

in C. P. Otero (ed.), *Noam Chomsky: Critical Assessments*, Routledge, 1994, pp. 653-690.

Conceptual Shifts in the Science of Grammar: 1951-1992*

Robert Freidin

Princeton University

Chomsky's work in linguistics over the past forty-one years has resulted in a number of fundamental changes in our conception of grammatical theory, its scope and practice. The way in which Chomsky's early work diverged from that of his teacher and mentor, Zellig Harris provides one interesting and important case study. Another concerns the shift in focus from systems of rules to systems of grammatical principles, and the related shift of emphasis from derivations to representations, both of which occurred during the 1970's and crystallized in Chomsky 1981. Most recently, Chomsky has attempted to give a more general interpretation to the interaction of rules and principles in Universal Grammar (UG) in terms of language design, and has demonstrated how this interpretation constitutes a minimalist program for linguistic theory. In addition to opening new lines of investigation, this work promises to reshape our understanding of the language faculty in essential ways.

1. The advent of transformational generative grammar

Chomsky's formal introduction to linguistics began in 1947 when he read the proofs of Harris's *Methods of Structural Linguistics* (1951, henceforth *Methods*). As Harris described the book in the introduction:

This volume presents methods of research used in descriptive, or more exactly, structural, linguistics. It is thus a discussion of the operations which the linguist may carry out in the course of his investigations, rather than a theory of the structural analyses which result from these investigations. The research methods are arranged here in the form of the successive procedures of analysis imposed by the working linguist upon his data.

The operations that Harris details in *Methods* are essentially taxonomic procedures of segmentation and classification.

The basic operations are those of segmentation and classification.

Segmentation is carried out at limits determined by the independence of the resulting segments in terms of some particular criterion. If X has a limited distribution in respect to Y, or if the occurrence of X depends upon (correlates completely with) the occurrence of a particular environment Z, we may therefore not have to recognize X as an independent segment at the level under discussion [footnote omitted--RF]. Classification is used to group together elements which substitute for or are complementary to one another [footnote omitted--RF]. (p. 367)

Harris goes on to say in a footnote that “the class of elements then becomes a new element of our description on the next higher level of inclusive representation.” These operations yield a grammar of lists.

In one of its simplest forms of presentation, a synchronic description of a language can consist essentially of a number of lists. (p. 376)

These include a segment-phoneme list, a phoneme distribution list, several morphophonemic lists, lists dealing with type and sequences of morphemes, a component and construction list, and a sentence list--the list of utterance structures.¹ The purpose of this grammar of lists is to state all the regularities in a language, derived from an exhaustive analysis of a presumably representative corpus of utterances.

In striking contrast, Chomsky’s first major work in linguistics, his master’s thesis *The Morphophonemics of Modern Hebrew* (henceforth *MMH*)², virtually bypasses these taxonomic procedures and focuses instead on the construction of a grammar that describes grammatical sentences in a language, of which the linguist’s analyzed corpus is merely a special subclass. In this work Chomsky equates the linguistic analysis of a language with “the process of converting an open set of sentences--the linguist’s incomplete and in general expandable corpus--into a closed set--the set of grammatical sentences--and of characterizing this latter set in some interesting way.”³ He goes on to distinguish linguistic method (e.g. as discussed in Harris’s book) from linguistic description.

Accordingly we might distinguish and consider separately two aspects of the linguistic analysis of a language, a process of 'discovery' consisting of the application of the mixture of formal and experimental procedures constituting linguistic method, and a process of 'description' consisting of the construction of a grammar describing the sentences which we know from step one to be grammatical, and framed in accordance with the criteria related to its special purposes. (p. 2)

In *MMH* Chomsky assumes that the two processes are sequential steps in linguistic analysis. Yet his work on Hebrew concerns only the second step.

The outline of Modern Hebrew grammar given below is an example of the second step in linguistic analysis, artificially isolated. (p. 3)

In attempting to motivate the distinction between discovery and description, Chomsky notes that for a linguistic grammar "considerations of elegance" are operative in both processes.

However it will still be useful to consider the processes of discovery and description separately. For the most reasonable way to approach to the investigation and analysis of the notions of simplicity in terms of which 'grammatical in \underline{L} ' is defined (i.e., those notions of elegance that are relevant to the very formulation of the procedures of linguistics) seems to be to assume, for some language, that the grammatical sentences are fixed (i.e., that the process of discovery has been completed) and to determine the effect on grammar-formulation of explicit considerations of simplicity imposed on the grammatical statement. [footnote omitted--RF] (p. 3)

In effect, these remarks indicate how Chomsky's initial work in linguistics is already moving in the direction of a theory of structural analyses (presumably, but not necessarily, derived from the application of linguistic method)--in contrast to Harris's *Methods* where the distinction between method and theory is made (cf. the first quote from *Methods* on page 1) but the issue of theory is not pursued.

Another point where Chomsky's early practice diverges from Harris's method concerns a distinction which Harris sets up in section 20.21 of *Methods*. Harris notes that an investigator in

linguistics faces a choice of purposes: to state “all the regularities which can be found in any stretch of speech, so as to show their interdependences (e.g. in order to predict successfully features of the language as a whole)”, or to synthesize “utterances in the language such as those constructed by native speakers” on the basis of some minimal information (p. 365). The procedures presented in *Methods* are geared most naturally to the former purpose. Moreover, as Chomsky notes in the introduction to *The Logical Structure of Linguistics Theory* (1955-56, published 1975--henceforth *LSLT*), “Harris did not elaborate on the suggestion that a grammar can be regarded as a device for “synthesizing utterances,” an idea that does not, strictly speaking, seem compatible with the general approach of *Methods*.” (p. 50:fn.45). In *MMH* Chomsky gives an explicit interpretation to Harris's notion of “synthesizing utterances” in terms of a generative grammar. Chomsky describes the origin of this work as follows:

Harris suggested that I undertake a systematic structural grammar of some language. I chose Hebrew, which I knew fairly well. For a time I worked with an informant and applied methods of structural linguistics as I was then coming to understand them. The results, however, seemed to me rather dull and unsatisfying. Having no very clear idea as to how to proceed further, I abandoned these efforts and did what seemed natural; namely, I tried to construct a system of rules for generating the phonetic forms of sentences, that is, what is now called a generative grammar. I thought it might be possible to devise a system of recursive rules to describe the form and structure of sentences, recasting the devices in Harris's methods for this purpose, [footnote omitted--RF] and thus perhaps to achieve the kind of explanatory force that I recalled from historical grammar. [footnote omitted--RF] (*LSLT*, p. 25)

Although the introductory section of *MMH* suggests that “the application of the mixture of formal and experimental procedures constituting linguistic method” (e.g. Harris's) constitute step one in linguistic analysis, Chomsky's actual practice demonstrated that step one could be bypassed.⁴ It is the implementation of “step two” in linguistic analysis that leads Chomsky to the formulation of a theory of structural analyses in *LSLT*.

Note that Chomsky's notion of "synthesizing utterances" differs significantly from Harris's. In section 20.3 (Description of the Language Structure), Harris characterizes linguistic analysis as follows:

The work of analysis leads right up to the statements which enable anyone to synthesize or predict utterances in a language. These statements form a deductive system with axiomatically defined initial elements and with theorems concerning the relations among them. The final theorems would indicate the structure of the utterances of the language in terms of the preceding parts of the system. (pp. 372-373)

Although the mention of "deductive system" has a distinctly contemporary ring--recall Chomsky's many discussions of the deductive structure of the current principles and parameters theory of UG⁵, Harris's notion of synthesizing utterances involves an actual synthetic process, starting with phonetic units and proceeding from them to the construction of phonemes, morphemes, and so on to the construction of utterances. For Chomsky, however, the notion "synthesizing utterances" is interpreted analytically. Chomsky is quite explicit about this in the introductory section of *MMH*, in the version submitted to the University of Pennsylvania:⁶

... the theoretical linguistic system constructed is available for descriptions of small parts of the realm of experience constituted by the totality of its individuals. In one of the most interesting cases, the part to be described is an individual language. As above, the system can be applied in two ways. One can begin with the elementary phonetic units, construct phonemes, morphemes, syntactic classes, etc., proceeding synthetically. Or one can state the most general unit (i.e., the sentence) in terms of its constituents (e.g., the particular phrases of the language), further analyze these into their constituents, etc., until finally every possible sentence is represented in terms of phonetic units, thus proceeding analytically. Again, the choice will depend on considerations of elegance and adequacy. (p. 3)

For Chomsky, *MMH* is "an attempt to carry out an analytic statement of Modern Hebrew grammar." In context, it constitutes a radical departure from the kind of linguistics Chomsky

had been learning as a graduate student. We have some indication of how radical a departure from the fact that Chomsky himself considered the work he was doing in *MMH* as “more or less as a private hobby, having no relationship to ‘real linguistics’” (*LSLT*, p. 29).⁷

Adopting the analytic approach to grammar construction led Chomsky to postulate abstract underlying morphological forms, which were simply outside the range of structures adducible via Harris’s methods of segmentation and classification. Note that the many of the intermediate derived forms are also beyond the range of Harris’s methods. Consider for example the sample derivation given in section 6, A.2 (where line 0 gives the underlying morphophonemic form):

- | | | |
|-----|------------------------------|-----------|
| 0. | kY ₂ m+a--á+m-+F | |
| 1. | kY ₂ m+a--á:+m-+F | [MR 2] |
| 2. | kY ₂ á:m+m-+F | [MR 3] |
| 3. | m+kaY ₂ á:m+F | [MR 8] |
| 4. | m+kY ₂ á:m+F | [MR 9] |
| 5. | m+kY ₂ á:m+F+A | [MR 13.1] |
| 6. | m+kY ₂ á:m+a+A | [MR 13.4] |
| 7. | kY ₂ á:m+aA | [MR 16] |
| 8. | ká:maA | [MR 22] |
| 9. | ká:má | [MR 35] |
| 10. | ka:má | [MR 37] |
| 11. | kamá | [MR 44] |

The derivation of *kama* under this analysis involves the sequential application of a set of morphophonemic rules where the output of one rule serves as the input of another. The derivations proceed analytically from a morphophonemic representation to a phonemic representation--in contrast to Harris’s synthetic procedure which begins with phonetic analysis and progresses to phonemic and then morphophonemic level of representation, presumably

without intermediate representations between the phonetic/phonemic or phonemic/morphophonemic levels. Recall that Chomsky's morphophonemic analysis is itself embedded in a broader syntactic analysis of Hebrew sentences, sketched in the introductory sections of *MMH* though not elaborated there.

Chomsky's use of abstract underlying representations (in contrast to Harris) is also a distinguishing feature of his syntax.

