

***The notion of derivations in
linguistics:
Syntax***

Howard Lasnik
U. Of Maryland
(MIT 1972)

The LSLT model:

Phrase structure component (context free; non-recursive); includes lexical insertion. Creates **P-markers** (which are set theoretic objects).

P(hrase)-markers:

Given a particular ΣF (PS) grammar and a particular terminal string (i.e., string of terminal symbols) (TS):

- a) Construct **all** of the equivalent PS derivations of TS.
- b) Collect all of the lines occurring in any of those equivalent derivations into a set. This set is the phrase marker (PM), a complete representation of the PS of TS, as far as the grammar is concerned.

(Two PS derivations are equivalent if and only if they involve the same rewrite rules the same number of times, but not necessarily in the same order.)

An artificial illustrative example:

Σ : S
F: S \rightarrow NP VP
NP \rightarrow he, Mary
VP \rightarrow V NP
V \rightarrow saw

Derivation 1: Line 1: S
Line 2: NP VP
Line 3: he VP
Line 4: he V NP
Line 5: he saw NP
Line 6: he saw Mary

Derivation 2: Line 1: S
Line 2: NP VP
Line 3: he VP
Line 4: he V NP
Line 5: he V Mary
Line 6: he saw Mary

Derivation 3: Line 1: S
Line 2: NP VP
Line 3: NP V NP
Line 4: he V NP
Line 5: he V Mary
Line 6: he saw Mary

Derivation 4: Line 1: S
Line 2: NP VP
Line 3: NP V NP
Line 4: he V NP
Line 5: he saw NP
Line 6: he saw Mary

Derivation 5: Line 1: S
Line 2: NP VP
Line 3: NP V NP
Line 4: NP saw NP
Line 5: he saw NP
Line 6: he saw Mary

Derivation 5: Line 1: S
Line 2: NP VP
Line 3: NP V NP
Line 4: NP saw NP
Line 5: NP saw Mary
Line 6: he saw Mary

PM =
{S, NP VP, he VP, he V NP, he V Mary,
he saw NP, NP V NP, NP saw NP,
NP saw Mary, NP V Mary, he saw Mary}

A **P**-marker must indicate what substrings of the terminal string are constituents, and for the ones that are, what the 'labels' of the constituents are. In LSLT, the fundamental predicate characterizing these relations is the *is a* relation.

In our simplified ex. the following are the **is a** relations:

he is an NP *Mary is an NP*
saw is a V *saw Mary is a VP*
he saw Mary is an S

By comparing certain members of the PM with the terminal string one by one, one can compute all the 'is a' relations in the following fashion.

Compare 'he saw Mary' with 'he VP':

he saw Mary

he VP

From this we deduce that 'saw Mary' is a VP. Now, compare 'he saw Mary' with 'he V Mary'

he saw Mary

he V Mary

This determines that 'saw' is a V. And so on.

In general, no one derivation will have all of the strings necessary for this algorithm to yield all of the is a relations. That's why LSLT specified that the PM contain all the strings in all of the derivations. In this regard, the LSLT theory of phrase structure was intensely derivational.

We don't need the entire P-marker to determine the 'isA' relations. Some members of the set, the **'monostings'**, consist of exactly one non-terminal symbol and any number of terminal symbols. Those are *S*, *NP VP*, *he VP*, *he V NP*, *he V Mary*, *he saw NP*, *NP V NP*, *NP saw NP*, *NP saw Mary*, *NP V Mary*, *he saw Mary*. (They are red and italicized below):

{*S*, *NP VP*, *he VP*, *he V NP*, *he V Mary*, *he saw NP*, *NP V NP*, *NP saw NP*, *NP saw Mary*, *NP V Mary*, *he saw Mary*}

For this reason, Lasnik and Kupin (1977) proposed a revised construct to represent phrase structure: the **Reduced Phrase Marker**, the set consisting just of the terminal string and all the monostings. [There was also a more significant change – to a **non**-derivational approach to phrase structure. More on this momentarily.]

McCawley (1968) argued against the LSLT derivational theory of phrase structure. One of his arguments was, roughly, that we could simplify the theory by creating constituent structure trees directly, rather than, he implies, by first going through all the equivalent derivations, etc.

