

chapter 7

Extending X-bar Theory to Functional Categories

Learning Objectives

After reading chapter 7 you should walk away having mastered the following ideas and skills:

1. Identify and distinguish subjects from predicate phrases.
2. Identify various kinds of T and C nodes.
3. Distinguish finite from non-finite clauses, using tests.
4. Identify embedded and root clauses, and distinguish specifier, adjunct or complement clauses.
5. Correctly use X-bar format for DPs, TPs, and CPs in tree drawing.
6. Explain the arguments for DPs, TPs, and CPs.
7. Identify subjects in all types of clauses and correctly place them in the specifier position of TP.

0. INTRODUCTION

In the last chapter, we looked at a refinement of our phrase structure rules that not only accounted for intermediate structure, but also generalized patterns across categories. This refinement is X-bar theory:

- 1) a) *Specifier rule:* $XP \rightarrow (YP) X' \text{ or } XP \rightarrow X' (YP)$
 b) *Adjunct rule:* $X' \rightarrow X' (ZP) \text{ or } X' \rightarrow (ZP) X'$
 c) *Complement rule:* $X' \rightarrow X (WP) \text{ or } X' \rightarrow (WP) X$

These rules not only generate most of the trees we need for the sentences of the world's languages, they also capture the additional properties of hierarchical structure found within the major constituents. This said, you may have noticed that this system is far from perfect. First, there is the status of specifiers. In particular, the specifier rule we proposed above requires that the specifier be a phrase- (XP-)level category. However, the only instances of specifiers we've looked at are determiners, which appear *not* to be phrasal. In this chapter, we will look at determiners and specifiers and propose a new category that fits X-bar theory: a *determiner phrase* (DP). We will see that determiners are not specifiers. Instead, we'll claim that the specifier position is used to mark a particular grammatical function: that of subjects. You'll see that specifiers (of all categories) are where subjects go.

Another troubling aspect of the X-bar theory is the exceptional CP and TP rules that we have yet to incorporate into the system:

- 2) $CP \rightarrow (C) TP$
 $TP \rightarrow NP (T) VP$

These rules do not fit X-bar theory. In the X-bar rules in (1), you'll note that the only obligatory element is the head. In the sentence rules in (2), the opposite is true: the only optional element is the head itself. In this chapter, we will look at how we can modify these so that they fit into the more general pattern.

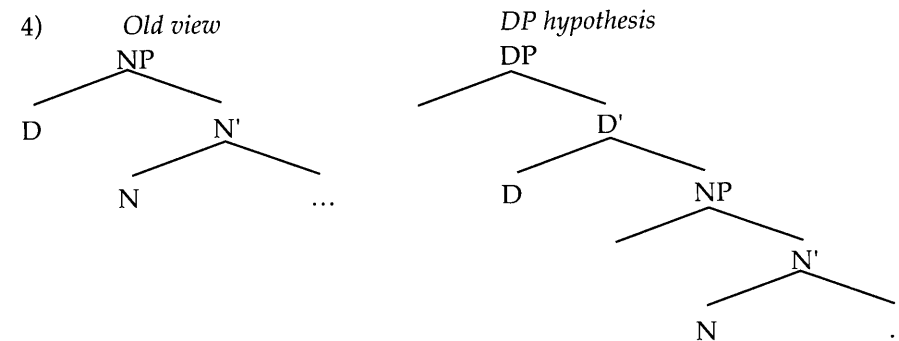
1. DETERMINER PHRASES (DPS)

In the last chapter, for lack of a better place to put them, we put determiners like *the, a, that, this, those, and these* in the specifiers of NPs. This, however, violates one of the basic principles underlying X-bar theory: All non-head material must be phrasal. Notice that this principle is a theoretical rather than an empirical requirement (i.e., it is motivated by the elegance of the theory and not by any data), but it is a nice idea from a mathematical point of view, and it would be good if we could show that it has some empirical basis.

One thing to note about determiners is that they are typically heads. Normally, there can only be one of them in an NP (this isn't true cross-linguistically, but for now let us limit ourselves to English):

- 3) *the that book

In other words, they don't seem to be phrasal. If our requirement says that the only thing that isn't a phrase in an NP is the N itself, then we have a problem. One solution, perhaps not obvious, to this is to claim that the determiner is not actually inside the NP. Instead, it heads its own phrasal projection. This was first proposed by Abney (1987):



Determiners, in this view, are not part of the NP. Instead the NP is the complement to the determiner head. This solution solves the theoretical problem of the non-phrasality of the D, but we still need empirical evidence in its favor.

One piece of evidence comes from the behavior of genitive (possessive) NPs. There are two kinds of possessive NPs. The first is of less interest to us. This one is often called the *free genitive* or *of-genitive*:

- 5) a) the coat of the panther
 b) the roof of the building
 c) the hat of the man standing over there

The free genitive uses the preposition *of* to mark the possessive relation between the two NPs. More important in terms of evidence for DP is the behavior of the other kind of possessive: the *construct* or *'s-genitive*.

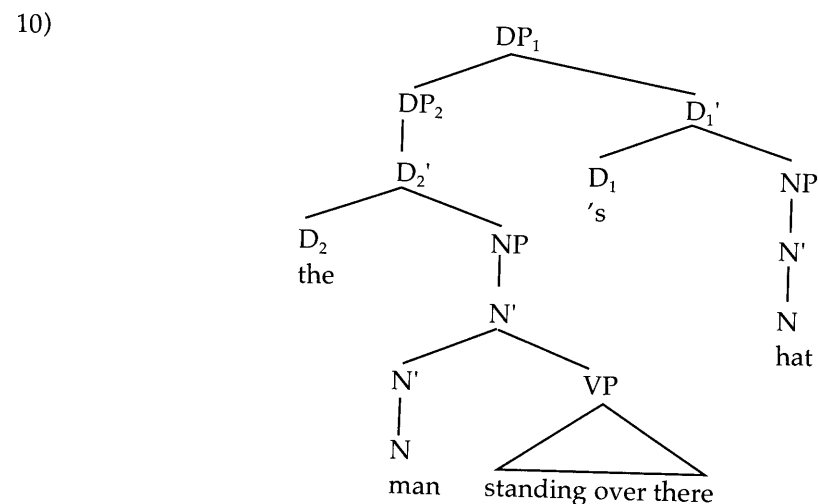
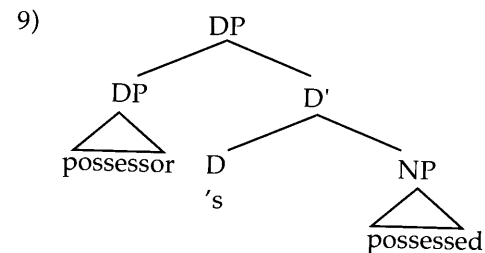
- 6) a) the panther's coat
 b) the building's roof
 c) the man standing over there's hat

There are a couple of important things to note about this construction. Notice first that the 's marker appears after the *entire* possessor NP. For example, it attaches to the whole phrase *the man standing over there* not just to the head *man* (7). This means that 's is not a suffix. Instead it seems to be a small word indicating possession. Next, note that it is in complementary distribution with (i.e., cannot co-occur with) determiners (8).

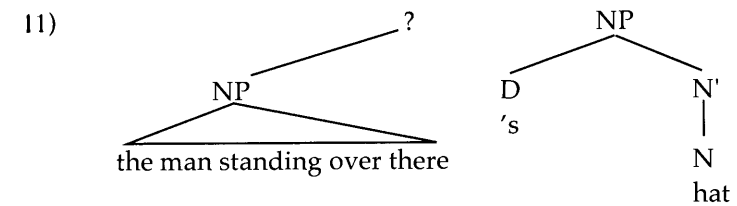
- 7) a) [the man standing over there]'s hat

- b) *the man's standing over there hat
- 8) a) *the building's the roof (cf. the roof of the building)
 b) *the panther's the coat (cf. the coat of the panther)
- c) *the man standing over there's the hat (cf. the hat of the man standing over there)

Unlike the *of*-genitive, the 's-genitive does not allow both the nouns to have a determiner. In other words, 's and determiners are in complementary distribution. As in other domains of linguistics, when two items are in complementary distribution, they are instances of the same thing. (Take for example, phonology, where when two phones are found in different environments – in complementary distribution – then they are allophones of the same phoneme.) Determiners like *the* and 's are different tokens of the same type. Assuming that 's is a determiner, and assuming the DP hypothesis holds true, we can now account for the positioning of the 's relative to the possessor. The 's occupies the head D position, and the possessor appears in its specifier (9–10)



The possessor [DP_2 *the man standing over there*] sits in the specifier of DP_1 , which is headed by 's. So 's follows the whole thing. Notice that with our old theory, in which determiners are specifiers of NP, there is no way at all to generate 's as a determiner and to also have the possessor NP preceding it.



The X-bar rules don't provide any place to attach this pre-determiner NP if determiners are specifiers.

Notice that in the tree in (10) there is a specifier of DP_1 (filled by DP_2). Note further that this specifier is phrasal (projects to an XP). Which means that it meets with our requirement that all non-head material be phrasal.

You might ask if by moving determiners out of the specifier we have completely destroyed the empirical justification for the specifier rule. Actually, we haven't. Again if you look closely at the tree in (10) we still have a specifier, it just isn't D_1 . Instead, it is the DP possessor (DP_2). Further, as we will see below, there are other related uses for the specifier positions. In particular, we will come to associate specifiers with subjects of various kinds of constituents.

You now have enough information to try WBE 1 & 2, GPS 1, and CPS 1 & 2.

