

- Finally, constraint-based parsers are often referred to as *one-stage* models because lexical, syntactic, and semantic processes are all viewed as taking place simultaneously (as opposed to lexical processing preceding syntactic processing preceding semantic processing, which is the general approach taken by two-stage models).
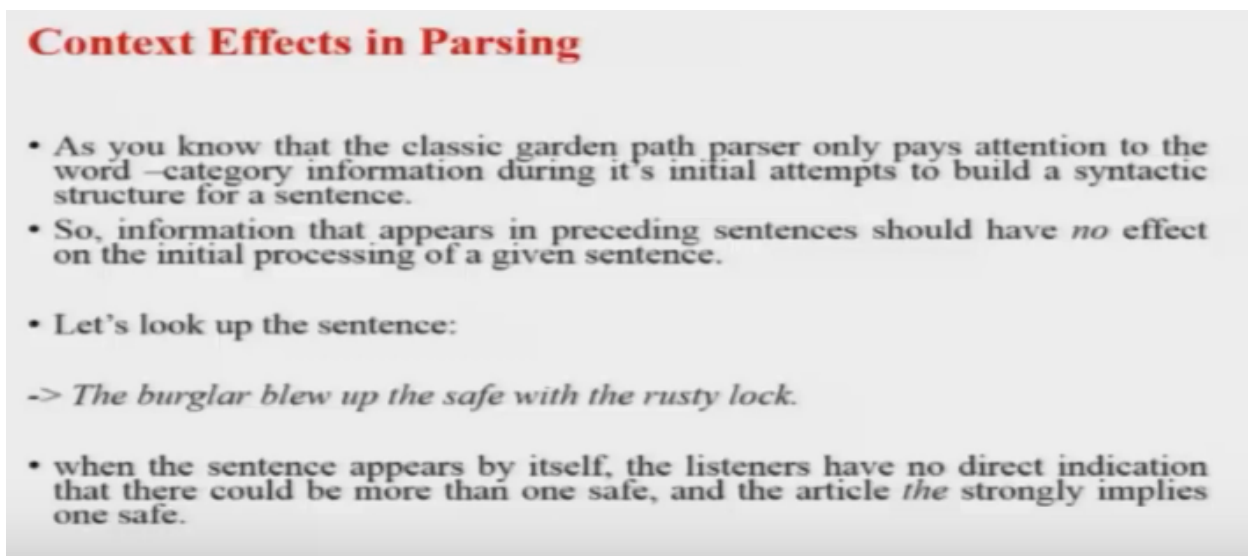
Another very critical difference between the garden path theory and the constraint based parser is that the garden path parser relies solely on word category information. If you remember, what was happening in the Lin Fraser's model, in the first stage the lexical processor was giving you word category information. Towards, giving you this is a conjunction, this is a noun, this is a verb, this is a determiner, this is another noun, this is another word, this is a preposition. So word category information, the garden path parser was relying only on the word category information for the inputs. But, the constraint based parsers, can draw on a much wider variety of cues. So, the based pastors, not only take word category information into account. But, they take all sorts of other information into account as well. And, it is on the basis of all of this information it is on, the basis of the summation of all of this information, that basically you know one gets to decide, which structures to build and the relative emphasis on each structure is kind of also organized in that sense. Okay?

So, I hope you get it, the whole idea is in a very simple sense, you have word category information, but you can have context information as well you have, you know visual information, phonological information, you can have so many information. On the basis of, so many information, you can create so

many structures. And, on the basis on say for example, if there are five converging evidence on one, whereas four converging evidence only on the second and three converging evidence on two. You can hierarchically organize this as well, this hierarchical organization; will be able to help you solve the whole competition as to which is the more plausible structure. So, the relative emphasis is also defined by the amount of evidence. We'll talk about this in more detail as you move ahead.

Now, constraint based parsers are often referred to as, one-stage models. There is no two-stage, there's no lexical processing in thematic interpretation at two-stages, all of that is happening one and at the same time. So, lexical syntactic and semantic processes are happening all viewed are basically happening simultaneously in these models. And, as supposed to say for example, lexical processing and thematic processing happening separately in the earlier model. So, this basically is one of the differences, if you see, with respect to the CBP models. From, now on I'm call them CBP models versus the garden path theory.

Refer Slide Time: (08:34)

So, let us kind of look at some of the cues that I was mentioning. The first cue that I can talk about is the effect of context. Now, context as you saw, in ambiguous, unambiguous word. When we talked about balance words, and ambiguous words, and you know bias words and unbiased words. And, talk about a neutral context and a biasing context. So, context kind of plays a very important role, not only in the interpretation of word meanings. But, even more so in the interpretation of how sentence has to be taken. So, it will be interesting, if we do not take context information into account. The earlier model the two-stage model from Linphrasal does not really talked so much about context. However, the constraint based models; the CBP models take context as an important source of information. So, let us look at now, how

context might be affecting somebody's interpretation, of particular sentences. What are the context based effects in passing, let's look at them.

Now, as you know I mean the classic garden path parser only pays attention toward category information, at during its initial attempts to build a syntactic structure. So, only later it will probably take on the other kind of information. Let's take an example, the burglar blew up the safe with a rusty lock is the sentence, and what you have to do is you have to kind of try and interpret it. Now, when the sentence appears just by itself in isolation, the burglar blew of the safe with the rusty lock and listeners have no direct indication as to what could be the meaning of this one. Okay? The article 'the' however strongly implies once if 'the safe' with the rusty law, not 'a safe' with the rusty lock. Okay? It was not that. It's not known that there are safe so many safes and each of the safes are with the rusty lock. So, one of the safes, one of a safe with rusty lock has been great. So 'the' is there, 'the' is saying a particular safe has been blown up. So, that you have to kind of you know keep in to mind.

Refer Slide Time: (10:41)

- So, this creates challenges for the listeners; in particular when they get to reading *the rusty lock*, they would need to revise some of their semantic assumptions, i.e. they have to change from assuming only a single safe to assuming only two safes; & they have to assume that the implicitly introduced safe or safes do not have rusty locks.

- These semantic changes have to be made regardless of the syntactic structure that listeners actually build for the sentence.

