Lecture 11. Neo-Structuralism: Chomsky's TG & Linguistic Universals

Weakness of American Structuralism

We remember that American structuralists made great achievements in observing and analysing the *basic units* that the system of language is made up of, focusing primarily on *phonemes* and *morphemes*. They shifted the focus of linguistic enquiry from written to the spoken language and developed valuable methodology (*discovery procedures*) for observing, analysing, and recording language systems by identifying their basic forms/units, such as *phonemes* & *morphemes*.

The American Structuralists ignored the role of meaning (semantics) in the structure of language. They believed that the basic units of language (phonemes & morphemes) could be identified strictly empirically, based on their *form* and without any reference to their *meanings*. Leonard Bloomfield insisted in his 5,000 page book 'Language' that the definitions used in grammar should be based on the **forms** of the language, ***not*** on ***the meanings*** of the forms, for meaning was not *observable* using rigid methods of analysis and was therefore 'the weak point in language study.' This, in fact, turned out to be the major 'weak point' in the structuralist theory itself, for what do we base definitions of forms on, if not on *the meanings* of those forms*?

N.B.:** According to de Saussure, the **Value** of the Linguistic Sign is the ***'collective meaning' assigned to Signs (i.e., to the ***connections*** between *Signifiers* and *Signifieds*), based on the ***difference*** of each Sign to all the others in the system.

Because **communication of meaning** (ideas) **is the very essence and purpose of language**, meaning cannot be left out of any systematic analysis/description of language.

Chomsky: Linguist / Philosopher / Political Scientist

This critical observation was made by arguably the most influential linguist of the 20th century, Noam Chomsky.

Noam Avram Chomsky, the American linguist, philosopher, and political activist, was born in Philadelphia, Pennsylvania (U.S.A.) in December 1928. His father was an expert in medieval Hebrew grammar. As an undergraduate at the University of Pennsylvania, Noam Chomsky was attracted to the study of formal linguistics partly through sharing the radical political views of his linguistics teacher, Zellig Harris. In 1955 Chomsky started teaching linguistics at MIT (Massachusetts Institute of Technology). In 1957 Noam Chomsky, then only twenty-nine, published his book **'*Syntactic Structures*'** and caused linguistics to take a new turn. This little book (less than 120 pages long) started a revolution in linguistics. Chomsky transformed linguistics from a relatively obscure discipline of interest mainly to PhD students and future missionaries into a major social science of direct relevance to psychologists, sociologists, anthropologists, philosophers, and others.

Chomsky has shifted attention away from detailed descriptions of actual utterances and started asking questions about the nature of the system that *produces* utterances. According to him, Structuralist approach was both far too ambitious and far too limited in scope:

- too ambitious, because it is impossible to describe all the rules of language from a mass of data – remember, language is ever-evolving, and it is impossible in principle to capture the essence of something that changes as you speak?
- too limited, because it describes only those utterances which happened to have been spoken (collected data).

A grammar, he claimed should be more than a description of old utterances – it should take into account possible future utterances. The traditional view that linguistics should provide a description of language based on a corpus of collected data/actual utterances does not account for the *productivity/creativity* of language.

Chomsky pointed out that all speakers of every language have internalised a set of rules that specify the sequences permitted in their language. In his view, the linguists' task is to discover those rules, which make up the grammar of the language in question. This kind of grammar which specifies the rules of all possible combinations (sequences, patterns) of a language is a **generative grammar.** Chomsky, therefore, was the founder of generative linguistics. In his words, a grammar is 'a device which generates all the grammatical sequences of a language and none of the ungrammatical ones.' This kind of grammar is totally explicit – nothing is left to the imagination. The rules must be precisely formulated in such a way that anyone, even an outsider, would be able to tell the well-formed utterances from the ungrammatical ones. The particular type of generative grammar proposed by Chomsky was the so-called **transformational grammar**.

Chomsky also brought the problem of **linguistic universals** back into focus. He pointed out that as all humans are rather similar, their internalised language mechanisms are likely to have important common properties. He saw the discovery of these universal elements and constructions (that are available to all human languages, irrespective of whether they actually occur or not) to be a major task for linguistics. Above all, he thought, linguists should specify the universal bounds, or **constraints** within which human language operates.

The constraints on human language are, in his view, inherited. Human beings may be pre-programmed with a basic knowledge of what languages are like, and how they work. Chomsky called this inherited core knowledge **Universal Grammar (UG)**, and tried to explore its make-up.

We briefly discussed the ***Nomos-Phusis*** Debate, which stems from the Ancient Greece and beyond, and its implications for existing theories of language acquisition. Here is just a little reminder of the major trends in the existing theories of language acquisition:

- Behavioristic: people are born as *tabula rasa* (clean slates) and are shaped totally by their environment.
- Generative/Nativist: people are born with innate knowledge.
- Functional: innate ability is shaped by environment.