However, the trees in Chomsky's presentations were almost entirely for expository purposes. Chomsky's theory was set theoretic, not graph theoretic, so no conversion to trees was necessary, or even relevant.

Another argument was more telling, and directly relevant to my concerns today: Chomsky's phrase structure rules were **ordered**, hence demanding derivations. But McCawley showed that the ordering wasn't necessary. (The ordering was invariably to capture contextual restrictions on particular lexical items, especially verbs. But once the model added a lexicon, as in Chomsky (1965), those ordering restrictions on PS rules became irrelevant.)

McCawley attributed to Richard Stanley a tree based theory that lacked derivations. Instead, it had “*node admissibility conditions*”.

But nothing in principle precludes a set-theoretic model without derivations. Lasnik and Kupin (1977) presented just such a model. In that model, PMs are simply sets of (mono)strings that satisfy certain conditions.

The transformational component of the Chomskian model is, of course, highly derivational. It includes an initial PM (or set thereof).

The **source** of the initial PM(s), derivational or not, is irrelevant. This is one among many Markovian aspects of transformational derivations.

In the original LSLT framework, the phrase structure derivations produced only simple mono-clausal structures, which could then be merged together by **generalized transformations** conjoining two (or possibly more) separate structures, or embedding one into another. The generalized transformations constituted the recursive component of the grammar.

The GTs could interact with the **singular T_s** (those operating on a single 'tree' rather than on a 'forest') in various ways. A major level of representation in this model was the **T-marker**, the record of all the T_s that applied and exactly how they interacted.

Chomsky (1965) innovation: Allow the phrase structure rule component itself to have a recursive character.

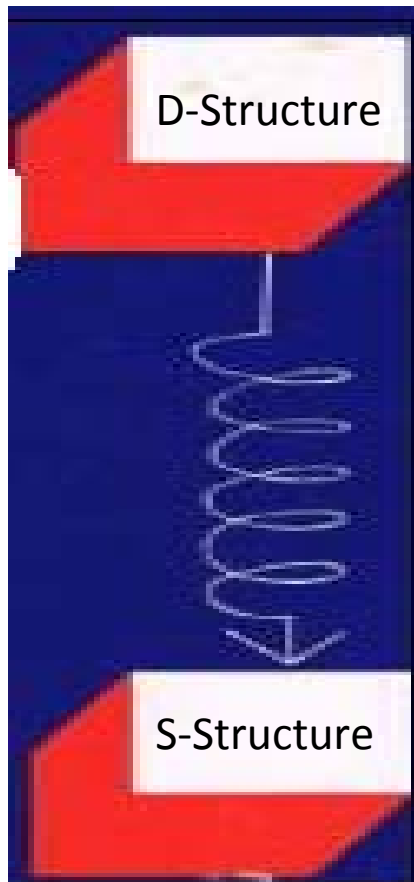
Chomsky's major arguments for this new organization were that it resulted in a simpler overall theory (eliminating one kind of operation), and at the same time it explained the absence of certain kinds of derivations that seemed not to be needed. The second of these points is our concern here.

Following Fillmore (1963), Chomsky argued that while there is extensive ordering among singular transformations "... there are no known cases of ordering among generalized transformation although such ordering is permitted by the theory of Transformation-markers." Chomsky (1965, p.133)

Further, Chomsky claimed that 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...."

The Aspects theory, with its elimination of generalized transformations, was a response to these concerns. As in the LSLT theory, there is still extensive ordering among singular transformations. In both frameworks, the set of singular transformations was seen as a linear sequence: an ordered list. (And, to the extent that the orderings were crucial, there was motivation for a derivational model, rather analogous to the opacity arguments for phonological derivations.)

But, as they say in the infomercials, there's more:



Given the full Aspects modification, this list of rules applies **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 P-marker

Thus, singular transformations apply to constituent sentences as if 'before' they are embedded, and to matrix sentences as if '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." Chomsky (1965, p.135)

In passing, I note that the cyclic principle was one of two major syntactic innovations of the mid 1960s borrowed from phonology.

The principle was first formulated in Chomsky et al. (1956), and applied in the phonological analysis of a variety of languages from the early 1960s on, most notably in SPE (Chomsky and Halle (1968)).