- If that is true, how can we make the sentence easier without changing it's syntactic structure?

Now, moving further, because there is 'the' there, it creates a bit of a challenge for the listener. What is the challenge? In particular, when the listeners or readers, get to reading or hearing 'the rusty lock' they would need to revise of their semantic assumptions. That is, they have to change from assuming only a single safe, to assuming only two safes. Okay? So, they have a kind of change from, they have to be two safes or at least more than one safe. Because, one of 'the safe' so, basically there could be one or two or three safes. 'The safe' with the rusty lock has been known of while others have been remained intact. So, this is something that, the people will need to take into account, when they are reading sentence or hearing this sentence. Okay? This, these semantic changes, the semantic updation that needs to take place, has to be made regardless of whatever syntactic structure you're coming up with. Okay?

So, you can, we've not talked about syntax, we will talk about that moving further. But, again kind of trying to explain this to you again. 'The burglar blew up a safe with the rusty lock' is one, there is, you know just only one safe and that has been blown, versus 'the burglar blew up the safe with a rusty lock' as if there are many other safes, and only that safe which has a rusty lock has been blown up. So, this is what you have to keep in mind. Mentioning 'the' there, kind of creates this possibility. So, that is why you have to update your assumptions to assuming that. There might be more than one and safe and only the safe which has the rusty lock has been blown up, while others have been you know have remained intact.

Refer Slide Time: (12:12)

- the problem can be solved by introducing background information.

- (17) *The burglar was planning his next job. He know that the warehouse had two safes. Although one was brand new from the factory, the other one had been sitting out in the rain for ten years.* **The burglar blew up the safe with the rusty lock.**

- Now, acc. to garden path theory, sentence (11) should still be difficult to process, as regardless of the story, the syntactic structure that you need for this sentence has a complicated syntactic structure & is hard to build.

Now, if this is true, how can we make the sentence easier, without changing its syntactic structure? How can we make the reading of it easier, one of the ways could be that you provide contextual information? You provide something please feeding to that information. So, how will you do it? Say for example, take the example seventeen here. The burglar was planning his next job. He knew that the warehouse had two safes. Although one was brand new from the factory, the other one had been sitting out in the rain for ten years. What did the burglar do? The burglar blew up the safe with a rusty lock. So, the safe which had the rusty lock was blown up. Apparently the burglar suspected that this old, you know slightly rusty is safe. Probably should be having all the goods, whereas the newer safe probably does not. Okay? So, intelligent burglar kind of you know you have to keep this in mind.

Refer Slide Time: (13:10)

Context Effects in Parsing

- As you know that the classic garden path parser only pays attention to the word –category information during it's initial attempts to build a syntactic structure for a sentence.
- So, information that appears in preceding sentences should have *no* effect on the initial processing of a given sentence.

- Let's look up the sentence:

-> *The burglar blew up the safe with the rusty lock.*

- when the sentence appears by itself, the listeners have no direct indication that there could be more than one safe, and the article *the* strongly implies one safe.

Now, according to the garden path theory, sentence eleven should still be difficult to process, as regardless of the story, the syntactic structure that you would need for the sentence, will be slightly complicated. Let us look at this. What is sentence somebody eleven? This one, the burglar blew up the safe with the rusty lock, this is sentence number eleven. Yeah! You have to kind of, see what kind of syntactic structure we can come up with.

Refer Slide Time: (13:20)

- But acc. to *referential context account*, the parser can use contextual information to decide which syntactic structure it will favor at a given point in time.

- In other words, "when there is a choice of syntactic structures, build the structure that is most consistent with your current semantic assumptions. If you have a choice of structures, build whichever one allows referring expressions to be unambiguous. "

- This means that sometimes the parser will build a more complicated structure when a simpler one is licensed by the grammar & consistent with the input.

- Also, in the case of (17), while the readers will need to build a more complicated structure than (11) but still should be easier to process. Confirmed.

Now, this was the garden path interpretation. However, if you look at the context based account, the referential context based account. What does that say? There says, the password can use the contextual information, to decide which syntactic structure it'll favor at a given point in time. It we kind of draw from the context, and in order to kind of see which contextual information is plausible here. And, take only that contextual information into account, and so, it will be easy to read. Okay? In other words, when

there is a choice of syntactic structures, build the structure that is most consistent with your current semantic assumptions. If you have a choice of structures, build whichever one follows allows referring expressions to be unambiguous. So, in case you can come up with multiple structures, you have to follow only this that structure that kind of allows the other things to become unambiguous. Because, the context is allowing this kind of integration will be unambiguous, you kind of going to take that syntactic structure. This means that, sometimes the parser will build a more complicated structure, when a simpler one is licensed by the grammar and consistent with the input..

So, even though, if you kind of take the assumption where the syntactic, I mean the safe with the rusty lock, will kind of lead to more complicated structure. Even though, it is kind of leading to a more complicated structure, it is still leading to disambiguation; it is still leading to something that kind of goes on with the context, and with your semantic assumptions. So, that is why, what people will do is, in the case of seventeen, while the readers we need to build a more complicated structure than a eleven. It will still be easier to process than eleven. That is the assumption here. Okay?

Refer Slide Time: (15:11)

- **Let's elaborate:**

  - Sentence (11) starts by saying *The burglar blew up the safe* ... As soon as listeners get to *the safe*, they try to figure out what *the safe* refers to.

  - Notice that the context in the mini- story has introduced two safes—a new one and an old one. By itself, *the safe* could refer to either of these two safes.

  - Thus, listeners need additional information to figure out which of the two safes the NP *the safe* is supposed to point to.

  - If listeners attach *with the rusty lock* to *the safe*, that will create a phrase that is semantically unambiguous and that fits well with the preceding story context.

