Chapter 14

Transformational Generative Grammar
Introduction

1. *Introduction*

Transformational Generative Grammar (TGG) is dated from 1957 and the publication of *Syntactic Structures*. Its author, Avram Noam Chomsky, was 28 years old when the book was published. He had been Zellig Harris's student at the University of Pennsylvania.

Thomas Wasow (1985) segments the subsequent history of TGG into **three phases**, which he in turn identifies with the **three adequacies** of TGG: observational, descriptive, and theoretical. The first phase extends from the initiation of TGG until the mid-1960's and Chomsky's *Aspects of the Theory of Syntax*. That period is devoid of semantics and the emphasis is upon pure form. Katz and Postal's *An Integrated Theory of Linguistic Descriptions* suggested an extension of the theory to incorporate some semantics, and Chomsky 1965 represents the codification of this addition and the inauguration of the **standard theory** and the second phase of classical TGG. The emphasis is now seen as being upon descriptive adequacy. That is, the concern is with grammars which do more than 'generate' grammatical sentences; grammars must accurately capture the speaker's intuitions about these sentences. This change is expressed as a concern with meaning and with appropriate structural descriptions, since the two are tied together. The distinction between competence and performance is established, with the former being the subject of the descriptively adequate grammars. This second period was the shortest of the three and it quickly gave way to the third which is concerned with constraining the theory which allows those descriptively adequate grammars. This produces the **extended standard theory**, also known as **government-binding** (or 'GB') theory. John Robert ('Haj') Ross (1968) represents one of the first steps in this direction. The goal of constraints is to gain explanatory adequacy. Chomsky (1982:3) evaluates the transition in this way:

> ... it [GB theory] develops directly and without radical break from earlier work in

transformational generative grammar, in particular, from research that falls within the framework of the Extended Standard Theory (EST).

In the same way that GB emerges from EST, so the first step of TGG grammar emerges from American Structuralism ... although to establish the theory as a distinct and 'better' theory, it was necessary to distance TGG from its origins and to emphasize its originality (Newmeyer 1980:20):

> The essence of Chomsky's revolution was his gift to the field of a truly scientific perspective [namely, "the theoretical assumptions first articulated in *Syntactic Structures*"].

The distancing of TGG from American Structuralism was in major proportion accomplished by polemics (Newmeyer 1980).

2. *Early Chomsky*

In 1955, Chomsky was arguing for an approach to grammar which was independent of semantics (Chomsky 1955:141):[1]

> The contention of this paper will be that semantic notions are really quite irrelevant to the problem of describing formal structure.

His position seemed to assume that not only was phonology independent of semantics (autonomous), but so was syntax independent.[2] This, of course, is a continuation of the general attitude of American Structural syntax, which found an extreme in Harris's (1951) advocation of a formal and meaningless linguistics. Chomsky (1955) makes his argument in terms of phonology. The argument progresses by the methodology of a negative proof: suggest an hypothesis, demonstrate that it must be false (or inconsistent or impossible), and then take the complement as proved. The hypothesis is (Chomsky 1955: 143):

> ...U[tterance]$_1$ is phonemically distinct from U$_2$ if and only if U$_1$ differs in meaning from U$_2$.

It is not difficult to show that distinctness in meaning is by itself no guarantee

---

[1] The author is perhaps 26 years of age at this point.

[2] The title of *Syntactic Structures*, Chapter 1, is "The Independence of Grammar".

of phonemic distinctness. Consider (Chomsky 1955:143):

(1)   I saw him by the bank.

is ambiguous because the lexical item *bank* represents a homonym 'bank as financial institution' and 'bank as geographical location'. Yet the phonology of (1) is not differentiated because of this. There are not two phonemically distinct utterances in (1); therefore, difference in meaning does not correlate with difference in phonology. Similarly, identity of meaning does not correlate with an identity of phonology (Chomsky 1955:143):

(2)   (a)   He is a bachelor.
       (b)   He is an unmarried man.