2. A DESCRIPTIVE TANGENT INTO CLAUSE TYPES

A *clause* is essentially a *subject* (usually a DP that has the property indicated by the predicate; this is what the clause is about) and a *predicate phrase* (a group of words that assign a property to the subject). The most obvious kind of clause is the simple sentence. In the following examples, the subject is indicated in italics and the predicate phrase is in bold:

- 12) a) *The boy* **ran.**
 b) *Howard* **is a linguistics student.**

As we'll see below, there are many other kinds of clauses. But we can use this as a working definition.

A clause that stands on its own is called a *root*, *matrix*, or *main clause*. Sometimes, however, we can find examples of clauses within clauses. Examples of this are seen below:

- 13) a) [Peter said [that Danny danced]].
 b) [Bill wants [Susan to leave]].

In each of these sentences there are two clauses. In sentence (13a), there is the clause (*that*) *Danny danced*, which is inside the root clause *Peter said that Danny danced*. In (13b), we have the clause *Susan to leave*, which has the subject *Susan*, and the predicate phrase (*to*) *leave*. This is contained within the main clause *Bill wants Susan to leave*. Both of these clauses within clauses are called *embedded clauses*. Another name for embedded clause is *subordinate clause*. The clause containing the embedded clause is still called the *main* or *root clause*.

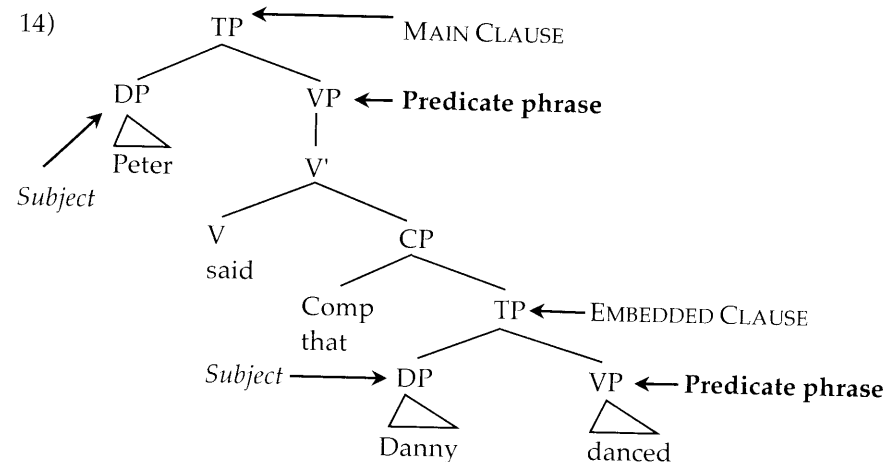
Embedded Clauses are Part of Main Clauses

A very common error among new syntacticians is to forget that embedded clauses are contained *within* main clauses. That is, when faced with identifying what the main clause is in a sentence like (i), most students will properly identify the embedded clause as (*that*) *Cathy loves him*, but will mistakenly claim that the main clause is only *Peter thinks*.

- i) Peter thinks that Cathy loves him.

This is completely incorrect. *Peter thinks* is not a constituent. The main clause is *everything* under the root TP node. So the main clause is *Peter thinks that Cathy loves him*. Be very careful about this.

Using the TP and CP rules we developed in chapter 3, the structure of a root clause containing an embedded clause is given below (I've obscured the irrelevant details with triangles):



In addition to the distinction between main and embedded clauses, we can also distinguish among specifier, complement, and adjunct clauses. Here are some examples of *complement clauses*:

- 15) a) Heidi said [that Art loves peanut butter].
 b) Colin asked [if they could get a mortgage].

These complement clauses (CPs) are sisters to the verb, and thus complements. Clauses can also appear in adjunct positions. Relative clauses are one example of *adjunct clauses*:

- 16) [The man [I saw get into the cab]] robbed the bank.

The relative clause in (16) [*I saw get into the cab*] modifies the head *man*. *Specifier clauses* are ones that serve as the subject of a sentence (why these are specifiers will be made clear below):

- 17) a) [[People selling their stocks] caused the crash of 1929].
 b) [[For Mary to love that boor] is a travesty].

To summarize, we have two basic kinds of clauses, main and embedded. Embedded clauses are contained within main clauses. Further, there are three types of embedded clauses: specifier clauses, complement clauses, and adjunct clauses. This is summarized in the following table:

18)

Main clauses	Embedded clauses		
	specifier clauses	complement clauses	adjunct clauses

There is another way of dividing up the clause-type pie. We class clauses into two groups depending upon whether they are tensed or not. Clauses with predicates that are tensed are sometimes called (obviously) *tensed clauses*, but you may more frequently find them called *finite clauses*. Clauses without a tensed verb are called *tenseless* or *non-finite clauses* (sometimes also *infinitival clauses*).¹

- 19) a) I said [that Mary signed my yearbook]. *tensed or finite*
 b) I want [Mary to sign my yearbook]. *tenseless or non finite*

There are a number of tests for distinguishing finite from non-finite clauses. These tests are taken from Radford (1988). The embedded clause in sentence (20a) is tensed. The one in (20b) is untensed. I have deliberately selected

¹ In many languages, the form of a verb found in a non-finite clause is called the *infinitive*. In English, infinitives are often marked with the auxiliary *to*, as in *to sign*.

a verb that is ambiguous between tensed and untensed in terms of its morphology (suffixes) here as an illustration:

- 20) a) I know [you eat asparagus]. *finite*
 b) I've never seen [you eat asparagus]. *non-finite*

One way to tell if a clause is finite or not is to look for agreement and tense morphology on the verb. These include the *-s* ending associated with third person nouns (*he eats*) and the past tense suffixes like *-ed*. The above examples don't show any such suffixes. However, if we change the tense to the past a difference emerges:

- 21) a) I know you ate asparagus. *finite*
 b) *I've never seen you ate asparagus. *non-finite*

Finite clauses allow past tense morphology (the *ate* form of the verb *eat*); non-finite clauses don't. The same effect is seen if you change the person of the subject in the embedded clause. Third person subjects trigger the *-s* ending. This is allowed only in finite clauses.

- 22) a) I know he eats asparagus. *finite*
 b) *I've never seen him eats asparagus. *non-finite*

The case on the subject of the noun is often a giveaway for determining whether or not a clause is finite. Case refers to the notions *nominative* and *accusative* introduced in chapter 1, repeated here:

23)

	Nominative		Accusative		Anaphoric	
	Singular	Plural	Singular	Plural	Singular	Plural
1	I	we	me	us	myself	ourselves
2	you	you	you	you	yourself	yourselves
3 masc	he	they	him	them	himself	themselves
3 fem	she		her		herself	
3 neut	it		it		itself	

If the clause is finite, then a subject pronoun will take the nominative case form:

- 24) I know *he* eats asparagus. *finite*

If the clause is non-finite then the subject will take the accusative form:

- 25) I've never seen *him* eat asparagus. *non-finite*

One test that works most of the time, but is not as reliable as the others, is to see if the subject is obligatory. If the subject is obligatory, then the clause is finite. If the subject is optional, or is not allowed at all, then it is non-finite.

(Note: this test only works for English; in many languages, such as Spanish, subjects of finite clauses are optional.)

- 26) a) I think that he eats asparagus. *finite*
 (cf. *I think that eats asparagus.)
 b) I want (him) to eat asparagus. *non-finite*
 (cf. I want to eat asparagus.)

Another way to tell if a clause is finite or not is by looking at the complementizer. The complementizer *for* is only found with non finite clauses. By contrast, *that* and *if* are only found with tensed clauses:

- 27) a) I wonder if he eats asparagus. *finite*
 b) I think that he eats asparagus. *finite*
 c) [For him to eat asparagus] is a travesty. *non-finite*
 d) I asked for him to eat the asparagus. *non-finite*

As a final test, we can note that finite and non-finite clauses take different kinds of T elements. The T in tensed clauses can contain auxiliaries and modals like *will*, *can*, *must*, *may*, *should*, *shall*, *is*, and *have*. By contrast the only auxiliary allowed in non-finite clauses is *to*.

- 28) a) I think [he *will* eat asparagus].
 b) I want him *to* eat asparagus. (cf. *I want him *will* eat asparagus.)

This last property gets at the heart of the distinction between finite and non finite clauses. In structural terms the difference between a finite and a non finite clause lies in terms of what kind of T the clause has. If a clause is finite it bears some tense feature (like $[\pm\text{PAST}]$ or $[\pm\text{FUTURE}]$). If it is non finite, it doesn't have any of these features. The question of how this works for clauses where there is no auxiliary we'll leave as a bit of a mystery for now, but will return to later in this chapter.

Let's summarize the discussion we've had thus far. We've been looking at a number of terms for describing various kinds of clauses. We defined clauses as a subject and a predicate phrase. We distinguished root or main clauses from embedded clauses. Embedded clauses come in three types: specifier clauses, complement clauses and adjunct clauses. The other dimension along which we can describe clauses is the finite/non finite distinction.

With this terminology under our belts, we'll now turn to the structure of clauses, and see if we can make them fit better into X-bar theory.

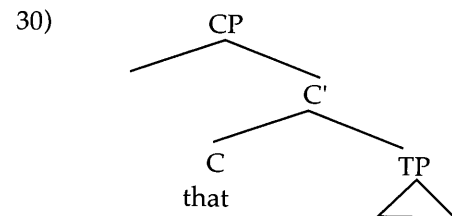
You now have enough information to try WBE 3-8 and GPS 2-4.

3. COMPLEMENTIZER PHRASES (CPs)

We've observed that the TP rule and the CP rule stand out, since they don't fit X-bar theory. In X-bar theory, the head is always obligatory. This is not true of these two rules:

- 29) a) $CP \rightarrow (C) TP$
 b) $TP \rightarrow DP (T) VP$

In fact, it is a fairly trivial matter to change these rules into X-bar-theoretic format. Let us deal with the CP rule first. If we take X-bar theory to extend to CPs, we can assimilate the rule in (29a) to get a tree like that in (30):



This CP structure has C as the head, a TP complement and an empty specifier position (this empty specifier position will become very important later for us when we do *wh*-movement in chapter 11).

