Contents

C-Movement in Minimalist Derivation: Recursion, Parsimony and Distal Governance in the Competence Grammar of $\Theta_{English}$

Austin K. Faulkner

18 October 2007

The essential spirit of Minimalism is to reduce the theoretical apparatus which we use to describe syntactic structure to a minimum (Radford 2004:73).

The following addresses some of the core aspects of what is now know as the Minimalist Program (**MP**), a development out of the "Principles & Parameters" (**P&P**) project of the '80s and a continuation of the ever evolving Generative Grammar paradigm (**GG**) of Universal Grammar (**UG**) reaching as far back as the early '60s. Although, as we shall see, both the P&P approach as well as the MP development out of P&P maintain that the Generative Grammar studies being done in the '60s and '70s were not only misguided in fact but, as it turns out, were misguided *in principle* (Lasnik in (Chomsky 1997:1-5)). The revised version of the principles of generation are now thought to be even more *minimal*, and for that, some speculate, all the more *optimal*.[1]

This paper is divided into three parts, each appropriately introducing the other. The first section attempts to characterize the locus of the discussion. More precisely, it is the task of the first section to clarify the distinction between a theory of Language versus a theory of a (particular) language. With this distinction in mind, we'll discuss how in the matrix of generative interfaces Deep-Structure (**D-Structure**) directionally maps into Surface-Structure (**S-Structure**) and its Phonological (**PF**) and Logical (**LF**) counterpart interfaces.[2] This, in turn, raises interesting questions about the exact nature of recursion in natural language. Section one stages discussion taken up in section two, wherein will be discussed (1) the basic set of *principles* and *parameter settings* placed on the particular language of English and (2) the *minimalist constraints* placed on these parametrized principles. The third and final section is crucial in that it takes into account how these principles generate the phenomena of Constituent Movement (**C-Movement**), where Constituent Movement is usually taken to include Auxiliary (**A-**) and Head Movements (**H-Movements**) of various sorts, and Do- and Wh-Support. In the tradition, these phenomena are simply classed as Transformations at S-Structure; moreover, recent MP research indicates that, in fact, all C-Movement may be characterized reducibly as H-Movement (Radford 2004:Chapters 3 and 5).

1 Recursion, GG, and the Matrix of Competence Interfaces

1.1 Operations over a Closed Set

At bottom, (natural) Language is said to be generative over a closed set. Not only is Language said to be generative; moreover, it is said to be *recursively* generative (Pinker 1999:8-9; Radford 2004:69); and, arguably, it is this added property that gives Language its splendor. As Steven Pinker writes in his *Words and Rules*: "A recursive grammar can generate sentences of any length, and thus can generate an infinite number of sentences. So a human being possessing a recursive grammar can express or understand an

[1] The theory of *optimality* is not something pursued here; however, it's worth mentioning in that speculation on the "optimality" of language—a satilite pursuit tangentially related to the Minimalist Program—suggests something of interest when considering the character and ultimate ambitions of MP (CF. Chomsky 1997).

[2] PF and LF traditionally are known as "Phonological Form" and "Logical Form" (Cf.Chomsky 1997; Pinker 1999:23)

infinite number of distinct thoughts, limited in practice only by stamina and mortality." And while, to be sure, the infinite extension of human thought by means of the iterative use of grammatical constructions is something to marvel at, what interests us here is not that, but rather something else. We are here interested in those components which work together to *enable* this ability or *competence*.[3] In this connection, Noam Chomsky writes in his *Aspects of a Theory of Syntax* of how "a real understanding of how a language can (in Humboldt's words) 'make infinite use of finite means' has developed [. . .] in the course of studies in the foundations of mathematics" (1965:8). With this connection in mind, briefly, we may resort to some of the basic concepts of Recursion Theory in mathematics to get an idea of GG's inception. Concepts illustrated below suggest how, *in principio*, one might begin to conceive of Language as an algorithm of sorts.

Definition 1 (Recursion) *Recursion occurs when an infinite set is specified in terms of a proper subset.*[4]

Taking this definition, let's look at examples of recursion. Below, there are two, one (1) adapted from number theory and the other (2) adapted from the theory of formal languages. In (1), where D is understood to be the closed set of numerals $[0, 1, 2, \ldots, n, \ldots, 7, 8, 9]$ and x is understood to range over that set D, we find that, where $D \subset SN$ ("D is a proper subset of SN"), D generates by recursion the infinite set of all numerals, such that the range of x over SN may be seen to be embodied in the interval $[0, 1, 2, \ldots, n, n + 1, \ldots]$.

Example 1 *Numerals (SN) of arithmetic*

(1.1) $if\ x \in D, then\ x \in SN;$
(1.2) $if\ x \in SN\ and\ y \in D, then\ xy \in SN;$
(1.3) $nothing\ else\ is\ a\ member\ of\ SN.$

And so, in this way, utilizing the ten-digit vocabulary of D, the base case in (1.1), the recursive step in (1.2), and the limiting constraint in (1.3), we derive all the whole numbers out to infinity.

Now, let's look at the other example of recursion below. Albeit, different in content, example (2) is no less recursive in structure than is example (1) in the number-theoretic terms outlined above. Setting up the example we have several definitions:[5]

Definition 2 (Axiomatic System) *An axiomatic system is an ordered triple* $(A,\ S,\ P)$ *in which*

(1.) $A\ is\ a\ finite\ set\ of\ symbols,\ called\ the$ *alphabet.*
(2.) $S\ is\ a\ set\ of\ strings\ on\ A,\ called\ the$ *axioms.*
(3.) $P\ is\ a\ set\ of\ n-place\ relations\ in\ A*,\ the\ set\ of\ all\ strings\ made\ from\ the\ alphabet\ A,\ where\ n\ \geq\ 2\ (i.e.,\ the\ n-tuples\ in\ P\ must\ be\ at\ least\ ordered\ pairs).\ The\ members\ of\ P\ are\ called$ *productions or rules.*

[3] The notion of *competence* and its intimate connection with the explanation of UG is all too familiar within the tradition of GG. Writes Noam Chomsky in his *Reflections on Language*:

> The proper way to exorcise the ghost in the machine is to determine the structure of the mind and its products. There is nothing essentially mysterious about the concept of an abstract structure, created by an innate faculty of mind, represented in some still-unknown way in the brain, and entering into *the system of capacities and dispositions* to act and interpret. On the contrary, a formulation along these lines, embodying the conceptual *competence-performance distinction* seems a prerequisite for a serious investigation of behavior. Human action can be understood only on the assumption that first-order capacities and families of dispositions to behave involve the use of cognitive structures that express systems of (unconscious) knowledge, belief, expectation, evaluation, judgement, and the like (Chomsky 1975:23-4; emphasis added).

Elsewhere, he writes, "Any interesting [**GG**] will be dealing, for the most part, with mental processes that are far beyond the level of actual or even potential consciousness" (1965:8; See also pp. 19). The notion suggested here, concerning the problem of "epistemic access", seems to me no less important today, especially in regards to assessing the overall plausibility of MP's account of GG's principled basis.

[4] Cf. Davis and Gillon (2004:23). For more on the relationship between Recursion Theory and Linguistics, see especially their chapter, "Linguistics and Logic", pp. 22-68. Also, just as a matter of interest, recall what Chomsky says in his *Aspects of a Theory Syntax*, "The syntactic component [D-Structure] specifies an infinite set of abstract formal objects" (1965:18).

[5] The following is adapted in part from Partee, Meulen, and Wall (1990:186,190).

Definition 3 (Derivational Linearity in A*) *Given an axiomatic system* $(A,\ S,\ P)$, *a linearly ordered sequence of strings* $y_1,\ y_2,\ .\ .\ .\ , y_m$ *is called a derivation or proof of* y_m *if and only if every string in the sequence is either* (1) *an axiom, or* (2) *follows one of the productions in* P *from one or more strings preceding it in the sequence. If there is a derivation of* y *in a given axiomatic system,* y *is called a theorem of that system.*

Definition 4 (Theory Specification (Θ_α-Specification)) *A set of axioms together with all the theorems derivable from them is called a theory, where a theory* Θ_α *is said to be equivalent to some subset* a *in the mathematical structure* $\mathsf{A}*$, *such that* $\Theta_\alpha \equiv \mathsf{a} \in \mathsf{A}*$, *as introduced above in Def.* (3).

Applying these Definitions (2-4), we get the following example from the theory of formal languages:

Example 2 *A recursive grammar over* $A \cup B$ (the union of sets A and B), *where* Θ_{FL}, *utilizing Definitions* (3) *and* (4) *above, is taken to be* $\Theta_{FL} = P$ *by* $(A \cup B)*$.

$$A = \{E, F\}$$
$$B = \{\&, \vee, \neg, \longrightarrow, \equiv, (,), \mathrm{p}, \prime\}$$

$$P = \begin{cases} \alpha F\beta \longrightarrow \alpha\neg F\beta \\ \alpha F\beta \longrightarrow \alpha(F\,\&\,F)\beta \\ \alpha F\beta \longrightarrow \alpha(F\ \vee\ F)\beta \\ \alpha F\beta \longrightarrow \alpha(F\ \longrightarrow\ F)\beta \\ \alpha F\beta \longrightarrow \alpha(F\ \equiv\ F)\beta \\ \alpha F\beta \longrightarrow \alpha E\beta \\ \alpha E\beta \longrightarrow \alpha E\prime\beta \\ \alpha E\beta \longrightarrow \alpha\, p\, \beta \end{cases}$$

where α and β are any strings on $(A \cup B)^*$.

