

## CAS LX 522 Syntax I

Week 4.  $\theta$  Theory and the Big Picture

### The forest vs. the trees

- Let's take a moment to look at what we've done and lay out where we're going.
- The underlying goal is to lay out a *model* of this subconscious knowledge of sentence structure that we have; a system which predicts what speakers find grammatical and ungrammatical.

### The forest vs. the trees

- All we have from the outset is *data* and *intuition* as to how the system works; we look at the data, we draw generalizations, we formulate hypotheses, and we look at more data to see how the predictions of our hypotheses fared.

### The forest vs. the trees

- We started out looking at sentences to see what makes a sentence?
- Looks like sentences need a subject, kind of centered around a noun, and a predicate, usually kind of centered around a verb.
- The subject part can have other stuff, not just the noun (adjectives, etc.), and the verb part can have other stuff, not just the verb (adverbs, etc.).
- So, we laid out some hypotheses as to what the subject and the predicate could contain.

### The forest vs. the trees

- We noticed that the things which can be subjects (we called then "noun phrases" because of the intuitive centrality of the noun) can also be objects of verbs or of prepositions, we noticed where the adjectives and prepositional phrases seem to be able occur with respect to the noun and so forth.
- Looking deeper, with the idea of constituency in mind, we continued to revise our hypotheses until we came up with rules for the noun phrase and the verb phrase and the other components of the sentence that seemed to share a lot of common properties.

### The forest vs. the trees

- The fact that pretty much *any* kind of phrase that we looked at seemed to have roughly the same properties suggested a further, bigger hypothesis about how Language works:
- X-bar theory: The hierarchical structure of sentences is constructed (only) of phrases that conform to the X-bar template.

## The forest vs. the trees

- Along the way, we discovered that if we assume X-bar theory is right, we probably had mis-named the phrase which can be the subject of a sentence or an object of the verb; based on the evidence from possessor phrases, we determined that what we thought was a “noun phrase”, headed by a noun, was actually a “determiner phrase” headed by a determiner (and *containing* a noun phrase headed by a noun).

## The forest vs. the trees

- In general, this is how we’ll progress; we consider some part of the data, form hypotheses based on the generalizations we see, and then look for data that we don’t account for.
- Right now, we have the basic tools we need to diagram the structures of sentences (categories, X-bar theory), but that’s by no means the end of the story...

## The forest vs. the trees

- There are several large areas we need to address. First of all, simply drawing a tree that conforms to X-bar theory doesn’t guarantee that we’re going to have a grammatical sentence.
- Theta theory and subcategorization are the major components of our final theory which help make sure that our structures are legitimate. These are going to be our main topics today.

## The forest vs. the trees

- Another major component of syntax is *movement*, which has many complex properties.
- In general, the idea is that sentences like these
  - John will leave.
  - Will John leave?
- Are *related* in a meaningful way.

## The forest vs. the trees

- The underlying view of the grammatical system has us *starting with* something like:
  - John will leave
- ...in either case, and if you are trying to form a yes-no question, you will additionally move *will* from where you see it above to where you see it below:
  - Will John — leave?



## The forest vs. the trees

- That means that there are *two levels* involved in the *generation* of a sentence (where we our system is supposed to, in the end, *generate* all and only the grammatical sentences of a language).
- There’s the first level (*John will leave*), which is sometimes called the *Deep Structure* or *D-Structure* or *DS* representation of the sentence.
- Then, there’s a second level, after any movement has happened (*Will John leave?*), and this is what we pronounce. This is sometimes called the *Surface Structure* or *S-Structure* or *SS* representation.

## The forest vs. the trees

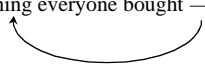
- In fact, there's even a *third* level; it's a level conceptually *after* the one we pronounce.
- Consider:
  - Everyone bought something.
  - ...I don't remember what that thing was, though.
  - ...but they all bought different things.

## The forest vs. the trees

- Everyone bought something
  - For every person *x*:  
for something *y*:  
*x* bought *y*.
  - For some thing *y*:  
for every person *x*:  
*x* bought *y*.
- These are renditions of the two meanings in a "logical form"; they differ in whether *everyone* or *someone* comes first.

