

Journal of Universal Language 14-2
September 2013, 79-112

The Minimalist Program and its New Insight to the Concept of Universal Grammar*

Nasser Al-Horais
Qassim University

Abstract

The prime concern of this paper is to introduce the art of the Minimalist Program to the theory of grammar. Besides presenting the basic principles and techniques that make this new theory

* I would like to express my deepest gratitude to Qassim University for supporting this work. Here I would like to extend my sincerest thanks and appreciation to the University's Rector Prof. Khalid Al-Humodi, and his Vice-Rectorate for Graduate Studies & Research: Prof. Abdulrahman Al-Wasel for having created the "right environment" for scientific research in Qassim University. Also, I extend my thanks to Prof. Yousef Al-Saleem, the Dean of Scientific Research for his unlimited efforts in supporting the scientific research environment in Qassim University. Many thanks also go to Anders Holmberg, Abdelkader Fassi Fehri, and Jeffrey Pool for insightful comments, suggestions, and discussions on an early version of this paper. I am grateful to Carmen Ríos García for her invaluable help during all phases of this work. Special thanks also go to the anonymous reviewers and the editor of this volume for clarifying my ideas and for significant editorial improvements. I am solely responsible for any mistakes or inadequacy.

Nasser Al-Horais
Arabic Language Department, Qassim University,
P.O.Box 6611, Buraidah 51425, Saudi Arabia
Phone: 00966556063595; Email: nasser.alhorais@gmail.com

Received August 11, 2013; Revised August 30, 2013; Accepted September 11, 2013.

moving away from the over-generation and filtering character of its predecessors, this paper discusses some new ideas articulated recently by Chomsky such as changing the function of movement and the Extended Projection Principle (EPP) feature, or proposing new theories such as Phases and Feature Inheritance. These new generative ideas can express the sharp contrast between the conceptions of Universal Grammar in the Minimalist Program and the previous generative theories.

Keywords: Economy Principles, Merge, Agree, Multiple Agree, EPP features, Phases, Feature Inheritance

1. Introduction

The Minimalist Program (MP, henceforward) is a major line of inquiry that has been developing inside Generative Grammar since the early nineties, when it was proposed by Chomsky (1993, 1995). In that time, Chomsky presented the MP as a program, not as a theory, but today, the MP lays out a very specific view of the basis of syntactic grammar that, when compared to other formalisms, is often taken to look very much like a theory that represents a dramatic break with earlier generative theories. Unlike all its predecessors that involve the postulation of ever more complex principles, the MP is a move away from excessive complexity as it strives to create a model of language that eliminates unnecessary steps in the representation of the derivation of a sentence (Chomsky 1995). One of leading ideas of the MP is that the language faculty is optimal realization of interface conditions, and thus, the Strongest Minimalist Thesis (SMT) is that language L is an optimal solution to interface conditions imposed on FL by performance systems, and the linguistic expressions generated by L must be legible to these external systems if they are usable (Chomsky 1995, 2004). Such a new insight leads minimalist researchers to a view of Universal Grammar (UG) that

is essentially different from the view in its predecessors, from which the MP originated.¹ This difference can be captured in the following Chomsky's quot (2007: 4):

Throughout the modern history of generative grammar, the problem of determining the character of FL [Faculty of Language] has been approached "from top down": How much must be attributed to UG to account for language acquisition? The MP seeks to approach the problem "from bottom up": How little can be attributed to UG while still accounting for the variety of I-languages attained, relying on third factor principles? The two approaches should, of course, converge, and should interact in the course of pursuing a common goal.

The aim of current paper is not to discuss the devices and properties attributed to UG in the MP and its predecessors, but to give a clear sketch picture of the key premises of the MP that lead to viewing definitely to the concept of UG by Chomsky and his followers. To this end, the paper is divided into three main sections. Section 2 outlines the most features distinguishing the MP from its predecessors, namely Government and Binding theory (GB), and Principles and Parameters Theory (PPT). The reason is to show that the MP is motivated not only by the search for the explanatory adequacy, but also for a certain level of formal simplicity. Section 3 considers 'the structure-building computation' that is viewed as a series of a number of operations which are considered to be the heart and soul of the MP. Section 4 discusses some new ideas articulated recently by Chomsky such as changing

¹ Some scholars, especially functionalists like Golumbia, believe that while the MP represents a dramatic break with earlier generative theories, it "adopts many of the assumptions and goals of the linguistic research projects that emerged before, alongside, and contrary to Chomsky's own, the ones which have come in the linguistic literature to be called functionalism" (Golumbia 2010: 28).

the function of movement and the Extended Projection Principle (EPP) feature, or proposing new theories such as Phases and Feature Inheritance in order to determine the least “costly” derivation and reduce the computational complexity.

2. The Novelty of the MP

2.1. Economy Principles

The MP is distinguished from its predecessors by its ‘derivational concept’ which provides principles for how an analysis is constructed, rather than providing filtering conditions that constrain output representations (Chomsky 1991). As the MP seeks to determine the least “costly” derivation and reduce the computational complexity, it should provide principles for how an analysis is constructed, rather than providing filtering conditions that constrain output representations. The main derivational constraints, which are considered to be the cornerstone of the MP, are the so-called ‘Economy Principles’ first made explicit in Chomsky (*ibid.*). To make this idea clear, Hornstein et al. (2005: 8) describe principles of economy as the practice of “placing a premium on least-effort notions as natural sources of grammatical principles.” As pointed out by Motut (2010), the hypothesis that UG itself is based on principles that favour more economical operations, derivations, and so on derives from Chomsky (1991: 130):

I think we can also perceive at least the outlines of certain still more general principles, which we might think of as ‘guidelines,’ in the sense that they are too vaguely formulated to merit the term ‘principles of UG.’ Some of these guidelines have a kind of ‘least effort’ flavour to them, in the sense that they legislate against

‘superfluous elements’ in representations and derivations.

