2

The Minimalist Programme as a general linguistic framework for Verb Movement

This chapter offers a survey of recent developments in the theory of sentence-structure within the Chomskyan generative framework. The emphasis is on the development of syntactic theory up to and including MP. MP is introduced and explained in two of Chomsky's recent publications, namely *A Minimalist Programme for Linguistic Theory* (1992) and *Bare Phrase Structure* (1994).

The explanation of MP in this chapter relies on the abovementioned papers while also taking account of the language-specific proposals in Broekhuis & Den Dikken (1993), Marantz (1995), Zwart (1993, 1996), and Oosthuizen & Waher (1996).

2.1 Developments in the theory of V movement from *Barriers* to MP

Since the publication of *Barriers* (Chomsky 1986b), there has been a growing awareness among linguists that language variation is probably to a large extent, if not exclusively, determined by functional categories. This awareness is best illustrated by the following quotation from Chomsky (1991: 418): "If substantive elements (verbs, nouns, etcetera) are drawn from an invariant universal vocabulary, then only functional elements will be parametrised". Parametric variation reduces, to a large extent, to variation in the lexical properties of the functional categories involved.

A clear-cut distinction may be drawn between functional categories and substantives in terms of their inherent/lexical properties (Borer 1984, Ouhalla 1990). Substantive categories are the so-called

"open" class. The open class contains the major lexical categories V, N and A.[1] The remaining categories, for example Comp, Infl, Aux, and Det,[2] are referred to as functional categories and belong to the closed class.

A number of properties are thought to distinguish between these two types of categories:

- The ability/inability to assign thematic roles is one of the fundamental respects in which functional categories and substantives differ. Substantives have thematic grids; functional categories do not.
- Functional categories have c-selectional properties,[3] while substantives lack these.
- A third difference lies in the fact that all functional categories have m-selectional (morphological selectional) properties. The m-selectional properties of bound categories include the information that the category in question is affixal as well as specifying the categories of elements to which it can attach/adjoin. Substantives have no m-selectional properties.
- A further lexical property of functional categories is grammatical features, including the ϕ-features of Chomsky (1981) (*viz* person, number and gender), tense features, wh-features and case features, while substantives have no grammatical features.

Typological word order differences can be accounted for in a principled way in terms of minimal parametric differences involving the lexical properties of certain functional categories. Any attempt to clas-

1 N - Noun
N's - Nouns
A/Adj - Adjective.

2 C/omp - Complement
Infl - Inflection
Aux - Auxiliary
Det - Determiner

3 Lexical items c-select (category-select) for their objects and complements. Concerning the verbs *reside* and *inhabit*, location is expressed through a prepositional phrase in the case of *reside* and directly with an NP object in the case of *inhabit*:
i) Anteaters resided in Southern California.
ii) Anteaters inhabited (in) Southern California.

sify languages along typological lines should take into consideration the properties of functional categories rather than those of substantives.

In a comparative study concerning the distribution of the inflected V in French and English, Pollock (1989) proposed the so-called Split Infl-hypothesis. According to this hypothesis, Inf divided into two functional heads, each with its own projection, Agreement (Agr) and Tense (T). See (1):

(1)

Infl should not be considered as one constituent with two different sets of features ([± Tense, ± Agr]). Instead, each of these sets of features is the syntactic head of a maximal projection, AgrP and TP.

The association between the verb stem and its morphology is established through the move of the verb head to the inflexion head position by means of a head to head move.

Other possibilities are noted: the verb may be associated with its morphology by Affix Hopping.[4] Various combinations of affix hopping exist and apply in various languages, namely:

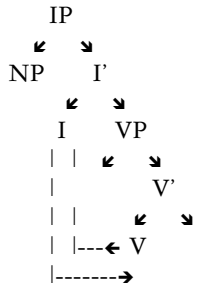
- association through affix hopping, with affixes lowering to the verb stem;
- association partly by affix hopping and partly by V movement, with the verb moving to the first functional head and the rest of the affixes lowering to the same position;
- complete association by means of V movement, with the verb moving to the first functional head position and as a result thereof to the second functional head.