We might ask how pervasive this rule is in our mental grammars. That is, do all clauses have CPs, or do only embedded clauses have CPs? On the surface, the answer to this question seems obvious: Only embedded clauses have CPs, since only embedded clauses appear to allow complementizers (31). However, there is evidence that all clauses, even root clauses like (31), require some kind of complementizer.

- 31) a) John thinks that the asparagus is yummy.
 b) *That the asparagus is yummy. (cf. Asparagus is yummy.)

32) Asparagus grows in California.

In particular, we'll claim that some sentences have null complementizers. Don't assume that I'm crazy. No matter how strange this proposal sounds, there is actually some good evidence that this is correct. The structure in (33) shows one of these null complementizers.

- 33) $[_{CP} [_{C} \emptyset] [_{TP} \text{Asparagus grows in California}]]]$

The evidence for this claim comes from cross-linguistic comparison of questions among languages. In particular, we'll focus on *yes/no questions* (see chapter 9 for more discussion on these). These are questions that can

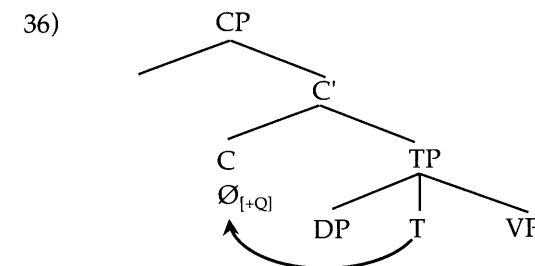
be answered with either *yes*, *no*, or *maybe*. Examples of yes/no questions in English are given below:

- 34) a) Did John leave?
 b) Have you seen Louis?

In English, to form a yes/no question you either insert some form of the verb *do* (*do*, *does*, *did*) before the subject, or you invert the subject and the auxiliary (*You have seen Louis.* \rightarrow *Have you seen Louis?*). This operation is called *subject-aux inversion* (more on this in chapter 9). In many other languages, however, yes/no questions are formed with a complementizer particle that precedes the verb. Take for example, Irish, which indicates yes/no questions with a special particle *Ar* (or its allomorph *An*):

- 35) *Ar thit Seán?*
 Q fall John
 "Did John fall?"

Languages like English that use subject-aux inversion don't have special complementizer question particles. If a language has complementizer question particles, then it won't have subject-aux inversion. The phenomena are in complementary distribution. It seems reasonable to claim, then, that question complementizers and subject-aux inversion are part of the same basic phenomenon. In order to make this concrete, let's make the following proposal: There is a question complementizer particle in English, just like there is in Irish. The difference is that in English this complementizer particle is null (has no phonological content). We will represent this *null complementizer* with the symbol $\emptyset_{[+Q]}$. It has no phonological content, but it must be realized or pronounced somehow. The way English satisfies this requirement is by moving T into the C head:



This results in the correct order, where the auxiliary (in T) now appears before the subject. By contrast, languages like Irish don't utilize this mechanism. Instead they have a particle that fills their [+Q] complementizer (like *Ar/An* in Irish).

English does, in fact, have an overt [+Q] complementizer, but it is only found in embedded questions. This complementizer is *if*. Unsurprisingly, subject-aux inversion is completely disallowed when *if* is present:

- 37) a) Fabio asked if Claus had run a marathon.
 b) *Fabio asked if had Claus run a marathon.
 c) *Fabio asked had if Claus run a marathon.
 d) ?Fabio asked had Claus run a marathon.

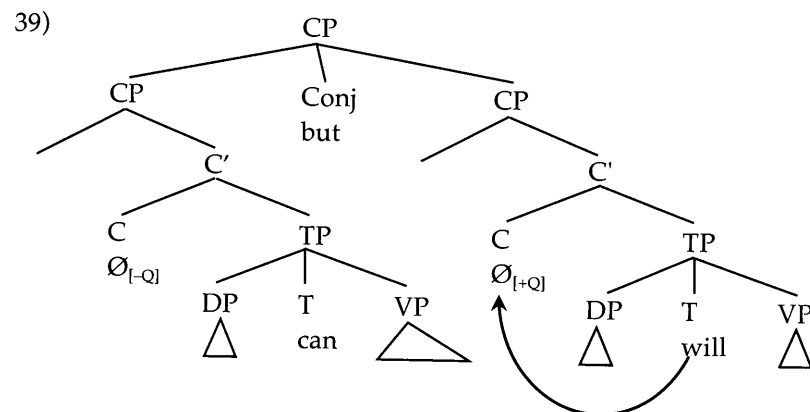
If occupies the [+Q] complementizer, so no subject-aux inversion is required (or allowed).

Given the existence of overt root complementizers in other languages and the evidence that subject-aux inversion patterns like these overt root complementizers, we can conclude that, for questions at least, there are complementizers (and CPs) present even in main clauses.

Of course, we haven't yet shown that non-question sentences have a root complementizer. For this, we need to add an extra step in the argument. You can only conjoin identical categories. If sentences showing subject-aux inversion use a null complementizer and if you can conjoin that question with a non-question (such as a statement), then that statement must also include a (null) complementizer and CP. It is indeed possible to conjoin a statement with a question:

- 38) [You can lead a horse to water] but [will it drink]?

Since the second clause here shows subject-aux inversion, we know there is a $\emptyset_{[+Q]}$ question complementizer present. By extension, we know that the clause it is conjoined with must *also* have a complementizer – this time, a non-question $\emptyset_{[-Q]}$. A CP can only be conjoined with another CP.



This is an argument for null complementizers attached to root clauses, even in simple statements. From this point forward, we will assume that there is

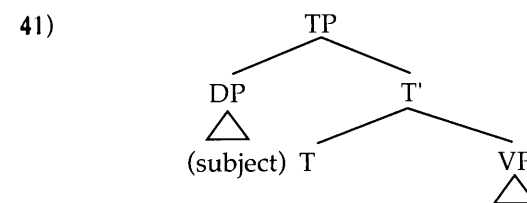
a CP on top of every clause. For brevity's sake, I may occasionally leave this CP off my trees, but the underlying assumption is that it is always there. You should always draw it in when you are drawing your trees.

4. TENSE PHRASES (TPs)

The other rule that doesn't fit the X-bar pattern is our S rule:

- 40) TP → DP (T) VP

Assimilating this rule to X-bar theory results in a structure like the following:



In this tree, the subject DP sits in the specifier of TP, and the VP is the complement. (This is our first clear instance where the notion of specifier corresponds to the notion of subject. We will consider some other cases below.) Again the problem here is that the element that we have designated as the head of the phrase (T) is apparently optional. In X-bar theory, heads are the only obligatory element.

In chapter 2, we equated T with auxiliary verbs. But we might ask what happens in clauses where there is no auxiliary: Is there a TP? Is there a T? Can we make the same claim we did for CPs that the C is obligatory? In order to answer this question, let's make the following observation: Tense inflection on a verb is in complementary distribution with auxiliaries (you never get both of them at the same time):

- 42) a) The roadrunner walks funny.
 b) The roadrunner is walking funny.
 c) *The roadrunner is walks/walkings funny.

Recall that when two elements are in complementary distribution then they are instances of the same category. This means that T can be realized both by auxiliaries and by inflectional endings on verbs. Similar evidence comes from coordination. Recall that you can only coordinate two items that are of the same category and bar level. In the following sentence, we are conjoining a T' that has an auxiliary with a T' that has a tensed verb. The tense inflection and auxiliary are italicized.

43) [_{TP} I [_T [_T kissed the toad] and [_T must go wash my mouth now]]].

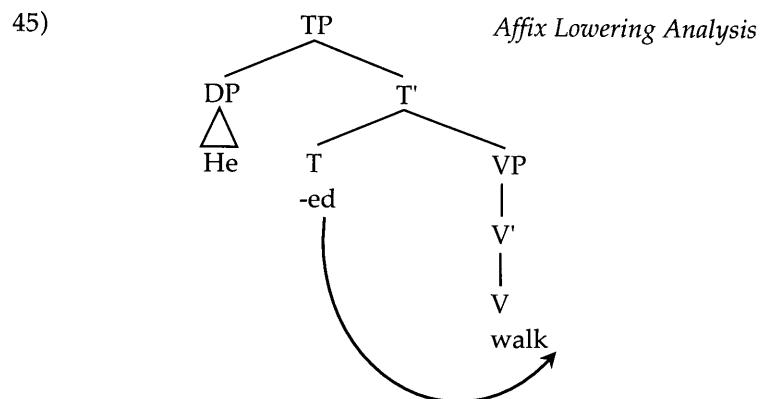
This evidence suggests that the two T's are identical in some deep sense. That is, they both involve a T node: one an auxiliary, the other a tense inflectional ending.

If you think about the basic order of the elements we seem to have argued ourselves into a corner. Auxiliaries appear on the left of verbs, and inflectional suffixes (like *-ed*, and *-s*) appear on the right:

- 44) a) He *will* go.
b) He goes.

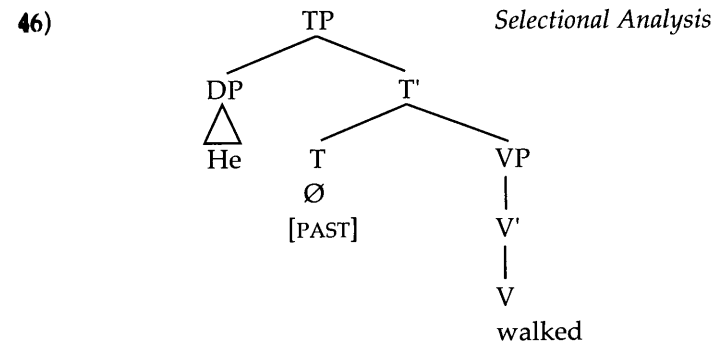
There are other differences between auxiliaries and inflectional suffixes. For example, auxiliaries, but not suffixes, undergo subject-aux inversion. If we are to claim that inflectional suffixes and auxiliaries are both instances of T we have to account for these differences.

