

Linguistics in Potsdam | 25

Hans Broekhuis, Ralf Vogel (eds.)

Optimality Theory and Minimalism:  
A Possible Convergence?

Linguistics in Potsdam 25 (2006)

Hans Broekhuis, Ralf Vogel (eds.):

Optimality Theory and Minimalism: A Possible Convergence?

Series editors: Heiner Drenhaus, Ruben van de Vijver, Ralf Vogel

— LiP issues do not appear according to strict schedule —

© 2006

ISSN: 1616-7392

ISBN:

## Preface

This issue of *Linguistics in Potsdam* contains a number of papers that grew out of the workshop *Descriptive and Empirical Adequacy in Linguistics* held in Berlin on December 17-19 December, 2005. One of the goals of this meeting was to bring together scholars working in various frameworks (with emphasis on the Minimalist Program and Optimality Theory) and to discuss matters concerning descriptive and empirical adequacy. Another explicit goal was to discuss the question whether Minimalism and Optimality Theory should be considered incompatible and, hence, competing theories, or whether the two frameworks should rather be considered complementary in certain respects (see <http://let.uvt.nl/deal05/call.html> for the call for papers). Five of the seven papers in this volume directly grew out of the oral presentations given at the workshop. Although Vieri Samek-Lodovici's paper was not part of the workshop, it can also be considered a result of the workshop since it pulls together some of his many comments during the discussion time. The paper by Eva Engels and Sten Vikner discusses a phenomenon that received much interest from both minimalist and optimality theoretic syntax in the recent years, Scandinavian object shift. The paper may serve as a practical example for a claim that is repeatedly made in this volume: minimalist and OT analyses, even where they might be competing, can fruitfully inform each other in a constructive manner, leading to a deeper understanding of syntactic phenomena.

Hans Broekhuis

Ralf Vogel

Leiden and Potsdam, Dec 18, 2006



## Contents

Pius ten Hacken:

The Nature, Use and Origin of Explanatory Adequacy . . . . . 9-32

Jane Grimshaw:

Last Resorts and Grammaticality . . . . . 33-41

Alona Soschen:

Natural Law: The Dynamics of Syntactic Representations in MP 43-76

Vieri Samek-Lodovici:

Optimality Theory and the Minimalist Program . . . . . 77-97

Ralf Vogel:

The Simple Generator . . . . . 99-136

Hans Broekhuis:

Derivations (MP) and Evaluations (OT) . . . . . 137-193

Eva Engels and Sten Vikner:

An Optimality-Theoretic Analysis of Scandinavian Object Shift  
and Remnant VP-Topicalisation . . . . . 195-231

# The Nature, Use and Origin of Explanatory Adequacy

*Pius ten Hacken*

Swansea University

If we want to compare the explanatory and descriptive adequacy of the MP and OT, the original definitions by Chomsky (1964) are of little direct use. However, a relativized version of both notions can be defined, which can be used to express a number of parallels between the study of individual I-languages and the language faculty. In any version of explanatory and descriptive adequacy, the two notions derive from the research programme and can only be achieved together. They can therefore not be used to characterize the difference in orientation between OT and the MP. Even if 'OT' is restricted to a particular theory in Chomskyan linguistics (to the exclusion of, for instance, its use in LFG), it cannot be said to be stronger in descriptive adequacy than in explanatory adequacy in the technical sense of these terms.

*Keywords: Levels of adequacy, Chomskyan linguistics, research programme, Minimalist Program, Optimality Theory*

## 1 Introduction

In the Call for Papers of the workshop on Descriptive and Explanatory Adequacy in Linguistics, the organizers assume an opposition between the Minimalist Program (MP) and Optimality Theory (OT) as formulated in (1).

- (1) a. The MP is strong in explanatory adequacy, but struggles to get a sufficient degree of descriptive adequacy.
- b. OT is strong in descriptive adequacy, but struggles to get a sufficient degree of explanatory adequacy.

In this paper, an analysis of the term *explanatory adequacy* and the correlated term *descriptive adequacy* will be proposed so that the statements in (1) can be seen in the proper perspective.

**Linguistics in Potsdam 25 (2006): 9–32**

Hans Broekhuis and Ralf Vogel (eds.):

Optimality Theory and Minimalism: a Possible Convergence?

©2006 Pius ten Hacken

## 2 The historical origins of *explanatory adequacy*

The first time *explanatory adequacy* was used as a term in generative linguistics was in Chomsky's address to the 1962 International Congress of Linguists, published also as Chomsky (1964). Chomsky compares two types of device that grammars may be thought to model, one for language processing and one for language acquisition (1964:26). He refers to the second one, represented in Fig. 1, as "(1b)" in the quote in (2).

- (2) a. a grammar that aims for observational adequacy is concerned merely to give an account of the primary data (e.g., the corpus) that is the input to the acquisition device (1b);  
 b. a grammar that aims for descriptive adequacy is concerned to give a correct account of the linguistic intuition of the native speaker; in other words, it is concerned with the output of the device (1b);  
 c. and a linguistic theory that aims for explanatory adequacy is concerned with the internal structure of the device (1b); that is, it aims to provide a principled basis, independent of any particular language, for the selection of the descriptively adequate grammar of each language. [Chomsky (1964:29)]

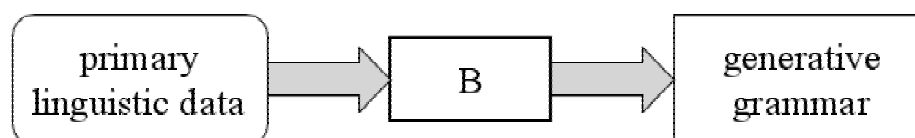


Fig. 1: Model of language acquisition

The approach to explanatory adequacy in (2) is to contrast it to observational and descriptive adequacy. Although (2) defines the three levels of adequacy in terms of Fig. 1, for a full understanding it is useful to refer also to the research programme of Chomskyan linguistics as represented in Fig. 2 (adapted from ten Hacken (2006:582)).

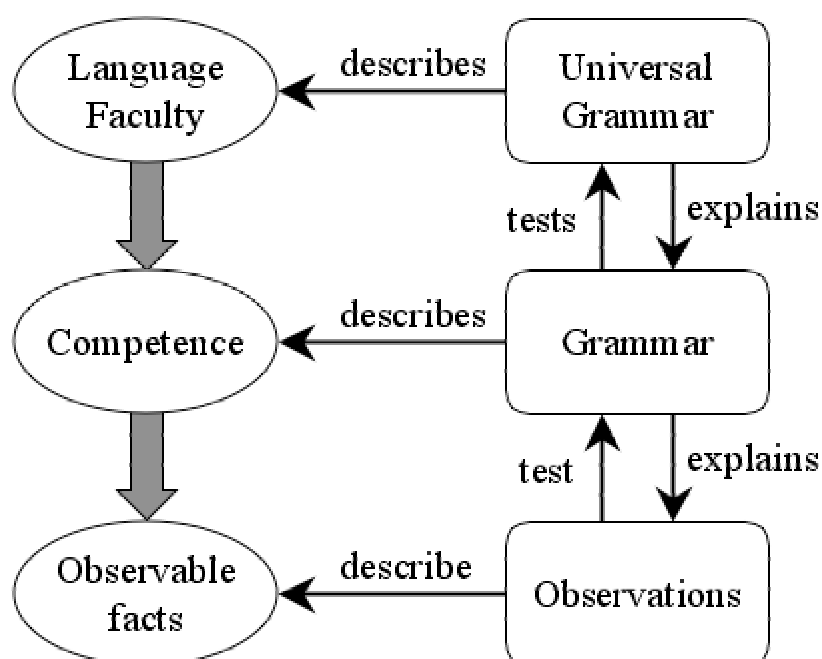


Fig. 2: The research programme of Chomskyan linguistics

In Fig. 2, the boxes on the left represent real-world phenomena and the boxes on the right theoretical constructs. Instead of the widespread ambiguous use of *grammar* and *universal grammar* (UG), these terms are here reserved for the theoretical concepts only. For the corresponding real-world items, *competence* (or *I-language*) and *language faculty* are used. The downward arrows indicate that the higher entity *underlies* the lower. This means that it is essential for its origin without determining all of its nature. In line with modern theory of cognition, e.g. Jackendoff (1989), it is assumed that observations are theory-driven constructs based on real-world facts. Therefore, instead of *data*, Fig. 2 depicts the observation of facts at the lowest level, where the facts belong to the outside world and the observations to the theoretical domain.

It is straightforward to match the three levels of adequacy in (2) with the three levels of theory in Fig. 2. Observational adequacy corresponds to a correct account of the observable facts, descriptive adequacy corresponds to a correct



account of the grammatical competence, and explanatory adequacy corresponds to a correct account of the language faculty.

### 3 The nature of explanatory adequacy in GB-theory

When Chomsky introduced the levels of adequacy in (2), his main argument was that observational adequacy is not an interesting goal to aim for (1964:52-55). There is an interesting relationship between the two higher levels of adequacy, hinted at in (3).

- (3) It is not necessary to achieve descriptive adequacy before raising questions of explanatory adequacy. [Chomsky (1965:36)]

Whereas (3) only states the temporal relationship as “not necessary”, the actual relationship is even stronger. A well-known theorem of mathematical linguistics is that for any finite set of data, an infinite set of context-free grammars can be devised. In the absence of any further evidence, we cannot even assume that the range of grammars to be considered is restricted to context-free grammars. Applying this insight to Fig. 2 raises a serious problem of indeterminacy. How can we discover which of the many possible grammars is the correct one? A central insight of Chomskyan linguistics depends on the analysis in (4).

- (4) The fundamental fact that must be faced in any investigation of language and linguistic behavior is the following: a native speaker of a language has the ability to comprehend an immense number of sentences that he has never previously heard and to produce, on the appropriate occasion, novel utterances that are similarly understandable to other native speakers. The basic questions that must be asked are the following:
1. What is the precise nature of this ability?
  2. How is it put to use?
  3. How does it arise in the individual?
- [Chomsky & Miller (1963:271)]

The three questions listed in (4) have often been repeated from the 1980s onwards, but usually in a different order. The reason is that Chomskyan linguistics uses question 3 in (4) as its primary tool to attack the indeterminacy problem, whereas question 2 is epistemologically side-tracked. This means that question 1, the basis for descriptive adequacy, can only be answered properly by simultaneously answering question 3. This is necessary to solve the indeterminacy problem. Therefore, (3) can be strengthened to the effect that descriptive adequacy can *only* be achieved as far as explanatory adequacy is achieved at the same time.

Questions 1 and 3 in (4) generate a tension because describing the observed I-languages in response to question 1 is easier with a more powerful grammar formalism, whereas explaining learnability in response to question 3 is easier with stronger constraints on the power of the grammar formalism. Question 2 is side-tracked in the sense that it does not have a role in constraining the theory. The only epistemological role assigned to the use of language is to produce the PLD in Fig. 1.

The transition from Chomsky's (1965) Standard Theory to Chomsky's (1981) GB-theory marks the solution of the tension between the two central questions. The difference in attitude is illustrated by the quotations in (5) and (6).

- (5) As a long-range task for general linguistics, we might set the problem of developing an account of this innate linguistic theory that provides the basis for language learning. [Chomsky (1965:25)]
- (6) What seems to me particularly exciting about the present period in linguistic research is that we can begin to see the glimmerings of what such a theory might be like. [Chomsky (1981:4)]

In (5), question 3 is set as a “long-range task”, whereas in (6) an answer to this question has come in sight. In Standard Theory, competence was described in

terms of rewrite rules and transformations. Individual grammars were not learnable on the basis of PLD and the few universals that had been identified. In GB-theory, the Principles and Parameters P&P model managed to relax the tension between the constraints of expressivity and learnability. In fact, the tension was relaxed to such an extent that it lost its original epistemological significance. By adding parameters, expressivity could be increased without immediately affecting learnability.

In conclusion, we can say that the attempt to operationalize explanatory adequacy as formulated in (2c) led to a more advanced framework that allows for a deeper explanation of the data. In doing so, it opened up a new range of questions. It would be wrong to say that the P&P model caused the loss of the tension between descriptive and explanatory adequacy. Instead, the indeterminacy problem re-emerged at a deeper level.

#### **4 The use of explanatory adequacy in the MP**

The transition from GB-theory to the MP can be considered from two perspectives. For the syntactician, the transition consists above all of the replacement of a constraint-based approach by an approach based on economy. Whereas in GB-theory, move  $\alpha$  is constrained by principles, in the MP movement must be motivated. This revolutionizes the way syntacticians formulate their accounts for linguistic phenomena, but it does not directly affect the way these accounts are interpreted in terms of Fig. 2.

The MP can also be seen as an attempt to tackle the second-order indeterminacy problem raised by the solution of the tension between descriptive and explanatory adequacy in Fig. 2. The central idea in this perspective is that the MP reaches for a deeper level of explanation. An important step in this effort

is the formulation of additional research questions. A new, extended list of questions is given in (7).

- (7) a. What exactly are these properties of things in the world?
- b. How do they arise in the individual
- c. and the species?
- d. How are they put to use in action and interpretation?
- e. How can organized matter have these properties (the new version of the unification problem)? [Chomsky (1993:46)]

The context of (7) is a presentation of the goals of linguistics as compared to other, more established sciences. The “properties of things in the world” in (7a) refers to I-language as a component of the human brain. The three questions in (4) are then (7a), (7d), and (7b), respectively. Two new questions are formulated. (7e) was added as a fourth question to the list by Chomsky (1988). In this more accessible work, he formulates it as “What are the physical mechanisms that serve as the material basis for this system of knowledge and for the use of this knowledge?” (1988:3), i.e. how is language realized in the brain. The fifth question, (7c), concerns the evolutionary origin of the language faculty.

In the same way as among the questions in (4), Chomsky selects one of the questions to determine the nature of explanation he wants to add to his framework, whereas the other question is epistemologically side-tracked. As also suggested by the order of questions in (7), the evolutionary origin of the language faculty is chosen to extend the model and the realization of the language faculty in the brain is treated in much the same way as the question of language use. The result can be represented as in Fig. 3.

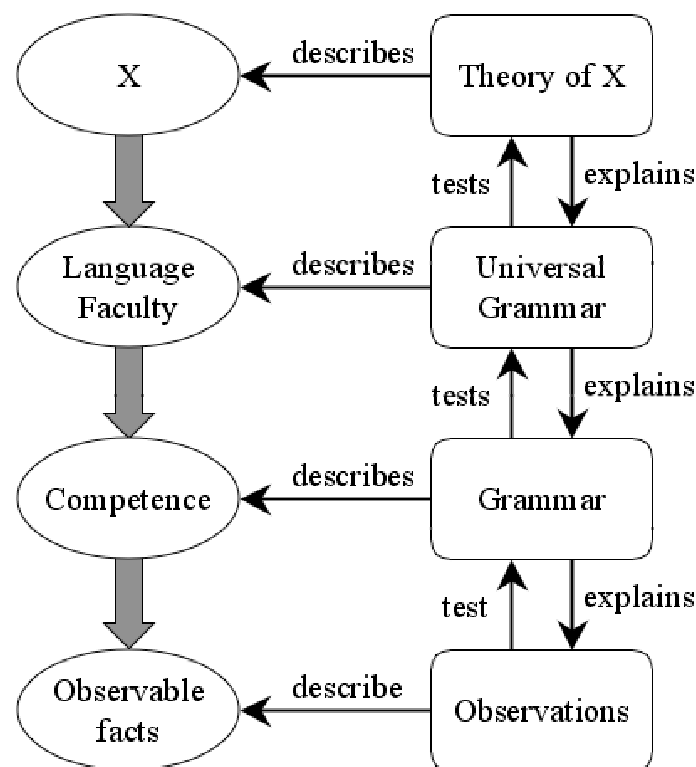


Fig. 3: Extended research programme of Chomskyan linguistics

The entire model of Fig. 2 is part of the model in Fig. 3. The extension adds a level with an entity *X* underlying the language faculty and a theory of *X*. In the same way as the language faculty underlies the competence in the sense that it is at its origin, the new real-world entity underlies the language faculty in the sense that it is at its origin. The two top levels in Fig. 2 represent language in the individual and the species. What underlies the origin of the language faculty in the human species must be a set of general biological principles. Since in the course of the twentieth century biology has been unified with chemistry and chemistry with physics, this level is no longer part of linguistics proper but belongs to the sciences in general. This the reason Chomsky (1993) refers to the “new version of the unification problem” in (7e).

At this point in time, we do not know what the real-world entity underlying the language faculty is. By extrapolating from Fig. 2, however, we can derive a number of relevant properties. An entity of this type is ideally

described by a theory of the appropriate kind. Of course, we do not have such a theory yet. We know, however, that it will explain Universal Grammar in a way parallel to how UG explains individual grammars. Conversely, UG can be used to test this new, high-level theory as it is one of the phenomena covered by it. We can use this additional level without knowing in any detail what the real-world entity and the theory describing it are, because in linguistics we are only interested in the effects they have on the language faculty and UG.

This extension of the model has repercussions for the discussion of the levels of adequacy. As formulated in (2), the levels of adequacy correspond to the levels of theory in Fig. 2. Since Fig. 2 is entirely subsumed in Fig. 3, GB-theory and the MP are equivalent in terms of (this variety of) explanatory adequacy. Explanatory adequacy is achieved if UG correctly describes the language faculty. The difference between the GB-theory and the MP can only be expressed by naming the type of adequacy corresponding to the additional level in Fig. 3. The two main possibilities how this can be done are formulated in (8).

- (8) a. Extending the three levels of adequacy in (2) by adding a new, higher level, e.g. *unificational adequacy*.
- b. Relativizing the opposition between descriptive and explanatory adequacy to the level of theory at which it applies.

Whereas (8a) is an adequate way to dispose of the naming problem, it does not add any insight. Therefore, option (8b) will be pursued here in order to explore the insight that can be gained by it.

## 5 Relativized explanatory adequacy

The general idea of relativized explanatory adequacy is that the opposition between descriptive and explanatory adequacy is formulated without referring to a particular level. This enables us to apply it to any level, highlighting the similarities and parallels between the individual levels.

### 5.1 The nature of relativized explanatory adequacy

Explanatory and descriptive adequacy can be relativized with respect to the level of application by the definitions in (9). These definitions refer to the underspecified architecture in Fig. 4, and use the notion of *level* as defined more formally in (10).