What we have here then is the specification of a formal language, Θ_{FL}, which is denumerably infinite, in terms of a finite specification schema (A, B, P), where A and B are Θ_{FL}'s closed set of terms and operations and P is an 8-tupled series of recursive steps over $(A \cup B)$, generating the theory (recall Def. (4) Θ_{FL}, the set of all *well-formed formulae* (*wff*) in $(A \cup B)^*$, such that (A, B, P) *specifies* all, and only those, *wff* $\in (A \cup B)^*$.

A couple of points are worth noting here, before we advance the loose thesis of the next two sections: The first of these two points is to clarify the notion of a theory Θ_α (or, alternatively, $a \in A^*$). Partee, Meulen, and Wall write in their *Mathematical Methods in Linguistics*:

> We may [. . .] define a *language* (over a vocabular A) as *any* subset of A^*. Since A^* is a denumerably infinite, it has cardinality $\aleph_0$; its power set, i.e., the set of all languages over A, has cardinality $2^{\aleph_0}$ and is thus non-denumerably infinite. Since the devices for characterizing languages which we will consider, *viz.*, formal grammars and automata, form denumerably infinite classes, it follows that there are infinitely many—in fact, non-denumerably infinitely many—*which have no grammar*. What this means is that there are languages which are such motley collections of strings that *they cannot be completely characterized by a finite device. The languages which are so characterizable exhibit a certain amount of order or pattern in their strings which allows these strings to be distinguished from others in A* by a grammar or automaton with finite resources.* (1990:435; added emphasis is mine save emphasis in the word "language")

What this would then suggest is that all (natural) languages form some subset of A^*, the set of all possible, denumerably infinite languages, where "language" is as defined above, being defined first and foremost—for any one $\mathrm{a} \in A^*$, where a is equal to some $l \in L$—by some criterion of "countable" *specifiability*. This we might call the criterion of Θ_α-*specifiability*, drawing on Def. (4). And yet, in the following paragraph, Partee, Meulen, and Wall seem inclined to maintain that the complexity of any one $l \in L$, where l is understood to be some one particular (natural) language as *governed* by principles of L, where L is specified in UG, is perhaps structurally too multivalent to be satisfactorily characterized in the simplistically computational terms of Θ_α-specifiability. (One might reasonably wonder, though, what the alternatives are. It is with this question in mind, that I've written this paper.) Writes Partee, Meulen, and Wall:

> The study of formal languages is essentially the investigation of a scale of complexity in [the patterning of strings]. [. . .] One linguistic application of these investigations is to try to locate natural languages on this scale. This is part of the overall task of linguistics to characterize as precisely as possible the class of (potential and actual) natural languages and to distinguish this class from the class of all language-like systems which could not be natural languages. One must keep clearly in mind the limitations of this enterprise, however, the principal one being that languages are regarded here simply as string sets. It is clear that *sentences of any natural languages have a great deal more structure than simply the concatenation of one element with another*. Thus, to establish a complexity scale for string sets and to place natural language on this scale may, *because of the neglect of other important structural properties*, be to classify natural language along an ultimately irrelevant dimension. Extending results from the study of formal languages into linguistic theory must therefore be done with great caution. (1990:435-7; emphasis added)

All this is simply to say, that in likening the principles of UG, as purportedly described, at least in part, by MP, *caveat emptor*; nevertheless, as I hinted at above, conceivability commitments seem to suggest that l-recursion over L must at least be *something* like the mechanisms of recursion observed in Examples (1) and (2) above. It's just not *exactly* clear *how* the sum total of mechanisms behind GG would otherwise operate, across the global generative matrix of competence interfaces, to produce, what Chomsky (1997:15-16) calls, "E-language" from "I-language". In whatever event, Chomsky does seem to buy the criterion of Θ_α-specifiability, when he writes, "A standard hypothesis is that one component of the mind/brain is a *parser* (18; emphasis in text). Taking up with this, as yet, ambiguous notion of a parser component in the mind/brain, we may pause a moment, before entering discussion of its role in the matrix of competence interfaces, and ask of our particular language, Θ_E (English), whether or not we can believe Θ_E to be specified, *in principle*, in terms stated above, in Definitions (2-4). Particularly, we may want to ask of Definition (3), whether or not, its possible for a Θ_α, where Θ_α is some one $l \in L$, to be specified in GG-supervenience relations by some Derivational Linearity—i.e., a monotonic function or deductive system—in A^*, as described in Def. (3) above—something on par with methods of deduction, where all Θ_E's *wff* are specified in terms of some closed set and a replacement schema, the likes of (A, B, P) in Example (2) above.[6] Essentially, the question reduces to one familiar in computational linguistic studies of semantics: *the compositionality hypothesis* (Cf. Cann (1993); Cruse (2004:65-79); Kamp and Reyle (1993); Janssen (1997:417-473)).[7] And that would make good sense, if, that is, the principles of GG can, in principle, be specified, in some desirably rigid way– say, by some explicit set of cohesive algorithms– across all competence interfaces and at all contributing competence components (i.e., at D-Structure, . . . ,n-Structure, . . . , LF, *et al*). Ultimately, this is a question I leave open, and rightly so, I think, for it seems a long ways off before any real answers to these questions can be brought to bear. In whatever event, it seems safe to start at this sort of "at least" notion of the conditions of recursion in any $a \in A^*$, where L and all its members are presumed to be proper subsets.

1.2 The Matrix of Competence Interfaces (MCI) and Language

As we'll be specifying the first-starters at L for all l in section two, it's helpful to know where L-Government resides in MP's theory of GG. The MCI will help with this, specifying not only L-Government's residence in

[6] Partee, Meulen, and Wall write, "A formal grammar (or simply, grammar) is essentially a deductive system of axioms and rules of inference, which generates the sentences [the *wff*] of a language [Θ_α] as its theorems" (1990:437). Here, there's a sense in which they want to open up this proclamation on the set of all grammars, not just formal grammars, as indicated in the hesitant parenthesis.

[7] Borrowing from Ronnie Cann (1993:4), we may define the *compositionality hypothesis* as follows:

> **The principle of compositionality**
> The meaning of an expression is a monotonic function of its parts and the way they are put together.

As suggested, this brings to mind (A, B, P), where the "parts" are specified in terms of A and "the way they are put together" is specified in terms of B and its use in P.

the mind/brain but also its scope. Below, we have a figure, Fig. 1., describing the "place" of *L*-Government at D-Structure and all its subject modules.

$$\begin{tikzcd} & \text{D-Structure} \arrow[d] & \text{Lexicon} \arrow[l, dashed] \\ \text{LF} & \text{S-Structure} \arrow[l, dashed] \arrow[d] & \\ & \text{PF } (\textit{MF}) & \end{tikzcd}$$

Figure 1: The Matrix of Competence Interfaces (Chomsky 1997:22; added material in italics is mine)

Already implicit in our discussion of grammars *in toto* we have employed the notion of "Language" versus "language", denoted L as distinguished from l (Cf. Chomsky (1997:7,25) for more on this important distinction). L is essentially that which predicates UG in GG, and, using Fig. 1. above, we note L is embodied in D-Structure as the basis of all l, generated at "higher" subject modules. Notes Chomsky, "The *base* of the syntactic component is a system of rules that generates a highly restricted (perhaps finite) set of *basic strings*, each with an associated structural description called a *base Prase-marker* [BPM]. These [BPMs] are elementary units of which deep structures are constituted" (1965:17; emphasis in text). In the following, as we look at aspects of MP's take on GG, we'll keep these notions in mind, especially as regards what is meant when Lasnik writes, "In contrast [to the tradition on grammar, including early GG], the P & P approach maintains that the basic ideas of the tradition, incorporated without great change in early generative grammar, are misguided in principle—in particular, that a language consists of rules for forming grammatical constructions (relative clauses, passives, etc.)" (Chomsky 1997:5). He goes on to write, "The P & P approach held that languages have no rules anything like the familiar sense, and no theoretically significant grammatical constructions except as taxonomic artifacts. There are universal principles and a finite array of options as to how they apply (parameters), but no language-particular rules and no grammatical constructions of the traditional sort within or across languages" (5-6).

We'll examine these assertions, especially those that insist MP stands in stark contrast to the tradition, and even to traditional GG itself. Also, we'll speculate on what is meant, when Chomsky writes, "The apparent eliminability of phrase structure rules became clear by the late 1960s, with the separation of the lexicon from the computational system and the development of *X-bar theory*" (1997:25). Concerning this separation, Lasnik describes,

> For each particular language, the cognitive system, we assume, consists of a computational system CS and a lexicon. The lexicon specifies the elements that CS selects and integrates to form linguistic expressions—(PF, LF) pairings, we assume. [. . .] Virtually all items of the lexicon belong to the *substantive categories*, which we will take to be noun, verb, adjective, and particle, putting aside many serious questions about their nature and interrelations. The other categories we will call *functional* (tense, complementizer, etc.). (Chomsky 1997:6; emphasis in text)

Further characterizing the CS, as depicted above in MCI, we may note, further borrowing from Chomsky, "The phonological component of a grammar determines the phonetic form [PF] of a sentence generated by the syntactic rules," thus creating "a certain phonetic representation. The semantic component determines the semantic interpretation of a sentence," creating "a certain semantic representation. [. . .] Both phonological and semantic components are therefore purely interpretive" (1965:16). Intermediating these levels of representation is S-Structure, at which Transformational Grammar (TF) takes place. "In addition to [CSs base in MCI at D-Structure], the syntactic component of a generative grammar contains a *transformational* subcomponent. [. . .] [t]he central idea" of which "is that [deep and surface structures] are, in fact, distinct and that the [S-Structure] is determined *by repeated application of certain formal operations [over BPMs] called 'grammatical transformations'*[. . .] (emphasis added)." In section two, the nature of BPMs will be our primary occupation, as viewed under the MP lense. Following that, in the last section, we'll take up TF and *L*-Government's distal scope over TF at S-Structure, as depicted by the merits of MP's take on GG.