"... it is natural to suppose that in general the phonetic shape of a complex unit ... will be determined by the inherent properties of its parts and the manner in which these parts are combined, and that similar rules will apply to units of different levels of complexity. These observations suggest a general principle for the application of rules of the phonological component, namely, what we will call the principle of the **'transformational cycle.'**" [SPE p.15]

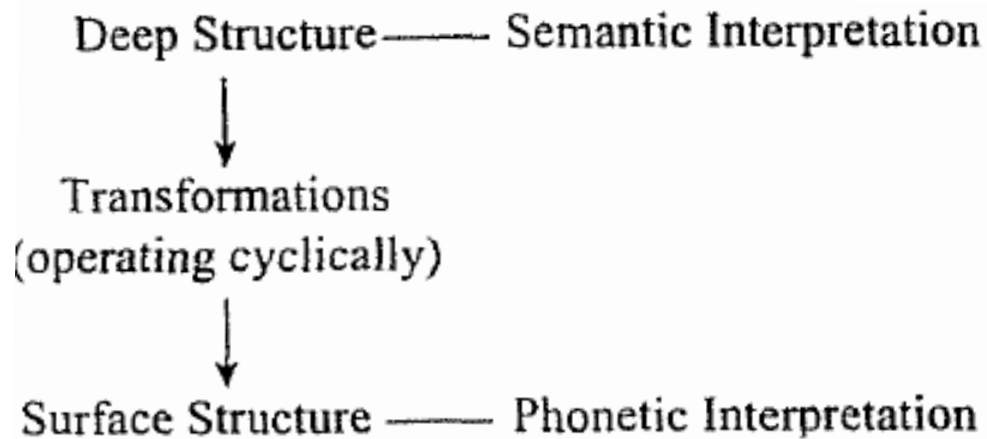
Interestingly, it was almost three decades before it became clear that the absence of certain kinds of derivations was not a real argument against generalized transformations.

For it was not recursion in the base that excluded the unwanted derivations: It was the cyclic principle. And there are a variety of ways that a version of the cyclic principle could be grafted onto a theory with generalized transformations, as work of the last decade and a half has amply demonstrated.

Derivations and interfaces

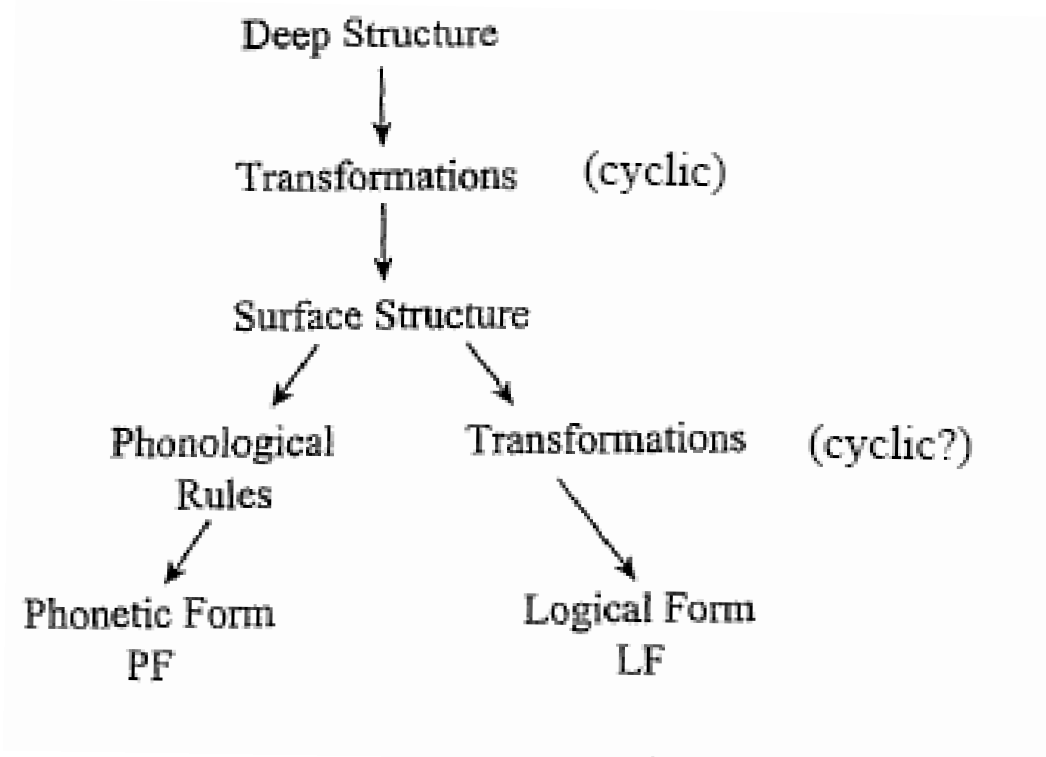
In LSLT the T-marker is the interface with semantic interpretation, and the final derived syntactic representation is the interface with morphophonemics.

