Studia Anglica Posnaniensia 35, 2000

SUBJECT AND PREDICATE AGREEMENT IN THE DEVELOPMENT OF GENERATIVE GRAMMAR

MICHAEL MOSS

University of Gdańsk

1. Introduction

In the following paper, I will attempt to assess the work undertaken until now to explain the phenomenon of agreement between the Subject and Predicate in a sentence. Some of the underlying questions that I will be asking are: When did generative linguistics break away from a string adjacent approach to agreement?; What kinds of mechanisms are proposed for communication between Subject and Predicate in absence of simple adjacency?; and finally, What questions still need to be answered in this area? I will go through the material chronologically in order to preserve the structure of the arguments as they were presented. Most of the material to be analyzed was produced by Noam Chomsky, although supporting and questioning materials have been introduced as appropriate.

2. The transformational approach to sentence construction

With *Syntactic structures* in 1957 Chomsky led linguistics into the realm of generative grammar. He outlined a program by which languages could be broken down to a set of rules which, while being a finite set, would allow for the production of an infinite set of properly formed utterances. The set of rules would also limit the language to properly formed sentences. The heart of the structure in 1957 looked like this:

(1) $\sum$: *Sentence* :

$X_1 \rightarrow Y_1$
:
$X_n \rightarrow Y_n$
} Phrase structure

T_1
:
T_j
} Transformational structure

$Z_1 \rightarrow W_1$
:
$Z_m \rightarrow W_m$
} Morphophonemics

(Chomsky 1957: 46).

Before going further, let us note how these ideas were developed. The phrase structure grammar was originally formalized in Chomsky (1957) as a finite number of rules of the form: $XAY \rightarrow XYZ$, where the arrow indicated 'is to be re-written as', and the following conditions were imposed:

(2) Condition (1) *X, Z,* and *Y* are strings of symbols (*X* or *Y* or both possibly null) but *A* is a single symbol.
Condition (2) *Z* is not null.
Condition (3) *A* is not identical with *Z* (Postal 1964: 143).

These conditions apply to both terminal and non-terminal constituents. Terminal constituents are those which can no longer be broken down by "re-write" rules. At this point, terminal constituents were labeled simply "morphemes". The non-terminal constituents, thus, are variables or higher symbols, which can be broken down. The maximum non-terminal constituent was (and still is) the sentence, given the symbol *S* (Postal 1964: 143).

Getting back to the model presented in (1), Chomsky presents the idea of the phrase structure grammar with *S* as the highest non-terminal constituent as *part* of the model describing natural languages. Here, the phrase structure constituent is made up of a group of unordered rules which "re-write" the constituents of a sentence. The constituents are broken down or "parsed" by the following kinds of rules:

(3) a. *Sentence* $\rightarrow$ *NP* + *VP*
b. *NP* $\rightarrow$ *T* + *N*

c. *VP* $\rightarrow$ *Verb* + *NP*
d. *T* $\rightarrow$ *the*
e. *N* $\rightarrow$ *man, ball,* etc.
f. *Verb* $\rightarrow$ *hit, took,* etc. (Chomsky 1957: 26)

It is important to note that Chomsky here refutes the idea that "phrase structure" grammars suffice to formulate adequate grammars for the description of natural languages.

The phrase structure model given above is powerful. It is a great improvement over traditional Markovian finite state machines,[1] which analyze the sentence in a linear process. Yet due to some of phrase structure's inherent similarities to a finite state machine, it quickly runs into difficulties itself. As such, it is still not powerful enough to adequately describe natural language on its own. The phrase structure model analyzes a non-terminal string *S* by the constituent parts, which make up the whole. As such, it does not run into the problems of linear analysis that Markovian models did. The phrase structure model is similar, however, in that it analyzes the constituents in a finite state manner: a single constituent is selected and a rule is applied changing the state of the machine. The next constituents are chosen and rules are applied, and so on until all the strings are "terminal". Its rules only operate on a finite state and as such do not allow for the complexity exhibited by language. Generally, three major problems were found with the phrase structure model:

(4) a. Phrase structure rules analyze *one* constituent at a time, i.e. no reference can be made to other constituents (terminal or non-terminal) in their fulfillment.
b. Phrase structure rules cannot make discontinuous analysis, they cannot take separated elements into consideration such as *have ... en*, or *be ... ing*.
c. Phrase structure rules are "top-to-bottom" – "the state of the machine is completely determined by the string it has just produced" (Chomsky 1957: 37-38).

To make the phrase structure model usable, special more insightful rules were admitted into the grammar called *transformational rules*. Grammatical transformations operate on entire strings, or even sets of strings instead of single constituents. As such, while still being "re-write" rules, they allow many constit-

[1] Such as the ability to place entire S structures into another S structure "infinitely" and still produce a gramatically correct sentence. For more detailed information, see Chomsky (1957).

uent-based rules to be executed at the same time. Transformations were introduced to eliminate the problems illustrated in (4), i.e., they are string-based, allow for discontinuous analysis and are "left-to-right". Some transformations are obligatory, while others are not.

It is important to note that the phrase structure component is not capable of producing sentences on its own. The transformational component provides some activity basic to the production of sentences. I shall illustrate below what happens step by step during the formation of a sentence using the model shown in (1).

The phrase structure of this grammar produces a terminal string using rules as listed in (3). That is, S is broken down into NP and VP. The noun phrase is expanded to a Noun marked as singular or plural noun and a determiner. Finally, the verb phrase is expanded to a *Verb* and an optional NP where *Verb* consists of *Aux* (helper or modal verb plus the element *C*) plus a verb. This process is finished when a terminal string is produced, i.e. "a sequence of morphemes, though not necessarily in the correct order" (Chomsky 1957: 46).

Having finished with the phrase structure rules, the sentence is subjected to the obligatory (and optional) transformations. One of these transformations (shown below in (5)) deals with the element *C*, mentioned above. *C* is a single symbol, and as such can be introduced by the phrase structure rules. What it represents, however, is too complex to be dealt with by the phrase structure due to its discontinuous nature and the dependencies on the presence or absence of other elements in the sentence:

(5) $$C \rightarrow \left\{ \begin{array}{l} \text{S in the context } NP_{sing}\_\_ \\ \varnothing \text{ in the context } NP_{pl}\_\_ \\ \text{past} \end{array} \right\}$$ (rule (29i) in Chomsky 1957).

Here *C*, the manditorily filled position occurring directly before *Verb* as part of the *Aux* rewriting rule, is replaced by a morpheme or tense, depending on the presence of other types of morphemes in the string. Thus the string *Bill + eat* would be rewritten as *Bill + S + eat*. The further placement of the morpheme *S* is also taken care of by a transformational rule:

(6) Let *Af* stand for any of the affixes *past, S, ∅, en, ing*. Let *v* stand for any *M* or *V*, or *have* or *be* (i.e. for any non-affix in the phrase *Verb*). Then:
$Af + \rightarrow v + Af\#$
where # is interpreted as word boundary ((29 ii) in Chomsky 1957: 39).

The real problem that Chomsky was trying to deal with here was the problem of dependencies and discontinuous elements. Verbal agreement with the Subject of the sentence is basic to any theory of language. Yet, the phrase structure rules

were not strong enough to deal with this problem in any graceful way. The transformations allowed the grammar to take two disjoint elements, analyze them and make a decision. Still, however, the transformations are bound to the strings of morphemes which result from the constituent analysis. As such, the transformations are completely string-dependent. The element *C* must be directly adjacent to an NP marked singular in order for the morpheme *S* to be inserted. Furthermore, the transformations need to "know" quite a bit about the derivation of the morphemes as they appear in the string. The element *C* is inserted after an NP, not just an N, and it is attached later to the *Verb* itself and not the entire VP. These relations are illustrated in the examples below:

(7) a. The man in the car drives around the block.
$_S[_{NP}[_{NP}[\text{The man}]_{NP\ PP}[\text{in the car}]_{PP}]_{NP\ VP}[_V[\text{drives}]_{V\ PP}[\text{around the block}]_{PP}]_{VP}]_S$

b. The phrase structure produces the following string of morphemes:
The+man+in+the+car+*C*+drive+around+the+block

c. (5) takes the NP (*the man in the car*) which is marked "sing" by the main NP "the man" and re-writes the abstract element *C* as the morpheme *S* which gives us:
The+man+in+the+car+S+drive+around+the+block

d. Affix hopping (6) then takes care of the order of *S* in the sentence producing
The+man+in+the+car+drive+*S*+around+the+block.

e. Finally, we take care of the word boundaries with Chomsky's (1957: 39) rule (29iii):
Replace + by # except in the context *v – Af*. Insert # initially and finally:
The#man#in#the#car#drive+S#the#car#around#the#block

Even in an example like this one, however, it must be noted that many problems immediately come to the foreground. First of all, if the nominal constituent which agrees with the verb is embedded in the subject NP as its non-final item, it is not explained how the information from the inner NP is transferred to the outer NP for the purposes of agreement. Secondly, the NP is a complex NP with a PP complement. How the NP and the PP are "joined" to form a complex NP is at this point left unanswered. These problems will, of course, be addressed in later works.