- (9) a. Descriptive adequacy is achieved relative to level  $i$  iff the theoretical entity at level  $i$  describes the real-world entity at level  $i$  adequately.
- b. Explanatory adequacy is achieved relative to level  $i$  iff there is a theoretical entity  $t$  at level  $i + 1$  such that  $t$  adequately describes the real-world entity at level  $i + 1$  which underlies the real-world entity at level  $i$ .
- (10) a. At each level  $i$ , a theoretical entity  $t_i$  describes a real-world entity  $r_i$ .
- b. At level 0,  $r_0$  are observable facts and  $t_0$  observations.
- c. For each  $r_i$  and  $r_{i+1}$ ,  $r_{i+1}$  underlies  $r_i$ .

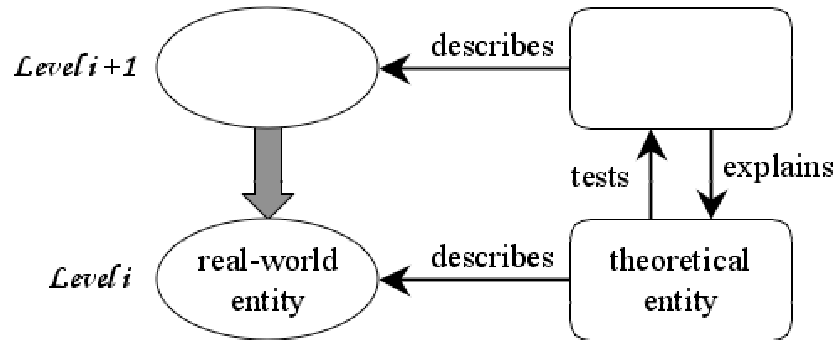


Fig. 4: Architecture for relativized descriptive and explanatory adequacy

In Fig. 4, levels  $i$  and  $i + 1$  are represented, but only for level  $i$  all elements are named, thus highlighting that this is the level with respect to which descriptive and explanatory adequacy are expressed. Informally stated, relativized descriptive adequacy concerns the arrow labeled “describes” in Fig. 4 and relativized explanatory adequacy the arrow labeled “explains”. However, whereas descriptive adequacy can be expressed directly as a relationship between the two elements at either end of the arrow as in (9a), explanatory adequacy has to refer to all four elements in Fig. 4. As expressed in (9b), the

identification of the element from which the “explains” arrow starts is not possible without mentioning the real-world entities at both level  $i$  and level  $i + 1$ .

The simplest application of (9) is the case where  $i = 1$ . In that case, the real-world entity in Fig. 4 is the competence in Fig. 2 and the theoretical entity the individual grammar. For  $i = 1$ , descriptive and explanatory adequacy according to (9) correspond directly to the concepts in (2b-c). Note that it is not necessary to specify the entities at level  $i + 1$ . Explanatory adequacy at level 1 requires that the competence is learnable in the way it is described by the grammar. It is a property of the description of the competence.

If  $i = 0$ , the focus of attention is the linguistic facts. Descriptive adequacy means that the facts are described correctly. Explanatory adequacy means that they are described such that they can be produced by an underlying competence. The difference from observational adequacy in (2a) is above all that level-0 descriptive and explanatory adequacy are properties of the way the data are treated, whereas observational adequacy is a property of the grammar.

The scope of the model can be extended to the MP when  $i = 2$ . At this level, descriptive adequacy means that the language faculty is described correctly and explanatory adequacy that it is described such that its emergence in the course of evolution was possible. It is essential to see the difference between level 2 descriptive adequacy and level 1 explanatory adequacy. Level 1 explanatory adequacy is a property of the competence, but level 2 descriptive adequacy is a property of the language faculty. An idealization such as instantaneous language acquisition is perfectly reasonable in the context of level 1 explanatory adequacy, but much less so in level 2 descriptive adequacy. At level 2, explanatory adequacy can be achieved without specifying the overarching theory at level 3. Only the influence on the emergence of the



language faculty should be specified, and this only to the extent that the logical problem of its emergence can be solved.

## **5.2 The use of relativized explanatory adequacy**

The reason why the concept of relativized explanatory adequacy is interesting is that there are interesting parallels between the applications to different levels. There are at least two areas where such parallels can be observed. First, at different levels there is an opposition between logical and practical problems. Second, at different levels there is a tension between descriptive and explanatory adequacy.

The opposition between a logical problem and a corresponding practical problem was identified by Hornstein & Lightfoot (1981) for (first) language acquisition. The practical problem is the version that is recognized in real life. The logical problem makes abstraction from a number of factors that complicate the problem. A formulation of the practical problem may seem simpler, because many factors are added to it by common-sense knowledge. The logical problem is how the child manages to construct a highly complex system on the basis of restricted input. The simple answer is to refer to the language faculty. A more interesting answer is a description of the properties of the language faculty that make first language acquisition possible by filling the logical information gap between the input and the acquired competence.

A similar opposition between logical and practical problems can be observed in various other areas. White (2003:22-56) discusses the logical problem of second language acquisition. Here there is also an information gap, but this time between on the one hand the L2 input (naturalistic or teaching) and the L1 competence, and on the other hand the interlanguage competence. This gap is less impressive than in the case of first language acquisition, because the second language learner has access to more types of data, e.g. explicit teaching

and negative evidence, and in general achieves a lower level of competence. Nevertheless, White argues that it cannot be bridged without assuming the involvement of the language faculty. An important class of evidence is those cases where the learner's interlanguage competence diverges in parameter settings both from the L1 competence and from the I-languages in the L2 speech community. Interlanguage competence and its acquisition therefore provide further evidence for the nature of the language faculty.

A third example where logical and practical problems have been distinguished is in the discussion of language change. Roberts & Roussou (2003:9ff.) see the logical problem as the question of how change can ever take place. Their answer is that change takes place when the parameter setting of the originating I-language cannot be obtained on the basis of the input the child gets in language acquisition.

If we compare these three logical problems, the last one stands out as not involving an information gap. Although it is a logical problem in the sense that it makes abstraction of certain superficial observations in order to make research in a particular area relevant to the study of the language faculty, it does so in a different way. As opposed to the first two it does not take the form of using a gap between input conditions and observed output to measure the contribution of the language faculty. The main epistemological difference between the first two problems is that first language acquisition is crucial for the existence of language in a way second language acquisition is not. Only first language acquisition is a problem that directly renders the architecture in Fig. 2. The other two provide external evidence.

The three problems considered so far all concern explanatory adequacy of level 1. Let us now consider pairs of logical and practical problems at other levels. At level 0, the question is how to account for the linguistic facts. The

logical problem is that of linguistic creativity. This is of course a well-known subject in early generative grammar, corresponding to Chomsky's (10).

- (10) The most striking aspect of linguistic competence is what we may call the 'creativity of language', that is, the speaker's ability to produce new sentences, sentences that are immediately understood by other speakers although they bear no physical resemblance to sentences which are 'familiar'. [Chomsky (1966:4)]

Chomsky does not explicitly distinguish a logical and a practical problem in (10), but there are clear parallels with the logical problem of language acquisition. There is an information gap between the input and the performance of speakers. This can be solved by assuming the existence of linguistic competence that transcends the input. In the context of (10) it is essential to distinguish the problem under consideration from the problem of the creative use of language, i.e. the use of linguistic competence, which, as Chomsky states, "still seems to elude our understanding" (1975:77). The creative use of language can be seen as the practical problem corresponding to the logical problem in (10). It is not accidental that linguistic competence is at the origin of creative use in the same way as the language faculty is at the origin of the competence.

When we pursue this line of reasoning to level 2, we expect to find a logical problem pertaining to the origin of the language faculty. In current discussion, the origin is interpreted as the evolutionary origin. In generative linguistics, we find two main positions as to the evolutionary origin of human language. One, represented by Jackendoff (2002:231-264), analyses the evolution as a succession of steps resulting in the language faculty as the human species has it now. This results in intermediate stages of which at least one receives a name, *protolanguage* (2002:238).

The other, represented by Hauser et al. (2002), approaches the evolution of the language faculty analytically in the same way as Chomsky approached the

problem of (10) analytically. Chomsky (1980:224-226) analyses what underlies linguistic performance into a number of interacting modules. One of them is grammatical competence. Another is what he calls “pragmatic competence” which “places language in the institutional setting of its use, relating intentions and purposes to the linguistic means at hand.” (1980:225). Yet other components, such as free will, remain outside the domain of analysis. Hauser et al. (2002) argue that the language faculty should also be analysed into components. They distinguish the *Faculty of Language in the Narrow sense* (FLN), as the main object of attention, as well as a number of other components that belong to the *FL in the Broad sense* (FLB), e.g. the conceptual-intentional module and the sensory-motor module (2002:1570-1). FLN is “the abstract linguistic computational system alone” (2002:1571), whereas the other two named components of FLB are interfaces to sound and meaning.

At first sight it is tempting to see the difference between the two views of evolution as one between a logical and a practical problem of language evolution. There are various reasons to be sceptical of such an approach, however. The logical problem of language acquisition makes the idealization of instantaneous acquisition, whereas more realistic studies of the process, cf. Guasti’s (2002) overview, would never make such an idealization. Hauser et al. (2002) also propose a kind of instantaneous evolution, based on their analysis of the transition from a species without to a species with the language faculty in (11).

- (11) For example, suppose we adopt the conception of hypothesis 3, oversimplifying radically, that the interface systems—sensory-motor and conceptual-intentional—are given, and the innovation that yielded the faculty of language was the evolution of the computational system that links them. [Hauser et al. (2002:1578)]

In (11), “hypothesis 3” refers to the hypothesis that only FLN distinguishes human language from communication systems used by other species. Even though (11) is formulated as an example of a possible supposition, the status of instantaneous evolution is not that of an idealization but of a hypothesis. Ten Hacken (to appear) gives a more elaborate analysis of the difference between the two analyses and the discussion between their proponents.

Let us now turn to the parallels between different levels in the tension between (relativized) descriptive and explanatory adequacy. For level 1, this tension is explained in section 3 above. The central properties are that expressivity and learnability exert opposite forces on the power of the constraints governing the way competence can be described. This tension was resolved by the new P&P framework.

At level 0, the tension concerns the way linguistic facts are described. Descriptive adequacy requires that all grammaticality judgements and other data can be accounted for. Explanatory adequacy means that this account has to be in terms of regularities rather than lists. This tension was solved at the start of generative linguistics when Chomsky proposed to use rewrite rules and transformation rules as a way to describe the mental component underlying these data. While it may seem trivial now, this tension is what makes observational adequacy as defined in (2a) not worthwhile as a goal.

At level 2, the tension between descriptive and explanatory adequacy concerns the language faculty. Expressivity now means having enough parameters available to describe the differences between I-languages. The constraining factor is that the language faculty must have emerged in the course of evolution. There are two approaches to making the evolution operational as a criterion. In Jackendoff’s (2002) approach, each intermediate stage is motivated by competitive advantages compared to the preceding stage. In Hauser et al.’s (2002) approach, the essential property that makes the language faculty

operational is a formally relatively simple element, FLN, for which evolution by exaptation is conceivable. It is the latter approach that is adopted in the MP. The new mechanism corresponding to rewrite rules and transformations at level 0 and principles and parameters at level 1 seems to be the set of economy principles. As described, for instance, by Chomsky (2000), these economy principles are meant to be derived from external necessity and to be sufficient to derive from the FLN all properties of the language faculty which we need in view of its function in the account of language.

In conclusion, relativized descriptive and explanatory adequacy highlights a number of parallels between different levels of explanatory depth. If the transition from Standard Theory to GB-theory is seen as reaching for a deeper level of explanation, the transition from GB-theory to the MP can be considered as having a similar impact.

## **6 The logical origin of explanatory adequacy**

For a theory to achieve explanatory adequacy, whether in the sense of (2c) or of (9b) two conditions have to be fulfilled. First, the theory has to address a question that asks for an explanation (of the relevant type). Second, it has to propose a plausible answer to this question. The type of question to be asked is exemplified by the questions listed under (4) and (7). From an epistemological point of view it is essential to separate the choice of a question and the effort to answer it. Once a question has been chosen, scientific practice can work along the empirical cycle in (12).

- (12) a. Select a set of data.
- b. Formulate appropriate generalizations about the data.
- c. Formulate a theory as a hypothesis about the system underlying these generalizations.
- d. Test the theory by deriving new generalizations and carrying out experiments.

The procedure in (12) is cyclic, because the experiments in (12d) will yield new data, extending the set in (12a), which will typically lead to additional or corrected generalizations in (12b) and an adaptation of the theory in (12c). The empirical cycle in (12) does not work in isolation. Without proper guidance, there are too many possible observations, too many possible generalizations, and too many possible theories. Ten Hacken (2006, to appear) elaborates this point and develops the notion of a *research programme*. A research programme is a set of assumptions that guides the selection and constrains the search space for each of the elements in (12) combined with criteria to evaluate the success of alternative theories. Chomskyan linguistics can be seen as a research programme in linguistics. Other research programmes are, for instance, Lexical-Functional Grammar (LFG) and Head-Driven Phrase Structure Grammar (HPSG).

It is not always easy to recognize the boundary between what belongs to the research programme and to the theory. In the case of Chomskyan linguistics, we are lucky in this respect, because the assumptions of the research programme have been discussed, attacked and defended to such an extent that they have been made much more explicit than is usual for research programmes. Moreover, Standard Theory, GB-theory and the MP constitute successive stages of the theory, which can be interpreted as operating within the same research programme. A strong motivation to consider them part of the same research programme is that they use the same criteria to measure success.

The question of the explanatory adequacy of the MP, as raised by (1a), can only be answered by means of the reference framework made available by the research programme. The research programme determines the question underlying explanatory adequacy. An isolated theory is like an isolated answer. The MP's potential for achieving explanatory adequacy derives from the research programme of Chomskyan linguistics. The extent to which this explanatory adequacy is realized can only be measured by means of the criteria

that are also part of the research programme of Chomskyan linguistics. Therefore, the research programme is the origin of the explanatory adequacy of a theory.

## 7 The opposition between the MP and OT

The MP is not only a theory that determines an approach to the questions asked by the research programme of Chomskyan linguistics. It also involves a formalism for the expression of the answers. The formalism is essential to achieve a sufficient degree of formalization to make claims precise and discuss them meaningfully within the research programme. Typical examples of components of formalisms are the ones listed in (13).

- (13) a. Phrase structure rules
- b. Feature structures
- c. Transformation rules
- d. Unification operations

In practice, a formalism usually consists of a selection of elements. Standard Theory uses (13a) and (13c), LFG uses (13a) and (13b), HPSG (13b) and (13d). The choice of formalism is to a large extent independent of the research programme. Whereas it is often possible to translate a theory from one formalism into another, it is much more difficult to translate a theory from one research programme to another one. The reason is that in one case, only new means of expression of the ideas have to be found, in the other a new underlying motivation.

The most striking difference between the MP and OT is the formalism used. MP uses *merge* as its central operation and trees as the way to represent the resulting structure. OT, as presented by Archangeli (1997), uses GEN and EVAL as its operations and represents the results in terms of tableaux.



Against this background, we can consider different interpretations of the nature of OT. If the description of the formalism exhausts its definition, we expect that it can be used in principle in various research programmes. In the same way as the mechanisms in (13), OT would only determine *how* theoretical statements are made, not *what* they are made about and *why* they are interesting. In this *formalism interpretation*, no questions as to the truth of OT arise. The only questions concern its expressivity. Alternatively, if OT is a proper theory, it has to be embedded in a research programme. In this *theory interpretation*, it makes sense to consider questions of truth and adequacy and use the evaluation criteria provided by the research programme.

The evidence as to whether the formalism interpretation or the theory interpretation of OT is correct is somewhat mixed. Archangeli (1997:2-4) suggests that the research programme of Chomskyan linguistics constitutes the background of OT. On the other hand, Bresnan claims that “LFG is actively being developed in an OT setting” (2001:122, fn. 1). In fact, in a single volume we find Legendre et al.’s (1998) analysis of *wh*-chains competing with an MP account and Bresnan’s (1998) analysis of weak cross-over effects in an LFG framework.

The best conclusion we can draw is that OT is different things to different people. As a technique it is used in different theoretical settings. Bresnan’s use of OT does not compete with the MP in any direct sense. We could assume that LFG as a research programme competes with Chomskyan linguistics, although, as discussed in detail by ten Hacken (to appear), such discussions are often indirect and usually problematic in nature. When Legendre et al. (1998:285-287) discuss the relationship between OT and MP, however, they treat them as alternative ways of accounting for the same phenomena. This is only possible if they make a different set of additional assumptions, both theoretical and meta-theoretical, than Bresnan.

There is a clear parallel between OT and phrase structure grammar. Phrase structure grammar is used in Standard Theory to generate deep structures and in LFG to generate c-structures. The formalism is basically the same, but it is used in different research programmes. In the MP the last vestiges of explicitly formulated phrase structure rules have long been dropped. The question whether phrase structure rules contribute to explanatory adequacy can only be addressed meaningfully within a particular research programme. That the MP has a higher degree of explanatory adequacy than Standard Theory depends on the relationship of both theories to the research programme of Chomskyan linguistics. Arguably, part of this increase in explanatory adequacy can be attributed to the abolition of phrase structure grammars in a general sense, but this evaluation depends more on the nature of Chomskyan linguistics than of phrase structure rules. It is not possible to use this evaluation as a basis for drawing a parallel conclusion as to the position of phrase structure rules in LFG. Similarly, the contribution of OT to explanatory adequacy can only be determined in the comparison of two full theories in the same research programme, not as an element of the formalism used both in Chomskyan linguistics and in LFG.

## 8 Conclusion

In this paper, an analysis of explanatory and descriptive adequacy has been proposed as an approach to the central claims in (1). It was first of all shown that the levels of adequacy as originally proposed by Chomsky (1964) do not apply in any transparent way to the MP. As they are formulated in (2) they lack the necessary generality. In order to solve this problem, the concept of relativized level of adequacy was introduced in section 5. This concept is more generally applicable and can be used also in the context of the MP. Adequacy of any level

cannot be assigned to a theory in isolation, however. It crucially depends on the embedding of the theory in a research programme.

At this point, we can summarize the nature, use, and origin of explanatory adequacy as follows. Explanatory adequacy means that, at a particular level of theoretical depth, the real-world entity at this level is adequately explained in terms of the level above it. The use of discussing explanatory adequacy at different levels is to show the logical connections between the historical stages of theoretical discussion in a field. Its origin is always a research programme. The role of the research programme is on the one hand to determine the questions with regard to which an explanation has to be provided, on the other hand to provide evaluation criteria to determine to what extent the potential explanatory adequacy is realized in a theory.