2 The Building-Blocks of Minimalist Derivation

"The P & P approach," writes Chomsky,

> aims to reduce descriptive statements to two categories: *language-invariant* and *language-particular*. The language-invariant statements are principles (including parameters, each on par with a principle of UG); the language-particular ones are specifications of particular values of parameters. [. . .] [Or, again,] the hypothesis is that all principles are assigned to UG and that language variation is restricted to certain operations as to how the principles apply. [. . .] [Therefore,] [t]he notion of construction, in the traditional sense, effectively disappears; it is perhaps useful for descriptive taxonomy but has no theoretical status. (1997:24-25)

Employing the distinction sketched above, it is the task of this section to tackle certain aspects of "language-invariance", while in the next section we'll take up the relationship language-invariance in L has to the "language-particular" features of Θ_E (English), as embodied in aspects of TF-constructions like Head Movement.

2.1 Primitive Structures: Operations over an Ordered Quadruple $\langle\delta, \langle\sigma, \langle\langle\lambda\rangle, \langle\pi(\mu)\rangle\rangle\rangle$

MP's description of UG, as was GG's, is essentially ordered. *Order*, as indicated in Definition (3), is a vital part of any grammar. In the case of MP, however, the ordering is more complex still in that several orders of functional mapping take place across the Matrix of Competence Interfaces. MCI is an ordered quadruple in that δ (D-Structure) directionally maps into σ (S-Structure); and, moreover, σ directionally maps into both λ (LF) and π (PF).[8] Upon "receiving" a value, the value of σ is then mapped into either λ or π, or both, thus giving one—or both, as the case may be—a unique value at LF or PF. In this section, we'll primarily be concerned with principles at δ, or D-Structure, where this component is traditionally understood under the P&P guise as possessing rudimentary principles of global government and parameterizations over these.[9]

Let us first begin by looking at some of the basic structures used as components in other more complex structures, generated from the D-Structure of MCI, in any $l \in L$, at particular parameterizations in δ. Basic among the principles of Minimalism is the principle of "Headedness". The Headedness Principle is defined below as follows:

> **Headedness Principle (HP)**
> *Every* syntactic structure is a projection of a head word.

A head word, it is said, has one and only one complement, such that, for every one head word, there is exactly one complement. This is true even in those cases wherein we find that one of these components in this concatenation is said to be null (that is, $\emptyset$).[10] As such, it is helpful to go ahead and define a result of the Headedness Principle outlined informally above.

---

[8] As in the figure of MCI in section (1.2) and the ordering schema above, I leave it open as to how the mapping takes place between σ and both λ and π. I do so for various reasons, principal among these being that the literature itself is unclear about this interfacing phenomenon, as it is also about the mapping into μ (MF), hence the added ambiguities above.

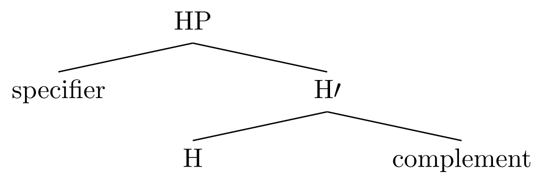
[9] Although this is not the place to discuss language acquisition, it should be said, that in the move from the initial state to the steady state in competence, it is thought that language acquisition is merely the setting of the values of parameters. Howard Lasnik writes in *The Minimalist Program*:

> In this [MP] context, language acquisition is interpreted as the process of fixing the value of parameters of the initial state in one of the permissible ways. A specific choice of parameter settings determines a *language* in the technical sense that concerns us here: an I-language in [Chomsky's sense], where I is understood to suggest "internal", "individual", and "intensional". (Chomsky 1997:6; See also Radford (2004:21-26) where he discusses both "lexical learning" and its role in parameter-setting.)

[10] This raises some interesting issues in Empty Category Theory (**ECT**) we can only make cursory light of in this essay (Cf. Chomsky (1997:129-66) for helpful exposition of ECT's crucial role in MP and especially Transformational Grammar).

Binarity Principle (BP)
Every syntactic structure is binary-branching.[11]

It only takes a moment to see that this follows as a direct result of the HP, in that, if *for every* syntactic structure σ in Σ—where $\Sigma = \Theta_\Sigma$ over L—σ in Σ is understood to be a projection of at least one (and only) one head word, then, because *every* head has at least one and only one complement, it follows that *every* syntactic structure σ in Σ is also binary-branching in Σ.[12] As we go on, we'll define these principles more explicitly. Also, as we proceed, we'll be enlisting other principles similarly basic in character, as well as their associated constraints. For now, however, having set out two of the primitive most principles in the set, we may display a generalized matrix exhibiting their structure, as in the phrase structure tree (**PST**) below in (1).



(1)

Analyzing the structure of the PST in (1), we see that both HP and BP have been complied with in the above structure. Branching down from the topmost "parent" node *HP*, we find two constituent "offspring", the *specifier*, which, in this case, happens to be a head, and the *H/* ("H-bar"), which happens to be the complement. Branching further still, we find as constituent offspring of the parent node *H/* both an *H*, obviously a grammatical head, and a *complement*, also obviously a complement. It should be noted, then, that, with each successive branching from the parent node (and, in this generalized matrix, there is only two), the branching array is binary in structure. By the BP, this holds by global governance across all structures, however high-level or specific at S-Structure. Additionally of interest here is the MP-theoretic assertion that, for any HP whatever, where *HP* is taken as a variable, ranging over all *phrase structures*, HP *contains* both a head and a complement. This is the case even in the H-bar extension exemplified above.

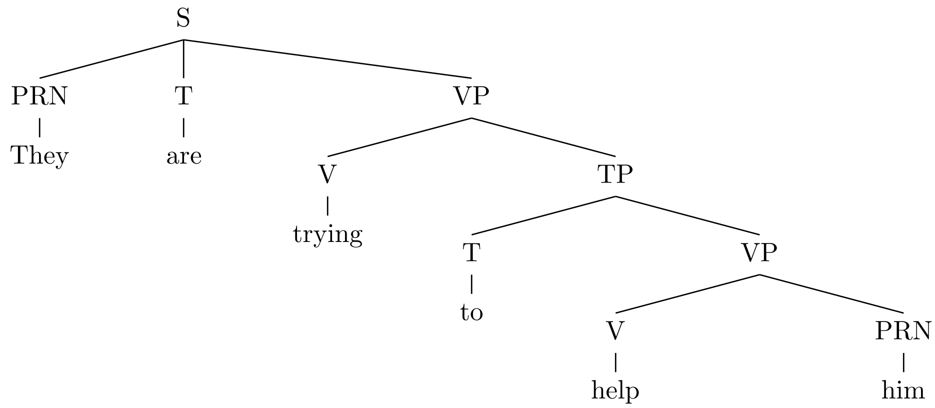
Another point of interest is worth noting here, in that, like both HP and BP, this principle is also primary most in governance, likewise scoping over all GG-generations, however high-level or specific at S-Structure. The notion is one already implicit in (1) above, known as the Extended Projection Principle, as defined below (Radford 2004:72).

Extended Projection Principle (EEP)
Every finite tensed constituent T must be extended into a TP (Tense Phrase) projection containing a subject.

To explain the Extended Projection Principle, there are a number of other key elements which also warrant explanation. I'll highlight these other key elements and also EPP's relationship with the two basic principles stated above—HP and BP—by way of a couple of examples. Prior to the Minimalist approach to D-Structure, a phrase structure tree would have represented the syntactic relationships in *They are trying to help him* something like this:

[11] Radford (2004:70) suggests that BP extends also into the Phonological (**PF**) and Morphological (**MF**) Structures as well, thus suggesting a principled consistency from the D-Structure computational interface to the computational interfaces of both PF and MF. BP would then seem to scope over merger operations in D-Structure $\Longrightarrow$ [S-Structure $\Longrightarrow$ [PF $\Longleftrightarrow$ MF]]. If this were to be the case, it would certainly lend desired credibility to the Minimalist Program and its explanations of UG.