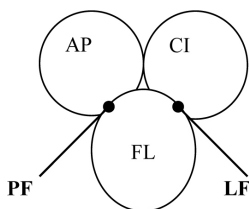
2.2. Levels of Representation

There are two and only two syntactic levels of representation, Logical Form (LF) and Phonological Form (PF). LF is the level of representation that interfaces with the Conceptual Intention system (CI). PF is the interface with the Articulatory-Perceptual system (AP). All conditions on syntactic representations hold at LF and/or PF (Chomsky 1995: 219). This new idea, which is considered to be one of the most significant cornerstones of Chomsky’s Minimalist Program, moves the MP away from the previous syntactic theories of generative grammar. In those theories, a grammar has four distinctive levels: D-Structure (DS), S-Structure (SS), Phonetic Form (PF), and Logical Form (LF), the reason behind proposing that PF and LF are the only available levels of representation, is that these two levels are conceptually required and also empirically sufficient, and, as argued by Chomsky, many of empirical reasons that led to adopting DS and SS can be addressed without postulating any levels other than PF and LF. To justify this reduction and other new assumptions proposed by this approach, he states that (*ibid.*: 168):

The language is embedded in performance systems that enable its expressions to be used for articulating, interpreting, referring, inquiring, reflecting, and other actions. We can think of the Structural Description [i.e., linguistic expression] as a complex of instructions for these performance systems, providing information relevant to their functions. While there no clear sense to the idea that language is “designed for use” or “well adapted to its functions,” we do expect to find connections between the properties of the language and the manner of its use.

Reducing the levels of presentation into the interface levels of PF and LF, however, leads the MP to assume that linguistic expressions, which are the optimal realizations of the interface conditions, where ‘optimality’ is determined by Economy conditions of UG (Chomsky 1993: 4), are generated in the Faculty of Language (FL); the linguistic component of the mind that has interfaces with AP system and LF; CI. This means that form and meaning are represented at these two interfaces (Chomsky 1995, 2007; Zeijlstra 2004). The former is the interface between FL and the AP system and the latter between FL and CI system. This can be represented in the diagrammatic form below (1).

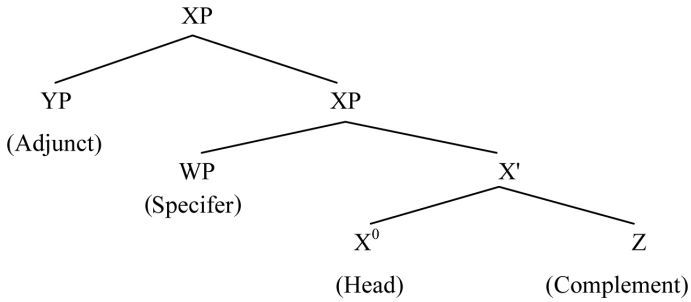
- (1) The linguistic component and its interfaces with other components (adopted from Zeijlstra 2004: 12)



2.3. Bare Phrase Structure

As the MP is a search for simplicity and an attempt to reduce the language specific rules including the principles and parameters, the crucial properties and relations must be stated in the simple and elementary terms of X-bar theory (Chomsky 1995: 172). X-bar theory traditionally asserts each phrase XP has one unique, structurally obligatory element: the head X^0 of the phrase. All phrases have a uniform organization in which heads, specifiers, complements, and adjuncts are structurally defined as shown by the schema in (2).

(2) X-bar structure



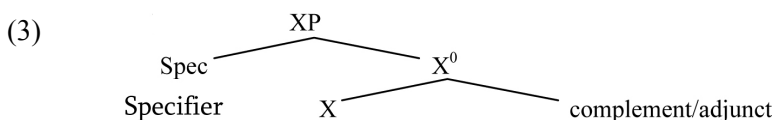
By contrast, the MP eschews all bar levels and deduces their effects from other principles of the grammar (Chomsky 1995). The specifier relation, then, is defined by the same structural relation to the head as a complement. To illustrate this more, Chomsky (ibid.: 241-249) eliminates the X-bar levels in favour of a Bare Phrase-Structure that satisfies *Inclusiveness Condition*² which has been formulated by Chomsky as follows:

. . . any structure formed by the computation . . . is constituted of elements already present in the lexical items selected for N [the numeration]; no new objects are added in the course of computation apart from rearrangements of lexical properties (in particular, no indices, bar levels in the X-bar theory, and so on . . .) (Chomsky 1995: 228, cf. Chomsky 2001: 2).

As for the reason that leads the MP to eliminate X-bar levels, Chomsky (1995: 249) states “ . . . there are no bar levels and no

² *Inclusiveness Condition* can be informally defined as follows: the LF must be built only from the features of the lexical items of N (i.e., no new features are introduced by the computational system) (Radford 2004: 94, Hornstein et al. 2006: 74).

distinction between lexical items and ‘heads’ projected from them. A consequence is that an item can be both an X^0 and an NP.” That is, all projections are the lexical items themselves (3).



What (3) implies is that phrases are composed from a binary relation in which one member serves as both the head and the label for the syntactic constituents. This is due to the fundamental difference between the MP and X-bar theory. The difference is that the MP, as demonstrated earlier, is derivational. That is, it is built from the bottom up, one by one with two main operations, Merge and Move as will be explained later in section 3.1.3. X-bar theory, on the other hand, is representational in which a structure for a given construction is built in one fell swoop, and then the lexical items are inserted into the structure.

An important point should be addressed here is that unlike all its predecessors, this new approach is not at all tolerant of assuming any non-branching phrasal categories. This is not, of course, surprising since binary branching³ forms the basis for the Merge operation (see section 3 below) which is central to the MP. If merge is the only tree-building operation, then it is impossible to have a “phrase” consisting of a single member. Binary branching, however, is the single most important principle limiting the complexity of the grammar in mainstream generative linguistics. “A theory that assumes Binary Branching as a guiding principle (e.g., MP) has the advantage of MINIMIZING the class of possible structures while ensuring that the relations between

³ Binary branching means that the branches split in two daughters and never more than two.

their constituents are unambiguous” (Guevara 2006, emphasis mine).

2.4. Full Interpretation

Full Interpretation (FI) requires all features that pass across the interface to receive an interpretation, and representations be minimal in a certain sense (Chomsky 1995: 130). That is, all features and elements have to get an interpretation at, or be deleted before, the interface levels PF and LF (i.e., no superfluous⁴ ‘uninterpretable’ at the interfaces) (ibid.: 27). For derivation to be *convergent* and *optimal*, FI must be satisfied by the derivation at both LF and PF by containing no uninterpretable features. Otherwise, the derivation crashes (ibid.: 219-220). Accordingly, FI is relativized to the two interface levels (see Kennedy 2000):

- A syntactic expression is PF-interpretable iff it can be assigned a phonological representation (i.e., iff it can “read” by the phonology).
- A syntactic expression is LF-interpretable iff it can be assigned a semantic representation (iff it can be read by the semantics).

