Acourse on:
TRADITIONAL GRAMMAR, IMMEDIATE CONSTITUENT ANALYSIS, PHRASE STRUCTURE GRAMMAR AND X-BAR SYNTAX THEORY

Prof Moubtassime Mohammed
S6 Linguistics Stream

1. INTRODUCTION

This course is a study of ICA, which is a method of sentence analysis first founded by Leonard Bloomfield and developed further by Rulon Wells. As we know, the practice of ICA is now widespread. This course starts with a brief introduction of ICA and explores what is ICA. It also gives illustration of how sentences are analyzed and divided into constituents in the large construction. Then, it discusses what is ICA, and its approach and frame. In this course, we will develop a capacity of how to analyze sentential structures through the use of ICA. The course will explore the advantages of ICA, but also discusses what are the limitations of ICA.

Daily usage of a language leads people to understand more about how the language is produced. The different understanding of a language leads to a different interpretation of the people who use it or even listen to it. Analyzing a human language is very important to minimize the miscalculation of the aspects inside the language. Letters, which form a word and words or phrases, which form a sentence are the parts that need to be examined. The analysis of those parts can be done in many ways. This course presents four of many ways to analyze a sentence. They are traditional grammar, Immediate Constituent analysis (ICA), Phrase Structure Grammar and X-bar theory. How the analysis works in understanding a sentence will be accompanied by a few examples that make them clearer. The course will find out the weaknesses and the strengths of those analyzed through the explanations and examples.

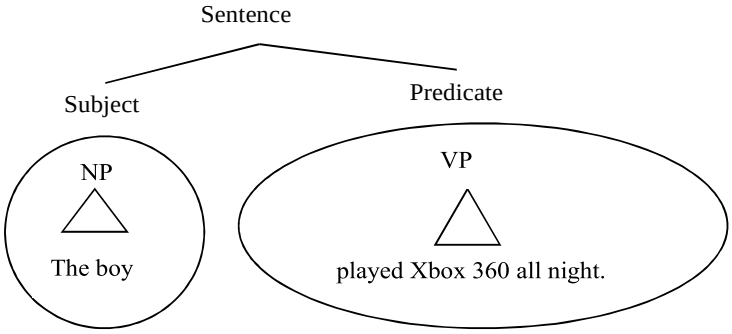
Sentences are used as a daily tool for everybody in doing a conversation with others. From those created every day, the correctness is the thing to be maintained. To know whether a sentence is correct or incorrect, we need to analyze it. Analyzing a sentence can be done in various ways. We have to know how the theory works first before we realize the strengths and weaknesses of it. The result can be very useful to understand how good the sentence is.

In the following, we will try to explain the analysis using traditional grammar, IC-analysis, Phrase Structure Grammar and X-bar theory. A few examples accompanying the explanations help us to understand of what are discussed. Realizing about the discussion leads to the understanding of strengths and weaknesses of those theories and we will discover that syntactic theories are a continuum (there are no breaks between these theories and others, as well).

2. TRADITIONAL GRAMMAR

Diachronically, Priscian as the most respected grammarian in the medieval period based his grammatical descriptions on the writing and speech of the best models available to him. With a few comments on his grammar the new pronunciation and grammar were preferred with the legalization of grammar in Latin language. This is the embryo that makes grammar as an important part of any language. The part of speeches were introduced in this period by grammarians as the noun which meant “substance” or “quality”, the verb which meant “action” or “passivity”, and so on.

Traditional grammar states the reason why certain grammatical features of a language occur and to explain how important the features are. This leads to the positive idea that traditional grammar gives a good explanation of the language. The traditional understanding of a sentence division is as subject and predicate or noun phrase (NP) and verb phrase (VP) as in



One way to analyze sentence structure is to think in terms of form and function. Form refers to a word class such as noun, verb, adjective, adverb, and preposition as well as types of phrases, such as prepositional phrase, nominal clause, and adverbial clause. Function word refers to a word that is important in grammar rather than its meaning in a sentence.

There are few forms of subject as a noun phrase in the sentence such as a bare noun phrase such as singular or plural noun phrases and a mass noun as in "*Students are at work",* a determiner phrase, a noun phrase which is preceded by a determiner (determiner + noun phrase) as in "*The teacher's car stopped outside our classroom",* a gerund, a phrase without a determiner. This gerund behaves like a noun phrase as in "*Watching movie is a pleasure"*, an infinitive which is preceded by question words, who, what, why, when, where, and how, as an embedded clause as in "*To speak is easier than to do"* and "*What to do is a start for everything new",* a full clause which is introduced by a complementizer as in "*That the students do not understand the lesson is not important"*, a direct quotation as in "*I like you is what she wants to hear from me"*, imperative mood which states "you" as the subject recipient as in "*Submit the paper immediately!",* the words like *it* or *there* which do not refer to anything or place "*It's dangerous* and *There is a student in the room*", and the word *it* which is used as co-referent with a subordinate clause that comes after it as in "*It is known by his friends that he never cheats in every test*".

This predicate is one of the two main parts of a sentence. It must contain a verb as in "She sings" and other sentence elements that can make the predicate complete. They can be as direct objects (She reads the syntax book), indirect objects (She gave me a book), or object of preposition (She watches to the movie), noun predicative (Jakarta elected him governor), and adjuncts (She met me in the library). These predicates provide information about the subject, such as what the subject is, what the subject is doing, or what the subject is like.

There are forms of subject and predicative, which must be connected by a linking verb. A *predicative nominal* such as *Pak Dwi is the lecture of Syntax* is the first form. *"the lecture of syntax"* here is a predicative nominal. It is a noun phrase that functions as the main predicate of the sentence. A *predicative adjective* as in *This lesson is attractive* is the second form. "*Attractive*" is the predicate adjective which functions as a predicate of the sentence.

Objects fall into three classes. They are direct objects, indirect objects, and objects preceded by a preposition. A direct object must be placed in a sentence when there is an indirect object in it. A direct object answers the question "What?" as in "The students watched The Dark Night Rises", while an indirect object which functions as the recipient of the direct object answers the question "To whom?" or "For whom?" as in "The students sent the lecturer a ticket", and in "The students listen to the radio", *radio* is the object of the preposition *to*, and the **prepositional object** of the verb *listen*.

An object is nominal which can take a number of forms. They are as a noun or noun phrase, as in "The student knows *his number*", an infinitive, as in "The student likes *to learn*", a gerund, as in "The student loves *doing the test*", and a noun clause, as in "The student does not know *that the paper is typed*."

In English, adverbials most commonly take the form of adverbs (*The students leave hurriedly*), temporal noun phrases (*The students were tested this morning*) or prepositional phrases (*The students presented in English*), or clauses of time, cause (*The students asked and answered in English because they were in English study class*), condition, manner, place, and contrast.

An adverbial is a construction that modifies, or describes, verbs. When an adverbial modifies a verb, it changes the meaning of that verb. Word groups that are also considered to be adverbials can also modify verbs: for example, a prepositional phrase, a noun phrase, a finite clause or a non-finite clause. In every sentence pattern, the adverbial is a clause element that tells where, when, why, or how. There can be more

than one adverbial in a sentence. In addition, the same adverbial can be moved to different positions in a sentence.

Adverbials are typically divided into four classes; adverbial complements, adjuncts, conjuncts, and disjuncts. Adverbial complements are adverbial that can make a sentence meaningless if removed as in "The student presents the paper in the class." Adjuncts are part of a sentence. They are not too important because if they are omitted, the sentence is still meaningful as in "One of the students helped me with an idea." Conjuncts are words that link two sentences together such as "She helped me, therefore I can do my thesis." Disjuncts are words that make comments on the meaning of the rest of the sentence, for example "*Suddenly, We passed all of the exams*."