According to this interpretation, in English only one Verb takes tense, i.e. the Verb appearing leftmost in the VP. Using Chomsky's own example (1957: 40-41), we start with the string:

(8) a. The+man+C+have+en+be+ing+read+the+book.
b. C is changed into S as above, which gives:
The+man+S+have+en+be+ing+read+the+book
c. Affixes are reordered:
The+man+have+S+be+en+read+ing+the book
d. The morphosyntactic processes take care of the rest:
The man has been reading the book

The grammatical transformations were used to simplify the grammar by taking certain "kernel" sentences and making standard stylistic changes to them. This meant fewer specific rules, and simplified matters by working on entire structures. The passive is a good example of this process:

(9) *Passive*
Structural analysis: $NP - Aux - V - NP$
Structural change: $X_1 + X_2 + X_3 + X_4 \rightarrow X_4 + X_2 + be + en + X_3 + by + X_1$
(Chomsky 1957, Appendix II, Rule 12)

This rule apparently changes *The man in the car drives the dogs around the block* into *The dogs are driven around the block by the man in the car*. The underlying structure is as follows:

(10) a. *Kernel Sentence*
$[\text{The man in the car}]_1$ + Aux_2 + drive_3 + $[\text{the dogs}]_4$ +
+ $[\text{around} + \text{the} + \text{block}]_X$

b. *Passive Transformation*
$[\text{the dogs}]_4$ + Aux_2 + be + en + drive_3 + $[\text{around} + \text{the block}]_X$ +
+ by + $[\text{the man in the car}]_1$

c. *Rule 29i* (=(5) above)
$[\text{the dogs}]_4$ + Aux_2 + be + en + drive_3 + $[\text{around} + \text{the block}]_X$ +
+ by + $[\text{the man in the car}]_1$

d. *Affix Hopping*
$[\text{the dogs}]_4$ + be + pres + drive_3 + en + $[\text{around} + \text{the} + \text{block}]_X$ +
+ by + $[\text{the man in the car}]_1$

e. *Morpholexical Transforms*
The dogs are driven around the block by the man in the car.

This example brings several questions to light. Namely, when is the kernel sentence actually formed? What are these rules actually operating on: single elements or bracketed elements? And finally, when is the subject/verb relation established? Explicitly, in order for the passive transformation to function correctly, it must occur very early in the process of sentence production. Otherwise,

the subject/verb relation does not function properly. It must, in fact, occur even before the kernel sentence is fully formed, or at least before the phrase structure rules are fully realized. Furthermore, there is the problem of what to do with the PP accompanying "drive". In the "kernel sentence", it is not attached to the verb as part of the VP, at least not in a "linear" sense, as the rest of the constituents are analyzed. It is only attached to the verb (as we can see in (10) above) in terms of the labeled bracketing as part of the VP.

Thus, it appears that, on the one hand, linear order is the deciding factor in the way the Phrase Structure and Transformations are performed: C is rewritten as S if C is directly to the right of an NP_{sing}. On the other hand, however, it is clear that structure plays an important role. In both (7) and (10) above, it is the entire NP which is taken as the subject of the sentence, and not just the leftmost N within that NP. There also seems to be a bracketing problem: the passive rule moves the object NP, found immediately to the right of V, to the beginning of the sentence to act as the subject NP. In fact, the object NP is *extracted* from the entire VP (as can be seen in (10) above), leaving the rest of the VP intact. The transformational rule, however, does not mention VP, but only V, so this kind of extraction is left unexplained.

According to Chomsky, the transformations operate on strings with a particular structure: "A transformation will operate on a string of symbols *with* a particular structural description ... and will convert it into a new string of symbols with a new structural description" (Chomsky 1964b: 222). He does not elaborate as to how sentences in which the subject and object vary according to number are analyzed.

Most importantly, however, we can see that at this stage of development, subject/verb agreement is determined by the physical location of both constituents. NP_{sing} triggers the change of *C* to *S*, which is then passed on to the adjacent verb. When the subject and verb must be physically adjacent in the sentence in order for their relationship to be established, I will refer to this as being *string adjacent* or *string adjacency*. As mentioned above, subject/verb agreement at this point is determined by a string adjacent relationship between an NP and an *Aux*, and not by a relation between the subject and the verb themselves.

3. Introduction of "grammatical relations"

Where earlier the relation between Subject and Verb/Predicate in a sentence was not well defined, in 1965 Chomsky made several important observations. Most importantly, he made the distinction between "grammatical functions" and "grammatical categories". *Subject of*, *Object of* and *Predicate* were now recognized as grammatical functions, which meant that they held a special relation between themselves. This is different from the grammatical categories, which only represent the category to which the individual items in a sentence belong

(N, V, S, VP, NP etc.). Grammatical categories exist no matter what relation the items are in with other items in the sentence.

(11) The notion "Subject", as distinct from the notion "NP", designates a *grammatical function* rather than a *grammatical category*. It is, in other words, an inherently relational notion. We say in traditional terms, that in (1) [Sincerity may frighten the boy.] *sincerity* is an NP (not that it is the NP of the sentence), and that it is (functions as) the *Subject-of* the sentence (not that it is a Subject). Functional notions like "Subject", "Predicate" are to be sharply distinguished from categorial notions such as "Noun Phrase", "Verb", a distinction that is not to be obscured by the occasional use of the same term for notions of both kinds. ... It is necessary only to make explicit the relational character of these notions by defining "Subject-of", for English, as the relation holding between the NP of a sentence of the form NP^Aux^VP and the whole sentence [footnote omitted], "Object-of" as the relation between the NP of a VP of the form V^VP and the whole VP etc. (Chomsky 1965: 68-69).

The relations, then, that are being discussed are between an item and a larger category containing it. Within an S there is both an NP and a VP. Within a VP there is a V and optionally an NP. The relations specifically were defined as follows:

(12) a. Thus subject-Verb can be defined as the relation between the Subject-of a Sentence and Main-Verb-of the Predicate-of the Sentence, where Subject-of, Main-Verb-of and Predicate-of are the notions of [12b]; and Verb-Object can be defined as the relation between the Main-Verb-of and the Direct-Object-of a VP. (Chomsky 1965: 73).

b. (i) *Subject – of:* [*NP,S*]

(ii) *Predicate – of:* [*NP,S*]

(iii) *Direct object – of:* [*NP,VP*]

(iv) *Main – Verb – of:* [*VP,V*] (Chomsky 1965: 71).

This is actually quite a radical way of presenting things. The Subject is only the subject of S. It is related to the VP of the sentence only through S. The relationship to the "main verb" of the sentence is even more distant as it is related through VP by S. This is not the same situation with the direct object, which is directly related with the VP. Although, strictly speaking, the direct object and the main verb are only directly related to the VP.

Despite the significant breakthrough in the inclusion of grammatical functions into the realm of generative grammar, the definitions of the relations are still quite inconclusive. As a matter of fact, it would seem that the string adjacency principle still applies. Several rules were introduced into the grammar which tried to compensate for the difficulty which phrase structure rules were having with the above mentioned relations. Among the most important was (57xiv), which has been reproduced below:

(13) $[+\mathrm{V}] \rightarrow \mathrm{CS} / \alpha \hat{}\, \mathrm{Aux} \pm (\mathrm{Det} \hat{}\, \beta)$

where α is an N and β is an N (Chomsky 1965: 107).
(CS means Complex Symbol. In this symbol all of the relevant syntactic information related to the Verb was also elaborated.)

This rule ensures that the Verb, Subject and Object are in the correct *adjacent* positions in the sentence. These rules were called *selectional rules*. Their function was defined as follows:

(14) ... a selectional rule, such as (57xiv) [=13 above], (57xv), defines a *selectional relation* between two positions in a sentence – for example, in the case of (57xiv), the position of the Verb and that of the immediately preceding or following Noun. Such selectional relations determine grammatical relations ... (Chomsky 1965: 113).

In a mechanical way, then, we now have the Subject, Verb and Object relations defined for us. But we still do not know how the subject and verb manage to agree. The CS is important as it links the Verb with the kinds of Nouns that can fill slots α and β. The CS further provides a way for the Verb to define what sorts of Subject and Object are permissible, and also to define which Nouns may fill the slot. However, the only criteria available here are: ±Abstract, ±Animate, ±Count, ±Human etc. The CS does not appear to account for the features relating Gender, Number or Case.

Features relating to the NP (and the N in particular) are taken up later in the work, and are not clearly linked to the formation of a sentence, but rather to the internal agreement within an NP. First it was decided that, in addition to the marked features mentioned above, each N must have a set of feature specifications for Gender, Number and Case. Then it was decided that the transformational component must be expanded to include rules "that alter and expand the matrix of features constituting a lexical item" (Chomsky 1965: 174). To take care of agreement between an Article and an N in an NP, the following rule was proposed:

(15)

$$\text{Article} \rightarrow \begin{bmatrix} \alpha\ \text{Gender} \\ \beta\ \text{Number} \\ \gamma\ \text{Case} \end{bmatrix} \quad /\!\!-\ \ldots \begin{bmatrix} +\text{N} \\ \alpha\ \text{Gender} \\ \beta\ \text{Number} \\ \gamma\ \text{Case} \end{bmatrix}$$

where Article ... N is an NP (Chomsky 1965: 175).

This allows an article to agree with the N within the same NP in Case, Number and Gender. Nothing is said, however, about the NP itself.

Further complicating the situation, there seems to be a conflict surrounding the entire issue of agreement and word order. The conflict is centered on the development of the Deep Structures (produced by the base) and the Surface Structures (produced by the transformational rules). On one hand, Chomsky states that the grammatical relations are determined by the base rules in the deep structure through concatenation-rules such as:

(16) $\text{S} \rightarrow \text{NP}\hat{}\text{VP}$

$\text{VP} \rightarrow \text{V}\hat{}\text{NP}$ (Chomsky 1964b: 124).