There are two possibilities out there in the literature. The older possibility, dating back to Chomsky's famous (1957) book, *Syntactic Structures*, is to claim that both inflectional suffixes and auxiliaries are indeed generated under T. They differ, however, in terms of whether they can stand alone or not. Auxiliaries are independent words and can stand alone. By contrast, suffixes like *-s* and *-ed* have to be attached to a verb. Much like the case of moving T to C in order to pronounce $\emptyset_{[+Q]}$, we might hypothesize that endings like *-s* and *-ed* can't be pronounced in isolation, so they move to attach to the verb. In particular they seem to lower onto the verb. The following tree shows how this would work for the simple sentence *He walked*. This sentence starts out as [*he -ed walk*] and then the *-ed* ending lowers to attach to the end of the verb:



A more modern approach doesn't have the affixes generated in T. Instead, there are two tense markers, both of which are null: $\emptyset_{[PAST]}$ and $\emptyset_{[PRESENT]}$. The $\emptyset_{[PAST]}$ simply requires that its VP complement be in a past tense

(or preterite) form and $\emptyset_{[PRESENT]}$ requires that its VP complement be in a present tense form.



We look at how heads can place restrictions on their complements and specifiers in the next chapter, and then we extend this more modern analysis to the very complicated English auxiliary system in chapter 9.

You can now try WBE 9, GPS 5 & 6, and CPS 3.

IDEAS, RULES, AND CONSTRAINTS INTRODUCED IN THIS CHAPTER

- i) **Determiner Phrase (DP):** D is not in the specifier of NP. D heads its own phrase: [_{DP} [_{D'} D NP]].
- ii) **Complementizer Phrase (CP):** C is the head of CP and is obligatory in all clauses, although sometimes phonologically null: [_{CP} [_C C TP]].
- iii) **Tense Phrase (TP):** T is the head of TP and is obligatory in all clauses. Sometimes it involves lowering of the affix to the V. The subject DP occupies the specifier position: [_{TP} DP_{subject} [_T T VP]].
- iv) **Free Genitive/of-Genitive:** Possessed of the possessor.
- v) **Construct Genitive/'s-Genitive:** Possessor 's possessed.
- vi) **Subject:** A DP that has the property indicated by the predicate phrase. What the sentence is about. In most sentences, this surfaces in the specifier of TP.
- vii) **Predicate Phrase:** A group of words that attributes a property to the subject. (In most sentences this is the VP, although not necessarily so.)
- viii) **Clause:** A subject and a predicate phrase (always a CP in our system).
- ix) **Root, Matrix, or Main Clause:** A clause (CP) that isn't dominated by anything.

- x) **Embedded Clause/Subordinate Clause:** A clause inside of another.
- xi) **Specifier Clause:** An embedded clause in a specifier position.
- xii) **Adjunct Clause:** An embedded clause in an adjunct position.
- xiii) **Complement Clause:** An embedded clause in a complement position.
- xiv) **Tenseless or Non-finite Clause:** A clause that isn't tensed (e.g., I want [*Mary to leave*]).
- xv) **Tensed or Finite Clause:** A clause that is tensed.
- xvi) **Yes/No Question:** A question that can be answered with a *yes*, a *no* or a *maybe*.
- xvii) **Subject-Aux Inversion:** A means of indicating a *yes/no* question. Involves movement of T to $\emptyset_{[+Q]}$ complementizer for morphophonological reasons.
- xviii) **Affix Lowering:** An old analysis of how past and present tense suffixes get on the verb: The lowering of inflectional suffixes to attach to their verb. Now largely replaced by an analysis where T is null and selects for a VP complement that is correctly inflected.

FURTHER READING: Abney (1987), Chomsky (1991), Emonds (1980), Pollock (1989)

GENERAL PROBLEM SETS

GPS1. TREE DRAWING: DPS

[Application of Skills; Basic]

Draw the DP trees for the following phrases:

- a) the kitten
- b) the very orange kitten
- c) the kitten's paw
- d) the paw of the kitten
- e) the kitten's mother's paw
- f) the kitten's left paw

GPS2. SUBJECTS AND PREDICATE PHRASES

[Data Analysis; Basic]

In each of the following clauses identify the subject and the predicate phrase. Some sentences contain multiple clauses. Be sure to identify the subjects and predicate phrases of *all* clauses.

- a) The peanut butter has got² moldy.
- b) The duffer's swing blasted the golf ball across the green.
- c) That Harry loves dancing is evidenced by his shiny tap shoes.
- d) The Brazilians pumped the oil across the river.

² You may prefer *gotten* to *got* here. The choice is dialect-dependent.

GPS3. CLAUSE TYPES

[Data Analysis; Basic]

The following sentences are "complex" in that they contain more than one clause. For each sentence, identify each clause. Remember, main clauses include embedded clauses. Identify the complementizer, the T, and the subject of the clause; be sure to identify even *null* (\emptyset) complementizers and Ts with suffixes in them. State whether each clause is a finite clause or a non-finite clause.

- a) Stalin may think that Roosevelt is a fool.
- b) Lenin believes the Tsar to be a power-hungry dictator.
- c) Brezhnev had said for Andropov to leave.
- d) Yeltsin saw Chernyenko holding the bag.

GPS4. ENGLISH THAT³

[Critical Thinking; Basic]

Discuss the status of the word *that* in each of the following two sentences. Explain the differences between the two sentences. If you assign a different category status to *that* in each sentence, explain why. Draw the tree (use X-bar theory) for each of the sentences.

- a) Robert thinks that students should eat asparagus.
- b) Robert thinks that student should eat asparagus.

GPS5. TREES I

[Application of Skills; Basic to Intermediate]

Draw the trees for the following sentences. Use X-bar theory and show all CPs, DPs, and TPs.

- a) The very young child walked from school to the store.
- b) Linguistics students like phonetics tutorials.
- c) John paid a dollar for a head of lettuce.
- d) Teenagers drive rather quickly.
- e) Martha said that Bill loved his Cheerios in the morning.
- f) Eloise wants you to study a new language. [assume *to* = T]
- g) For Maurice to quarrel with Joel frightened Maggie.
- h) John's drum will always bother me.

GPS6. TREES II

[Application of Skills; Basic to Intermediate]

- 1) Go back to chapter 3, GPS 1 and GPS 2 and draw the trees using X-bar theory, including DPs.
- 2) Go back to chapter 3, GPS 6, and draw the trees using X-bar theory, including DPs, TPs, and CPs.

³ Thanks to Eithne Guilfoyle for contributing this problem set.

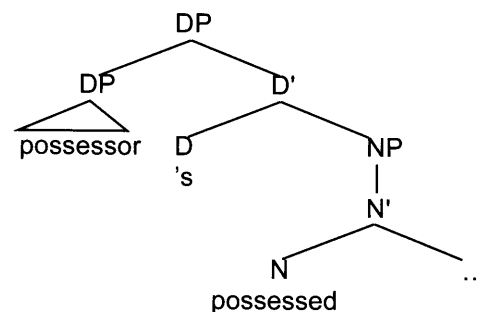
- 3) Go back to chapter 4, GPS 1, and draw the trees using X-bar theory, including DPs, TPs, and CPs.

CHALLENGE PROBLEM SETS

CHALLENGE PROBLEM SET 1: HUNGARIAN DPs

[Data Analysis; Challenge]

In the text above, we argued that the structure of genitive constructions in English looks like:



Consider the following data from Hungarian. Does the possessor DP appear in the same place as the English ones? Assume the determiners *az* and *a* modify the *possessed* noun, not the possessor. The ending on the word *kalapom/kalapod/kalapja* varies depending upon the possessor. This does not affect the answer to this question. (Data from Szabolcsi 1994.)

- a) *az én kalapom*
the I hat-1SG
"my hat"
- b) *a te kalapod*
the you hat-2SG
"your hat"

Hungarian has another possessive construction, seen in (c).

- c) *Marinak a kalapja*
Mary the hat-3SG
"Mary's hat"

Where is the possessor DP in (c)? Explain your answer.

CHALLENGE PROBLEM SET 2: NPI LICENSERS

[Data Analysis and Critical Thinking; Challenge]

The adverb *ever* is a negative polarity item. Negative polarity items must stand in a c-command relationship with a negative licensor. Assume that the properties of the head uniquely determine the properties of a phrase. Explain

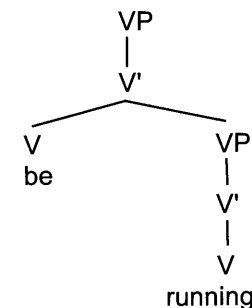
how the following sentences are an argument for the subject being a DP rather than an NP:

- a) No man has ever beaten the centaur.
b) *Some man has ever beaten the centaur.
c) *Every man has ever beaten the centaur.