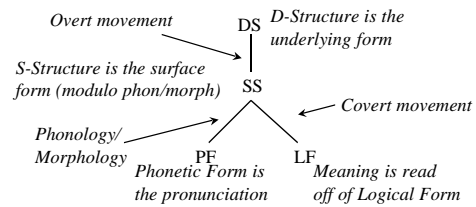
## The forest vs. the trees

- The idea is that *after* S-Structure there can be more movement to yield the *Logical Form* (or *LF*).
- There are two possibilities for *Everyone bought something*. It's ambiguous, like *I saw the man on the hill with the binoculars*.
- We might say that one one meaning (every...some...) no movement occurs, but on the other meaning (some...every...) *something* moves over *everyone*.
- something everyone bought —



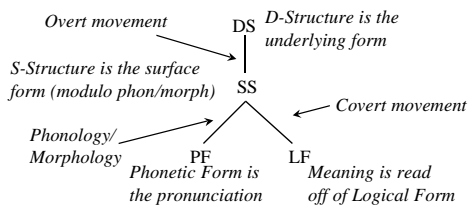
## The Y model

- This overall view of grammar has this shape (something like an inverted Y)



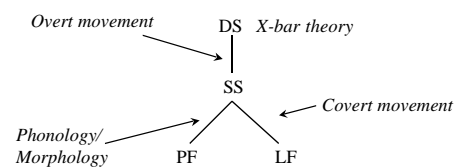
## The Y model

- We haven't been making distinctions, but we have generally been considering sentences that did not contain any (obvious) overt movement. Basically, we have been characterizing SS/DS.



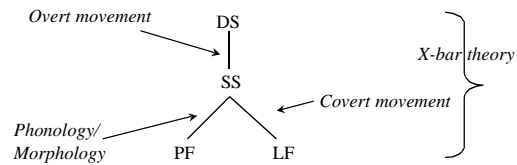
## The Y model

- Given this, we can only say that X-bar theory applies to SS/DS. However, we will make an additional assumption: Movement is structure preserving.



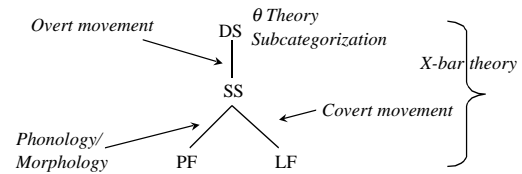
## The Y model

- By movement is structure preserving, we mean that movement will never change an X-bar compliant structure into an X-bar noncompliant structure. X-bar theory constrains DS and all representations created by movement (SS, LF).



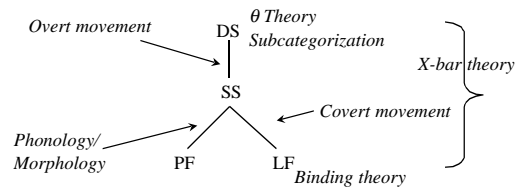
## The Y model

- Theta theory and subcategorization will constrain additional aspects of DS (for example, the requirement that *hit* has a DP object).



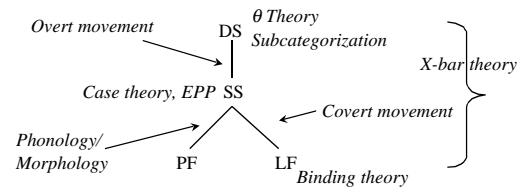
## The Y model

- Binding Theory concerning the *interpretation* of noun phrases (DPs) like *him*, *himself*, and *Bill*, are constraints on the form *LF* takes.



## The Y model

- Case Theory concerning the placement of noun phrases (DPs) within a sentence will turn out to be basically a set of constraints on SS.



## The plan

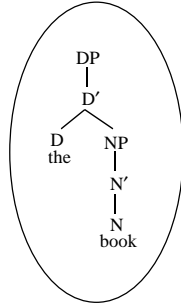
- This is an overview of the components of the grammar (to a good first approximation, of course), and the plan from here will be to work our way through the components ( $\theta$ -theory, movement, Case theory, Binding theory).
- Today:  $\theta$ -theory and subcategorization.

## But first, some clarifications...

- The introduction of the DP last week seemed to cause some unrest and confusion.
- As mentioned a few minutes ago, what makes this confusing is in part just an issue of *labeling*. We, sensibly enough, called the kind of phrase that can serve as a subject or an object, a "noun phrase."

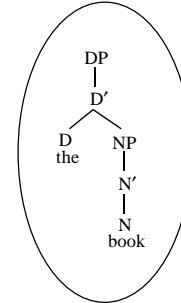
### But first, some clarifications...

- We discovered that as we explored the phrase of which the noun is the head (the NP), we shouldn't include determiners like *the* (or the possessive 's) inside; rather, the D is outside the NP.



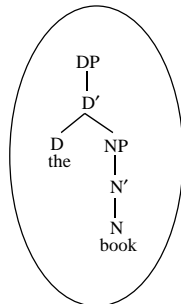
### But first, some clarifications...