These are "two expressions with the same meaning", and that identity does not predict identity in terms of phonology. Thus, semantic difference does not predict phonological difference nor does semantic identity predict phonological identity; therefore (Chomsky 1955:144):

> ... if we take meaning seriously enough to assign meanings correctly to utterances, we seem to learn very little about phonemic distinctness. It is important to add that when we run into a real problem of establishing phonemic distinctness, we in fact do not rely upon meaning in any way .[3]

If the theory is not dependent upon semantics, it may be based upon **form**. And at this point, Chomsky mirrors the views of Harris, his teacher, in rejecting intuition about these forms as an intrusion of semantics (and intuition) into formal description (Chomsky 1955:149):

> The major goal of methodological work in linguistics is to enable us to avoid intuition about linguistic form, replacing it by some explicit and systematic account ... I have argued that the appeal to meaning is actually an appeal to intuition, and hence is to be avoided in linguistic analysis. But the study of meaning is an essential task of linguistics. It is certainly important to find some way of describing language in use. However, this is not the study of linguistic form.[4]

---

[3] Chomsky (1955:143) suggests Harris's (1951:32-33) pair test.

[4] The attitude towards intuition, whether the speaker's or the linguist's, continues the American Structuralist negative evaluation of such. Recall the rejection of the strand of

The preliminary equation of meaning with use and the stigmatization of meaning then leads to the emphasis of a meaningless form (Chomsky 1955:150):

> If the same sort of argument can be applied to other proposals for recourse to meaning, as I think it can, then we are forced to conclude that at least at the present stage of our knowledge, the theory of linguistic form does not appear to have semantic foundations.

2.1   *Syntactic Structures*

The stage is now set for taking language to be a *set of sentences* (Chomsky 1957:11, 13):

> We can determine the adequacy of a linguistic theory by developing rigorously and precisely the form of grammar corresponding to the set of levels contained within this theory, and then investigating the possibility of constructing simple and revealing grammars of this form for natural languages ... From now on I will consider a *language* to be a set (finite or infinite) of sentences, each finite in length and constructed out of a finite set of elements.

The language may be a set of sentences, but the description of the language is not. The description, or grammar, of a language bears the relation *project* (Chomsky 1957:15) or *generate* to the sentences which constitute language:

> In this respect, a grammar mirrors the behavior of the speaker who, on the basis of a finite and accidental experience with language, can **produce or understand an indefinite number of new sentences** [Emphasis mine, PWD].

This idea is not new with TGG and represents a continuation of one of the

---

theorizing represented by Sapir and his students as well as the emphasis on objectivity. Chomsky's teacher Harris (1946:161) comments on syntax and intuition:

> One of the chief objectives of syntactic analysis is a compact description of the structure of utterances in the given language. The paucity of explicit methods in this work has made syntactic analysis a tedious and often **largely intuitive task**, a collection of observations whose relevance is not certain and whose interrelation is not clear ... In many of the descriptions that have been written, the lack of explicit methods has permitted the use of diverse and undefined terms and a reliance on **semantic** rather than formal differentiations [Emphasis mine, PWD].

Chomsky's "language in use" anticipates "performance" as "linguistic form" anticipates "competence". But at this point, meaning is associated with the equivalent of performance, language in use.

concerns of American Structuralism (Hockett 1954:217):

> A grammatical description built according to the plan outlined ... [in the Item and Arrangement model] sets forth principles by which one can **generate any number of utterances** [Emphasis mine, PWD] in the language; in this sense, it is operationally comparable to the structure of that portion of a human being which enables him to produce utterances in a language, i.e., to speak.

The concerns and goals of TGG and American Structuralism are very much the same, as is the starting point ... with respect to syntax:

(i)   The primary data are form (phonology and syntax),
(ii)   Their patterns are independent of meaning,
(iii)   A description of grammar generates the same sentences/utterances which a speaker of the language can produce.

TGG takes the description of a sentence, which is a *string* of words, to be represented (in part) in its *phrase structure*. And a phrase structure appears in turn very much like the descriptions provided by American Structuralist IC-analysis. Chomsky (1957:27) provides an early depiction of such a description. Cf. Figure 1. A construction is now equivalent to the phrase
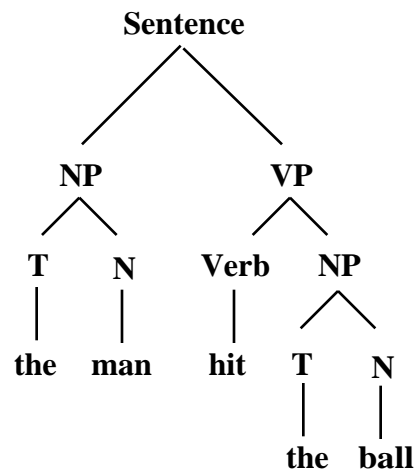