Chomsky's recent work, the so-called **Minimalist Program**, has become more and more abstract, specifying only the broad general principles, the mere 'skeleton of human language. He compares his interests to those of a scientist who is not content to just watch apples falling to the ground – he wants to understand the principle of gravity. In this, he is following the current trend of looking for the 'Theory of Everything' that would sum up the entire Universe in a single equation.

Chomsky has been the major linguistic influence in the second half of the 20th century. He still has many devoted followers, but he also has critics who argue that he overemphasizes the constraints within which human language operates. It has been impossible to identify any firm boundaries, however: so many times the proposed constraints were broken by some newly discovered languages that do not obey the set limitations. And so, the linguistic 'Theory of Everything' is still in the making – will it ever be found? That is the question.

Re-cap:
Be that as it may, Chomsky is one of the most influential figures in 20th century linguistics. He

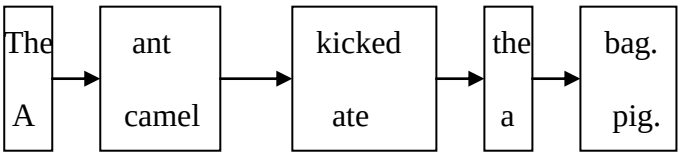
- Initiated the era of **generative linguistics**, which focuses on the rules that underlie our knowledge of language (the 'rules of the game')
- Reawakened people's interest in **language universals**: this topic had become unfashionable in the early part of the century, when it was commonly assumed that 'Languages differ without limit and in unpredictable ways.' Chomsky argued that linguists should concentrate not so much on finding out what all languages have in common, but on discovering the **constraints** within which language operates.

Transformational Grammar and Linguistic Universals

Chomsky did not just make vague statements about the need for generative grammars and universal constraints – he put forward some detailed proposals for a universal framework. He has, however, changed his mind on many facets of his original theory put forward in 1957, which started off as a **transformational grammar**. Let us try to understand how he came up with this particular type of grammar, and take a look at its main characteristics. Then we shall try and see why he changed his mind on some of the issues involved, and what his new proposals are.

Imagine you are Chomsky in the mid-50s, trying to set up a universal grammatical framework. Where is the right place to start? A way to get started is to write a grammar of a language you know – say, English. If we can do that, then we can see to what extent other languages might fit into the same framework. In order to write down the grammar of English, we would need to make a guess, or hypothesis about the rules internalised by speakers of the language, and then test the validity of this hypothesis by checking it against some raw data (English sentences). If the rules we hypothesised do not produce grammatical English sentences, then we'd have to make the necessary adjustments.

Let us start with an over-simplified hypothesis: are words linked together in a linear fashion, like a simple chain?



However, if we look at some other sentences, we shall discover that this hypothesis does not work in all cases – a word is not necessarily dependent on the words next to it. Often there are 'long-distance' relationships, as in

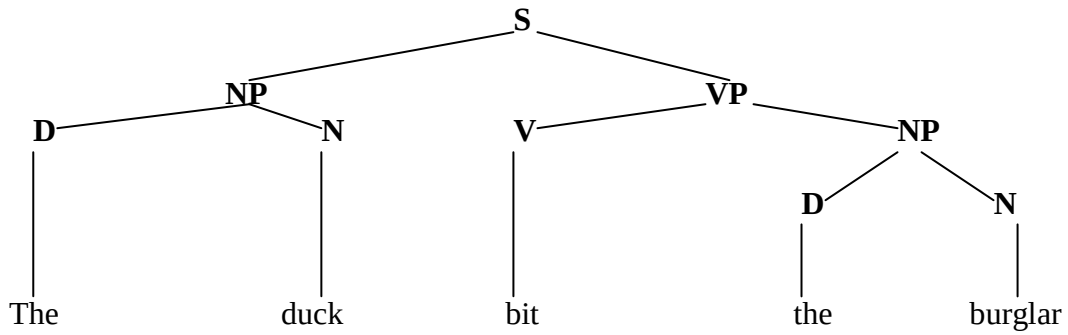
Either *behave yourself,* ***or*** *leave the room.*
Peter *fell and hurt* ***himself****.*

This calls for a 'layered' sentence structure that would show the intrinsic connection between *either* and *or* and *Peter* and *himself.* This means assuming that languages have several basic sentence patterns, each with a number of different 'slots' which can be expanded in various ways. In English, noun phrase (NP) followed by a verb phrase (VP) is a basic English sentence type – also 'expandable':

NP	VP
Ducks	bite.
Ducks	bite burglars.
The duck	bit the burglar.