By 1952, I was working on generative grammar of English, and shortly obtained results that I found quite exciting, though they were entirely divorced from the systems of procedures of analysis on which I was working at the same time; in particular, results on the system of auxiliary verbs in simple declaratives, interrogatives, negatives, passives, and on complex verb constructions such as "want (*NP*) to (*VP*)," "consider *NP* (to be) *Predicate*." [footnote omitted--RF] As in the case of my earlier work on morphophonemics of Hebrew, it was possible, so it became clear, to discover systems of rules that made sense of the distribution of forms, principles that served to explain the collection of superficially chaotic and anomalous facts. In this case, too, investigation led to more abstract underlying structures that were far removed from anything that might be obtained by systematic application of procedures of analysis of the sort that I was investigating. *LSLT*, pp. 30-31.

The systems of rules Chomsky is referring to here are of course those of phrase structure and transformational structure as he was then developing them.⁸

It is well known that Chomsky's theory of transformations is very different from Harris's.⁹ One of the most salient differences between the Harris and Chomsky theories is that grammatical transformations for Harris were essentially equivalence relations among sentences, whereas for Chomsky they were rules of grammar. Thus Harris's first published discussion of grammatical transformations in section 2.33 of Harris 1952 uses the term in the title of the section but not in the text of that section. Instead, he discusses the procedures of establishing equivalences and providing several examples:

But what is ‘equivalence’? Two ELEMENTS are equivalent if they occur in the same environment within the sentence. Two SENTENCES in a text are equivalent simply if they both occur in the text (unless we discover structural details fine enough to show that two sentences are equivalent only if they occur in similar structural positions in the text). Similarly, two sentences in a language are equivalent if they both occur in the language. In particular, we will say that sentences of the form A are equivalent to sentences of the form B , if for each sentence A we can find a sentence B containing the same morphemes except for differences due to the difference in form between A and B . For example, $N_1 V N_2$ is equivalent to N_2 is V -en by N_1 because for any sentence like *Casals plays the cello* we can find a sentence *The cello is played by Casals*.¹⁰

In Harris 1957 he gives a more formal definition of what constitutes a transformational relationship:

If two or more constructions (or sequences of constructions) which contain the same n classes (whatever else they may contain) occur with the same n -tuples of members of these classes in the same sentence environment (see below), we say that the constructions are transforms of each other, and that each may be derived from any other of them by a particular transformation. (p. 147)

Moreover, this relationship is predicated in terms of strings rather than constituent structure, as in Harris 1965 where he explicitly rejects constituent structure analysis in favor of string analysis. Thus for Harris, a transformation is a relation that holds between pairs of sentences.

In striking contrast, Chomsky defines a transformation as a mapping of a string with its phrase structure interpretation (i.e., a phrase-marker) onto another phrase-marker. Such mappings may iterate and thus cannot be simply relations between sentences. Furthermore, they crucially rely on constituent structure analysis, unlike Harrisian transformations. Whereas Harris uses transformations as a means for describing the syntax of texts, Chomsky employs transformational analysis to get at a more explanatory analysis of the structure of language. As he notes in *Syntactic Structures*, for example, “transformational analysis brings out the fact that negatives and interrogatives have fundamentally the same ‘structure,’ and it can make use of

this fact to simplify the description of English syntax” (pp. 64-65). This kind of insight, which involves positing abstract underlying structures and intermediate stages of derivation is simply beyond the power of Harris’s transformational analysis, which is restricted to relations between pairs of surface strings.

Perhaps the greatest difference between Harris and Chomsky concerns their interpretation of linguistic analysis. From the very outset, Chomsky adopts a realist interpretation of grammar. Language is somehow a real phenomenon of the physical world and as such has a particular structure--i.e., it is structured one way and not just any way. Thus there is a fact of the matter for all grammatical descriptions because they make empirical claims. Harris, in contrast, does not adopt the realist position.

The methods described here do not eliminate non-uniqueness in linguistic descriptions. It is possible for different linguists, working on the same material, to set up different phonemic and morphemic elements, to break phonemes into simultaneous components or not to do so, to equate two sequences of morphemes as being mutually substitutable or not to do so. The only result of such differences will be a correlative difference in the final statement as to what the utterances consist of. The use of these procedures is merely to make explicit what choices each linguist makes, so that if two analysts comes out with different phoneme lists for a given language we should have exact statements of what positional variants were assigned by each to what phonemes and wherein lay their differences of assignment.

The methods presented here are consistent, but not the only possible ones of arranging linguistic description. (p. 2)¹¹

It is obvious from this passage that for Harris there is no fact-of-the-matter in linguistic descriptions. He goes on to say that “the particular way of arranging the facts about a language which is offered here will undoubtedly prove more convenient for some languages than for others.” Thus they avoid “the undesirable effect of forcing all languages to fit a single Procrustean bed, and of hiding their differences by imposing on all of them alike a single set of logical categories. If such categories were applied, especially to the meanings of forms in

various languages, it would be easy to extract parallel results from no matter how divergent forms of speech....” Because Harris’s procedures “are merely ways of arranging the original data; and since they go only by formal distinctions there is no opportunity for uncontrolled interpreting of the data or for forcing of the meaning.” As a result, “the data, when arranged according to these procedures, will show different structures for different languages.” Thus Harris seems to deny the generality of his methods of analysis as part of a theory of linguistic structure. His focus here appears to be on the apparent diversity of linguistic forms rather than on some underlying unity.

In rather striking contrast, consider the opening paragraph of chapter 2 “The Nature of Linguistic Theory” of *LSLT*:

Descriptive linguistics is concerned with three fundamental problems. On the one hand, the descriptive linguist is interested in constructing grammars for particular languages. At the same time, he is interested in giving a general theory of linguistic structure of which each of these grammars is an exemplification. Finally, he must be concerned with the problems of justifying and validating the results of his inquiries, and demonstrating that the grammars that he constructs are in some sense the correct ones. All three of these problems will occupy us in this investigation of linguistic structure.
(p. 77)

The agenda that Chomsky maps out here is clearly more ambitious than Harris’s. In addition to providing linguistic descriptions of particular languages, linguistic theory under Chomsky’s conception must provide a general theory of linguistic structure--something which Harris says his procedures will not provide. And furthermore linguistic theory must deal with the problems of justification and validation under the realist interpretation (which follows from the assumption that a grammar can be correct or incorrect).

Chomsky’s realist interpretation of linguistic theory is grounded in the psychological interpretation of grammar: a grammar exists as part of the real world in the mind/brain of a speaker and constitutes that speaker’s knowledge of his/her language. Although this

psychological interpretation is not discussed in Chomsky's published writings until 1965 (in the first chapter of *Aspects*, written in 1958-59¹²) Chomsky has noted that during the period in which he wrote *LSLT* (that is, circa 1955), "it lay in the immediate background of my own thinking."¹³ However, he refrained from raising the issue at that time because it seemed to him "too audacious." The earliest reference to the psychological interpretation of generative grammar occurs in the last section of Robert Lees's review of *Syntactic Structures in Language* (1957). The final section (titled "Learning Theory") begins:

Perhaps the most baffling and certainly in the long run by far the most interesting implications of Chomsky's theories will be found in their cohesions with the field of human psychology.

Lees goes on to discuss Chomsky's model of grammar as a scientific theory of human linguistic behavior. He concludes the section and the review by raising the fundamental issue of language acquisition.

We come now to the point of this lengthy discussion of induction versus theory construction. Though it is possible, it is certainly not an easy task for the psychologist to explain the mechanism by means of which a child, confronted with a vast and perplexing array of different stimuli, manages to learn certain things which can be generalized by induction from repeated occurrences. We would not ordinarily suppose that young children are capable of constructing scientific theories. Yet in the case of this typically human and culturally universal phenomenon of speech, the simplest model that we can construct to account for it reveals that a grammar is of the same order as a predictive theory. If we are to account adequately for the indubitable fact that a child by the age of five or six has somehow reconstructed for himself the theory of his language, it would seem that our notions of human learning are due for some considerable sophistication.

To Chomsky it seemed obvious "that language acquisition is based on the child's discovery of what from a formal point of view is a deep and abstract theory---a generative grammar of his language--many of the concepts and principles of which are only remotely related to experience

by long and intricate chains of unconscious quasi-inferential steps.”¹⁴ Furthermore, Chomsky assumed that much of this theory must be innate in the mind/brain of the language learner. “A consideration of the character of the grammar that is acquired, the degenerate quality and narrowly limited extent of the available data, the striking uniformity of the resulting grammars, and their independence of intelligence, motivation, and emotional state, over wide ranges of variation, leave little hope that much of the structure of language can be learned by an organism initially uninformed as to its general character.”¹⁵

The psychological interpretation of grammar leads to a clarification of the notion “language”-though it should be noted that this clarification does not appear in print until Chomsky 1986.¹⁶ The psychological interpretation concerns the internalized grammar of a speaker, a structure of the mind/brain--what Chomsky calls an I-language, where “I” is intended to suggest internal, individual, and intensional. In contrast, structural and descriptive linguistics is concerned with a rather different notion of language, “as a collection of actions, or utterances, or linguistic forms (words, sentences) paired with meanings, or as a system of linguistic forms or events.”¹⁷--what Chomsky calls an E-language, where “E” is intended to suggest external and extensional. The latter notion has no status under the psychological interpretation.¹⁸

In Chomsky’s earliest work on generative grammar, before he explicitly adopted the psychological interpretation, the discussion of the concept “language” tends in the direction of E-language--not surprising for someone trained in the tradition of structural linguistics. Consider, for example, the way Chomsky defines language and grammar in *LSLT*.

We define a *language* to be a set (in general, infinite) of strings in a finite alphabet, each string being of finite length. We define a *grammar* of the language *L* as a finite device which generates all and only the strings of *L* in a determinate way. (p. 71)¹⁹

Given such definitions, a primary goal of linguistic analysis is to construct a system of rules from which the language can be derived. This assumes that a language is in some sense a well defined object--a false assumption simply because there is no way independent of the grammar

postulated by the linguist to determine exactly all and only the grammatical sentences of a language.

Under the psychological interpretation, grammar construction is subject to a further constraint. In addition to accounting for linguistic phenomena, a grammar must meet a criterion of learnability. Thus the system of rules in a grammar must plausibly be learnable given the primary language data available to a child learning a first language. Given that the theory of grammar under the psychological interpretation constitutes a theory of the language faculty, the innate mental structure a child employs in grammar construction, the more linguistic phenomena that can be derived from the theory of grammar, the less that needs to be accounted for by constructing language specific rules. Therefore, the psychological interpretation leads to a shift in focus from rule systems for particular languages to the more general properties of the language faculty. What happened in actual practice in the development of generative grammar was that investigations of general constraints on rule systems lead to a significant simplification of the rule systems themselves, as will be discussed in the following section.