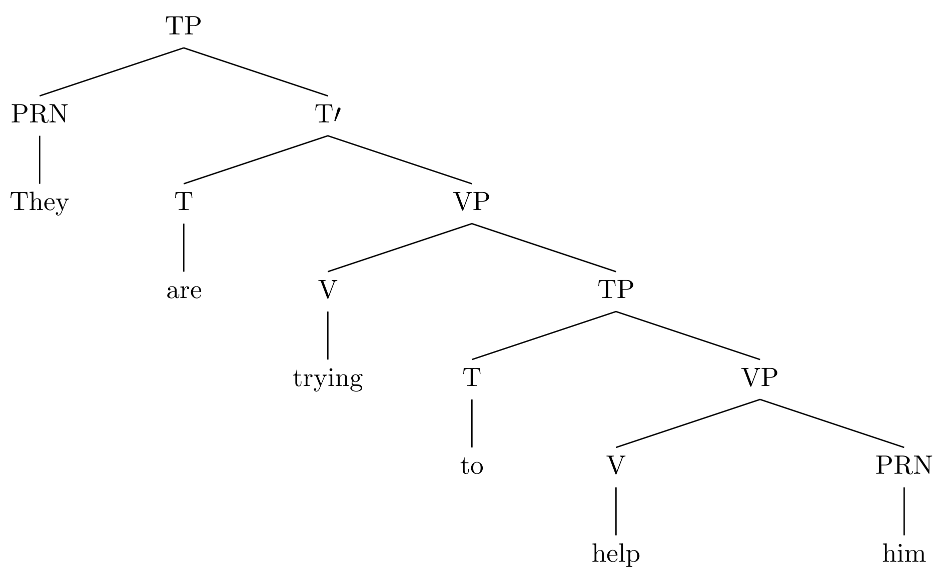
[12] We'll discuss later whether or not it is possible for a given σ in Σ to have more than one head and still comply with BP.



(2)

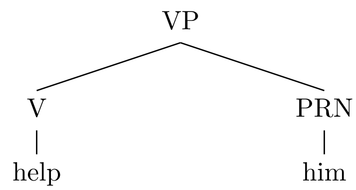
Without going into matters of justification, as to why the MP-principles “explain” the material “better”, we may proceed on the widely acceded assumption, that *parsimony* is a virtue of theory in general in all those cases—no matter the complexity of the phenomenon under consideration—wherein we find that the simplest answer, in point of fact, *better* accounts for the considered phenomenon than some alternative, less simple explanation. That said, if upon consideration, we find that the PST in (2) above may be *reduced* to principles of derivation more basic in governance, we may then wish to adapt our present understanding of grammatical phenomena to these known and unknown members of the *closed set*, L, where the presumption of their constituting a closed set is simply the assumption that this number of primary most MP-principles is finite in countability—that is, the number of MP-principles in L meets Θ_α-specification—and is open to *recursive access*.

Re-rendering structure (2) above to comply with the principles already discussed, we find the following structure:



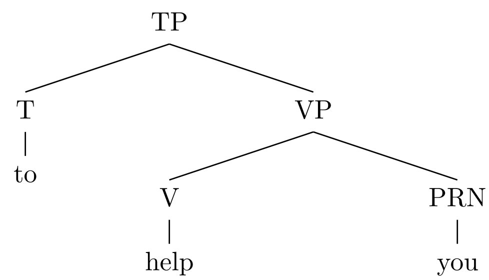
(3)

To clarify the structure above, we may better define the notions of constituent and head. Below, we see the simple verb phrase structure



(4)

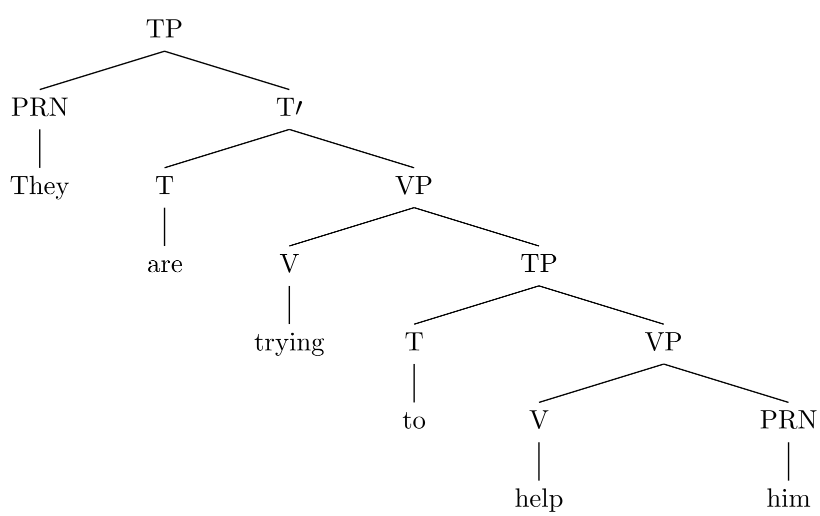
in which there are two constituents, a verb and a pronoun. The head of this verb phrase is the verb *help*, about which we may say *help* is a *minimal projection* and VP is its *maximal projection*. *Help* is a minimal projection, in that in structure (4), *help* is a constituent which is not a projection of some other word. VP is *help*'s maximal projection, in that *help*, a V-constituent, is contained within this larger constituent, a VP-constituent, and there is no larger constituent than VP containing the same head word *help* (Cf. Radford (2004:473) for more on these and the following notions). Now, were we to extend the structure in (4) at bit as in (5) below



(5)

the conditions of projection change. While it is still the case in (5) that *help* is a minimal projection, VP is no longer *help*'s maximal projection, because, now there is a larger constituent, TP, containing the V-constituent *help*. To see this, note that, just as the TP constituent contains its head T, the word *to*, so also it contains VP, which, in turn, is a larger constituent, containing its head V, the word *help*. Therefore, we can begin to see that the notion of *containment* is transductive, much as we might expect from our discussions about Derivational Linearity.[13]

Extending this analysis all the way out into a clause, we get the structure



(6)

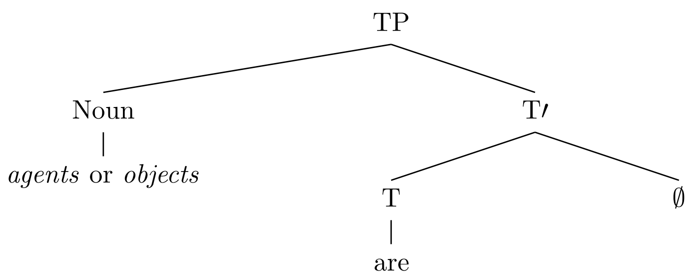
[13] Transduction is a result of monotonic structures.

already seen above in (3). We might still wish to ask how it is these principles operate in the full extension. And to answer, we look again at the Principle of Extended Projection. Applying the principles of analysis considered already to the structure in (6), we see that, where *They*, a PRN-constituent, is considered to be the head of TP, the largest constituent in the structure, TP is not only the maximal projection of *They* but also of the whole structure. For consider, if S is the largest constituent containing the the head *They* and if it is clear that T′ nor any other phrase constituent contains *They*, while at the same time being larger than TP, then, by definition, how can TP not be the maximal projection of the whole structure? If TP contains, by pairwise merger operation, both *They* and T′, then TP contains both *They* and *are trying to help him*, because T′ contains both *are* and VP, where VP is understood to contain both *trying* and TP; and, TP is understood to contain both *to* and VP; and, VP is understood to contain both *help* and *you*, thus accounting for the whole clause in the series of merger operations: [*We*$_{\text{PRN}}$ [*are*$_{\text{T}}$ [*trying*$_{\text{V}}$ [*to*$_{\text{T}}$ [*help*$_{\text{V}}$ [*you*$_{\text{PRN}}$]]]]]].

Recall now the Principle of Extended Projection, which states:

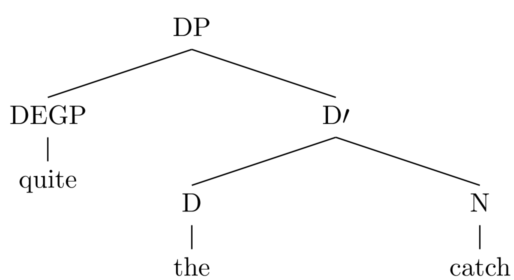
> **Extended Projection Principle (EEP)**
> *Every* finite tensed constituent T must be extended into a TP (Tense Phrase) projection containing a subject.

Explicating EPP in terms of (6), note that, in the case of *are*, where *are* is the head of T′, we have a head which necessitates *extension* into a TP, such that T′ is considered the *intermediate projection* of *are* and TP is considered its maximal projection. This analysis follows from our discussion of containment relations from above, with one exception. In the case of *are*, *are* necessitates a subject because of certain *selectational features* (Chomsky 1997:32-35; Radford 2004:68,476), and this would be the case even if we had a structure like (7) below:



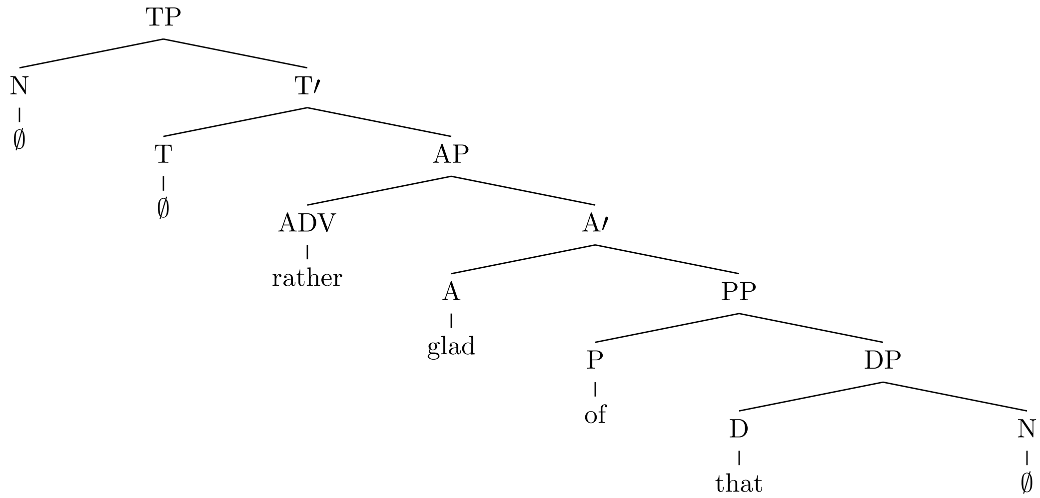
(7)

Other X-bar theoretic structures, operating on the template of (1) above, are as follows: In the case of (8) below, we see that even determiner phrases might have bar-dependencies.