The principle of Full Interpretation is perhaps the most widely used by Economy Principles mentioned in section 2.1. In this regard, Chomsky & Lasnik (1993: 23) say the following:

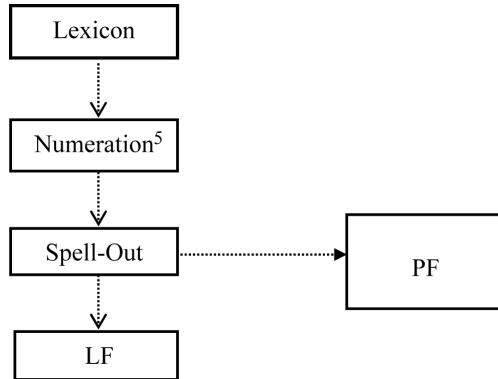
The principles [of UG] have further structure . . . There are also certain general ideas that appear to have wide applicability, among them, principles of economy stating that there can be no superfluous symbols in representations (the principle of Full Interpretation, FI) or superfluous steps in derivations.

⁴ That is, every operation must have a purpose.

3. Syntactic Derivation

The diagrammatic form in (3) below, shows how the syntactic operations derivate within the minimalist framework. In practice, a set of lexical items enter a numeration N , which is a set of pairs (LI, i) , where LI is an item of lexicon and i the number of times that LI is selected from N to be included in a given derivation (Chomsky 1995: 226). It should be noted that each time an item is taken from the numeration, i is reduced by one. At the end of a derivation, the numeration must be empty and every index of every lexical item must be reduced to zero. Otherwise, the derivation crashes (ibid.: 228). At a certain point during the derivation, a derived structure is spelled out. Spell-Out is a technical term meaning that the strictly syntactic, structure-building part of the derivation is completed, at which point the derivation splits and goes off in two directions. On one hand, it is mapped onto the PF-component, to undergo phonological rules (i.e., assimilation, contraction, deletion, and so on), to eventually end up as PF, the representation which is the interface of the grammar with the system controlling articulation and perception. On the other hand, the fully constructed syntactic structure ends up as LF, the interface of the grammar with the cognitive system dealing with meaning (i.e., logical inferences, determining truth, and so on). “Ideally, Spell-Out applies freely and without restriction: if it applies at the wrong point or sends the wrong information to one of the interfaces, the derivation crashes. Spell-Out is not a level of representation that the grammar can refer to” (Kennedy 2000).

(4) Model of grammar (Chomsky 1995: 219)



In the next subsections, I explain, from a minimalist perspective, what these derivational operations that form the syntactic objects are.

3.1. Derivational Operations

In the MP, the derivation, or in more technical terms, “the structure-building computation,” is viewed as a series of a number of operations. It consists of a small set of basic operations: Select, Merge, Agree, Move, and Transfer, aiming at determining the least “costly” derivation in terms of computation.

3.1.1. Select and Merge

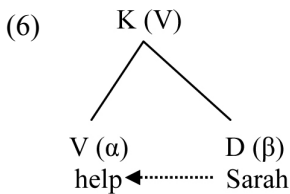
One of the important concepts or the key elements in the MP is that of Select and Merge. The two operations “are necessary

⁵ Chomsky (1995: 226) defines numeration as “a set of pairs (LI, *i*), where LI is an item of lexicon and *i* is its index, understood to be the number of items that LI is selected” to be included in a given derivation.

components of any theory of natural language.” Both operations “are ‘costless’; they do not fall within the domain of discussion of convergence and economy” (Chomsky 1995: 226). Merge can be defined as a binary operation that by Select takes per operational step two constituents “from the numeration [N] and turns them into one constituent that carries the same label as that of the dominating item” (Zeijlstra 2004: 14).⁶ This definition simply reflects the fact that Merge needs at least two arguments to form them into a constituent. The reason behind this is that Merge must be *recursive* (i.e., there can be any number of merge operations) and hence merging two objects is the minimum required to get recursivity. Moreover, merging two arguments meets the requirement that all branching must be binary as explained in section 2.3 above (see Hornstein et al. 2005: 209-210). In technical terms, the operation Merge is defined in (5):

$$(5) \text{ Merge: } K = \{\alpha/\beta \ \{\alpha,\beta\}\}^7 \quad (\text{Zeijlstra 2004: 15})$$

This definition in (5) can be empirically illustrated more by the following example in (6).

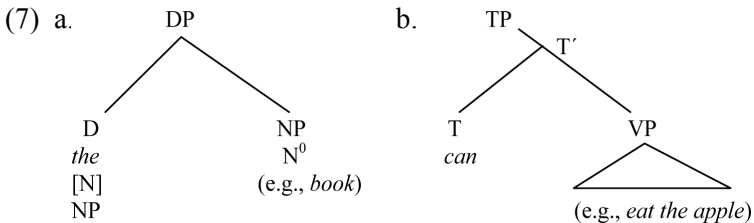


(6) asserts the simplicity of Merge operation, so all what Merge

⁶ The Headedness Principle (Radford 2004: 70): “Every syntactic structure is the projection of a head word.”

⁷ K is a newly-formed constituent that is labeled after its head which can be either α or β as illustrated in (6).

does is taking two elements, say: A and B, and putting them together to create a more complex structure. Accordingly, the MP assumes that it is conceptually necessary that phrases and sentences are built up from words by a series of merger operations. Words have selection features which stipulate which category or categories they can merge with. For instance: *the*, which a head D, has an uninterpretable selectional (N) feature, which signifies that it must combine with a noun or NP to form a DP. The selectional [N] feature of D is deleted upon merger with the NP complement as illustrated in (7a). The modal *can* has a selection feature (V), which must combine with a verb or VP as shown in (7b), and so on (Radford 2004: 58-59).



It is important, however, to note that if selectional steps are taken correctly, merged elements converge, if not, they crash, and then, we have ungrammatical structure.

3.1.2. Agree

One of the integral derivational operations in the MP is that of Agree. It is an operation that “establishes a relation between two elements if they share certain grammatical features” (Kremers 2003: 6). This operation consists of two elements: *probe* and *goal*. The *probe* in order to be able to enter into an agree-relation must be active. It can be active if and only if it has an unvalued feature so that it can value its features by probing for an active *goal* that has the same matching features but valued.