The way the tension between descriptive and explanatory adequacy is constructed as an opposition between different theories, as in (1), is highly problematic. The problem of indeterminacy as discussed in section 3, implies that they can only be reached together. If the MP and OT in (1) are to be understood as comparable entities, this implies a particular interpretation of OT, because explanatory adequacy does not come from a theory, but from a research programme. Assuming that both are theories in Chomskyan linguistics, we exclude from consideration any use of OT in other research programmes, e.g. LFG. In the context of Chomskyan linguistics we can compare the extent to which explanatory adequacy is realized by the MP and OT, but we cannot construct this as a contrast in which the MP is better at explanatory adequacy and OT better at descriptive adequacy. As Chomsky (1998:117) emphasizes, *explanatory adequacy* is a technical term and should not be confused with the potential of a theory to provide explanations. The most plausible interpretation of (1) is then that description is easier in OT than in the MP, considered as two

alternative theories in Chomskyan linguistics. *Description* in this informal use is only vaguely related to *descriptive adequacy* in the technical sense.

## References

- Archangeli, Diana (1997), 'Optimality Theory: An Introduction to Linguistics in the 1990s', in Archangeli & Langendoen (eds.), *Optimality Theory: An Overview*, Oxford: Blackwell, p. 1-32.
- Barbosa, Pilar; Fox, Danny; Hagstrom, Paul; McGinnis, Martha & Pesetsky, David (eds.) (1998), *Is the Best Good Enough? Optimality and Competition in Syntax*, Cambridge (Mass.): MIT Press.
- Bresnan, Joan (1998), 'Morphology Competes with Syntax: Explaining Typological Variation in Weak Crossover Effects', in Barbosa et al. (eds.), p. 59-92.
- Bresnan, Joan (2001), *Lexical-Functional Syntax*, Oxford: Blackwell.
- Chomsky, Noam & Miller, George A. (1963), 'Introduction to the Formal Analysis of Natural Languages', in Luce, R. Duncan; Bush, Robert R. & Galanter, Eugene (eds.), *Handbook of Mathematical Psychology*, New York: Wiley, vol. 2, p. 269-321.
- Chomsky, Noam (1964), *Current Issues in Linguistic Theory*, Den Haag: Mouton.
- Chomsky, Noam (1965), *Aspects of the Theory of Syntax*, Cambridge (Mass.): MIT Press.
- Chomsky, Noam (1966), 'Topics in the Theory of Generative Grammar', in Sebeok, Thomas A. (ed.), *Current Trends in Linguistics Volume III: Theoretical Foundations*, Den Haag: Mouton, p. 1-60.
- Chomsky, Noam (1975), *Reflections on Language*, s.l.: Fontana.
- Chomsky, Noam (1980), *Rules and Representations*, New York: Columbia University Press.
- Chomsky, Noam (1981), *Lectures on Government and Binding*, Dordrecht: Foris.
- Chomsky, Noam (1988), *Language and Problems of Knowledge*, Cambridge (Mass.): MIT Press.
- Chomsky, Noam (1993), *Language and Thought*, Wakefield RI: Moyer Bell.
- Chomsky, Noam (1998), 'Some Observations of Economy in Generative Grammar', in Barbosa et al. (eds.), p. 115-127.

- Chomsky, Noam (2000), 'Minimalist Inquiries: The Framework', in Martin, Roger; Michaels, David & Uriagereka, Juan (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, Cambridge MA: MIT Press, p. 89-155.
- Guasti, Maria Teresa (2002), *Language Acquisition: The Growth of Grammar*, Cambridge MA: MIT Press.
- ten Hacken, Pius (2006), 'Formalism/Formalist Linguistics', in Brown, Keith (ed.), *Encyclopedia of Language and Linguistics, 2nd Edition*, Oxford: Elsevier, vol. 4, p. 558-564.
- ten Hacken, Pius (to appear), *Chomskyan Linguistics and its Competitors*, London: Equinox, to appear in 2007.
- Hauser, Marc D.; Chomsky, Noam & Fitch, W. Tecumseh (2002), 'The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?', *Science* 298:1569-1579.
- Hornstein, Norbert & Lightfoot, David (1981), 'Introduction', in Hornstein & Lightfoot (eds.), *Explanation in Linguistics: The logical problem of language acquisition*, London & New York: Longman, p. 9-31.
- Jackendoff, Ray (1989) 'What is a concept, that a person can grasp it?' *Mind and Language* 4:68-102.
- Jackendoff, Ray (2002), *Foundations of Language: Brain, Meaning, Grammar, Evolution*, Oxford: Oxford University Press.
- Legendre, Géraldine; Smolensky, Paul & Wilson, Colin (1998), 'When is Less More? Faithfulness and Minimal Links in *wh*-Chains', in Barbosa et al. (eds.), p. 249-289.
- Roberts, Ian G. & Roussou, Anna (2003), *Syntactic Change: A Minimalist Approach to Grammaticalization*, Cambridge: Cambridge University Press.
- White, Lydia (2003), *Second Language Acquisition and Universal Grammar*, Cambridge: Cambridge University Press.

*Pius ten Hacken*  
*Department of Translation*  
*School of Arts*  
*Swansea University*  
*Singleton Park*  
*Swansea*  
*SA2 8PP*  
*United Kingdom*  
*p.ten-hacken@swansea.ac.uk*

# Last Resorts and Grammaticality<sup>\*</sup>

*Jane Grimshaw*

Rutgers University

A “last resort” is argued to be nothing more than a winning, i.e. grammatical form, once it is understood in terms of competition between alternative candidates. It is a theorem of OT that we find last resort effects, since it follows from the nature of competition and constraint interaction.

*Economy, Optimality Theory, Competition*

## 1 Introduction

Every winning candidate (grammatical output) is a last resort under OT. It is the best one (or within the set of the best ones) left after the impossible have been eliminated: the last one standing.

I will refer to the observation that grammatical structures have a last resort appearance as the Last Resort (LR) effect. The LR effect is a theorem of OT. There is no “Last Resort” principle or constraint. There are no stipulations that a particular process or structure is possible only as a last resort. The LR effect follows from the very theory of competition and constraint interaction that determines grammatical well-formedness.

At the heart of OT (Prince and Smolensky 1993/2002, 2004) is a principle which chooses an optimum among a set of competitors which do not violate the same constraints. Other components of OT, in particular GEN, the set of

---

<sup>\*</sup> Thanks, as always, to Alan Prince, Vieri Samek-Lodovici, Roger Schwarzschild and Sten Vikner.

constraints, and the nature of inputs, are to be studied empirically, under the logical structure imposed by this model of optimization. OT is a theory of constraint interaction. It defines a class of grammatical theories. The following statement of the principle follows that given in Grimshaw (1997):

An optimal output form for a given input is selected from among the class of competitors in the following way: a form that, for every pairwise competition involving it, best satisfies the highest-ranking constraint on which the competitors conflict, is optimal.

Since the winning candidate is the one that “best satisfies” the constraints in the way just defined, it is not (necessarily) perfect. If all candidates but one are eliminated, the remaining candidate must be grammatical.

A number of structures or processes have been declared in the literature to be LRs.<sup>1</sup> Collins (2001) reviews some examples. The essence of the hypothesis is that operations or structures are possible only when necessary.

## 2 Illustrations

The analysis of *do* support in Grimshaw (1997) illustrates the point that the optimum in a grammatical evaluation in OT is the last resort. A skeletal version of the proposal goes as follows. The constraints relevant for *do* are:

FULL-INT	A syntactic element has a meaning
OB-HD	A projection has a head
NO-LEX-MVT	A lexical head cannot move

The English Ranking is:  
 NOLEXMVT, OBHD >> FULLINT

---

<sup>1</sup> “A court of last resort, is one which decides, definitely, without appeal or writ of error, or any other examination whatever, a suit or action, or some other matter, which has been submitted to its judgment, and over which it has jurisdiction.” <http://www.new-york-lawyer.ws/law-dictionary/label.htm>

- (1) a. Which books will they t read t?  
b. \* Which books they will read t?
- (2) a. Which books did they t read t?  
b. \* Which books read they?
- (3) a. They read books  
b. \* They did read books

The interrogative in (2a), with *do* support, is optimal because it best satisfies the constraints in their English ranking. It violates FULLINT, but the presence of *do* makes it possible to satisfy the other two constraints under consideration. Choosing the optimum with *do* is no different from choosing any other optimum. Importantly, the constraint rankings which select *do* in this configuration are those of the grammar of English, and must therefore be consistent with every other grammatical property of the language. The only language particular aspect of *do* support is the ranking which is responsible for it. If both OB-HD and FULL-INT dominate NO-LEX-MVT, the result is a system in which a main verb will raise, and no *do*-like morpheme will appear. The constraints themselves are the constraints of universal grammar. Nothing more need be said.

(4) Matrix Interrogatives with no auxiliary

Input: <read(x, y), x=they, y=which books, past>

Candidates	NOLEX MVT	OBHD	FULLINT
a. [CP which books e [IP DP [VP read t ]]]		*!	
☞ b. [CP which books do <sub>i</sub> [IPDP t <sub>i</sub> [VP read t ]]]			*
c. [CP which books read <sub>i</sub> [IP DP will [VP t <sub>i</sub> t ]]]	*!		

My point here is that it is not just epenthesis that is a last resort phenomenon. All choice of grammatical forms is, even the order of elements



within simple phrases of English. The analysis which follows is from Grimshaw (2001, 2002), and analyzes phrases in terms of the three alignment constraints: SPECLEFT, HDLEFT and COMPLEFT. These constraints are violated when a specifier, head or complement respectively is not aligned with the left edge of a phrase, designated as “HP” in the tableau. The references above contain the definitions of the constraints. Of the six logically possible candidates I consider only those in which the head (H) and the complement form a constituent, and are therefore adjacent.<sup>2</sup>

(5) English: orders for Specifier, Head and Complement

	SPECLEFT	HDLEFT	COMPLEFT
☞ a. [ <sub>HP</sub> Spec H Comp]		*	**
b. [ <sub>HP</sub> Spec Comp H]		**!	*
c. [ <sub>HP</sub> H Comp Spec]	*!*		*
d. [ <sub>HP</sub> Comp H Spec]	*!*	*	

The ranking of these constraints for English is:  
SPECLEFT >> HDLEFT >> COMPLEFT.

(5) illustrates the fact that the winner, candidate a., is the choice which remains after SPECLEFT has eliminated all of the specifier-final options, and HDLEFT has eliminated the head final option. The winning grammatical candidate, is thus the last resort.

What these instances show is that the status of a grammatical structure as a LR is not a matter to be declared as an aside – a statement of grammatical fact

---

<sup>2</sup> If GEN imposes this grouping, then there need be no constraint preferring Head Comp adjacency or constituency. If GEN does not impose this grouping, it must be imposed by a constraint, or perhaps a combination of constraints. If either HDLEFT or COMPLEFT dominates the grouping constraint(s), the specifier will separate the head from the complement.

that is outside the scope of, or in addition to, the grammatical structure of a particular language, or indeed Universal Grammar as a whole. The set of possible LRs is the set of possible optima, i.e. those that can be selected under rankings of the constraints of UG. The set of LRs found in a given language is the set of optima that are in fact selected under the actual ranking of the universal constraints, as set in the grammar of the language. The fact that the solution (i.e. the LR solution) for the structure of phrases in English is specifier-head-complement is the consequence of the ranking of the relevant markedness constraints. The fact that *do* support is the solution for filling a complementizer when a sentence contains no meaningful auxiliary is also the consequence of ranking among particular markedness constraints. In sum, LRs do not exist as grammatical structures that have a special status. There are winning candidates and that is it. Even a perfect candidate, should one exist, is an LR. Again, it is the candidate which best satisfies the constraints.

If the above is correct, why is it that the notion of a LR has significant appeal? The fundamental reason, I think, is that in a theory which does not work by comparative evaluation of alternative forms, the LR status of all grammatical structures is NOT a theorem. Consider Government-Binding theory, for example. A sentence containing *do* support violates any principle which requires that elements in syntactic structure have meaning. Hence it must be accommodated by appeal to some further notion, in this case the hypothesis that extra options are available to particular languages, as a last resort (see Chomsky 1991). This notion falls outside the theory proper, and is not connected in any way to properties of the theory, apart from the fact that it fixes mismatches between the predictions of the theory and empirical observation.<sup>3</sup>

---

<sup>3</sup> The theory-external character of LR is reflected in its frequent positioning between quotation marks. On p.269 Lasnik et al 2005, the term appears 5 times in the penultimate paragraph, in quotes each time.

There is another reason why the concept of LR has a credible status. This is the simple observation that some grammatical phenomena seem to be rarer than others, or perhaps limited to languages with particular properties, or in other ways “marked”. In the case of *do* and perhaps the behavior of *much* (Corver 1997) this may be accurate, although unrelated languages behave similarly, as we can see in the role of the verb *suru* in Japanese. The question is whether this motivates treating them as different in some fundamental way from other grammatical devices, and the sentences containing them as different in some fundamental way from other grammatical structures. Epenthesis violates faithfulness, whether it is phonological or syntactic.<sup>4</sup> There is a clear difference between phonological and syntactic systems in their tolerance of unfaithful mappings between inputs and outputs. In syntactic optimizations unfaithful candidates are winners only under particular circumstances. It is not impossible for them to be winners, however. Examples of proposals which crucially posit unfaithful optima include Legendre et al (1998), Grimshaw (2000).

These cases have in common the fact that they involve lack of faithfulness to grammatical, and not lexical, information in the input. Information such as “+wh” for example, or “+plural”. While this has not been formalized, recoverability evidently prevents mass deletions in order to better satisfy markedness constraints, such as the alignment constraints discussed above. Since unfaithfulness is so limited, it is striking when it is found.

In these terms, what is a default? If an LR is a structure which best satisfies the constraints as ranked in a grammar, isn’t a default exactly the same?

---

<sup>4</sup> The precise nature of the faithfulness violation involved in *do* support depends on some details of the analysis which I will not go into here. In Grimshaw 1997 I suggested that *do* is the verb of choice for epenthesis because it has the least meaning, hence to strip away its meaning is to minimally violate faithfulness. Since *do* is not in the input its presence in the output constitutes a faithfulness violation in addition.

Yes. However, in the case of epenthesis discussed above, we can separate the two as follows. The selection of a candidate with epenthesis as the optimum is determined by the ranking of constraints on configurations. The selection of *do* as the item to epenthesize is a function of markedness constraints evaluating *do* itself. The first set of constraints mandates the insertion of something, the second set of constraints decides what will be inserted. We call this a “default”. This is comparable to the choice of an epenthetic vowel: constraints on syllable structure require the presence of a vowel, other constraints determine which vowel in fact occurs.

### **3 Last resorts as winning candidates**

A fundamental prediction of the claim that LRs are nothing more than winning candidates is that the set of winning candidates across languages is the set of all candidates, minus the set of candidates which are harmonically bounded (Samek-Lodovici and Prince 2005). I will call this the “real winners”. The set of last resorts must be the same. The logic of the argument is that LRs are chosen in exactly the same way as grammatical candidates, since this is what they are. There is no reason, then, to expect to find a universally identifiable set of LR “strategies”. (More accurately, the universally identifiable set will be identical to the real winners). On the contrary, since languages with differently ranked constraints vary in their choice of grammatical structures, the range of possible last resorts must also be variable. Movement may be the choice of one grammar and no movement the choice of another. Insertion of *do* is the choice in some grammars and not in others. This point lies at the heart of deriving LR effects from constraint interaction. LRs will be entirely determined by the grammar of a given language, and it will not be possible to draw up a list of them which is invariant across languages. “.. the first one now will later be last”.

## 4 Conclusion

This paper further develops a line of research which examines the status of economy and related notions in OT. In Grimshaw (2001, 2002) I argued that “Economy of Structure” is a theorem of OT. It follows from the nature of constraints on phrases, including the alignment constraints discussed in Section 2. These constraints conflict with constraints requiring the presence of a specifier and a head (the constraint OBHD from Section 2). Because of the conflict, every phrase is guaranteed to violate at least one constraint: the fewer phrases the better. In Grimshaw (2006) I sketched an argument that “Economy of Movement” is also a theorem of OT. It follows from faithfulness and markedness constraints, which inevitably penalize chains, since they are inevitably unfaithful.

The core hypothesis of this work is that these effects are not due to “Principles” which are added to a theory to regulate its effects. Rather they follow from the very factors that determine syntactic well-formedness in the first place. The nature of optimality theoretic competition enforces what we call economy, without any assistance from us. Nothing is possible unless it is necessary.

## References

- Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In R. Freidin ed. *Principles and Parameters in Comparative Grammar*. 417-54. MIT Press.
- Corver, Norbert. 1997. *Much* support as a last resort. *Linguistic Inquiry* 28. 119-164.
- Grimshaw, Jane. 1997. Projection heads and optimality, *Linguistic Inquiry* 28. 373-422.
- Grimshaw, Jane. 2000. Optimal clitic positions and the lexicon in Romance clitic systems”, in Géraldine Legendre, Jane Grimshaw and Sten Vikner eds. *Optimality Theoretic Syntax*, MIT Press.

- Grimshaw, Jane. 2001. Economy of structure in OT. Rutgers Optimality Archive 434-0601.
- Grimshaw, Jane. 2002. Economy of Structure in OT, in *Papers in Optimality Theory II*. University of Massachusetts Occasional Papers 26. 81-120.
- Grimshaw, Jane. 2006. Chains as unfaithful optima. ROA 844.-0606. Note: this is a chapter of ROA-844, *Wondering at the Natural Fecundity of Things: Essays in Honor of Alan Prince*.
- Lasnik, Howard, Juan Uriagareka and Cedric Boeckx. 2005. *A Course in Minimalist Syntax*. Blackwell.
- Legendre, Géraldine, Paul Smolensky and Colin Wilson. 1998. When is less more? Faithfulness and minimal links in Wh-Chains. In *Is the Best Good Enough? Optimality and Competition in Syntax* (P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis, and D. Pesetsky, eds). MIT Press. 249-289.
- Prince, Alan & Paul Smolensky. 1993/2002. *Optimality Theory: Constraint Interaction in Generative Grammar*. RuCCS TR-2, Rutgers Center for Cognitive Science. CU-CS-696-93, University of Colorado Department of Computer Science. ROA 537-0802.
- Prince, Alan, and Paul Smolensky. 2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Blackwell: Malden, MA; Oxford, UK; Carlton, Aus.
- Samek-Lodovici, Vieri and Alan Prince. 2005. Fundamental Properties of Harmonic Bounding. ROA 785-1205.