CHALLENGE PROBLEM SET 3: ENGLISH MODALS AND AUXILIARIES

[Data Analysis and Critical Thinking; Challenge]

In traditional grammar, two different kinds of T are found: modals and auxiliaries. Modals include words like *can*, *must*, *should*, *would*, *could*, *may*, *will* and in some dialects *shall*. Auxiliary verbs, by contrast, include such words as *have* and *be*. In this book, we've treated both modals and auxiliaries as T. An alternative is that only modals are really of category T, and that auxiliaries are real verbs. Auxiliary and verb combinations are actually a stacked set of VPs:



Construct an argument in favor of the idea that modals are of category T, but auxiliaries are really verbs. Assume the following: You may have as many V categories as you like, but there is only one T in any tensed clause's tree.

chapter 8

Constraining X-bar: Theta Theory

Learning Objectives

After reading chapter 8 you should walk away having mastered the following ideas and skills:

- 1) Distinguish between thematic relation and theta role.
- 2) Identify the thematic relations agent, theme, goal, source, experiencer, location, instrument, recipient, benefactor.
- 3) Explain how X-bar theory over-generates.
- 4) Explain the structure of the lexicon.
- 5) Draw the theta grids for a predicate.
- 6) Apply the theta criterion to a sentence as a filter to X-bar theory.
- 7) Distinguish sentences with expletive subjects from ones with theta-role-bearing subjects.
- 8) Explain the Extended Projection Principle and how it accounts for expletives.
- 9) Explain the ordering of the EPP with the theta criterion in the context of the model we are developing.

0. INTRODUCTION

In chapters 6 and 7, we developed a very simple and general theory of phrase structure: X-bar theory. Using only three rules, this theory accounts

for the distinction between adjuncts, complements, and specifiers. It incorporates the more articulated view of sentence hierarchy required by constituency tests, and it captures cross-categorical generalizations (i.e., the fact that all kinds of phrases – NPs, VPs, APs, PPs, CPs, DPs, and TPs – have the same basic properties). Most importantly, it allows us to draw trees for most of the sentences of any language.

This said, there is a significant problem with X-bar theory: it also generates sentences that are not acceptable or grammatical. Take for example the following pairs of grammatical and ungrammatical sentences:

- 1) a) Rosemary hates New York.
b) *Rosemary hates.
- 2) a) Jennie smiled.
b) *Jennie smiled the breadbox.
- 3) a) Traci gave the whale a jawbreaker.
b) *Traci gave the whale.
c) *Traci gave a jawbreaker.

Sentence (1b) should be perfectly acceptable (compare it to *Rosemary ran*). X-bar theory says that complements are optional. Therefore, direct objects, which are complements, should always be optional. The opposite type of fact is seen in the pair in (2). X-bar theory optionally allows a complement. So having a direct object here should be fine too. The same kind of effect is seen in (3), where both the direct object and indirect object are obligatory – contra X-bar theory.

What seems to be at work here is that certain verbs require objects and others don't. It appears to be a property of the *particular* verb. Information about the peculiar or particular properties of verbs is contained in our mental dictionary or *lexicon*. In this chapter, we'll look at how we can use the lexicon to constrain X-bar theory so that it doesn't predict the existence of ungrammatical sentences.

1. SOME BASIC TERMINOLOGY

In chapter 2, we discussed how different verb types take a different number of arguments. For example, an intransitive verb like *leave* takes a single DP, which is the subject. A transitive verb such as *hit* takes a DP subject and a DP object. Below are the subcategories we came up with in chapter 2 (substituting DP for NP):

Subcategory	Example
$V_{[DP_]}$ (intransitive)	<i>leave</i>
$V_{[DP_ DP]}$ (transitive type 1)	<i>hit</i>
$V_{[DP_ \{DP/CP\}]}$ (transitive type 2)	<i>ask</i>
$V_{[DP_ DP DP]}$ (ditransitive type 1)	<i>spare</i>
$V_{[DP_ DP PP]}$ (ditransitive type 2)	<i>put</i>
$V_{[DP_ DP \{DP/PP\}]}$ (ditransitive type 3)	<i>give</i>
$V_{[DP_ DP \{DP/PP/CP\}]}$ (ditransitive type 4)	<i>tell</i>

In addition to these restrictions, we also find semantic restrictions on what can appear in particular positions:

- 5) a) #My comb hates raisinettes.
b) #A bolt of lightning killed the rock.

There is something decidedly strange about these sentences. Combs can't hate anything and rocks can't be killed. These semantic criteria are called *selectional restrictions*.

In the next section, we'll look at the theory of thematic relations, which is a particular way of representing selectional and subcategorizational restrictions.

2. THEMATIC RELATIONS AND THETA ROLES

One way of encoding selectional restrictions is through the use of what are called *thematic relations*. These are particular semantic terms that are used to describe the role that the argument plays with respect to the predicate. This section describes some common thematic relations (this list is by no means exhaustive, and the particular definitions are not universally accepted).

The initiator or doer of an action is called the *agent*. In the following sentences, *Ryan* and *Michael* are agents.

- 6) a) *Ryan* hit Andrew.
b) *Michael* accidentally broke the glass.

Agents are most frequently subjects, but they can also appear in other positions.

Arguments that feel or perceive events are called *experiencers*. Experiencers can appear in a number of argument positions, including subject and object:

- 7) a) *Leah* likes cookies.
 b) *Lorenzo* saw the eclipse.
 c) Syntax frightens *Kenny*.

Experiencers are normally only found with verbs that involve a psychological component or express a notion that can be felt by a living being. For example, the subjects of verbs of perception (*see, perceive, hear, taste, feel, smell, etc.*), subjects and objects of verbs of emotion (*frighten, fear, dishearten, etc.*), and verbs of cognition (*know, understand, etc.*), among others can be experiencers. There is a temptation among new syntacticians to extend the notion "experiencer" to all sorts of arguments. For example, in the sentence *The rock fell on Terry*, does Terry experience the event? Of course he does, but he's not an "experiencer" in the technical sense we are using here. What about *the wall* in *The car hit the wall*? We want to limit the experiencer thematic relation to entities that can experience events, and only with predicates where the experience is a critical part of the meaning of the verb (like those listed above). In the end, as we'll see below, the exact thematic relations involved for a given predicate are not critical to the syntactic analysis we give these verbs, but it's good to be as precise as we possibly can be.

In some languages, the difference between agents and experiencers is marked grammatically on the nouns (usually as case markers), or sometimes with morphology on the verb. Take Korean¹ for example. If you have a verb like *culkew* which roughly means "enjoy", the normal morphology has the subject marked with dative case (8) – which is surprising since subjects typically take the nominative. This is the morphology typically associated with experiencer predicates.

- 8) Ku yeca-ekey sopwung-i culkew-ess-ta
 she.DAT picnic-NOM enjoy-PAST-DECL
 "She enjoyed the picnic."

But when the verb is marked with the suffix *eha*, meaning roughly "do", the case marking on the arguments shifts and the subject is marked with nominative case as in (9). Korean speakers report that sentences like (9) have a more agentive feel, and that the subject is overtly expressing her enjoyment, rather than just feeling it inside.

- 9) Ku yeca-ka sopwung-lul culkew-eha-ess-ta
 she.NOM picnic-ACC enjoy-do-PAST-DECL
 "She enjoyed the picnic."

Entities that undergo actions or are moved, experienced, or perceived are called *themes*.

- 10) a) Alyssa kept *her syntax book*.
 b) The arrow hit *Ben*.
 c) The syntactician hates *phonology*.

The entity towards which motion takes place is called a *goal*. Goals may involve abstract motion.

- 11) a) Doug went *to Chicago*.
 b) *Dave* was given the piña colada mix.
 c) An evil thought struck *Dave*.

There is a special kind of goal called *recipient*. Recipients only occur with verbs that denote a change of possession:

- 12) a) Mikaela gave *Jessica* the book.
 b) *Daniel* received a scolding from Hanna.

The opposite of a goal is the *source*. This is the entity from which a motion originates:

- 13) a) *Bob* gave Steve the syntax assignment.
 b) Stacy came directly *from sociolinguistics class*.

The place where the action occurs is called the *location*:

- 14) a) Andrew is *in Tucson's finest apartment*.
 b) We're all *at school*.

The object with which an action is performed is called the *instrument*:

- 15) a) Chris hacked the computer apart *with an axe*.
 b) *This key* will open the door to the linguistics building.

Finally, the one for whose benefit an event took place is called the *beneficiary*:

- 16) a) He bought these flowers *for Aaron*.
 b) She cooked *Matt* dinner.

You now have enough information to try WBE 1–7 and GPS 1 & 2.

¹ Thanks to Hyun-Kyoung Jung for helpful discussion of these facts.

Notice that any given DP can have more than one thematic relation. In the following sentence, the DP *Jason* bears the thematic relations of agent and source (at the very least).

17) *Jason* gave the books to Anna.

There is no one-to-one relationship between thematic relations and arguments. However, linguists have a special construct called a *theta role* (or *θ role*) that does map one-to-one with arguments. Theta roles are bundles of thematic relations that cluster on one argument. In (17) above, *Jason* gets two thematic relations (agent and source), but only one theta role (the one that contains the agent and source thematic relations). Somewhat confusingly, syntacticians often refer to particular theta roles by the most prominent thematic relation that they contain. So you might hear a syntactician refer to the "agent theta role" of [_{DP} *Jason*]. Strictly speaking, this is incorrect: Agent refers to a thematic relation, whereas the theta role is a bundle of thematic relations. But the practice is common, so we'll do it here. Remember, thematic relations are things like agent, theme, goal, etc., but theta roles are bundles of thematic relations assigned to a particular argument.

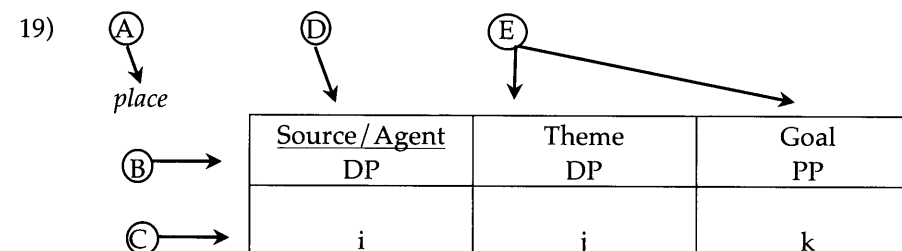
Let's now see how we can use these theta roles to represent the argument structure of a verb. Take a ditransitive verb like *place*. *Place* requires three arguments: a subject that must be an agent (the placer), a direct object, which represents the theme (the thing being placed), and an indirect object, which represents a location or goal (the thing on which the theme is being placed). Any variation from this results in ungrammaticality:

- 18) a) John placed the flute on the table.
 b) *placed the flute on the table.
 c) *John placed on the table.
 d) *John placed the flute.
 e) *John placed the flute the violin on the table.²
 f) *The rock placed the sky with the fork.
 g) *John placed the flute the table.