Such rules not only state the contents of the given category, but the order in which the elements must appear.[2] On the other hand, the domains [NP, S], [NP, VP] etc. seem in reality to be determined only at the surface structure.

Problems appear when the surface structure does not accord with the deep structure in terms of constituent order. Such a situation produces what Chomsky refers to as *logical* elements, i.e. the Logical Subject or the Logical Object with verbs like *strike*, *please* and *frighten* in the deep structures of sentences as: *he was struck by a bullet*, *he is easy to please*, and *he frightens easily. He* in these sentences is termed the Logical Object since its role in the deep structure is the object being acted on: *a bullet struck him*, *it is easy to please him*, *it is easy to frighten him* (Chomsky 1964: 221-222). The fact that *he* in these sentences is the subject was seen to be the result of the passive transformation acting on the deep structures, producing a new constituent order. Thus as stated earlier in Chomsky (1957), agreement is taken care of during the transformational level of the grammar, only after the surface structure has been developed enough to determine the

[2] Other models based on a set model of presentation using rules such as: $\text{S} \rightarrow \{\text{NP, VP}\}$ and $\text{VP} \rightarrow \{\text{V, VP}\}$, in which the order of the elements is "free", were rejected at this stage due to the fact that the transformational component would then be enlarged by a second set of "ordering" rules, which would make the system as a whole more complicated. Work using the set model of Phrase Structure rules was developed, however, starting with Gazdar (1981), and others following.

final word order. Agreement is not directly related to the deep structure except in the logical representation of the sentence.

Thus in *Aspects*, agreement is the result of a transformation which is performed directly on the surface structure, which, thus, takes place only after all other transformations have taken place, and the surface structure is set in place. Rules defining the relation of *Subject-of*, *Verb*, *Object-of* have been introduced, but the mechanisms describing the agreement between these elements has mysteriously been left alone. In *Syntactic structures*, it is evident that an NP can be marked for Number (NP_{sing}). In *Aspects*, the N (and subsequently the Article) is also marked for Number. Perhaps the NP and VP could be "allowed" to agree using a similar rule to (15) above, where S is the domain in which the NP and VP must be found (NP...VP is an S). Such a rule could be formulated as:

(17)

$$\text{VP} \rightarrow \begin{bmatrix} \alpha\ \text{Gender} \\ \beta\ \text{Number} \end{bmatrix} / \begin{bmatrix} +\text{NP} \\ +\text{Nominative} \\ \alpha\ \text{Gender} \\ \beta\ \text{Number} \end{bmatrix} \ldots —$$

where NP...VP is an S.

Two problems arise with this formula: (1) Can a constituent pass its markings up to a major category (N to NP); (2) can a major category assign properties to its constituents (VP to V). These questions obviously mirror the same problem. We can ask them slightly differently. In the above formula, we can get the NP and the VP to agree, but how do we allow the constituents within the NP and VP to convey the necessary information over the category "boundary". Furthermore, using the definitions in (12), *Subject-of* and *Predicate-of* are only defined in terms of S. Communication between the categories themselves is not mentioned in these definitions, so how do we deal with the passage of information through S itself?

4. The role of subcategorization, the lexicon and the $\overline{\text{X}}$ template

In 1970 Chomsky published *Remarks on nominalization* in which he introduced the "Lexicalist Hypothesis" and proposed that the Phrase Structure component be simplified using the X-bar theory to universalize phrase structures. This is a turning point in the development of generative grammar. It marks a departure from analyzing the sentence as a structure built on string adjacent relations to one based on previously known structures. Using the X-bar theory, structures could be used as building blocks held together by underlying relations.

In a trade, Chomsky argues that the transformational component of the grammar can be simplified by enriching the base. The model presented in (1) has been modified, the Phrase Structure rules are now part of a sub-component of

the syntax called *the base*, which produces the kernel sentences. The transformational rules are in a separate sub-component called the *transformational sub-component*, which performs transformations on the kernel sentences. These modifications were introduced in Chomsky (1965: 17). In 1970, the base was further modified by splitting it into the *lexicon*, and the *categorial component*. The most important feature of the lexicon is that its entries are "item[s] with certain fixed selectional and strict subcategorization features, which [are] free with respect to the categorial features [noun] and [verb]" (Chomsky 1970a: 21). That is to say, a lexical entry is marked for possible complements. "The fact that *refuse* takes a noun phrase complement or a reduced sentential complement, either as a noun or as a verb, is expressed by the feature structure of the "neutral" lexical entry, as are selectional properties" (Chomsky 1970a:21). Thus the base is a bi-directional system of rules describing the infinite set of possible phrase constructions, with the lexicon setting restrictions on the individual complement options open to each word chosen.

Such a move was extremely helpful as it took an immense burden off of the transformational component. Structures such as **John's easiness to please* were now ungrammatical because of the inherent qualities of the word *easiness*, which is marked as not being able to take a S or $\overline{S}$ complement. As was suggested in 1965, the lexicon is now seen as the holding place for individual and idiosyncratic information related to each word used in the language, allowing the words themselves to determine the "correctness" of the structure that they are used in.

The transformational component has been substantially simplified, reduced to more general rules such as Agent-postposing and NP-preposing for the passive operation. Furthermore, transformations are no longer bound to a specific category. That is, the passive operation works equally well on S of the form $NP - Aux - V - NP - by\ \Delta$, as it does on NP of the form $Det - N - NP - by\ \Delta$.

The introduction of "agent" and "causative" show that Chomsky at this stage was already considering what would later be called *θ-roles*. These two ideas give another level of depth to the structure of the sentence independent of the Phrase Structure order and the "grammatical relations". For instance, the agentive quality of a phrase in a sentence is not affected by either its physical place in the structure S or its relation to S ([NP, S] or [NP, VP]). It is important to note that the lexicon was capable of assigning such features.

By enriching each lexical position with subcategorization frames, limiting its complement options, the transformational component could be dramatically simplified. Next, the decision "to replace *categories* systematically by *features* that can enter into complex symbols" (Chomsky 1970a: 52, emphasis mine) allowed for the simplification of the phrase structure component. Now all phrases

(AP, NP and VP) were reduced to $\overline{\overline{X}}$, which was broken down by the following rule (Chomsky's 1970a: 48):

(18) $\overline{\overline{X}} \rightarrow [\text{Spec}, \overline{X}]\overline{X}$

where [Spec, $\overline{N}$] will be analyzed as the determiner, [Spec, $\overline{V}$] as the auxiliary (perhaps with adverbials associated), and [Spec, $\overline{A}$] perhaps as the system of qualifying elements associated with adjective phrases (comparative structures, *very*, etc.) (Chomsky 1970a: 52).

This gave the phrase structure component a much simplified model to follow, reducing and simplifying its functions in the grammar. It also meant that language was based on structures such as $\overline{\overline{X}}$, which each had very similar attributes regardless of what X was replaced by (A, N, V, P, S, etc.). The Lexicalist Hypothesis and its implications allowed Chomsky to move away from a string analysis of linguistic data to a more hierarchical approach. The introduction of the $\overline{X}$-structure allowed for the redundancies in the transformational and phrase structure components to be greatly reduced.

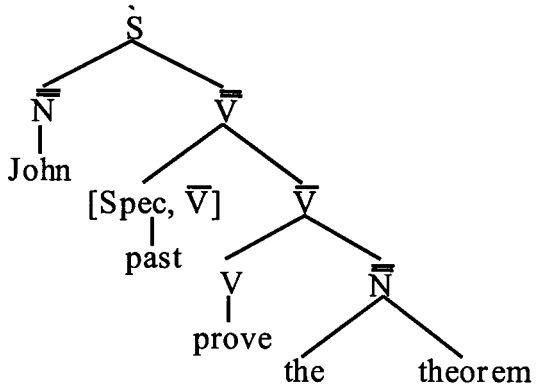
The introduction of the $\overline{X}$-structures in conjunction with the simplified transformational component answers some of the difficulties that we encountered earlier in (10). Thus, in making the active sentence *The man in the car drives the dogs around the block* passive, NP-preposing places *the dogs* into the [NP, S] position, and Agent-postposing places the entire agent NP into the passive *by* construction. The V *drives* and the PP *around the block* have no other open slots to appear in, and thus must be adjacent. It does not, however, deal with any of the other problems mentioned concerning agreement.

Furthermore, although this is mentioned in passing, it is important to notice that with the $\overline{X}$-structures, we are also introduced to the notion of percolation:

(19) One might extend this operation of *there*-insertion, introducing the complex symbol [*there*, +NP, α *plural*] (α = + or α= –), where the third term in the proper analysis (*a man*, in the cited example) is [α *plural*], plurality now being regarded as a feature that *ascends from a head noun to the NP node dominating it* (Chomsky 1970a: note 31, emphasis added).

We now have a much more powerful method of analyzing the grammatical relations. Let us take Chomsky's (1970a: 53) example (52):

(20)

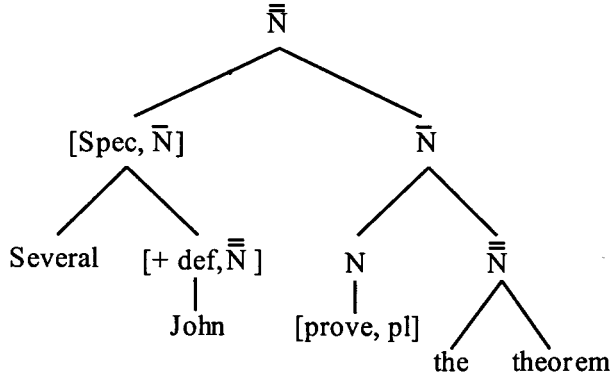


John proved the theorem.