The purpose of traditional grammar is to prescribe the way people speak, or simply, to tell people how to speak and let people know the correct way of their speaking. That is why traditional grammar is said to be prescriptive. Traditional grammar is normative as the different usage from the rules is said to be ungrammatical. This makes people difficult to produce a language for daily conversation when they have to think about the correct from first. It is different with modern linguistic which is said to be descriptive, to describe the way people speak, which deals with pragmatics, sociolinguistics, etc. In the prescriptive way, traditional grammar pays more attention to the written form of language which makes it restricted mainly to syntax, the way of words making patterns to form sentences.

Traditional grammar cannot resolve the ambiguity existing in the grammatical forms. Its methods are inaccurate, incomplete and inconsistent, and the descriptions are inexplicit and intuitive. The sentence *On Friday* (adjunct), *the students* (the subject argument) *did* (the predicate) *the test* (the object argument) *in Palem room* (adjunct) uses adjuncts of time and place. The adjunct *in Palem room* modifies the verb *did* in which case it is "The student who did the test while the student in Palem room" or the noun phrase *the test* in which case it is "the test which is in Palem room". It means that the adjunct *in Palem room* is ambiguous.

Basically, it is the most widespread, influential, and best understood method of discussing the language. It is consistently applied by most of those who teach and study it as it is available in every language. It gives a fairly thorough and consistent analysis of the declarative sentence in written and spoken discourse. The theory contained in it can be explained, and this is the vehicle for ordinary students and scholars to master the language successfully for centuries. This study can justify fairly within the forms of speech and writing which makes all languages is equal and share aspects beside their uniqueness.

The differentiation shown in morphology and syntax is as the result of inadequate notion of modification and inadequate distinction of part of speech. The researcher can never understand the usages of the sentence as "ungrammatical", "exceptions", and "idioms". It is difficult to know whether a language is more sophisticated and cultured than the other or not.

3. IMMEDIATE CONSTITUENT ANALYSIS

Carni (2006) gave a definition of constituent as a syntactically unit combined with other syntactically unit to form a construction. He mentioned that each constituent can be classified based on its arrangement of internal constituent itself and with the grammatical function related with other constituent. It is clear that constituent is a part or component of a construction.

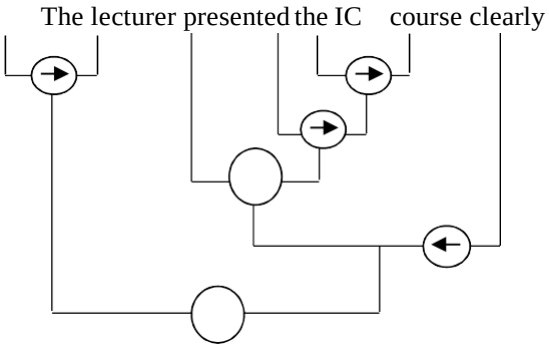
Constituent Analysis is a sentence analysis by cutting into smaller units. To know whether a structure is constituent or not we need to have a constituency test. This test can be done with topicalization (...the student is doing the test), clefting (It is the student who is doing the test), pseudoclefting (The one who is doing the test is the student), pro-form substitution (She is doing the test), and answer ellipsis (Who is doing the test? -the student). One can safely assume that the noun phrase *The student* in the example sentence is a constituent.

Immediate constituent analysis or IC analysis is a sentence analysis method that was first mentioned by Leonard Bloomfield. He did not divide a sentence into "subject" and "predicate", but he replaced them with the terms "actor" and "action" because he did not create a terminology for phrasal categories.

IC-Analysis in Phrase Structure Grammar

IC analysis is a very important tool for syntactically analysis. The technique applies that a sentence must be analyzed into immediate constituents actor and action. The construction inside the immediate constituents must be further analyzed until the final constituents are reached that is a word.

Firstly, a sentence is divided into major parts or immediate constituents. These constituents are divided into further immediate constituents. This continues until each constituent consists of only one meaningful part of a word. All is presented in a diagram, usually in trees form, that reveals the hierarchical immediate constituent structure.

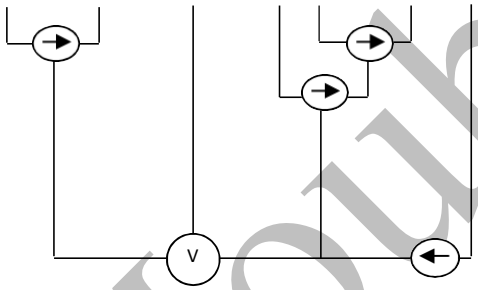


This tree illustrates the manner in which the entire sentences is divided first into the two immediate constituents *the lecturer* and *presented the IC course clearly.* These two constituents are further divided into the immediate constituents *the* and *lecturer* and *presented the IC course* and *clearly*. Then *presented* and *the IC course.* The last one is *the* and *IC course*.

Each individual word is a constituent by definition. This is a significant aspect of IC analysis is phrase structure grammar. A word as the smallest constituent is the final process of IC-analysis. A different process happens in a dependency grammar where individual words are not as the final result of constituents.

IC-Analysis in Dependency Grammar

The finite verb functions as the root of all sentence structure. There is no initial binary actor-action division of the clause. A finite verb phrase (VP) constituent and many individual words are not qualified as constituents in the IC-analysis as in the sample below.



The lecturer presented the IC course clearly

While the structures that IC-analysis identifies for dependency and constituency grammars differ in significant ways, as the two trees just produced illustrate, both views of sentence structure are acknowledging constituents.

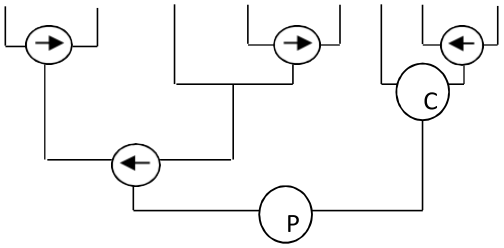
Immediate constituent produces the preference for binary analyses. Breaking down sentences into two and only two immediate constituents was one example of the work. Obviously, in morphological analysis, many words can also be divided in two. Thus, English *incompleteness* consists of the suffix *–ness* and the word *completeness,* which in turn consists of the prefix *in-* and *complete*. In a bigger example of a sentence, we can make components in first cutting in constituent analysis as in *The teacher teaches a lesson.* This sentence has two immediate constituents, i.e. *The teacher* and *teaches a lesson*. *The teacher teaches a lesson* has six ultimate constituents i.e. *The, teach, –er, teach, –es, a lesson.*

It can also be used for any languages. Here is an example taken from Indonesian language. The sentence "Pemuda itu mengerjakan skripsi" has two immediate constituents, "pemuda itu" and "mengerjakan skripsi". There are seven ultimate constituents of the sentence i.e. *pe-, muda, itu, me-, kerja, -kan, skripsi*.

It helps translators recognize autonomous units, any stretch of language that can be translated as a single unit. The clause can be broken into three distinct autonomous units. *The classes/are divided/into*

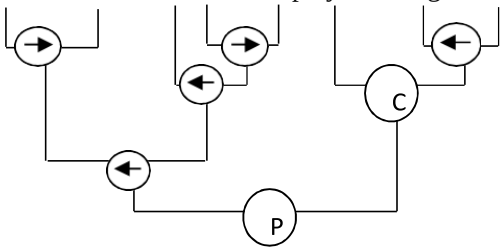
three is broken up into noun phrase (determiner + noun), verb phrase (verb + particle), and prepositional phrase.