In order to give this operation more substance, Chomsky (1995) made an obvious distinction for syntactic features. In this regard, he divides Syntactic Features into two sorts: those with a semantic interpretation (e.g., a pronoun with the features [3M, SG] refers to different elements than a pronoun with the features [3F, PL]), and those with a purely syntactic function “formal features” (Chomsky 2001: 10). The former features are called interpretable, which enter the computation valued, while the latter are uninterpretable which enter the computation unvalued, but are valued during the computation. Thus, at Spell-Out, all features must be valued. Table 1 below lists the basic uninterpretable and interpretable features:

Table 1. The Basic Uninterpretable and Interpretable Features

Uninterpretable	Interpretable
Φ -features on T, v , C . . . tense features on V case features on DP EPP features (D) on T, C, v , Neg . . .	Φ -features on DPs tense features on T

The way that uninterpretable features capture syntactic dependencies is that they have the following property: an uninterpretable feature must be checked by a matching feature (whether interpretable or not). This means that every syntactic dependency will be triggered by the presence of an uninterpretable feature. The uninterpretable features cannot be given an interpretation at the interfaces and hence they have to be eliminated before semantic representation. The requirement is triggered, as illustrated in section 1, by the interface condition FI in which “there can be no superfluous symbols in representations” (Chomsky 1995: 27).

A good example to show how Agree operates is the example of subject-verb agreement. In this regard, it has been observed that

subject-verb agreement and nominative case go together. This is captured by assuming that agreement and case are assigned under the same Agree operation, where Agree, as mentioned above, can be defined as the relation between a so-called probe (which can be T), searching for a category to value its unvalued phi-features (i.e., an expression it can agree with), a so-called goal (which can be a subject DP). The DP which the probe T needs to value T's unvalued-features⁸ is at the same time the DP that T assigns nominative case to. It should be noted that the DP probed by T is always the closest one which does not already have a valued case-feature. That is usually the subject. Overall, noun and pronoun expressions are case-marked by the closest case-assigner which c-commands them. Accordingly, there is a mutual feature-valuing relation between T and a D/DP: T receives Φ -feature values from D/DP, in return D/DP receives a case-feature value. The operation Agree can be formulated in (8) (Chomsky 2000):

- (8) The relation Agree is established between a probe and goal iff:
- i. the probe has one interpretable and one uninterpretable feature, F and uG, and the goal has the same features but with reversed values for interpretability, uF and G and
 - ii. the probe c-commands the goal and
 - iii. there is no element closer to the probe than the goal with the relevant feature-values.

Once again, it should be recalled that when the Agree relation is successfully established, the uninterpretable features are removed from the narrow syntax “being handed over to morphology/phonology, the derivation of PF” (Holmberg 2005: 7), “as they are phonetic effects” (Chomsky 2001: 5). They cannot

⁸ We mean by Φ -features (or phi-features) the person, number, and gender features of a category.

survive until LF. If, for some reason, they cannot be eliminated in the course of the derivation of LF, the derivation crashes.

3.1.2.1. Agree and Some Constraints on Computation

A pressing issue relates to our discussion about Agree is that if the goal is not active, for instance, by not having unvalued features (i.e., a subject DP with unvalued case feature), the operation agree fails. The probe cannot alternatively go down the derivation and look for another element that can serve as an eligible goal. If the goal tries to do such an operation, this will violate the Defective Intervention Effect Principle which prohibits an establishment of an AGREE relation when a closer but inactive goal intervenes between a probe and another goal in the configuration (9):

- (9) *AGREE (α , γ), α is a probe and β is a matching goal, and β is inactive due to a prior Agree with some other probe⁹ (Hiraiwa 2001: 69).

Moreover, according to ‘Phase Impenetrability Condition’ in (10) (Chomsky 2000: 108, added illustrations from Boeckx & Grohmann 2004: 4), Agree cannot hold between a (root node) probe and a goal within the domain of a lower phase head. In this connection Chomsky (2005: 12) emphasizes that “for minimal computation, the probe should search the smallest domain to find the goal: its c-command domain.” That is, only the phase head and its specifiers are active for Agree.

⁹ That is, “an element β (c-commanding γ and c-commanded by α) blocks the establishment of an Agree-relation between two other elements α and γ even if β itself could not agree with α ” (Boeckx 2003: 17).

- (10) In a phase α with head H, the domain of H [=complement of H] is not accessible to operations outside α [=HP], only H and its edge [=H plus any/all of its specifiers] are accessible to such operations.

To simplify things, the principle in (10) above states that ‘every goal has to have a probe in the phase.’

3.1.2.2. Multiple Agree

Multiple Agree relation is a theory that was articulated by Hiraiwa (2001) in which one head can probe more than one goal at a simultaneous point in the derivation. The theory is proposed as a refined theory of multiple feature-checking. One example to show how multiple agree relation operates is ‘Japanese licensing of multiple nominative case’ in raising constructions by the matrix light verb. Another example is ‘Icelandic licensing of multiple accusatives.’ Hiraiwa (ibid.: 69-70) formalizes Multiple Agree as follows (11):

- (11) MULTIPLE AGREE with a single probe is a single simultaneous syntactic operation; AGREE applies to all the matched goals at the same derivational point derivationally simultaneously. MULTIPLE MOVE (movement of multiple goals into multiple specifiers of the same probe H) is also a single simultaneous syntactic operation that applies to all the AGREED goals.

Based on the formulation in (11), two issues must be demonstrated about this theory. One is that Multiple Agree requires an assistance of the operation Move that is substantively a combination of the operations Copy and Merge as illustrated earlier. This supports the view, suggested by Chomsky (2000), which considers Move to be a ‘reflex’ of Agree. (That is, Move

is preconditioned by Agree.) A second one relates to the nature of Multiple Agree. What Hiraiwa means by Multiple Agree or multiple feature-checking is a single syntactic operation Multiple Agree, not multiple instances of an operation agree.

3.1.3. Move

Move is an operation that is derived from Merge (Chomsky 1995: 348). That because, Move does the following steps: (i) given the syntactic object Σ with the terms K and \hat{a} , Move targets K, (ii) raises \hat{a} , and (iii) merges \hat{a} with K to form the new syntactic object (Chomsky *ibid.*: 250). If merge is not a part of move, it is, then, pure merge.