- The implication of this is that subjects like *the student* or objects like *the book* were never NPs at all—they were DPs which *contain* NPs.
- Of course, NPs still *exist*! And everything we had previously discovered about them is still true. The data hasn't changed. It's only that NPs are *inside* of DPs.



### But first, some clarifications...

- One note about DP and the old term "noun phrase": You will find that people are not as precise about DP as they should be—even the textbook will frequently refer to "noun phrase" or even "NP" when it really means "DP".
- The term "noun phrase" (and its abbreviation) "NP" had become very entrenched in the vocabulary of linguistics—you'll just have to be awake as you read. Most of the time, people mean DP.



### But first, some clarifications...

- Perhaps a little more shocking is the basic idea of X-bar theory, which was probably not fully driven home last time.
- The logic was like this: Looking at NP, VP, and so forth, we found that the shape of the phrases is pretty much the same. This suggested a fundamental property of language, a generalization that holds over *any kind of phrase*.

### But first, some clarifications...

- The shape of a phrase is given by these three rules, where you can fill in X, Y, Z, and W with any category (N, V, Adj, ...):
  - XP: (ZP) X' (specifier rule)
  - X': (YP) X' or X' (YP) (adjunct rule)
  - X': X (WP) (complement rule)

### But first, some clarifications...

- NP: (ZP) N' (specifier rule)
- N': (YP) N' or N' (YP) (adjunct rule)
- N': N (WP) (complement rule)
- VP: (ZP) V' (specifier rule)
- V': (YP) V' or V' (YP) (adjunct rule)
- V': V (WP) (complement rule)

### But first, some clarifications...

- The thing is: X-bar theory has now *taken over* a lot of the function that our NP, VP rules had.
- The radical view to take on this is that there is only X-bar theory—*there is no NP rule, there is no VP rule, no AdjP rule, etc.*
- We can build trees with X-bar theory alone, without any category-specific rules like NP.

### But first, some clarifications...

- That sounds economical, but let's think about what the NP rules said:
  - NP: N'
  - N': AdjP N'
  - N': N' PP
  - N': N (PP)
- What here *can't* we simply derive from X-bar theory by substituting N for X?

### But first, some clarifications...

- NP: N'
- N': AdjP N'
- N': N' PP
- N': N (PP)
- X-bar theory does not *restrict complements to being a PP* (it allows any phrase to be a complement, for example VP, TP, DP, AdjP). It does not restrict left-adjuncts to be AdjPs, or right-adjuncts to PPs. X-bar theory makes *no category-specific* statements.

### But first, some clarifications...

- So, if X-bar theory is taking over the role of our NP, VP, PP rules, we are still left with the question of how the other restrictions get there.
- That is, we *have made progress*, we can now use a single set of rules to describe *any* kind of phrase. It is a higher kind of generalization, with a lot more coverage.
- But it leaves us with the information that we accumulated while constructing the NP rules that still needs to be predicted.

### And now, $\theta$ -theory

- To understand  $\theta$ -theory, we'll need to back away from the issue at hand (to start from the beginning of the topic), but what we're going to end up with is a system for ensuring that only the right kinds of things appear in NPs, VPs—to take care of parts of the NP, VP rule which *isn't* covered by X-bar theory.

### Verbs and arguments

- Verbs come in several kinds...
- Some have only a subject, they can't have an object—the intransitive verbs.
  - Sleep: Bill slept; \*Bill slept the book.
- Some need an object—the transitive verbs.
  - Hit: \*Bill hit; Bill hit the pillow.
- Some need two objects—ditransitive verbs.
  - Put: \*Bill put; \*Bill put the book; Bill put the book on the table.

## Verbs and arguments

- The “participants” in an event denoted by the verb are the *arguments* of that verb.
- Some verbs require one argument (subject), some require two arguments (subject and object), some require three arguments (subject, indirect object, direct object).

## Predicates

- We will consider verbs to be *predicates* which define properties of and/or relations between the *arguments*.
  - Bill hit the ball
  - There was a hitting, Bill did the hitting, the ball was affected by the hitting.
- Different arguments have different *roles* in the event. (e.g., The hitter, the hittee)

## Subcategorization

- Not all transitive verbs (that take just one argument) can take the same kind of argument.
  - Sue knows [<sub>DP</sub> the answer ]
  - Sue knows [<sub>CP</sub> that Bill left early ]
  - Sue hit [<sub>DP</sub> the ball ]
  - \*Sue hit [<sub>CP</sub> that Bill left early]
- So *know* can take either a DP or a CP as its object argument; *hit* can only take a DP as its object argument.