(8)

And again, we see also that in some adjective phrases bar-dependencies might exist, as in (9).



(9)

Before moving on to transformations over these primitive structures, two things should be mentioned. The content of *basic Phrase markers* [BPMs] is not altogether clear. No where in the literature that I've come across is there specified specific content for BPMs. Characterizing the nature of BPMs is most often done tentatively and in an indirect manner. Wishing to be no more ambitious than they, I will do likewise, merely making suggestions, based upon my interpretation of what MP is up to in its sketches of L, where L is taken to be equivalent to UG. The PSTs above are all Θ_E parameterizations on MP-principles, if for no other reason than that they are right-branching. English just so happens to be a *head-first* parameterization on L, whereas other languages, e.g., Japanese, happen to be *head-last* parametrized. Radford writes, "Many other logically possible orderings of heads with respect to complements appear not to be found in natural language grammars" (2004:20). This suggests the notion of constraints. Is it the case—as it certainly appears to be—that Kaynean (Chomsky 1997:131) binary branching is constrained to one of two choices: (1) head-first or (2) head-last, for any Θ_α, where a given Θ_α is understood to be equivalent to some one $l \in L$? We can't be sure, but certain probability indicators are there.

Again regarding the nature of BPMs, Thomas Klammer, Muriel Schulz, and Angela Volpe (2007:197-237) seem to suggest that there are five basic sentence types, perhaps on par with something like Chomsky's "kernel sentences" (1965:17; 1997:23), with the following structures:

(1.) $[\,\mathrm{NP} + \mathrm{MVP}_{\,\mathrm{intrans}} + (\mathrm{ADVP})\,]$
(2.) $[\,\mathrm{NP} + \mathrm{MVP}_{\,\mathrm{be}} + \mathrm{ADVP}_{\,\mathrm{time\backslash place}}\,]$
(3.) $[\,\mathrm{NP} + \mathrm{MVP}_{\,\mathrm{link}} + \mathrm{ADJP}\,]$
(4.) $[\,\mathrm{NP}_{\,1} + \mathrm{MVP}_{\,\mathrm{link}} + \mathrm{NP}_{\,1}\,]$
(5.) $[\,\mathrm{NP}_{\,1} + \mathrm{MVP}_{\,\mathrm{trans}} + \mathrm{NP}_{\,2}\,]$

Perhaps this could be the case, after all, in their text, they are analyzing Θ_E, and so, we might step back a bit and maintain that their right-branching characterization of BPMs is a mere epiphenomenon, specific to Θ_E parameterization. However, one might reasonably wonder, if, by the lights of Θ-*Theory*, one may be able to reduce most, if not all, of these to a construction even more simple structure.[14] So, for example, following this line of analysis, one concludes that ((1.)-(5.)) above reduce more simply to

(1.) $[\,\mathrm{NP} + \mathrm{MVP}_{\,\mathrm{any\ type}}\,]$

where NP and $\mathrm{MVP}_{\,\mathrm{any\ type}}$ are taken to be open to either head-first or head-last parameterization at the initial state of acquisition.

[14] Θ-Theory deals with notions like "agent", "patient", "goal", and the like (Chomsky 1997:31).

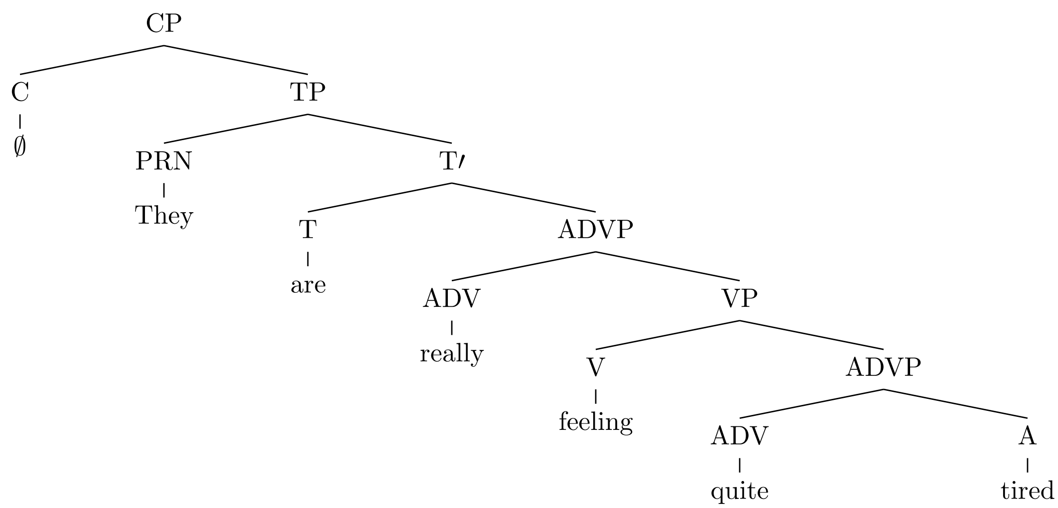
I leave these speculations aside, reiterating only what has been mentioned already in section (1.2) above:

> For each particular language, the cognitive system, we assume, consists of a computational system CS and a lexicon. The lexicon specifies the elements that CS selects and integrates to form linguistic expressions—(PF, LF) pairings, we assume. [. . .] Virtually all items of the lexicon belong to the *substantive categories*, which we will take to be noun, verb, adjective, and particle, putting aside many serious questions about their nature and interrelations. The other categories we will call *functional* (tense, complementizer, etc.). (Chomsky 1997:6; emphasis in text)

Mulling over the suggestions in this passage, perhaps we want to say that L is Θ_{α}-specified if and only if MP-principles—such as they are—operate over a set of substantive categories, wherein, conceivably, all selectational features and ϕ-features are accounted for (Chomsky 1997:35 ; Radford 2004:58-60). This seems plausible enough for our investigations, and with that we move to CP analysis of main clauses, ultimately leaving this an open question.

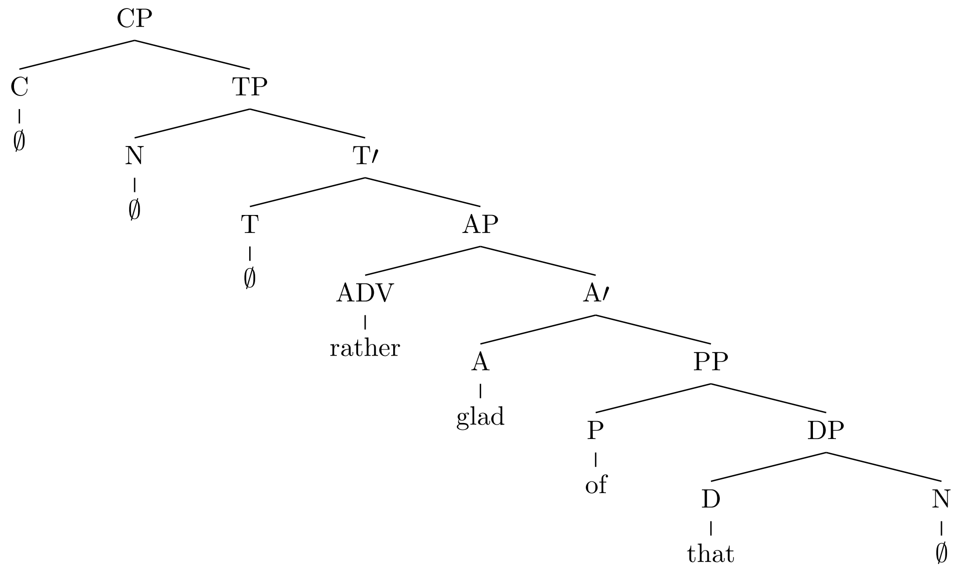
2.2 Main Clauses as Complementiser Phrases (CP): A Primitive?