As illustrated in section 1 above, Move operation is guided by economy conditions, which involve economy of derivation and representation. It always takes the shortest route. At each step of derivation the principle of economy allows only a minimum of transformational activity. Hence, Chomsky (1993) introduced the Minimal Link Condition (MLC) (12) as an economy condition on the operation Move to preclude the longer movement to occur if there is a shorter legitimate movement. As shown in (12) below, this condition concerns about a locality restriction on syntactic movement: Movement of α to a target K is blocked by β , if β is closer to K and could enter the same checking relation.

(12) Minimal Link Condition (MLC):

K attracts α only if there is no β , β closer to K than α , such that K attracts β (Chomsky 1995: 311).

In an early version of the MP, movement is driven by the need for a morphological requirement to be satisfied, and hence, some element, at certain point, is required to move to check some feature in a syntactic structure, and hence movement is crucial in order to “enable a previously uncheckable feature to get checked”

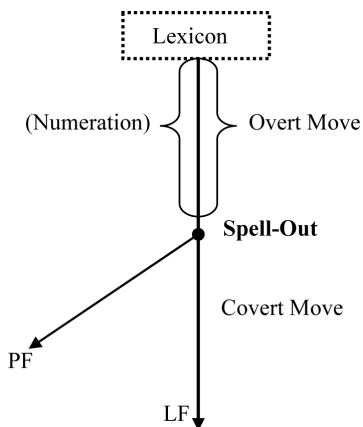
(Chomsky 1995: 261). The features which need checking include structural Case, phi-feature of T and other agreeing categories, and so on as illustrated in Table 1 above. The checking is accomplished when a category needing a feature value is in construction with some other elements in the sentence that can supply that feature value as explained in more depth by operation Agree earlier. This reflects the fact that one significant role of Move is combining Merge and Agree. It merges Y to XP and Y becomes the specifier of XP after the checking features is accomplished by Agree. It, also, serves to allow an element to transfer a feature necessary to satisfy some constraint (Weinberg 1999). In order to give this operation more substance, Chomsky (1993) proposes the two following principles:

- Greed: A constituent does not move unless it has to in order to satisfy some requirement that it has.
- Procrastinate: Movement occurs as last as possible in the derivation.

An important point of detail to note about the nature of this operation, in early version of the MP, is that movement can be occurred prior to Spell-Out or in LF (i.e., after Spell-Out). The former type is called ‘overt movement’ and the head of the chain it creates is pronounced. The latter, however, is called ‘covert movement’ and the tail of the chain it creates is pronounced. Chomsky (1995: 262-265) argues that overt movement is for satisfaction of morphological properties (formal features) such as moving an entire X (head movement) or XP (phrase movement), whereas covert movement would be expected to be restricted to feature raising such as *wh*-movement, expletive replacement, and anaphor raising. Both these two types, however, are maintained in the MP. But since the framework is economy-driven, the overt movement is unwelcome because it is costly in terms of economy conditions. However, the covert movement is preferred since it is

cost-free as shown in (13) below (see Culicover 1997: 350).

(13) Overt and Covert Move



More recently, Chomsky (2000 and subsequent works) dismisses with the previous idea that Move is driven by the feature checking (Chomsky 1995: 253). In the more recent Agree-based framework discussed earlier, movement occurs only to satisfy the EPP feature, whereas Case/agreement are licensed in the subject’s base-position. Accordingly, the EPP is the sole reason for movement (see section 4), since Agree enables other relations to be satisfied without displacing anything. This new insight is further underscored in Chomsky’s discussion of “phases” (see section 5 below).

3.1.3.1. Move and Copy Theory

In this subsection, I introduce the idea that Move is not a primitive operation, but, rather, the combination of the operations Copy and Merge (Hornstein et al. 2005: 214).¹⁰ The Copy Theory

¹⁰ The copy theory of movement indeed involves a form of merger operation by

assumes that a trace of a moved constituent is actually a copy of moved element that reemerges later with another element at the edges of successive phases. That is, movement leaves behind a copy of the moved element, instead of replacing it by an indexed trace. When the narrow syntactic derivation is completed, language specific PF conditions determine which copy is privileged for pronunciation (Chomsky 2000, 2001). Such an assumption made by this theory indeed indicates that the operation Move is simply the sequence of operations Copy and Move.

More significantly, the interesting motivations for treating operation Copy as one of the operation Move components, as well as Merge, had to do with interpretation phenomena. If traces are copies, reconstruction effects may be captured at LF without the need to postulate non-interface levels of representation (Chomsky 1993). Thus, it can be claimed that the Copy Theory of movement provides strong evidence that PF and LF are the only available levels of representation as illustrated in section 2.2 above. Moreover, Assuming a trace is actually a copy of the moved element and hence it is a syntactic object built based on features of the numeration and not a theoretical prime inserted in the course of the computation is compatible with the Inclusiveness Condition which requires that the machinery of syntax does not introduce any new features not already contained in the lexical items as explained in section 2.3 above (Hornstein et al. 2005: 213). Accordingly, since we cannot add anything, we know that the copy of moved element must be something that we got from the lexicon.

To sum up, the copy theory provides an option not available in trace theory, namely that the lower rather than the higher member

which the moved copy that has been merged in one position is subsequently merged in another position. As a result of this, it has been proposed that “remerge is simply a notation for the copy theory as originally formulated in the most elementary terms” (Chomsky 2005: 7, n. 17).

of a non-trivial chain may be phonetically realized. This can be seen by comparing *wh*-movement in English with *wh*-movement in Chinese and Korean, for example. In English the copy in spec CP is pronounced, whereas in Chinese and Korean the copy in the first-merged ‘the original’ position is pronounced (see Hornstein 2001).¹¹

3.1.4. Transfer

Chomsky (2001) introduces the notion of Transfer as an operation that could be constructed to be different from Spell-Out. However, in his late work, in particular of (2004), Chomsky apparently deals with Transfer as another name for Spell-Out as indicated by his definition of Transfer (Chomsky 2004: 107, bracketed illustrations are taken from Chomsky’s discussion):

- (14) TRANSFER hands D [derivation] - NS [narrow syntax] over to Φ [the phonological component] and to Σ [the semantic component].

Grohmann (2007: 102) dissociates the operation Transfer from the operation Spell-Out. He proposes that “transfer takes a sub-part of the derivation and ships it to PF cyclically (where operations like building prosodic domains apply); whereas Spell-Out feeds the sensorimotor system [articulatory-perceptual] once the PF-branch is complete, uniquely (i.e., once the derivation has assembled all Prolific Domains¹²).”