Such a grammar (called **phrase structure** grammar) is made up of phrase structure rules, normally in the form of **rewrite rules** which show the progressive expansion:

S → NP → VP
VP → V → (NP)
NP → D → N



This sort of 'expansion mechanism' (embedding) seems to be built into any grammar. However, according to Chomsky, this model of representation is inadequate, because it has two serious flaws:

1. **Clumsiness**: we need an enormous number of rules to generate all the sentences of English.
2. **Ambiguity**: it groups together sentences which are dissimilar, and separates others, which are similar.

Take, for example, the sentences

Romeo is anxious to help.
Romeo is difficult to help.

Any speaker of English will see the *semantic* difference between the two: in the first one, Romeo wants to do the helping, whereas in the second one it is somebody else who wants to help Romeo. Yet, the 'slot' pattern of both sentences is identical:

NP	V	ADJ	INF
Romeo	is	anxious	to help
Romeo	is	difficult	to help

We have a similar problem with the sentence *Romeo is ready to eat* ☺: the slot model cannot show the radically different interpretations.

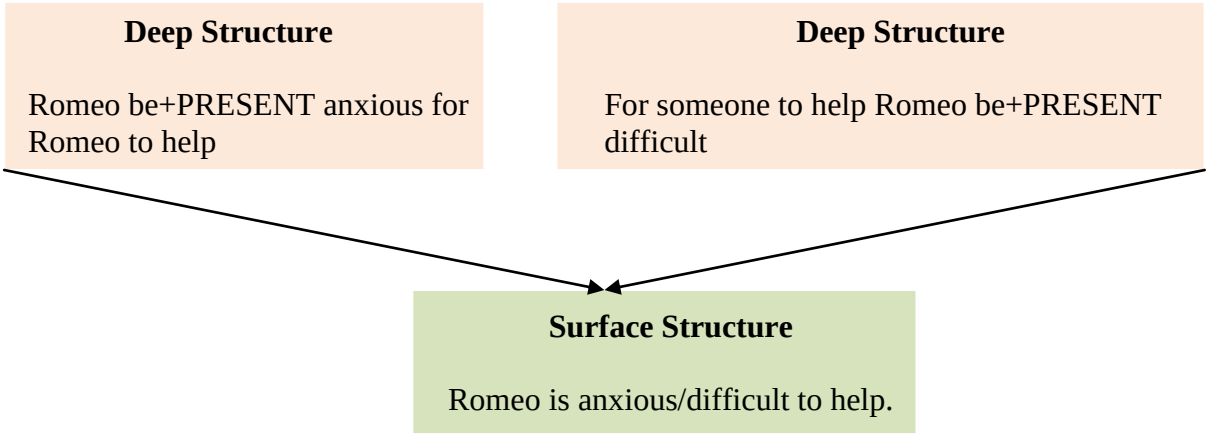
The reverse problem occurs in sentences such as:

To wait longer is useless. / It is useless to wait longer.
Yesterday it rained. / It rained yesterday.

Regarded as quite similar by English speakers, these sentences have different structures (slot patterns). Chomsky argued that a grammar which provides only one structure for ambiguous statements, and different structures for similar sentences, was a bad grammar. He suggested a solution to these problems: the **transformational** model.

Deep & Surface Structures

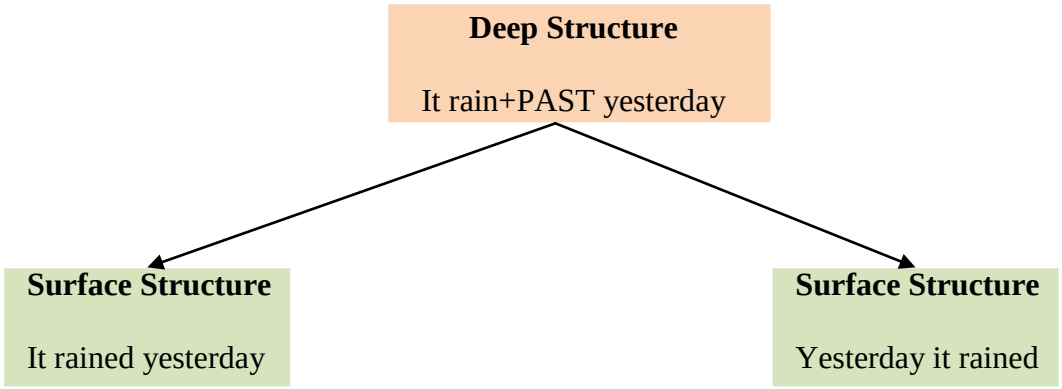
Chomsky suggested that every sentence has two levels of structure: one on the **surface**, and another, which is **deep** and abstract. This accounts for ambiguity by suggesting that some sentences have similar **surface structures**, but different **deep structures:**



This is a simplified version of Chomsky's **Standard Theory** of transformational grammar, outlined in his book *Aspects of the Theory of Syntax* (1965).