Let us commence with some of Pollock's recent ideas based on work by Emonds (1978) on verbal inflection in English-type and French-type languages. Emonds's basic idea is that in French-type languages, V raises to I, whereas in English-type languages, I lowers

4 Affix Hopping: The affix moves for amalgamation to V, for example: -ing + go becomes go + -ing.

to V, as in (2). That explains why VP-adverbs are preverbal in English and postverbal in French.

(2)

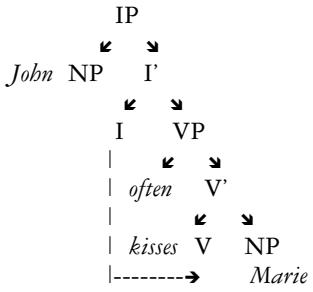


The structures in (5) and (6) indicate the derivation of sentences with VP-adverbs in English (3) and French (4) respectively:

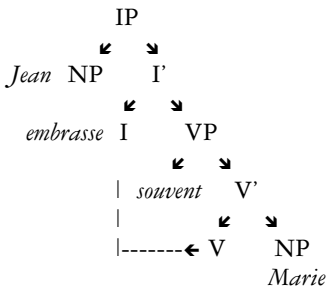
(3) *John often kisses Marie*

(4) *Jean embrasse souvent Marie*

(5)



(6)



According to Chomsky (1991: 422-5) the distinction is not between raising in French and lowering in English, but some other difference that requires French verbs and English auxiliaries to be raised while denying this possibility to other verbs in English.

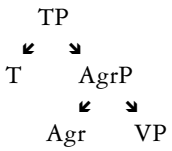
It has been postulated that the Agr element is "stronger" in French than in English. A strong Agr attracts verbs. A morphological feature suggested by Lasnik (1989) pertains here, namely; only a strong Agr can accept a "heavy" element such as a verb, though any Agr can accept a "light" element such as an auxiliary. Another possibility, developed by Pollock (1989) is that the difference reduces to θ-theory: a strong Agr allows an adjoined element to head a θ-chain, but a weak Agr does not.

In the hierarchy of T and Agr where one is relative to the other, several views exist. A number of questions arise about the status of Agr in the Agr-system. Pollock (1989), on the one hand, assumes that Agr is dominated by T. Belletti (1990), on the other hand, indicates that in a number of languages where it is possible to obtain relevant evidence, Agr lies "outside" T in the verbal morphology.

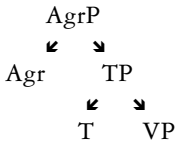
Baker (1985) adduces the Mirror Principle, in which the morphological order (the order of Infl) reflects the syntactic order, and vice versa, also coming to the conclusion that T is lower than Agr. In the context of Baker's explanation, Campbell (1991) argues for a particular view of the distribution and syntactic function of T and Agr, namely: T is a species of Agr, but it agrees with a temporal argument, rather than with the subject of a clause. The properties of T and Agr in English, German and mainland Scandinavian are considered in the light of the syntactic properties of agreement categories in general. The most important specific claim is that while T is lower than Agr in past-tense sentences in English and German, the reverse is true in present-tense sentences.

Thus, (7) is the basic structure of a declarative clause in the present tense, and (8) in the past tense:

(7)



(8)



According to Ouhalla (1990), the order of Agr and T in the derived verbal composition of one language group differs from that in another. Thus, languages divide into two typological groups according to whether Agr occurs "inside" or "outside" T. In languages such as Berber and Arabic, Agr is "inside" T, while the opposite is true of languages such as Italian and Chichewa.

In pursuing the basic lines of Pollock's (1989) analysis, Chomsky (1991) finds two Agr elements: the subject-agreement element AgrS and the object-agreement AgrO. Generally, AgrO should be close to V, and AgrS close to the subject, therefore more remote from V. The development of functional projection caused Chomsky to take a new direction, namely the MP.

2.2 The Minimalist Programme (1992)

The MP continues the trend in syntactic theory which began in the late seventies: the move from specific grammatical rules describing particular syntactic constructions to general principles interacting to explain syntactic phenomena.

A number of conceptual issues lead from standard GB (Government & Binding) theories to the MP:

- Conditions such as "least effort" and "last resort" suggest a striving for the most economical or minimal way of satisfying principles. The MP relies explicitly on such "principles of economy" in evaluating derivations.