Some hold that all "main clauses are CPs headed by a force" and tense "marking complementiser" (Radford 2004:127). Here, we deal with this notion merely as an aside before looking at how primitives are transformed at S-Structure. Considering there are three different complementisers—that, for, and if—CP analysis of a generalized main clause matrix gives us



(10)

which, depending upon which complementiser has been given a null spell-out at PF, we have any one of three different values for (10): (1) [Declarative + Finite Force], if the head of CP, $\emptyset$, equals *that* at S-Structure; (2) [Interrogative + Finite Force], if $\emptyset$ equals *if* at S-Structure; or (3) [Irrealis + Infinite Force], if $\emptyset$ equals *for* at S-Structure (Radford 2004:52-57,127-28). These different value settings are of crucial importance for Transformations at σ. So, for instance, again looking at (9) above, CP analysis of (9) would generate the PST in (11) below:



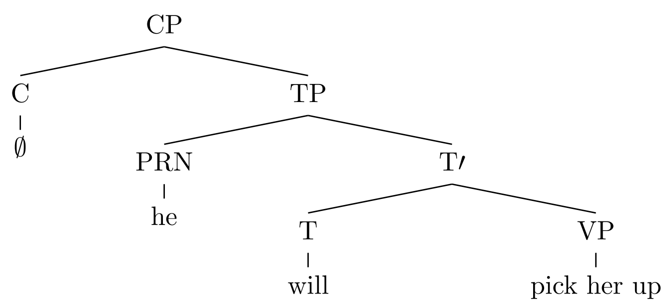
(11)

3 Constituent Movement and *L*-Government: Aspects H-Movement

The notion of C-Movement as *copy-and-deletion* is a revised version of something known as *trace theory* in the P& P literature. Here, we'll begin synopsizing aspects of H-Movement by first taking a look into the mechanism of transformational Aux-movement.

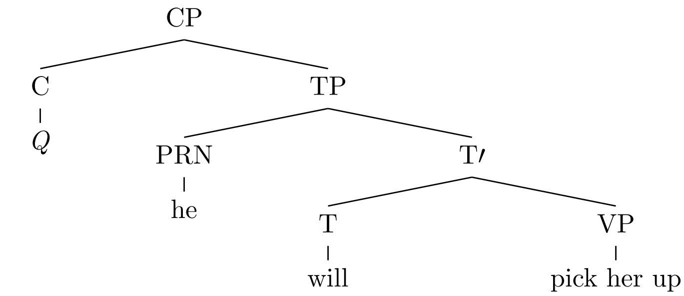
3.1 Elements of TF

To begin, we have a presumably [Dec-Force] main clause, in which case the null spell-out ∅ at PF equals *that*.



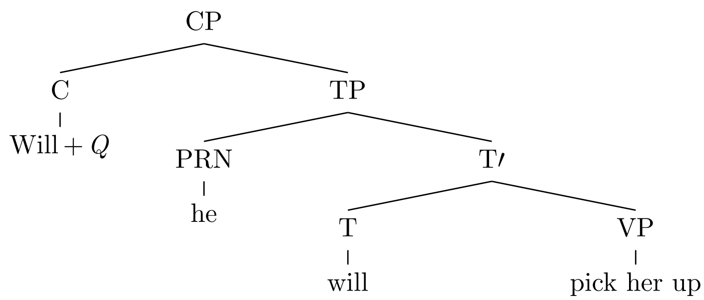
(12)

But what if the null spell-out ∅ at PF is not equal to *that* but rather, say, *if*. In this case, then, instead of (12) above, we would have something like the following structure:



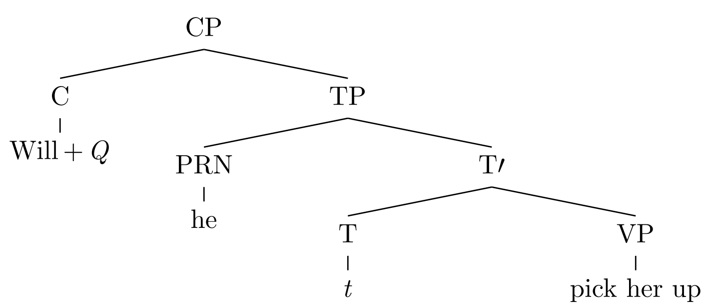
(13)

How then is (13) to transform? It must be remembered that Q is still the null spell-out $\emptyset$ at PF but is nevertheless [Int-Force]. That said, which head is to move, and where? To answer, we find the transformed version of (13) in PST (14) below:



(14)

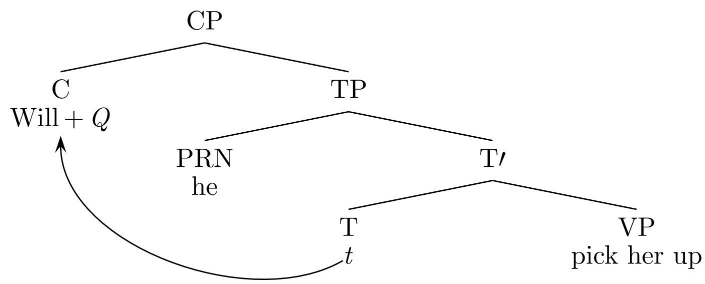
But something is still not right. Now the TF-construction is * *Will he will pick here up*, clearly an aberrant transformation over the supposed [Dec-Force] of (12) above. So, what is to be done? Well, this is where the two merger operations *copy-merge* and *copy-deletion* come in. We have already seen the effects of the first operation, copy-merge, in (14). All that remains is to observe its role in standardizing the transformation over (12).



(15)

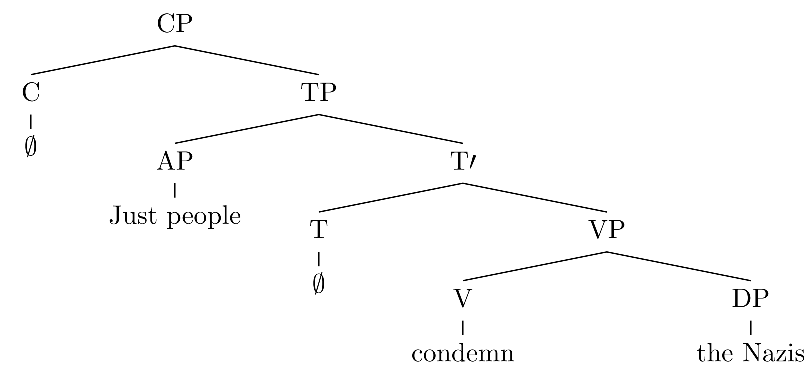
Giving (15) a full interpretation and we get

(16)



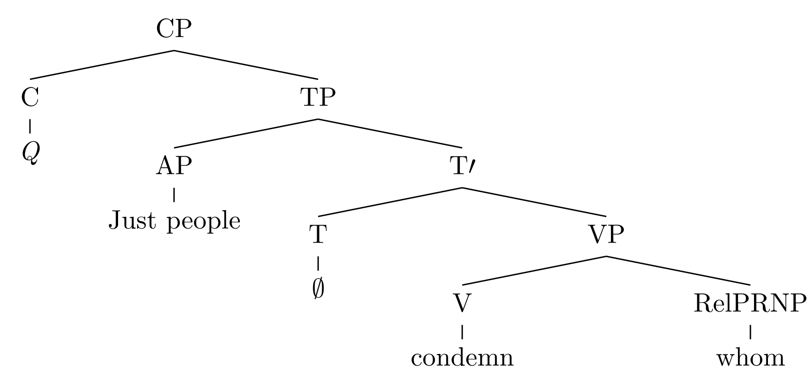
wherein we see $[\, Will_{\mathrm{T}} + Q_{\ \mathrm{C}}\ [he_{\ \mathrm{PRN}}\ [t_{\mathrm{null\ T}}\ [\ pick\ her\ up_{\ \mathrm{VP}}?\ ]]]]$. This is standard analysis for all H-Movement of Auxiliary heads at T ∈ T′. We may wish, then, to see what H-Movement is in those cases of Wh-Transformation.

(17)

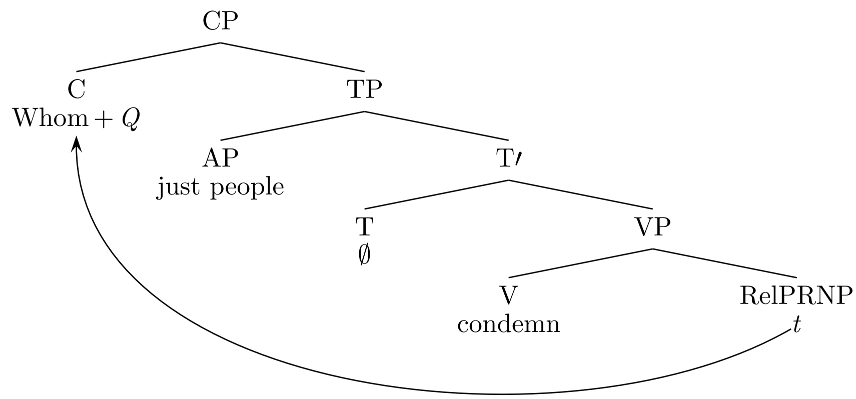


Now, supposing, in (17), the null spell-out ∅ at PF to be equivalent to *if* at S-Structure, then what we get is something like (13) above. Supposing further that (17) were to be Wh-Transformed at σ, the initial step would be as follows

(18)

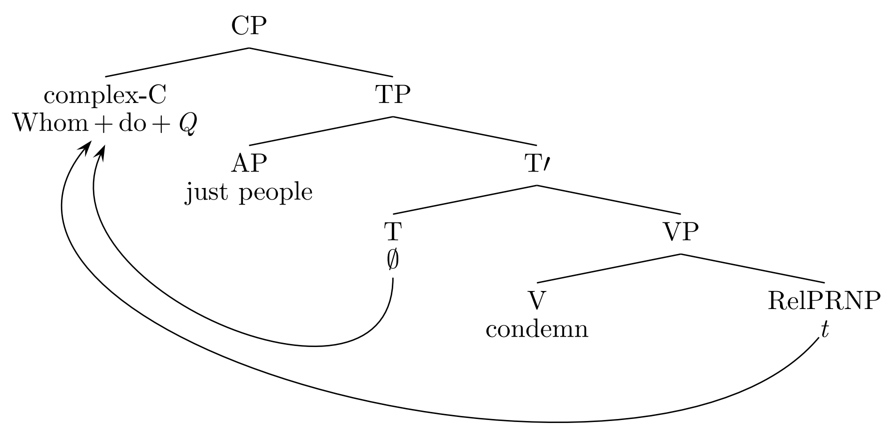


Thus, moving through copy-deletion algorithm, involving first copy-merge and then copy-delete, we get the following PST in (19):