A similar explanation accounts for the ambiguity in *Romeo is ready to eat*, where two different deep structures are realised by a single surface structure.

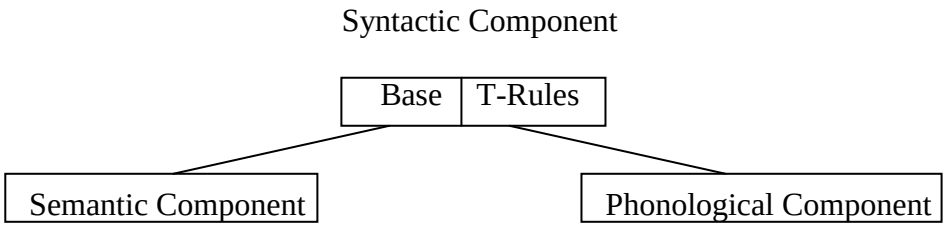
However, the reverse situation occurs in sentences of the other type, *Yesterday it rained* and *It rained yesterday*; here two different surface structures have the same deep structure:



If every sentence has two levels of structure, then it is necessary to link the two levels in some way, to be able to 'process' the sentence. Chomsky said that deep structures are related to surface structures by way of **transformations**. A deep structure is transformed into its related surface structure by the application of one or more transformations. For example, the sentence *It rained yesterday* would require only one transformation: the attachment of the tense to the end of the verb. But the sentence *Yesterday it rained* requires a second one also, the one which moves the adverb *yesterday* from the end of the sentence to the beginning.

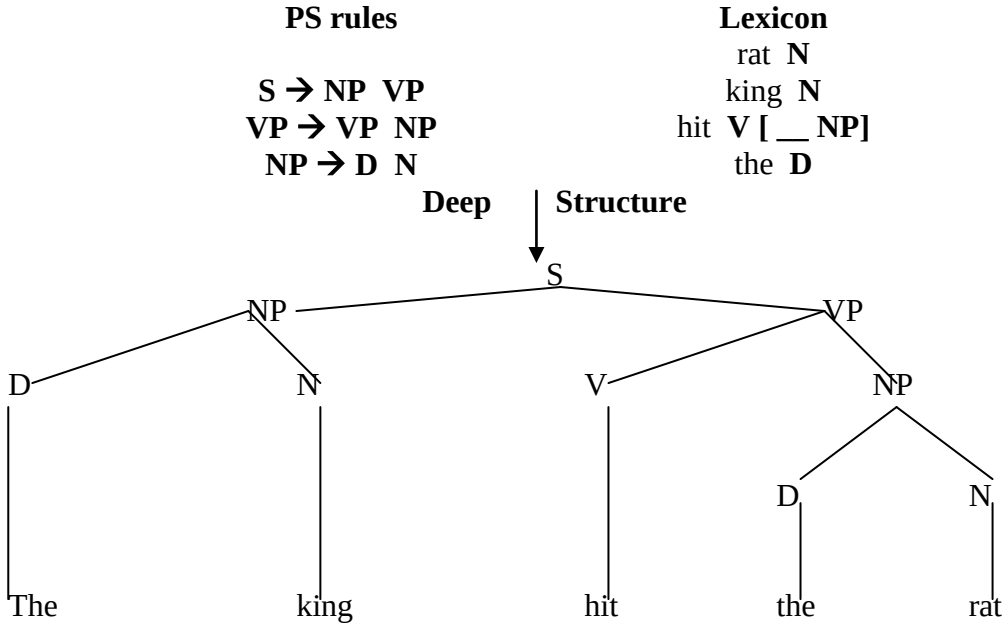
Thus, **Transformational Grammar** is a grammar which sets up two levels of structure, and relates these levels by means of operations known as **transformations**.

Like most other grammars, it has three major components: syntax, phonology, and semantics. It differs from other grammars in that it splits the syntactic component into two parts: the **base**, and the **transformational rules**:



In the Standard Theory, the base contained phrase structure (PS) rules for the formation of deep structures, and also a lexicon, from which words were slotted into the output of the PS rules:

Base (simplified)



The deep structure then passed to the transformational rules, to be converted into the surface structures. At this point, the surface structure of a sentence was still abstract: it did not yet have a phonetic form. Phonological component then converted each surface structure into a phonetic shape/representation. Meanwhile, transformations could not change meaning, so the deep structures were fed directly into the semantic component, which gave a semantic interpretation of each (See diagram on p. 7).