*Jane Grimshaw*  
*Department of Linguistics*  
*Rutgers University*  
*18 Seminary Place*  
*New Brunswick NJ 08901*

*grimshaw@ruccs.rutgers.edu*

*<http://rulinguistics101.org/page/grimshaw.html>*

# **Natural Law: The Dynamics of Syntactic Representations in MP<sup>\*</sup>**

*Alona Soschen*

MIT

This work concentrates on the requirements of the computational system of HL, by developing the idea that Natural Law applies to universal syntactic principles. The systems of efficient growth are for the continuation of motion and maximal distance between the elements. The condition of maximization accounts for the properties of syntactic trees - binary branching, labeling, and the EPP. NL justifies the basic principle of organization in *Merge*: it provides a functional explanation of phase formation and thematic domains. In Optimality Theory, it accounts for the selection of a particular word order in languages. A comprehensive and definitive understanding of the principles underlying MP will eventually lead to a more advanced design of OT.

*Keywords: minimalism, optimality, Natural Law, phases, applicatives*

## **1 Introduction**

Minimalist Program and Optimality Theory are complementary approaches. Their implementation exemplifies a natural tendency to

---

<sup>\*</sup> This research was supported by the grant from SSHRCC #756-2004-0019. The author wishes to acknowledge the following people for their comments, suggestions, and discussions that influenced this research at different stages: Cedric Boeckx, Hans Broekhuis, Sylvain Bromberger, Andrew Carnie, Noam Chomsky, Pius ten Hacken, Irene Heim, Alec Marantz, Clemens Mayr, David Pesetsky, Paul Pietroski, Conor Quinn, Henk van Riemsdijk, Edward Rubin, Adam Szczegielniak, and Kenneth Wexler. Naturally, all unresolved issues are solely mine.

proceed from descriptive methods toward a more generalized, explanatory adequate theory.

At this point, OT- system should be reformulated in such a way that the operations of the computational system of human language CHL can provide an account for the syntactic constraints. So far, proposals of how to go about this incorporation are lacking. The present state of disjunction leaves both MP and OT with certain flaws. It is not clear what basic operations underlie CHL, and how these operations relate to the output of the system. This inconsistency can be cancelled by incorporating the insights of MP and OT into a larger, highly desirable comprehensive scheme. A question that remains unanswered is the implication of the crucial difference between OT (static) and MP (dynamic) models - OT does not involve movement, and MP does. While MP is concerned with structural aspects and derivational procedures of the generator, OT is designed to assess the resulting syntactic representations. According to the view expressed in this paper, the proposals of what rules apply on the generator should precede an adequate formulation of interface/ output conditions that follow from more basic assumptions.

Natural Law (NL) exemplified in the growth of organisms as the Fibonacci (Fib) sequence has serious consequences for the theory of syntax. Similar to other structures that comply with NL, tree structures are maximized. The principle of maximization applied to the sequence of nodes in syntactic trees provides a functional explanation of binary branching, labeling, and the Extended Projection Principle (EPP) – the requirement for a sentence to have a subject. This explains why languages tend to have filled specifiers and complements, and why the number of arguments found in natural languages is limited the way it is.



The maximization requirement that every head must have an XP complement creates a problem at the bottom-most layer of a syntactic tree: it eliminates a line in a tree with only terminals. The solution to this problem lies in redefining binarity to include *level 0*, which follows directly from the functional pressure of cyclic derivation: each successive element combines with two already merged elements, not with one. For example, merging 1 with 2 (which is a sum of 1 and 1) yields a new element 3. However, merging two elements none of which is a sum – such as 0 and 2 – will not yield a new element. ‘Zero’-branching is exemplified e.g. as X-labeled elements in conjunctions. Furthermore, determining whether a node is an XP or an X in terms of a Fib sequence depends on whether the element is a result of Merge or not. In addition, a node has to be immediately dominated by a node bearing a different label. This clarifies the notion of labeling, and answers the question of what labels can be disposed of in syntax.

Redefining syntactic representations in terms of NL leads to the discussion of phasal properties of xPs, in Chomsky's sense (2001, 2004, 2005). A ‘maximal thematic domain’ requires a single pair of dyadic structures: the lower part constitutes a relation between individuals, and the upper part relates individuals to events. It is shown that passives of double object constructions (with obligatory arguments) and Applicative constructions (with optional arguments) follow the same pattern of derivation. NL explains why XP should be a well-defined space in a derivation, and argument representations are constructed a certain way. The cross-linguistic analysis offered in this paper leads toward the definition of both minimal and maximal syntactic domains.

This paper offers new ideas concerning minimal requirements imposed by CHL, and represents movement as a crucial part of the dynamic model of MP. The proposals of what rules apply in the process of generation of syntactic structures will allow OT to evaluate the resulting syntactic representations, and adequately formulate the output conditions.

## 1.1 Natural Law

The Fibonacci series is one of the most interesting mathematical curiosities that pervade the natural world. The series was invented around 1200 by Leonardo Fibonacci. In the series, each new term is the sum of the two that precede it:  $X(n) = X(n-1) + X(n-2)$ , 0,1,1,2,3,5,8,13,21,... The limit ratio between the terms is an irrational number .618034..., ‘Golden Ratio’ (GR). For centuries, it has been recognized that e.g. plants have a fixed number of leaves and petals. Early approaches to FS in nature were purely descriptive; they just sorted out the geometry of patterns. Recently, a theory of plant growth (phyllotaxis) explained the observed arrangements as following from space filling (Douady & Couder, 1992).<sup>1</sup> This system follows from simple dynamics that impose constraints on the arrangement of elements to satisfy conditions on efficient packing. Fib numbers are evident in the growth of every living organism.<sup>2</sup>

<sup>1</sup> The Fib sequence is related to maximizing space. As a consequence of simple dynamics, successive elements form at equally spaced intervals of time on the edge of a small circle, representing the apex. The repulsion between elements ensures that the radial motion continues and that each new element appears as far as possible from its immediate successors.

<sup>2</sup> In humans, Golden Ratio appears e.g. in the geometry of the DNA molecule. On a cellular level, the ‘13’ Fib-number present in the structure of microtubules (cytoskeletons and conveyor belts inside the cells) may be useful in signal transmission and processing. The brain and nervous systems have the same type of

## 1.2 The discrete infinity of language

The faculty of language (FL) in the broad sense (FLB) includes a sensory-motor system, a conceptual-intentional system, and the computational mechanisms for recursion; FL in the narrow sense (FLN) only includes recursion (Hauser et al, 2002). A highly specific property of the *discrete infinity* makes FLN crucially different from other discrete systems found in nature.<sup>3</sup> This is the most elementary rule of syntax; there are neither *n*-and-a-half words nor *n*-and-a half-word sentences. Furthermore, there is no limit to the length of a meaningful string of words; there are ten-word sentences, twenty-word sentences and so on indefinitely. This property is exemplified e.g. in a well-known nursery rhyme where each sentence  $X_k$  with a number of words  $n$  is succeeded by a sentence  $X_{k+1}$  with a number of words  $n+m$ :  $X_{k+1}(n) = X_k(n+m)$ .<sup>4</sup> In contrast, the Fib sequence in other biological systems exhibits *discrete finiteness*. Discretely infinite syntactic recursion is a species-specific property of the human mind. Consequently, finding out more about the principles underlying recursion will provide us with the clue to the structure of mental representations. Hauser et al. argues that FLN may have evolved for reasons other than language. In this article, rather than trying to identify the driving force

---

cellular building units, so the response curve of the central nervous system may also have the Fib sequence at its base.

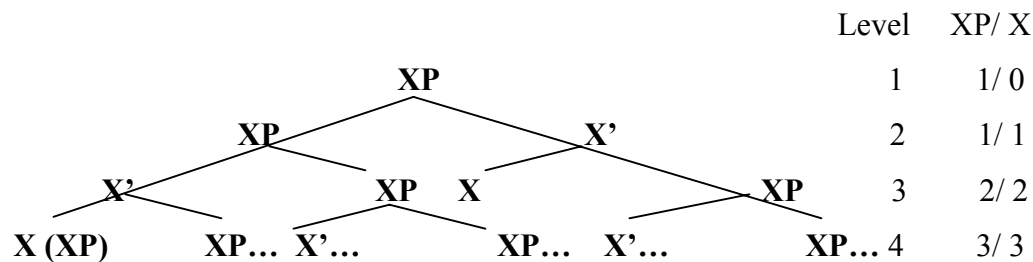
<sup>3</sup> The only other system of this kind - arithmetical capacity – can also be a part of FLN (Chomsky 2000).

<sup>4</sup> (i) The discrete infinity of language / ‘The House That Jack Built’  
 $X_{k+1}(n) = X_k(n+m)$ :  $X_2(n) = X_1(n+4), \dots, X_5(n) = X_4(n+4), X_6(n) = X_5(n+8), \dots$   
 (ii) Various kinds of flowers have a fixed number of petals. For each kind  $K$  of a flower (a, b, c, d, e,...), there is a fixed number of petals  $X$  that corresponds to a Fib number, e.g.  $K_a=X(3)$ ,  $K_b=X(5)$ ,  $K_c=X(8)$ ,  $K_d=X(13)$ ,  $K_e=X(21)$ ,  $K_f=X(34)$ ,  $K_g=X(56)$ .

behind the evolvement of FLN, we will approach FLN as a part of a more general system.

## 2 Maximization of syntactic trees

Recently, it was shown that syntactic structures exhibit certain mathematical properties (Carnie et al. 2005). Similar to other systems that comply with NL, tree structures are maximized in such a way that they result in a sequence of categories that corresponds to FS.



**Fig. 1**

The tree is generated by merging two elements; the next operation adds a newly introduced element to the already formed pair. Each item is merged only once; every subject/specifier and every object/complement position is filled. In the traditional sense of Chomskyan X-bar theory, a label immediately dominated by the projection of another category is an XP(hrase).<sup>5</sup> Other non-terminal nodes are annotated as X'. For example,

<sup>5</sup> The Fib-sequence in a tree is related to the fact that each node dominates exactly one maximal projection. Possibly, hierarchical structures created by pair-Merge (adjunction) comply with NL as well. This gives rise to the following question. It is not clear how the Narrow Syntax can determine that pair-Merge is required, rather than the default set-Merge. Rubin (2003) proposes the (obligatory) existence of a functional category, Mod, in the structure of adjuncts ([Mod [[YP]Adjunct]]) that is parallel in nature to functional categories in clauses.

*Adam loves Eve* may have a representation  $[_{XP} [_{XP} \text{Adam} ] [_X \text{loves} _{X'} [_{XP} \text{Eve} ] ]]$  where XPs are phrases, X's are intermediate projections, and Xs are 'heads'. Count XPs in each line of this derivation, and you will receive a partial FS (1, 1, 2, 3,...). If  $XP(n)$  is the number of XPs in the  $n$ th level, then  $XP(n) = \text{Fib}(n)$ . This property is true of all trees that are maximized by having specifiers and complements filled.

What is the reason behind compositionality that motivates combining exactly two terms in a set? The requirement to achieve tree maximization explains why the trees are constructed out of binary units. If Merge were allowed to optionally select three terms and combine them into a ternary structure, then FS of maximal categories would disappear. The sequence where each term  $A_n$  combines with the two that precede it is 1, 1, 1, 3, 5, 9, 17, 31, 57,... The ternary branching system shows a Fib-like sequence; however, the arrangement of elements displays a ratio different from GR, which fails to meet the condition of optimization. As a result, ternary branching or any operation that merges more than two syntactic elements is disallowed.<sup>6</sup>

NL provides an external motivation for Merge to distinguish between syntactic labels in a particular way. Determining whether a node is XP or X follows directly from the functional pressure of cyclic derivation. The Fib-based system distinguishes between *a sum of terms* and *a single term* (XP/ X), rather than XP/ X' or X'/ X. For example, level 4 has three XPs and three non-XPs: two X's and one X (cf. Level 3 – 1 X', 1 X, 2 XPs; Level 2 – 1 X', 1 XP). The assumption that syntactic structures have an

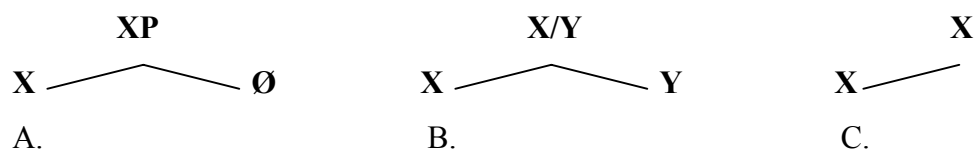
<sup>6</sup> Chomsky (2006) asserts that "Merge cannot create objects in which some object W is shared by the merged elements X, Y. It has been argued that such objects exist. If so, that is a departure from SMT, hence a complication of UG."

intermediate X' projection is a stipulation: basic syntactic representations are *monadic* (cf. the *dyadic* model of X-bar theory).<sup>7</sup>

## 2.1 Zero-Merge

The requirement to have specifier and complement positions filled faces a problem: it creates a ‘bottomless’ tree by eliminating a line with only terminal Xs. However, real sentences always have an ending point. The solution to this problem lies in redefining syntactic binarity to include zero-branching – in other words, to start FS with 0 instead of 1. This follows directly from the requirement of NL: each successive element is combined with a sum of already merged elements, not with one: merging 2 with 1 {1, 0} yields a new element 3, while merging two elements one of which is not a sum (2+0) does not. In the present system, singleton sets are indispensable for recursion.

The newly introduced type of merge, Zero-merge (Ø-M) distinguishes between *entities* {1}/X and *singleton sets* {1, 0}/XP at the bottom of the tree. New terms are created in the process of merging *sums* with *entities* to ensure continuation of motion; in (fig. 2), (i) and (ii) are the instances of Merge while (iii) is not. When the sum of terms is present at each step, it provides the ‘bottom level’ in the syntactic tree.



**Fig. 2**

<sup>7</sup> See also Collins (2002).

The suggestion to regard an empty element as functional in Merge has serious consequences for the theory of binary branching. The minimal building block that enters into linguistic computation is re-evaluated to include  $\emptyset$ -Merge, and identified as the product of  $\emptyset$ -Merge.<sup>8</sup> As a result, binarity is preserved, while there is no problem caused by the requirement to have specifier and complement positions filled. XPs and Xs can be disambiguated, which eliminates the necessity to proceed with further branching below the bottom level. Furthermore, labels X and XP are not syntactic primitives.<sup>9</sup> There exist numerous instances of label-switching between X and XP; for example, *that* may behave as X and XP in the same sentence.<sup>10</sup> The analysis along the lines of NL clarifies the notion of labeling, and answers the question of why labels can be disposed of in syntax. The idea that constituent structures are labeled appears to be a stipulation - this part of Merge should be abandoned in favor of a more explanatory adequate rule. As the grammar evolves toward a more generalized syntactic representation, the only necessary mechanism is not the one that determines which node is XP and which is X or X', but the one that determines whether a node is a result of Merge or not. Thus,

- *Determining whether a bottom node is XP or X depends on whether the element undergoes  $\emptyset$ -Merge.*
- *Determining whether a node is XP or X depends on whether the element is the result of Merge.*

<sup>8</sup> For the discussion of zero-branching constructions (bare nouns in conjunctions) see Roodenburg (2004).

<sup>9</sup> Heads can behave like Phrases and visa-versa, according to Carnie (2000), Collins (2002), and Chomsky (2004, 2005).

<sup>10</sup> (i)  $_{XP}$  That  $_{X}$  that is, is;  $_{XP}$  that  $_{X}$  that is not, is not - we all know  $_{XP}$  that.

## 2.2 Argument structure

Why is the number of arguments limited in a certain way? *Eve<sub>1</sub> laughs*, *Eve<sub>1</sub> kissed Adam<sub>2</sub>*, and *Eve<sub>1</sub> gave Adam<sub>2</sub> an apple<sub>3</sub>* are the only possibilities. In contrast, sentences *Eve<sub>1</sub> gave Adam<sub>2</sub> an apple<sub>3</sub> a pear<sub>4</sub>* and *gave Adam<sub>1</sub> an apple<sub>2</sub>* are ungrammatical: in the former, there is an extra argument; in the latter, one argument is missing. If we agree that syntactic principles follow from more general rules, we can make suggestions as to why thematic domains have a fixed number of nodes.

Merge is operation responsible for the construction of elementary trees and combination of these pieces into larger structures. Strong Minimalist Thesis entails that Merge of  $\alpha$ ,  $\beta$  is unconstrained. Under External Merge (EM),  $\alpha$  and  $\beta$  are separate objects; under Internal Merge (IM), one is part of the other, and Merge yields the property of *displacement* (Chomsky 2001). The argument structure is the product of EM. The pressure for the tree to be maximized justifies the basic principle of organization in both types of Merge. Move is just one of the forms of Merge: EM induces IM by virtue of the fact that already conjoined elements have to be *displaced* to occupy maximally advantageous positions in the tree.

The application of Fib-rule makes some interesting predictions about the constraints on EM. Assume that  $\emptyset$ -Merge ( $\emptyset$ -M) is the operation responsible for constructing elementary argument-centered representations prior to lexical selection.<sup>11</sup>  $\emptyset$ -M is relevant for distinguishing between

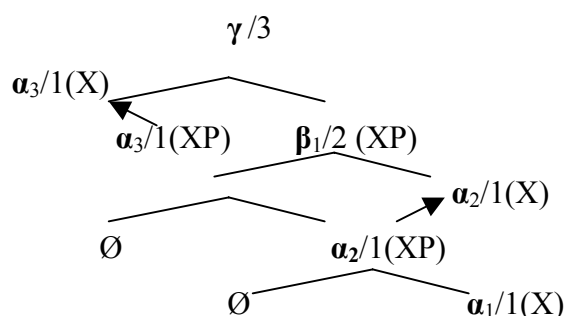
---

<sup>11</sup> Chomsky (2006) specifies that there exist other argument-based constructs such as e.g. Pritchett's (1992) theta-driven model of perception, 'relevant to the use of



entities  $\{1\}/X$  (*single terms*) and singleton sets  $\{1, 0\}/XP$  (*sums*). Determining whether a node is XP or X follows directly from the functional pressure of cyclic derivation to merge terms of *different* types only.