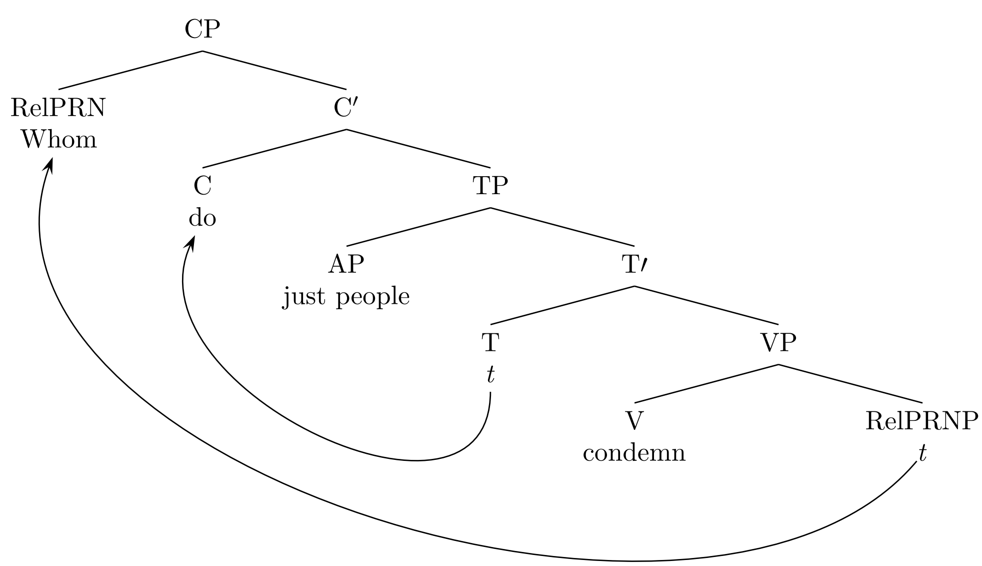
(19)

Nevertheless, we still have some more transformations to make, in that * *Whom just people ∅ condemn* t.



(20)

Perhaps then this is cause to consider again the effects of X-bar theoretic principles like EPP, particularly as they affect CP structure in PSTs like (20) above. Recalling PSTs (1) and (6-8) we might induce that the revised version would look something like this:



(21)

Reasonably then we might speculate that not only does ∅ receive a non-null spell-out at PF after Do-raising but also, upon the double operation of copy-deletion, t receives a null spell-out at PF, after (21)'s fully algorithmized Wh-Transformation.

Reaching back into L for all l, then, we need to specify a few more principles and their constraints:

> **Head Movement Constraint (HMC)**
> Movement from one head position to another is only possibly between a geven head and the closest head which asymmetrically c-commands it (i.e., between a head and the next highest head in the structure containing it).[15]

Having specified this we need to clarify the notion c-command. We may recall discussion above of the containment and its importance in projection and bar-dependencies. Generally, c-command is said to be the governing factor in Movement permissions, as specified in HMC, where c-command is defined as follows:

> **C-command**
> A constituent X c-commands its sister constituent Y and any constituent Z which is contained within Y.[16]

As a corrolary to this c-command relation and its constraint on permissible H-Movements, we find there to be another significant member in L for all l, which governs movement manner.

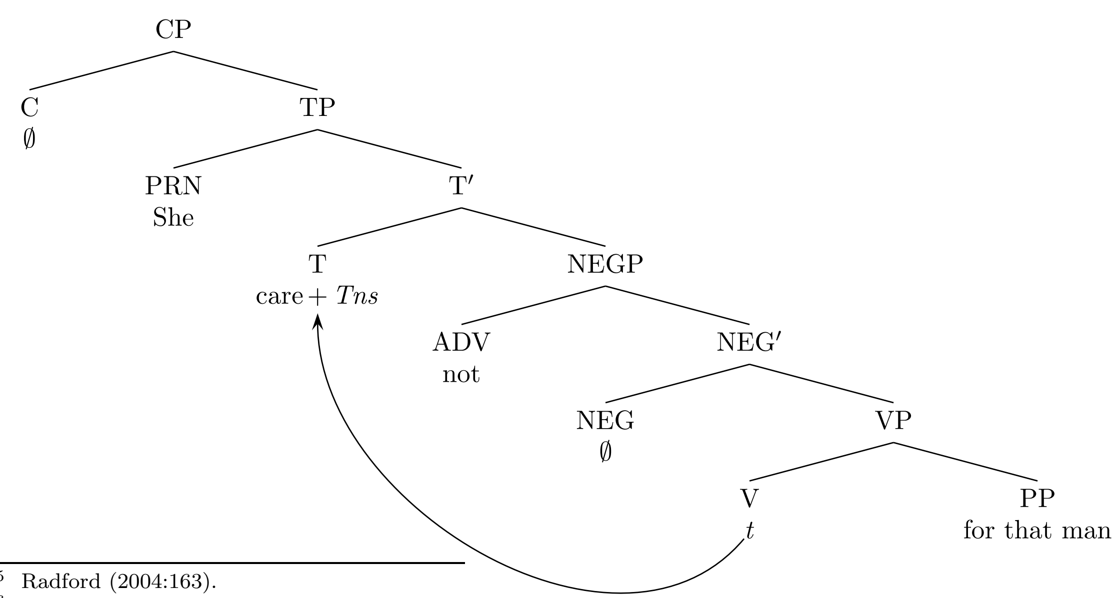
> **Strict Cyclicity Principle (SCP)**
> At a stage of derivation where a given projection HP is being cycled/processed, only operations affecting the head H of some HP and some other constituent of HP can apply.[17]

And, in general, across the MCI, there is the

> **Inclusiveness Condition**
> No new information can be introduced in the course of syntactic computation.[18]

a constraint placed on all computation in an order quadruple, particularly in the δ and σ modules and interfaces.

Having defined the above, the following PST is positioned here for expository purposes; note that the operations in (22) are prevented by L-Government and L-Binding for all l in A^*:

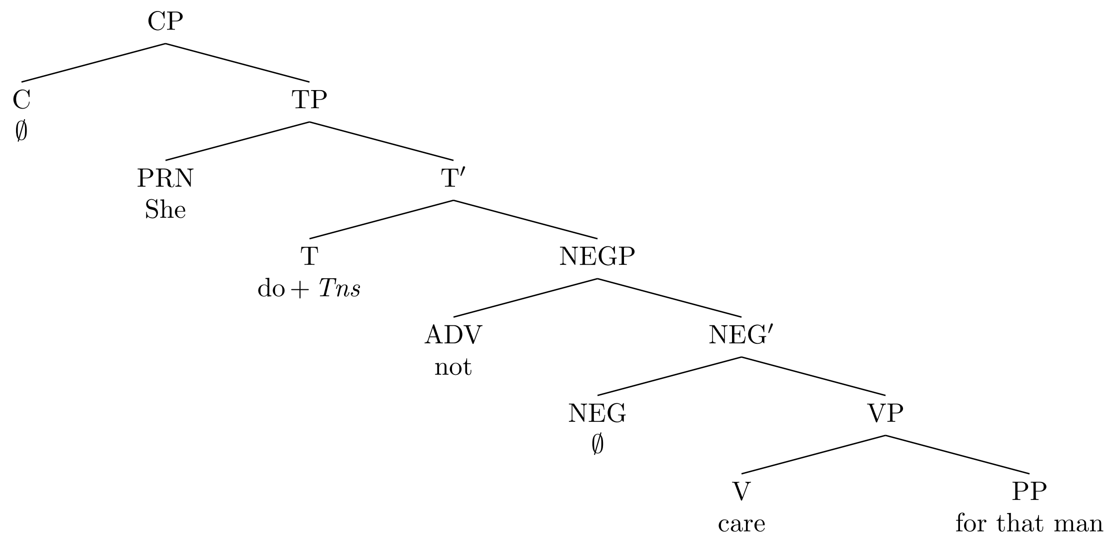


(22)

[15] Radford (2004:163).
[16] Radford (2004:91; Cf. pp. 139 for more on this).
[17] Radford (2004:173; Cf. pp. 200 for a corrolary principle to SCP).
[18] Radford (2004:94).