Figure 1: *A typical phrase structure.*

structure which organizes the string of lexical items; and the pattern of syntax

lies in these hierarchies. Since TGG continues the negative bias towards semantics (cf. Chomsky 1955 above and Chomsky 1957:92-105), the hierarchy which was the innovation and contribution of American Structuralism is taken over.[5] Structures such as those in Figure 1 are not, however, the primitives of the syntax of a language. In a paper entitled "From Morpheme to Utterance", Harris (1946) had proposed the use of **equations** to "indicate substitutability". The expressions to the left in (1) - (6) are taken from Harris (1946.166ff.). The expressions to the right are their rough equivalences in TGG phrase structure (cf. [7] - [12] and [13] -[16]):

(1)	$N^4V^4 = N^4V^4$	$S \longrightarrow NP + VP$
(2)	$I = N^4$	$NP \longrightarrow N$
(3)	$V^3\text{-}V_v = V^4$	$VP \longrightarrow Aux + Verb$
(3)	$V = V^1 = V^2 = V^3$	$Aux \longrightarrow C\ (M)\ (have + en)\ (be + ing)$
(4)	$I = she$	$N \longrightarrow she$
(5)	$V_v = \text{-}ed$	$C \longrightarrow Past$
(6)	$V = gain$	$Verb \longrightarrow gain$

This selection of equations is represented in an order reverse from the one in which Harris lists them, i.e., from the top-down in place of from the bottom-up. Harris is working from morpheme to utterance and therefore begins with *she*, *-ed*, *gain*, etc. The turn around is to match the sequence used in TGG, which then is more akin to IC-analysis (Chomsky 1957:26):

> Customarily [i.e., in American Structuralism], linguistic description on the syntactic
> level is formulated in terms of constituent analysis (parsing).

The equations of Harris can then be adapted in the following way (Chomsky 1957:26):

(7)	Sentence $\longrightarrow NP + VP$
(8)	$NP \longrightarrow T + N$

---

[5] Harris (1952a, 1952b) is the acknowledged source (Chomsky 1962:124):

> The approach to syntax that I want to discuss here developed directly out of the
> attempts of Z. S. Harris to extend methods of linguistic analysis to the analysis of the
> structure of discourse. This research brought to light a serious inadequacy of modern
> linguistic theory, namely, its inability to account for such systematic relations between
> sentences as the active-passive relation.

Cf. below.

(9) VP —> Verb + NP
(10) T —> *the*
(11) N —> *man*, *ball*, etc.
(12) Verb —> *hit*, *took*, etc.

And because the statements of (7) - (12) produce such structures as those in Figure 1, they are called *phrase structure rules*, and the (or any) grammar which is of this sort is a *phrase structure grammar*. And that is the syntax of American Structuralism ... a phrase structure grammar. The shared view of grammar is another continuation (iv) from American Structuralism.

2.2 *Transformations*

   But there must be more to TGG, or it will remain a notational variant of the syntax from which it arose. The 'more' is the relation of *transformation*. One of the early pieces of evidence for the necessity of this relation was the formal behavior of the verbal affixes in English: *-s*, *-ed*, *-en*, and *-ing*. Chomsky (1957:39 & 1962:131) presents this set of PS rules for rewriting Verb in English:

(13) Verb   —> Aux + V
(14) Aux —> C(M)   (*have+en*) (*be+ing*)
(15) M —> *can*, *will*, *may*, *shall*, *must*
(16) C —> past, present

Notice the position of **C**, the representation of past and present in these rules. It precedes the **M**[odals]. But consider these sentences:

(17) Klaus open*ed* the window.
(18) Klaus *was* opening the window.
(19) Klaus ha*d* opened the window.
(20) Klaus m*ight* open the window.