In contrast with what is found in other natural systems of efficient growth, once a syntactic constituent is formed, it cannot be broken up into parts. The Impenetrability Condition IC induces type-shift from *sums* to *entities* at each level in the tree. For example, at the point where 3 is merged with 5, element 5 is the *a* of 2 and 3, but 3 is a single entity. As is shown in (fig. 3),  $\alpha_2/1$  is shifted from *singleton set*  $\{\alpha_1, 0\}$  (XP) to *entity*  $\alpha_2$  (X) and merged with  $\alpha_3$  (XP). The type of  $\alpha_3/1$  is shifted from *singleton set*  $\{\alpha_2, 0\}$  (XP) to *entity*  $\alpha_3$  (X) and merged with  $\beta_1$  (XP).<sup>12</sup>



**Fig. 3**

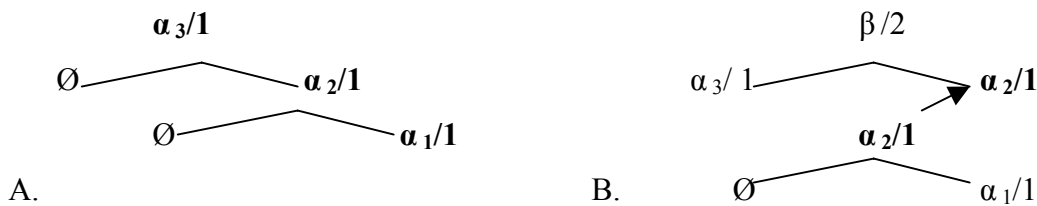
There is a limited array of possibilities for the Fib-like argument tree depending on the number of positions available to a term adjoining the tree. This operation either returns the same value as its input ( $\emptyset$ -Merge), or the cycle results in a new element (N-Merge). The recursively applied rule

language'. In such and similar models, a verb is theta-role assigner. In a Fib-like EM, the only function that matters is the one that identifies arguments.

<sup>12</sup> Throughout the paper, the author complies with Chomskyan X-bar model to build representations that are not in 'real time'.

adjoins each new element to the one that has a higher ranking in a bottom-up manner, starting with the term that is ‘ $\emptyset$ -merged first’.<sup>13</sup>

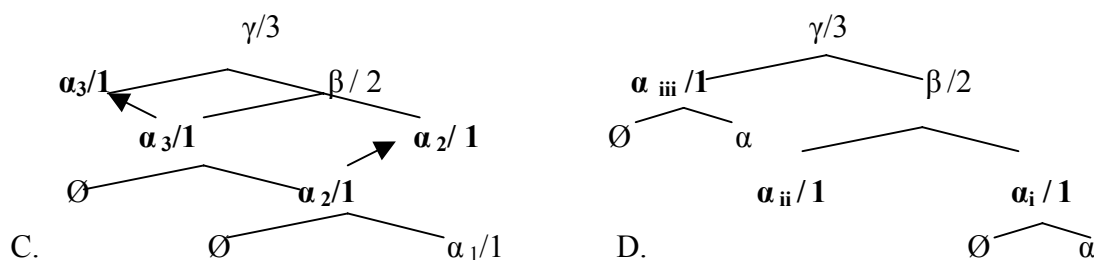
- Term  $\alpha_1$  can be  $\emptyset$ -merged *ad infinitum*. The function returns the same term as its input. The result is zero-branching structures (fig. 4, A).
- $\emptyset$ -merged  $\alpha_1$  is type-shifted to  $\alpha_2$  and N-merged with  $\alpha_3$ . The process creates a single argument position of intransitive (unergative and unaccusative) verbs, e.g. *Eve<sub>1</sub> laughs*, *The cup<sub>1</sub> broke* (fig. 4, B).<sup>14</sup>
- Terms  $\alpha_2$  and  $\alpha_3$  assume positions where each can be merged with a non-empty entity, the result is two positions (fig. 4, C).
- There are three positions to accommodate term 1 (i, ii, and iii). This may explain why in double object constructions the number of arguments is limited to three (*Eve<sub>1</sub> gave Adam<sub>2</sub> an apple<sub>3</sub>*) (fig. 4, D).



<sup>13</sup> Term A may undergo  $\emptyset$ -Merge either first or second. The supporting evidence comes from Japanese that threatens the same NP as any of the two. In (i), the argument position of *the girl* is ‘ $\emptyset$ -merged second’ in the matrix clause and ‘ $\emptyset$ -merged first’ in the subordinate clause.

(i) Yoko-ga kodomo-o koosaten -de mikaketa onnanoko-ni koe-o kaketa  
 Yoko-NOM child -ACC intersection-LOC saw girl -DAT called  
 ‘Yoko called the girl who saw the child at the intersection’  
 (Pritchett 1992)

<sup>14</sup> Certain verbs of spatial configuration such as *lean* are unergative with an agentive subject but unaccusative when they take a non-agentive subject (Levin and Rappaport Hovav 1995).

**Fig. 4**

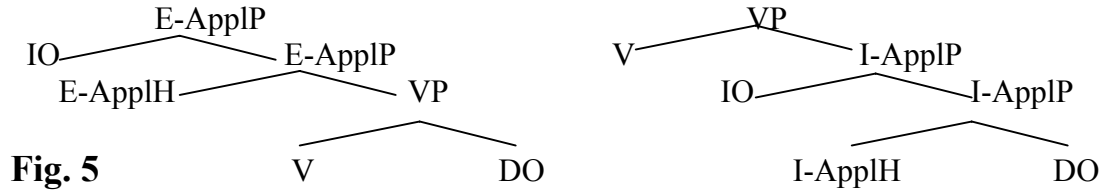
## 2.3 Maximal thematic domains

We have shown so far that the NL-logic can be applied to the analysis of EM to provide an account for the number of argument positions. The argument structure is built upon *hierarchical* relations. Hierarchy is assumed to be automatic for recursive operations (Chomsky 2005).

The applicative and double object constructions of the kind *John baked Mary a cake* and *John gave Mary a cake* are essential for the analysis of maximal thematic domains.<sup>15</sup> Recent research on argument structure has resulted in a complex representation that consists of two levels: one involves two individuals, and another expresses an individual-event relation. Sentences *John baked/ gave* [*Mary*]<sub>individual</sub> [*a cake*]<sub>individual</sub> are the first type, and [*John baked a cake*]<sub>event</sub> [*for Mary*]<sub>individual</sub> / [*John gave a cake*]<sub>event</sub> [*to Mary*]<sub>individual</sub> are the second. It was suggested that a relation between individuals is established by means of the Individual Appl

<sup>15</sup> See Marantz (2003), McGinnis (2001), Pylkkänen (2001, 2003).

Head I-ApplH in I-ApplP, and by means of the Event Appl Head E-ApplH in E-ApplP (fig. 5).<sup>16</sup>



When the trees are maximized and all positions are filled, the sum of heads, specifiers, and complements yields a maximal space of 13 (the Fibonacci number):

- (1) [XP Y<sub>EP</sub> [Y<sub>E</sub> Y<sub>E</sub>' [XP vP [v v' [XP VP [V V' XP ]]]]] Y<sub>E</sub> E-Appl H
- (2) [XP vP [v v' [XP VP [V V' [XP Y<sub>IP</sub> [Y<sub>I</sub> Y<sub>I</sub>' XP]]]]] Y<sub>I</sub> I-ApplH

In theory, there is a strong possibility that maximal thematic domains are constructed to accommodate all possible argument configurations represented in (fig. 6). There does not seem to be any intrinsic reason semantically or morpho-phonologically as to why thematic domains of this kind should be maximal spaces with a particular number of nodes. However, from a broader perspective, there is a sense in which the domains under discussion are maximal (see Part 1).

<sup>16</sup> This classification is viewed as necessary to provide an account for the difference in semantic interpretation. See e.g. Erteshik-Shir (1979) and Snyder (2003) on the semantics of the English to-dative and double object constructions with 'give'.

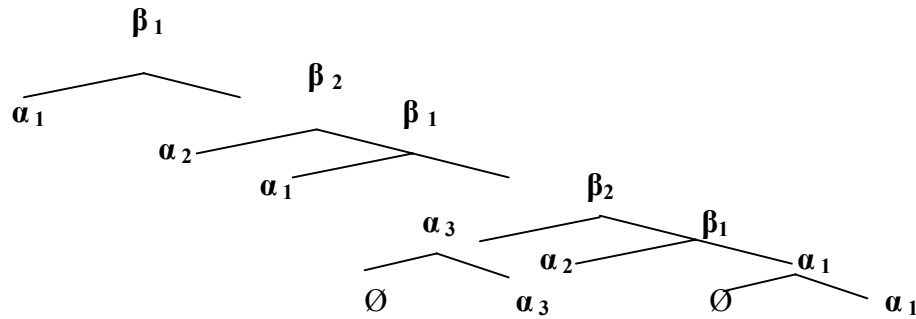


Fig. 6

### 3. Merge and displacement in syntax

Syntactic representations are characterized by two operations: *Merge* and *Displacement*. As was already shown, EM creates a hierarchical structure with a maximal number 3 as the number of arguments. Application of NL not only makes interesting predictions about the constraints on EM but also explains the properties of IM. *Displacement*, which is relevant at the point of pronunciation, assigns the order to lexical items LIs. It is possible that maximization requirement exemplified as the Fib-law justifies the principle of organization in IM replacing *hierarchy* with *dependency* relation between sisters that invariably involve an antecedent and a dependent.

The explanation of IM is very straightforward if we assume that derivations proceed by phases and movement depends on the qualification of phrases as phases.<sup>17</sup> According to Chomsky's Phase Impenetrability Condition PIC, only the Edge and the Head of a phase are visible to later syntactic operations; the domain is opaque. At the end of each phase, derivations are sent off to PF (Spell-Out) and LF (Interpretation). Are phases propositional? According to Chomsky (who suggests that *vP* and *CP*

<sup>17</sup> See Chomsky (1995, 2004, 2006) for the discussion of phase formation. See also Boskovic (2002), Epstein and Seely (2002), Legate (2003), Müller (2004), Suranyi (2004), and Wexler (2004).

are phases, while VP and TP are not) the answer is most probably *yes*. Only a fully fledged phrase can qualify as a phase. *Bill likes Mary* is possible because there is an additional position *x* in Spec, *v*P to accommodate NP *Bill*. This position is projected by a phasal Head *v* in  $[x_{\text{Bill}} \text{ } v\text{P} \text{ } v \text{ } v' \text{ } [_{\text{VP}} \text{ loves Mary } ]]$ . In contrast, *likes Mary* is not a phase as no position *x* is available to accommodate NP *Bill*: representations of the kind  $_{\text{VP}}[x \text{ loves Mary}]$  is not feasible. As was already discussed, ternary branching or any operation that merges more than two syntactic elements is disallowed in syntax. In this paper, phases are primarily characterized by their ability to induce a cycle by projecting extra Spec positions, to ensure continuation of movement in derivations. Syntactic phase formation is regarded as language-specific in this article: phases are redefined as *maximal* (propositional) and *minimal* (non-propositional) constituents. It follows then that any X can in principle head a phase.

### 3.1 Minimal and maximal phases

In the linguistic literature, it was maintained that only the relation between individuals and events constitutes a (propositional) phase, to provide an account of passive formation in the Applicative and Double Object constructions (McGinnis, 2001). It was concluded that the absence of an extra Spec-position in I-Appl Phrase disqualifies it from phases, by blocking direct object (DO) movement. Sentences of the kind *A cake was baked  $t_{\text{cake}}$  for Mary* and *A cake was given  $t_{\text{cake}}$  to Mary* are grammatical (DO-movement of NP *a cake* to Spec, E-ApplP), while *A cake was baked Mary  $t_{\text{cake}}$*  and *A cake was given Mary  $t_{\text{cake}}$*  are not. However, I-Applicatives behave like phases in other languages, by allowing DO-movement and

blocking IO-movement in passives.<sup>18</sup> Synthetic (inflectional) languages such as e.g. Italian and Hebrew I-ApplPs exhibit the properties of *minimal (min)*-phases, analytical languages such as English and Icelandic lack I-ApplP phases, and both groups are characterized by *maximal (max)*-phases such as *vP* and E-ApplP.<sup>19</sup> The absence of *min*-phases is characteristic of languages with fixed word order, where subject and object have to be ordered with respect to the verb. This is in contrast with languages that establish relations between words by means of inflections.

### 3.2 Phase parallelism and ECM

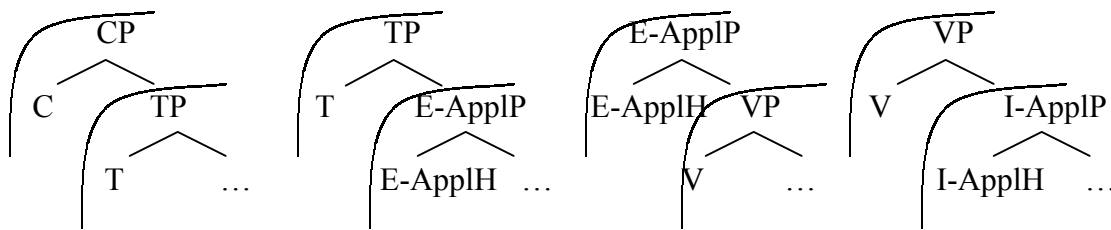
A certain class of verbs assigns structural case to an embedded subject in Exceptional Case Marking constructions in sentences such as *Eve wanted<sub>Acc</sub> Adam to taste an apple* where NP *Adam* is assigned Acc Case by the matrix verb *want*. This fact was accounted for in terms of CP-reduction. If this is a universally accessible rule, it is not clear why many languages – Spanish, Hebrew, and Russian among them – lack ECM. The explanation of this contrast lies in the distribution of the language-specific types of phases.<sup>20</sup>

- 
- <sup>18</sup> (i) [ VP V [ DO I-ApplP [ IO I-Appl' [ I-Appl' I-Appl *t<sub>DO</sub>* ]]] Italian, Russian,  
I-ApplP minimal phase Hebrew, Kinyarwanda  
(ii) [ IO *vP* *v* [ VP V [*t<sub>IO</sub>* I-ApplP [I-Appl' I-Appl DO ]]] English, Icelandic  
*vP* maximal phase  
(iii) [ DO E-ApplP [ PP E-Appl' [ E-Appl' E-Appl [VP V *t* ]]] I/R/H/K/E/I  
E-ApplP maximal phase

<sup>19</sup> There is further evidence that syntactic structures that express relations between two individuals should be considered more basic than those expressing a relation involving events. In languages with phasal I-ApplPs, sentences such as *A boy tore a girl a skirt*, *My friend broke me glasses*, *She fixed her neighbor a car*, and *A daughter washed her mother the dishes* are regular grammatical structures.

<sup>20</sup> Once the lower *T<sub>inf</sub>P*-phase is complete, subject NP in Spec, *T<sub>inf</sub>P* requires Nominative Case that cannot be assigned in this position due to the properties of

Phrases can be compared along the lines of their configurations if any syntactic phrase may in principle constitute a phase. For example, [CP C [TP T]] is parallel [VP V [I-ApplP I-ApplH]], because both have a no-label dyadic pair [XP X [XP X]] at their base as (Fig. 10). If this is true, one may expect to identify other *min*-phases in languages with I-ApplP phases, such as e.g. TP and VP.<sup>21</sup>



**Fig. 7**

The absence of ECM can be accounted for if in languages characterized by *min*-phases TPs constitute phases as well. For the same reason, these languages lack Optional Infinitival (OI) Stage.<sup>22</sup> English-speaking children at some stage between 1;10-2;7 on occasion omit TPs by producing sentences such as “Mary like John”, while they have no problems forming CPs (“Who Mary like?”). Cross-linguistic data shows that this stage is absent in Polish, Russian, Italian, and Spanish. Evidently, *min*-phases cannot be omitted even at an early stage of language development. The

T<sub>inf</sub>. The conflict between Case requirements and phasal status of TP cannot be resolved, and derivation crashes. In English, TP is not a phase, and subject moves to object position of matrix verb to receive Accusative Case. When Nominative Case assignment is unnecessary (e.g. in *Eve wanted to taste an apple*), derivation survives in a language with *min*-phases.

<sup>21</sup> Recall that in the present system, phases are characterized solely by their capacity to project extra Spec positions.

<sup>22</sup> See Wexler (1998) for the discussion of OIs.



cross-linguistic distribution of OIs in child language is consistent with the proposed universal phase parallelism and existence of two types of phases.

### 3.3 The Strict Cycle Condition

Chomsky (1973) states that ‘no rule can apply to a domain dominated by a cyclic node A in such a way as to affect solely a proper subdomain of A dominated by a node B which is also a cyclic node’. This condition is borne out in languages with *min*-phases that allow DO-movement in (3): IO-movement in (4) is blocked.

- (3) [ VP V [ DO I-ApplP [ IO I-Appl' [ I-Appl' I-Appl  $t_{DO}$  ]]]]  
I-ApplP minimal phase

- (4) # [ IO  $vP$   $v$  [ VP V [  $t_{IO}$  [ DO I-ApplP [I-Appl' I-Appl  $t_{DO}$  ]]]]]]  
 $vP$  maximal phase

From a more general perspective, in a system where  $X(n) = X(n-1) + X(n-2)$ , GR between the terms is preserved only when each term is combined with the one that immediately precedes it. Once a phase is complete, there is no possibility to extract yet another element from its domain. For example, 5 is a sum of 3 and 2. If the sum were formed by adding 1 to 3 etc., sequence would yield (1, 1, 2, 3, 4, 6, 9,...), violating GR.

### 3.4 Spell-Out and interpretation of phases

The next important question is how PF (Spell-Out) and LF (Interpretation) are derived in a language system that possesses both types of phases – *max*-/propositional and *min*-/non-propositional. As was already stated, PIC

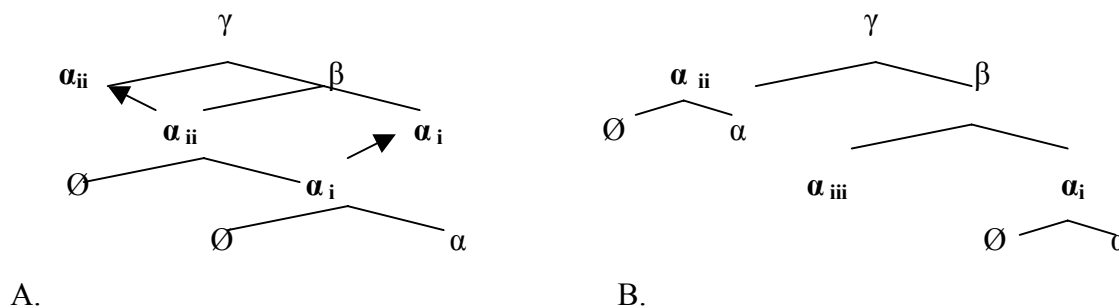
requires that only the Edge and the Head of a phase are visible to later syntactic operations; the domain is opaque. At the end of each phase, it is sent off to PF and LF.