## Selection

- Verbs also exert semantic control of the kinds of arguments they allow: selection.
- For example, many verbs can only have a volitional (agentive) subject:
  - Bill likes pizza. Bill kicked the stone.
  - #Pizza likes anchovies. #The stone kicked Bill.

## The lexicon

- A major component of our knowledge of a language is knowing the words and the properties of those words. This knowledge is referred to as the lexicon.
- In the lexicon, we have the words (lexical items) stored with their properties, like:
  - Syntactic category (N, V, Adj, P, C, T, ...)
  - Number of arguments required
  - Subcategorization requirements (syntax)
  - Selectional requirements (semantics)
  - Pronunciation
  - ...
- These pretty much just have to be learned separately for each verb in the language.

## Thematic relations

- It has come to be standard practice to think of the restrictions (both subcategorization and selection) in terms of the *thematic relation* that the argument has to the verb—the role it plays in the event.
- One thematic relation is *agent* of an action, like *Bill* in:
  - Bill kicked the ball.

## Thematic relations

- There are lot of possible thematic relations; here are some common ones:
- Agent: initiator or doer in the event
- Theme: affected by the event, or undergoes the action
  - Bill kicked the ball.
- Experiencer: feel or perceive the event
  - Bill likes pizza.

## Thematic relations

- Goal:
  - Bill ran to Copley Square.
  - Bill gave the book to Mary. (Recipient)
- Source:
  - Bill took a pencil from the pile.
- Instrument:
  - Bill ate the burrito with a plastic spork.
- Benefactive:
  - Bill cooked dinner for Mary.
- Location:
  - Bill sits under the tree on Wednesdays.

## Thematic relations

- Armed with these terms, we can describe the semantic connection between the verb and its arguments.
  - Ray gave a grape to Bill.
    - Ray: Agent, Source, ...
    - A grape: Theme
    - Bill: Goal, Recipient, ...

## $\theta$ -roles

- An argument can participate in several thematic relations with the verb (e.g., Agent, Goal).
- In the syntax, we assign a special connection to the verb called a “ $\theta$ -role”, which is a *collection* of thematic relations.
- For the purposes of syntax, the  $\theta$ -role (the *collection* of relations) is much more central than the actual relations in the collection.

## $\theta$ -roles

- We will often need to make reference to a particular  $\theta$ -role, and we will often do this by referring to the most prominent relation in the collection.
- For example, in *Bill hit the ball*, we say that *Bill* has the “Agent  $\theta$ -role”, meaning it has a  $\theta$ -role containing the Agent relation, perhaps among others.

## The Theta Criterion

- Although an argument can have any number of thematic relations in the  $\theta$ -role...
- Each argument has exactly one  $\theta$ -role.
- On the other side, verbs (as we’ve seen) are recorded in the lexicon with the number of participants they require; each participant must have a  $\theta$ -role as well.



## The Theta Criterion

- Verbs have a certain number of  $\theta$ -roles to *assign* (e.g., say has two), and each of those must be assigned to an argument.
- Meanwhile, every argument needs to have exactly one  $\theta$ -role (it needs to have at least one, it can't have more than one).
- This requirement that there be a one-to-one match between the  $\theta$ -roles a verb has to assign and the arguments receiving  $\theta$ -roles is the Theta Criterion.

## Theta Grids

- We can formalize the information about  $\theta$ -roles in the lexical entry for a verb by using a *theta grid*, like so:

give

Source/Agent	Theme	Goal
i	j	k

- The columns each represent a  $\theta$ -role, the indices in the lower row will serve as our connection to the actual arguments; e.g.
- John<sup>i</sup> gave [the book]<sup>j</sup> [to Mary]<sup>k</sup>.

## Theta Grids

- John<sup>i</sup> gave [the book]<sup>j</sup> [to Mary]<sup>k</sup>.

give

Source/Agent	Theme	Goal
<u>i</u>	j	k

The first  $\theta$ -role is assigned to the subject. It is the *external*  $\theta$ -role. It is often designated by underlining it.

The other  $\theta$ -role are *internal*  $\theta$ -roles.

## Theta Grids

- One important thing to note about theta grids is that *adjuncts are never in the theta grid*.

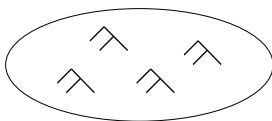
give

Source/Agent	Theme	Goal
i	j	k

- Adjuncts are related to the verb via thematic relations (e.g., instrument, location, etc.), but an adjunct does not get a  $\theta$ -role. They are optional.
- The  $\theta$ -roles in the theta grid are obligatory.