2. From Rules to Principles

In *LSLT* Chomsky has two related goals, to develop a general theory of structural analyses and to show how this theory applies to natural languages. The initial work on generative grammar is therefore focused on rule systems for the description of language phenomena. Thus *LSLT* is organized with a chapter of phrase structure analysis followed by one on the phrase structure of English (chapters VII and VIII respectively) and a chapter on transformational analysis followed by one on the transformational analysis of English (chapters IX and X respectively). The focus on rule systems predominates in the early work on generative grammar into the early 1960s (especially in *Syntactic Structures* and Chomsky's contribution to the 1958 Texas Conference on Problems of Linguistic Analysis in English, published in 1962).

2.1 A brief history of the passive transformation

Consider for example, the analysis of passive sentences and the passive transformation, perhaps the most central transformation within early transformational analysis. In *SS* we find the following formulation:

(1) *Passive* -- optional:

Structural analysis: $NP - Aux - V - NP$

Structural change: $X_1 - X_2 - X_3 - X_4 \rightarrow X_4 - X_2 + be + en - X_3 - by + X_1$

The passive transformation as formulated in (1) performs several distinct operations. It permutes a NP subject in front of an auxiliary phrase²⁰ with a NP following the verb (usually construed to be the object of the verb). In addition it adjoins the passive auxiliary *be+en* to the right of the auxiliary phrase, and also adjoins the grammatical formative *by* to the left of the permuted subject NP. The rule is stipulated to be optional, indicating that it need not apply even if its structural description is met in the course of a derivation.²¹ Given this formulation, an active sentence and its corresponding passive counterpart would have the same underlying deep structure. Thus the transformational analysis expressed the relation between active and passive sentential structures in terms of a common deep structure.

In *Aspects* Chomsky tries to improve on the analysis of the passive transformation by proposing that the passive *by*-phrase constituted an instance of a manner adverbial. The motivation comes from an observation that passivization seemed possible only for verbs that could occur with manner adverbs.²²

These observations suggest that the Manner Adverbial should have as one of its realizations a “dummy element” signifying that the passive transformation must obligatorily apply. That is, we may have the rule (55) as a rewriting rule of the base and may formulate the passive transformation so as to apply to strings of the form (56), with an elementary transformation that substitutes the first NP for the dummy element *passive* and places the second NP in the position of the first NP:

(55) $Manner \rightarrow by\ passive$

(56) $NP - Aux - V - \dots - NP - \dots by\ passive - \dots$

(where the leftmost ... [a string variable--RF] in (56) requires further specification--e.g., it cannot contain an NP). (pp. 103-104)

Thus the *by*-phrase would exist in underlying structure, thereby eliminating the need to insert a language specific grammatical element (*by*) using the passive transformation. The analysis also solved the problem of the derived constituent structure of the *by*-phrase, which in *SS* had to be specified by an ad hoc rule of derived constituent structure.²³ It assumed that markers such as *passive* were “drawn from a fixed, universal, language-independent set” (p. 223).

The analysis of the passive auxiliary *be + en* is not discussed explicitly in *Aspects*. Note that the description of the passive transformation quoted above does not mention the insertion of the auxiliary as part of the transformation. This leaves open the possibility that the passive auxiliary is introduced into the derivation via a phrase structure rule. Yet it is not clear from the *Aspects* phrase structure analysis of Aux that the passive auxiliary could be introduced via a phrase structure rule. The rewrite rule for Aux given in (57.xvi) on page 107 is essentially the same as the one given in *SS* on page 111 (i.e. $Aux \rightarrow C (M) (have + en) (be + en)$) except that C is replaced by T (for tense) and the two aspectual auxiliaries are covered by a single designation “Aspect.” However it is unlikely that the passive auxiliary was assumed to be introduced via a phrase structure rule given the discussion in *SS* concerning the problems created by such an analysis.²⁴

In “Remarks on Nominalization” (written in 1967, published in 1970--henceforth *Remarks*), the analysis of passives and the passive transformation undergoes an even more radical revision. In *Remarks* Chomsky considers the correspondence between sentences and nominals, as illustrated in (2-3).

- (2) a. the enemy destroyed the city
b. the city was destroyed by the enemy
- (3) a. the enemy's destruction of the city
b. the city's destruction by the enemy

c. the destruction of the city by the enemy

Earlier work on nominalizations (i.e. Lees 1960) assumed that the nominals (3a-b) were derived via a nominalization transformation which applied to the sentential forms (2a-b) respectively.

Chomsky rejects the transformational analysis of nominals on the grounds that such transformations do not have the general character of other standard transformational rules.

They are not productive--that is, not all verbs have a corresponding derived nominal.

Furthermore, the relation between a derived nominal and its associated verb is idiosyncratic.

Consider, for example, such nominals as *laughter, marriage, construction, actions, activities, revolution, belief, doubt, conversion, permutation, trial, residence, qualifications, specifications*, and so on, with their individual ranges of meaning and varied semantic relations to the base forms. There are a few subregularities that have frequently been noted, but the range of variation and its rather accidental character are typical of lexical structure. To accommodate these facts within the transformational approach (assuming, as above, that it is the grammatical relations in the deep structure that determine meaning) it is necessary to resort to the artifice of assigning a range of meanings to the base form, stipulating that with certain semantic features the form must nominalize and with others it cannot. Furthermore, the appeal to this highly unsatisfactory device, which reduces the hypothesis that transformations do not have semantic content to near vacuity, would have to be quite extensive. (p. 19)

In addition, nominalization transformations, unlike other transformations, require the ability to change categorial analysis--for example, changing a verb into a noun and correspondingly a sentential structure into a NP. However, given the theory of syntactic features developed in *Aspects*, Chomsky is able to formulate the lexicalist hypothesis which can accommodate the analysis of derived nominals in the lexicon and thereby achieve a simplification of the transformational component. The lexicalist hypothesis prohibits feature changing by transformation as well as transformations whose application depends on idiosyncratic features of lexical items. It assumes that idiosyncratic properties of individual lexical items are captured in the lexicon.

Though the nominal (3b) is not transformationally derived from the sentence (2b), it is transformationally related to the nominal (3a) in much the same way that the passive construction in (2b) is transformationally related to the corresponding active construction (2a). The semantic relation between *city* and *destruction*, as well as *enemy* and *destruction*, is identical in both nominals even though their surface syntax is different. Chomsky analyzes the nominal (2b) as derived from an underlying structure resembling (2a).

(4) $[_{NP} [_{Det} [_{NP} \text{the enemy}]] [\text{destroy, +N}] [_{NP} \text{the city}] \text{ by } \Delta]$

Δ indicates a base generated empty category (NP). Thus in (4) *the city* bears the relation object of the nominal and *the enemy* bears the relation subject of the nominal in both (3a) and (3b). These grammatical relations are identical to those in the corresponding sentential forms (2a-b). Putting aside the insertions of 's in front of the nominal and *of* after the nominal, (3b) is derived from (4) via transformational operations that substitute the NP *the enemy* for Δ and fill the empty subject position with *the city* (via another substitution operation). Chomsky establishes the independence of the two operations by pointing out that Agent postposing may apply without the application of NP preposing (even when it could apply, as in (3c)), and that more generally, Agent postposing does not require the presence of an object as illustrated by the derivation of (5a) from (5b).

(5) a. $[_{NP} [_{Det} \text{the}] [\text{offer}] \text{ by John}]$
 b. $[_{NP} [_{Det} [_{NP} \text{John}]] [\text{offer, +N}] \text{ by } \Delta]$

If the *by*-phrase in (5b) were missing, then Agent postposing could not apply and *John's offer* would be derived. Furthermore, if the absence of the *by*-phrase in derived nominals is treated as the absence of the phrase in underlying structure, given its optionality in phrase structure, then NP preposing could apply in the absence of Agent postposing, as illustrated in (6a) (derived from (6b)).

(6) a. $[_{NP} [_{Det} [_{NP} \text{John's}]] [\text{appointment}] \text{ to the committee}]$

b. $[_{NP} [_{Det} [_{NP} \Delta]]] [\text{appoint, +N}] [_{NP} \text{John}] \text{ to the committee }]$

If NP preposing does not apply to (6b), then *the appointment of John to the committee* is derived.

Chomsky characterizes the two rules that apply to nominal constructions as follows:

Agent-postposing is simply a generalization of one of the components of the passive transformation. NP-preposing is similar to, and may fall under a generalization of, the other component. (1970:43)

Note that he still assumes the existence of a passive transformation even though he has just demonstrated that the two movement components of the rule that applies to sentential structures apply independently in nominal constructions. Presumably this is because he still assumes that the passive auxiliary must be inserted via the transformation given his arguments in *SS* against base-generating the passive auxiliary. Thus the analysis in *Remarks* identifies two subparts of the passive transformation as grammatical transformations with respect to nominals and as “elementary” transformations (in the sense of *LSLT*) with respect to sentential structures.²⁵ A similar analysis occurs in Hasegawa 1968 which identifies the postposing of agents (i.e. Agent-postposing) in gerundive nominals (e.g. *the polishing of the lenses by the workers*) as an independent grammatical transformation.²⁶

With hindsight, we now know that the analysis in *Remarks* constitutes the first important step towards a more general theory of transformations that does not allow the formulation of construction specific transformations like T_{passive} . It establishes that the movement component of the passive transformation is essentially a pair of substitution operations in which a lexical NP is substituted for an empty NP. The next step occurs with the realization that the rule of NP-preposing is actually an instance of a more general rule “Move NP”--or more accurately, Substitute NP. Thus much of the work of the early 1970s was directed towards an analysis of NP movement which generalized the NP-preposing component of the passive transformation to instances of subject-to-subject raising, as illustrated in (7), where t_i indicates the underlying position of the coindexed matrix subject and $[_\alpha$ indicates a clause boundary.

- (7) a. John_i is likely [_{α} t_i to win the election]
b. Bill_i seems [_{α} t_i to be ahead in the polls]
c. Mary_i is expected [_{α} t_i to defeat the incumbent]

(Note that the matrix predicate in each case is distinct—a predicative adjective in (7a), a verb in (7b), and a passive predicate in (7c).) Given a rule Move NP, the structural description for the transformation can be reduced to the bare essentials.

- (8) (*vbl*, NP, *vbl*, NP, *vbl*) [Chomsky 1976:(7)]

Only the movement site and the landing site are given as constant terms. The syntactic distance between the two positions must be mediated by other conditions which are not stated as part of the transformation—for example, the framework of conditions developed in Chomsky 1973, 1976. As it turned out, the development of a system of conditions on rules of grammar made it possible to formulate the more general theory of transformations, under the assumption that these conditions constituted general principles (i.e. of UG).