¹¹ Another wide assumed idea is that unlike English, *wh*-movement in these two languages applies later in the covert part of the syntax, after the Spell-Out, thus it does not show in PF (Sabel 2000).

¹² Grohmann (2003: 75): “A Prolific Domain is a contextually defined part of the computational system,
i. which provides the interfaces with the information relevant to the context and
ii. which consists of internal structure, interacting with derivational operations.”

I will hold Chomsky's view that treats both two operations as one operation since this view is compatible with the core assumption of the MP, discussed in the begging of this paper, which requires that syntactic representations formed in the course of a derivation should be as simple as possible, consisting of a minimal number of syntactic operations. Accordingly, I adapt the latest definition of the operation of Transfer, to my knowledge, proposed by Legate (2002: 2) in (15) below:

- (15) 'Transfer' is a macro-operation that ultimately sends the derivation to PF and LF [for interpretation], but that contains a number of prior sub-operations whose application is claimed to be simultaneous. One of these sub-operations eliminates features that were valued during the phase from the derivation proceeding to LF; these features are retained in the derivation proceeding to PF.

4. The EPP Feature

The Extended Projection Principle (EPP), which is connected to an uninterpretable feature, has been played a significant role in syntactic theory ever since Chomsky (1981, 1982, 1986) proposed it. The general strategy in this section is to present a brief rundown on the interesting development of the essence of the EPP feature in the literature, showing the significant role of EPP feature in the center of theorizing within the MP and with the two of its immediate predecessors, namely 'GB' and 'PPT' and indicating, at the same time, that the syntactic theory has had a great difficulty in finding a better understanding for this feature. As concluded by Butler (2004: 1) "EPP has a long and chequered history; its universality and indeed existence have been defended and denied with equal vehemence."

4.1. Background

Within the frameworks of GB and PPT, the EPP feature started out as a universal requirement for clausal subject, requiring that there be a subject in every clause or perhaps that certain functional heads have a specifier (Chomsky 1986 and elsewhere). More specifically, the EPP engendered a specifier position on IP. However, under Minimalism such a requirement is implausible, since specifiers are not obligatory. Instead, two new understandings for the EPP feature have been proposed. In the earliest work of the MP, it was assumed that the EPP is implemented as a [D] feature with a matching feature in nominal expressions.¹³ For instance, this feature can be located at T which is checked as a result of Merge or subject Move into Spec, TP (Chomsky 1995). This means that the EPP is an independent feature hosted by T as well as Case-feature and phi-features. That is, the EPP is a result of a feature-checking requirement which is checked by the subject of finite clauses.

Recently, Chomsky (2000) argues contra the idea of treating the EPP feature as a [D]. Alternatively, he perceives the EPP as a selectional feature, uninterpretable and nonsemantic, satisfied only as a result of dislocation; specifically, movement and re-merge the NP/DP at the spec of TP to check the [EPP] feature on the T head. The reason behind this approach could be that Chomsky wants to expand the role of EPP to be implicated “in a range of other mysteries beyond the necessity of ‘subject’ (like the apparent need for ‘very’ successive cyclic A-bar movement operations)” (Grohmann et al. 2000: 154).

More recently and surprisingly, several works have brought us

¹³ Chomsky (1995: 199): “The Extended Projection Principle, which requires that [Spec, IP] be realized (perhaps by an empty category), reduces to a morphological property of T: strong or weak NP-features.” Later, he adds “the Extended Projection Principle (EPP) plausibly reduces to a strong D-feature of I” (p. 232).

back to Chomsky's original conception of the EPP feature (1981, 1982, 1986) and proposed a restoration of EPP as an original formulation 'Extended Projection Principle' (Chomsky 2000, Holmberg 2000, Lasnik 2001). That is, the EPP property should be treated "as a requirement to have an overly filled specifier" (Bošković 2007: 186), or as summed up by Lasnik (2001: 357) ". . . the EPP has nothing to do with features checking in the sense of Chomsky (1995). Rather, in a return to the earliest view, it is the requirement that certain functional heads must have a specifier." Based on this approach, Chomsky (2000, 2001) reaches the conviction that movement is only driven by the EPP feature. How this? First, although the EPP feature is uninterpretable like phi-features and structural Case, it differs from those two features in being a selectional. Thus, unlike the EPP feature, they never induce movement. Second, according to Agree-based framework discussed in section 3.1.2, Case and agreement are licensed in the subject's base-position. What, actually, checked is only the EPP since it requires "second Merge" (i.e., that something be moved and merged as Spec, TP). "This move has an interesting consequence: the EPP is the sole reason for movement, since Agree enables other relations to be satisfied without displacing anything" (Grohmann et al. 2000: 164). In Chomsky (2000, 2006, 2008), when the notion "phases" has been introduced, the EPP feature has a new name: "Edge feature" (see section 5 below). In next subsection, I shall explain how languages vary in the EPP feature type they require.

4.2. The Universality of EPP

As proposed by Chomsky (1981, 1982, 1995, and subsequent work by others), the EPP may be a universal feature whose realization is parameterized across languages. Nevertheless, languages seem to vary in terms of the EPP feature type they require. According to Alboiu (2000), the EPP feature can be

cross-linguistically divided among languages, into three main EPP feature types: these are [T], [D], and [V]. In T-type EPP languages as in Niuean¹⁴ and probably French (Pollock 1989)¹⁵, the EPP is satisfied by selecting the predicate and merging it as Spec, TP (when the predicate is realized as XP), or as T (when the predicate is realized as X⁰). In D-type EPP languages, such as English, the EPP feature is erased by selecting an agreeing XP (i.e., the subject) and merging it as Spec, TP. Thus, the EPP feature, in this language, is dependent on the probe-goal relation: the category which is selected by the probe T as its goal is also the category which the EPP feature then attracts (a copy of) to Spec TP. In line with Chomsky's (1995) idea that movement is driven by the EPP, The [D] feature, however, is satisfied either by a subject in Spec, IP (Radford 2004) or by moving the finite verb with its nominal features to I (i.e., null-subject languages) (Holmberg 2005, cf. Alexiadou & Anagnostopoulou 1999). In some D-type EPP languages, such as Turkish, this feature is obligatory, otherwise the derivation crashes. That is, in Turkish sentences, Spec, TP must be occupied by a moved DP (i.e., specific nominal) to that position, and in the absence of a specific subject, another nominal is required to move to the spec, TP to satisfy this feature (see Cagri 2005 for more discussion).¹⁶

In V-type EPP languages, however, like Romanian, and perhaps Arabic, the EPP feature selects the lexical verb which always undergoes raising to I⁰. Thus, in Romanian, for example, the EPP

¹⁴ The Niuean language or Niue language (Niuean: *ko e vagahau Niuē*) is a Polynesian language, belonging to the Malayo-Polynesian subgroup of the Austronesian languages (Massam & Smallwood 1997).