Let us assume that (possibly all) languages have *max*-phases (such as CP, vP, and E-ApplP), while some languages also have *min*-phases (such as TP, VP, and I-ApplP). At the end of derivation, *max*-phases are sent both to PF and LF. One example is ‘garden-path’ sentences (Gibson 2000). Sentence  $_{CP1}$ [*The horse raced past the barn*] is interpreted as complete; the resultant derivation is sent to PF and LF. In  $_{CP2}$ [ $_{NP}$ [*The horse raced past the barn*]*fell*],  $_{CP1}$  is reinterpreted as NP and *max*-( $_{CP2}$ ) phase is sent to PF and LF.<sup>23</sup>

According to Epstein and Seely (2002), some features of LIs are illegitimate at one or the other interface. For instance, the pronoun *him* seems synonymous with *he*, even though their PF interpretations are distinct. It was assumed that unvalued lexical features are illegible at both LF and PF; valuation, however, is a necessary but not sufficient condition for LF convergence. The Case feature of a DP/N may be valued by the operation Agree, but a valued Case feature is by hypothesis still not interpretable at LF, and can be interpreted only at PF. Consider *John left his girlfriend with a baby* vs. *John left his girlfriend with a smile on his face*. Such and similar sentences (inspired by Chomsky’s examples) exemplify the Case feature valuation of a DP (*his girlfriend*, in this particular case) by Agree; however, the interpretation of the former varies depending on the semantics of matrix V, in contrast with the latter that has

<sup>23</sup> Note that in languages with *min*-phases such reinterpretation is expected to be blocked. By the time *max*-phase CP is complete, *min*-phase NP is already fully incorporated.

only one interpretation. In the EM label-three representation, the distinction between *John left his girlfriend with a smile*) and *John left his girlfriend* (\**with a baby*) is obvious: the first has two participants (fig. 8 A) and the second three (fig. 8 B). Possibly, a rule that determines the number of arguments and their hierarchy applies at each step in the derivation including *min*-phases, up till a complete LF is accessed at the level of *max*-phase.



**Fig. 8**

Chomsky (2001) identifies *vP* and *CP* as fully-fledged phases that are spelled-out cyclically and relatively independent at the interface. Epstein and Seely (2002) find this specification problematic: how do we know they are independent at the interface if Spell-Out applies before the interface is reached? The explanation is as follows. These phases are categories within which all theta roles are discharged, evidence that the underlying argument-based structure is preserved throughout derivations. To conclude,

- Phases can be compared along the lines of their label-free configurations.
- Heads of phases carry *edge-feature* that induces movement.

- All Ls have max-phases; certain (possibly synthetic) Ls also have min-phases.<sup>24</sup>
- At the end of derivation, maximal phases are sent to PF and LF.

#### 4. Argument-centered representations

A relation between individuals may constitute a phase, and induce movement (recursion). This means that *the core syntactic representations do not necessarily require a verb*. Certain languages have a very restricted number of verbs - for example, Australian language Jingulu has only three verbs: *do*, *go*, and *come*. Igbo (Ibo), a language spoken by approximately 18 million speakers in Nigeria, does not use verbs at all. A hierarchical linearization of arguments in the absence of verbs is exhibited in Igbo clusters. These clusters have the structure -gbá plus a noun: -gbá egwú *dance a dance*, egwú *dance*; -gbá igwè *ride a bicycle*, igwè *bicycle*; -gbá ákú/ egbè *shoot*, àkú *arrow*, égbè *gun*; gbá úkwú *kick*, úkwú *foot*; -gbá ọsọ *run a race*, ọsọ *race*; -gbá motò *travel with a vehicle*, motò *vehicle*, etc.

The structure termed ‘inherent complement verb’ (ICV) in Igbo linguistics has always been problematic for the analysis. The first characteristic that differentiates the use of ICV from light verbs in other languages is that it is a regular linguistic means. The second is that these structures do not have any simple verb equivalent. The root *gbá* is the only

<sup>24</sup> For the reasons already specified, Ls with *min*-phases always have *max*-phases, while the *max*-phase group may in principle (but not necessarily) have *min*-phases. The example seems to be Icelandic that has both ECM found in languages with *max*-phases and Dative experiencer constructions DEC such as (lit.) *John-Nom to-me-Dat likes* meaning *I like John*, DEC are characteristic of languages with *min*-phases (  $_{I-AppIP} [ NP_{John} NP_{me} ]$  ). English might have DP-phases and possibly PP-phases (  $_{PP} [ To\ him ]$ , *science is everything* ).

root in Igbo ‘devoid of meaning’, and the most productive one (Uchechukwu, p.c.). Other roots (e.g. *-tu*, *-kpa*, and *-ma*) check semantic features of the nouns they are combined with, such as ‘animacy’ and ‘shape’.<sup>25</sup> Similarly, the inflected *-gbá* roots are not semantically empty: e.g. *-do* is a suffix that expresses ‘fixation of the activity’ in *-gbá-do*.

As a matter of fact, *gbá* cannot be considered equal to light verb.<sup>26</sup> The semantic meaning of *-gbá*-clusters encodes the intrinsic connection between two key arguments, agent and theme, based on the *primary function of the theme with respect to the agent*. For example, the basic function of a car with respect to an agent is to carry passengers. Accordingly, *-gbá motò* means ‘travel with a vehicle’ – it does not mean ‘repair a vehicle’, or ‘sell a vehicle’. The intrinsic hierarchy of arguments supports the idea that the Relational Rel-(Appl) Head is expressed overtly as *-gbá* in Igbo. The agent is  $\emptyset$ -merged first *in situ* and then moved to Spec, RelP:

(5) [ Spec Rel-ApplP [ Rel-Appl’ **Rel-ApplH** (*-gbá*) [ [  $\alpha$ ,  $\emptyset$  ], [  $\beta$ ,  $\emptyset$  ] ] ] ]

(6) [  $\alpha$  Rel-ApplP [ Rel-Appl’ **Rel-ApplH** (*-gbá*) [  $t_\alpha$ , [  $\beta$ ,  $\emptyset$  ] ] ] ]

<sup>25</sup> This semantic feature checking is similar to SER/ESTAR alternation in Spanish and Portuguese. The choice of a particular (semantically empty) copula is consistent with (+/-) *permanency* feature of the predicate: SER is chosen over ESTAR when ‘sourness’ is a permanent property of the subject:

(i) a. Os limões são ácidos. ‘The lemons are [SER] sour.’ Portuguese  
 b. \*Os limões estão ácidos. ‘The lemons are [ESTAR] sour.’  
 (ii) a. \*As maçãs são ácidas. ‘The apples are [SER] sour.’  
 b. As maçãs estão ácidas. ‘The apples are [ESTAR] sour.’  
 (Costa 1998)

<sup>26</sup> In expressions *take a leap*, *take a leak* etc. there is no sharp divide between word and phrasal special meanings (Marantz 1997).

Further evidence for the identification of arguments prior to lexical selection comes from the analysis of verb formation (Hale & Keyser 2002). Conflation of N and V in verbs *to saddle* and *to shelf* is possible only from complement position, which results in *to saddle the horse* and *to shelf the book* (vs. # *to horse the saddle*, # *to book the shelf*). Nouns *saddle* and *shelf* can participate in the N/V conflation, but *horse* and *book* cannot because the hierarchical selection of themes (*horse*, *book*) precedes lexical formation.

The argument-centered logic of minimal syntactic units relies heavily on the data from language acquisition. It is well known that nouns are acquired first by children who have ‘perfect grammar’, equipped with the innate principles of universal syntax that allow them to master any language. Deprived of formal linguistic input, children of deaf parents simultaneously invent iconic languages in which the gesture for *give* is associated with three noun phrases, the gesture for *kick* with two, and the gesture for *sleep* with one (Lidz and Gletman 2004). Child language abounds in ‘verbless’ and ‘copulaless’ constructions. These structures are preserved in English as e.g. small clauses in *We consider<sub>SC</sub> [Mary a good friend]*. In many languages, copulas such as *is* in *Mary is my friend* are absent. Across language systems, nouns have a special status that ranks them higher than verbs.

The requirement of EM to disregard *order* in favor of *hierarchy* is evident in the following.<sup>27</sup> When asked to complete a sentence, the readers preferred conjuncts with a shared subject over object conjuncts, and both

---

<sup>27</sup> Kayne’s (1994) Linear Correspondence Axiom derives linear order from strict asymmetric c-command. Linearization applies only at the level relevant for pronunciation – the Spell-Out level (Chomsky 2000).

over clause conjuncts (Hoeks & Hendriks 2005). *The model embraced the designer and laughed* was chosen over *The model embraced the designer and the photographer*. Both of those sentences were ranked higher than the one that had conjoined clauses, such as *The model embraced the designer, and the photographer opened a bottle of expensive champagne*. The first type was selected because of the same agent for both verbs; the theme is ranked next. The preference is determined by the structure that identifies arguments first, before a verb is introduced.

In the propositional setting, verbs cannot be disposed of. In the Fib-terms, any two successive elements may be merged to form a part of recursive system. If certain types of phases are defined as non-propositional, IM can be analyzed as an (*edge*-)feature-driven mechanism, while in EM RelApplH establishes hierarchy of arguments  $\alpha$  and  $\beta$  in RelApplP, depending on whether  $\alpha$  or  $\beta$  is  $\emptyset$ -merged first.

#### 4.1 Word order

Grammatical linguistic expression is the optimal solution - the reason why a particular word order (Subject first) is preferred across languages. The hierarchy of nominal arguments is evident in the word order: SO (subject, object) order remains constant in the majority of languages (96%, Table 1). SOV order (rather than SVO) is the predominant one. The canonical word ordering in optimal terms is SOV, SVO, VSO, VOS, OVS, and OSV. Table 1 shows that the highest preference is given to languages that are either Subject and Object first, or Subject first. Furthermore, it is evident that language systems are symmetrical (SOV/ VSO, SVO/ OVS, VSO/ OSV), which confirms the idea of SO/OS parallelism.

---

ORDER	NUMBER	PERCENT
(SOV)	497	47
(SVO)	435	41
(VSO)	85	8
(VOS)	26	2
(OVS)	9	.9
(OSV)	4	.4
Lack dominant WO	172	
	total	1228

Table 1. Word order (Dryer, 2005)

It may be argued that even though S+O (in SO languages) and O+S (in OS languages) display syntactic independence such as moving as a constituent, it is far from being typical or unmarked. This can be explained if movement is re-evaluated as the ‘internal’ version of Merge, thus not an ‘imperfection’ of language. Internally merged elements A, B have to be independent to occupy maximally advantageous positions in a syntactic tree. The symmetrical representation of arguments underlying EM assigns an equal status to both, the reason why conjoined  $\emptyset$ -merged elements (such as bare nouns in conjunctions) can move as one constituent only.

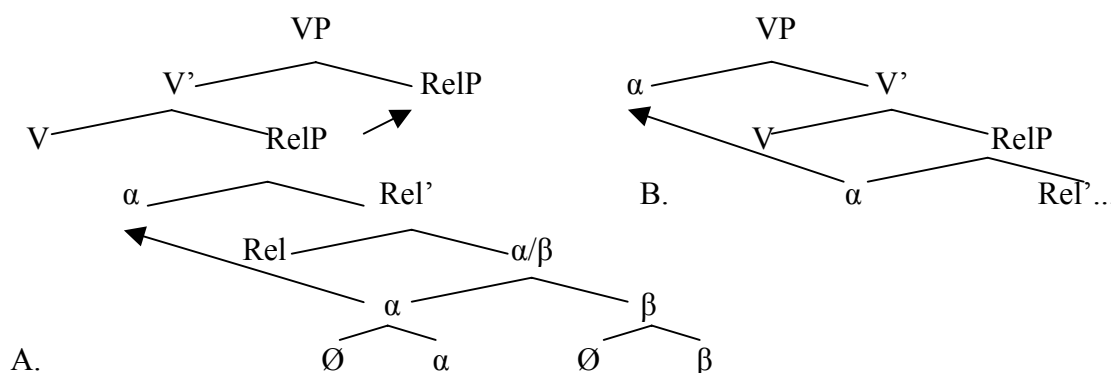
The introduction of R-function as a means of hierarchical prioritization is offered as an account for the ranking of word order across languages. The structure  $\alpha/\beta$  is symmetrical;  $\alpha$  and  $\beta$  share an equal chance for movement. The Rel(ational) Head RelH establishes a hierarchy of elements in the Relational Phrase RelP. In the present system, the choice of which element is ranked higher depends on which sum is merged first. If  $\alpha$  is  $\emptyset$ -merged with first, then  $\alpha$  is *displaced* first.

We have assumed that R takes a pair  $\{\alpha, \beta\}$  where each element has an equal status as its argument. The output of the function is the ordered



pair – either  $\langle \alpha, \beta \rangle$  or  $\langle \beta, \alpha \rangle$ , depending on whether  $\alpha$  or  $\beta$  is zero-merged first. According to Table 1,  $\langle \alpha, \beta \rangle$  is preferred to  $\langle \beta, \alpha \rangle$ . In a hierarchical organization of arguments, Subject-Object is preferred to Object-Subject. Further linearization proceeds in the following manner. Once S and O are ordered by RelH, SO undergoes second (Verb)-linearization. It has two options, where the first option is ranked higher than the second:

- The constituent SO is displaced. The resulting order is either  $\langle \alpha, \beta, \gamma \rangle$  or  $\langle \gamma, \alpha, \beta \rangle$  ( $\gamma$  is V).  $\langle \alpha, \beta, \gamma \rangle$  (SO-Verb) is preferred to  $\langle \gamma, \alpha, \beta \rangle$  (Verb-SO) (fig. 9, A).
- S is displaced. The resulting word order is  $\langle \alpha, \gamma, \beta \rangle$  (SVO). (fig. 9, B).



**Fig. 9**

In Object-first languages, R takes as its complement a pair  $\{\alpha, \beta\}$  with an output of the ordered pair  $\langle \beta, \alpha \rangle$  (OS), then verb merges with  $\langle \beta, \alpha \rangle$ . These are the two options:

- The whole constituent OS is merged with V. The order  $\langle \gamma, \beta, \alpha \rangle$  (VOS) is preferred to  $\langle \beta, \alpha, \gamma \rangle$  (OSV).
- The first constituent O is merged with V:  $\langle \beta, \gamma, \alpha \rangle$  (OVS).

## 4.2 Symmetrical conjunction

The conclusion we have arrived at is that a minimal syntactic domain (phase) can be defined in non-propositional terms, such as a relation between individuals. The analysis under development shifts the focus from verb to the noun, from propositional to the non-propositional logic of syntactic representations. As was already shown, a lower part [XP X] of [VP V [XP X]] represents a phase in certain languages, contrary to what had been previously assumed. In the present system of NL application, there is every reason to believe that in a non-linear representation that involves Merge only, this relation is *symmetrical conjunction* of the basic form  $\{\{\alpha, \emptyset\}, \{\beta, \emptyset\}\}$ .<sup>28</sup> Recall that  $\emptyset$ -Merge at the bottom level of the tree is necessitated by the requirement to induce a progressive cycle implemented by sums rather than single elements;  $\{\{\alpha\}, \{\beta\}\}$  is preferred over  $\{\alpha, \beta\}$ .

It is well known that conjuncts behave differently from other syntactic structures that can be derived from X-bar schema. Linguistic evidence attests to the fact that certain LIs selected from numeration LEX to participate in conjunctions are  $\emptyset$ -branching (non-maximal) projections such as e.g. prepositional Heads (*up and down the road*) and bare nouns (*cat and dog, knife and fork*). Movement of an entire conjunct out of a coordinate structure and movement of a subpart of a conjunct are prohibited. Conjunctions are syntactic primitives characterized by

---

<sup>28</sup> See Moro (2000) on the possibility of symmetry at base structure, resolved into asymmetry by Spell-Out. Kratzer's (1996) argumentation that subject should be introduced by a separate predicate opposes the view presented here.

symmetry, while *displacement* obeys the requirement to obtain a linear (asymmetric) order. The key requirement of CHL now includes a non-propositional configuration. As a result, the true structure of language can be characterized within a remarkably weak formalism.

## 5 Some implications for OT

In OT, variations among languages are attributed to differences in the constraint rankings which restrict linguistic expressions (Prince & Smolensky 1993, 1997). Given an underlying representation, a generator function produces a (potentially infinite) set of realizations, and a process of optimization picks the representation that minimally violates the constraints. Conflicts result in the satisfaction of higher ranked constraints at the expense of their lower ranked adversaries. Optimality Theory gives rise to a variety of specific formal instantiations depending on the types of representations and constraints invoked, but it is a largely unresolved question just what sort of formalism is appropriate for OT syntax.

A grammatical linguistic expression is the optimal solution. However, there has been no account for the preference of a particular word order (SO) in language systems. One possibility is that there are alignment constraints that involve the subject and the object, and the verb and the arguments. If this is the case, then a ranking of these constraints is responsible for the word order. Table 2 shows that the highest preference is given to languages that are either Subject and Object first, or Subject first. Furthermore, it is evident that language systems are symmetrical (SOV/ VSO, SVO/ OVS, VSO/ OSV), which confirms the idea of a parallelism of arguments at the basic level of syntactic representations.

Table 2. SO- and V-linearization

ORDER	SO-linearization		V-linearization					
	SO	OS	(SO)V	V(SO)	(OS)V	V(OS)	(S)V	(O)V
SOV	*		*					
SVO	*		*					
VSO	*		*					
VOS		*	*					
OVS		*	*					
OSV		*	*					

## 6 Summary and conclusions

Both OT and MP attempt to uncover the true structure of language which can be characterized within a remarkably weak formal system. Conjunctivism says that absolutely all relevant syntactic concatenation expresses conjunction; as is further developed to handle an increasingly broad range of constructions and theoretical considerations, it will inevitably become more complex.

The discussion concentrated on the ways to identify minimal requirements imposed by CHL by developing the idea that general physical laws underlie universal syntactic principles. In the present system, the external motivations of UG define the structure of atomic (indispensable) syntactic units. The argument structure was assessed depending on the number of positions available to element(s) adjoining a Fib-like syntactic tree. The minimal building block that enters into linguistic computation was re-evaluated to include  $\emptyset$ -Merge, and identified as the product of  $\emptyset$ -Merge. As a result, binarity was preserved, while labels XPs and X were disambiguated on the bottom line of the tree.

---

*The model outlined in this paper is argument-centered.* The idea under development is different from the existing approaches to the analysis of syntactic representations in that it shifts the focus from verb to the noun, from propositional to the non-propositional logic of syntactic representations. Conjunctions are identified as the core syntactic representations characterized by symmetry, and movement as a requirement to obtain a linear ordering. Movement depends on the qualification of phrases as phases, constituents characterized by *edge*-feature, in compliance with Phase Impenetrability Condition. Whether a phase is *maximal* (propositional) or *minimal* (non-propositional) is language-specific. All languages have *maximal* phases; in addition, synthetic (inflected) languages have *minimal* (i.e. Individual Applicative) phases. Label-free phases can be compared along the lines of their configurations, which in its turn provided an account of why languages with *minimal* phases lack ECM.