Examples (18b–e) show that having either too many or too few arguments results in ungrammaticality. Example (18f) shows that using DPs with the wrong theta roles does the same (*the rock* can't be an agent; *the sky* can't be a theme – it can't be given to anyone; and *with the fork* is an instrument,

² This sentence would be okay if there were a conjunction between *the flute* and *the violin*. What does this tell us about what conjunction does to theta roles?

not a goal). (18g) shows us that the category of the argument is important (this we already knew from chapter 2) – the goal argument of the verb *place* must be a PP. It appears as if the verb *place* requires three arguments, that bear precisely the theta roles of agent (DP), theme (DP), and goal (PP). We represent this formally in terms of what is called a *theta grid*.³



This grid consists of several parts. First of all, we have the name of the predicate (A). Next, for each argument that the predicate requires, there is a column (with two rows). Each of these columns represents a theta role. Notice that a column can have more than one thematic relation in it (but only one theta role). The number of columns corresponds exactly to the number of arguments the predicate requires. The first row (B) tells you the thematic relations and the categories associated with each of these theta roles. The second row (C) gives you what are called indices (singular: index) for each theta role. These are *not* the same as the indices in binding theory. When a predicate appears in an actual sentence, we mark the DP bearing the particular theta role with that index. Applying our grid to sentence (18), we get the following indexed sentence:

20) John_i placed [the flute]_j [on the table]_k.

The _i index maps the agent theta role to *John*. The _j index maps the theme theta role to *the flute*, etc.

Theta roles actually come in two types. The first is the *external theta role* (D). This is the one assigned to the subject. External theta roles are usually indicated by underlining the name of the theta role in the theta grid (e.g., Source/Agent in (19)). The other kind are *internal theta roles* (E). These are the theta roles assigned to the object and indirect object. There is a semantic reason for the distinction between internal and external theta roles (see Marantz 1984 for extensive discussion), but we will leave that issue aside here. We will have use for the external/internal distinction in chapter 10,

³ There are many ways to formalize theta grids, but I adopt here the indexing box method that Haegeman (1994) uses, since it seems to be the most transparent.

when we do DP movement. For now, however, you should simply indicate which argument is the subject by underlining its name.

If you look carefully at the theta grid in (19) you'll notice that it only contains a specifier (subject) and complements (direct object and indirect object). There are no adjuncts listed in the theta grid. Adjuncts seem to be entirely optional:

- 21) a) John put the book on the table (with a pair of tongs). *instrument*
 b) (In the classroom) John put the book on the table. *location*

This corresponds to our observation in chapter 6 that you can have as many or as few adjuncts as you like, but the number of complements and specifiers is more restricted. *Adjuncts are never arguments, and they never appear in theta grids.*

You can now try WBE 8 and GPS 3 & 4.

Up until now, we have been representing our grammar solely through the mechanism of rules (phrase structure, then X-bar rules). In order to stop X-bar rules from over-generating, we need a constraint. Constraints are like filters. They take the output of rules, and throw away any that don't meet the constraint's requirements. In essence, we are going to allow the X-bar rules to wildly over-generate, and produce ungrammatical sentences. Those sentences, however, will be thrown out by our constraint. The constraint we are going to use is called the *theta criterion*. The theta criterion ensures that there is a strict match between the number and types of arguments in a sentence and the theta grid.

22) *The Theta Criterion*

- a) Each argument is assigned one and only one theta role.
 b) Each theta role is assigned to one and only one argument.

This constraint requires that there is a strict one-to-one match between argument DPs and theta roles. You can't have more arguments than you have theta roles, and you can't have more theta roles than you have DPs. Furthermore, since theta roles express particular thematic relations, the arguments will have to be of appropriate semantic types for the sentence to pass the constraint.

Let's look at some examples to see how this works. Consider the verb *love*. It has the theta grid given in (23). I haven't written in the indices here, because we'll add them when we compare the grid to a particular sentence.

23) *love*

Experiencer	Theme
DP	DP

When a sentence containing the predicate *love* is produced, we apply indices to each of the arguments and match those arguments to theta roles in the grid. The sentence in (22) is grammatical with the correct number of arguments. It is matched to the theta grid in (23). There is a one to one matching between arguments and theta roles. So the theta criterion is satisfied, and the sentence is allowed to pass through the filter and surface.

24) Megan_i loves Kevin_j.

25) *love*

Experiencer	Theme
DP	DP
i	j

Contrast this with the ungrammatical sentence in (24):

26) *Megan_i loves.

This sentence lacks a theme argument, as seen in the following theta grid:

27) *love*

Experiencer	Theme
DP	DP
i	

The theme theta role is not assigned to an argument (there is no index in its lower box). This violates the second condition of the theta criterion: Every theta role is assigned to an argument. There is not a one-to-one matching of the theta roles to the arguments in this sentence. Since the theta criterion is violated, the sentence is filtered out (marked as ungrammatical). Notice, our X-bar rules *can* generate this sentence; it is ruled ungrammatical by our constraint.

The next sentence shows the opposite problem: A sentence with too many arguments.

28) *Megan_i loves Jason_j Kevin_k.

29) *love*

Experiencer	Theme	
DP	DP	
i	j	k

Here, the argument *Kevin* doesn't get a theta role. There are only two theta roles to be assigned, but there are three arguments. This violates the first part of the theta criterion: the requirement that every argument have a theta role. Again, the theta criterion filters out this sentence as ungrammatical.

To summarize, we can constrain the output of the X-bar rules using a semantic tool: theta roles. The theta criterion is a constraint or filter that rules out otherwise well-formed sentences. The theta criterion requires that there be a strict one-to-one matching between the number and kind of theta roles and the number and kind of arguments.

You now have enough information to try WBE 9, GPS 5, and CPS 1.

3. THE LEXICON

Let's take a step back from these details and look at the big picture. We have developed a model of grammar where we have three simple rules (the X-bar rules) that can generate a hierarchical constituent structure. These rules are constrained by the theta criterion, which uses the semantic notion of theta roles. Recall that our theory of syntax is meant to be a cognitive theory, so let's consider the question of where these rules and these theta roles are stored in the mind. Chomsky proposes that the part of the mind devoted to language is essentially divided into two parts. One part, which he calls the *computational component*, contains all the rules and constraints. This part of the mind does the work of building sentences and filtering out any ill-formed ones. The computational component can't work in a vacuum, however. It needs access to information about theta roles and the like. Chomsky claims that this information is stored in the *lexicon*, the other part of the human language faculty. The lexicon is your mental dictionary or list of words (and their properties). If you think about it, this is the obvious place for theta grids to be stored. Which theta role is assigned to which argument is a property of each predicate. It is information that must be associated with that predicate and that predicate only. The obvious place to store information about particular words (or more properly *lexical items*) is in the lexicon.

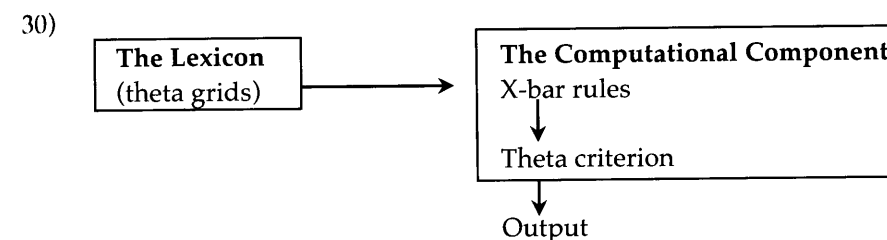
The lexicon contains all the irregular and memorized parts of language. Each lexical entry (dictionary entry) must contain at least the following information):

- the meaning of the word
- the syntactic category of the word (N, V, A, P, T, C, etc.)
- the pronunciation of the word

- exceptional information of all kinds (such as morphological irregularities)
- the theta grid (argument structure).

When you learn a new word, you memorize all this information.

On an abstract level we can diagram the grammatical system as looking something like:



The lexicon feeds into the computational component, which then combines words and generates sentences. The fact that lexical information affects the form of the sentence is formalized in what we call the *Projection Principle*:

31) *The Projection Principle*

Lexical information (such as theta roles) is syntactically represented at all levels.

4. EXPLETIVES AND THE EXTENDED PROJECTION PRINCIPLE

Before leaving the topic of the lexicon, I'd like to point out two special classes of predicates. Consider first the following "weather" verbs. These predicates don't seem to assign any theta roles:

- 32) a) It rained.
b) It snowed.
c) It hailed.

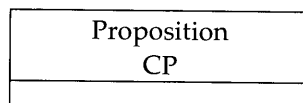
What theta role does the pronoun *it* get in these sentences? If you are having a problem figuring this out, ask yourself what *it* refers to in the above sentences. It appears as if *it* doesn't refer to anything. In syntax, we refer to pronouns like this as *expletive* or *pleonastic pronouns*. These pronouns don't get a theta role (which of course is a violation of the theta criterion – a point we will return to below). The theta grid for weather verbs is empty. They don't assign any theta roles.

There is another class of predicates that take expletive pronouns. These are predicates that optionally take a CP subject:

33) [_{CP} That Bill loves chocolate] is likely.