In Aspects, we have:



Already in Aspects, it was pointed out that some aspects of semantic interpretation (basically all except thematic relations) depend on derived structure. With the introduction of trace theory in the early 1970's, D-structure was no longer crucial even for thematic relations.

That led to the 'Extended Standard Theory' and GB model:



BUT, there had already been suggestions in the late 1960's and very early 1970's for links to semantic and phonological interpretation not just at levels but internal to the derivation.

Particularly important were the work of Bresnan (1971) wrt phonology (specifically sentence intonation) and Jackendoff (1969, 1972) wrt semantics (specifically anaphora).

For both, the derivation-internal points were end of cycle structures.

The picture then would look like:



(Some) semantic interpretation



(Some) semantic and phonological interpretation



(Some) semantic and phonological interpretation



(Some) semantic and phonological interpretation



(Some) semantic and phonological interpretation



(Some) semantic and phonological interpretation



(Some) semantic and phonological interpretation

Going further back in time, recall the Aspects arguments against GTs.

The one based on non-occurring derivations turned out to miss the mark, as it was the cyclic principle, not recursion in the base that blocked those derivations.

The simplicity argument took for granted that PS rules exist, and that the Aspects modification thus reduced the class of basic operations from 3 to 2 (just singularly Ts and PS rules).

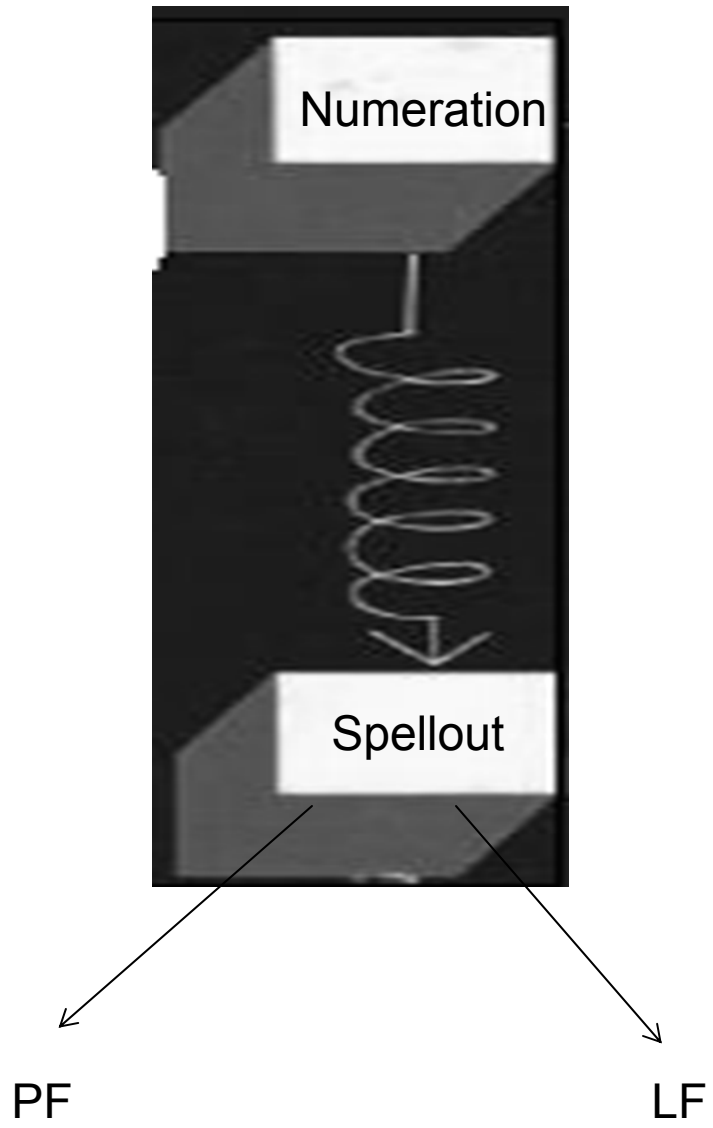
But if there are no PS rules (as suggested by McCawley and again by Lasnik and Kupin) ...