The word order in IC-analysis cannot be disturbed. This can be seen when a sentence composed of the same words is in different word order as in:

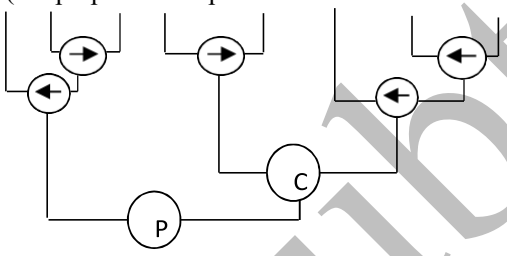


1. The students played card games on the floor
(the prepositional phrase "on the floor" modifies the verb phrase)

2. The students on the floor played card games



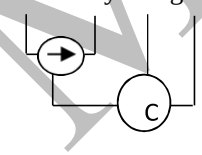
(the prepositional phrase "on the floor" modifies the noun phrase)



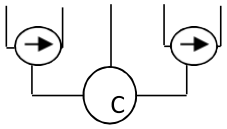
3. On the floor the students played card games
(the prepositional phrase "on the floor" modifies the rest of the sentence)

Unlike the traditional grammar, the IC-analysis can account for ambiguities and distinguish them. "good boy and girl" can be paraphrased in two ways. It can be "good boy and girl of all ages". It can also be "good boy and good girl". The expansion of "good boy" is as a single morpheme or "boy and girl" as a single morpheme. The recognition of two different IC analysis shown by the tree diagram:

Good boy and girl

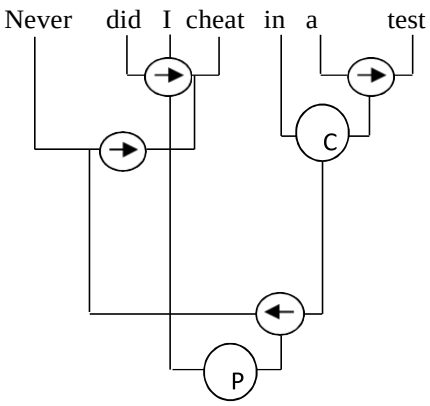


Good boy and good girl



The IC-analysis makes an early assumption about grammatical status of the elements. The words "going to leave" can be cut in two ways. They are "going/to leave" and "going to/leave". In IC-analysis "to leave" is the best possibility because it is a constituent. It forces the language user to accept "to leave" as a nominal element comparing with "going home" in the case that both have the same grammatical type.

The end result of IC analysis is mostly presented in a diagram revealing the hierarchical immediate constituents structure of a sentence. However, when the structure is unusual, the diagram may become complex. The IC-analysis cannot make a binary division when the elements in the sentence are separated in the sequence or discontinuous. In a "Never did I cheat in a test" the word "did" is nearer to "cheat" than to "I".



IC-analysis should be based on a linear string. The sequences of single morpheme in a sentence are broken when it is discontinuous. Permitting the discontinuity makes it nonlinear.

The IC-analysis cannot analyze the phrase further than just a level of words. The phrase like "civil law" which in practice means law that deals with the rights of private citizens rather than with crime cannot go further into a meaningful way when it is analyzed in IC-analysis. The phrase comes out in meaning clearly when it is cut in the way like "civil-law".

In brief, **I.C. Analysis – Definition:** is an implicit assumption that linguistic structure, especially syntactic structures are layered structures amenable to analysis by progressive dichotomous cutting.

In another word I.C. Analysis is that system of grammatical analysis that breakup sentences into sequential layers, or constituents until in the final layer, and every constituent consists of only a word or meaningful part of a word.

Approach of I.C. Analysis: The initial emphasis was upon pure segmentation, simply breakup the sentence into its constituent parts without, at first, knowing what these parts were: Generally the section is binary except in some cases where section into three or more points is allowed.

Immediate Constituent analysis has the following advantages and disadvantages:

ADVANTAGES OF IC ANALYSIS

A. **Identification of the layers of relationship in a construction**: IC analysis discovers the layers of relationship in a construction. English syntax is based on this ability of structures to function within larger structures, which are, in turn, serving other functions in still larger, more complex structures(sentences). Composing a more complex sentence such as

The pretty girl put on her red and blue coat kissed her mother and left.

demonstrates the nature of relationship that must be negotiated if a hearer or a reader is to understand such a sentence. Actually anyone who is capable of understanding the meaning of the sentence obviously has the mental capacity to keep all those relationships afloat as he hears or reads the sentence.

B. **Fixity of word order:** In IC analysis the word order is not disturbed in any way. This advantage is best demonstrated by sorting the relationship found in the following sentences which are composed of the same words but which are different in word order:

1. The boy played marbles on his knees.

2. The boy on his knees played marbles.

3. On his knees the boy played marbles.

Class work: Draw tree diagrams for 1, 2 and 3.

These sentences may be said to be stylistically different. In the first, the prepositional phrase "on his knees" modifies the verb phrase; in the second, it modifies the noun phrase; in the third it modifies all the rest of the sentence. Yet in the word order within the structure "on his knees" does not change.

C. **To account for ambiguities and distinguish them:** A famous example 'old men and women' can be paraphrased in two ways; it is either "old men and women of all ages" or "old men and old women". The principle of expansion here allows us two interpretations. Either 'old men' is an expansion of a single morpheme (e.g. men or boy) or 'men and women' is an expansion of a single morpheme (e.g. people or men). This would allow us to recognize two different ICA shown by the tree diagram:

1. old men and women

2. old men and women

LIMITATIONS OF ICA:

(A) Immediate constituent analysis has its limitations:

It cannot analyse discontinuous constituents, for example: She is taller than her sister.

In this sentence, the sequence **–er than** cannot be covered by ICA.

(B) ICA stops at the level of words: In ICA, it is tacitly assumed that there will be no division into pieces smaller than words (morphemes).

(C)Unbalanced Bracketing: IC analysis does not refer to our grammatical knowledge. So it does not take us very far and without the help of labeled bracketing we cannot point out the source of ambiguity in many sentences. The labeled bracketing can be used to differentiate the two possibilities in an example that is often against IC analysis.

Flying planes can be dangerous. Here, in one case 'flying' is the head of the noun phrase while on the other hand is 'planes'. It cannot help explain the ambiguity that exist in those sentences (see the difference between lexical ambiguity and structural ambiguity) as ICA does not use grammatical labels on the tree nodes of the brackets.

4. Phrase Structure Grammar

- In the previous class, we discussed generally
  - how ICA functions,
  - how it helps, in part, to determine the constituent parts of a sentence.
  - we saw how we can use it to generate an infinite number of sentences of a language.
  - This is possible because the rules allow several distinct but equivalent derivations of sentences.
  - ICA has both cons and pros
- Two important functions of Phrase Structure Grammar/Rules:
  - they state the constituents that make up a sentence, and
  - they tell us in what basic order the constituents are to be arranged.

Today, we shall concentrate on English phrase structure rules and their application to the structure of English. Examples from other languages are also interesting.

Origins of phrase structure analysis

To understand the properties of modern phrase structure grammars, it is useful to place their development in a wider formal and historical context. Phrase structure grammars and associated notions of phrase structure analysis have their proximate origins in models of Immediate Constituent (IC) analysis. Although inspired by the programmatic syntactic remarks in Bloomfield, these models were principally developed by Bloomfield's successors, most actively in the decade between the publication of Wells and the advent of transformational analyses in Harris and Chomsky. The central intuition underlying models of IC analysis was that the structure of an expression could be exhibited by dividing the expression into parts (its immediate constituents), further subdividing these parts, and

continuing until syntactically indivisible units were obtained. This style of analysis was motivated in part by a belief in the syntactic relations, in particular the view that the most important relations held between immediate constituents.