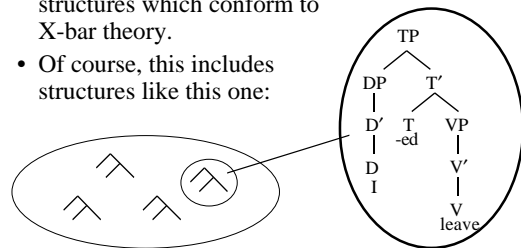
## How this works

- The Theta Criterion is a *constraint*, a *filter* on structures.
- There is an (infinitely big) set of structures which satisfy the requirements of X-bar theory. Here's a picture of it.



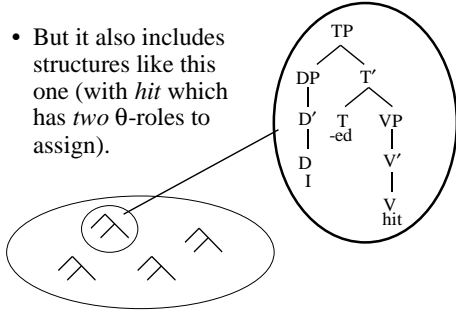
## How this works

- In here are *all* of the structures which conform to X-bar theory.
- Of course, this includes structures like this one:



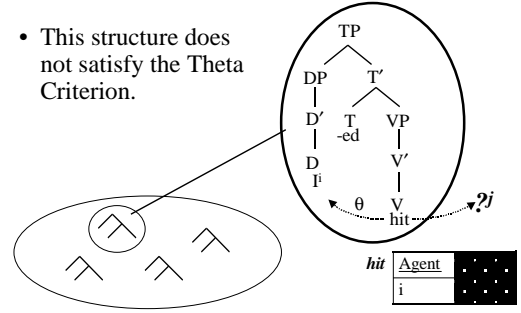
### How this works

- But it also includes structures like this one (with *hit* which has *two*  $\theta$ -roles to assign).



### How this works

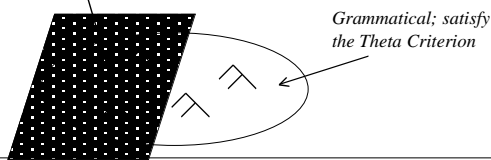
- This structure does not satisfy the Theta Criterion.



### How this works

- We can split the set of possible X-bar structures into two parts, those which satisfy the Theta Criterion and those which don't.

Ungrammatical; don't satisfy the Theta Criterion



### How this works

- In general, the model is one of *free generation* of (sets of) structures and movements, constrained by a variety of constraints (X-bar theory, the Theta Criterion, and many others that we will meet—the Case Filter, the Extended Projection Principle, Binding Theory, ...).
- Anything that satisfies the constraints is grammatical, anything that doesn't isn't grammatical.

### The Projection Principle

- The idea that lexical information directly constrains the validity of structures via categorial information, *argument structure* (theta grids), is embodied in the Projection Principle:
- The Projection Principle  
Lexical information (theta roles, etc.) is syntactically represented at all levels (DS, SS, LF)

### The Theta Criterion in action

- An example: *push*.

*push*

Agent	Theme
i	j

- Bill<sup>i</sup> pushed the shopping cart<sup>j</sup>.  
– Fine, *push* assigns two  $\theta$ -roles, one (the external  $\theta$ -role) is assigned to *Bill*, the other (the internal  $\theta$ -role) is assigned to *the shopping cart*. There are two arguments here, each gets a  $\theta$ -role.
- \*Bill<sup>i</sup> pushed.
- \*Bill<sup>i</sup> pushed the shopping cart<sup>j</sup> the corner<sup>?</sup>.

## The Theta Criterion in action

- An example: *cough*.

<i>cough</i>	Agent
	i

- Bill<sup>i</sup> coughed.
  - Fine, *cough* assigns one  $\theta$ -role (the external  $\theta$ -role), to *Bill*. There are one arguments here, and it gets a  $\theta$ -role.
- \*Bill<sup>i</sup> coughed the shopping cart<sup>?</sup>.

## Complications abound

- Things aren't really as simple as it might seem so far (have you already noticed)?
  - Bill ran.
  - Bill ran a mile.
  - Bill danced.
  - Bill danced a happy little jig.
  - Bill ate.
  - Bill ate a sandwich.

## Bill ran (a mile)

- So, *run* appears to be able to be used either as an intransitive verb (*Bill ran*) or as a transitive verb (*Bill ran a mile*).
- We will assume when you're building a sentence you choose the type of verb ahead of time (so, *run* is listed in the lexicon with two possible theta grids, chosen at the outset). We could notate this *run*<sub>1</sub> (intransitive) and *run*<sub>2</sub> (transitive). That is:
  - \*Bill ran<sub>2</sub>.
  - \*Bill ran<sub>1</sub> a mile.