All of the lexical information for John [NP, S] is passed up from the lexical item N to the N̿ through percolation. Furthermore, all of the information for *Aux* is passed up to the [VP, S] node (here the [Spec, V̄]). Thus, if we were to use a transformational rule like the one presented in (17), we could now account for the ability of N̿ and V̿ under S to acquire the appropriate information about their constituents. Through subcategorization, of course, we can also account for the ability of the verb *prove* to "take" an abstract Object as its complement. Further, if we adopt the terms *agent* and *agent-postposing*, we can then account for the passive construction *The theorem was proved by John.*

We also have the observation that S and NP have a very similar internal relation. Take, for instance:

(21)



Several of John's proofs of the theorem (Chomsky 1970a: 53).

Here Chomsky clarifies that in this NP, *John* and *proofs* are the "heads" of their respective phrases.[3] Selectional features limit the semantic context and subcategorizational features limit the structural context in which the heads of the major categories in the grammatical relations can be found. Thus the NP reflects the structure of S with *John* playing the function of *Subject-of* i.e., [NP, NP].

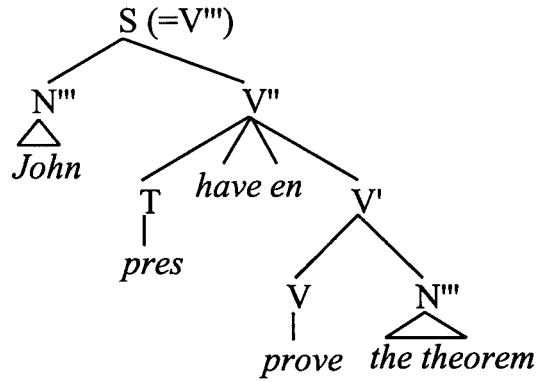
Jackendoff, however, is not satisfied with the structure presented by Chomsky, saying that the subject in the NP and the S need to be in identical structures. Thus, he proposes several modifications to the basic X̄-template, with the intent of separating S, and improving on the structural symmetry between the two structures and the representation of grammatical relations. First, he disposes of the [Spec, X̄] position, pointing out that it is not a syntactic category. Further, he universally imposes a three tiered structure on all categories, i.e. "X'''" is a *major lexical category*" in his Uniform Three Level Hypothesis

(22) $$X^n \rightarrow (C_1)\ldots(C_j) - X^{n-1} - (C_{j+1})\ldots(C_k)$$

where $1 \leq n \leq 3$, and for all C_i, either $C_i = Y'''$ for some lexical category Y, or C_i is a specified grammatical formative (Jackendoff 1977: 255).

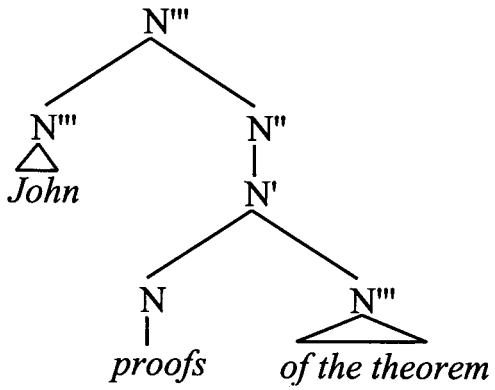
This allows him to formulate the following structures in which the subject of an NP and that of an S are in identical structural positions (Jackendoff's 1977 (14a) and (14b)):

(23) a.



[3] The idea of head seems to be taken from Chomsky (1965) where *head* is apparently defined as the *lexical category* (N, A, V, P) dominated by a *major category* (NP, VP, AP, PP). (See Chomsky (1965: 74, 106, 113-114) for more information.)

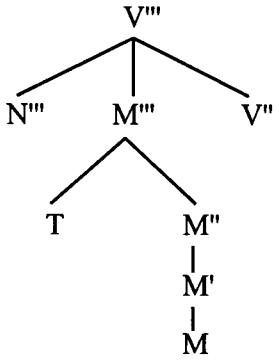
b.



(Jackendoff 1977: 258).

Further modifications to the *Remarks on nominalization* structure is to formulate a category M which replaces the Aux position. M" branches into Tense and Modal giving the following breakdown:

(24) a. V''' → N''' – M''' – V"
b. M''' → T – M"
c. M" → M'
d. M' → M



(Jackendoff 1977: 264).

So, Jackendoff formalizes the structure of Aux under $\bar{X}$-theory, and brings Tense and Modals back to the level of S, at a level equal to both the predicate and subject. However, his structure does raise an important question. According to this model, S is equal to V''' making V the ultimate head of sentence. How then, does a sentence differ from a VP?

5. The introduction of conditions into generative grammar

In the late 60's, the Standard theory (as set forth in *Aspects*) was extended in view of the problems dealing with semantics and their relationship with Σ. The extension was stated as:

(25) ... semantic interpretation is held to be determined by the pair (deep structure, surface structure) of Σ, rather than by the deep structure alone; further, it is proposed that insofar as grammatical relations play a role in determining meaning, it is the grammatical relations of the deep structure that are relevant (as before), but that such matters as scope of "logical elements" and quantifiers, coreference, focus and certain kinds of presupposition, and certain other properties, are determined by rules that take surface structure (more precisely, phonetically interpreted surface structure) into account (Chomsky 1970b: 134).

Importantly, the idea of cyclic nodes was introduced – specifically regarding NP and S as the bearers of Subject. Furthermore, when the NP is the result of a nominalization, it also keeps all of the arguments of the original predicate, including the complements as well as the Subject.[4]

(26) a. Within the extended standard theory, as developed in the references cited earlier, both NP and S are nodes to which cyclic operations apply, and the notion "subject of" is defined not only in S but also in such NPs as [26b], where *John*, in all cases, is the "subject", in an extended sense of the term:

b.
(i) *John's refusal to leave*
(ii) *John's picture of Bill*
(iii) *John's strategy for victory* (Chomsky 1973: 239).

Thus, grammatical relations are produced at deep structure, and they are brought to the surface structure (and sometimes modified) by transformations. All other semantic operations are taken care of at the surface level. Additionally, as we saw in Chomsky (1970a), NP and S are both seen to have the grammatical relation: *Subject*. This will be of importance in the following developments.

[4] All nouns are able to open the subject position in the NP of which they are the head. Deverbal nouns are additionally able to open positions for internal arguments inherited from the underlying verb.

6. *Conditions on transformations* – restricting movements

Until *Remarks on nominalization*, much of grammar was taken to be of a "sentence by sentence" nature. While some generalizations had been noted, and transformations formed in an attempt to explain their existence, it was found that there were too many variations available and that linguistic structures were too rich to be accounted for using cyclic transformations, which would only re-order the constituents from so called kernel sentences. Many such sentences and constructions were noted in groundbreaking work done by John Ross (1967). The introduction of $\overline{X}$ theory in 1970 brought standardized and universal structures to generative linguistics. Instead of working on strings of words, constantly re-building everything from scratch, linguistic analysis was now dealing with a group of known structures, which existed as the architecture of linguistic structures. The problem now was to produce conditions capable of governing and explaining their interaction.

In 1973, Chomsky's *Conditions on transformations* attempted to apply such conditions to limit the application of transformations in grammar within the framework of the extended standard theory. The most basic of these conditions, the "A-over-A" condition, was introduced in a slightly different form much earlier, in 1964, and subsequently analyzed by many people, most notably Ross (1967) and Horn (1974). The 1973 version is reproduced here (Chomsky's (3)):

(27) If a transformation applies to a structure of the form
$[_{\alpha} \ldots [_{A} \ldots] \ldots]$
where α is a cyclic node, then it must be so interpreted as to apply to the maximal phrase of the type *A* (Chomsky 1973: 235).

With the "A-over-A" condition, Chomsky was able to introduce the idea of domains and what would eventually lead to binding. At first glance, it is not easy to grasp the power of this condition. By nature, it is abstract enough to operate over all types of syntactic transformations. Fundamentally, it constrains a transformation from applying to a non-maximal constituent within an NP or an S. That is to say, a transformation cannot apply to a single NP within a coordinate NP such as [$_{NP}$ John and Bill], but must be applied to the whole NP. Using Chomsky's examples, the "A-over-A" condition prevents:

(28) *John and Bill saw Mary,*

from undergoing the passive transformation to produce:

(29) *John and Mary was seen by Bill.*

This has special import for Subject/Verb agreement. Namely, a finite V within a VP cannot agree with an embedded NP, but only with an entire NP. Further conditions, introduced in the same work, will indicate more precisely the domain in which the Subject NP can be found.