Let's kind of look into a little bit more detail, I am kind of giving you the theoretical idea. Let's look into a little bit more detail of how the processing is happening here. Sentence eleven starts by saying, 'The burglar blew of the safe. As soon as the listeners get to the safe, they try to figure where what the safe refers to. What is this safe thing? The context in the mini story that we, just introduced in seventeen, has introduced two safes new one that has come out from the factory and an old one that is rusting in the rain. By itself, the safe could refer to either of the two safes. Okay? But, 'The Safe' is written there. Because, the noun phrase is 'the safe', you kind of interpret. So, if listeners attach the with 'the rusty lock' to 'the safe' it will create a phrase that is semantically unambiguous. So, you know, because semantically the

rusty lock one is kind of, the one that is you know, rotting in the rain. That fits well with the preceding story context as well. And, that is why reading of sentence number seventeen is easier. You might need to kind of go back and forth here, and see what seventeen here and what eleven is. So, what you say need to do it, if you kind of go back and you look at the slide where seventeen is mentioned which is.

Refer Slide Time: (16:20).

- the problem can be solved by introducing background information.

- (17) *The burglar was planning his next job. He know that the warehouse had two safes. Although one was brand new from the factory, the other one had been sitting out in the rain for ten years.* **The burglar blew up the safe with the rusty lock.**

- Now, acc. to garden path theory, sentence (11) should still be difficult to process, as regardless of the story, the syntactic structure that you need for this sentence has a complicated syntactic structure & is hard to build.

Here, you see 'The burglar blew up the safe with the rusty lock.' So, if you kind of attach with the rusty lock to 'the safe' it becomes a more plausible story. Because, the story is already necessitated by the earlier context, you know? That, although one was brand new from the factory, the other one had been sitting out in the rain for years, you know? So, that kind of very makes this one very easy to follow. Even though, seventeen obviously will lead to a more complicated structure, there are so many notes you need to build here.

Refer Slide Time; (16:53)

- So, this creates challenges for the listeners; in particular when they get to reading *the rusty lock*, they would need to revise some of their semantic assumptions, i.e. they have to change from assuming only a single safe to assuming only two safes; & they have to assume that the implicitly introduced safe or safes do not have rusty locks.

- These semantic changes have to be made regardless of the syntactic structure that listeners actually build for the sentence.

- If that is true, how can we make the sentence easier without changing it's syntactic structure?

Eleven, although hit will be much more simpler to read. But, because there is context and there is that kind of help in seventeen. And, the information is completely unambiguous, that is why we are saying, reading seventeen by the referential context account is going to be simpler and easier.

Refer Slide Time: (17:13)

- If they build the simpler syntactic structure, *the safe* will remain ambiguous—it could refer to either of the safes introduced previously in the story.

- Referential theory predicts that, in the context of stories like (17), comprehenders will build the more complicated structure rather than the simpler one for sentences like (11).

- As a result, sentences like (11) should be very easy to process despite their complicated syntax when they appear in stories like (17).

- This prediction was confirmed when people's reading times were measured. When sentence (11) appeared by itself, people slowed down when they read *the rusty lock*. When sentence (11) appeared in the context of story (17), people did *not* slow down when they read *rusty lock*.

Let us move further and see what is exactly is happening. If they are going with a simpler syntactic structure. That is the one that they will construct with sentence eleven, 'The Safe' will remain ambiguous. There is no context; you cannot attach 'the safe' with anything. This could refer to either of the safes produced previously in the story. Okay? So, you don't know which safe are we talking about? Obviously, there's also not no story there. So, a referential theory, the context theory predicts that, in the case of stories like seventeen, comprehend us will build a more complicated structure, rather than the simpler one for like that is possible in sentence eleven. As a result sentence like eleven should be very easy to process despite their complicated syntax, when they appear in a stories like seventeen. So, sentence eleven will be easier to read, when it is in context with seventeen. Okay? So, that is, that is the idea. On their own it will be difficult to read however. This prediction was confirmed, when people's reading times was measured, and when sentence eleven appeared by itself, people were slower in reading it. As compared to when sentence eleven, appeared in the context of the story that was there in seventeen. People did not slow down, when they read the rest along. Because, the context kind of already made sense. Therefore, of the sentence for them.

Refer Slide Time: (18:32)

So, contrary to what the garden path theory predicted, the parser does seem to pay attention to the information that the context makes available. At least some of the times, inorder to make decisions about which syntactic structures you need to build, for understanding a new sentence. Okay?

Refer Slide Time: (18:51)

Subcategory Frequency Effects in Parsing

- The garden path parser uses only word category information to make initial decisions about which syntactic structures it will build but words can provide more information than that.

- For example: consider the words *took & put*.

- Both of these words belong to the same syntactic category – verb.

- But other than having different meanings, are these two words equivalent?

Let's move ahead, there's also information about subcategories or subcategory frequency that kind of can be useful in processing in parsing. Okay? Parsing processing. What is subcategory, if you remember again from the developmental chapter we talked about verb subcategory? You know, what kind of information goes with a particular verb that is what we're going to talk about. Okay? The garden path parser uses only word category information to make initial decisions, about which syntactic structures it will initially build. Also, however basically taking just the word category information. But, the words can also provide more information than just about their category. So, that also needs to be understood.

Refer slide Time: (19:47)

- (18) Dr. Phil took.*
- (19) Dr. Phil took a nap.

- *took* is the kind of verb that needs a *post-verbal argument*. a direct object, it is *obligatorily transitive*.

- (20) Dr. Phil put.
- (21) Dr. Phil put a book.*
- (22) Dr. Phil put a book on the shelf.

- *put* requires both an object and a goal. Hence it is in a different *verb subcategory*, requires thus a different syntactic structure.

For example, you take these words, 'took' and 'put' both of these words belong to the same category, both are verbs. However, both of them have different meanings. So, other than having different meanings, are these two equivalents let us look at them. 'Dr. Phil took'. Or 'Dr. Phil took a nap,' 'took' is the kind of verb that needs a post-verbal argument. It needs a direct object, after it. If you say, 'I took' this next part is a question mark, you have to tell what did I take. 'Dr. Phil took' you have to say, took a nap, took a pen, took a ball, whatever. Okay? So, 'took' is basically something that is called an "obligatory transitive."In a different example, let us look at 'Dr. Phil put,' 'Dr. Phil put a book,' 'Dr. Phil put a book on the shelf'. 'Put' basically requires both an object and a go. It   is not only require, it is not only require an object. It also requires a goal to that object. So, I put what did I put and where did I put, it both of these things are important. So, if you see not only they are different, 'took' and 'put' not only they are different in meaning, but they are different in the grammatical requirements that the proves. 'Took' requires, it's an obligatory transitive, it needs an object, there are words that do not, you know that are in transitive, that do not need. So, I slept, it's alright. I slept on the car, I can supply that information. Okay? But, 'Dr. Phil took' needs that information or 'Dr. Phil put' needs both of the information. He put a book in the shelf or you know, those, those kind of things are there. So, because they are posing different syntactic requirements. They will lead to different syntactic structures being formed, that is the idea.