- Another aspect of the MP is the impetus towards minimising the number of principles involved. Within Principle and Parameter syntax, increasing weight has been placed on conditions of well-formedness at the so-called "interface" levels of PF and LF. On the LF side, the principle of Full Interpretation (FI) might rule out the presence of "excess" constituents in a structure. On the PF side, FI might reject representations containing symbols without any phonetic realisation.

By means of the elimination of the independent levels of D-structure and S-structure and of principles applying at these levels, the grammar itself may be simplified. The major changes in the move to the MP are the following:

- constituents move for a reason, not freely;
- grammaticality depends on a comparison of derivations, not on the evaluation of a particular derivation in isolation, and
- principles apply only at the interface levels of PF and LF.

The most important similarities and differences between the MP and the previous stages in the theory are the following:

- In the MP as in the previous stage of the theory, the grammar is a derivation system. A sentence is first composed in a basic form and then completed by a process of projecting, merging and moving until the final stage is reached.
- In the MP, as opposed to the previous stage of the theory, the operation that results in the representation, and the operation responsible for the other work (move, merge and project) are the same. This is the operation Generalised Transformation (GT).
- In the MP a move exists because the elements must be formally licensed. In contrast to the previous stage in the theory, only formal licensing is needed for the move to take place. However, elements can never be formally licensed in a position obtained in the first representation.
- In a previous stage of the theory, a move could take place before or after the point of derivation where the instruction is issued after PF system. By contrast in the MP a move takes place rather to the particular point of derivation. An overt move is now only an option.

- In a previous stage of the theory, the number of overt moves was seen as differing from language to language, but in the MP the presence or absence of overt moves is the only instance of parametric variation in syntax between languages.

2.2.1 The Generalised Transformation Operation

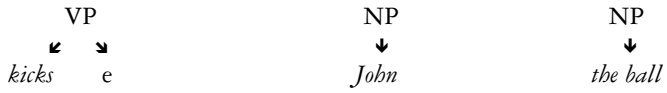
Representations are put together in a tree-diagram by the operation GT. The operations of GT can be illustrated by means of the English sentence in (9):

(9) *John kicks the ball*

The syntactic derivation of sentence (9) starts with a selection of substantive heads — the V "kicks", and the NPs "John" and "the ball", each of which is fully inflected with its particular morphological features (case, tense, and congruence) already added. As a transitive V, "kicks" selects two arguments: one can be accorded the thematic role of Theme and the other that of Agent. The course of the derivation may be set out as follows:

The operation of projection creates a VP with a vacant position (e) which can be filled by an argument accorded the Theme role; the VP and the NP's are independent (10):

(10)



The operation of merging places the NP "the ball" in the vacant VP-position (11):

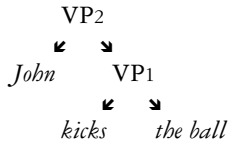
(11)

The operation of projection creates a new VP with a vacant position for an argument accorded the Agent role (12):

(12)

The operation of merging inserts the NP "John" into the vacant VP2-position (13). The NP "John" forms the specifier (Spec), and the NP "the ball" the complement (Comp) of the head "kicks".

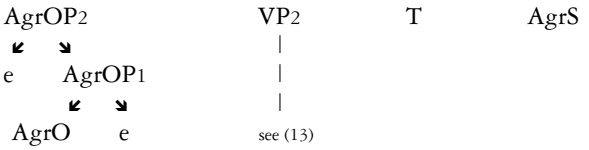
(13)



The licensing of the morphological features of "kicks", "John" and "the ball" in (13) is implied. At least three functional heads are selected, projected and merged with VP2 to finally constitute a single structure.