Deep Structure

Chomsky did not base his claim of there being two levels of structures simply on the native speakers' intuitions. His most important arguments were based on **movement** of sentence constituents, for example:

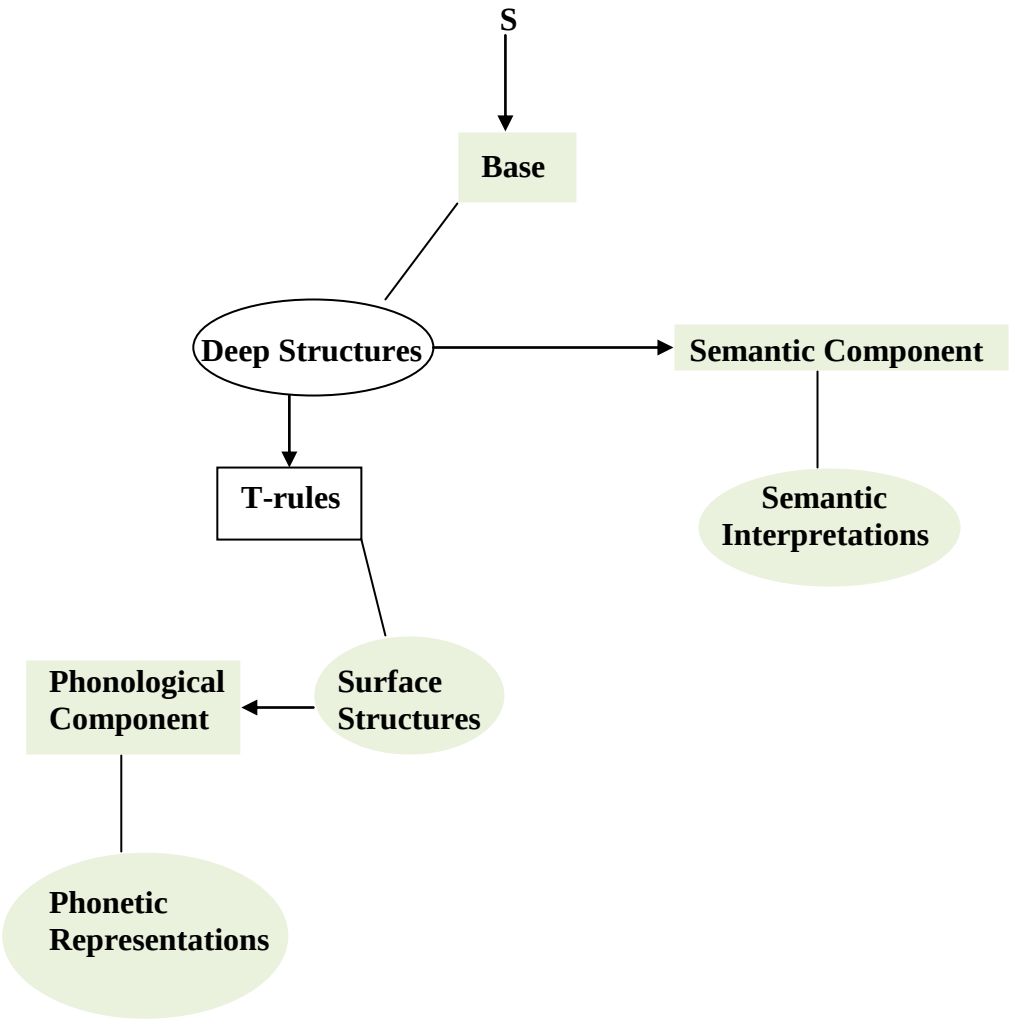
Janet put the lizard in a drawer. (This verb needs both the NP and the PP)

**Janet put in the drawer*
**Janet put the lizard*

Now look at these sentences:

***What** did J. put [___] in the drawer?*
***What** did J. put the lizard in [___]?*

We can conclude that this is a general syntactic rule: **'In order to form one common kind of question, substitute *what* in place of a NP, and move it in front of the sentence.'**



We can then hypothesize that the deep structure of the sentences is something like:

Question: *Janet put **what** in the drawer,* or *Janet put the lizard in **what***
A transformation would then move ***what*** in front: *What did Janet put in the drawer? / What did J. put the lizard in?*

These arguments convinced many people that there were indeed two levels of structures – deep and surface, linked by transformations.

Transformations, unlike the rewrite rules, had two parts to each:

1. The **applicability check (Structural Analysis – SA)**
2. **Structural Change (SC)** (change brought about)

Consider, for example, the transformation which moved adverbs to the front of a sentence, **T-adverb preposing**, as in:

Peter shrieked suddenly. → *Suddenly Peter shrieked.*

SA (applicability check) was necessary to make sure that the sentence contained an adverb: 'Is there an adverb?' → 'Move the adverb to the front.'