OBJECTIVES

- At the end of this unit, you should be able to:
  - know what English phrase structure rules are
  - how they apply in the derivation of English sentence structure.

The Phrase Structure Rules of English

- We can best begin the description of English phrase structure rules by taking a very simple example of a phrase structure of grammar, such as is given in (1) below:

1. (i) S□NP VP
   (ii) VP□V NP
   (iii) NP □ Det N
   (iv) Det □the
   (v) N □ cat, mouse
   (vi) V □ catch

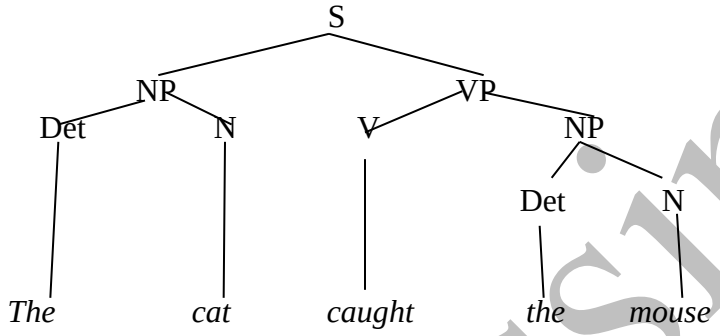
- This grammar consists of six phrase structure rules. ('Phrase structure' will henceforth be abbreviated 'PS').
  - Each rule is an instruction to 'rewrite' the symbol on the left of the arrow as the string of one or more symbols given on the right. The symbols used in the grammar are divided into two discrete sets (apart from the arrow): *non-terminal symbols* occur on the left of some rule or rules, *terminal* ones do not.
  - The terminal symbols represent morphemes, the non-terminal ones syntactic categories (with the following abbreviations: S = sentence, NP = noun phrase, VP = verb phrase, V = verb, Det = determiner, N = noun).
- Such rules permit the construction of sentences in a purely mechanical way in accordance with our goal of complete explicitness.
  - The procedure is illustrated in (2), where the succession of 'strings' of symbols constitutes what is called a PS derivation.

2. (i) #S#
   (ii) # NP VP#
   (iii) #NP V NP#
   (iv) #Det N V NP#
   (v) #Det N V Det N#

- What do you understand by this?
  - In (2i), # indicates sentence boundary.
  - S is designated as the initial symbol of the grammar, reflecting the fact that we are concerned with devising a grammar to generate sentences.
  - Thus all PS derivations generated by this grammar (and others we shall consider) have #S# as the first line.
  - Each subsequent line in a derivation derives from the one preceding it by one application of a PS rule.
- That means:
  - Line (2ii) derives from line (2i) by rewriting S as NP followed by VP according to rule (1i);
  - line (2iii) derives from (2ii) by rewriting VP as V NP by rule (1ii),
  - the initial NP of (2ii) being simply copied down unchanged into line (iii), and so on.

- In moving from one line to the next we replace just one symbol, and copy the remainder unchanged.
    - A derivation is complete, or terminated, when its last line contains no non-terminal symbols.
    - The last line of the derivation can then be converted into the appropriate phonological representation by means of morphophonemic rules corresponding to *The cat caught the mouse*.
- In addition to generating a set of sentences, such a grammar will automatically assign a structure to each.
    - For example, the syntactic structure assigned to the sentence *The cat caught the mouse* by grammar (1) may be represented in a tree diagram (3):

3.



- This kind of representation is known as a *phrase marker* (henceforth 'PM').
    - The phrase marker is a set of strings that assigns a phrase structure interpretation to a terminal string (the string of lexical categories that terminates the phrase structure derivation).

- labeled bracketing:
    - Apart from using a tree diagram, phrase markers can be represented using labeled bracketing.
    - In a labeled bracketing, whole constituents are enclosed in square brackets.
    - Such constituents are labeled according to the name of the constituent in which they are enclosed.
    - It should be noted that once a bracket is opened, it must be closed otherwise the analysis will be incorrect.
    - Consider our example (3).

3. The cat caught the mouse

[s[NP[The] [cat]] [VP[caught] [NP[the] [mouse]]]]

Det N v Det N

- The grammar discussed so far can generate sentences such as 4(a), (b), (c) but not 5(a), (b), (c):

4. (a) The man has the book
   (b) John passed the exam
   (c) Sheila owns a truck
5. (a) The man will have the book
   (b) John may pass the exam
   (c) Sheila can own a truck

- Is it possible to make up a new rule that will generate the sentences in (5)?
    - Indeed it is, and this new rule is formulated as (6):

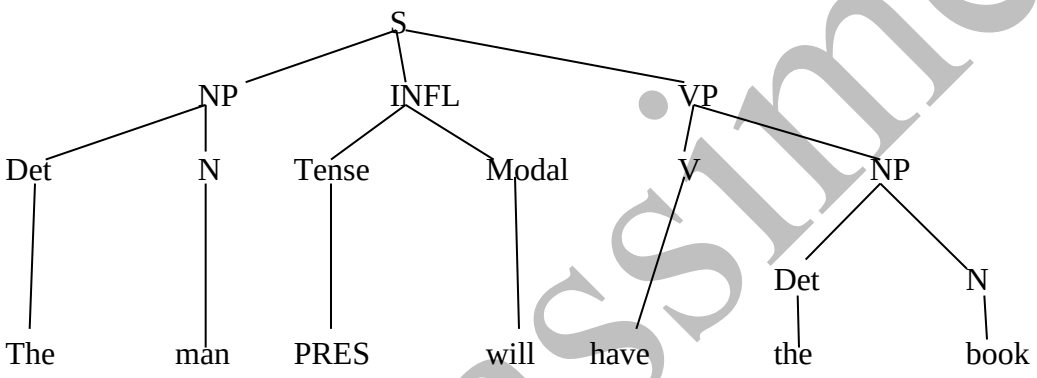
6. S □NP INFL VP

- As in (1) above, S = sentence, NP = noun phrase, INFL = inflection (this is the same as Aux = auxiliary used in some texts), VP = verb phrase.
    - Each of these phrasal categories are broken down into its component parts as illustrated in (1).
    - As can be observed, the only difference with (1) is the addition of the phrasal category INFL. This can be expanded to give (7):

7. INFL□Tense (M) (have + en) (be + ing)

- In (7) *tense* could be *present* or *past*, M = modal (*will, would, can, could, shall, should, may, might*, etc). Brackets still indicate optionality.
    - We can, thus, represent 5(a) in a tree diagram as follows:

5. (a) The man will have the book



Class work: Draw similar tree diagrams for

- John may have passed the exam.
- Sheila may have been doing the work.

Note that they include *have + en* and *be + ing*.

- After the transformations have applied, we now have the surface structure derivations.