2.2. From conditions to Move α

In discussing the early work on the system of conditions—in particular, the Tensed-S, Specified Subject, and Subjacency conditions, Chomsky makes the following important observations:

If these principles can be substantiated or improved, the class of potential grammars is vastly reduced. Many potential rules are eliminated outright. Furthermore, by limiting the possible application of rules, principles of the sort discussed make it unnecessary to make available in the theory of transformations as rich an apparatus as would otherwise be needed to delimit the application of particular rules. Thus, the principles constrain the variety of grammars by reducing the “expressive power” of grammatical rules.

We might even set ourselves the goal, still distant but perhaps attainable, of so restricting the apparatus of the theory of transformations that rules can only be given in the form “move NP,” with other conditions on their

application expressed either as general conditions on rules or as properties of initial phrase markers, or as properties of surface structures.” (1975a:111-112)

Work on the system of grammatical principles progressed so rapidly, that in Chomsky 1976 the rule Move-NP is taken as an established result.²⁷ In particular, work on the syntax of Romance languages (e.g. French and Portuguese) demonstrated that not only did Chomsky’s Tensed-S and Specified Subject conditions apply crosslinguistically, but that they applied to constructions not found in English--e.g. clitic constructions (see Kayne 1975, Quicoli 1976a,b).

To see how the Tensed-S Condition (henceforth TSC) and the Specified Subject Condition (henceforth SSC) mediate the syntactic distance between the movement and landing site with respect to Move NP, consider the following derived structures.

(9) a. finite clause complements:

i. *John_i was reported [_α t_i had recommended Mary]]

ii. *Mary_i was reported [_α John had recommended t_i]]

b. infinitival complements:

i. John_i was reported [_α t_i to have recommended Mary]

ii. *Mary_i was reported [_α John to have recommended t_i]

In essence, the TSC prohibits the application of a rule that links two positions across a finite clause boundary. Thus the application of Move NP in (9.a.i-ii) violates this condition. The SSC prohibits the application of a rule that links two positions across a syntactic subject--hence the application of Move NP in (9.a.ii) and (9.b.ii) is excluded by the condition. In contrast, the application of Move NP in (9.b.i) violates neither condition (or any other principle of grammar) and therefore is allowed.

Under the trace theory of movement rules²⁸, where t_i constitutes an empty category left behind by the movement operation (e.g. substitution)²⁹, it became possible to relate NP movement phenomena to anaphor binding by treating trace-binding as another instance of bound anaphora. Under this analysis it was possible to explain why NP-preposing in passive

sentences was obligatory. As noted by Fiengo (1974, 1977), the trace binding relation in structures like (10) violates an independently motivated constraint on anaphor binding which accounts for (11).

(10) * t_i was [_{VP} criticized George by the committee_i]

(11) *himself_i [_{VP} criticized George_i]

Given that this constraint on anaphor binding is part of UG, we have derived a piece of the behavior of NP-preposing from a general principle of grammar and therefore need not stipulate it in the formulation of the transformation. Moreover, as demonstrated in Chomsky 1976, the TSC and SSC apply to anaphor binding as well as NP movement.

(12) a. finite clause complements:

i. *John_i believes [_α himself_i admired Mary]]

ii. *John_i believes [_α Mary admired himself_i]]

b. infinitival complements:

i. John_i believes [_α himself_i to admire Mary]]

ii. *John_i believes [_α Mary to admire himself_i]]

In (12a.i-ii) anaphor binding violates the TSC; whereas in (12.a.ii) and (12.b.ii) the binding relation violates the SSC. In this way the paradigms in (9) and (12) are analogues. As a consequence, it appears that much, if not all, of the behavior of NP movement can be derived from general principles of UG, and therefore that the movement transformation may be formulated simply as Move NP.

In Chomsky 1976 the TSC and SSC are treated as conditions on the application of rules. Thus in (9) they block the application of Move NP, whereas in (12) they block a rule of interpretation that links the anaphor *himself* to the antecedent *John*. Another way to interpret these conditions is as conditions on representations--or more specifically, as conditions on binding, assuming that a NP trace is the empty category analogue of a lexical anaphor.

Restricting conditions (4) and (5) [the Propositional Island Condition (PIC)³⁰ and the SSC respectively--RF], now, to rules of construal, we interpret them as applying to transformational rules as filters, in effect; the result of applying a transformational movement rule may or may not yield an appropriate case of “bound anaphora.” (Chomsky 1977:77)

Under this interpretation Move NP (or the rule that links an anaphor to an antecedent) applies freely. If the result is deviant, then the derived representation will violate some principle of grammar.³¹ The latter interpretation captures a significant generalization and is in addition motivated by a redundancy argument, as demonstrated in Freidin 1978. If we interpret these principles as conditions on the application of rules, then a further condition on the application of rules, the Strict Cycle Condition (Chomsky 1973), is required to rule out one derivation of such deviant sentences as (13).

(13) *John was reported that book to have been given

The Strict Cycle Condition (henceforth SCC) prohibits the application of a rule within a cyclic subdomain of a current cycle. The derivation of (13) which violates the SCC is given in (14), where *np* is construed as a base generated empty NP.

- (14)a. [_{α} *np* was reported [_{α'} *np* to have been given John that book]]
 b. [_{α} *np* was reported [_{α'} John_{*i*} to have been given e_{*i*} that book]]
 c. [_{α} John_{*i*} was reported [_{α'} e_{*i*} to have been given e_{*i*} that book]]
 d. [_{α} John_{*i*} was reported [_{α'} that book_{*j*} to have been given e_{*i*} e_{*j*}]]

On the first transformational cycle, α' , *John* is preposed to the complement subject position. Then on the second transformational cycle, α , *John* is raised to matrix subject position. If the derivation stops here, we have a well-formed sentence *John was reported to have been given that book*.

However, if Move NP applies again within α' , now a cyclic subdomain of α , and substitutes *that book* for the trace of *John* in complement subject position, the resulting derivation violates the

SCC. If, on the other hand, the SSC is interpreted as a condition on representations, then (14d) violates the SSC because *John* now binds its only trace across a subject even though no application of Move NP violated the SSC.³² Under this interpretation we derive the empirical effect of the SCC. If all the empirical effects of the SCC can be derived in this way, as argued in Freidin 1978, then the SCC is redundant with respect to those other principles of grammar from which its effects are derived. This result, as well as the binding theoretic interpretation of the TSC and SSC, suggests a shift in focus from conditions on rule application to conditions on syntactic representations, thereby reinforcing the shift from rules to principles.³³

If this reduction of the “passive transformation” to the rule Move NP is to succeed fully, then some account of the ‘special’ behavior of the passive auxiliary *be + en* is required. Recall that in Chomsky 1976 the analysis in which the passive auxiliary is introduced via the rule of NP-preposing is explicitly rejected. Instead it is assumed that this auxiliary “derives from an independent source” (p. 173). If this independent source is a phrase structure rule, then the argument in SS against base generating the passive auxiliary must be addressed. As Chomsky noted, the occurrence of the passive auxiliary is, unlike other auxiliary verbs, subject to a number of ad hoc restrictions:

- (15)a. It can be selected only if it is followed by a transitive V.
- b. It cannot be selected if NP follows V, even if V is transitive.
 - c. It must be selected if a transitive V is followed by an agentive *by*-phrase.

(15a-a) rule out sentential constructions like (16a-c) respectively.

- (16)a. *the explosion was occurred at noon
- b. *Adam was praised Bernie
 - c. *the candidate praised by the committee

Under the earlier transformational analysis of passives, these restrictions were accounted for by inserting the auxiliary as part of the passive transformation. However, if the passive auxiliary is

base generated, then these restrictions must be handled by other ways--optimally, as the effects of UG principles.

Such an account can be derived from a theory of predicate/argument relations developed during the 1970s. The theory concerns the assignment of semantic functions (or thematic relations (henceforth θ -roles)) to arguments of predicates. In active sentences the assignment of θ -roles to arguments is straightforward.

(17) Bernie praised Adam

In (17) the predicate *praised* assigns two θ -roles, one to the subject (θ_S) and one to the object (θ_O).

In (18), the passive counterpart to (17), θ -role assignment works somewhat differently.

(18) Adam_i was praised e_i by Bernie

The syntactic subject *Adam* is assigned θ_O by virtue of the trace it binds.³⁴ The object of the *by*-phrase, *Bernie*, is assigned θ_S by the preposition *by*.³⁵ It is standardly assumed that the passive morphology on the verb *praised* prevents it from assigning θ_S to the syntactic subject position as in the active construction. Furthermore it is generally assumed that any licit application of NP-preposing moves a NP to a nonthematic position--i.e. one that is not directly assigned a θ -role.³⁶ The possibility of NP movement between two thematic positions, as in (19), is prohibited by a general principle (20) governing the distribution of semantic functions.

(19) *Adam_i praised e_i

(20) Functional Uniqueness: an argument may bear only one semantic function.

Given the analysis of the active/passive pair (17-18), *Adam* would be assigned θ_S by virtue of occupying the subject position and would also be assigned θ_O via its trace in object position.

Because the sentence *Adam praised* has no interpretation 'Adam praised himself', the structure in (19) is not licit--as predicted by the Functional Uniqueness principle. A further condition (21)

rules out the possibility of a construction in which an argument bears no θ -role, as would occur if every syntactic position in a passive construction were filled with a lexical NP, as in (22).

(21) Functional Relatedness: each argument must bear a semantic function.

(22) *Adam was praised Barbie by Bernie

In (22) θ_S is assigned to *Bernie* and θ_O is assigned to *Barbie*. *Adam* would receive no θ -role assignment in violation of (21). Thus Functional Relatedness accounts for restriction (15b) on the distribution of the passive auxiliary. Notice that it would also account for restriction (15a) under the assumption that passive morphology blocks θ -role assignment to the syntactic subject position.

The final restriction, (15c), falls under another principle of predicate/argument structure--namely, that each semantic function of a predicate must be uniquely assigned (that is, assigned only once)--henceforth the principle of Unique Assignment. Given that θ_S is assigned to the syntactic subject position by a verb lacking passive morphology or by the preposition *by*, it would seem that in (16c) θ_S is assigned twice to two distinct arguments in violation of the Unique Assignment principle.³⁷ Thus, given the principles of Functional Relatedness and Unique Assignment, the apparently ad hoc restrictions on the distribution of the passive auxiliary follow from the theory of UG. These two principles plus Functional Uniqueness form a subcomponent of UG referred to as θ -theory.³⁸

Given the principle of Unique Assignment it would be possible to base generate a *by*-phrase containing a lexical object rather than one with an empty object position which is filled by a rule of Agent postposing--that is, an instance of Move NP. One argument for base generating the object of the *by*-phrase in place concerns the derivation of nominals like (23).