In sum, this paper offered new ideas concerning the key requirements imposed by CHL, such as minimal syntactic domains where a relation between two elements is established in a non-propositional configuration. In OT terms, grammatical linguistic expression is the optimal solution - the reason why a particular (S>O) word order is preferred in language systems. A better understanding of the general principles underlying CHL will eventually lead to a more advanced design of Optimality Theory.

## References

- Boskovic, Z. 2002. A-movement and the EPP. *Syntax* 5: 167-218.
- Carnie, Andrew, Medeiros D., and C. Boeckx. 2005. Some Consequences of Natural Law in Syntactic Structure. Ms. University of Arizona, Harvard University
- Carnie, Andrew. 2000. On the Definition of X0 and XP. *Syntax* 3: 59-106.
- Chomsky, Noam. 1973. Conditions on transformations. In S. Anderson and P. Kiparsky (Eds.), *A Festschrift for Morris Halle*, pp. 232-286. New York: Holdt, Winehart and Winston.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge: MIT.
- Chomsky, Noam. 2000. *The Architecture of Language*. Mukherji et al, eds., Oxford University Press.
- Chomsky, Noam. 2001. Derivation by Phase. In M. Kenstowicz, ed. Ken Hale: *A Life in Language*. Cambridge, Mass.: MIT Press, 1-52.
- Chomsky, Noam. 2004. Beyond Explanatory Adequacy. In A. Belletti, ed. *Structures and Beyond: The Cartography of Syntactic Structures*, Vol. 3. Oxford: Oxford University Press, 104-131.
- Chomsky, Noam. 2005. On Phases, to appear in C.P. Otero et al, eds., *Foundational Issues in Linguistic Theory*, MIT.
- Chomsky, Noam. 2006. Approaching UG from Below. Ms., MIT
- Collins, Chris. 2002. Eliminating Labels. in Samuel Epstein and Daniel Seely (eds.), *Derivation and Explanation in the Minimalist Program*. Oxford: Blackwell Publishing.
- Costa, J., 1998. L'opposition SER/ESTAR en Portugais, in "Etre" et "avoir" (syntaxe, semantique, typologie), ed. Rouveret, A., Saint-Denis, PUV.
- Douady, S., and Couder, Y. (1992) Phyllotaxis as a physical self-organized growth process. *Physical Review Letters* 68, 2098-2101.
- Dryer, Matthew S. 2005. Order of subject, object, and verb. In *The World Atlas of Language Structures*, edited by Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie. Oxford University Press.
- Epstein, S. D., and T. D. Seely. 2002. Rule Applications as Cycles in a Level-Free Syntax. In *Derivation and Explanation in the Minimalist Program*, eds. S. D. Epstein and T. D. Seely. Oxford: Blackwell. 65-89.

- 
- Erteschik-Shir, Nomi. 1979. Discourse constraints on dative movement. *Syntax and Semantics*, Vol. 12: Discourse and Syntax. Academic Press, Inc.
- Gibson, E. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. In: A. Marantz, Y. Miyashita & W. O'Neil (Eds.), *Image, Language, Brain. Papers from the first Mind Articulation Project Symposium*. Cambridge, MA: MIT Press.
- Hale, Ken and S. J. Keyser. 2002. *Prolegomenon to a Theory of Argument Structure*. Linguistic monograph 39. Cambridge, Mass: MIT.
- Hauser, Marc D., N. Chomsky, and W. T. Fitch. 2002. The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science* 22 November 2002: Vol. 298. no. 5598, 1569 – 1579.
- Hoeks, John and P. Hendriks. 2005. Optimality Theory and human sentence processing: Towards a cross-modular analysis of coordination. Ms., University of Groningen.
- Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge: MIT Press.
- Kratzer, A. 1996. Severing the external argument from its verb. In J. Rooryck and L. Zaring (eds.) *Phrase Structure and the Lexicon*. Dordrecht: Kluwer Academic Publishers. 109-137.
- Legate, J. A. 2003. Some interface properties of the phase. *Linguistic Inquiry* 34.3.
- Levin, Beth and M. Rappaport Hovav. 1995. *Unaccusativity*. Cambridge, MA: MIT Press.
- Lidz, Jeffrey and L.R. Gleitman. 2004. Yes, we still need universal grammar. *Cognition* 94, 85-95.
- Marantz, Alec. 1993. Implications of asymmetries in double object construction. In S. Mchombo (ed.), *Theoretical aspect of Bantu grammar*, CSLI Publications, Stanford, California, 113-150.
- Marantz, Alec. 1997. No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. A. Dimitriadis, L. Siegel, et al., eds., *University of Pennsylvania Working Papers in Linguistics*, Vol. 4.2, Proceedings of the 21st Annual Penn Linguistics Colloquium, 1997, pp. 201-225.
- McGinnis, Martha. 2001. Phases and the syntax of I-Applicatives. In *Proceedings of NELS 31*. University of Massachusetts, Amherst: GLSA.
- Moro, Andrea. 2000. *Dynamic Antisymmetry*. Cambridge, MA: The MIT Press.

- 
- Müller, G. 2004. Phrase Impenetrability and Wh-Intervention. In Arthur Stepanov, Gisbert Fanselow, and Ralf Vogel, eds. *Minimality Effects in Syntax*. Berlin: Mouton de Gruyter. 289-326.
- Pritchett, Bradley I. 1992. *Grammatical competence and parsing performance*. Chicago and London: University of Chicago press.
- Pylkkänen, Liina. 2001. What applicative heads apply to. In *Proceedings of the 24<sup>th</sup> Penn Linguistics Colloquium*, eds. M. Fox, A. Williams & E. Kaiser. Department of Linguistics, University of Pennsylvania.
- Pylkkänen, Liina. 2003. Introducing arguments. Doctoral Dissertation, MIT.
- Richards, Norvin. 2001. *Movement in language: interactions and architectures*. Oxford: Oxford University Press.
- Roodenburg, J. 2004. French bare arguments are not extinct: the case of coordinated bare nouns. *Linguistic Inquiry*, 35/2, 301-313.
- Rubin, E. 2003. Determining pair-merge. *Linguistic Inquiry*, Vol. 34/4, pp. 660-668.
- Snyder, K.M. 2003. The Relationship between Form and Function in Ditransitive Constructions, Doctoral dissertation, University of Pennsylvania, Philadelphia, PA.
- Suranyi, B. 2004. The left periphery and Cyclic Spellout: the case of Hungarian. In: D. Adger, C. de Cat and G. Tsoulash (eds.) *Peripheries and Their Effects*. Dordrecht: Kluwer. 49-73.
- Uchechukwu, Chinedu. 2004. Igbo verb clusters: The ontology of the -gbá cluster. In CML conference proceedings, Varna, Bulgaria.
- Uriagereka, Juan. 1997. *Rhyme and Reason*. Cambridge: MIT.
- Wexler, Kenneth. 1998. Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua* 106, 23-79.
- Wexler, Kenneth. 2004. Theory of phasal development: perfection in child grammar. MIT Working Papers in Linguistics 48, 159-209.

*Alona Soschen*

*MIT Department of Linguistics and Philosophy*

*77 Massachusetts Ave. 32-D808*

*Cambridge, MA 02139-4307*

*United States*

[soschen@mit.edu](mailto:soschen@mit.edu)



# Optimality Theory and the Minimalist Program

*Vieri Samek-Lodovici*

Italian Department – University College London

## 1 Introduction

The Minimalist Program (Chomsky 1995, 2000) and Optimality Theory (Prince and Smolensky 1993, 2004) are not alternative theories logically inconsistent with each other. Optimality Theory is a theory of how universal constraints of grammar interact (Prince and Smolensky 1993, Grimshaw 2005). Minimalism, as Chomsky notes (2000:41), is a research *program* –not a theory– investigating to what extent the language faculty provides an optimal design for the satisfaction of conditions at the interface with the sensory-motor system (PF) and the system of thought (LF). It is thus possible to pursue an OT-perspective of human grammar while maintaining minimalist goals, a fact highlighted by many contributors to the DEAL 2005 conference at ZAS (Berlin) on the relation between OT and Minimalism and also explicitly pointed out by Chomsky (2000:141). In this paper I argue that an OT-approach to grammar is actually essential to minimalist investigations, because it dramatically widens the set of linguistic properties potentially reducible to interface conditions while at the same time dispensing with interface-external language specific provisos. The discussion will hopefully also dispel some common misconceptions about OT.

## 2 Crosslinguistic Variation

One of the most evident empirical properties of human language is its crosslinguistic variation. Current minimalist theorizing –e.g. Chomsky (1995, 2000)–

**Linguistics in Potsdam 25 (2006): 77–97**

Hans Broekhuis and Ralf Vogel (eds.):

Optimality Theory and Minimalism: a Possible Convergence?

©2006 Vieri Samek-Lodovici

excludes crosslinguistic variation from the set of properties and imperfections that call for a minimalist explanation. Variation is instead assumed to follow from differentiations in the feature-bundles associated with lexical items in each language's lexicon, allowing for parametric variation in feature strength (weak vs strong) and/or variation in feature distribution (presence of EPP-features, presence of dislocating features).

Some important consequences follow from this assumption: to begin with, variation is left unrelated to interface conditions. Interface conditions might possibly motivate the existence of interpretable and even uninterpretable features as part of the optimal design of grammar, but they cannot derive the differences in strength and distribution associated with crosslinguistic variation. It follows that variation is also modelled as accidental. The parametric properties of features could be eliminated with no consequences on the optimal design of  $C_{HL}$ . The very existence of variation is thus unexpected: since a relation with interface conditions is excluded a priori, the parametric properties responsible for it are left with no linguistic motivation. It is unclear why variation occurs at all.

This state of affairs appears at odds with minimalist goals. In its strictest possible interpretation a minimalist approach to language should see a pervasive property like crosslinguistic variation emerge naturally from interface conditions. Whether this ambitious goal can be achieved or not depends on our initial assumptions about the nature of grammar constraints and their interaction. If the universal constraints of grammar never conflict with each others, then grammatical status inevitably coincides with their simultaneous satisfaction, and since the set of structures satisfying this condition is necessarily invariant across all languages unless something else is added to differentiate them, it becomes inevitable to account for crosslinguistic variation via language-specific parametric properties.

If on the other hand universal constraints are allowed to conflict with one another, as maintained in OT, crosslinguistic variation becomes a predicted out-

come, merely reflecting all the possible alternative resolutions of the conflicts among UG constraints.<sup>1</sup> Under this view, crosslinguistic variation is no longer accidental. Rather, it is entailed by the universal constraints of grammar themselves, which directly determine (i) *whether* variation occurs: it only occurs whenever two or more constraints conflict, never when constraints do not conflict; (ii) *where* it occurs: it occurs only with respect to those structures and properties on which the constraints conflict; (iii) *how* it occurs: the different structural aspects and properties found across distinct languages are themselves entirely determined by UG constraints, not by language-specific provisos.<sup>2</sup>

Deriving crosslinguistic variation as an inevitable consequence of constraint interaction is highly desirable also because it deepens the explicative power of our generative models. As concisely but effectively stated by Edwin Williams in his DEAL 2005 contribution, “*deepening explanation [...] arises when previously unrelated parts of a theory become predictively interrelated – the ‘constants’ of the theory are thereby reduced, making the correct theory more inevitable [...]*” (Williams 2005). By making crosslinguistic variation a predicted property, OT relates it to UG constraints in the strictest possible way, reducing the need for unnecessary theoretical constants such as language specific devices and provisos.

The explicative power of constraint conflict also emerges when considering

<sup>1</sup> Variation is of course contingent on the assumption that conflicting constraints can re-rank freely. Free re-ranking follows from the null hypothesis that no ranking is superior to any other.

<sup>2</sup> Under OT, individual grammars coincide with specific rankings of UG-constraints. The structure selected as grammatical by each grammar is the one that best satisfies UG-constraints under the corresponding ranking. More precisely, it is that structure A that beats any conceivable alternative B on the ranking at hand, i.e. such that for any B, A beats B on the highest constraint on which the two perform differently (Prince and Smolensky 1993, 2004; Grimshaw 2005). When two or more constraints conflict, their possible rankings determine all the available conflict resolutions, with each distinct ranking selecting a distinct optimal structure. The properties of the optimal structure remain shaped by the UG-constraints that selected it.

the number of distinct languages derived by a set of  $N$  conditions. With  $N$  binary parameters we may at most derive  $2^N$  distinct languages, whereas  $N$  conflicting constraints may give rise to  $N!$  languages.<sup>3</sup> As  $N$  increases,  $N!$  rapidly becomes a vastly larger number than  $2^N$ . For example 6 conditions determine  $2^6=64$  languages with binary parameters against  $6!=720$  potential languages with conflicting constraints. With 8 conditions the numbers become 256 vs. 40,320. It follows that on purely logical grounds an OT approach to constraint interaction potentially reduces the variation manifested across human languages to a far more restricted number of grammar conditions than allowed by parametric devices, providing a clear measure of their explicative power.<sup>4</sup>

The arguments just examined provide compelling theoretical motivation for investigating an OT approach to constraint interaction. They hold independently of minimalist goals, yet they appear essential to a minimalist perspective given their potential for reducing all aspects of human grammar, crosslinguistic variation included, to the conflicting interaction of constraints at the PF and LF

<sup>3</sup> The above figures presuppose  $N$  constraints conflicting with each others. Distinct languages only arise when constraints conflict. When they do not conflict their ranking is irrelevant, since it no longer affects the choice of optimal form. It is therefore incorrect to assume that  $N$  constraints always predict a cross-linguistic typology of  $N!$  languages. The overall size of the typology depends on the number of conflicts and the number of constraints involved in each conflict. This does not affect the explicative power of constraint conflict, since it remains true that a set of  $M$  crosslinguistic variants will potentially be reducible to a smaller set of conflicting constraints than binary parameters.

<sup>4</sup> The striking differences between  $2^N$  and  $N!$  should also dispel the misconception that reranked constraints are parameters in disguise. On the non-equivalence between parameters and pairs of opposite constraints see also Grimshaw (1997), and Samek-Lodovici (1998).

interfaces.<sup>5</sup>

### 3 Constraint Conflict

Having considered the above theoretical motivations favoring an OT perspective, we may ask whether they are supported by the empirical evidence available to us. Obviously, the very existence of crosslinguistic variation provides a first important piece of empirical support since as we saw variation is expected if constraints conflict and unexpected if they do not. There is also a great variety of highly complex linguistic paradigms that find a simple and principled explanation once examined in terms of constraint conflict. See for example the numerous OT-syntax analyses downloadable at the Rutgers Optimality Archives at [roa.rutgers.edu](http://roa.rutgers.edu). Several analyses are also available in the following volumes: Legendre, Grimshaw, and Vikner (2001); Fanselow and Féry (2002); Samek-Lodovici (forthcoming); Barbosa, Fox, Hagstrom, McGinnis, Pesetsky (1998); and Beckman, Walsh Dickey, Urbanczyk (1995).

Here, I will only consider two specific cases that I find particularly significant for the kind of constraint conflicts involved.

#### 3.1 Conflict between Prosody and Syntax

Verb movement aside, the syntactic and prosodic properties of simple clauses with overt subjects in Italian and English are very similar. When the entire clause constitutes new information focus we observe SVO order with rightmost prosodic prominence in both languages, as shown in (1) (focused phrases are subscripted by ‘f’. Prosodic prominence is marked as ‘\*’. The word marked by

---

<sup>5</sup> An even more ambitious project is pursued in Smolensky and Geraldine (2006), where the OT articulation of human grammar is viewed as directly emerging from the connectionist architecture of the human brain.



The crucial insight to meet this challenge was provided by Zubizarreta (1998), who analyzed rightmost focus in Romance as arising from the need to keep prosodic prominence rightmost and focus stressed. Armed with this insight and constraint conflict, we may account for the entire paradigm in terms of three simple constraints: (i) the syntactic constraint EPP forcing subject to raise to specTP (Grimshaw 1997, Chomsky 1982); (ii) the prosodic constraint H-I requiring the prosodic head of the intonational phrase IP to align with the IP's right boundary (Selkirk 1995, Truckenbrodt 1995); and (iii) the constraint Stress-Focus requiring focused phrases to carry the highest prominence in their domain (Jackendoff 1972, Truckenbrodt 1995, Zubizarreta 1998. In (1) and (2) the focus domain coincides with the entire clause).

When the entire clause is focused Stress-Focus is trivially satisfied independently of the position of the prosodic peak '\*'. Consequently EPP and H-I can be satisfied independently of one another, giving rise to the preverbal subjects and rightmost stress of sentences (1a) and (1b). When focus applies to the subject, however, the need to satisfy Stress-Focus –here ranked highest– unleashes a conflict between EPP and H-I. Grammars ranking H-I higher than EPP, like Italian, strand the subject in rightmost position as in (2a) to satisfy H-I, even if this forces a violation of the lower ranked EPP. Grammars ranking EPP higher than H-I, like English, raise the subject to specTP as in (2b) to satisfy EPP, even if this forces a misaligned prosodic peak in the intonational phrase IP which violates the lower ranked H-I.<sup>6</sup>

<sup>6</sup> An extended and more detailed analysis consistent with the simpler version provided here is provided in Samek-Lodovici (2005). The analysis employs finer grained prosodic and syntactic structures and derives a wider range of empirical data from Italian, English, German, French, and Bantu languages. A reduction of Italian clause-initial and clause-internal focus to prosodically induced clause-final focus is available in Samek-Lodovici (2006), showing how aside for the marginalization cases examined by Cardinaletti (2000, 2001), focus is always clause-final in Italian while post-focus phrases are always right-dislocated and clause-external. Even clause-initial focus is actually formed by clause-final focus followed by an entire dislocated clause.



The conflict between EPP and H-I thus properly predicts where the above focus patterns converge and diverge, while its resolutions accurately determine how the structures diverge when they diverge. All these predictions follow with no appeal to language specific properties and devices. The relevant constraints remain invariant in both languages and are fairly non-controversial, rooted in a long tradition of generative analyses. They are also clearly active in both languages, since they are necessary in both to determine the location of subjects and stress under clause-wide focus.