The two main conditions introduced in Chomsky (1973), the Tensed-S Condition and the Specified Subject Condition (SSC), are reproduced below:

(30) *Tensed-S Condition*
No rule can involve *X, Y* in the structure
$\ldots X \ldots [_{\alpha} \ldots Y \ldots] \ldots$

where α is a tensed sentence (Chomsky 1973: 238);

(31) *Specified Subject Condition (SSC)*
No rule can involve *X, Y* in the structure
$\ldots X \ldots [_{\alpha} \ldots Z \ldots$ — *WYV*$] \ldots$
where *Z* is the specified subject of *WYV* in α (Chomsky 1973: 239).[5]

That is to say, they are conditions which constrain transformations which operate on constituents in two distinct domains. What is interesting about α in these conditions is that it must have a subject or a finite verb. In one condition α is a tensed sentence and in the other it is a domain containing a specified subject, i.e. S or NP. In other words, a verb may not agree with more than one subject and a subject may not cause changes within the domain of another subject [NP, S]. A subject may agree with more than one Verb as long as all of the verbs are contained in a VP at the same level as the subject NP, i.e. within the same S. As Chomsky points out later in the paper, the SSC tightens the domain between the Subject NP and the finite V by forcing the verb to "seek" out the nearest NP as its subject.

(32) Thus the Specified Subject Condition, in some cases, has the effect of reducing ambiguity, or, to put it differently, of increasing the reliability of a reasonable perceptual strategy that seeks the nearest NP to a verb (or the head noun of a nominal phrase) as its subject (Chomsky 1973: 257).

This is especially true when the SSC is combined with the Subjacency/Superiority relations:

(33) *Superiority*
A is "superior" to *B* if every major category dominating MMC(*A*) dominates MMC(*B*) as well but not conversely, *where MMC(X) is the minimal*

[5] There seems to be a problem with this rule, however. As we know, the NP in [NP, VP]$_S$ is not the *Subject of* VP, but of S. Thus it appears that in (31), *Z* is actually the *Subject of* α and not of *WYV*. Many thanks to Piotr Ruszkiewicz for this information.

major category dominating X (X itself, if X is a major category) (Chomsky 1973: 246 note 27), (emphasis mine);

(34) *Subjacency*
Where *X* is superior to *Y*, *Y* is "subjacent" to *X* if there is at most one cyclic category *C* MMC(*Y*) such that *C* contains MMC(*Y*) and *C* does not contain *X* (and therefore does not contain MMC(*X*) (Chomsky 1973: 250),

which allow one to specify more clearly the domain in which a Subject NP and finite V must be found. A subject cannot agree with a subjacent VP, nor can it agree with a superior VP. Seeking out the nearest NP as a subject, however, is quite a different approach than that found in Chomsky (1957). No longer are Subject and Predicate forced to be adjacent in the surface structure string. The domain of Subject and Predicate has been set to S, as was laid out first in Chomsky (1965).

Chomsky (1973) introduces a further modification to the general organization of the sentence. S has been expanded in its dependents as a category. Not only are Subject and Verb directly linked to this category. Now S ($\bar{S}$) is also the carrier of Tense, (excluding the other cyclic category NP):

(35) The structure $[_{NP}$ NP V_T ...] is excluded, where V_T is an element containing tense (T) (Chomsky 1973: 275).

Tense was formerly located in the specifier to V position under the predicate VP. Now it has been moved into a tertiary position between NP and VP in $\bar{S}$.[6] The structure originally attributed to Bresnan is given as follows (Chomsky's 54):

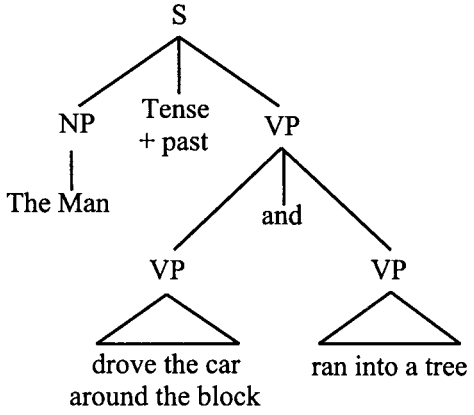
(36) $S \rightarrow COMP\ S'$
$S' \rightarrow NP\ Aux\ VP$
⋮

Bringing Chomsky (1965) back into the picture, the following two relations were proposed: *Subject-of* (NP, S) and *Predicate-of* (VP, S). Thus according to the conditions stated in 1973, NP and VP can agree only when the S for both is the same. But any rule that we try to formulate to relate an NP to a VP under S will violate the *A-over-A* principle. Thus it would appear that agreement can only be applied to S (or $\bar{S}$). Furthermore, if S ($\bar{S}$) is the only category to carry tense, we can derive the fact that *tense* must be a factor which holds for a subject and a verb in a sentence. A mirror image of the grammatical relations now comes into play, such a definition would state that an S ($\bar{S}$) must have both a *Subject* and a *Predicate*. Note, for instance, that the verb is barred from being in a pair with a noun in any other clause in the sentence by the SSC and the Tensed S Condition. However, Wasow and Roeper (1972, note 38), have pointed out that S requires a subject while NP does not.

[6] At this stage, it is important to remember that S dominated $\bar{S}$. Only later was the hierarchy changed so that S would be a category like all the others.

The question remains, what should be done to account for single subject/multiple verb sentences? The 1965 relations only state that S must be the same for the NP and VP in question. The SSC also makes it clear that a subject NP can relate to several verbs as long as they are dominated by the same S. What does this structure actually look like?

(37) *The man drove the car around the block and ran into a tree.*



Such a structure was proposed for sentences composed of multiple S nodes, using the following formula:

(38) $S \rightarrow \left\{\begin{matrix}\text{and} \\ \text{or}\end{matrix}\right\} S^n$, where $n \geq 2$ (Ross 1967: 91).[7]

This rules can also be applied to S, NP, VP or V (Ross 1967: 92). Gazdar proposed a solution which would ensure binary branching, under which *and* and *or* are sister elements to the conjunct constituent using the following formula:

[7] This formula is made complete when used together with "copying" and "Chomsky Adjunction", which when used together produce a structure in which the conjunction is first placed before each S in the sequence, and then the first conjunction to occur is erased. For more information, see Ross (1967).

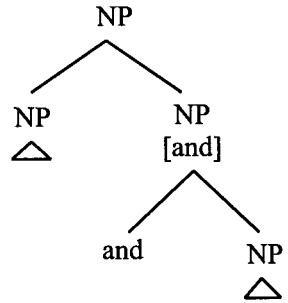
(39)

$$\alpha \rightarrow \alpha_1 \ldots \left\{ \begin{matrix} \text{or} \\ \text{and} \end{matrix} \right\} \alpha_n$$

where α is any syntactic category (Gazdar 1981: 157).

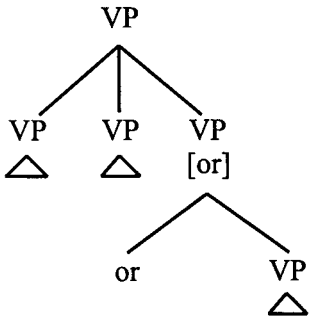
In practice, Gazdar's formula produces conjunct categories with the following structures:

(40)



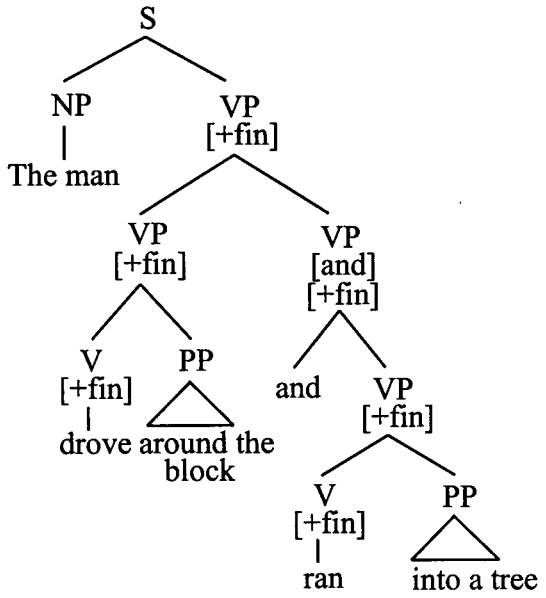
Gazdar also carries this structure over to the VP, producing the following representation:

(41)



Equally interesting is Gazdar's approach to tense. Instead of leaving tense as a third element resting in between the NP and VP, he places the feature [±finite] under the VP and then under the V, in a similar fashion to the conjunctions above. This means that NP and VP are still in relation with S, and that NP must *c*-command VP, and that it must be a direct sister to VP with a [+finite] feature marking for the subject/verb relation to take place. In other words, agreement applies to an NP and a VP which are directly dominated by S and not $\bar{S}$, and the feature [+finite] is passed down to VP's directly dominated by the topmost VP. Using Gazdar's mapping, our sentence from (37) would take on the following structure:

(42)



Central to these arguments is the idea that features from the constituent lexical categories percolate up to the main categories and then farther up to the categories holding the [and] feature. Percolation allows the features such as [+finite] to move from the V all the way up to the uppermost VP. Presumably, it also allows the features of the head N (singular, masculine, human etc.) to travel all the way up to the main NP along with any other features from the specifier node. If tense is part of the *Aux* node, as a third node under S, then it must be able to send information to the constituents that are also dominated by S (NP and VP). This would account for verbs taking the correct form (i.e. *drove* instead of *drive*, and *ran* instead of *run* in (42)). How does the information about the subject get to the predicate? In earlier stages, the transformations dealt with this problem saying that if a verb was adjacent to a noun with the form *Sing*, then the element *S* was added. Now that we have moved away from string adjacency, we must be able to account for agreement using some kind of structural explanation.