Refer Slide Time: (21:22)

- (23) Dr. Phil was reading.
- (24) Dr. Phil was reading a story.
- (25) Dr. Phil was reading a little girl a story.

- Verbs like *reading* can have multiple subcategories, including transitive, intransitive & ditransitive.

- So *was reading* has a number of subcategory possibilities, including intransitive, transitive, and ditransitive; and each of these subcategory possibilities is associated with a different syntactic structure. There are, in fact, many verbs that are flexible in this way. (See if you can think of a few.)

You can take another one. So, 'Dr. Phil was reading,' fine. 'Dr. Phil was reading a story,' also fine. 'Dr. Phil was reading a little girl sorry,' also fine. So, we have another kind of example, which is, which can have multiple subcategories, it can be intransitive, transitive and also ditransitive. So, one of the things that we kind of get out of this discussion is that you instead of only the word category. The kind of words, you know the kind of word subcategory is also important. So, you know 'took,' 'put' and 'reading' are all verbs. But, what kind of subcategory they belong to. Are they transitive verbs, are they intransitive verbs, are they ditransitive verbs, all of that is also required. And, the knowledge of all of that is going to help you in parsing the sentence that you are reading or listening. Okay? So, was reading has a number of subcategory possibilities, including intransitive, transitive and ditransitive. And, each of these subcategory possibilities is associated with a different kind of structure. So, whichever one you take, you will need to create a different syntactic structure to go with it. There are in fact many works that are very flexible in this kind of way, so you need to know this.

Refer Slide Time: (22:38).

- Constraint based theory says that the structural information is associated with individual words in the lexicon and this information influences which structural hypotheses will be pursued as sentences are being processed.

- In particular, a constraint-based parser will use subcategory information to determine which structural analysis to favor when more than one structure is consistent with the input. How does this work?

Now, how does the constraint based theory take this into account? The CBP theory says that the structural information that is associated with each individual word in the lexicon and this information influences the

way structural hypotheses will be generated. So, why you are kind of you know categorizing words into these particular categories. You will also know that, what kind of structure they come up with, what kind of structure they are permissible to use with? And, this information will obviously be taken into account, when you're generating these so many multiple possible structures. In particular, a CBP parser basically will use subcategory information to determine which kind of structural analysis to favor, when more than one structure is consistent with the input. Okay?

Refer Slide Time: (23:24)

- (26) The student saw the answer...
  - ...to the last question.

- (np) the student (vp) (v) saw) (np) the answer.

- (28) The student saw the answer...
  - ...was in the back of the book.

- In that case, *the answer* does not represent the direct object of *saw*. Instead, *the answer* is the subject of the verb *was*, and the sentence should be structured as in (29).

- In sentence (28), the part *the answer was in the back of the book* is called a *sentence complement*. *The answer was in the back of the book* is a sentence complement because it really is a sentence that could appear all by itself and because the whole thing is the post-verbal complement of *saw*.

- (29) (np) the student (vp) (v) saw **(s')** (np) the answer (vp) was in the back of the book.

So, let's, let's kind of look at this through an example. There is the sentence, Okay?'The students saw the answer,' that's all right. 'The students saw the answer to the last question', there's also one way of completing this. Now, the structure will become, 'the student' is the noun phrase, 'saw the answer' is the verb phrase. And, within that 'saw' is the verb, 'the answer' is the noun phrase. This is one kind of category. You can kind of draw it into a tree, if you want to. The other sentences, 'the student saw the answer, was in the back of the book'. So, this is a different kind of a sentence, a different kind of structure will be needed. Okay?

Now, in the second case, the answer does not represent the direct object of the verb 'saw'. Because, what is basically there because you kind of saw the answer, the answer is kind so the answer was in the back of the book, so the answer kind of goes with the next clause, Okay. Yeah! Instead the answer is the subject of the verb was and the sentence should be structured as in twenty nine. So, if you look at the house twenty nine is structured here. The noun phrase is 'the student,' verb phrases 'saw the answer' and then another verb phrase is 'was in the back of the book'.

Refer Slide Time: (24:33)

So, the structure the tree will kind of look very different so you can see here. Twenty seven can be structured like this, 'the student saw the answer,' so a noun phrase, verb phrase and within the verb phrase 'the saw' is the main verb and the answer is the noun.

Refer Slide Time: (24:46)
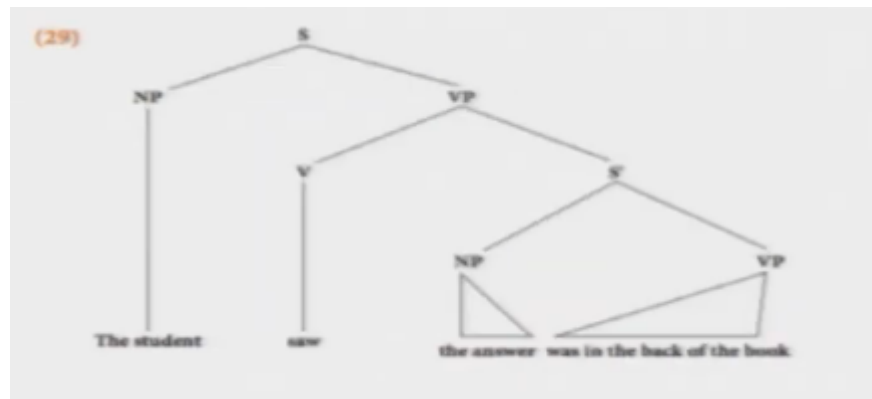
In a second possible way of organizing this, you can see it here. 'The student saw the answer, was in the back of the book.' So, this is one full sentence compliment. It's another sentence, it's sort of this is a combined sentence. This is a whole sentence in itself. The answer was in the back of the book, 'the students saw' is another sentence and they have been combined. Okay? So, this one's elements do not really kind of go and mesh with this once an element. Okay? But you need to know, you know, the verb category information will basically be able to tell you how to organize this.