(23) the criticism of the chairman by the committee

Under trace theory, the Agent-postposing analysis leaves behind an NP trace in prenominal position. The derivation of (23) would require the substitution of that trace with the determiner

the, in violation of the Nondistinctness Condition on substitutions. Note that this analysis presupposes that there is only one prenominal position to be occupied by either a determiner or a possessional NP, a standard assumption during the 1970s.³⁹ Granting basic assumptions about the analysis of categories, a noun (and hence the NP it projects) is distinct from a determiner. If we eliminate the Agent-postposing analysis in nominals, then there is little reason to retain it for sentential constructions.⁴⁰

In eliminating the application of Move NP that postposes the underlying subject from the derivation of sentential passive constructions, we can no longer rely on binding theory to explain the obligatory nature of NP-preposing as discussed above. Instead, the obligatory character of NP-preposing is subsumed under Case theory, another subcomponent of UG which was developed during the late 1970s.

The basic idea behind Case theory, due to Vergnaud 1977, is that there are two kinds of NP positions, those that are marked for Case and those that are not, and only Case marked NPs can be lexically realized. Chomsky 1980 formulates this idea in terms of a filter (24), generally known as the Case Filter:

(24) *N, where N has no Case. (p. 25)

Given the Case Filter, a lexical NP which occurs in an underlying Caseless position must move to a Case-marked position. This accounts for the syntax of predicate adjectives which take sentential complements.

- (25) a. *it is likely [_α the situation to improve]
 b. the situation_i is likely [_α e_i to improve]
 c. it is likely [_α that the situation will improve]

In (25a) the complement subject *the situation* occurs in a Caseless position, in contrast to the finite clause complement subject in (25c). However, in (25b) the complement subject has been moved to the subject position of a finite clause (cf. (25c)) where it can receive Case. Thus (25a) violates

the Case Filter, whereas (25b-c do not). The NP-preposing that occurs in the derivation of (25b) is therefore required to satisfy the Case Filter.

This Case theoretic analysis generalizes to NP-preposing in sentential constructions containing predicates with passive morphology, as illustrated by the raising paradigm given in (26).

- (26) a. we expect [_α more women to win elections]
b. more women_i are expected [_α e_i to win elections]
c. *it is expected [_α more women to win elections]
d. it is expected [_α that more women will win elections]

(26a) shows that the subject of the infinitival complement of the verb *expect* is in a Case-marked position--presumably because of its syntactic relationship to the matrix verb since the subject of an infinitival is not itself a Case-marked position, as illustrated in (25a). However, the addition of passive morphology to the matrix verb renders the complement subject position Caseless, hence (26c) violates the Case Filter. Thus the complement subject *more women* must move to a Case-marked position to satisfy the Case Filter. As in (25b), movement to the subject position of a finite clause places the complement subject in a Case-marked position, thereby satisfying the Case Filter. The comparison of the two paradigms (25) and (26) suggests that passive predicates, in contrast to their nonpassive counterparts, share the same property as predicate adjectives--namely an inability to Case-mark a NP.⁴¹ "Suppose that the unique property of passive morphology is that it in effect "absorbs" Case: one NP in the VP with the passive verb as head is not assigned Case under government by this verb" (Chomsky 1981:124).⁴² The analysis generalizes without qualification to simplex passive constructions. Therefore, the obligatory character of NP-preposing follows from a principle of UG, the Case Filter.

Given the Case Filter, the Functional Uniqueness principle of the θ -Criterion, the TSC and the SSC, UG severely limits the syntactic behavior of the rule Move NP. NP movement is obligatory when an NP occurs in an underlying Caseless position. It must wind up in a Case-marked

position at S-structure. NP movement cannot occur between two θ -marked positions, nor can it cross a finite clause boundary or a c-commanding subject.

The framework of principles developed over the 1970s demonstrated that the behavior of transformations which had been stipulated in the structural descriptions and structural changes of more elaborate rules could be derived from general principles of UG.⁴³ As a result, it was possible to dispense with structural descriptions for transformations altogether. Under this framework, a transformation could be given as a structural change which indicated the category (or categories--e.g. in the case of adjunction) affected and the elementary operation involved. In later work, circa 1980, Chomsky generalizes transformations still further by replacing specific reference to categories (e.g. NP or *wh*-phrase) with a variable α and characterizing the two distinct elementary operations which can perform movements (i.e. substitution and adjunction) under the designation "move"--hence the rule "Move α ".⁴⁴

This change of view concerning transformations is discussed in Chomsky's *Lectures on Government and Binding* (1981, henceforth *LGB*).

In early work in generative grammar it was assumed, as in traditional grammar, that there are rules such as "passive," "relativization," "question-formation," etc. These rules were considered to be decomposable into more fundamental elements: elementary transformations that can compound in various ways, and structural conditions (in the technical sense of transformational grammar) that are themselves formed from more elementary constituents. In subsequent work, in accordance with the sound methodological principle of reducing the range and variety of grammars to the minimum, these possibilities of compounding were gradually reduced, approaching the rule Move- α as a limit. But the idea of decomposing rules such as "passive," etc., remained, though now interpreted in a rather different way. These "rules" are decomposed into the more fundamental elements of the subsystems of rules and principles (1) and (2). (p. 7)

In *LGB* the system of rules has the following subcomponents:

- (27) [= (1)] (i) lexicon

- (ii) syntax
 - (a) categorial component
 - (b) transformational component
- (iii) PF-component
- (iv) LF-component

and the system of principles includes the following:

- (28) [= (2)]
 - (i) bounding theory
 - (ii) government theory
 - (iii) θ -theory
 - (iv) binding theory
 - (v) Case theory
 - (vi) control theory

Given this view, descriptive categories like “passive” have no status under the theory of grammar. Virtually all the properties of the “passive transformation” have been derived from general principles of UG, including the movement operation itself, which is neither a language particular nor construction specific rule.⁴⁵

2.3 Towards a theory of language design

More recently Chomsky (1991) has suggested that the theory of grammar developed over the past two decades can be interpreted in terms of more general guidelines for language design. In essence, these guidelines legislate against “superfluous steps” in derivations and “superfluous elements” in representations. For example, the principle of Full Interpretation (Chomsky 1986) requires that every element in Phonetic Form (PF) and Logical Form (LF), the representations interfacing with systems of language use, must receive an appropriate interpretation. From Full Interpretation it follows that LF representations for natural languages may not contain vacuous quantifiers or arguments that are not functionally related to some predicate in the representation (as in (22)). Therefore the functional relatedness requirement of

the θ -Criterion is subsumed under the principle of Full Interpretation, a more general requirement that representations be minimal in some sense. In this way Full Interpretation functions as a principle which determines the economy of representations.

With respect to the economy of derivations, Chomsky 1991 proposes a “last resort” condition on movement operations to the effect that movement cannot apply to a constituent unless the nonapplication results in the violation of some principle of grammar. For example in the derivation of sentential passive constructions the last resort condition predicts that the object of the *by*-phrase is exempt from movement since it is both Case-marked and θ -marked in its underlying position, in contrast to the object of the passive predicate which is Caseless in its underlying position. Notice that the addition of this condition to UG makes a new prediction--namely that there are no languages in which a θ -marked and Case-marked NP can move to a Caseless position which receives no θ -marking. Without the last resort condition such movements are possible under the principles of UG. The last resort condition on movements comes under the more general heading of a “least effort” guideline for derivations.⁴⁶

The last resort condition on movements also subsumes some of the empirical effects of other UG conditions--in particular, the functional uniqueness requirement of the θ -criterion and the TSC and SSC. Consider the standard case of a Functional Uniqueness violation ((19) above, repeated here):

(19) *Adam_i praised e_i

(19) violates the last resort condition on movements because the underlying object *Adam* is both Case-marked and θ -marked in that position and therefore the movement is presumably not required by any principle of UG.⁴⁷ The same sort of analysis applies in the standard cases of TSC and SSC violations (e.g. (9a.i-ii) and (9b.ii) above). In each case the last resort condition on movement is violated because the NP moved is both Case-marked and θ -marked in its underlying position and therefore is not required to move by any UG condition. Although it remains to be demonstrated, it seems highly probable that the last resort condition on movements

may subsume virtually all the empirical effects of the TSC and SSC as well as Functional Uniqueness.

In his most recent work (1992) Chomsky develops the idea that linguistic derivations and representations should be minimal in a certain sense into a more ambitious program for linguistic theory. In this paper he proposes a minimalist program for linguistic theory based on a number of plausible, though at this point speculative, assumptions. The primary assumption that motivates much of this program is that only PF and LF count as levels of representation with respect to UG.

UG must determine the class of possible languages. It must specify the properties of SDs and of the symbolic representations that enter into them. In particular, it must specify the interface levels (A-P [articulatory-perceptual], C-I [conceptual-intentional]), the elements that constitute these levels, and the computations by which they are constructed. A particularly simple design for language would take the (conceptually necessary) interface levels to be the only levels. That assumption will be part of the “minimalist” program I would like to explore here. (p. 3)

The assumption is motivated by the fact that most principles of UG apply to either LF or PF representations. Part of the minimalist program involves a demonstration that virtually all principles of UG apply at PF or LF and that apparent counterexamples of principles that hold at S-structure or D-structure can be reanalyzed so that only PF and LF are the relevant levels.

Conditions on representations -- those of Binding Theory, Case Theory, Theta Theory, etc. -- hold only at the interface, and are motivated by properties of the interface, perhaps properly understood as modes of interpretation by performance systems. (Chomsky 1992:6)

To the extent that conditions on the economy of representations and derivations restrict the kinds of representations that occur at the interface levels, these conditions are crucial to carrying out the minimalist program.

The discussion of the role of economy in grammatical description dates back to the advent of generative grammar--i.e. Chomsky's *MMH*. There Chomsky identifies two kinds of criteria of adequacy for grammars--one concerning the correct description of the structure of the language under analysis, and the other concerning requirements imposed by its special purposes, "or, in the case of a linguistic grammar having no such special purposes, requirements of simplicity, economy, compactness, etc." (p. 1). In a footnote, Chomsky adds the following clarification:

Such considerations are in general not trivial or 'merely esthetic'. It has been recognized of philosophical systems, and it is, I think, no less true of grammatical systems, that the motives behind the demand for economy are in many ways the same as those behind the demand that there be a system at all. Cf. Goodman (1943).

In other words, a grammar is not merely a description of a language; it is moreover an explanatory theory about the structure of a language--i.e., why a language has the properties it does. It is in this context that considerations of economy, etc. come into play.⁴⁸

In the earliest work on generative grammar, notions of economy and simplicity were considered as part of an evaluation measure, itself part of the theory of grammar, which ranks the potential grammars available (assuming, of course, that the theory allows for more than one grammar that is consistent with the linguistic data). Under the psychological interpretation of grammar, where linguistic theory provides an explanation for language acquisition, an evaluation procedure was thought to be an indispensable element of the theory.