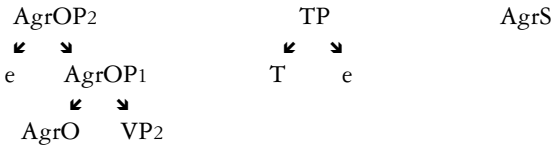
The operation of projection creates an AgrOP1 with a vacant Comp-position, as well as an AgrOP2 with a vacant Spec-position; the structures AgrOP2, VP2, T and AgrS are independent (14):

(14)



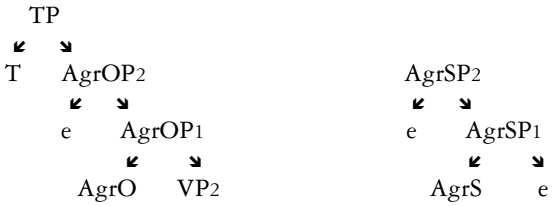
The operation of merging places VP2 in the vacant AgrOP1-position, while the operation of projection creates a TP with an empty Comp-position (15):

(15)



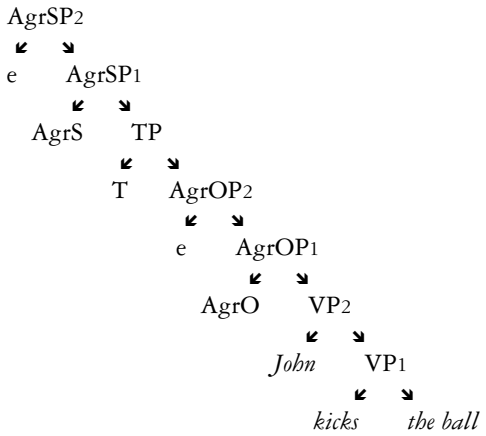
The operation of merging places AgrOP2 in the vacant TP-position, while the operation of projection creates an AgrSP1 with an empty Comp-position as well as an AgrSP2 with an empty Spec-position (16):

(16)



The operation of merging places TP in the empty AgrSP1 position (17):

(17)



The extension of a target constituent (VP), the projection of an empty element, and the combination of the empty element by means of a second phrase-marker are all part of one undivided process. GT always adds material extraneous to existing phrase-markers. According to Chomsky (1992) it is impossible to add material within a phrase marker.

2.2.2 Checking morphological features

At the heart of the MP are what Chomsky calls "morphological features". These include features associated with tense, case and agree-

ment. Items from lexical categories such as V, N and Adj are fully inflected in the lexicon. The addition of tense and agreement morphology, for instance, to a verb in the lexicon involves the simultaneous addition of tense, case, and agreement features.

These features, then, if potentially visible to these interfaces, must be eliminated prior to PF and LF. Failure to eliminate morphological features prior to an interface at which they are visible causes a derivation to "crash" (fail to converge) at this interface.

The functional categories of Agr and T in this system are the loci of agreement and tense features which may check off or eliminate the corresponding features of a verb which moves up and adjoins these categories. Agr and T also contain case and ϕ-features (person, gender, number) which they may check off against features of NP's that raise to Spec-Agr and T. In the MP, these functional nodes never contain items from the lexicon; they are not positions at which inflectional affixes are inserted. Rather, inflectional affixes are attached to items of the lexical categories in the lexicon. The functional nodes of Agr and T serve only to carry the morphological (inflectional) features necessary to check off the features of N and V.

The AgrP above TP is labelled AgrS (for "subject agreement") and the AgrP above VP AgrO (for "object agreement"). However, Chomsky claims that there is no difference between the sorts of features found at these two Agr positions and hence no difference in the category labels of the two positions. That is, there is no intrinsic difference between the features of AgrS and AgrO, only a positional difference between the two Agr nodes. The contrast between the Nom(inative) case usually licensed by AgrS and the Acc(usative) case associated with AgrO results from the raising of T to AgrS and V to AgrO; the cases are associated with features of T and V respectively. This is not to say that the Agr features in AgrS and AgrO will be identical in every sentence. The Agr features in AgrS will include the ϕ-features of the subject NP that is raised to Spec of AgrO. However, the ϕ-features themselves — i.e. the Agr features — will come from the same pool of features and will not be identifiable as subject or object Agr features except by the location of the Agr node in which they reside.

The requirement for checking morphological features causes the movement of lexical elements to positions in the functional domain. Structure (17) is the result of inflected elements licensed outside the lexical domain.

To check its Agr and Tense features, the V will be raised to AgrO, to T, and to AgrS. The corresponding tense and agreement features on T and on Agr are known as the V-features of T and Agr, *ie* the features that check off the features of the V. These functional nodes, as well as the verb, also have N-features, or features that check off features on NP arguments.

