ASPECTS OF THE THEORY OF PHRASE STRUCTURE

HOWARD LASNIK
University of Maryland

1 Generalized Transformations vs. Recursion in the Base

There is a certain irony implicit in the Chomsky (1957) argument for the necessity of Σ, F (context free rewriting) grammars. Chomsky shows straightforwardly that formal languages with unbounded mirror image properties go beyond the bounds of finite state Markov description. Then, upon suggesting that human languages have just such properties, he proposes Σ, F grammars as the necessary extension, since such grammars are admirably well suited to characterizing mirror image languages. The irony resides in the fact that the theory of human language grammar that Chomsky then presents, following the one articulated in great detail in Chomsky (1955), restricts the power of the Σ, F module in precisely such a way that it cannot handle the phenomena at issue. In particular, the theory explicitly and completely disallows recursion in this module (pp. 517-519). In this model, the infinitude of human languages is the responsibility of generalized transformations - operations melding separate phrase markers together into one phrase marker.

Chomsky (1965) resolves this irony by removing the restriction on the Σ, F module, and thus allowing recursion 'in the base.' Though Chomsky (1955) had indicated that the restriction constituted a simplification, he didn't actually offer any arguments to that effect. One might actually argue that removing this restriction is a simplification. After all, it seems to be a stipulation. Further, it is not trivial to determine whether there is recursion in the base. Certainly, the determination is trivial if there is a rule like (1):

(1)   A → BA

But recursion might involve a pair of rules rather than any one rule:

(2)   A → BC
        C → DA

Or a trio:

(3)   A → BC
        C → DE
        E → FA

In fact, there is no limit on how large the minimal group of rules might be that yield recursion.

Chomsky (1965) offered two major arguments for the revised approach (a major part of what he dubbed the 'standard theory'). First, Chomsky claimed that the theory of transformational grammar is simplified by this change, the simplification being that the notion "generalized transformation" is eliminated entirely, at no apparent cost. Thus, in place of three syntactic operations - Σ, F rules, singular transformations, and generalized transformations - we have just the first two. Further, the construct "Transformation-marker" is eliminated, as its major work was to show exactly how the separate trees combine into one, but now that is transparently represented in the initial phrase marker, the 'deep structure.'

Chomsky's second argument is even more interesting. He argues that while there is extensive ordering among singular transformations (situations where a derivation produces an unacceptable sentence if two transformations are applied in reverse order), "there are no known cases of ordering among generalized embedding transformations although such ordering is permitted by the theory of Transformation-markers" (1965: 133). Further, while there are many cases of singular transformations that must apply to a constituent sentence before it is embedded, or that must apply to a 'matrix' sentence after another sentence is embedded in it, "there are no really convincing cases of singular transformations that must apply to a matrix sentence before a sentence transform is embedded in it…"

In both frameworks, the set of singular transformations was seen as a linear sequence: an ordered list. Given the *Aspects* modification, this list of rules is claimed to apply 'cyclically,' first operating on the most deeply embedded clause, then the next most deeply embedded, and so on, working 'up the tree' until they apply on the highest clause, the entire generalized phrase marker. Thus, singular transformations apply to constituent sentences 'before' they are embedded, and to matrix sentences 'after' embedding has taken place. "The ordering possibilities that are permitted by the theory of Transformation-markers but apparently never put to use are now excluded in principle" (1965: 135).

It is important to note that within the Minimalist program, Chomsky argues against a level of deep structure, and Chomsky (1993) argues for generalized transformations as the sole structure-creating operation, responsible even for the structure of single clause sentences. Furthermore, singular transformational operations are interspersed with these generalized transformational operations, again roughly as in the much earlier model. But what of the powerful Chomsky (1965) argument against such a model, that it allowed derivations that never actually occur in human languages. The model with recursion in the base excluded those unwanted derivations. However, on closer inspection, it was not actually elimination of generalized transformations that had this limiting effect. Rather, it was the stipulation that transformations operate strictly cyclically, starting on the most deeply embedded clause and proceeding monotonically up the tree. Chomsky (1993) observes that a condition with the same effect can be imposed on the operation of generalized transformations and their interaction with singular transformations. This condition, often called the 'extension condition,' simply requires that a transformational operation 'extends' the tree upwards. This guarantees the same sort of

monotonic derivations as those permitted by Chomsky (1965).