What appears to be the mark of past occurs in various positions: after the verb *open*, after the A[uxiliaries] *be* and *have*, and after the M[odal] *may*. The positioning is variable, and its correct placement in terms of a PS grammar will require its occurrence in multiple or variable locations, although it actually appears only once. The regularity is that it follows the first element to its right in the category Verb, whatever that may be. Cf. (14). The optionality of the members of Aux, indicated by the parentheses, now embodies the

variation. What we need is a rule that displaces C one position to the right in **Verb** and which affixes **C** to that category. This same rule can also be made to operate upon *-en* (the past passive participle) and upon *-ing* the active participle to relocate them to their appropriate positions (Chomsky 1957:61ff.). Expression of such formal regularities requires that we allow a rule to move a category from one location in a phrase structure to another.[6] Compare Figures 2 and 3.
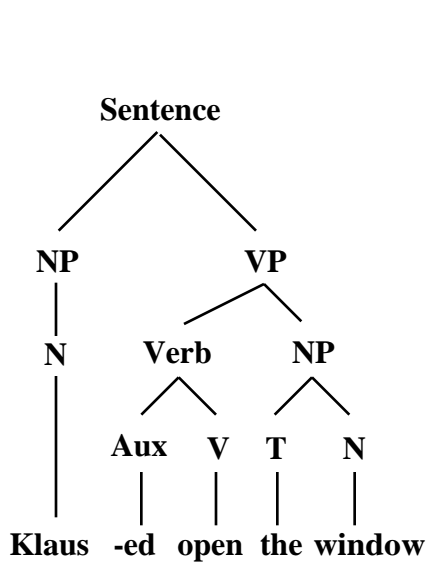

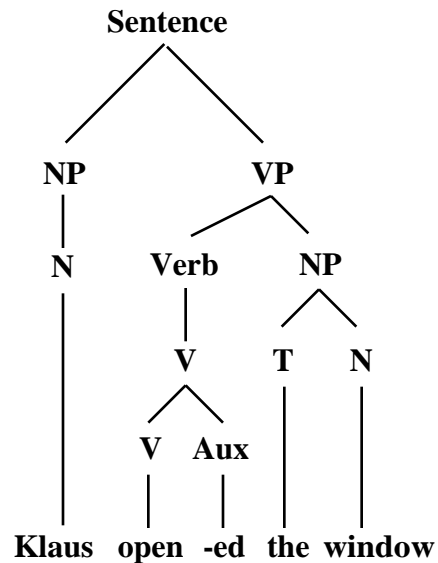Figure 2: *Before movement.*   Figure 3: *After movement.*

The other ... and perhaps the principle ... example to justify the notion of transformation was the relation between the active and the passive expressions in English. And that derives directly from Harris (1952a). Harris (1952a, 1952b) attempted to extend his method of grammatical analysis to stretches of

---

[6] Hockett (1954:211) evaluates a static Item and Arrangement (IA) model with a dynamic Item and Process (IP) model (which is what TGG is):

> The younger model, IA, has been formulated at least in part of because of a feeling of dissatisfaction with the 'moving-part' or 'historical' analogy implicit in IP ... For example ..., if it be said that the English past-tense form *baked* is 'formed' from *bake* by a 'process' of 'suffixation', then no matter what disclaimer of historicity is made, it is impossible not to conclude that some kind of priority is being assigned to *bake*, as against either *baked* or the suffix. And if this priority is not historical, what is it? Supporters of IP have not answered that question satisfactorily.

speech larger than a single sentence. The assumption which motivated the extension was similar to that which underlies his analysis of identification of morphemic variants, an assumption that is embodied in Hockett's later observation that morphs are not native to phonology. And it is their foreign nature allows them to be detected in phonology as potential units for further analysis. Harris (1952a:3) observes:

> Although we cannot state the distribution of sentences (or, in general, any inter-sentence relation) when we are given an arbitrary conglomeration of sentences in a language, we can get quite definite results about certain relations across sentence boundaries when we consider just the sentences of a particular connected discourse — that is, the sentences spoken or written in succession by one or more persons in a single situation. This restriction to connected discourse does not detract from the usefulness of the analysis, since all language occurrences are internally connected. Language does not occur in stray words or sentences, but in connected discourse ...

On this assumption of asymmetrical distribution, Harris proceeds as he did in working from morpheme variant to morpheme and from morpheme to utterance. The same analytical method works all the way from morph to discourse. In this procedure, one does not deal with unique forms/tokens but with *equivalence classes* (Harris 1952a:7):

> After discovering which sequences occur in equivalent environments, we can group all of them together into one equivalence class.[7]

Harris (1952a:16-17) considers a text which contains these sentences:

> Casals who is self-exiled from Spain, stopped performing after the fascist victory ...The press failed to say why he stopped performing, etc. But he has stated publicly why he is self-exiled, etc. ... The self-exiled Casals is waiting across the Pyrenees for the fall of Franco ...

and the equivalence classes of Figure 4 are proposed.