¹⁵ It is worth pointing out that French requires verb raising to I⁰, alongside subjects in Spec, IP and expletives. Presumably, this means that French has a 'mixed' type EPP, namely, both a D-type and a V-type EPP feature (Pollock 1989, cf. Alboiu 2000).

¹⁶ Hence, Cagri assumes that in Turkish, the EPP feature of T can be satisfied by a DP, but not by an NP.

feature is assumed “to be equivalent to a strong [V] feature on I^0 . This strong [V] feature attracts verb movement to I^0 , thus ‘activating’ the IP domain” (Alboiu 2000).

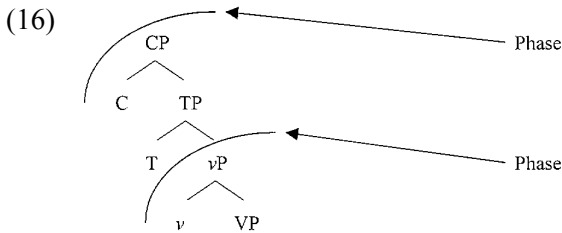
5. Phases and Feature Inheritance

The ultimate concern of this section is to capture the concept of “phases” in the latest version of the MP. Based on his suggestion in (Chomsky 2000), Chomsky (2001: 12, cf. Legate 2003) defines the notion of ‘phase’ as follows, “the phases are ‘propositional [in nature]’¹⁷: verbal phrases with full argument structure and CP with force indicators, but not TP alone or ‘weak’ verbal configurations lacking external arguments (passive, unaccusative)” in (16). From this definition, he assumes that “substantive categories are selected by functional categories: V by a little verb, T by C. If so, phases are CP [including tense and force] and v^*P ¹⁸ [having all θ -roles]”.¹⁹ His justifications for taking CP and v^*P as phases are that CP behaves as a complete clausal complex containing essential elements of the clause (e.g., the force markers, topic, focus markers, and so on) and v^*P represents a complete thematic (argument structure) complex, including a subject in a specifier position. Moreover, the phases can be fronted, extraposed, and serve as response fragments (Chomsky 2000, 2001).

¹⁷ That is, “a phase is the closest syntactic counterpart to a proposition” (Kremers 2003: 9).

¹⁸ To be distinguished from unaccusative v , Chomsky (2005, 2007) marks transitive little v with*.

¹⁹ Legate (2003: 1) provides an interesting definition of Phases according to how they are used in Chomsky’s system: “a phase is a self-contained subsection of the derivation, beginning with a numeration and ending with Spell-Out. At the point of Spell-Out, the complement of the phase-defining head phase is sent to each of the PF and LF components for interpretation.”



However, once the derivation within a given phase has been completed, the phase arguments become impenetrable to further syntactic operations. This results in the Phase-impenetrability Condition (PIC) in (10), repeated here as (17).

- (17) In a phase α with head H, the domain of H [=complement of H] is not accessible to operations outside α [=HP], only H and its edge [=H plus any/all of its specifiers] are accessible to such operations.

As indicated by PIC in (17) above, phases are syntactically independent. “The derivation of a syntactic structure takes place phase by phase” and sends each one separately “to PF to be spelled out. Once it has been spelled out, it can be merged into another syntactic structure, but because it has already been spelled out, it has been stripped of its syntactic information [i.e., no longer accessible to the syntax].” (Kremers 2003: 10). In this way, the edge of a phase is syntactically transparent, while the complement of a phase head is syntactically opaque. To put things differently, the complement of the head of a phase is out of reach for further computations, but its edge is accessible to operations like agreement and movement as the phase heads C and *v contain two types of features: Agree features (Φ -features) and the Edge feature. The latter is the current version of the “generalized EPP” of Chomsky (2000, 2005, 2007), and triggers movements to the specifier position of the phase head.

It should be, however, noted that phases and PIC are synonyms of ‘bounding nodes/governing categories/barriers’ and ‘subjacency,’ respectively, in the pre-minimalist era.

VPs and TPs are excluded to serve as phases,²⁰ because VP lacks a subject and hence cannot be the syntactic counterpart of a proposition. As for TP, T fails to define a phase boundary along with C, although it seems to be “the locus of the Φ -features that are involved in the Nominative-agreement system, and raising of the external argument subject or unaccusative/passive object to SPEC-T” (Chomsky 2005: 9). The question, then, becomes, why T cannot be treated as a phase as well as C and *v*. In addition to violating (PIC) in (17) because T is part of a clause, Chomsky (ibid.) notes that there is antecedent reason to sustain that TP is not a phase. The reason is that Tense and Φ -features, which appear to be determined by T, are, in fact, determined by C. These features are inherited in T from the head of the phase C. The antecedent reason is that “in the lexicon, T lacks these features. T manifests the basic tense features if and only if it is selected by C . . . ; if not, it is a raising (or ECM) infinitival, lacking Φ -features and basic tense. So it makes sense to assume that Agree- and Tense-features are inherited from C, the phase head” (ibid.: 10). The same can be said about the phase head *v** that transmits its Inheritance features (accusative Case and Φ -features) to V as illustrated by (18) below, taken from Al-Horais (2013: 338).

²⁰ Some scholars argue that unaccusative VPs, passive VPs (Svenonius 2004, Hiraiwa 2005) and DPs (Kremers 2003, Legate 2003) are phases as well.

possibly better, than before.” And as a result, the grammar of a language consists of only those rules, categories, and representations which are conceptually necessary for a language to function the way it does to meet the requirement that the derivation *convergent* and *optimal* (Chomsky 1995: 227).