Formal version:

SA X – ADV
SC X – ADV → ADV – X

Here, X is a 'variable.'

In the **Standard Theory** there were about two dozens of these transformations, each applying to a specific structure. Some of them moved things around, and some deleted items (for example, Come! **T-imperative**). Yet another type of transformations added items (a **T-there-insertion**).

In the late 1960s, linguists hoped to succeed in compiling a full list of all the transformations of English, and make a complete specification of how they worked – this, however proved unrealisable.

1. Define a **transformational grammar.**
2. What arguments can be put forward to support the claim that languages have deep structures as well as surface structures?
3. How many parts do T-rules consist of, and what is the purpose of each of these?

Trouble with Transformations

- T-rules Change *Meanings*? Two Views: EST & GS
- Trace Theory: REST
- T-rules Reduced to Common Denominator: Movement
- Back to Basics: Universal Grammar
- Government and Binding
- The Bare Bones: the Minimalist Program

T-rules Change Meanings? Two Views: EST & GS

Transformational grammar seemed to have transformed linguistics: all linguists had to do was to identify the T-rules that transformed deep structures into surface structures *without changing the meaning of the deep structure.* That was the strongest constraint placed on transformations within the Standard Model of TG - T-rules were not allowed to change the meaning of the Deep Structure, otherwise we would end up with transformations that would change

'*Peter kept a crocodile in his cupboard*'

into, say,

'*Our pig adores avocados*'! ☺

Problems with sentences containing quantifiers:

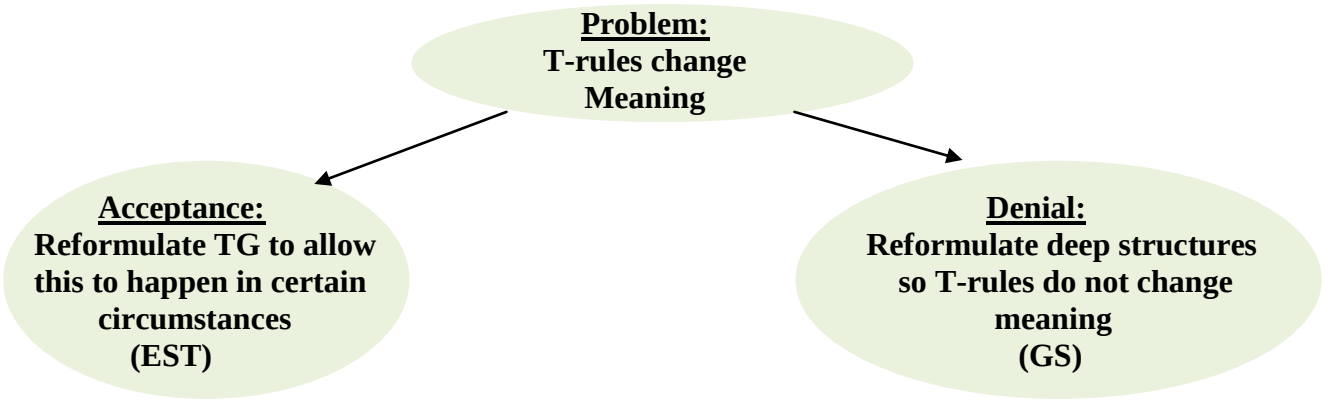
T-passive:
Many women do not do housework. → *Housework is not done by many women.*
Many babies do not drink milk. → *Milk is not drunk by many babies.*

1. Many X do not Y, although many still do.
2. Many X do not Y, and hardly any still do.

T-conjunction reduction:
Few men are rich and few men are famous. → *Few men are rich and famous.*
Few chickens are big and few chickens are intelligent. → *Few chickens are big and intelligent.*

[T-conjunction reduction has produced a sentence with a different meaning! This is not allowed in **TG**]

Two Views: Extended Standard Theory (EST) Vs. Generative Semantics (GS)



Generative Semantics:
Generative semanticists denied that surface structures could affect the meaning of the *underlying structures* (DS) and continued to modify their underlying structures - until they became indistinguishable from semantic structures. This made Deep Structures & Semantic Structures (separate components of TG) fuse together: the base would generate a set of underlying structures which *was* the semantic structure (that is why they became known as **generative semanticists**).
Problem w/ GS: overemphasis on semantics; denial of the role of syntax in expressing meanings.

Trace Theory - Revised Extended Standard Theory (REST):
Eventually people recognised that at least some Surface Structures *did affect* the interpretation of the meaning → **REST**.