The predicate *is likely* assigns one theta role. It takes one argument (the clause). (We will tentatively notate these clausal arguments with the theta role *proposition*, but will refine this in the next chapter.)

34) *is likely*



You'll note that in (34) the theta role is not underlined. This is because the clause bearing the theta role of proposition is a complement. This can be seen in the following example:

35) It is likely that Bill likes chocolate.

In this sentence, we again have an expletive *it*, which gets no theta role.

In order to maintain the theta criterion, we need to account for these expletive DPs without theta roles. Expletive pronouns usually appear in subject position. When *it* appears in other positions, it usually bears a theta role:

- 36) a) I love *it*. (*it* is a theme)
 b) I put a book on *it*. (*it* is a goal or location)

Expletives seem to appear where there is no theta marked DP (or CP) that fills the subject position. This is encoded in a revised version of the Projection Principle: The *Extended Projection Principle* (EPP):

37) *Extended Projection Principle* (EPP)

All clauses must have subjects (i.e. the specifier of TP must be filled by a DP or CP) and lexical information is expressed at all levels.

The EPP works like the theta criterion. It is a constraint on the output of the X-bar rules. It requires that every sentence have a subject. Next, we must account for the fact that expletives violate the theta criterion.

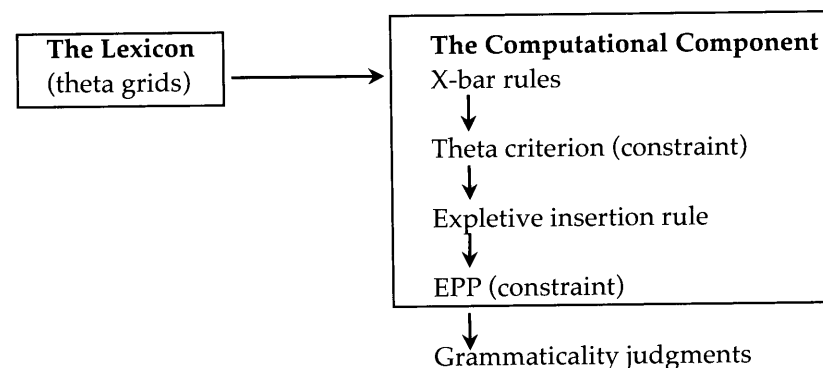
One way of doing this is by claiming that expletives are not generated by the X-bar rules. Instead, they are inserted by a special *expletive insertion* rule:

38) *Expletive insertion rule*

Insert an expletive pronoun into the specifier of TP.

This rule applies when there is no other subject. If there is no theta marked subject and no expletive subject, then the EPP will filter the sentence out.

The way in which we get around the theta criterion is by *ordering* the expletive insertion rule after the theta criterion has applied.



Since expletives are inserted *after* the theta criterion has applied, they can't be filtered out by it.

The model we've drawn here is very preliminary. In the next chapter, we will introduce a new kind of rule (the transformation – of which expletive insertion is a very special case) that will cause us to significantly revise this diagram.

Two Kinds of It

There are two *it* pronouns in English. One is the expletive found with weather verbs. The other is the neuter pronoun *it* found in sentences like:

- i) It bit me on the leg.

If you contrast the *it* in (i) with the ones in the weather verbs, you'll see that the *it* in (i) does take a theta role (agent) and does refer to something (probably an insect or some other animal). Not every sentence with an *it* involves an expletive.

You now have enough information to try WBE 10, GPS 6, and CPS 2–5.

5. CONCLUSION

We started this chapter off with the observation that while X-bar rules capture important facts about constituency and cross-categorical generalizations, they over-generate (that is, they generate ungrammatical sentences). One way of constraining X-bar theory is by invoking lexical restrictions on sentences, such that particular predicates have specific argument structures, in the form of theta grids. The theta criterion rules out

any sentence where the number and type of arguments don't match up one to one with the number and type of theta roles in the theta grid.

We also looked at one apparent exception to the theta criterion: theta-role-less expletive pronouns. These pronouns only show up when there is no other subject, and are forced by the EPP. They escape the theta criterion by being inserted after the theta criterion has filtered out the output of X-bar rules.

By using lexical information (like theta roles) we're able to stop the X-bar rules from generating sentences that are ungrammatical. Unfortunately, as we'll see in the next chapter, there are also many sentences that the X-bar rules *cannot* generate. In order to account for these, we'll introduce a further theoretical tool: the movement rule.

IDEAS, RULES, AND CONSTRAINTS INTRODUCED IN THIS CHAPTER

- i) **Selectional Restrictions:** Semantic restrictions on arguments.
- ii) **Thematic Relations:** Semantic relations between a predicate and an argument – used as a means of encoding selectional restrictions.
- iii) **Agent:** The doer of an action (under some definitions must be capable of volition).
- iv) **Experiencer:** The argument that perceives an event or state.
- v) **Theme:** The element that is perceived, experienced, or undergoing the action or change of state
- vi) **Goal:** The end point of a movement.
- vii) **Recipient:** A special kind of goal, found with verbs of possession
- viii) **Source:** The starting point of a movement.
- ix) **Location:** The place where an action or state occurs.
- x) **Instrument:** A tool with which an action is performed.
- xi) **Beneficiary:** The entity for whose benefit the action is performed.
- xii) **Proposition:** The thematic relation assigned to clauses.
- xiii) **Theta Role:** A bundle of thematic relations associated with a particular argument (DPs, PPs, or CPs).
- xiv) **Theta Grid:** The schematic representation of the argument structure of a predicate, where the theta roles are listed.
- xv) **External Theta Role:** The theta role associated with subjects.
- xvi) **Internal Theta Role:** The theta role associated with other arguments.
- xvii) **The Theta Criterion:**
 - a) Each argument is assigned one and only one theta role.
 - b) Each theta role is assigned to one and only one argument.
- xviii) **Lexical Item:** Another way of saying "word". A lexical item is an entry in the mental dictionary.

- xix) **The Projection Principle:** Lexical information (like theta roles) is syntactically represented at all levels.
- xx) **Expletive (or Pleonastic) Pronoun:** A pronoun (usually *it* or *there*) without a theta role. Usually found in subject position.
- xxi) **Extended Projection Principle (EPP):** All clauses must have subjects. Lexical information is syntactically represented.
- xxii) **Expletive Insertion:** Insert an expletive pronoun into the specifier of TP.
- xxiii) **The Lexicon:** The mental dictionary or list of words. Contains all irregular and memorized information about language, including the argument structure (theta grid) of predicates.
- xxiv) **The Computational Component:** The combinatorial, rule-based part of the mind. Where the rules and filters are found.

FURTHER READING: Gruber (1965), Haegeman (1994), Marantz (1984), Williams (1980, 1994)

GENERAL PROBLEM SETS

GPS1. IDENTIFYING THEMATIC RELATIONS

[Data Analysis and Application of Skills; Basic]

Part 1: Identify the thematic relations associated with each DP or embedded CP in the following sentences. Each DP or CP may have more than one thematic relation associated with it.

- a) Shannon sent Dan an email.
- b) Jerid thinks that Sumayya cooked some beef waffles for him.
- c) Stacy hit a baseball to Yosuke.
- d) Jaime danced a jig.
- e) Yuko rubbed the pizza with a garlic clove.
- f) It is raining in San Francisco.

Part 2: Draw the trees for (b–f). Use CPs, DPs, and TPs.

GPS2. WARLPIRI⁴

[Data Analysis; Basic]

Consider the following data from Warlpiri:

- a) Lungkarda ka ngulya-ngka nguna-mi.
bluetongue AUX burrow-A lie-NON.PAST
"The bluetongue skink is lying in the burrow."

⁴ The data for this problem set comes from Ken Hale via Barb Brunson.

- b) Nantuwu ka karru-kurra parnka-mi.
horse AUX creek-B run-NON.PAST
"The horse is running to the creek."
- c) Karli ka pirli-ngirli wanti-mi.
boomerang AUX stone-C fall-NON.PAST
"The boomerang is falling from the stone."
- d) Kurdu-ngku ka-jana pirli yurutu-wana yirra-rni.
child-D AUX stone road-E put.NON.PAST
"The child is putting stones along the road."

What is the meaning of *each* of the affixes (suffixes) glossed with -A, -B, -C, -D, and -E? Can you relate these suffixes to thematic relations? Which ones?

GPS3. THETA GRIDS

[Data Analysis; Basic]

For each of the sentences below identify each of the predicates (including non-verbal predicates like *is likely*). Provide the theta grid for each. Don't forget: include only arguments in the theta grid; DPs and PPs that are adjuncts are not included. Index each DP, PP, CP argument with the theta role it takes. Assume that there are two different verbs *give* (each with its own theta grid) to account for (c) and (d); two different verbs *eat* (each with its own theta grid) for (e) and (f); and two different verbs *ask* for (i) and (j).

- The stodgy professor left with his teaching assistant.
- I played a tune on my iPod.
- Molly gave Calvin a kiss.
- Mercedes gave a test to the students in the lecture hall.
- Pangur ate a cat treat.
- Susan ate yesterday at the restaurant.
- Gwen saw a fire truck.
- Gwen looked at a fire truck.
- Michael asked a question.
- Adam asked if Hyacinth likes pineapples.
- It is sunny in the dining room.
- I feel it is unfortunate that television is so vulgar these days.
- That Angus hates sushi is mysterious.

GPS4. SINHALA⁵

[Data Analysis; Basic/Intermediate]

Two forms of the Sinhala verb appear in the data below and are identified in the glosses as A or B. (Data from Gair 1970.)

⁵ This problem is loosely based on one given to me by Barb Brunson. However, the data and questions have been altered. The data in this version of the problem set is taken directly from Gair, with some minor modifications to the glosses.