Refer Slide Time: (25:21)

- So, when listeners get to "the answer" in (26) and (27), they face a choice between the structures in (30):
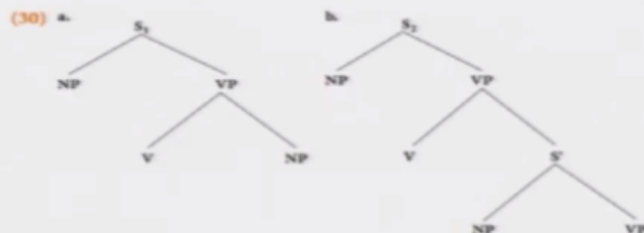
So, in the listeners get to the answer in twenty six and twenty seven, they face a choice between structures here. So, thirty (a) and thirty (b). Okay? Thirty (a) is how we organize twenty seven and thirty (b) is how we organize and as number twenty nine.

Refer Slide Time: (25:35)

- Garden path theory predicts that people should prefer the left-hand structure
  - because it is simpler than the right-hand structure
  - and because pursuing that structure allows comprehenders to continue working on the current VP.

- So garden path theory predicts that sentences like (28) should be harder to understand than sentences like (26), and this is true in general.

- Constraint-based theory also predicts that (28) should be harder than (26), but for a different reason. It turns out that both theories are correct in this instance—sentences like (28) really are harder to process than sentences like (26) (Frazier & Rayner, 1982; Pickering & Traxler, 1998).

Now, the garden for theory predicts that, the people should prefer the left-hand structure. Which is the one with, simpler nodes minimal attachments. The because it is simpler than the right-hand structure as I said and because pursuing the structure allows the comprehenders to continue working on the same Clause. You know, working on the same verb phrase. So, the garden path theory says what? It says, the sentences like twenty eight, should be harder to understand then sentence like twenty six, and obviously because sentence twenty eight is a compound sentence. You can say that, okay, that's true in general. But, is that true? Because of the reasons that the gp, garden path theory is saying or because of something else. The

CBP theory actually also says that, sentence number twenty eight should be harder, but it says that that is for a different reason. It turns out that both theories are correct in this instance sentences like twenty eight are actually found to be really harder to read than sentence like twenty six. That is alright. But, what is the underlying mechanism? Let us look at the underlying mechanism. Why is it, why is reading twenty eight harder?

Refer Slide Time: (26:40)

- **Constraint based theory says that (26) should be easier than (28), because:**
  - For the verb *saw*, the parser would predict a *direct object*, and hence (26) should be easier.
  - But the cbp has to change its mind for s', because it is a new instance.

- **Constraint-based theory assumes that people pay attention to *subcategory preference* information.**

- **Subcategory preference information reflects the likelihood that a given structure and a given verb go together.**
  - Consider the verb *saw* again. Suppose you know that 9 times out of 10 in the past *saw* was followed by a direct object (as in 26).

So, constrain ways theory says, twenty-six should be easier than twenty-eight. Because, for the verb 'saw' the passage should a direct object, so the answer is the direct object. The CBP says, their partner has to change its mind for sentence complicated complement because it's a new instance. As soon as it goes here, so it still needs to change its mind. You know, because there's a new compliment coming. So, CBP assumes that people pay attention the word subcategory preference information. So, saw the answer and you have to kind of look at what is the verb saw usually coming up with, is it coming up with direct object or is it coming up with these relative clause-like things.

Now, subcategory preference information, reflects the likelihood that a given structure and given verb will go together. Let us take an example, consider the verb saw again suppose you know, that nine times out of ten in the past saw was followed by a direct object, okay? You would know, that 'I saw,' I saw a duck, I saw a train, I saw man, I saw you know any, anything. So, you know that nine times out of ten saw will come up with a direct object. So, when you kind of are reading twenty-eight, you are surprised you have to reassess everything. That is why you will be slower on twenty-eight as compared to twenty-six. That is the CBP explanation.

Refer Slide Time: (28:55)

- the constraint-based parser takes information about the past—e.g., the likelihood that a given structure will appear when a given verb appears—and uses it to predict the future.

- So, according to constraint- based theory, (28) is hard because the parser predicts that a direct object is coming, and so the parser assigns more weight to the syntactic structure that allows *saw* to have a direct object.

- When the sentence actually provides the input for a different structure (sentence complement in this case), the constraint-based parser has to change its mind.
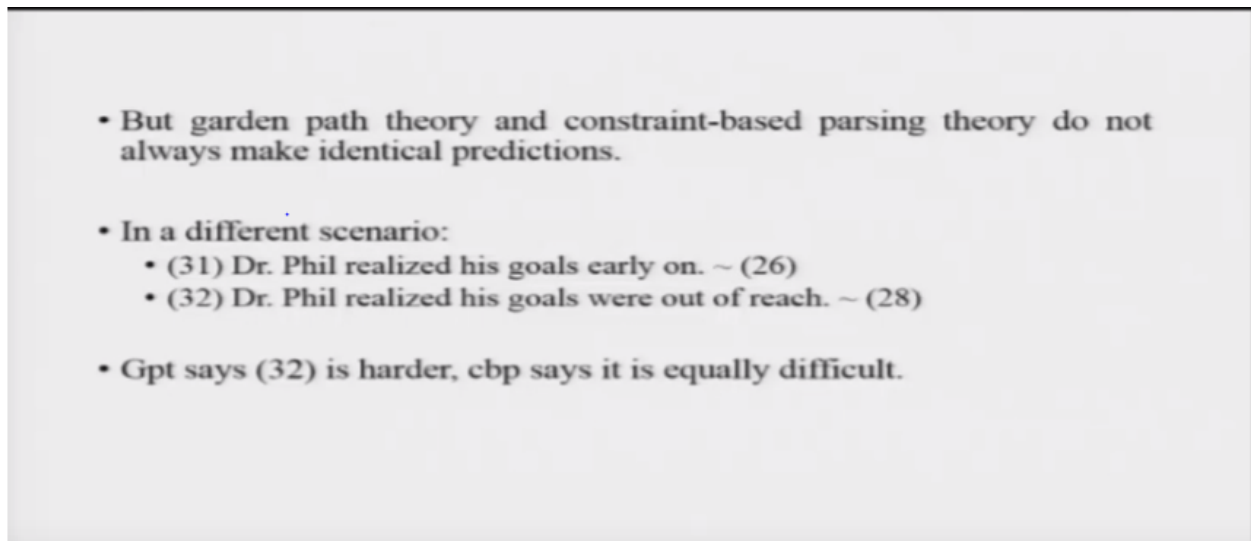
So, the CBP parser says it takes information about the past, how is, how or in what kind of structures has the word saw appeared, and what is the likelihood that saw will appear with the direct object or a sentence comprehend, sentence complement. Now, they likelihood that again, so because of that calculation it will get surprised in twenty-eight, so according to the CBP twenty-eight is hard because the parser predicts a direct object to be coming but, the direct object does not come, an entire sentence compliment come. The student saw the answer was in the back in the book, so then the full sentence is coming, that is creating complications here. The students saw the answer, this one saw the duck, this one saw the bat that would have been much easier. So, when the sentence actually provides the input for a different structure the constraint-based parser has to change its mind obviously when the sentence compliment comes in, Okay?