REST focused on *where* items were moved from the deep structure in the process of its transformation into a surface structure. It was assumed that items leave a faint **trace** (**t**) of its previous location:

Q **DS:** *Ann find+Past* ***what*** *in her bag.* → **SS:** ***What*** *(did) Ann find* ***t*** *in her bag?*

T-rules Reduced to Common Denominator: Movement

Remember, the main constraint on T-rules in the Standard Theory was that they could not change meaning: all the meaning was in the DS, and it had to be retained in the SS.

REST version of TG maintained that Surface Structures alone provided the semantic interpretation.

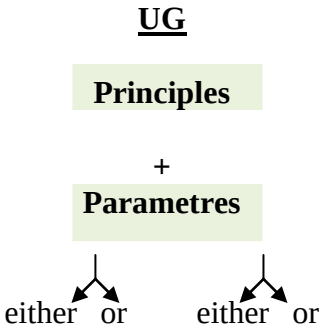
⇨ Therefore, back to square one: too many T-rules/Surface Structures to describe (same as Phrase Structure Grammar rules). It therefore became important to provide *firm guidelines as to what could move where*. REST tried to constrain transformations, prevent random movement of items in all directions.

Suggestion 1: T-rules must follow modified phrase structure rules = be '**structure-preserving**.'
Suggestion 2: Limit the *distance* which items can travel: *Absence of proof is not proof of absence.*
→ ***Absence is not proof of absence of proof.***

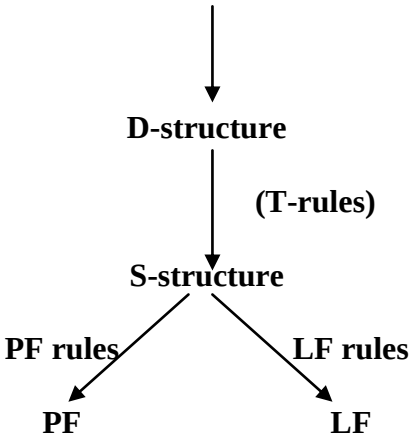
T-rules were revised and their number was drastically reduced down to **one major T- rule**: **'anything can be moved,'** but with strong constraints specifying what could be moved where. This led to a new version of Transformational Grammar – **Government & Binding (GB).**

Back to Basics: Universal Grammar

Chomsky has become increasingly concerned with the **learnability problem**. How do children manage to learn language so efficiently? They must, according to Chomsky, be born equipped with **UG**, a basic outline knowledge of language properties:



(Re: implicational universals: an animal w/ feathers and a beak is likely to have two legs.)



New terminology:

Deep Structures of the 1960s → D-structures of the 1980s
Surface Structures "—"--- " → S-Structures of the 1980s
Semantic Representation → LF (Logical Form)
Phonetic Representation → PF (Phonetic Form)

However, not one of these levels corresponds exactly to the relevant level in the old TG, and the rules which operate here are not the same: for example, both LF rules and LF contain a lot of material which belonged strictly to the syntax in a Standard TG.

Government and Binding Theory

Lectures on Government & Binding (1981): the name stuck (**GB**). This is an attempt to specify exactly which part of trees influence one another. The concept of **c-command [constituent command]** specified relationships and restrictions between different parts of the sentence.

Anna had a dream about herself. → **Herself had a dream about Anna.*
The students argued with one another. → **Each other argued with the students.*

Rule: words which refer back to others (**anaphors**) must be c-commanded by the words they refer back to (their **antecedents**).

Binding is strongly interlinked with the notion of c-command. Briefly, a **binding principle** states that when two NPs are co-indexed/refer to the same thing or person (as in *Anna hurt herself*), then the antecedent (Anna) must command the anaphor (herself). If so, then it has been **properly bound**.

**Herself hurt Anna.* ← improperly bound.

This seems like common sense, but the need for specifying the structural relationships between NPs becomes apparent as sentences get more complex, i.e.:

Who did Ann claim hurt herself? → ***SS:*** *Who did Ann claim* **t** *hurt herself.*

To summarize, the GB approach was particularly concerned about relationships between constituents. It identified 'the bosses' with power to command others, with the ultimate purpose of expressing clearly and simply which nodes on a tree were interlinked. In other words, GB was concerned with HOW all words in a sentence related to all the others.

We see here a relative shift in perspective:

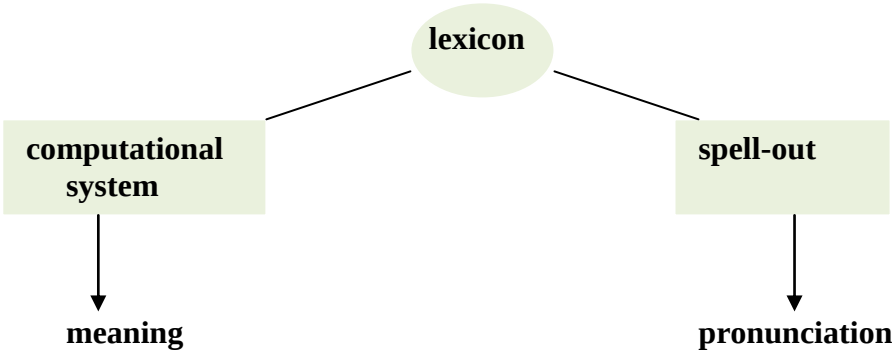
- **TG** was a device which specified what was/was not a well-formed sentence, whereas
- **GB** focused on the general principles and relationships which exist within language.