---

[7] Recall the focus classes of immediate constituent analysis.

C	S_2	S_1

Casals, who is self-exiled from Spain, stopped performing after the fascist victory.

__	S_2	C	S_3

The self-exiled Casals is waiting across the Pyrenees for the fall of Franco.
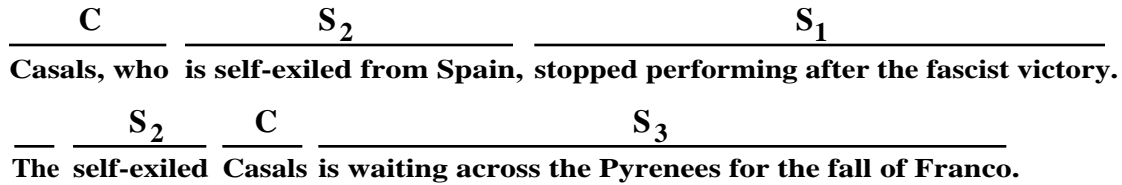
Figure 4: *Equivalence classes in discourse.*

Without quibbling about their establishment, we can see that a certain relation exists between CS_2S_1 and S_2CS_3 as does Harris (1952a:17):

> Let us now say that any sentence X_1AX_2 can be 'transformed' [There it is!!, PWD] into *A is X_1: AX_2*. This means that if X_1AX_2 occurs in the text, then *A is X_1: AX_2* also occurs in the text. In that case we will consider X_1AX_2 equivalent to *A is X_1: AX_2*; as a new structure our maverick has disappeared. We replace S_2CS_3 by the transformationally equivalent *C is S_2* and CS_3, both of which occur elsewhere in the same text.

The notion of transformation is then a formal one based entirely on the methodological technique of equivalence classes and — be it noted — discourse. But it is possible to (Harris 1952a:17, 20-21):

> ... proceed on this basis even to transformations which are not already justified by the text, provided they do not conflict with the text ... Grammatical equivalence can be investigated more systematically if we introduce a technique of experimental variation ... [which] can be used in the language outside the text [Read 'performance', PWD], where we have the **right**, **as speakers** [Read 'competence'. Emphasis mine, PWD], to create any social situation which might favor another speaker's uttering one rather than another of the many sentences at his disposal.

And we have justification ('right') for the later movement from situated discourse (performance) to isolated competence.

The discussion of the active-passive pairing now occurs in this context (Harris 1952a:3-4):

> ...we want to use the method in order to find out all that we can about a particular text ... To this end we would use only those statements of the grammar of the language which are true for any sentence of a given form. For example, given any English sentence of the form N_1VN_2 (e.g. *The boss fired Jim*), we can get a sentence

with the noun phrases in the reverse order N_2—N_1 (*Jim — the boss*) by changing the suffixes around the verb: *Jim was fired by the boss*. The justification for using such grammatical information in the analysis of a text is that since it is applicable to any N_1VN_2 sentence in English it must also be applicable to any N_1VN_2 sentence in the particular text before us, provided only that this is written in English ... It merely transforms certain sentences of the text into grammatically equivalent sentences ... in such a way that the application of the discourse-analysis method becomes more convenient [!] ...

Harris (1952a:19) does not conclude that such pairs mean the same since he is not concerned with meaning; the two are equivalent in that "they occur in similar structural positions in the text".

3. *Conclusion*
    The focus on the relation between sentences arises first in discourse study. Neither the emphasis on such relations nor the notion of transformation for its expression is Chomsky's, as (more generally in American Structralism) the notion phrase structure rule (equation) is not Chomsky's contribution. It ***all*** originates in Harris's work and is adapted by Chomsky.[8] The motivation for transformation was originally one of textual equivalence and restricted to one specific text. The idea of the 'technique of experimental' variation (again Harris's) allowed the possibility of generalizing (abstracting) the relation beyond the text and for the retreat from discourse and from parole to the safer confines of the sentence and language. That is Chomsky's contribution.

[Version: April 23, 2014]

---

[8] "[the term] ... 'grammatical transformation' itself, is Harris'. It is, like 'transformation,' being used here in a sense somewhat different from his, and in a different general framework" (Chomsky 1962:136). Most importantly, it has nothing to do with discourse now.