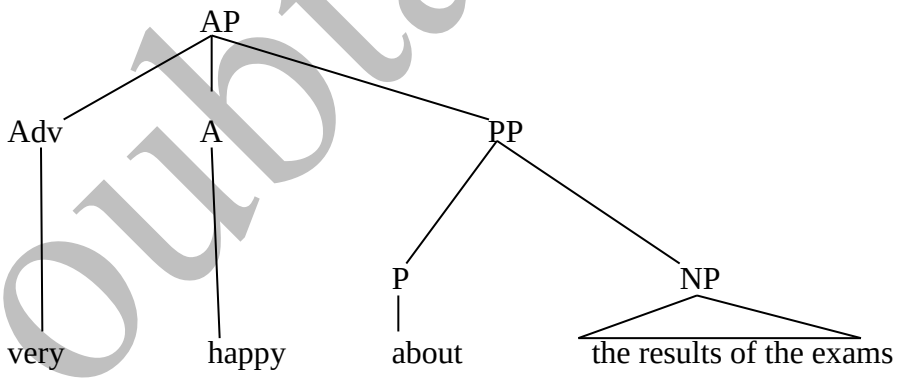
Dominance Relations

- Before we consider how lexical items are introduced into phrase markers, it will be useful to discuss the relations between the elements in phrase markers.
    - The relation between a category and its constituents is called *dominance*. That is, a category *dominates* all of its constituents.
    - For example, the category VP in (3) dominates the category V, NP, Det and N.
    - However, the relation between VP and NP is different from the relation between VP and N.
    - VP *immediately dominates* NP, but it does not immediately dominate N since another category, NP, intervenes between VP and N. We may say that a category X immediately dominates a category Y when X dominates Y and every other category dominating Y dominates X. When X immediately dominates Y, then Y is an immediate constituent of X.
- In example (3) above, both NP and N are constituents of VP, but only NP is an immediate constituent of VP.
    - Dominance relations between categories in a structural description determine the *hierarchical structure* of sentences.
    - If two categories in a phrase marker are not in dominance relation, then they are in a *precedence* relation, for example, the subject NP in (3) precedes VP.

- Precedence relations determine the linear (left-to-right) ordering of elements in the phrase marker and hence determine the *linear structure* of sentences.

- In terms of tree diagram in (3) above, two kinds of branching are possible:
  - *Unary branching* (where a category immediately dominates only one constituent) and *binary branching* (where a category has two or more immediate constituents).
- When a category has just two immediate constituents, it constitutes an instance of binary branching.
- A *node* is a point on the phrase marker where there is a category label. This category label can be phrasal (e.g. NP, VP) or lexical (e.g. N, V, Adj.)
- A *path* is the line that links a node to the other. In (3) the line from NP to N is a path, and so is VP to V or VP to NP.
  - As we shall see later, the branching property of structural descriptions plays an important role in defining principles of syntactic theory.
- Some phrasal categories are defined in terms of the lexical categories they contain (e.g. VP in terms of V).
  - The lexical category stands as the *head* of its phrasal category, and the phrasal category is designated as the *phrasal projection* of the lexical category.
  - Whether or not every lexical category projects its own phrasal category is open to discussion.
- We will assume, however, that it is not the case. For example, Det in (3) does not project its own phrasal category.
  - Rather, it is a lexical constituent of the phrasal projection of N (=NP).
  - A similar relationship holds between the adverb *very* and the adjective *happy* in the adjective phrase (AP) in *very happy about the results of the exams* as represented in (9).

9.



- Lexical constituents of phrasal projections like Det and Adv act as modifiers of the lexical head of the projection.
  - Each can be referred to as a *specifier* of the lexical head it modifies. Phrasal constituents of phrasal projections are designated as *complements* of the phrasal head.
    - Thus, in (3) the object NP is a complement of V; in (9) the PP is a complement of A, and the NP is a complement of P.
- In general, the non-head constituents of phrasal projections are optional.
  - For example, NP's can lack determiners, as in *Frank admires Aditi*.
  - Parentheses are used to express the optionality of constituents; for example, the notation in (10) indicates that Det is an optional constituent of NP.

10. NP (Det) N

- Similarly, the fact that some verbs occur without NP objects can be expressed as in (11):

11. VP V (NP)

- Braces are used to collapse rules where a disjunctive choice of constituents is involved. For example, English VP's may be of the form:
    - V-NP-NP (*told John a story*) or of the form, V-NP-S (*told John Sheila was leaving*).
- Since English does not have VP's of the form V-NP-NP-S (or V-NP-S-NP), it appears that the occurrence of a second NP complement in a VP precludes the occurrence of an S complement, and conversely.
    - This disjunction can be represented by placing NP and S in braces to indicate that only one can be chosen in the expansion of VP, as shown in (12):

12. VP□V (NP) $\left[\left\{\begin{matrix}\text{NP}\\ \text{S}\end{matrix}\right\}\right]$

Lexical Insertion

- So far, we have discussed the formalism for rules that generate constituent structure for sentences.
    - We must now relate constituent structures to sentences – that is, strings of words in a language.
    - This relationship is established at the level of lexical categories and lexical items.
    - In what follows, we will see that a lexical item is just an instantiation of a lexical category – from which it follows that a lexical item is not a constituent of a lexical category.
    - A lexical item consists of three separate parts that concern sound, meaning and structure.
    - Consider, for example, the kinds of information that would be contained in a lexical entry for *man*.

13.
$$\left(\begin{matrix}\textit{Man}\\ \text{mæn}\\ \text{+N}\\ \text{+ animate}\\ \text{- plural}\\ \text{+ human}\\ \text{+ masculine}\end{matrix}\right)$$

- Explanation
    - The phonemic transcription (/mæn/) gives a representation of the lexical item *man* in terms of its sound.
    - The [+N] is a categorial feature indicating that *man* belongs to the class of nouns. Categorial features provide syntactic information about lexical items.
    - Finally, there is a group of features that relate to the semantic interpretation of *man*.
    - Thus, *man* is [-plural] as opposed to *men*, which is [+plural]; [+masculine] as opposed to *woman*, which is [-masculine]; [+animate] as opposed to *plant*, which is [-animate]; and [+human] as opposed to *bull*, which is [- human].
- Representations such as (13) show that lexical items can be represented, at least partially, in terms of sets of features that relate them to their phonological, syntactic, and semantic forms.
- *Lexical insertion* can be formulated as an operation that substitutes a fully specified lexical feature matrix (the lexical item) for a partially specified matrix (the categorial feature alone).
    - So the derivation of a sentence containing the lexical item *man* would include an operation that substitutes (13) for a [+N] in the phrase marker.
- Lexical insertion of verbs is sensitive to syntactic context.
    - Thus, a verb like *sleep* cannot be inserted into a phrase marker where V is followed by an object NP as in (14b); nor can a verb like *mention* be inserted into a phrase

marker where V is not followed by an object NP as in (15b). In the case of *put*, V must be followed by both an object NP and a PP as in (16a):

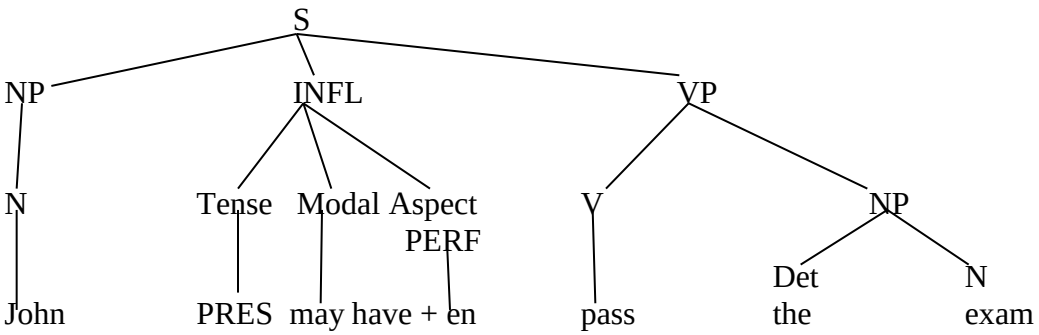
14. (a) Mary slept
    (b) *Mary slept the room
15. (a) Mary mentioned the book
    (b) *Mary mentioned
16. (a) Mary put the money in the wallet
    (b) *Mary put the money