Refer Slide Time: (28:56)

- Garden path theory and constraint-based parsing theory both make the same prediction for sentences (26) and (28), and they both provide an explanation for people's actual observed behavior.

So, what we see the garden bar theory unconstrained bass theory both make the connected assumption. Both say that twenty-eight is harder, but they say it for different reasons. Twenty-eight is harder for the garden path theory because it'sa more complicated structure twenty-eight is harder for the CBP theory because the verb subcategory information is violated. So, yeah! You have to kind of understand that. Now, it's again it is that you know and in this case the garden path theory and the CBP were making the identical predictions. But, in some cases they will not make that into a prediction as well.

Refer Slide Time: (29:34)

- But garden path theory and constraint-based parsing theory do not always make identical predictions.

- In a different scenario:
  - (31) Dr. Phil realized his goals early on. ~ (26)
  - (32) Dr. Phil realized his goals were out of reach. ~ (28)

- Gpt says (32) is harder, cbp says it is equally difficult.

Let us see where, let's take a difference intense. Dr. Phil realize his goals early on it's very similar to twenty-six, Dr. Phil if a realize his goals were out of reach. It's similar to twenty-eight. Dr. Phil realize his goals were out of reach is the sentence compliment. Gpt the garden path theory says, thirty-two is harder, CBP however says, it's equally difficult. Now, very quickly I will just tell you my guesses and then we will see this in more detail. Dr. Phil realized his goals early on, is direct object, realized his goals you know, realize comes with his direct object his goals in thirty-two. Dr. Phil realized his goals were out of reach. That's a sentence compliment. That goals are not a direct object of realized, okay? So, obviously thirty-two leads to a more complicated structure, thirty-one leads to a simpler structure, so Gpt will say thirty-one is easier to understand. Thirty-one is harder to understand. CBP on the other hand says both thirty-one and thirty-two are equal you know, in difficulty. Let us see how that really happens.

Refer Slide Time: (30:37)

- Acc. to cbp, (32) ~(31) as the subcategory information in (32) points readers towards the correct syntactic structure right away. Unlike the verb *saw*, the verb *realized* appears with a sentence compliment s', 90% of the time; so the parser is ready and does not face trouble while processing (32).

- (28) was predicted to be harder, as the parser could not predict the s' because of *saw*. Confirmed.

- This might be an example for the *tuning hypothesis*, i.e. *"the structural ambiguities are resolved on the basis of stored records relating to the prevalence of the resolution of comparable ambiguities in the past."*

According to CBP thirty-two and thirty-one will be similar, because the subcategory information in thirty-two points the readers towards the correct syntactic structure right away. Unlike, the verb saw, the verb realized usually is followed by sentence compliments, because thread realized is usually followed by the sentence compliment there is no surprise or no near this nothing out of the blue coming in for the readers and that is why they will be reading it with equal different equal ease or difficulty as you would say. Twenty-eight was predicted to be harder, because the parser could not predict the sentence compliment earlier but here the parser can predict the sentence compliment and that is why it's being equally easy. This might be an example for what is called the tuning hypothesis. The tuning hypothesis says the structural ambiguities are resolved in the basis of stored representations or stored records relating to the prevalence of the resolution of comparable ambiguities in the past.

Basically, if you've had practice with similar kind of ambiguous structures you will be able to solve forthcoming or future ambiguous structures on the basis of your knowledge of solving them. It's just practice you know way simply. Now, these are two effects we've talked about context effects, we've talked about you know verbs sub category information. Let us talk about one more effect that seems to be useful in parsing.

Refer Slide Time: (32:02)

Cross – Linguistic Influences in Parsing

- Data from cross-linguistic research may be informative about sentence processing techniques.

- One aspect of cross-linguistic research focuses on:
  - Structural preferences *w.r.t* structural frequencies.

(33) *Someone shot the (female) servant of the actress who was standing on the balcony with her (male) spouse.*

(34) *Someone shot the (male) servant of the actress who standing on the balcony with her (male) spouse.*

Let us talk about cross linguistic influences in passing, see a lot of us most of us nowadays are bilinguals, multilingual. Okay? Because, we know so many of these languages, because each language has a different set of syntactic rules and structures. It is now increasingly felt plausible and possible, that syntactical information from two languages that you know, will interact with each other. Okay? Suppose I know Hindi and I know English. Hindi syntax will probably affect my reading, of English syntax and vice-versa. Let us look at this in more detail now. Now, data from cross linguistic research has shown that its sentence processing kind of takes cues from both languages. Now, one aspect of cross linguistic research and across linguistics and this comprehension research focuses on structural preferences within languages. Okay? With respect to structural frequency, so in particular languages particular structures are more frequent, whereas in particular other languages those structures might be less frequent. So, there is structure A, structure B, structure C; language A, language B, language C. Structures A and B could be more frequent in a language A, but not frequent in language B or why you know, or you can kind of create so many of these combinations. Okay?