The one remaining *Aspects* argument against generalized transformations can also be straightforwardly addressed. Chomsky had argued that eliminating generalized transformations yields a simplified theory, with one class of complex operations jettisoned in favor of an expanded role for a component that was independently necessary, the phrase structure rule component. This was a very good argument. But now that 1965 argument can be reversed on itself: Eliminate phrase structure rules!

2 Redundancy Between Subcategorization and Phrase Structure (PS) Rules

Aspects also contains the germ of another argument for the elimination of phrase structure rules. Chomsky (p. 96) discusses the phrase structure rules for VP, such as the following, with a representative VP using each rule:

(4) (i)   VP → V   (*elapse*)
     (ii) VP → V NP   (*bring the book*)
     (iii) VP → V NP *that*-S [i.e., CP] (*persuade John that there was no hope*)
     (iv) VP → V PP   (*decide on a new course of action*)
     etc.

Now among the many major lasting innovations in *Aspects* was the introduction of the lexicon into syntactic theory. (In the *LSLT* model, the source of words and grammatical morphemes was phrase structure rules.) Lexical entries included all that is particular and idiosyncratic about specific lexical items. Among that information was a subcategorization frame indicating for each verb, for example, exactly what kind of VP it can be inserted into. For example, the subcategorization frames for *elapse* and *persuade* were essentially as follows:

(5) *elapse*   [__]
     persuade   [__ NP CP]

After introducing this technology, Chomsky, commenting on (4), observes (p. 96) "Corresponding to each such string dominated by VP, there is a strict subcategorization of Verbs." This is as things must be, of course, since what would it mean for a particular language to have a particular VP rule if that language had no verb that could ever be inserted into that VP? The converse is equally true. How could a particular language have a particular verb if there were no phrase structure rule to create the VP structure required by that verb? What Chomsky doesn't point out is that there is thus total redundancy between subcategorization and PS rules at (what became) the V-bar level. This redundancy strongly suggests that one or the other should be eliminated. Further, since lexical information must be stated somewhere, it is PS rules that should go. At the time, this seemed improbable since phrase structure was so closely tied to phrase structure rules and phrase structure derivations. But eventually proposals began to appear for phrase structure without phrase structure rules or derivations. For example, McCawley (1968) attributes to Richard Stanley a theory of phrase structure in which the notion of 'derivation' is dispensed with: the base component is simply a set of node admissibility conditions. Lasnik and

Kupin (1977) offer a set-theoretic analogue of the Stanley/McCawley graph-theoretic proposal. For Lasnik and Kupin, any set of mono-strings (strings containing at most one non-terminal element) is a (reduced) phrase-marker, as long as it satisfies certain representational conditions on its completeness, consistency, etc. And just a bit later, in very influential work, Stowell (1981) argued extensively for the superfluity of phrase structure rules in the creation of phrase structure representations.

3 Linear Order in the Phrase Structure Component?

Chomsky observes in *Aspects* (p. 123) that the PS rules carry out two separate functions: They define hierarchical relations (crucial to the specification of grammatical relations), and they determine the linear ordering of elements in deep structure. He then suggests that "At least the first of these functions appears to be carried out in a very general and perhaps universal way by these rules." Referencing Curry (1961) and Saumjan and Soboleva (1963), he then makes the following striking comment:

> "It has been suggested several times that these two functions of the categorial component be more sharply separated, and that the second, perhaps, be eliminated completely."

So, an old concern but also quite a modern one. However Chomsky's position at the time was quite different from what it became in the 1990s. First, he summarizes the Curry and Saumjan-Soboleva position:

> "They propose, in essence, that in place of such rules as (69), the categorial component should contain the corresponding rules (70), where the element on the right is a set rather than a string:
>
> (69)   S $\rightarrow$ NP $\frown$ VP
>   VP $\rightarrow$ V $\frown$ NP
> (70)   S $\rightarrow$ {NP, VP}
>   VP $\rightarrow$ {V, NP}
>
> In (70), no order is assigned to the elements on the right-hand side of the rule; thus {NP,VP} = {VP, NP}, although NP $\frown$ VP $\neq$ VP $\frown$ NP. The rules (70) can be used to define grammatical relations in exactly the way indicated for the rules (69). The rules (69) convey more information than the corresponding rules (70), since they not only define an abstract system of grammatical relations but also assign an abstract underlying order to the elements." (p. 124)

Chomsky then reports that

> "Proponents of set-systems such as (70) have argued that such systems are more 'abstract' than concatenation-systems such as (69), and can lead to a study of grammatical relations that is independent of order, this being a phenomenon that belongs only to surface structure."