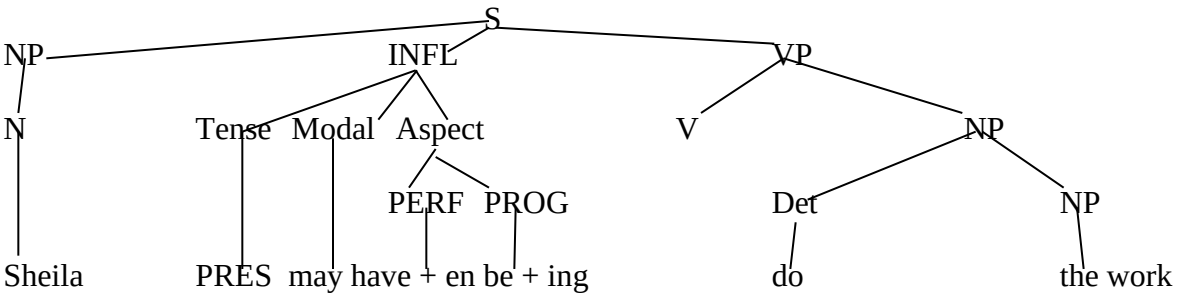
- To account for this context-sensitivity of lexical insertion, the grammar must be able to distinguish among three distinct subcategories of verbs:
  - (i) verbs that may not occur with an NP object (*slept*),
  - (ii) verbs that must occur with an NP object (*mention*), and
  - (iii) verbs that must occur with both an NP object an a PP (*put*).
    - These contextual features are designated as *subcategorization features*.
- A subcategorization feature constitutes a lexical property of a specific lexical item.
  - When the subcategorization feature and actual syntactic context of the lexical item match, we can say that this lexical property is *satisfied*.
  - When they do not match, the lexical property is not satisfied.
- It is the fact that the lexical properties are not satisfied that is responsible for the ungrammaticality of (14b), (15b) and (16b).
  - The grammar should, therefore, contain a principle that requires lexical properties to be satisfied.
  - That is, the grammar consisting of phrase structure rules, a lexical insertion transformation, and the principle of lexical satisfaction will designate as ungrammatical any sentence containing a lexical item whose subcategorization feature does not match its actual context in a phrase marker.

- Today, we have discussed the phrase structure rules of English.
  - We have also tried to explain certain terms that will help you follow the discussion – dominance relations and lexical insertion.
  - Study these rules and try to use them in generating your own sentences.
  - It is by so doing that you will acquaint yourself of their operations.

Class work:

- Try to find out the limitations of PSG, and how we can solve.
- Those who have really managed to find out the disadvantages of PSG will also find out these limitations can be solved by the use of Transformational Grammar.
- Read about Transformational Grammar and find our how it works.





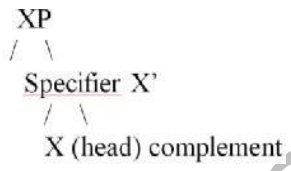
5. X – BAR THEORY

X-bar theory is discussed in almost all modern textbooks of syntax. It attempts to identify syntactic features. It claims that among their phrasal categories, all those languages share certain structural similarity that does not appear in traditional grammar for English. X-bar theory was first proposed by Noam Chomsky in 1970 and further developed by Ray Jackendoff in 1977. An X-bar theoretic understanding of sentence structure is possible in a constituency-based grammar only. It is not possible in a dependency-based grammar.

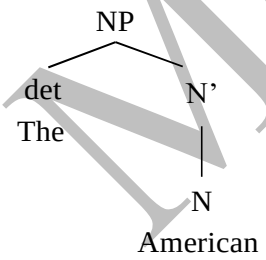
The letter X is used to signify part of speeches; when analyzing a specific utterance, specific categories are assigned. Thus, the X may become an N for noun, a V for verb, an A for adjective, or a P for preposition.

The term *X-bar* is derived from the notation representing this new structure. Certain structures are represented by X (an X with an over bar). Because this is difficult to typeset, this is often written as X′. In English, however, this is still read as "X bar". The notation XP stands for *X Phrase*, and is equivalent to *X-bar-bar* (X with a double over bar), written X″, usually read aloud as *X double bar*.

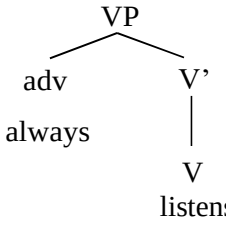
Three level structures are needed to express the relationship between head and their complements. Under the highest node of any phrase (XP) will be a specifier, which is optional, to the left which modifies everything generated under X' on the right.



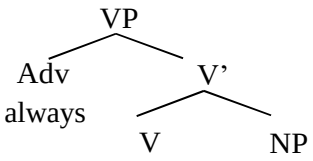
X-bar distinguishes syntactically among complements, specifiers, adjuncts. Here is an example with this structure which will generate the following NP:



The specifier *the* modifies the N *American*

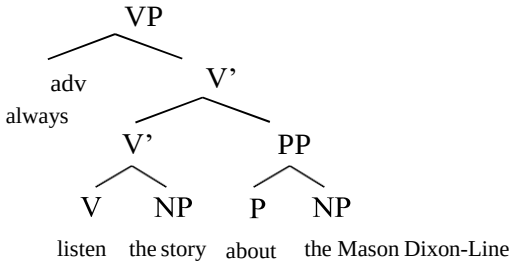


And the following VP The specifier *always* modifies the V *talks*

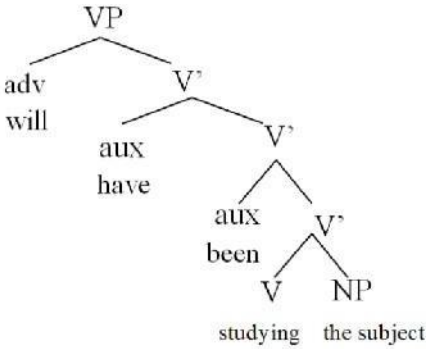


The expansion of X' level includes complements or adjuncts.

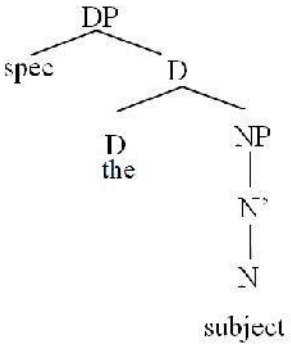
The adverb *always* modifies the V'. Under V' and to the left is the head of the phrase with its complement to the right. V' is the node which an adjunct is attached.

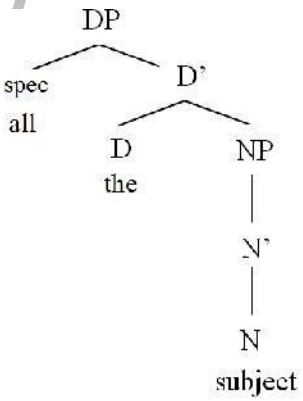


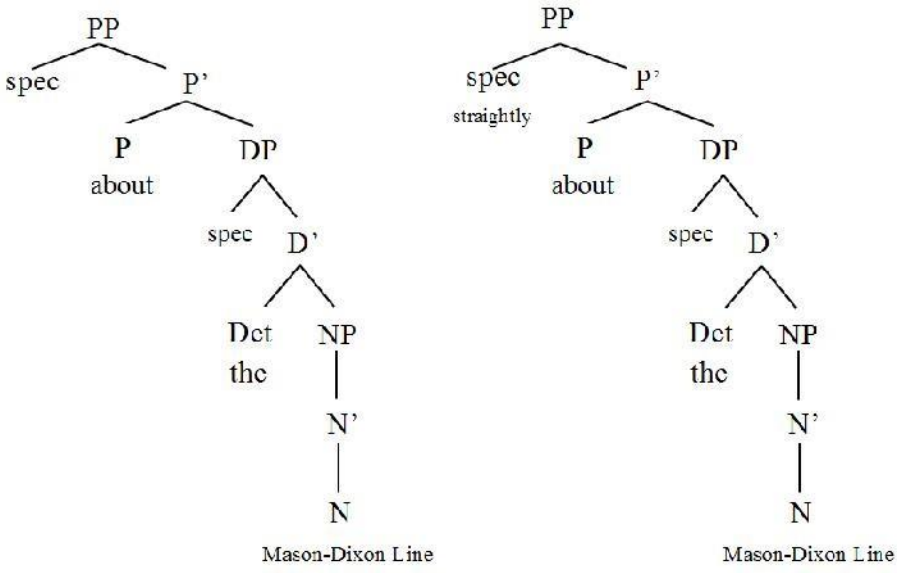
The repeated X' node happens in the following example:



Determiner under the N cannot be the head of the NP. To make the theory to be consistent, the specifier position is allowed to be empty. In this theory, NP is analyzed as DP (Determiner Phrase). In the following structure NP is as a complement to D.



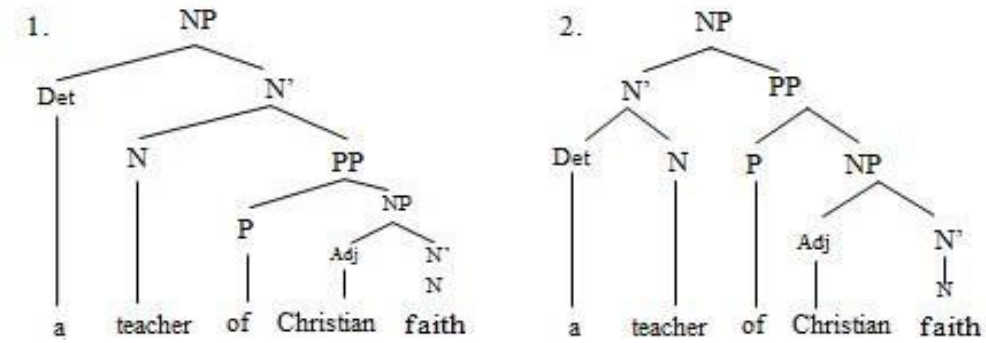




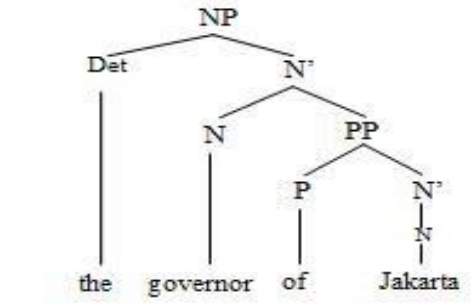
It is obvious that the X-bar theory offers a unified approach to all phrasal structures and simplifies the concepts of syntactic categories as well as syntactic nature of complements and adjuncts. The tree diagrams above can distinguish syntactically among complements, specifiers, and adjuncts in a phrase.

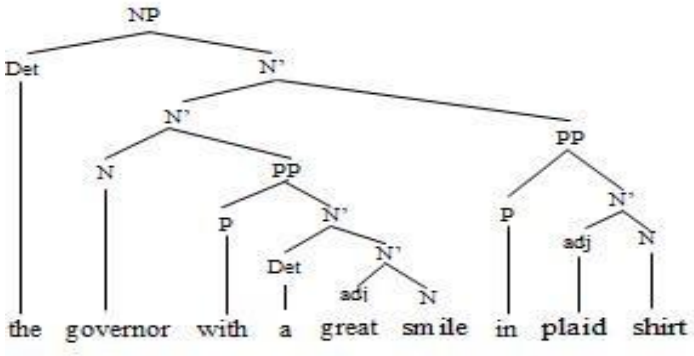
From the above descriptions about x-bar theory, analyzing sentences using an x-bar can avoid the repetitive use of phrasal categories (NP, VP, PP, AdvP or AdjP) within one sentence. We can differentiate different relationships of words in a noun phrase. Correctly represents constituents smaller than XP, bigger than X. The nature of the relationships of different NP's which serve as post modifiers for a deverbal noun can be determined.

Unlike traditional grammar, when we use X-bar, we can recognize ambiguity. A phrase "a teacher of Christian faith" can be clearly differentiated in x-bar syntax in the following way. In the first tree structure, the meaning of the sentence is the teacher teaches Christian faith. In the second one, the teacher is a person who has a Christian faith. The complement is closest to the head noun that is sister of the N, and the adjunct is sister of the N'.

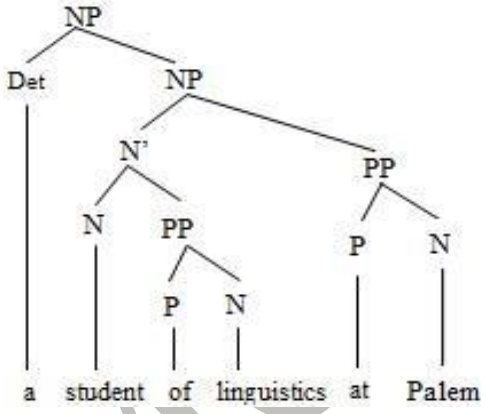


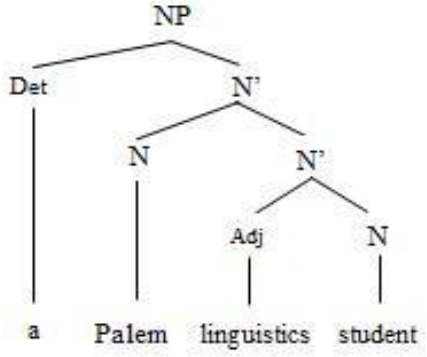
We can distinguish between a complement and adjunct in the distributional properties of PP (prepositional phrase). For example, it is more grammatical to have one complement in a phrase "the governor of Jakarta" than "the governor of Jakarta of Indonesia". We can also say a phrase which has more than one adjunct as in "the governor without moustache with a great smile in plaid shirt". The complement always precedes the adjunct. It is acceptable to say "the governor with a great smile in plaid shirt" and not "the governor in plaid shirt with a great smile." There is a possible way to coordinate either adjuncts or complements together as in "the governor with a great smile and plaid shirt" and not "the governor with a great smile and in plaid shirt".





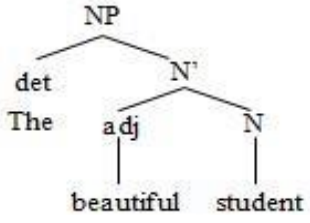
We can describe the occurrence of premodifiers, which are divided into determiners, attributes, and complements, using the x-bar syntax. This can be seen in comparing the following phrases:

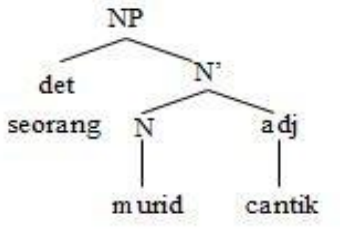




We see that PP *of linguistics* is a complement (modifies N) whereas the PP *at Palem* is an adjunct (modifies N'). In the second phrase, we see that *Linguistics* is a complement because it is the sister of the N *student*, whereas *Palem* would be an attribute because it is the sister (and daughter) of N'.

The x-bar syntax is also useful to analyze languages other than English. The theory takes into account N – ADJ in Indonesian as well as to ADJ – N in English. The semicolon that separates the constituents in the rule shows that the linear order is not fixed.