## Passive

- The *passive* is something which appears to directly affect the theta grid of a verb; consider:
  - Bill ate a sandwich.
  - The sandwich was eaten.
- Eat* has two  $\theta$ -roles to assign. The *-en* suffix on *eaten* (or on any verb) seems to turn a transitive verb into an intransitive verb; *eaten* (passive) has only one  $\theta$ -role to assign. In fact, it's the  $\theta$ -role that was the *internal*  $\theta$ -role for *eat*.

## Lexical derivation

- Specifically, we can say that the *-en* suffix attached to a verb *removes* the external  $\theta$ -role (in some sense which we'll clarify later).

<i>eat</i>	Agent	Theme
	i	j

<i>eat+en</i>	Agent	Theme
	i	j

## Lexical derivation

- There are several other *derivational suffixes* of this kind, that alter lexically encoded properties in predictable ways; for example, there are several which change the syntactic category.
  - ion*: turns V to N (*translation*)
  - ize*: turns N to V (*colonize*)
  - ish*: turns N to Adj (*sheepish*)
- These are for our purposes considered to be "pre-syntax" (so their effects have already occurred to the elements in the terminal nodes of the trees).

## Lexical derivation

- As a side note, those category-changing suffixes often leave the *theta grid* (more or less) intact.
- *Destroy*: V,  $\theta$ : Agent, Theme.
  - Homer destroyed the toaster.
- *Destruction* (a noun, from *destroy* + *ion*)
  - Homer's destruction of the toaster.
- See? The complement of the verb (toaster) is now the complement of a noun (with the of we usually see with noun complements), the destroyer (Homer) takes the form of a "possessor". There are many complexities here that we'll save for later (probably Syntax II), but it's an interesting point.

## Bill ate (a sandwich)

- Now, back to the issue of "either transitive or intransitive" verbs (like *run*).
- The thematic role played by *a mile* in *Bill ran a mile* isn't really a Theme (the mile wasn't affected by Bill's running of it), but a sandwich in *Bill ate a sandwich* is pretty canonical Theme.
- Some verbs with canonical Themes of this kind can nevertheless appear without them (*Bill ate, Bill drank, Bill kicked, ...*).

## Bill ate (a sandwich)

- We could treat these in the same way we treated *ran* (by supposing that *eat* has two theta grids to choose from), but we might also look at it another way.
- There are a number of languages which seem to have an "antipassive" construction, which is sort of like the English passive except that it seems to be the *internal*  $\theta$ -role which gets "removed". This is often detectable through some kind of marking on the verb. Like English *-en* indicates passive.

## Bill ate (a sandwich)

- Given this crosslinguistic parallel, many syntacticians instead assume that *eat* exhibits the same phenomenon in English:
  - Bill ate a sandwich.
  - Bill ate+ $\emptyset$ .
- That is, English has an antipassive morpheme, but it is a zero morpheme (not entirely unlike the zero morpheme that can create verbs from nouns; e.g., *xerox, impact, shelf, corral, ...*)

## The EPP

- With the Theta Criterion in our toolbox, let's take a look at a special kind of sentence (which will turn out to tell us something important about syntax).
  - It rained.
  - It snowed.
- How many  $\theta$ -roles does *rain* assign?
- If we think about *it*, it doesn't really mean *anything* at all. It is not a participant in the event; it really *can't* be getting a  $\theta$ -role. (cf. also Spanish).

## The EPP

- So, the theta grid for *rain* really looks like this:

## The EPP

- Given the Theta Criterion and the fact that *rain* doesn't have any  $\theta$ -roles to assign, what's *it* doing there? And why doesn't *it* violate the Theta Criterion?
- As to the first question, the conclusion that syntacticians have come to is that the *it* is there due to a separate constraint, which goes by the name EPP.

## The EPP

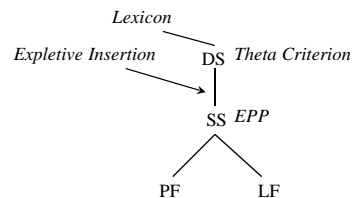
- The EPP**  
All clauses have subjects.
- The idea is that there must be something in the subject position (SpecTP) of every clause.
- Because *rain* has no arguments (no  $\theta$ -roles), a special, contentless pronoun (*it*) has to be inserted in order to have a grammatical sentence. This kind of "empty *it*" is called an *expletive* or a *pleonastic pronoun*.
- Expletive Insertion**  
Insert an expletive pronoun into the specifier of TP.