Aux is one of the most important elements which has survived from the very beginning. In *Syntactic structures*, *Aux* was introduced as a part of Verb, which was later broken down into the tense and modal information in a sentence:

(43) (i) $Verb \rightarrow Aux + V$

(ii) $V \rightarrow$ *hit, take, walk, read, etc.*

(iii) $Aux \rightarrow X(M)(have+en)(be+ing)(be+en)$

(iv) $M \rightarrow will, can, may, shall, must$ (Chomsky 1957: 39).

In *Aspects, Aux* is initially taken out of VP and placed directly under S:

(44) S → NP^Aux^VP (Chomsky 1965: 68),

only to be put back into the domain of VP by attaching it to VP under S to form an intermediate stage called a "Predicate Phrase":

(45) (i) S → NP ^ Predicate-Phrase

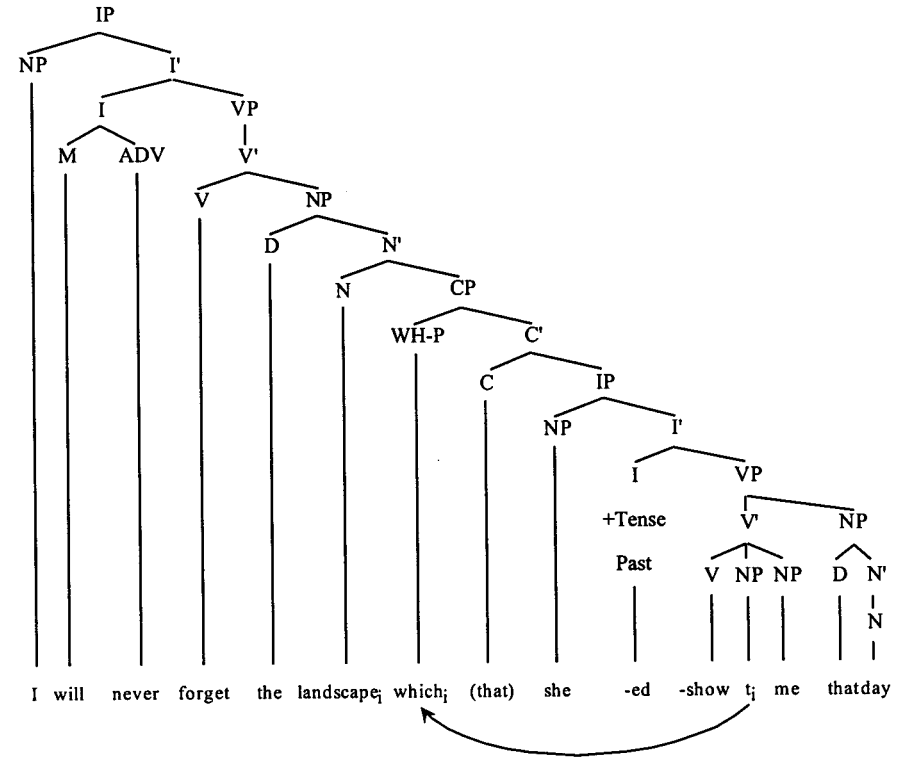
(ii) Predicate-Phrase → Aux ^ VP (Place) (Time)

(xvi) Aux → Tense (M) (Aspect) (Chomsky 1965: 106-107).

As we have seen, *Aux* was moved back under VP in *Remarks on nominalization*. In *Conditions on transformations*, however, it was moved to an intermediary stage under $\overline{S}$, where it was equal with both NP and VP. As we shall see, this position is important and powerful. *Aux* seems to hold much information which is important to the formation of the sentence, but it cannot be formalized. Perhaps Jackendoff's structuralization of *Aux* as a separate syntactic category M is closest to the mark.

The *Aux* position in itself also has import in our discussion of string- versus structure-adjacency. The *Aux* position was important, because it explained the fact that in a given sentence only one Verb receives tense, and that the tense carrying verb is not necessarily even part of the VP. This situation occurs very frequently in English: *I have seen the Tower of London/I will see the Tower of London*. In these sentences, *Aux* inserted Tense, auxiliary verbs and modals in the proper place, as well as the *-en* affix for the past participle and other affixes. If the tense carrier did not directly follow the Subject NP, Affix Hopping was used to carry the information across. Such structures were commonly used well into the late 1970's. That is to say, even though the body of theory was moving away from string adjacency, *Aux* and its key role in tense/modal insertion remained as a convenient explanation to questions that were hard to answer. In terms of string adjacency, it forces tense and other auxiliary constituents to directly follow the Subject NP under S. Even in today's analysis of sentence structure, remnants of string adjacency (*Aux* and Affix Hopping) can be found:

(46)



(Saameño Aivar 1998: 116)

The importance of *Conditions on transformations* was that it moved linguistic analysis away from the idea that syntactic (and grammatical) relations are based on the physical positions occupied by the constituents and the relations produced by these positions. From this point on, the relations between constituents are understood to be more abstract, defined by domains in which constituents must appear or are barred from appearing relative to each other in terms of the structures defined by the $\overline{X}$-theory. More will be said about these domains in following sections.

7. Further reduction of the transformational component

In *On binding*, Chomsky takes many of the concepts and relations presented in 1971, condenses and enhances them, and makes some preliminary movements toward a unified theory. In this paper, Chomsky makes the following observations:

(47) a. SSC and the PIC can be merged together to form the Opacity Condition,

b. The transformational component can be reduced to two rules
- Move-α
- The Case Assignment rule

c. Grammatical relations are syntactic and need not bear relation to the "semantic" interpretations of such ideas.

d. The LF component is the seat of most "structure building" activities.

In terms of our present inquiry into the relationship between Subject and Predicate, several major observations were made in this paper, one of the most important being that the subject and the predicate are incorporated in the *c*-command domain of Tense:

(48) Let us assume, for the moment, that the basic expansion of $\bar{S}$ and S is (17a), as in Emonds (1976), so that Tense *c*-commands both the subject and the predicate of S; and let us assume further that NP is the subject of $\bar{S}$ in (17a) and of NP in (17b):

(17) a. $[_{\bar{S}}\text{COMP}[_{S}\text{NP Tense VP}]]$

b. $[_{NP}\text{NP }\bar{N}]$

(Chomsky 1980: 57).

the definition of *c-command* at this point being:

(49) β is said to *c-command* α if β does not contain α (and therefore $\beta \neq \alpha$) and α is dominated by the first branching category dominating β; then α is in the *domain* of β (Chomsky 1980: 57).

Furthermore, the grammatical relation *Subject-of* is more precisely defined, by structurally identifying the Subject of S, $\bar{S}$ or NP as the least embedded NP in the structure.

(50) Thus, we understand 'subject of α' to refer to the least embedded or most prominent NP in α ($\bar{S}$ or NP), in a configurational language such as English, and we understand 'tense of α' to refer to the occurrence of Tense that is, correspondingly, least embedded in α (Chomsky 1980: 57).

Now we have an answer to the problem noted in the Chomsky (1957) transformations that we inspected earlier (reproduced below for convenience). Firstly, the $\bar{X}$-template allows us to create complex NPs with PP complements.

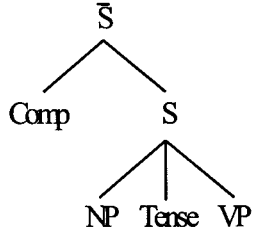
Secondly, percolation allows feature information to travel from the lexical category X to the $\bar{\bar{X}}$ position.

(51) *The man in the car drove around the block.*

At first it seems that we need to be able to pick out *the man* and not *the car* for agreement. However, any attempt that we try to make to specify one NP within a complex NP will violate the *A-over-A* condition. Thus the only relation needed is that NP be directly dominated by S or NP.

Taking this definition into consideration, such a relation in which Tense *c*-commands both Subject-of and Predicate and the former being the most prominent of its type in S ($\bar{S}$ or NP), could only have the following structure:

(52)



Here, the "traditional" *Aux* has been replaced by Tense. Furthermore, Tense assumes the very powerful position of *c*-commanding both NP and VP (subject and predicate) under S.

We still have the problem of the VP and the NP communicating. How, for instance does the Subject receive/take Nominative Case? In *On Binding*, the N in the NP is selected for Nominative case assignment for two reasons: (1) it is the lexical head of the (Subject) NP and (2) the NP is governed by tense. Government is understood in terms of *c*-command as follows:

(53) α is *governed* by β if α is *c*-commanded by β and no major category or major category boundary appears between α and β (Chomsky 1980: 75).

Finally, we seem to begin to be able to use the grammatical relations set down in 1965 to their full extent. Earlier in generative grammar, the phrase structure component produced structures which insured adjacency to allow for agreement. The transformational component shifted elements within the sentence to produce new sentences with the proper adjacency to allow agreement between the individual elements. Now, through the use of conditions such as *c*-command and government, we are able to pick out individual elements from the major categories to identify their roles.