References

- Al-Horais, N. 2013. A Minimalist Approach to the Internal Structure of Small Clauses. *The Linguistics Journal* 7.1, 320-347.
- Alboiu, G. 2000. The Features of Movement in Romanian. Ph.D Dissertation. University of Manitoba.
- Alexiadou, A. & E. Anagnostopoulou. 1999. EPP without Spec, IP. In D. Adger et al. (eds.), *Specifiers: Minimalist Approaches* 93-110. Oxford: Oxford University Press.
- Boeckx, C. 2003. Intricacies of Icelandic Agreement. Ms., University of Maryland & Harvard University.
- Boeckx, C. & K. Grohmann. 2004. Putting Phases into Perspective. [Word file]. Retrieved Dec 10, 2006. Available at URL <http://www.punksinscience.org/kleanthes/papers/bg_ppp.pdf>.
- Bošković, Ž. 2007. A-Movement and the EPP. In Z. Boskovic & H. Lasnik (eds.), *Minimalist Syntax: The Essential Readings* 186-211. Oxford: Blackwell Publishing Ltd.
- Butler, J. 2004. On Having Arguments and Agreeing: Semantic EPP. *York Papers in Linguistics (Series II)* 1, 1-27.
- Cagri, I. 2005. Minimality and Turkish Relative Clauses. Ph.D Dissertation. University of Maryland.
- Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- _____. 1982. *Some Concepts and Consequents of the Theory of Government and Binding*. Cambridge, MA: MIT Press.

- _____. 1986. *Barriers*. Cambridge, MA: MIT Press.
- _____. 1991. Some Notes on Economy of Derivation and Representation. In R. Freidin (ed.), *Principles and Parameters in Comparative Syntax* 417-457. Cambridge, MA: MIT Press.
- _____. 1993. A Minimalist Program for Linguistic Theory. In K. Hale & S. Keyser (eds.), *The View from Building 20: Essays in Honor of Sylvain Bromberger* 1-52. Cambridge, MA: MIT Press.
- _____. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- _____. 2000. Minimalist Inquiries. In R. Martin et al. (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik* 89-155. Cambridge, MA: MIT Press.
- _____. 2001. Derivation by Phase. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language* 53-122. Cambridge, MA: MIT Press.
- _____. 2004. Beyond Explanatory Adequacy. In A. Belletti (ed.), *Structures and Beyond* 104-131. Oxford: Oxford University Press.
- _____. 2005. On Phases. Ms., Massachusetts Institute of Technology.
- _____. 2007. Approaching UG from Below. In U. Sauerland & H. Gärtner (eds.), *Interfaces+Recursion=Language? Chomsky's Minimalism and the View from Syntax-Semantics* 1-30. Berlin: Mouton de Gruyter.
- Chomsky, N. & H. Lasnik. 1993. The Theory of Principles and Parameters. In J. Jacobs et al. (eds.), *Syntax: An International Handbook of Contemporary Research* 506-569. Berlin: Walter de Gruyter.
- Culicover, P. 1997. *Principles and Parameters: An Introduction to Syntactic Theory*. New York: Oxford University Press.
- Golumbia, D. 2010. Minimalism is Functionalism. *Language Sciences* 32, 28-42.
- Grohmann, K. 2003. *Prolific Domains: On the Anti-Locality of*

- Movement Dependencies*. Amsterdam: John Benjamins.
- _____. 2007. The Road to PF. In E. Agathopoulou et al. (eds.), *Proceedings of the 17th International Symposium on Theoretical and Applied Linguistics* 94-104. Thessaloniki: Aristotle University of Thessaloniki, School of English.
- Grohmann, K. et al. 2000. No More EPP. *Proceedings of WCCFL XIX* 153-166. Somerville, MA: Cascadilla Press.
- Guevara, E. 2006. Binary Branching: Evidence from Morphology. Paper presented at the 32th International Generative Grammar Meeting. Firenze.
- Hiraiwa, K. 2001. Multiple Agree and the Defective Intervention Constraint in Japanese. *MIT Working Papers in Linguistics* 40, 67-80.
- _____. 2005. Dimensions of Symmetry in Syntax: Agreement and Clausal Architecture. Ph.D Dissertation. Massachusetts Institute of Technology.
- Holmberg, A. 2000. Scandinavian Stylistic Fronting: How Any Category Can Become an Expletive. *Linguistic Inquiry* 31.3, 445-483.
- _____. 2005. Is There a Little Pro? Evidence from Finnish. *Linguistic Inquiry* 36.4, 533-564.
- Hornstein, N. 2001. *Move: A Minimalist Theory of Construal*. Malden, MA: Blackwell.
- Hornstein, N. et al. 2005. *Understanding Minimalism*. Cambridge: Cambridge University Press.
- Kennedy, C. 2000. Topics in Minimalist Syntax. [PDF file]. Retrieved Nov 10, 2011. Available at URL <<http://www.ling.nwu.edu/~kennedy/Handouts/intro-handout.pdf>>.
- Kremers, J. 2003. The Noun Phrase in Arabic: A Minimalist Approach. Ph.D Dissertation. University of Nijmegen.
- Legate, J. 2002. Phases in "Beyond Explanatory Adequacy". [PDF file]. Retrieved Jan 16, 2010. Available at URL <<http://www.ling.udel.edu/jlegate/beyond.pdf>>.
- _____. 2003. Some Interface Properties of the Phase. *Linguistic*

- Inquiry* 34.3, 506-516.
- Lasnik, H. 2001. A Note on the EPP. *Linguistic Inquiry* 32, 356-362.
- Massam, D. & C. Smallwood. 1997. Essential Features of Predication in English and Niuean. *Proceedings of NELS XXVII* 263-272. Amherst, MA: GLSA.
- Motut, A. 2010. Merge over Move and the Empirical Force of Economy in Minimalism. *Toronto Working Papers in Linguistics (TWPL)* 33, 1-54.
- Pollock, J. 1989. Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry* 20, 365-424.
- Radford, A. 2004. *Minimalist Syntax: Exploring the Structure of English*. Cambridge: Cambridge University Press.
- Sabel, J. 2000. Partial Wh-Movement and the Typology of Wh-Questions. In U. Lutz et al. (eds.), *Wh-Scope Marking* 409-446. Amsterdam: John Benjamins.
- Svenonius, P. 2004. On the Edge. In D. Adger et al. (eds.), *Peripheries: Syntactic Edges and their Effects* 259-288. Dordrecht: Kluwer.
- Weinberg, A. 1999. Minimalist Theory of Human Sentence Processing. Retrieved June 10th, 2012. Available at URL <<http://www.umiacs.umd.edu/users/weinberg/lamp-024.html>>.
- Zeijlstra, H. 2004. Sentential Negation and Negative Concord. Ph.D Dissertation. University of Amsterdam.