## The EPP

- As for the question of why *it* doesn't cause *it rained* to be a violation of the Theta Criterion, the solution we will adopt is an *ordering solution*.
- The idea is this: *First*, we check the Theta Criterion, and *then* we insert *it* (if necessary in order to satisfy the EPP). So *it* isn't even there when we evaluate the Theta Criterion.

## The EPP

- This is how this looks in the Y model we're building up.



## It is likely...

- Another place we see an expletive pronoun is with verbs like *is likely*.
  - It is likely that Bill left.
- Think about the semantic role that *it* plays in this sentence, and you'll see that it too is "empty", an expletive pronoun. However *likely* does have a  $\theta$ -role to assign, it assigns a  $\theta$ -role to its complement, the CP.

## It is likely...

*likely*

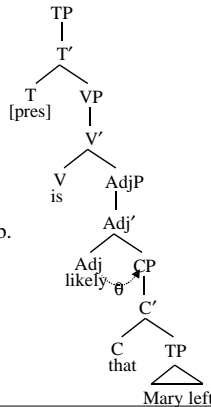
Proposition
<u>i</u>

- Note that in the theta grid for *likely*, we have a single  $\theta$ -role, but *it is not underlined*—*likely* has a single, *internal*  $\theta$ -role to assign.
- So, *likely* assigns a Proposition  $\theta$ -role to the CP in its complement, but the subject position is still empty and therefore needs to be filled with an expletive pronoun.

## It is likely...

- So, we have a partial tree like this, at DS, which satisfies the Theta Criterion.

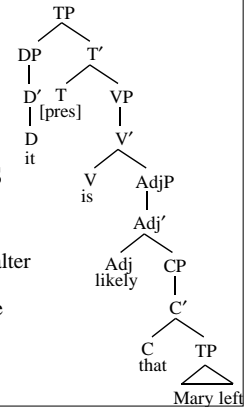
– *Note:* The textbook basically treats *is-likely* as if it were a verb. However, really *likely* is an Adj, and *is* is an auxiliary verb. It is *likely* (despite not being a verb; remember *destruction*) that has the  $\theta$ -roles to assign.



## It is likely...

- Then, Expletive Insertion applies, inserting *it* into SpecTP, resulting in this SS representation... which satisfies the EPP.

– (Basically, anyway... we'll alter one thing about this SS representation next week; the auxiliary verb *is* moves to T)



## That is likely.

- In this connection, consider also:
  - That Mary left is likely.
  - That is likely.
- In the first case, we have a CP in subject position instead of *it*. But of course the EPP is satisfied anyway because SpecTP is filled.
  - There is a quirk about the  $\theta$ -role—recall that we said that *likely* has only an internal  $\theta$ -role to assign, which now appears to be assigned to the subject (like an external  $\theta$ -role). We will return to this *apparent* (but only apparent) quirk in two weeks, when we talk about passives.

## That is likely.

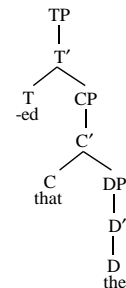
- That is likely.
- In this second case, we have *that* in subject position. Here, *that* is first of all not an expletive—it has a definite role to play, it's a pronoun standing in for a proposition (such as “Mary left”).
- It's worth pointing out that this *that* is not the complementizer *that* either, it's a real pronoun. There are two *that*'s in English, one that refers to something (a D) and one that introduces embedded clauses (a C).
- *It* is the same way; there is a real pronoun *it* (*I saw it*) and an expletive *it* (*It's hot in here*), different things.

## $\theta$ -Theory

- So,  $\theta$ -theory (theta grids and the Theta Criterion) and the EPP are two ways in which we narrow down the *overgeneration* of X-bar theory.
  - (*overgeneration* is the problem that there are many trees that comply with X-bar theory but yet are not grammatical)

## $\theta$ -Theory

- This still leaves open a couple of things.
- So far we have only talked about  $\theta$ -roles assigned by *lexical categories* (verbs, primarily, but sometimes nouns, adjectives, etc.).
- This doesn't provide an obvious way to rule out structures like this, though.



## Subcategorization

- The way this is generally thought of is as a matter of *subcategorization* (recall, subcategorization is a lexical property that specifies the syntactic category of its complement).
  - C *subcategorizes for* TP,
  - T *subcategorizes for* VP.
- These count as lexical properties, and thus can fall under the Projection Principle.