An early, or middle-aged, Minimalist model of structure building where **everything** is done by transformations.

GTs don't just merge clausal structures with clausal structures, but even lexical items with lexical items. The derivation begins with a selection of lexical items (the '[numeration](#)'), which are then combined; these combinations are then combined with other lexical items or other combinations, and so on.

Meanwhile, interspersed with these instances of 'external merge' we also have 'internal merge', the classic singulary movement transformation. All these operations are constrained by some version of cyclicity (for example, always merge at the 'root' of the 'tree').

Eventually (the point of '[spellout](#)'), the derivation branches to PF and LF.



There are, though, a number of reasons for thinking that Bresnan and Jackendoff were fundamentally correct about the derivational nature of the interface relations. Among many examples, I could mention Barss (1986) on 'reconstruction' into intermediate positions; McCoskey (1991) and Torrego (1984) on intermediate footprints of wh-movement; Uriagereka (1999) on deriving certain island constraints; Fox and Pesetsky (2003) on forcing successive cyclic movement; and Merchant (2001), based on Perlmutter (1971), on distinctions between representational and derivational constraints.

Grafting Bresnan's and Jackendoff's ideas onto the early minimalist model, we wind up with a 'Multiple Spellout' (Uriagereka (1999)) or Derivation by Phase (Chomsky (2001)) picture.



Semantic and phonological interpretation
Semantic and phonological interpretation
Semantic and phonological interpretation
Semantic and phonological interpretation
Semantic and phonological interpretation

What I find fascinating about all these developments is that the same ideas (and sometimes even arguments) keep coming back in slightly revised forms, and combining in slightly different ways. Does this mean these ideas must be right? Or is it just a poverty of imagination situation? I hope I'll last long enough to find out. Maybe at the 100th anniversary?

- Barss, Andrew. 1986. *Chains and anaphoric dependence: On reconstruction and its implications*. Doctoral dissertation, MIT, Cambridge, Mass.
- Bresnan, Joan W. 1971. Sentence stress and syntactic transformations. *Language* 47: 257-281.
- 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. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT press.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1-52. Cambridge, Mass.: MIT Press.
- Chomsky, Noam and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.
- Chomsky, Noam, Morris Halle and Fred Lukoff. 1956. On accent and juncture in English. In *For Roman Jakobson*, ed. M. Halle, H. Lunt, and H. MacLean, 65-80. The Hague: Mouton & Co.
- Fillmore, Charles J. 1963. The position of embedding transformations in a grammar. *Word* 19: 208-231.
- Fox, Danny and David Pesetsky. 2003. Cyclic linearization and the typology of movement. Ms. MIT, Cambridge, Mass.
- Jackendoff, Ray. 1969. *Some rules of semantic interpretation for English*. Doctoral dissertation, MIT, Cambridge, Mass.
- Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambridge, Mass.: MIT Press.
- Lasnik, Howard and Joseph J. Kupin. 1977. A restrictive theory of transformational grammar. *Theoretical Linguistics* 4: 173-196. [Reprinted in *Essays on restrictiveness and learnability*, Howard Lasnik, 17-41. Dordrecht: Kluwer, 1990]
- McCawley, James. 1968. Concerning the base component of a transformational grammar. *Foundations of Language* 4: 243-269.
- McCloskey, James. 1991. Resumptive pronouns, A'-binding, and levels of representation in Irish. In *Syntax and Semantics 23: The Syntax of the Modern Celtic Languages*, ed. Randall Hendrick, 199-248. New York: Academic Press.
- Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.
- Perlmutter, David. 1971. *Deep and surface constraints in syntax*. New York: Holt, Rinehart and Winston.
- Torrego, Esther. 1984. On inversion in Spanish and some of its effects. *Linguistic Inquiry* 15: 103-129.
- Uriagereka, Juan. 1999. Multiple spell-out. In *Working minimalism*, ed. Samuel David Epstein and Norbert Hornstein, 251-282. Cambridge, Mass.: MIT Press.