However, in writing the x-bar theory, we have to generate the articulated trees to replace the flat structure. It has three rules to be generated because it has three levels (NP as the phrase level, N' as the intermediate level, and N as the word/head level). NP is divided into (det) and N', N' is divided into (AP) N' or N' (PP), and N (PP). They are all binary branching, and all elements in x-bar rules are the projections of the head N. The NP represents the maximal projections and the N' represents the intermediate projections. This surely needs lots of space.

6. CONCLUSION

There are hierarchical orders in analyzing sentences. In traditional grammar, the hierarchy can be seen from the parts of sentences. That is to say, words are the constituents of the phrases and phrases are the constituents of sentences or clauses. The hierarchical order in immediate constituent analysis and X-bar theory can be seen clearly from the tree structures made as the last result.

At this point, it can be said that sentences have a hierarchical constituent structure in which words are grouped together into phrases and phrases into sentences or clauses. Furthermore, it is necessary to know the features of the categories as constituents of the own sentence. The features can be used to analyze sentences using those three ways explained above.

From the explanations given along with examples about those three theories chosen, each has its own way to analyze a sentence. The theories used to analyze sentences have their own strengths and weaknesses. They are understood as the problems that have to be dealt by language analyst. This also makes us realize that there is no perfect theory in analyzing a sentence.

REFERENCES

Bingfu, L. U. (2008). Direct constituent analysis: A reflection on the debates of sentence analysis methods [J]. Studies of the Chinese Language, 2, 006.

Carnie, Andrew.(2006), Syntax: A Generative introduction 2nd edition, Blackwell Publishing,.

Chomsky, N. (2002). Syntactic structures. Walter de Gruyter.

Damodar Thakur, (1998). Linguistics Simplified Syntax, ISBN: 97881- 7709-316-2.

Downing, Angela and Philip Locke (2006), English Grammar: A University Course: 2nd Edition, New York: Rout ledge

Haegeman, L. (1994), Introduction to Government and Binding Theory, Oxford, Blackwell.

Hockett, C. F. (1958). A course in modern linguistics. Language Learning, 8(3-4), 73-75.

Longacre, R. E. (1960). String constituent analysis. Language, 63-88.

Peters Jr, P. S., & Ritchie, R. W. (1972). Context-sensitive immediate constituent analysis: context-free languages revisited. Mathematical systems theory, 6(3), 324-333.

Pike, K. L. (1943). Taxemes and immediate constituents. Language, 65-82.

Pushpinder Syal, D.V. Jinda, 2007: An Introduction to Linguistics, ISBN: 978-81-203-3216-4

Radford, Andrew (1988), Transformational Grammar, New York: Cambridge University Press.

Stark, B. R. (1972). The Bloomfieldian Model. Lingua, 30, 385-421.

Stowell, Timothy A. (1981), Origins of Phrase Structure, Cambridge MA: MIT dissertation

Street, J. C. (1967). Methodology in immediate constituent analysis. In Approaches in linguistic methodology (pp. 89-114)

Wells, R. S. (1947). Immediate constituents. Language, 23(2), 81-117.

Syntax: Terms

Syntax

the module of the grammar that accounts for the **structure** (= the form) of **grammatically acceptable sentences**

Sentence

the "**highest**" (i.e., largest) syntactic unit

- the **lowest** (i.e., smallest) syntactic units are **words**
- the **intermediate** syntactic units are the **phrases**

Phrase

a word or group of words functioning as a syntactic unit between the level of individual words and the sentence as a whole

▪ **noun phrase (NP)** ▪ **verb phrase (VP)** ▪ **adjective phrase (AdjP)** ▪ **adverb phrase (AdvP)**
▪ **prepositional phrase (PrepP)**

Phrasal head

the **central word** of a phrase whose grammatical category defines the type of phrase

▪ **noun** → NP ▪ **verb** → VP ▪ **adjective** → AdjP ▪ **adverb** → AdvP
▪ **preposition** → PrepP

Noun phrase (NP)

phrasal (syntactic) category: contains a noun or pronoun as its head, and functions as the subject or as various objects in a sentence

Verb phrase (VP)

phrasal (syntactic) category: contains a verb as its head along with its complements such as noun phrases and prepositional phrases

Adjective phrase (AdjP)

phrasal (syntactic) category: contains an adjective as its head

Adverb phrase (AdvP)

phrasal (syntactic) category: contains an adverb as its head

Prepositional phrase (PrepP)

phrasal (syntactic) category: consists of preposition – the phrasal head – and a noun phrase

Adnominal prepositional phrase

occur **in NPs**, where they modify the head N or Prn

The old desk **by the window** suddenly collapsed. (= which desk?)

Adverbial prepositional phrase

occur **in VPs**, where they modify the head V

The old desk suddenly collapsed **after lunch**. (= when?)

Noun (**N**)

- syntactic category (part of speech) of words that function as the head of a noun phrase
- name (refer to) entities (things, people, ideas, concepts) in the external world

desk, **love**, **reading**

Adjective (**Adj**)

- syntactic category (part of speech) of words that function as the head of an adjective phrase
- express **attributes** of nouns

old desk, **real** love, **slow** reading

Verb (**V**)

- syntactic category (part of speech) of words that function as the head of a verb phrase
- denote actions, states, and sensations

The old desk **collapsed**. / Their love **died**. / Slow reading **is** the best.

Transitive verbs (**V_t**)

take a direct object - express actions which an **active subject** NP (= agent) does to a **passive object** NP
(= patient)

Robin **wrote** the letter.

Intransitive verbs (**V_i**)

take no direct object - express actions which a **passive subject** NP (= patient) experiences itself

Jack **slept** until noon. The bridge **collapsed**.

Adverbs (**Adv**)

- syntactic category (part of speech) of words that function as the head of an adverb phrase
- express **attributes** of Vs

suddenly collapsed, died **slowly**, is **often**

Function words

words that have no clear lexical meaning, but have grammatical functions

Pronoun (**Prn**)

syntactic category of words that replace NPs

it, **that**, **they**, **my**, **him**, etc.

Preposition (**Prep**)

- syntactic category of words that function as the head of a prepositional phrase
- relate NPs in various ways to other sentence constituents

by, **for**, **in**

Determiner (**Det**)

- syntactic (also functional) category of words and expressions which when combined with a noun form a noun phrase
- specify nouns in various ways

indefinite **a, an**, definite **the**, possession **their**, **many** quantity

Verbal Auxiliaries (**Aux**)

together with Vs, form complex verb expressions

have, do, be can, will

The bridge **has** collapsed.
Their love **is** dying.
Slow reading **can** be fun.
Do you know her?

Degree Words (**Deg**)

specify the degree of Adj or Adv

The **really** old desk collapsed **quite** suddenly.
Their **very** weak love finally died**rather** quickly.
Slow reading is **terribly** boring.

Conjunctions (**Conj**)

link other constituents in a sentence

The old desk collapsed, **but** their love remained true.
Slow reading **and** slow eating are great fun.

Phrase structure rules

principles of grammar that specify the constituency of syntactic categories

S → NP VP
NP → Det (AdjP) N (PrepP)

Phrase structure tree

- a tree diagram with syntactic categories at each node
- reveals both the linear and hierarchical structure of phrases and sentences

Deep structure

- any phrase structure tree generated by the phrase structure rules of transformational grammar
- the basic syntactic structures of grammar

Surface structure

- the structure that results from applying transformational rules to a deep structure
- the actual utterances in language use

Transformation rule, transformation

- syntactic rule that applies to the deep structure of a sentence and derives a new structure by moving or inserting elements