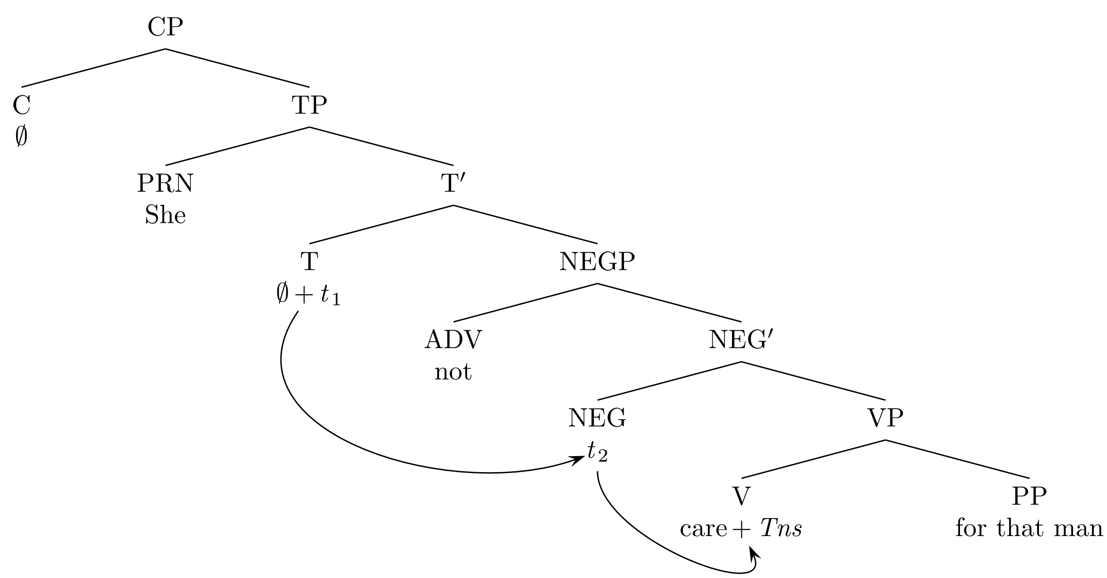
This is so because of SCP and HMC, two intimately related principles in L. So, where we have the base case, perhaps some approximation on some one BPM, as discussed above, we have the non-Transformed PST in (23):

(23)

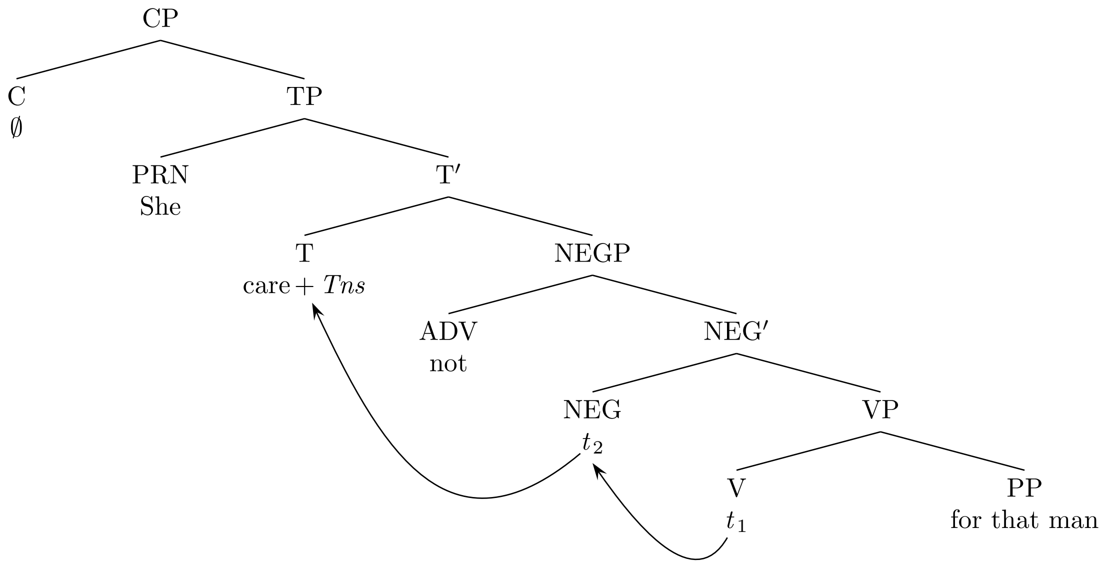


Alternatively, instead of (23) above, presumably, we might have (24) by SCP operations which result in *Affix Hopping*, as in the PST below:

(24)



In whatever the case, again considering the H-Movement of the V-constituent *care* as in the SCP and HMC violations in (22) above, getting it right, algorithmically speaking, where the displayed PSTs are here seen as *wff* productions in Θ_E, we derive:



(25)

4 Concluding Remarks

Oppurtunities for further research are myriad in number and aspect. To be sure, what's presented here is by no means an exhaustive account. Much more needs to be said about the MP-principles presented in the last section and about the way they work together. The account given above is incompete and, at best, introductory. The same limitation holds for the mathematics presented in this paper. More needs to be said of the various types of mathematical recursion, and along what lines these are to be distinguished (i.e., recursion over the complex numbers versus recursion over the reals; discrete recursion versus real-valued recursion). Interesting too would be a deeper analysis of the role these mathematical structures play in developing our intuition of the recursive structures found in natural language, particularly those structures posited by UG analysis. In fact, whole sections could be devoted to precisely these subjects. As an aside, we might also like to couple these discussions together with an investigation into research done on formal languages—into the disparate structural types and functions these are said to possess. For now, suffice it to allow me to point the way to further research in these feilds. In the bibliographic contents below, you'll find an asterisk by the texts which expand upon the scope of the present discussion.

Pursuing such research holds the delightful hope that, by gathering together a deeper, more thorough, understanding of the structure of our language—and its relationship to certain mathematical constraints—we may further comprehend our place in the world, and the way we're bound to it. Human nature, a complex of harry components, may insightfully be studied through the, oftentimes, surprising constraints under which we talk with one another.

References

[1] *J-M. Autebert, J. Berstel, L. Boasson, "Context-Free Languages and Pushdown Automata", *Handbook of Formal Languages*, (Eds.) G. Rozenberg and A. Salomaa, vol. I, 105-172.

[2] *N. Asher, A. Lascarides, *Logics of Conversation*, Cambridge University Press, New York, 2003.

[3] *H. Barendregt, "The Impact of the Lambda Calculus in Logic and Computer Science", *The Bulletin of Symbolic Logic*, 2, 3, (1997) 181-215.

[4] *J. V. Benthem, A. ter Meulen (eds.) *The Handbook of Logic and Language*, M.I.T. Press, Cambridge, 1997.

[5] *M. Bergmann, J. Moor, J. Nelson, *The Logic Book*, McGraw-Hill, New York, 1998.

[6] *R. E. Byerly, "Recursion Theory and the Lambda-Calculus", *The Journal of Symbolic Logic*, 1, 47, (1982) 67-83.

[7] *W. Buszkowski, "Mathematical Linguistics and Proof Theory", *Handbook of Logic and Language*, (Eds.) J. V. Benthem and A. ter Meulen, 685-733.
.

[8] N. Chomsky, *Aspects of a Theory of Syntax*, M.I.T. Press, Cambridge, 1965.

[9] ——, *Reflections on Language*, The New Press, New York, 1975.

[10] ——, *The Minimalist Program*, M.I.T. Press, Cambridge, 1997.

[11] H. B. Enderton, *A Mathematical Introduction to Logic*, Academic Press, New York, 1972.

[12] *F. Grecseg, M. Steinby, "Tree Languages", *Handbook of Formal Languages*, (Eds.) G. Rozenberg and A. Salomaa, vol. III, 1-49.

[13] T. P. Klammer, M. R. Schulz, A. D. Volpe, *Analyzing English Grammar*, Longman-Pearson, New York, 2007.

[14] B. H. Partee, A. ter Meulen, R. E. Wall, *Mathematical Methods in Linguistics*, Kluwer Academic Publishers, Dordrecht, Netherlands, 1990.

[15] S. Pinker, *Words and Rules: The Ingredients of Language*, Basic Books, New York, 1999.

[16] *A. Mateescu, A. Salomaa, "Formal Languages: an Introduction and a Synopsis", *Handbook of Formal Languages*, (Eds.) G. Rozenberg and A. Salomaa, vol. I, 1-38.

[17] *——, “Aspects of Classical Language Theory”, *Handbook of Formal Languages*, (Eds.) G. Rozenberg and A. Salomaa, vol. I, 175-246.

[18] *M. Moortgat, “Categorical Type Logics”, *Handbook of Logic and Language*, (Eds.) J. V. Benthem and A. ter Meulen, 116-171.

[19] A. Radford, *Minimalist Syntax: Exploring the Structure of English*, Cambridge Textbooks in Linguistics, Cambridge University Press, New York, 2004.

[20] K. Ross, C. R. B. Wright, *Discrete Mathematics*, Prentice Hall, Upper Saddle Riv., New Jersey, 2003.

[21] *G. Rozenberg, A. Salomaa (eds.) *Handook of Formal Languages*, Springer-Verlag, Berlin, vols. I, II, III, 1997.

[22] *J. R. Shoenfield, *Mathematical Logic*, Addison-Wesley, Reading, Mass., 1967.

[23] *W. Thomas, “Languages, Automata, and Logic”, *Handbook of Formal Languages*, (Eds.) G. Rozenberg and A. Salomaa, vol. III, 389-449.

[24] *R. H. Thomason, “Nonmonotonicity in Linguistics”, *Handbook of Logic and Language*, (Eds.) J. V. Benthem and A. ter Meulen, 779-825.

[25] *R. Turner, “Types”, *Handbook of Logic and Language*, (Eds.) J. V. Benthem and A. ter Meulen, 419-470.

[26] *S. Yu, “Regular Languages”, *Handbook of Formal Languages*, (Eds.) G. Rozenberg and A. Salomaa, vol. I, 51-105.