The Bare Bones: Minimalist Program

Trying to discover the linguistic equivalent of the law of gravity, Chomsky has tried to cut down his linguistic theory to the bare bones of language. In his latest work, ***Minimalist Program*:**

- **The on/off switch setting of the Principles and Parameters (P&P) concept is the main framework feature of UG which was retained**
- **Two levels of structure have been abolished: no more deep/surface structures, or D-/S-structures.**

In this 'minimalist' model, the lexicon feeds a 'computational system' which checks that the word combinations/linguistic patterns agree with basic linguistic principles. The end product is meaning on the one hand, and pronunciation (or 'form' - OT), on the other:



The linguistic Principles which guide the system are still only sketchy, but they are essentially principles of 'economy' or 'simplicity.' The most straightforward is the **Shortest Move**.
Consider the sentence, *Anna asked Peter to find her keys.* Suppose you want to ask *who* Anna had asked and *what* she wanted to be found: *Anna asked who to find what?*

Normally, any word beginning with *wh-* is brought to the front of the sentence. But in this case, only the word which moves the shortest distance can come forward: it is possible to say,

***Who** has Anna asked to find what?*, but it is impossible to say,
****What** has Anna asked who to find?*

This is the type of broad-ranging linguistic principle which Chomsky is hoping to identify, though 'current formulation of such ideas still leaves substantial gaps,' as he himself admits.
Chomsky's model of language is not the only model around at present. However, it still has more adherents than any other model – which is why we have looked at it in preference to others.

:

1. What is meant by **Principles and Parameters (P&P)**?
3. What do the terms **D-structure, S-structure, LF** and **PF** mean?
2. What do the terms **government** and **c-command** mean?
3. What is the **minimalist program** and its **two major features**?

Appendix I: Some linguistic terms for your REFERENCE ☺

What is reference?

Definition

Here are two senses for *reference:*

1. **Reference is the symbolic relationship that a linguistic expression has with the concrete object or abstraction it represents.**
2. **Reference is the relationship of one linguistic expression to another, in which one provides the information necessary to interpret the other.**

Example

A pronoun *refers* to the noun antecedent that is used to interpret it.

Kinds

Here are some kinds of reference:

- What is coreference?
- What is endophora?
- What is exophora?

What is endophora?

Definition

Endophora is coreference of an expression with another expression either before it or after it. One expression provides the information necessary to interpret the other.

Discussion

The endophoric relationship is often spoken of as one expression "referring to" another.

Examples (English)

- *A well-dressed man* was speaking; *he* had a foreign accent.
- If you need *one,* there's *a towel* in the top drawer.

What is anaphora?

Definition

Anaphora is coreference of one expression with its antecedent. The antecedent provides the information necessary for the expression's interpretation.

This is often understood as an expression "referring" back to the antecedent.

Discussion

The term *anaphora* is also sometimes used to include both anaphora, as defined here, and cataphora. When it is used that way, it becomes synonymous with *endophora*.

Example (English)

In the following sequence, the relationship of the pronoun *he* to the noun phrase *a well-dressed man* is an example of ***anaphora***:

A well-dressed man was speaking; *he* had a foreign accent.

Anaphora is a kind of **endophora**

What is cataphora?

Definition

Cataphora is the coreference of one expression with another expression which follows it. The following expression provides the information necessary for interpretation of the preceding one.

This is often understood as an expression "referring" forward to another expression.

Example (English)

- In the following sentence, the relationship of *one* to *a towel* is an example of cataphora:
    - If you need *one*, there's *a towel* in the top drawer.

Generic

Cataphora is a kind of

- What is endophora?

What is exophora?

Definition

Exophora is reference of an expression directly to an extralinguistic referent.

The referent does not require another expression for its interpretation.

Kinds of exophora:

- deixis
- homophora

What is homophora?

Definition

Homophora is reference that depends on cultural knowledge or other general knowledge, rather than on specific features of a particular context.

Examples (English)

- The use of *the President* in the U.S.
- The use of *the sun*
- The use of *the baby* between parents to refer to their own baby
- The use of *bees* to refer generically to the class of bees in *Bees make honey*