- Provide a complete theta grid for each of the verbs in the following data. Be sure to primarily look at the second line of each piece of data, not the English translation.
- Using indexes identify what theta role is assigned to each DP.
- Discuss briefly (no more than two sentences) what kind of DP the suffix *-tə* attaches to.
- What is the difference between *mamə* and *maʔə*? (Hint: the answer to this question is related to the answer to question (3).)
- In terms of theta roles, what is the difference between the A and the B verb forms?
 - Mamə kawi kiənəwa.
I poetry tell-A
"I recite poetry."
 - Maʔə kawi kiəwenəwa.
I poetry tell-B
"I started reciting poetry (despite myself)."
 - Lamea kataawə ahanəwa.
child story hear-A
"The child listens to the story."
 - Lameaʔə kataawə əhenəwa.
child story hear-B
"The child hears the story."
 - Mamə naʔənəwa.
I dance-A
"I dance."
 - Maʔə nəʔənəwa.
I dance-B
"I dance (I can't help but do so)."
 - Həmə irida mə mamə koləmbə yanəwa.
every Sunday EMPH I Columbo go-A
"Every Sunday I deliberately go to Colombo."
 - Həmə irida mə maʔə koləmbə yəwenəwa.
every Sunday EMPH I Columbo go-B
"Every Sunday I experience going to Colombo."
 - Malli nitərəmə aṅḍənəwa.
brother always cries-A
"Brother always cries."
 - Malliʔə nitərəmə əṅḍənəwa.
brother always cries-B
"Brother always bursts out crying without control."

- k) Mame untə baninəwa.
 I them scold-A
 "I deliberately scold them."
- l) Maṭə untə bænenəwa.
 I them scold-B
 "I experienced scolding them."
- m) Apiṭə pansələ peenəwa.
 we temple see-B
 "We saw the temple."

GPS5. THETA CRITERION

[Data Analysis; Intermediate]

Show how each of the following sentences is a violation of the theta criterion. Use theta grids to explain your answers.

- a) *Rosemary hates.
 b) *Jennie smiled the breadbox.
 c) *Traci gave the whale.
 d) *Traci gave a jawbreaker.
 e) *placed the flute on the table.
 f) *John placed on the table.
 g) *John placed the flute.
 h) *John placed the flute the violin on the table.
 i) *The rock placed the sky with the fork.
 j) *John placed the flute the table.

GPS6. EXPLETIVES

[Application of Knowledge; Intermediate]

In Suzette Haden Elgin's science fiction novel *Yonder Comes the Other End of Time* (Daw Books, New York 1986) there is a planet called Ozark, where magic is commonplace. Magic is performed by applying generative grammar rules to the real world. In the following passage, a group of magicians has just put up a wall around an entire kingdom, and a local magician and the hero (Coyote Jones from Earth) are discussing how it happened.

"... And you understand how it's done?"

"Certainly. Don't you?"

Coyote admitted that he didn't understand it at all, and got an odd look for his candor. But as the two of them made their way back to Castle Smith, with Willow leading this time, the Magician of Rank explained it to him.

"***It would be an Insertion Transformation***," he said, as casually as Coyote would have discussed the workings of his flyer. "No different, except in scale from causing a flower to appear where there wasn't one before or some such baby trick ..."

(Haden Elgin 1986: 159)

In the bold italic sentence above ("***It would be an Insertion Transformation***"), has the expletive insertion transformation applied? What evidence do you have for your claim? If not, how do you know (what's your evidence)? Be careful – the question is not whether they inserted a wall, but whether expletive insertion happened *in the bold-faced sentence* above!

CHALLENGE PROBLEM SETS**CHALLENGE PROBLEM SET 1: IRISH AND THE THETA CRITERION**

[Data Analysis and Application of Skills; Challenge]

What problems does each of the following examples present for the theta criterion? (As a starting point, it may help to draw the theta grid for each verb and show what DP gets what role.) Please, not more than 3–4 sentences of discussion per example.

- a) an fear a bhfaca mé é
 the man who saw I him
 "the man who I saw"
- b) Rinceamar.
 dance.1PL
 "We danced."
- c) Ba-mhaith liom an teach a thógail.
 COND-good with-me the house its building
 "I would like to build the house."

CHALLENGE PROBLEM SET 2: OBJECT EXPLETIVES

[Critical Thinking; Challenge]

In the text above, it was observed that theta-role-less expletives primarily appear in subject position. Consider the following sentence. Is *it* here an expletive?

I hate it that you're always late.

How could you tell?

CHALLENGE PROBLEM SET 3: PASSIVES

[Data Analysis; Challenge]

Part 1: Write up the theta grids for the verbs in the following sentences. Assume there are two verbs *give* (*give*₁ is seen in (d), *give*₂ in (e)).

- a) John bit the apple.
 b) Susan forgave Louis.
 c) The jockey rides the horse.
 d) Phillip gave the medal to the soldier.
 e) Phillip gave the soldier the medal.

Part 2: English has a suffix *-en* that, when attached to verbs, changes the structure of the sentence associated with them. This is called the **passive** morpheme. The following sentences are the passive equivalents of the sentences in part 1. The bracketed PPs starting with *by* are optional.

- f) The apple was bitten (by John).
- g) Louis was forgiven (by Susan).
- h) The horse was ridden (by the jockey).
- i) The medal was given to the soldier (by Phillip).
- j) The soldier was given the medal (by Phillip).

Describe in your own words what the *-en* passive suffix does to the theta grids of verbs. Pay careful attention to the last two examples, and to the optionality of the *by*-phrases.

CHALLENGE PROBLEM SET 4: HIAKI -WA⁶

[Data Analysis and Critical Thinking; Challenge]

Part 1: Consider the function of the suffix *-wa* in Hiaki (also known as Yaqui), a language spoken in Southern Arizona and Mexico. Look carefully at the data below, and figure out what effect this suffix has on the theta grids of Hiaki verbs. What English phenomenon is this similar to? (Data from Escalante 1990 and Jelinek and Escalante 2003.)

- a) Peo Huan-ta chochon-ak.⁷
Pete John-ACC punch-PERF
"Pete punched John."
- a') Huan chochon-wa-k.
John punch-WA-PERF
"John was punched."
- b) 'Ume uusi-m uka kuchu-ta kuchi'i-m-mea bwa'a-ka.
the children-PL the-ACC fish-ACC knife-PL-INST eat-PERF
"The children ate the fish with knives."
- b') 'U kuchu kuchi'i-m-mea bwa'a-wa-k.
the fish knife-PL-INST eat-WA-PERF
"The fish was eaten with knives."
- c) Peo bwiika.
Pete sing
"Pete is singing."

⁶ Thanks to Heidi Harley for contributing this problem set.

⁷ Sometimes when *-wa* attaches to a verb, the form of the root changes (usually /e/ becomes /i/). This is a morphophonological phenomenon that you don't need to worry about. ACC refers to accusative case, INST means instrument, and PERF means perfective aspect (aspect plays no role in the answer to this problem). There is no nominative suffix in Hiaki.

- c') Bwiik-wa.
sing-WA
"Singing is happening." or "There is singing going on." or "Someone is singing."

Part 2: Not all verbs allow *-wa*. Consider the following pairs of sentences, which show verbs that don't allow *-wa*. In terms of theta grids, what do these sentences have in common with each other that differentiates them from the ones that allow *-wa* (above in part 1)?

- a) 'U wikia chukte.
the rope come.loose
"The rope is coming loose."
- a') *Chukti-wa.
come.loose-WA
"Coming loose is happening." or "There is coming loose going on." or "Something is coming loose."
- b) 'U kaaro nasonte.
the car damage
"The car is damaged."
- b') *Nasonti-wa.
damage-WA
"Damage is happening." or "There is damage going on." or "Something is getting damaged."
- c) 'U kari veete-k.
The house burn-PERF
"The house burned."
- c') *Veeti-wa-k.
Burn-WA-PERF
"Burning happened." or "There was burning going on." or "Something is getting burned."
- d) 'U vachi bwase'e.
The corn cook
"The corn is cooking."
- d') *Bwase'i-wa.
cook-WA
"Cooking happened." or "There was cooking going on." or "Something is being cooked."

Part 3: The data in (e) and (e') below might throw a wrench in the hypothesis you developed above in part 2. Explain why these data are problematic for your analysis in part 2.

- e) Ume uusim sawaria-ta-mak koko-n
 The.PL children.PL yellow.fever-ACC-with die.PL-IMPF
 "The children were dying of yellow fever."
- e') Sawaria-ta-mak koko-wa-n
 Yellow.fever-ACC-with die.pl-WA-*impf*
 "People were dying with yellow fever" or
 "There was dying with yellow fever."

CHALLENGE PROBLEM SET 5: ANTIPASSIVES

[Data Analysis and Critical Thinking; Challenge]

In many languages there is an operation that changes the theta grid of certain verbs. This operation is called the **antipassive**.

Part 1: Here is some data from Inupiaq, an Inuit language of Canada and Alaska. Explain what adding the antipassive morpheme does to the theta grid of the verb. Verbs in Inupiaq agree with both their subjects and their objects. 3SUBJ-3OBJ means that the verb agrees with both a 3rd person subject and a 3rd person object. 3 means that the verb only agrees with a 3rd person subject. (Data from Seiler 1978.)

- a) Aṅuti-m umiaq qiñig-aa tirrag-mi. *Active*
 man-ERG boat-ABS see-3SUBJ.3OBJ beach-at
 "The man sees the boat at the beach."
- b) Aṅun (umiag-mik) qiñiq-tuq tirrag-mi. *Antipassive*
 man-ABS boat-INST see-3 beach-at
 "The man sees (with a boat) at the beach."

Part 2: The following is some data from English. This might also be called an antipassive construction. How is it similar or different from the Inupiaq antipassive?

- c) I ate a basket of apples.
 d) I ate.