That sounds very modern indeed. But interestingly, Chomsky proceeds to strongly reject that argument:

> "The greater abstractness of set-systems, so far as grammatical relations are concerned, is a myth. Thus the grammatical relations defined by (70) are neither more nor less 'abstract' or 'order-independent' than those defined by (69); in fact, the systems of grammatical relations defined in the two cases are identical. A priori, there is no way of determining which theory is correct; it is an entirely empirical question, and the evidence presently available is overwhelmingly in favor of concatenation-systems over set-systems, for the theory of the categorial component. In fact, no proponent of a set-system has given any indication of how the abstract underlying unordered structures are converted into actual strings with surface structures. Hence, the problem of giving empirical support to this theory has not yet been faced." (pp. 124-125)

This was evidently true at the time, and for quite a while after, but decades later an intriguing proposal did appear precisely addressing the question of "how the abstract underlying unordered structures are converted into actual strings." Based on the important insights of Kayne (1994), Chomsky (1995) argues that syntactic structures are, indeed, unordered and that the observed linear ordering arises via Chomsky's version of Kayne's Linear Correspondence Axiom, which in the mapping to PF maps asymmetric c-command into linear precedence. Following a tradition advanced within generative grammar by Reinhart (1976), Chomsky indicates that "There is no clear evidence that order plays a role at LF or in the [transformational] computation" (p. 334).

4 Conclusion

As with so many topics in syntactic theory, and with so many of his writings, Chomsky in *Aspects* set the agenda for generations of research. The lexicon, the syntactic cycle, the status of linear order are as important now as when they were first discussed. In all of these cases, the perspective has been altered, occasionally even reversed, by later developments and discoveries. But it is fair to say that it was the arguments and analyses of *Aspects* that made many of these later advances possible.

References

Chomsky, Noam. 1955. *The logical structure of linguistic theory*. Ms. Harvard University, Cambridge, Mass. and MIT, Cambridge, Mass. [Revised 1956 version published in part by Plenum, New York, 1975; University of Chicago Press, Chicago, 1985]

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, eds. Kenneth Hale and Samuel J. Keyser, 1-52. Cambridge, Mass.: MIT Press. [Reprinted in Noam Chomsky, *The minimalist program*, 167-217. Cambridge, Mass.: MIT Press, 1995]

Chomsky, Noam. 1995. Categories and transformations. In *The minimalist program*, 219-394. Cambridge, Mass.: MIT Press.

Curry, Haskell Brooks. 1961. Some logical aspects of grammatical structure. In *Structure of language and Its mathematical aspects, Proceedings of the Twelfth Symposium in Applied Mathematics*, ed. Roman Jakobson, 56-68. Providence, RI: American Mathematical Society.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.

Lasnik, Howard and Joseph J. Kupin. 1977. A restrictive theory of transformational grammar. *Theoretical Linguistics* 4: 173-196. [Reprinted in Howard Lasnik, *Essays on restrictiveness and learnablity*, 17-41. Dordrecht: Kluwer, 1990]

McCawley, James. 1968. Concerning the base component of a transformational grammar. *Foundations of Language* 4: 243-269.

Reinhart, Tanya. 1976. *The syntactic domain of anaphora*. Doctoral dissertation, MIT.

Saumjan, Sebastian and Polina Arkad'evna Soboleva. 1963. *Applikativnaja porozdajuscaja model' i iscislenie transformacij v russkom jazyke*. Moscow: Izdatel'stvo Akademii Nauk SSSR.

Stowell, Tim. 1981. *Origins of phrase structure*. Doctoral dissertation, MIT.