The structure resulting from the rule S → *NP Tense VP* shows that the NP and the VP must be structurally adjacent to Tense (no major categories may in-

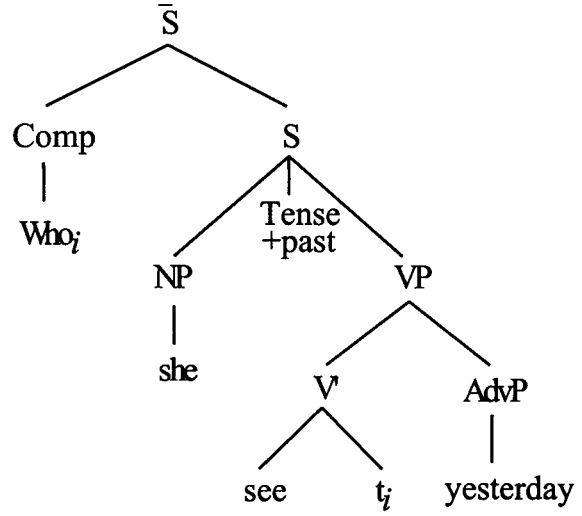
tervene in the tree structure). The nominative N, however, no longer must be string adjacent to the V with which it "agrees":

(54) $[_S[_{NP}[_N$The man$][_{PP}$ in the car$]][_{VP}[_V$drives$][_{PP}$around the block$]]]$

It would appear that we have broken free of string adjacency. Given the structure $S \rightarrow$ *NP Tense VP*, any number of syntactic categories which come in between the head of the Subject NP and the head of the Predicate (VP) are part of either the Subject NP [NP, S] or the Predicate [VP, S] and as such they are incapable of barring the Subject NP and the Predicate from being directly dominated by S.

This is made even more powerful with the introduction of Move-α and trace theory. Move-α allows constituents of the sentence to freely move within the sentence providing that there is an appropriate landing site. Having moved, the constituent must leave a trace, which is fully covered by the government and binding conditions. That is to say, *wh*-elements can move from the end of the sentence to the Comp position while still retaining their position as Object to the predicate:

(55)



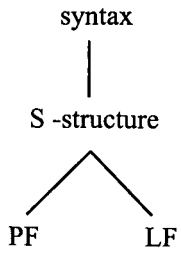
Who$_i$ did she see t_i yesterday?[8]

[8] For a more detailed discussion of this topic, see Pollock (1989) and Iatridou (1990).

8. LF, θ-theory and the dismissal of S

One of the first things to be described in the Government and Binding (GB) framework is the organization of the various components of Universal Grammar. The now familiar chart was introduced showing the relationship between syntax, the S-structures and the Phonetic and Logical forms:

(56)



The syntax consists of a base, which is made up of a categorial component and the lexicon. D-structures (deep structures) are created by the base and turned into S-structures through the rule Move-α (Chomsky 1981: 18). In this model two components will be of considerable interest: the D-structures and the Logical Form (LF). Let us start with the latter, keeping in mind the role of logical representations as a reflection of deep structure representations in the *Aspects* model. In 1981, the LF is said to derive structures using "something like the standard quantifier-variable notation." That is to say, we can use (57) as a derivation for the sentence *Bill saw someone*:

(57) There is a person *x*, Bill saw *x*.

This is fine if we are only worried about the logical relation between the direct object and the verb. What happens when we wish to derive the LF for the Subject of a sentence? Chomsky specifically states that *all* nouns (except idiom chunks and non-argument *it*) in a sentence must receive a θ-role, and that Subjects, being neither of the omitted categories, must also receive θ-roles. This is stated most clearly in the θ-criterion:

(58) Each argument bears one and only one θ-role, and each θ-role is assigned to one and only one argument (Chomsky 1981: 36).

Subject, however, is problematic for Chomsky, who at this point sees θ-roles mainly as resulting from the subcategorization features of a lexical head or the complement position in an $\overline{X}$-structure. S is not a lexical category, and thus not a lexical head, and has no subcategorization features. Subject is also not in the

complement position to VP. Still, Theta theory is meant to cover all NP's and as such the Subject must also be covered:

(59) Furthermore, a θ-role may (though it need not) be assigned in the position of subject, whether of NP or S, a position not associated with a subcategorization feature of a lexical head (Chomsky 1981: 36).

Williams (1977) suggested that a one-place lambda operator be used to formulate the LF derivations:

(60) John saw everybody $\Rightarrow$ John λx ($\forall y$ (x saw y))

(taken from Williams 1994: 37-38).[9]

The problem that Williams wanted to show, at that point, was that the Subject is the *only* argument *external* to the verb, and thus not in the *domain* of the verb (remember, its relation is strictly to S and not VP). The lambda calculus provides a structure in which the Subject of a sentence may bind the predicate, and the predicate its local arguments, but the Subject is left external to the verb's domain.

Chomsky, while seeing that the Subject is not in the same relation to the verb as the verb's complements, presents the relation at S between the Subject and predicate traditionally:

(61) S $\rightarrow$ NP INFL VP (Chomsky 1981: 25 (his 2.1.(25))).[10]

This means that NP must be present at S, and that it must receive a theta role from VP, not from S. All three elements are necessary for the construction of a sentence. The order of the elements, however, is no longer an issue.

(62) ... let us now adopt the position that 2.1.(25) [= (61) above] is the only base rule analyzing S in UG (order aside) (Chomsky 1981: 41).

To alleviate the *domain* problem expressed in (61), Chomsky introduces the so-called GF-θ, which is derived at the D-structure. GF-θ's are θ-roles which are coupled with Grammatical Functions (such as *Subject-of*, *Object-of* etc.) and which are necessarily assigned at D-structure. It is important to note that GF's are still the same in GB as they were in *Aspects* ([NP, S], [VP, S], etc.).

[9] At present, Williams has dispensed with the lambda calculus. I am using his example here for explanatory purposes only.

[10] It should also be noted that, at this stage, the PS component has been reduced completely to this rule alone, letting the $\overline{X}$-template take care of the rest of the structural development of the sentence.

Accordingly, the LF and the D-structures should present a similar representation of the GF's independently of the S-structures. The sentences *The barbarians destroyed Rome* and *Rome was destroyed by the Barbarians* should have the same logical representation of Agent and Patient, with a diverging representation of order at the S-structure:

(63) $\vee y[\wedge x[barbarian\{x\}] \leftrightarrow x = y] \wedge [(\hat{}destroy'\{x\})\{y\}]$ [11]

However, just as in 1957, the DS and LF are not capable of producing sentences without the aid of the transformational component, i.e. Move-α. A sentence is produced when all of the constituents are in the proper order, allowing for the GFs to be established and other constraints to be held.

As the sentence is converted from Deep to S-structure, the constituents are assigned θ-roles. Assignment of QI-roles occurs in the following environments:

(64) a. $[\ldots_{\gamma}\ldots\alpha\ldots\beta\ldots]$

b. $[\ldots_{\gamma}\ldots\beta\ldots\alpha\ldots]$

where α is an immediate constituent of γ, and, thus, α c-commands β (Chomsky 1981: 36).

Traditionally, this is the structure for subcategorization. Whenever α is a lexical item, then β is subcategorized by it. This was mentioned above as it was introduced in *Remarks on nominalization*. Chomsky (1981: 37) states θ-roles and subcategorization are very closely linked. He goes as far as to say that *whenever* α subcategorizes β , it also assigns β a θ-role. The difference is that in (64), for subcategorization, either α or β must be a lexical item, whereas θ-role assignment can also occur between phrasal categories. In other words, γ can equal S and α=VP, which then θ-marks (assigns a θ-role) to the subject NP=β. Two distinctions have been made concerning θ-marking. If the item α subcategorizes β, then α *directly* θ-*marks* β. When phrasal categories are doing the marking (as when VP QI-marks the NP in [NP, S] position), then the lexical head of the marking category α *indirectly* θ-*marks* β.[12]

[11] The use here of $\wedge$ and $\vee$ to represent the universal and existential quantifiers $\forall$ and $\exists$ is taken after Dowty in his use of Montague's original notation. Furthermore, to return briefly to our discussion of conjunctive sentences, the logical representation of a conjunct sentence using lambda operators is also revealing. The sentence *A woman is walking and talking* can be represented with lambda operators as:

(a) $\lambda P \vee x[woman'(x) \wedge P\{x\}(\hat{}\lambda x_2[walk'(x_2) \wedge talk'(x_2)])$ (Dowty 1974: 160-161).

The above example is used to illustrate how two verbs can have a single external argument.

[12] In all instances of θ-marking, if α θ-marks β, then it also θ-marks a category C such that C or a trace of C occupies the position β (Chomsky 1981: 38).

There are some problems with this presentation of subcategorization and θ-roles. First of all, all lexical items are capable of subcategorization. θ-roles, however, can only be assigned by verbs (or VP), prepositions and nouns which are the result of nominalization. A further example of the problems which arise in linking subcategorization and θ-roles can be seen with modal verbs which subcategorize VPs, but do not assign them θ-roles (Piotr Ruszkiewicz personal communication).

In addition to θ-marking, Chomsky (1981) also presents a model of Case assignment. According to this model, "normally, case is assigned to an NP by a category that governs it" (Chomsky 1981: 50). Potential governors are the lexical categories and INFL. That is to say lexical categories assign case to their complements, and INFL when marked [+tense] assigns case to the subject. Once case is assigned to a category, it percolates down to its head (Chomsky 1981: 49). INFL, when marked [+tense] (and sometimes in the infinitive) is broken down into two categories, the new one being AGR, which is made up of the features person, gender, number, and is basically nominal in character (Chomsky 1981: 52). That is to say, INFL = [[±Tense], (Agr)], and AGR has the features of PRO (Chomsky 1981: 209).