To acquire language, a child must devise a hypothesis compatible with presented data -- he must select from the store of potential grammars a specific one that is appropriate to the data available to him. It is logically possible that the data might be sufficiently rich and the class of potential grammars sufficiently limited so that no more than a single permitted grammar will be compatible with the available data at the moment of successful language acquisition, in our idealized "instantaneous" model (cf. notes 19 and 20). In this case, no evaluation procedure will be necessary as a part of linguistic theory -- that is, as an innate property of an organism or a

device capable of language acquisition. It is rather difficult to imagine how in detail this logical possibility might be realized, and all concrete attempts to formulate an empirically adequate linguistic theory certainly leave ample room for mutually inconsistent grammars, all compatible with primary data of any conceivable sort. All such theories therefore require supplementation by an evaluation measure if language acquisition is to be accounted for and selection of specific grammars is to be justified; and I shall continue to assume tentatively, as heretofore, that this is an empirical fact about the innate human *faculté de langage* and consequently about general linguistic theory as well. (*Aspects*, pp. 36-37)

From this perspective the relationship between linguistic theory and the particular grammars it provides is that of an evaluation procedure. Given a corpus (e.g. the primary linguistic data a child is exposed to) and a set of possible grammars, linguistic theory ranks the grammars and selects the most highly valued one.

As noted in *SS*, “the strongest requirement that could be placed on the relation between a theory of linguistic structure and particular grammars is that the theory must provide a practical and mechanical method for actually constructing the grammar, given a corpus of utterances” (pp. 50-51). A theory of this sort constitutes a discovery procedure for grammars. Early work on generative grammar assumed that linguistic theory would not be able to meet this requirement, and that only the weaker requirement of providing an evaluation procedure for grammars was attainable.⁴⁹

However, with the development of the principles and parameters framework during the past decade and a half⁵⁰ an evaluation measure for grammars has become essentially superfluous. With the reduction of the transformational component to a set of elementary operations that are part of UG and with the elimination of phrase structure rules (see note 45), it appears that linguistic theory allows only a severely limited number of grammars--perhaps only one for each given language. Under the minimalist program of Chomsky 1992 an even stronger assumption is proposed: “there is only one computational system and one lexicon,” apart from a limited kind of variety found in the PF component and in the lexicon (including association of concepts with

phonological matrices (what Chomsky refers to as Saussurean arbitrariness), properties of grammatical formatives like inflection, and detectable lexical properties like the linear orientation of heads and complements (the head parameter)). If this is correct, then current linguistic theory now provides a discovery procedure for grammars and has thereby achieved, in Chomsky's own words, "a scientific advance of the highest importance."⁵¹

References

- Babby, L. (1989). Subjectlessness, External Subcategorization, and the Projection Principle. *Zbornik Matice Srpske Za Filologiju i Lingvistiku*, XXXII/2 (Novi Sad), 7-39.
- Bloomfield, L. (1933). *Language*. New York: Holt.
- Bloomfield, L. (1939). Menomini Morphophonemics. In C. F. Hockett (Eds.), *A Leonard Bloomfield Anthology* Bloomington, IN: Indiana University Press, 1970.
- Bresnan, J. (1972) *Theory of Complementation in English Syntax*. Doctoral dissertation, M.I.T. [published New York: Garland 1979].
- Browning, M. A. (1991). Bounding Conditions on Representation. *Linguistic Inquiry*, 22, 541-562.
- Chomsky, N. (1951). *Morphophonemics of Modern Hebrew*. [published New York: Garland 1979].
- Chomsky, N. (1955-56). *The Logical Structure of Linguistic Theory*. [published New York: Plenum 1975].
- Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.
- Chomsky, N. (1962). A Transformational Approach to Syntax. In A. A. Hill (Eds.), *Proceedings of the Third Texas Conference on Problems of Linguistic Analysis in English, 1958* Austin, TX: University of Texas (reprinted in J.

- Fodor and J. Katz (eds.), (1964), *The Structure of Language: Readings in the Philosophy of Language*, Englewood Cliffs, NJ, Prentice-Hall).
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (1970). Remarks on Nominalizations. In N. Chomsky, *Studies on Semantics in Generative Grammar*, The Hague: Mouton, 1972.
- Chomsky, N. (1973). Conditions on Transformations. In S. R. Anderson & P. Kiparsky (Eds.), *A Festschrift for Morris Halle* New York: Holt, Rinehart and Winston.
- Chomsky, N. (1975a). *Reflections on Language*. New York: Pantheon.
- Chomsky, N. (1975b). “The Amherst Lectures.” Paris: Departement de Recherches Linguistiques, Université de Paris VII: unpublished transcription of 1974 lecture notes.
- Chomsky, N. (1976). Conditions on Rules of Grammar. *Linguistic Analysis*, 2, 303-351.
- Chomsky, N. (1977). On wh-Movement. In P. Culicover, T. Wasow, & A. Akmajian (Eds.), *Formal Syntax* New York: Academic Press.
- Chomsky, N. (1979). *Language and Responsibility*. New York: Pantheon.
- Chomsky, N. (1980). On Binding. *Linguistic Inquiry*, 11, 1-46.
- Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, N. (1982). *The Generative Enterprise. A discussion with Riny Huybregts and Henk van Riemsdijk [1979-1980]*. Dordrecht: Foris.
- Chomsky, N. (1986). *Knowledge of Language*. New York: Praeger.
- Chomsky, N. (1991). Some Notes on the Economy of Derivation and Representation. In R. Freidin (Eds.), *Principles and Parameters in Comparative Grammar* Cambridge, MA: MIT Press.
- Chomsky, N. (1992). A Minimalist Program for Linguistic Theory. *MIT Occasional Papers in Linguistics*, 1.

- Chomsky, N., & Lasnik, H. (1977b). Filters and Control. *Linguistic Inquiry*, 8, 425-504.
- Chomsky, N., & Lasnik, H. (1991). Principles and Parameters Theory. To appear in A. v. S. J. Jacobs W. Stemefeld, and T. Vennemann (Eds.), *Syntax: An International Handbook of Contemporary Research* Berlin: de Gruyter.
- Dresher, E., & Hornstein, N. (1979). Trace Theory and NP Movement Rules. *Linguistic Inquiry*, 10, 65-82.
- Emonds, J. (1970) *Root and Structure Preserving Transformations*. Doctoral dissertation, M.I.T.
- Emonds, J. (1976). *A Transformational Approach to English Syntax*. New York: Academic Press.
- Fiengo, R. (1974) *Semantic Conditions on Surface Structure*. Doctoral dissertation, M.I.T.
- Fiengo, R. (1977). On Trace Theory. *Linguistic Inquiry*, 8, 35-61.
- Freidin, R. (1975). The Analysis of Passives. *Language*, 51, 384-405.
- Freidin, R. (1978). Cyclicity and the Theory of Grammar. *Linguistic Inquiry*, 9, 519-549.
- Freidin, R. (1992). The Principles and Parameters Framework of Generative Grammar. To appear in R. E. Asher (Eds.), *Encyclopedia of Languages and Linguistics* Edinburgh: Pergamon.
- Harris, Z. (1946). From Morpheme to Utterance. *Language*, 22, 161-183 (reprinted in Z. Harris (1981)).
- Harris, Z. (1951). *Methods in Structural Linguistics*. Chicago: University of Chicago Press.
- Harris, Z. (1952a). Co-occurrence and Transformation in Linguistic Structure. *Language*, 33, 283-340 (reprinted in Z. Harris (1981)).

- Harris, Z. (1952b). Discourse Analysis. *Language*, 28, 1-30 (reprinted in Z. Harris (1981)).
- Harris, Z. (1965). Transformational Theory. *Language*, 41, 363-401 (reprinted in Z. Harris (1981)).
- Harris, Z. (1981). *Papers on Syntax*. Dordrecht: D. Reidel.
- Hasegawa, K. (1968). The Passive Construction in English. *Language*, 44, 230-243.
- Hornstein, N. (1977). S and the X' Convention. *Linguistic Analysis*, 3, 137-176.
- Itkonen, E. (1978). *Grammatical Theory and Metascience: a critical investigation into the methodological and philosophical foundations of 'autonomous' linguistics*. Amsterdam: Benjamins.
- Katz, J., & Postal, P. (1964). *An Integrated Theory of Linguistic Descriptions*. Cambridge, MA: M.I.T. Press.
- Kayne, R. (1975). *French Syntax: The Transformational Cycle*. Cambridge, Massachusetts: MIT Press.
- Kuroda, Y. (1988). Harris and the Reality of Language. *Center for Research in Language Newsletter*, 3, 4-12.
- Lasnik, H. (1988). Subjects and the θ -criterion. *Natural Language and Linguistic Theory*, 6, 1-17.
- Lasnik, H., & Saito, M. (1984). On the Nature of Proper Government. *Linguistic Inquiry*, 15, 235-290.
- Lasnik, H., & Saito, M. (1992). *Move α* . Cambridge: MIT Press.
- Lees, R. (1957). Review of Chomsky (1957). *Language*, 33, 375-408 (reprinted in G. Harman (ed.) *On Noam Chomsky*, 2nd ed., University of Massachusetts Press, Amherest, MA, pp. 34-79).
- Lees, R. (1960). *The Grammar of English Nominalizations*. The Hague: Mouton.
- Newmeyer, F. (1986). *Linguistic Theory in America* (2nd ed.). New York: Academic Press.

- Otero, C. P. (1992). The Cognitive Revolution and the Study of Language: Looking Back to See Ahead. In H. Campos & F. Martínez-Gil (Eds.), *Current Studies in Spanish Linguistics* Washington, D. C.: Georgetown University Press.
- Otero, C. P. (in prep.). *Chomsky's Revolution: Cognitivism and Anarchism*. Blackwell.
- Perlmutter, D. (1971). *Deep and Surface Constraints in Syntax*. New York: Holt, Rinehart and Winston.
- Quicoli, A. C. (1976a). Conditions on Clitic Movement in Portuguese. *Linguistic Analysis*, 2, 199-223.
- Quicoli, A. C. (1976b). Conditions on Quantifier Movement in French. *Linguistic Inquiry* 7, 583-607.
- Selkirk, E. (1972) *The Phrase Phonology of English and French*. Doctoral dissertation, MIT [published by Garland, 1980].
- Speas, M. (1990). *Phrase Structure in Natural Language*. Dordrecht: Kluwer.
- Stowell, T. (1981) *Origins of Phrase Structure*. Doctoral dissertation, MIT.
- Vergnaud, J. R. (1977). letter to Chomsky and Lasnik - April 17. In R. Freidin & A. C. Quicoli (Eds.), *Fundamental Readings in Generative Grammar* Cambridge, MA: M.I.T. Press (forthcoming).
- Wasow, T. (1972) *Anaphoric Relations in English*. Doctoral dissertation, MIT
- Wasow, T. (1979). *Anaphora in Generative Grammar*. Ghent: E. Story-Scientia.
[revised version of Wasow 1972].