2.2.3 Principles

The derivation of a sentence is subject to the principles of economy. The three major principles to be discussed here are Shortest move, Greed and Procrastination.

Shortest move is the most technically specific principle of economy. The basic idea is that a constituent must move to the nearest position of the right kind above its source position. The evaluation of shortest move involves a comparison of all possible derivations, not just those that converge.

Although the MP has no specific S-structure level, there is a point in the computation of a grammatical representation where the derivation splits and heads toward the two interface levels, PF and LF. This point, called "spell-out", determines which movements will affect the utterance of a sentence (those that occur before spell-out) and which will not (those that occur after spell-out, on the way to LF).

Procrastination is a principle which prefers derivations that defer movements until after spell-out, so that the result does not affect PF. Procrastination is evaluated over convergent derivations; in effect, then, a derivation may violate procrastination in order to converge. In English, for example, main verbs do not raise to T before spell-out. This behaviour of English contrasts with French. English verbs thus obey procrastination, waiting until after spell-out to raise to T, which they must do by LF to check off their tense features. French verbs must violate procrastination to ensure convergence. The assumption is that French tense features, unlike those of English, are

"strong" and visible at PF if not checked off. Because English tense features are weak, English may wait until LF to check them off on main verbs. The principle of procrastination states that if a move can wait, it must do so.

The principle of greed states that a constituent may not move to satisfy the needs of some other constituent; movement is motivated for self-centred reasons, to satisfy the needs of the moving constituent. The constituent should move only to check off its own features.

2.2.4 Parametric variation: the strength of features

The question arises here as to why and what ways languages differ in their overt syntax. Within the MP, differences among languages are attributed to differences between the features of their lexical items and specifically between the features of lexical items belonging to the functional categories Agr and T. Features of the functional categories are said to be either "strong" or "weak" with respect to their visibility at the PF interface. The strength of features thus becomes responsible for rendering overt syntactic movement necessary. "Strong" Agr features are visible at PF if they are not checked off before the interface. "Weak" features are not visible.

Chomsky assumes that the significant parametric differences between languages are limited to lexical differences, specifically, to differences in the features of the lexical elements that occupy the functional category nodes. For basic parametric differences in constituent (word) order, Chomsky looks to the Agr and T nodes and to their N-features and V-features.

The N-features are those checked off against an NP in Spec-AgrP and the V-features are those checked off against a V that adjoins a functional head. These features may be either weak (invisible at PF even if unchecked) or strong (visible at PF if unchecked) in any language. For Agr and T independently, there are four possible combinations of weak and strong N- and V-features. Since the strength of these features on Agr is at least conceptually independent of their strength on T, the four combinations of features on Agr can combine freely with the four combinations on T to yield sixteen possible language types, as defined by the strength of their morphological features on Agr and T.

Languages can thus differ parametrically with regard to the strength/weakness of the feature of the functional head. In a language where the head has a strong feature, there is a forced overt syntactic move of an element Y with the same features down to the control domain of X to eliminate the feature before spell-out takes place. On the other hand, in a language where the head X has a weak feature, there is no forced overt move of Y downward. In the MP the move of Y is therefore excluded by the principle of procrastination.

A typical example of the above is the difference between English and French with regard to the locus of the V-move to Infl. Lexical V's in French undergo forced raising to Infl; a position to the left of the adverb, while in English lexical V's are not raised in the overt syntax:

(18)

(a) *Jean embrasse souvent Marie.*

(b) **Jean souvent embrasse Marie.*

(c) **John kisses often Marie*.

(d) *John often kisses Marie*.

The difference between the two languages can now be brought into relation with a parametric difference in the strength of the V-features of the inflected head position which attracts the V. In French the AgrO has a strong V-feature, while in English the AgrO node has a weak feature. In English, the overt syntactic move to AgrO will not be necessary. After spell-out the English V will be raised to AgrO to eliminate the interpretable V-features on LF.

It is a prerequisite of the MP that the strong/weak distinction be the only instance of parametric variation among languages. This implies that parametric variation is limited to functional categories. In a language such as BA it will be necessary to define the strong/weak distinction in features of Agr and T in order to define the order of difference between BA and other languages. More detail about the working of the MP will be provided in the chapters which follow, as the various aspects of the MP relevant to BA will be dealt with less abstractly.