## Subcategorization

- In a sense, we could also look at this as an extension of  $\theta$ -roles; we could say that T has a  $\theta$ -role which can only be assigned to VP, for example, but the intuition that drove our original postulation of thematic relations is no longer available to guide us for functional categories.
- We can keep this as an option for later, but for the moment we'll just think of this as an issue of straightforward syntactic subcategorization.

## Subcategorization

- One other possibility which we won't directly pursue here but which has been pursued in recent syntactic theory is that what rules *\*did that the* out is that it must be possible to read the meaning off of the LF structure, and *\*did that the* is simply not meaningful—it is “gibberish”, it cannot be assigned a meaning, even if it is otherwise syntactically well-formed.
- This puts the problem in the “semanticists' court” in a sense; it would no longer be a problem of syntax to say why *\*did that the* is ungrammatical, but a problem of semantics. Either one could be right, perhaps it's even a combination of both. For now, we'll stick to syntax and subcategorization.

## A couple of loose ends

- The assignment of  $\theta$ -roles is considered to be part of the initial construction of the structure—when the DS is constructed by putting together lexical items into an X-bar compliant structure, this is where the Theta Criterion needs to be satisfied. We mentioned this in connection to the expletive pronoun—the Theta Criterion needs to apply *before* Expletive Insertion. Just to highlight this:
  - The Theta Criterion applies at DS.

## A couple of loose ends

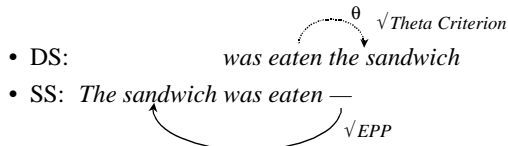
- Perhaps you noticed this, but let's think about the passive again. In the lexicon, the *-en* suffix takes a verb and “strips off” the external  $\theta$ -role.
  - Mike ate the sandwich.      Eat: Agent Theme
  - The sandwich was eaten.      Eaten: ~~Agent~~ Theme
- Now, the external  $\theta$ -role is the one that is assigned to the subject position—yet it looks like in the passive, the internal  $\theta$ -role is appearing there. So, does the passive *-en* “promote” the internal  $\theta$ -role to an external  $\theta$ -role?

## A couple of loose ends

- It turns out the answer is *no*, that this really is the argument which receives the internal  $\theta$ -role that is appearing in subject position. We're going to explore this in much more detail, but consider:
  - The internal  $\theta$ -role is *always* assigned inside the VP.
  - The Theta Criterion applies at DS.
  - The EPP applies at SS.
  - What happens between DS and SS is *movement*.
- So what's happening in the passive?

## A couple of loose ends

- Taken together, this suggests that between DS and SS, the Theme argument *moves* from the object position to the subject position (in order to satisfy the EPP).



## A couple of loose ends

- This also leaves open an interesting possibility with respect to intransitive verbs. Intransitive verbs have a theta grid with a single  $\theta$ -role to assign. Like *walk*, say.
- *Walk*: Agent.
  - (Agents are pretty much *always* external arguments)
- So, you can have verbs with only a single external  $\theta$ -role, and the passive *-en* morpheme can “create” verbs with only a single internal  $\theta$ -role.
- Might there be intransitive verbs that start out with only a single internal  $\theta$ -role?

## A couple of loose ends

- *Why, yes...* Here’s an example:
- *Fall*: Theme.
- How would we suppose these would act? They’re essentially “inherently passive”—they don’t have an *-en* morpheme, but instead they *start out* without an external  $\theta$ -role.
  - (Actually, we saw something that assigns only an internal  $\theta$ -role already when we considered *is likely* earlier, but *is likely* acts differently in that it allows Expletive Insertion in order to satisfy the EPP—With *fall*, you have only the movement option: \*It fell Bill.)

## Unaccusatives vs. unergatives

- There are many reasons to think that verbs like *fall* have only an internal argument.
- First, the subject is really a Theme as far as thematic relations go, it is *affected*, not an agent.
- Another interesting piece of evidence comes from Romance languages like French, where passives and verbs like *fall* acts similarly, and differently from other (truly agentive) intransitive verbs.
  - *Jean est tombé*. ‘John fell.’ (past unaccusative)
  - *Le fromage a été mangé*. ‘The cheese was eaten.’ (passive)
  - *Jean a marché*. ‘John walked.’ (past unergative)

## Unaccusatives vs. unergatives

- The point is really that we can distinguish two types of single-argument (intransitive) verbs in terms of their theta grid with respect to whether they have an external  $\theta$ -role to assign or not. Their (highly unintuitive) names, for the record, are:
- Unaccusatives: Have one, internal  $\theta$ -role.
- Unergatives: Have one, external  $\theta$ -role.