An important modification was also made to the definition of Subject. Chomsky (1981) introduces a category SUBJECT, which is:

(65) a. the subject of an infinitive
b. an NP (or AGR if present in INFL)
c. a small clause (Chomsky 1981: 209)

The AGR position is included as being the most prominent nominal element in S. Chomsky himself then hints that this makes INFL a kind of head of S, but does not make any further comments. Further, AGR is coindexed with the NP it governs.[13] That is to say, [AGR, S] and [NP, S] must have the same features ("when NP and a pronominal (pronoun or PRO) are coindexed, they must share the appropriate features" Chomsky 1981: 211). This then explains how the features gender, person and number find their way into the AGR. If indeed AGR is head of S, and S, NP are the only governing categories, then we may also have a way of explaining how the features make their way down to the head of the predicate.

The VP, thus, *selects* (*θ-marks*) the Subject of S (and so the head of the Predicate indirectly θ-marks the Subject NP). The question of how the Subject and

[13] The exact nature of this coindexing is not clear. In certain cases normal coindexing will violate condition C of the binding theory which states that "An R-expression is free" (Chomsky 1981: 188). I will not go into the proposed solutions here.

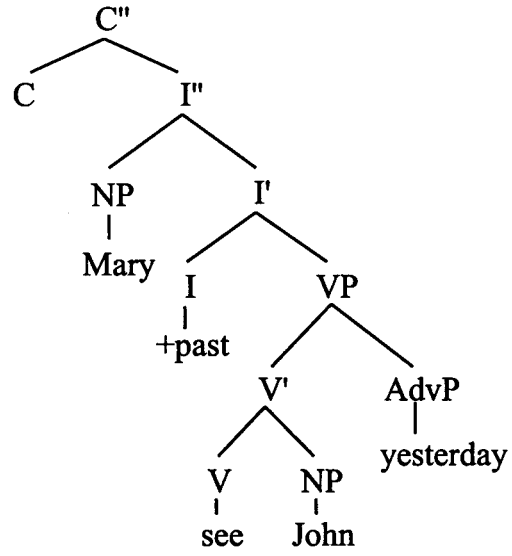
Predicate of a sentence communicate has finally been brought to the surface. How the Subject receives Nominative case has also been addressed. Several questions do still remain. Having transmitted the feature data (i.e. person, gender, and number) from the NP to the AGR element, how are they then transferred to the VP? If INFL governs the subject NP and the VP and AGR, then does that make it head of S? Perhaps INFL is a separate element in which the features of the sentence such as gender, number, person, aspect, tense, time, etc. are set by the speaker at the moment of speech, assigning the features to the elements of the sentence that it governs?

In Chomsky (1986), some important changes to the $\overline{X}$-model were proposed, which have some bearing on the above questions. Namely, S and S' have been replaced by I" and C" (I = INFL and C=complementizer) respectively:

(66) a. $S = I'' = [NP.[_{I'}[_{VP}V\ldots]]]$
b. $S' = C'' = [\ldots_{C'}C\ I'']]$ (Chomsky 1986: 3).

Such a structure provides an interesting insight into the general structure of a sentence. Consider the following:

(67)



Mary saw John yesterday?

The structure of the sentence is clearly laid out. The system still runs into problems, however, when we try to analyze English questions. The problem of *do*-support and other auxiliary verb structures do not fit conveniently into this

framework. Much recent work has gone into trying to explain how English structures can be explained using this new framework.

The structure in (67) is also interesting as it confirms one of the above questions. INFL has become the head of S (S=I"). If coindexing between NP and INFL holds true, then I" will acquire all of the relevant agreement features from NP. It also will have all of the tense and aspect information. Perhaps it assigns these features to the predicate in a method like case assignment – selecting the appropriate complement (or complements in a conjunction), and marking it for tense and aspect as well as agreement features. We could call such a relation Feature Marking. For the moment, this question still remains unanswered.

9. Conclusion

In this paper, I have attempted to survey the body of the theory of generative grammar as presented by Chomsky in order to establish what kinds of mechanisms are understood to produce the phenomenon we normally call "Subject-Predicate Agreement". Although other kinds of agreement do exist, and presumably operate under similar mechanisms, this work has concentrated on the agreement between the Subject and the Predicate in a sentence. Several conclusions can be brought to light. To start with, generative linguistics was for many years stuck with the idea that surface structures were responsible for agreement through string adjacency. That is, a noun agrees with a verb when they are physically adjacent in the string of the sentence. At first this was mechanically realized through the use of transformations, which moved the words in a sentence around, matching them with the needed order. Even with the introduction of $\overline{X}$-theory, and a more relational approach to syntactic structures, presence of this underlying assumption was exhibited well into the 1970's as the *Aux* position. The Government and Binding model has given us new insight into the relations between the syntactic structures and the Grammatical Functions that give us what we understand to be language. It is still hesitant, however, about some of these most basic of questions: How do a Noun and a Verb agree? Where do tense, aspect and time come from, how are they interjected into the structure?, and at what point in the development? Is adjectival agreement the same as subject-predicate agreement?

The purpose of this paper has been to survey the development of ideas surrounding subject-predicate agreement. In conclusion we can state the following.

Apart from lexical and syntactic categories, there are grammatical relations held between individual elements in a sentence. These are called Grammatical Functions. They can be defined in terms of the following sets of relations between an element and a whole structure: Subject [NP, S], Predicate, [VP, S], Object [VP, NP]. NP can also have a subject: [NP, NP], linking it as a category with S. Some means of "communication" between the players has been established.

The head of a VP selects the NP under S (=I') and assigns a θ-role. Nominative case is assigned to the Subject [NP, S] by the AGR element of INFL which governs the [NP, S] position. At the same time, some kind of coindexing takes place, which gives INFL the same features as the Subject NP. Percolation has been introduced which allows information about the individual elements to travel up and down from a category's maximal projection to the head, and vice versa. Presumably some similar device works between a head and its constituents. Information about tense and aspect are in INFL. Their presence there and how they are transferred to the Predicate is not explained. A more detailed study may reveal more about the relations and transfer of information between Subject and Predicate in the domain Sentence.

REFERENCES

Akmajian, Adrian – Peter Culicover – Thomas Wasow (eds.)
1977 *Formal syntax*. New York: Academic Press.

Anderson, Stephen R. – Paul Kiparsky (eds.)
1973 *A Festschrift for Morris Halle*. New York: Holt, Rinehart and Winston.

Bresnan, Joan Wanda
1972 Theory of complementation in English syntax. [Unpublished Ph.D. dissertation, Cambridge, Mass.: MIT.]

Chomsky, Noam
1957 *Syntactic structures*. The Hague: Mouton.
1964a "On the notion 'rule of grammar'", in: Jerry A. Fodor – Jerrold J. Katz (eds.), 119-136.
1964b "A transformational approach to syntax", in: Jerry A. Fodor – Jerrold J. Katz (eds.), 211-245.
1965 *Aspects of the theory of syntax*. Cambridge Mass.: The MIT Press.
1970a "Remarks on nominalization". in: Roderick A. Jacobs – Peter S. Rosenbaum (eds.), 184-221. [Reprinted in Noam Chomsky, 1972, 11-61.]
1970b "Some empirical issues in the theory of transformational grammar". [Reprinted in Noam Chomsky, 1972, 120-202.]
1972 *Studies in semantics and generative grammar*. The Hague: Mouton.
1973 "Conditions on transformations", in: Stephen R. Anderson – Paul Kiparsky (eds.), 232-286.
1980 "On binding", in: Frank Heny (ed.), 47-103.
1981 *Lectures on government and binding*. The Hague: Mouton.
1986 *Barriers*. Cambridge, Mass.: The MIT Press.

Dowty, David R.
1974 *A guide to Montague's PTQ*. Bloomington, Ind.: Indiana University Linguistics Club.

Emonds, Joseph
1976 *A transformational approach to English syntax*. New York: Academic Press.

Fodor, Jerry A. – Jerrold J. Katz (eds.)
1964 *The structure of language: readings in the philosophy of language*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Heny, Frank (ed.)
1981 *Binding and filtering*. London: Croom Helm.

Horn, George M.
1974 The noun phrase constraint. [Unpublished Ph.D. dissertation, University of Massachusetts at Amherst.]

Gazdar, Gerald
1981 "Unbounded dependencies and coordinate structure", *Linguistic Inquiry* 12: 155-184.
Iatridou, Sabine
1990 "About Agr(P)", *Linguistic Inquiry* 21: 551-577.
Jackendoff, Ray
1976 "Constraints on phrase structure rules", in: Adrian Akmajian – Peter Culicover – Thomas Wasow (eds.), 249-283.
Jacobs, Roderick A. – Peter S. Rosenbaum (eds.)
1970 *Readings in English transformational grammar.* Waltham, Mass.: Ginn and Co.
Postal, Paul M.
1964 "Limitations of phrase structure grammars", in: Jerry A. Fodor – Jerrold J. Katz (eds.), 137-151.
Ross, John Robert
1967 Constraints on variables in syntax. [Unpublished Ph.D. dissertation, Cambridge Mass.: MIT.]
Pollock, Jean-Yves
1989 "Verb movement, universal grammar, and the structure of IP", *Linguistic Inquiry* 20: 365-424.
Saameño Aivar, Emilio
1998 "Levels of complexity in relative clauses", *Studia Anglica Posnanensia* 32: 115-126.
Wasow, T. – T. Roeper
1972 "On the subject of gerunds", *Foundations of Language* 8: 44-61.
Williams, Edwin
1977 "Discourse and logical form", *Linguistic Inquiry* 8: 101-139.
1994 *Thematic structure in syntax.* Cambridge Mass.: The MIT Press.