^NNOTES

*I am indebted to Carlos Otero for discussion of the material covered here and also for comments on an earlier draft, and to Larry Solan for his comments on the paper. For a discussion of Chomsky's theory of grammar in the more

general context of the cognitive revolution, see Otero 1992. See also Newmeyer 1986 for further discussion of the history of generative grammar.

¹ See Harris 1951, Appendix to §20.3 for details.

² This work, submitted in 1951, was based on his undergraduate thesis of 1949. A version of this work was published by Garland in 1979. In particular, the introductions of the 1951 and 1979 versions differ substantially. See below for further discussion, and also Otero (in preparation) for more extensive discussion of the differences between the two versions.

³ *MMH*:1. In a footnote to the quotation he points out that the closed set is not necessarily finite. “Thus the resulting grammar will contain a recursive specification of a denumerable set of sentences.” [*MMH*:67]

⁴ In spite of what he had achieved, Chomsky continued to believe in the importance of taxonomic discovery procedures for linguistic analysis for another two years. It wasn’t until 1953 that he “abandoned any hope of formulating taxonomic ‘discovery procedure’ and turned [his] attention entirely to the problems of generative grammar, in theory and in application. It was at that point that [he] began writing *LSLT*, bringing together and extending the work [he] had begun on various aspects of generative grammar, but now with conviction as well as enthusiasm.” (*LSLT*, p. 33)

⁵ See Chomsky and Lasnik 1991 and Freidin 1992 for discussion.

⁶ It should be noted that the 1979 version published by Garland Press differs from the University of Pennsylvania manuscript. For example, the introductory section of the manuscript version raises issues concerning the relation between logic and linguistics which have been deleted from the published version. See Otero (forthcoming) for a more detailed discussion of the differences between the two versions.

⁷ In the 1975 introduction to *LSLT* Chomsky cites Bloomfield's "Menomini Morphophonemics" (1939) as a potential precursor of generative grammar (one that he was unaware of when he wrote *MMH*)--presumably because the notion of rule ordering is implicit in the work. A comparison of the *MMH* and Bloomfield's paper is instructive. The latter is essentially an informal sketch in contrast to the formal rigor of Chomsky's work where rule ordering is an explicit central topic.

⁸ In the case of phrase structure, Chomsky rules are modeled on Harris's morpheme-to-utterance formulas (see Harris 1946, and chapter 16 of *Methods*).

⁹ However, a detailed account of the early history of transformational theory remains to be written. But see Chomsky 1977/1979, chapter 5 for some helpful discussion.

¹⁰ It is perhaps worth noting here that Chomsky also contributed to this initial work on transformations, as indicated in Harris's first footnote: "It is a pleasure to acknowledge here the cooperation of three men who have collaborated with me in developing the method and in analyzing various texts: Fred Lukoff, Noam Chomsky, and A. F. Brown." And in the first footnote in the published version (1957) of his 1955 Presidential address to the Linguistic Society of America, Harris makes the following comment: "From a time when this work was still at an early stage, Noam Chomsky has been carrying out partly related studies of transformations and their position in linguistic analysis: see his dissertation, *Transformational Analysis* (University of Pennsylvania, 1955); and his paper 'Three Models for the Description of Language', *IRE Transactions on Information Theory*, IT-2, No. 3 (1956), 113-24; now also his book *Syntactic Structures*, The Hague 1957. My many

conversations with Chomsky have sharpened the work presented here, in addition to being a great pleasure in themselves.”

The notion of transformation as equivalence relation explains the otherwise curious usage of the term in *MMH*, where phrase structure rules are referred to as “transformation statements”--see p. 6.

¹¹ See Kuroda 1988 for an attempt to interpret Harris’s transformational theory as a species of realism. However, unless Harris’s thinking underwent some radical change between the late 1940’s and 1965, it seems unlikely that he adopted a realist interpretation of linguistic theory. He clearly never accepted the psychological interpretation of grammar that Chomsky put forth in *Aspects*. Note also that the comparison of Harris and Chomsky in Itkonen 1978, §3.5 fails precisely because the author has missed the fundamental difference vis a vis realism.

¹² See Huybregts and van Riemsdijk 1982:62.

¹³ *LSLT*, p. 35.

¹⁴ Chomsky 1965, p. 58.

¹⁵ Chomsky 1965, p. 58.

¹⁶ See also Chomsky & Lasnik 1992 for additional discussion.

¹⁷ Chomsky 1986, p. 19.

¹⁸ It is worth noting here that the rejection of the psychological interpretation of grammar by the structuralists was based on a perception of the fundamental diversity of human languages--a perception which seems to be wrong, given the current direction of contemporary generative grammar. Thus for example, Bloomfield criticizes Herman Paul’s *Prinzipien der Sprachgeschichte* (1880; fifth edition, 1920) for its “insistence upon ‘psychological’ interpretation.” He claims that Paul’s discussion of language

in terms of mental processes “add nothing to the discussion, but only obscure it.”

Paul and most of his contemporaries dealt only with Indo-European languages, and, what with their neglect of descriptive problems, refused to work with languages whose history was unknown. This limitation cut them off from a knowledge of foreign types of grammatical structure, which would have opened their eyes to the fact that even the fundamental features of Indo-European grammar, such as, especially, the part-of-speech system, are by no means universal in human speech. Believing these features to be universal, they resorted, whenever they dealt with fundamentals, to philosophical and psychological pseudo-explanations. (1933:17)

These criticisms do not carry over to contemporary generative grammar.

¹⁹ In *Syntactic Structures* Chomsky defines language as “a set (finite or infinite) of sentences, each finite in length and constructed out of a finite set of elements” (p. 13). It is this characterization that allows Chomsky to equate human languages with formalized systems of mathematics. Thus, he continues: “All natural languages in their spoken or written form are languages in this sense, since each natural language has a finite number of phonemes (or letters in its alphabet) and each sentence is representable as a finite sequence of these phonemes (or letters), though there are infinitely many sentences. Similarly, the set of ‘sentences’ of some formalized system of mathematics can be considered a language.” Under the psychological interpretation, there is no similarity between a natural language and a formalized system of mathematics.

²⁰ The Auxiliary Phrase need not contain an actual auxiliary verb under Chomsky’s analysis. At deep structure Aux will be at least a tense morpheme or its equivalent (e.g. in imperatives).

²¹ The stipulation was necessary under the SS analysis since there were other rules like Affix Hopping (called the Auxiliary transformation) which had to be designated as obligatory. See below for further discussion of the optional/obligatory distinction and how it was eliminated from the formulation of transformations.

²² However, this analysis turned out to be untenable because there were some verbs (e.g. *know* and *think*) which did not allow manner adverbs freely and yet had passive counterparts. Furthermore, analyzing the passive *by*-phrase as a manner adverb appeared to make a false prediction--that passive predicates couldn't take manner adverbs. To account for the fact that passives could take manner adverbs it would be necessary to have two manner adverb phrases in underlying structure. But that would require the addition of an ad hoc restriction against neither of them being realized as the passive *by*-phrase. See Freidin 1975 for details.

Note also that Chomsky's analysis here is an adaptation of the one proposed in Katz and Postal 1964 which was the first to suggest that abstract markers in underlying structure distinguished one sentence type from another (e.g. question vs. imperative vs. passive).

²³ See Aspects, p. 104. In SS Chomsky addresses the issues of derived constituent structure as follows:

We have not discussed the manner in which transformations impose constituent structure, although we have indicated that this is necessary; in particular, so that transformations can be compounded. One of the general conditions on derived constituent structure will be the following:

(77) If *X* is a *Z* in the phrase structure grammar, and a string *Y* formed by a transformation is of the same structural form as *X*, then *Y* is also a *Z*.

In particular, even when passives are deleted from the kernel we will want to say that the *by*-phrase (as in “the food was eaten – by the man”) is a prepositional phrase (*PP*) in the passive sentence. (77) permits this, since we know from the kernel grammar that *by* + *NP* is a *PP*. (77) is not stated with sufficient accuracy, but it can be elaborated as one of a set of conditions on derived constituent structure. (pp. 73-74)

²⁴ For example, the phrase structure analysis would require that if the passive auxiliary occurred then the passive marker must also occur. Given the *Aspects* analysis of passives, the requirement could not be stated in a non ad hoc fashion. For further discussion of the analysis of the passive auxiliary, see Freidin 1975.

²⁵ It is worth noting here that in *LSLT* the passive transformation is also decomposed into its component parts.

The passive transformation T_p is based on an underlying elementary transformation t_p which is the product of a permutation and a deformation. That is, we form a passive by interchanging the subject and object (by a permutation) and adding *be* + *en* between the auxiliary and the verb, and *by* after the verb. T_p will be applied to any string of the form $NP_1 - VP_A - \langle D \rangle V_T - NP_2$ and will convert it into a string of the form $NP_2 - VP_A - be + en - \langle D \rangle V_T - by + NP_1$. (p. 449)

In *LSLT* (p. 450) the permutation underlying the passive transformation is given as:

$$\pi_p : Y_1 - Y_2 - Y_3 - Y_4 \rightarrow Y_4 - Y_2 - Y_3 - Y_1$$

and the underlying deformation as:

$$\delta_p : Y_1 - Y_2 - Y_3 - Y_4 \rightarrow Y_1 - Y_2 - be + en + Y_3 - by + Y_4$$

The Remarks analysis shows that the permutation itself can be broken down into two separate components, thereby suggesting that such permutations are not elementary operations, but rather the result of two elementary operations.

²⁶ However Hasegawa's interpretation of this result is less conservative than Chomsky's.

Thus, wherever we can extract such cross-transformational elementary operations, it would be much simpler to reinterpret them as transformations (mapping rules) than to set up a far greater number of 'transformations' consisting of various combinations of these elementary operations (e.g. if there are three such elementary operations [that is, substitution, adjunction, and deletion--RF], seven different 'transformations' are theoretically possible).
(p. 235)

Moreover, he goes on to propose that the structural change of a transformation be limited to a single elementary operation as a means of limiting the expressive power of transformations (see §5). This proposal was, as we now know, ahead of its time. Hasegawa offers it as a conjecture. His analysis of passive constructions in itself provides very little motivation for the conjecture and does not discuss the consequences of adopting it other than to point out that the expressive power of transformations would be substantially reduced.