So, let us kind of try and understand this in with the help of this example sentence thirty-three short the female servant of the actress who was standing on the balcony with her spouse, someone shot the servant of the actress who was standing on the balcony with her spouse. Okay? In the first there is this sense of you know, you say, the males of a female servant. The second you express you know, male servant. Okay? And, that kind of a spouse can be male or female you know, it's basically male in both cases. Because, her spouse is there, those spouses will be male. Let's read it again someone shot we are talking about female someone shot the female servant of the actress who was standing on the balcony with her spouse, female servant of the actress who are standing in the balcony with her spouse, someone shot the male servant of the bulk of the actress was standing on the balcony with her spouse. Okay?

- Sentence (33) is globally ambiguous because the relative clause, "who was standing in the balcony with her husband" could go with either "(female) servant" or "actress".

- Sentence (34) is temporarily ambiguous, because the relative clause in ENGLISH, would be tied to the second of the two nouns, i.e. "actress".

- But SPANISH and FRENCH speakers, prefer attaching the relative clause to the first noun, i.e. "servant".
  - Attachment to the first noun, appears to be the more *frequent* option in SPANISH and FRENCH.

Now, if you again go back and read sentence thirty-three, which is the first sentence we find that it is globally ambiguous. Because, who was standing in the balcony with her husband could apply to both the female servant and the actress. In sentence thirty-four it is just temporarily ambiguous because as soon as you read the entire sentence you will know that their relative Clause is basically will apply to the second one actress. Also, you will know, that the relative clause in English would be tied to the second of the two now so the it is basically observed that in English the relative Clause is really tied to the more recent or the second one. Okay? But Spanish and French speakers they prefer attaching the relative clause to the first noun. So, basically what is happening is if you were, you know, if you are a French or a Spanish reader in you reading this in English, you will want to attach the relative clause to the first noun to the servant and then because the initial when the servant is kind of also because this event is female it can also take up this they'll create to an ambiguity, and the first one to the spouse can actually not be attached to the servant, because that's a male servant. That'll kind of lead to a problem, Okay? So, attachment to the first noun appears to be the more frequent option in Spanish and French, but not a very frequent option in English.

- Experimental outcomes for (33) & (34) appear to support the idea that more frequent structures appear to be easier to process.

- Confirms with the cbp assumption that people keep track of how often they encounter particular kinds of sentences.

- However, other factors like *animacy* and *concreteness* may also provide additional constraints; indicate the role of more *fine-grained information*. E.g. Brysbaert et al.'s Dutch examples.

Experimental outcomes people have done these experiments sentences through thirty-three and thirty-four and they appear to support the idea that more frequent structures appear to be easier to process. So, for a reader for him attaching the first class is more frequent. They will basically find that kind of organization easier for those readers and were a attaching to the second clause is more frequent, they will kind of find it easier processing according to that, okay. It confirms the CBP assumption that people do keep track of how often they encounter particular kinds of sentences. You know, the whole knowledge of frequency basically comes from number of times we have come across these kinds of sentences. So, if you have come across a particular kind of sentences many a times and if applied the same solution a time attaching to the first or attaching to the second clause and that solution has worked it kind of gives you a sense of practice and as long as the you know, the practice can be applied to a new sentence. It gives you the sense of you know, frequency. It gives you a sense of. Okay? This is how this works, I'm happy with it.

However, it has been shown that not only frequency, but there could be other factors that could decide where you're going to attach the relative clause to, those factors can also be factored in into your representation or choice of the structure, to such factors are animacy and concreteness. Let us see, so Brysbaert et al.'s group Marc Brysbaert you know, there is this kind of analysis and what they said that there is this fine-grained information animacy concreteness etc. Which will kind of figure in along with the information about structural frequencies?

Refer Slide Time: (37:09)

- Marc Brysbaert and Don Mitchell measured Dutch speakers' eye movements while they read the Dutch equivalents of (33) and (34) (Brysbaert & Mitchell, 1996).

- The eye movements indicated that Dutch speakers had more trouble interpreting the test sentences when the relative clause went with the first noun than when it went with the second noun.

- But when researchers looked at a database of Dutch sentences (that came from newspaper and magazine articles), they found that relative clauses went with the first noun more often than they went with the second.

- So, the more frequent structure appeared to be more difficult to process, contrary to what constraint-based and other frequency dependent parsing theories would predict.

So, Brysbaert and Don Mitchell they measured Dutch speaker's eye movements while they were reading Dutch equivalence of sentences thirty-three and thirty-four. The eye movements indicated that the Dutch speakers had more trouble interpreting the test sentences, when the relative Clause went with the first noun then when it went with the second noun. But when researchers looked at the database of Dutch sentences, they found the relative clause went with the first noun more often than when they went for the second. So, this is kind of counter intuitive. If by frequency, usually the relative Clause goes with the first noun so participants should find it easier, but in this case using that the participants are not finding it easier. I am repeating, the eye movements indicated that the Dutch speakers had more trouble interpreting the sentences, when the relative clause went with the first noun as compared to when the relative clause went with the second noun. By frequency estimates researchers found, that in Dutch the relative Clause goes more often with the first noun and less often with the second noun. So, by the frequency idea it should have been perfectly easy to read, however it is not the case why so? Let us wonder about that. So, the more frequent structure appeared to be more difficult to process contrary to what the CBP and other frequency dependent parsing accounts would suggest.

Refer Slide Time: (38:36)

- However, when other researchers analyzed the test sentences and the sentences from the database, they found that semantic factors like animacy and concreteness were more important than position in determining where the modifying relative clauses should go (Desmet, De Baecke, Drieghe, Brysbaert, & Vonk, 2006).

- So, when more *fine-grained* information was taken into account, reading time could be predicted by detailed frequency information.

However, when you look in more closely even researchers looked in more closely and they analyze the test sentences and the sentences from the database they found that semantic factors like animation concreteness were actually playing a part. So, those lying apart even more important than the position in determining then frequency in determining where the modifying relative clauses will be attached. So, what basically happens is, when more fine-grained information is taken into account reading times could be better predicted than just taking into account the structural frequencies.

Refer Slide Time: (39:13)

- How does the parser decide whether something is frequent or infrequent?
  - If we just count all sentences, simple active voice sentences will be the most frequent (example active voice sentence: *John kissed Mary*).
  - The parser should therefore favor the direct-object interpretation of any sentence that starts with a noun phrase and a verb phrase.

- But if we start counting up which structures go with an individual verb, then the parser should favor the sentence complement interpretation of any sentence that starts with a noun phrase followed by the verb *realized* followed by another noun.

How did the partial decide whether something is frequent or in frequent? If we just count all the sentences in simple active voice sentences, basically will be the more frequent. The passes should therefore favor the direct object interpretation of any sentence that starts with a noun phrase and a verb phrase, so typically you're assuming that the parser is accumulating all of this knowledge, it kind of gets to the point that okay, these are direct interpretations, I mean obviously active sentences are much more frequent as

compared to passive sentences. So, every time an active sentence comes processing should be faster, every time a passive sentence is comes processing should be relatively slower. Is that really the case? Let us look.

But, if we start counting up structures which go with an individual verb then the passage should favor the sentence compliment interpretation of any sentence. That starts with a noun phrase followed by the verb other noun. So, basically what will happen here is, that you will kind of start favoring structures wherein there is another sentence complain, say for example realize followed by another noun is more plausible here or, if we start counting up the likelihood of specific verb noun combinations, you know, verb subcategory information.

Refer Slide Time: (40:19)

- Or, if we start counting up the likelihood of specific verb–noun combinations, then the parser should switch back to favoring the direct- object interpretation of any sentence that starts with a noun, the verb *realized,* and the noun *goals.*

- Likewise, if we start factoring in animacy, then any sentence that starts with an inanimate noun should reduce the likelihood of a simple, active structure.

- This problem goes by the name *the grain size problem* (coined by Don Mitchell in an article in 1987).

Then the passage should switch back to favoring the direct object interpretation of any sentence, that starts with a noun say for example the verb realized and the noun goals so they could kind of go back to that. However, if you start factoring in animacy, then any sentence that starts with an inanimate noun should reduce the likelihood of a simple active structure. Because, simple active structures do not start with inanimate nouns. Okay? So, this problem is referred to as the grain size problem. What is it? That you're attaching to what and in what manner, so you have to kind of take that into account.

Refer Slide Time: (41:01)

Now, the grain, grain size problem kind of says, that languages offer multiple levels of analysis and people can potentially keep track of this statistics at any level of analysis. So, they will know at the level of sentence structure, they know at the level of verbs, they'll know at the level of which noun goes with which verb all of that. Okay? And, the degree to which a structural alternative is preferred can differ at different grains. So, at one level of analysis this interpretation seems plausible at, another level of analyze is this structure seems prominent and, and another level of structure the other structure can seem prominent. Okay? Now, how do you kind of get around this mesh? You know, how do you solve, which is the correct way? Okay. A possible solution that they offer is that the parser should not keep any statistics at all it does not keep any frequency information at all, as the two stage models say. Another the solution however, could be, that the parser keeps track at different grains and combines data from different games to come at a sum or a vague estimate of what is the most probable structure. Okay? Say, for example at the sentence level of analysis this is the most probable structure. At the animacy level of analysis this is the most probable structure. What I can do is easily average the two and compare the other two, another two with each of them, and the one which gives me the highest likelihood of correctness is the one that I will choose and go of it.

Refer Slide Time: (42:32)

- So, in our example involving *realized*, the parser will give some weight to the fact that the most common structure in the language is subject-verb-object, it will also give some weight to the fact that the most likely structure for any sentence with the verb *realized* in it is the sentence complement structure, but if it gets *realized* followed by *goals*, the parser will pay attention to the fact that, at this very fine grain, *goals* is a really good direct object for *realized*, and will therefore boost the activation of the syntactic structure that goes with that interpretation.

So, in our example, involving realized yeah!

Refer Slide Time: (42:37)

- But garden path theory and constraint-based parsing theory do not always make identical predictions.

- In a different scenario:
  - (31) Dr. Phil realized his goals early on. ~ (26)
  - (32) Dr. Phil realized his goals were out of reach. ~ (28)

- Gpt says (32) is harder, cbp says it is equally difficult.

Dr. Phil realized, his goals early on what says, Dr. Phil realized, his goals were out of reach.

Refer Slide Time: (42:43)

- So, in our example involving *realized*, the parser will give some weight to the fact that the most common structure in the language is subject-verb-object, it will also give some weight to the fact that the most likely structure for any sentence with the verb *realized* in it is the sentence complement structure, but if it gets *realized* followed by *goals*, the parser will pay attention to the fact that, at this very fine grain, *goals* is a really good direct object for *realized*, and will therefore boost the activation of the syntactic structure that goes with that interpretation.

So, in our example involving the world realized the parser will give some weight to the fact that the most common structure in the languages subject-verb-object, It was also give some weight to the fact that the most likely structure for any sentence of the verb realized is the one having the sentence complement and not the one having direct object. But, if it gets the realized followed by goals the parser will pay also attention to the fact that in this very fine-grain, goals is really a good direct object for realized. Dr. Phil realized, so rom or something would not really fit there because that's an inanimate thing. Goals is inanimate kind of fits very well with realized. Okay?

And, will therefore the boost the activation of the syntactic structure that goes with that interpretation, so at three levels you see one level the subject-verb-object is there, the other level you know that you know after realize you should expect a sentence compliment, and the other you'll also know that okay, realized is being followed by goals, goals are a good object for realized. If it were an animate thing that would not be a good object for the realized and it will create problems. So, on the combination of all three of these possible interpretations, is when you understand this is the one that, I need to boost the activation of and finally select. So, I hope this was not very complicated, I'm sure and saying again, you have to kind of, you know, go over this again and again a little bit to understand. But this is we still kind of will be talking about some of the other factors in constraint based parsing models. In the next class I hope this is making some sense. Thank you.