²⁷ Note that in Chomsky 1976 the rule Move NP is taken to be "a natural consequence of assuming the condition of minimal factorization"--which prohibits any transformation whose structural description mentions "two successive categorial terms unless one is satisfied by a factor changed by the rule" (pp. 172-173). However, without general conditions which mediate the behavior of Move NP, it would not be possible to maintain as strong a condition as minimal factorization.

²⁸ See Chomsky 1973:fn.49, 1975a, 1975b for the initial discussions. See also Selkirk 1972, Wasow 1972, 1979, and Fiengo 1974, 1977 for some early extensions and applications of trace theory.

²⁹ In Chomsky 1980 it is noted that traces in syntactic representations result if transformations may not compound elementary operations.

Movement of the category α is assumed to “leave behind” the category [$_{\alpha} e$], in accordance with trace theory. This assumption was implicit in earlier versions of transformational grammar, and becomes explicit when compounding of elementary transformations is forbidden.... (p. 4)

Note that trace theory now motivates Hasegawa’s conjecture as discussed in footnote 26.

³⁰ The PIC is a refinement on the TSC--see Chomsky 1977 for details.

³¹ Chomsky goes on to suggest that “it might be appropriate to give a similar interpretation to the subjacency condition for movement rules.” See Freidin 1978 and Browning 1991 for further discussion.

³² Of course there is a derivation of (12) from (13a) in which the application of Move NP violates the SSC--namely, the one where *that book* is moved into complement subject position first and then *John* is moved long distance into matrix subject position.

³³ Note that the discussion of filtering in generative grammar goes back to *Aspects* where the ‘filtering function’ of transformations is discussed. The first discussion of filters distinct from transformations occurs in David Perlmutter’s 1968 doctoral dissertation, published as Perlmutter 1971. See also Chomsky and Lasnik 1977 for further discussion of filters and their transformational character.

Regarding the redundancy argument concerning the SCC, a caveat is in order. The analysis given here and also in Freidin 1978 assumes that any lexical NP may substitute for a trace. This may be a dubious assumption, depending on how nondistinctness of categories is determined. If having a different index contributes to the determination of distinctness of categories, then two NPs with different indices will be distinct even if all other features are identical. If so, then the Nondistinctness Condition on substitutions will automatically prohibit the substitution of a lexical NP for a trace which bears a different index. (This might provide an explanation for the trace erasure principle of Dresher & Hornstein 1979 (that is, only designated elements (essentially nonreferential *it* and existential *there*) can erase traces) provided such pleonastic elements are nondistinct from a NP trace.) Under this analysis the empirical effects of the SCC follow from the Nondistinctness Condition on substitutions, a basic condition on an elementary transformational operation. (Notice that we obtain the same result with the copying theory of movement operations, whereby a moved phrase leaves behind a copy of itself (which is later deleted in PF)--see Chomsky 1992.) While this interpretation still supports the shift from rules to principles, it does not suggest any shift of focus from derivations to representations since the Nondistinctness Condition is fundamentally a condition on the application of an elementary operation. Furthermore, the analysis demonstrates that strict cyclicity follows from one of the most fundamental properties of movement transformations--hence is as deeply embedded in the theory of transformations as it is possible to be. As for the argument that the TSC and SSC are properly interpreted as conditions on representations rather than conditions on derivations, the redundancy of the SCC no longer applies. Instead, we are left with the argument that trace-

binding is just another instance of anaphor binding and that the binding theory applies to binding relations irrespective of how they are formed.

³⁴ Alternatively, it could have been assigned θ_O in underlying structure in the object position of the verb.

³⁵ See Lasnik 1988 for further discussion.

³⁶ See Freidin 1978 for discussion.

³⁷ The alternative is to assume that once a θ -role has been assigned it is no longer available for other assignments--that is, Unique Assignment is just a consequence of θ -role assignment. In this case (15c) constitutes yet another violation of the Functional Relatedness principle since one or the other position would fail to be assigned a θ -role.

³⁸ Functional Uniqueness and Functional Relatedness were originally proposed in Freidin 1978. The principle of Unique Assignment was originally proposed in Freidin 1975:fn.20 (though without that designation). The three principles are combined with a fourth (requiring every θ -role of a predicate to be assigned) as the θ -Criterion of Chomsky 1981:36.

A reasonable criterion of adequacy for LF is (4) [footnote suppressed--RF]:

(4) Each argument bears one and only one θ -role, and each θ -role is assigned to one and only one argument.

I will refer to (4) as the “ θ -criterion.” An argument is assigned a θ -role by virtue of the θ -position that it or its trace occupies at LF.

For variant formulations see Chomsky 1981:335 and Chomsky 1986.

³⁹ The disjoint distribution of the determiner and possessional NP was expressed in phrase structure grammar using the curly brace notation. Under contemporary analysis, in which there are no phrase structure rules, this

disjoint distribution remains to be explained. Ideally there is some explanation at the level of UG for the nonoccurrence of constructions as in (i).

- (i) a. [_{NP} [_{Det} the] [_{NP} the boy's] book]
b. [_{NP} [_{Det} the] [_{NP} Sam's] book]

⁴⁰ See Bresnan 1972, Emonds 1976, and Hornstein 1977 for analyses which eliminate Agent-postposing in NPs. For example, Hornstein 1977:144f. offers the following empirical argument for base generating the *by*-phrase with its lexical object in nominal constructions like (i).

- (i) John's photograph of Mary by Warhol

Given that the possessional NP *John's* must be base generated in prenominal position, this position cannot be the source for the object of the *by*-phrase. Notice that this argument is not entirely compelling since the underlying source for the possessional NP *John's* could be as postnominal position as in (ii).

- (ii) the photograph of Mary by Warhol of John's

Although this construction may be somewhat awkward, there are certainly other perfectly acceptable nominals in which possessional NP occurs postnominally--e.g. *that photograph by Warhol of John's* or *that photograph of John's*. Under this analysis it would be possible for the NP *Warhol* to occur underlyingly in prenominal position and undergo Agent-postposing.

However, the movement of the possessional NP to prenominal position would involve the substitution of a trace--which would be prohibited by the Nondistinctness Condition on substitutions (as mentioned in footnote 33).

Therefore even if the possessional NP is moved to prenominal position, there is

no motivation for (in fact, motivation against) postulating a movement analysis for the derivation of the *by*-phrase. This argument generalizes to sentential forms as well.

⁴¹ See Chomsky 1955 and Freidin 1975 for some discussion of the adjectival character of passive predicates.

⁴² For further discussion of Case absorption see Jaeggli 1986. See also Babby 1989 for a critical discussion of Case absorption with respect to Slavic languages and a proposal that deals with the crosslinguistic variation in terms of an optional subject parameter.

⁴³ A similar analysis was developed for *wh*-movement in Chomsky 1973, 1975a, 1976, 1977. See Freidin 1978:542 for further discussion.

⁴⁴ Lasnik & Saito 1984 take the formulation one step further as “Affect α ” which includes deletions. See also Lasnik & Saito 1992.

⁴⁵ A similar sort of reduction has been proposed for phrase structure rules. Because phrase structure rules tend to be language specific stipulations, whereas the concept of phrasal projection and the constituent schema for projections are part of UG, it is generally assumed that phrase structure rules can be eliminated entirely. Instead the principles of X-bar theory and other principles of UG--in particular, Case theory and θ -theory--predict the details of phrase structure. The general idea is that whatever is stipulated in phrase structure rules is redundant because this follows from various parts of the UG in conjunction with the lexical properties of words in a language. See Stowell 1981 and Speas 1990, among others.

⁴⁶ See Chomsky 1991 for a detailed discussion.

⁴⁷ However, a more bizarre construction like (i) might provide independent motivation for a functional uniqueness principle.

- (i) a. np believed [$_{\alpha}$ it to have been promoted Henry]
 b. *Henry_i believed [$_{\alpha}$ it to have been promoted e_i]

Given (i.a) as the underlying structure of (i.b), where np indicates an empty subject position, the last resort condition does not block the movement from complement object position to matrix subject position. The underlying complement object *Henry* is Caseless and therefore must move to a Case-marked position. Since the complement subject position is filled with pleonastic *it*, the NP *Henry* must move into the empty matrix subject position to receive Case. In this position it also receives an additional θ -role, in violation of Functional Uniqueness. The validity of this analysis depends on the presence of pleonastic elements in underlying structure, which is not obvious. In fact given the principle of Full Interpretation, pleonastic elements like nonreferential *it* should not occur in LF representations since they receive no interpretation at LF. We might therefore assume that such elements are inserted at PF after movement operations (governed by the last resort condition) have applied. Under this analysis, (i) does not arise.

⁴⁸ It is worth pointing out here that in *MMH* Chomsky's notion of simplicity bears some similarity to the more current discussions of economy.

For the formulation of any relative precise notion of simplicity, it is necessary that the general structure of the grammar be more or less fixed, as well as the notations by means of which it is constructed. We want the notion of simplicity to be broad enough to comprehend all those aspects of simplicity of grammar which enter into consideration when linguistic elements are set up. Thus we want the reduction of the number of elements and statements, any generalizations, and, to generalize the notion of generalization itself, any similarity in the form of non-identical statements, to increase the total simplicity of the grammar. As a first approximation to the notion of

simplicity, we will here consider shortness of grammar as a measure of simplicity, and will use such notations as will permit similar statements to be coalesced. (p. 5)

To avoid circularity, the notation must be fixed in advance and neutral to any particular grammar.

Given the fixed notation, the criteria of simplicity governing the ordering of statements are as follows: that the shorter grammar is the simpler, and that among equally short grammars, the simplest is that in which the average length of derivation of sentences is least. (p. 6)

In current work, the “shortness” of grammars and of derivations is driven by substantive principles of UG.

⁴⁹ See Otero 1992:fn.83 for some additional illuminating discussion.

⁵⁰ See Chomsky and Lasnik 1991 and Freidin 1992 for discussion.

⁵¹ Cited in Otero 1992:fn.83. It is worth noting in this regard that Chomsky considers the discussion of discovery procedures to be the major contribution of structural linguistics, among its other significant contributions.

Structural linguistics has enormously broadened the scope of information available to us and has extended immeasurably the reliability of such data. It has shown that there are structural relations in language that can be studied abstractly. It has raised the precision of discourse about language to entirely new levels. But I think that its major contribution may prove to be one for which, paradoxically, it has been very severely criticized. I refer to the careful and serious attempt to construct “discovery procedures,” those techniques of segmentation and classification to which Saussure referred. This attempt was a failure--I think that is now generally understood. It was a failure because such techniques are at best limited to the phenomena of surface structure and cannot, therefore, reveal the mechanisms that underlie the creative aspect of language use and the expression of semantic content. But what remains of fundamental importance is that this attempt was directed to the basic

question in the study of language, which was for the first time formulated in a clear and intelligible way. (*Language and Mind*, p.22)