These desirable properties of the analysis are tightly linked with constraint conflict. As soon as we stipulate that constraints do not conflict we immediately lose the potential for a unified analysis rooted in UG-constraints alone. Since EPP and H-I remain necessary to derive the preverbal subject and rightmost stress of (1a) and (1b), the consequences of our stipulation emerge in the accounts for (2a) and (2b), which must now be made consistent with the satisfaction of both constraints despite clear evidence of the contrary. This is exactly the problem faced by the analysis in Zubizarreta (1998), where the lack of a theory of constraint conflict forces the introduction of two parametric devices. The first makes unfocused phrases prosodically invisible in English (but not in Italian, where their visibility is crucial for the analysis of rightmost focus). This reduces the IP in (2b) to the size of the focused subject alone, thus ensuring that stress is assessed as rightmost even in this case and satisfying the conditions equivalent to H-I in Zubizarreta's analysis. The second parametric device occurs in the grammar of Italian, where it ensures that the conflict between the conditions equivalent to EPP and H-I unleashed by focusing of non-final constituents is detected and resolved via the necessary syntactic re-arrangements.

Judging from the analysis of focus alone the benefits of a conflict based analysis are apparent, since it provides a unified analysis of the attested convergent and divergent patterns with no appeal to language-specific provisos. Under a minimalist perspective we may also wish to ask whether the conflict-



based analysis just examined is consistent with the criteria informing minimalist inquiries. These criteria are likely to require further investigation on how constraints like Stress-Focus, EPP, and H-I relate to interface conditions at PF and LF but they do not entail any specific assumption about the possible conflicts holding among these conditions. Since the sensory-motor and conceptual system serve largely independent goals there is no reason to exclude a priori the possibility of conflicting interface requirements. As far as I can see minimalist goals remain here best served by an analysis based on constraint conflict.

Before concluding, note how Zubizarreta's insights on Romance focus also show that PF cannot be insulated in a sub-system of its own external to narrow syntax as proposed in Chomsky (2000:118) because the constraints governing prosodic prominence clearly affect syntactic structures. In my opinion, this is a welcome result for a minimalist perspective, because it ties a type of syntactic dislocation to constraints governing prosodic properties, providing precisely the kind of genuinely non-syntactic requirements impacting syntax that are expected under a strict minimalist interface-based approach. Even in this respect, an analysis based on the conflict between prosodic and syntactic constraints appears to positively contribute to the minimalist enterprise, assigning a more concrete role to PF-interface constraints than originally envisaged.

### **3.2 Conflict between Economy Principles**

A second particularly interesting case of constraint conflict concerns the tension between structural and movement economy discussed by Cardinaletti and Starke (1994, 1999) in their crosslinguistic study of pronominal forms. Using data from a great variety of languages, including Italian, French, Slovak, and Gun (an African language of the Kwa family), they make four important observations: (i) weak pronominal forms are structurally simpler than their strong counterparts, lacking one or more of the top functional projections found in

the structure of strong forms; (ii) weak forms must obtain/check the functional features not already available in their simpler functional shell by raising to appropriate positions of the clause (e.g. spec AgrP to get Case); (iii) there is thus an inverse relation between the richness of a pronominal form's structural representation and the length of its chain, with simpler forms requiring longer chains; (iv) despite their longer chains, weak forms are always preferred to strong forms: strong forms are possible only where weak forms are excluded by independent factors.

On the basis of (iv), Cardinaletti and Starke propose the existence of an 'Economy of Representation' principle requiring minimization of structure. The challenge here concerns how to best model the conflict between Economy of Representation and Economy of Movement highlighted by the inverse relation between structure and chain-length observed in (iii). Under a conflict-based theory of constraint interaction such as OT the solution is straightforward, because economy is always a general property determined by the optimality-theoretic interaction of simple constraints (Grimshaw 1997, 2005; Prince 1997:2; Burzio 2000:209,216; McCarthy 2002:40; Smolensky, Legendre and Tesar 2006:505, 531).

Economy of Movement, for example, need not be stated as such because it follows from the conflict between the constraints that require movement to specific positions of the clause, – henceforth collectively identified as 'Check-F' whether defined in terms of feature checking or not – and the constraint Stay (Grimshaw 1997) violated by any instance of movement. When Stay is ranked lower than Check-F, the structure selected as optimal is the one that best meets Check-F while ensuring the lowest number of Stay violations, effectively minimizing movement.<sup>7</sup>

<sup>7</sup> Interestingly Chomsky (2000:132) describes the concept of 'feature strength' in the model developed in Chomsky (1995) as 'introduced to force violation of Procrastinate', confirming the violable nature of early minimalist economy principles. Optimality accounts like the one sketched above explicitly identify the constraint that is violated and the higher ranked

A similar analysis can be given for Cardinaletti and Starke's Economy of Structure, with a general constraint \*Struc (Zoll 1993, Prince and Smolensky 1993/2004) that penalizes any structure not required for the parsing of the initial array of lexical and functional items. Cardinaletti and Starke's observed preference for weaker pronominal forms then simply reflects the ranking Check-F  $\gg$  \*Struc  $\gg$  Stay (Check-F outranking \*Struc, and \*Struc outranking Stay). Under this ranking, weaker pronominal forms with less structure are preferred to more complex ones even at the cost of increased Stay violations, as observed in Cardinaletti and Starke's points (iii) and (iv) above. Such a preference however is subordinated to identical performance on the higher ranked Check-F constraint: whenever the weaker form underperforms the stronger one on Check-F, the stronger form is preferred, completing the account for point (iv).

The same tension between structural and movement complexity is not as readily accounted for in models that disallow constraint conflict. Consider for example Cardinaletti and Starke's analysis, cast in terms of the interaction between feature checking and economy principles in accord with the minimalist model proposed in Chomsky (1995). While they acknowledge the apparent contradictory nature of the two economy principles under discussion, they are also convinced that the tension can be dissolved by letting Economy of Representation (henceforth ER) apply at the point of lexical insertion and therefore prior to Economy of Movement (EM) (Cardinaletti and Starke 1999:202). At first sight, the proposed serialization might appear to deliver the desired result: first ER selects the least structured pronoun and then EM imposes the derivation with the shortest chain among those involving that particular pronoun. Since EM only examines derivations involving the same pronoun, the one selected by ER, the conflict between the two principles appears to have been dissolved.

---

constraints that force its violation. These accounts, however, have constraint conflict and constraint violability as their prerequisite and thus remain precluded to any theory of grammaticality based on the simultaneous satisfaction of all UG constraints.

The problems emerge when examining the assessment of ER, which cannot occur prior and independently of EM. The choice between a weak and a strong form depends in fact on the availability of a non-crashing derivation for the weak form since only in this case the weak form should be preferred to the strong one by ER. Assessing ER thus requires unfolding the derivation of the weak pronoun, a process that includes the assessment of EM. Consequently, there is no genuine point in the overall derivation where ER is truly assessed prior and independently of EM. On the contrary, EM is an integral part of the assessment of ER, so much so that a final unfolding of the derivation past the assessment of ER becomes redundant. Serialization thus fails as a strategy to avoid the conflict between ER and EM.

A possible alternative conflict-free solution can be provided via an explicit model of ER's assessment along the lines just examined above. This makes it possible to confine EM to derivations that share the same pronominal form, hence preventing the conflict with ER via the explicit subordination of EM to ER. The obvious question raised by this last solution is what determines the subordination of one principle to another. The answer is once again constraint conflict. In OT, conflict is a primitive, and the subordinate status of a constraint relative to another follows from the impossibility of satisfying both. Subordination is encoded via constraint ranking, and assessed in a unified and principled way via optimization. The opposite is true in the non-OT account outlined above, where subordination is an accidental property built-in in the assessment procedure for ER, envisaging a system where different principles are assigned different assessment procedures depending on their relation with each others.

Cardinaletti and Starke's analysis was conceived under the early minimalist system of Chomsky (1995) which allowed for economy principles. The revised crash-proof minimalist model proposed in Chomsky (2000) aims at disposing of economy principles too by a careful design of the operations involved in syntactic derivations, the domain to which they apply, and the order in which they

occur. For example, Move is defined in terms of the Agree and Merge operations plus an additional operation necessary to select the phrase that pied-pipes with the moving head (Chomsky 2000:135). The higher intrinsic complexity of Move relative to Agree and Merge is then assumed to prevent undesired movement. For example, ‘*a proof*’ will not move to specTP in “*there was a proof discovered*” whenever the expletive ‘*there*’ is present in the relevant array because merging of ‘*there*’ only requires Agree and Merge alone and no additional projection selection (Chomsky 2000:138).<sup>8</sup> Even this revised system, however, does not seem to be able to provide an analysis for the subordinate relation between movement and structural economy uncovered by Cardinaletti and Starke while keeping a principled and unified account of economy effects. The problem remains how to account for the ungrammaticality of a strong pronominal form when a weak form is possible. The conceivable solutions appear to contradict significant aspects of the design of C<sub>HL</sub> proposed in Chomsky (2000). The most obvious one involves a (potentially phase-internal) explicit comparison of distinct derivations, selecting the non-crashing derivation with the least structured pronominal form, thus introducing back in the system transderiva-

---

<sup>8</sup> The definition of Move  $\beta$  in Chomsky (2000), repeated below, has Agree followed by the selection operation (ib), followed by Merge. It remains unclear exactly how the complexity of Move alone can favor merging of the expletive over raising of ‘*a proof*’ in the derivation of ‘*there was a proof discovered*’. The initial Agree operation, step (ia), is shared by both derivations (Chomsky 2000:123, 135). Once step (ia) has been performed the correct derivation is contingent on proceeding with Merge of the expletive rather than performing the selection operation in step (ib), which would eventually yield the raised subject of ‘*a proof was discovered*’. The correct choice does not appear entailed by the complexity of Move, but rather by the assumption that Merge of array items always precedes the phrase selection operation in (ib).

- (i) Definition of Move  $\beta$  (Chomsky 2000:135).
  - a. A Probe P in the label L of  $\beta$  locates the closest matching [goal] G in its domain.
  - b. A feature G’ of the label containing G selects a phrase  $\beta$  as a candidate for “pied-piping”.
  - c.  $\beta$  is merged to a category K.

tional comparisons. Another solution, possibly more in tune with the spirit of the proposed system, would have to assume some degree of freedom in the merging of the feature bundles represented by the items in the lexical array. The derivation could then let weak pronominal forms that leave unparsed the feature bundles associated with higher functional layers proceed with their derivation. The derivation would then backtrack to a structurally more complex form that parses those same feature bundles whenever the derivation of the weaker form does not converge. Transderivational comparisons are then avoided at the cost of backtracking. While the details of each solution would have to be further investigated,<sup>9</sup> neither of them accounts in a uniform and principled manner for the property of economy shared by the two principles proposed by Cardinaletti and Starke. Economy of movement is assumed to follow from the relation holding between Move, Agree and Merge, whereas structural economy would have to follow from transderivational comparison or backtracking.

In conclusion, the attempts to model economy while disallowing constraint conflict appear unable to provide a fully general and principled analysis of the various instantiations of economy in human grammar. In contrast, allowing for constraint conflict and defining grammaticality accordingly enables OT to capture the notion of economy in its full generality, letting its specific applications emerge from very simple constraints whose subordination relations are explicitly encoded in a language's constraint ranking. Constraint-specific assumptions and provisos are dispensed with; all constraints are assessed in exactly the same way, examining only the structures at hand with no reference to the evaluation of other constraints. These would appear to be highly desirable properties for a minimalist perspective, making it possible to pursue a view of UG where conflicting universal constraints are dictated by legibility conditions at the PF and

<sup>9</sup> This is particularly true for the second solution, where Merge of array items has to wait past the attempted derivation of the weak pronoun. This contradicts the crucial assumption that Merge preempts Move (Chomsky 2000). See also the above footnote.



LF interfaces, while economy in its various manifestations emerges unstipulated from their interaction.

## 4 OT and Minimalism

The above sections provided some theoretical and empirical motivation for pursuing the minimalist program while supported by a formally precise theory of constraint conflict and interaction such as OT. Conversely, we may ask what a minimalist perspective would bring to OT-based inquiries.

A minimalist perspective would encourage a deeper understanding of universal constraints with the ultimate goal of linking them directly to interface conditions. OT's fundamental tenet that crosslinguistic variation follows from constraint conflict already forces a better understanding of UG constraints because it makes it impossible to derive variation through the escape hatch of language specific properties and devices. Successive analyses of similar phenomena within the OT-literature show a welcome trend towards ever simpler constraints; this increases the explicative power of the analysis and possibly comes closer to identifying constraints dictated by interface conditions alone as required by Minimalism. A particularly clear example of this trend is provided by Grimshaw's (2001) analysis of structural and movement economy. Rather than viewing them as separate phenomena emerging from the constraints *Stay* and *\*Struc* introduced above, Grimshaw derives both from a fixed set of five simple constraints: two of them respectively require the presence of specs and heads in phrasal projections while the remaining three require specs, heads, and complements to occur leftmost in their projection. Under this system, every projection is bound to violate some constraints. Consequently, any representation involving structure not required by higher ranked constraints loses against competing representations lacking such unneeded structure, yielding economy of representation. Likewise, since movement operations increase structure by

building additional copies of a constituent they too always violate some of the proposed constraints. It follows that movement operations that are not necessary to meet the demands of higher ranked constraints are suboptimal too, deriving economy of movement. We may still wonder about how to relate Grimshaw's constraints to interface conditions, but the explicative depth of our linguistic analysis has increased because economy of structure and economy of representation are now predicted epiphenomena determined by constraint conflict.

A minimalist perspective on OT might also lead to investigating how exactly the form selected as optimal by OT-optimization is identified. In this respect many linguists incorrectly believe that OT-optimization requires the human mind to actively generate an infinite number of competing structures, an impossible task in the finite time of linguistic exchanges. The error lies in interpreting optimization tableaux as a procedure to compute the optimal structure (hence contingent on the generation of all suboptimal alternatives) rather than as demonstrations of the optimal status of the selected form, relative to any other conceivable structural alternative hypothetically generable by a maximally unconstrained procedure 'GEN' responsible for structural composition. The issue then becomes whether computing optimal status relative to an infinite set of potential alternatives (mostly left ungenerated) with finite means and within finite time is psychologically feasible. Humans are clearly able to do that. We know that zero is lower than any other positive integer with no need to first enumerate all positive integers. We also know that number 21 is the least common multiple of 3 and 7 despite the infinitely many others available. We even know that even numbers are a subset of all integers despite both sets being infinite. In all these cases, and the infinitely many others that can be easily conceived, our mind appears able to reason in terms of the invariant properties and relations of the objects involved rather than by sheer enumeration and comparison. The identification of grammatical expressions as optimal solutions to possible rankings of UG constraints is likely to follow the same kind of reasoning. For example,



a ranking with Stay placed highest necessarily selects structures lacking movement; this property is sufficient to confine to suboptimal status every possible structure involving movement. There are infinitively many of them, but none of them needs to be actively generated to determine that they are all suboptimal (on the misconception of infinite generation and other common misconceptions see also Prince and Smolensky 1993:197, and Smolensky, Legendre, and Tesar 2006:523).

Misconceptions aside, the issue of how optimal structures are identified is a valid one. Tesar (1995) shows how dynamic programming provides a solution to this problem depending on the complexity of the constraints involved. He also applies this technique to the theory of syllable structure providing an algorithm that correctly computes the optimal structure among an infinite set of potential competitors for any given ranking of five specific constraints. Riggle (2004) goes even further providing a fully general solution to the above issue cast in terms of finite state automata (FSA). FSAs representing specific OT-constraints are combined together into a single larger FSA for which Riggle provides a general algorithm that efficiently computes the optimal forms selected across all possible constraint rankings.

Finally, the properties of OT-optimization themselves provide some useful tools in guiding the identification process. For example, any given set of structures identifies an infinite set of alternatives that are necessarily suboptimal because inevitably beaten by one or more of the original structures on any possible constraint ranking (Samek-Lodovici and Prince 1999, 2002). These alternatives need not be generated since the optimal form cannot be among them.<sup>10</sup>

In conclusion, whether UG constraints conflict or not is an empirical issue.

---

<sup>10</sup> The above discussion also shows why it is incorrect to view  $C_{HL}$  as a possible model for GEN. GEN defines the set of possible linguistic structures among which a constraint ranking selects the grammatical ones; it does not itself identify the optimal structure.  $C_{HL}$  on the other hand is expected to do just that, building the grammatical structure once provided with a suitable array of lexical items.

If they do, and they do appear to do so, a formally precise theory of their interaction becomes necessary for a proper understanding of grammar because simultaneous satisfaction of all constraints ceases to be a viable definition of grammaticality. This reason alone provides a strong motivation for pursuing an OT-perspective of human grammar, while further theoretical and empirical reasons have been offered in the above sections. The pursuit of minimalist goals does not presuppose a specific type of constraint interaction. It is fully consistent with an OT approach to constraint interaction, and as I argued in this paper it can greatly benefit from OT for an appropriate analysis of defining aspects of human language such as crosslinguistic variation, the syntactic impact of prosodic requirements, and economy in all its manifestations.

## References

- Aissen, J. 2003. Differential Object Marking: Iconicity vs. Economy. *Natural Language and Linguistics Theory* 21:435-483.
- Barbosa, P., Fox, D., Hagstrom, P., McGinnis, M., and D. Pesetsky. 1998. *Is the Best Good Enough. Optimality and Competition in Syntax*. MIT Press, Cambridge, MA.
- Beckman, J., Walsh Dickey, L., and S. Urbanczyk. 1995. *Papers in Optimality Theory*. University of Massachusetts Occasional Papers 18. GLSA, UMASS, Amherst, MA.
- Burzio, L. 2000. The Rise of Optimality Theory. In: Cheng, L., and Sybesma, R. (Eds.), *The First Glot International State-of-the-Article Book*. Mouton de Gruyter. New York. pp. 199-220.
- Cardinaletti, A., 2001. A second thought on emarginazione: destressing vs. “right dislocation”. In: Cinque, G., Salvi, G. (Eds.), *Current Studies in Italian Syntax: Essays Offered to Lorenzo Renzi*. Elsevier, Amsterdam, pp. 117–135.

- Cardinaletti, A., 2002. Against optional and null clitics. Right dislocation vs. marginalization. *Studia Linguistica* 56, 29–57.
- Cardinaletti, A. and M. Starke. 1994. *The Typology of Structural Deficiency. A Case Study of the Three Classes of Pronouns*. Ms., MIT, Boston.
- Cardinaletti, A. and M. Starke. 1999. *The Typology of Structural Deficiency. A Case Study of the Three Classes of Pronouns*. In: van Riemsdijk, H. (Ed.), *Clitics in the Languages of Europe*. Mouton de Gruyter. New York. pp. 145-234.
- Chomsky, N. 1981. Lectures on Government and Binding. Foris, Dordrecht.
- Chomsky, N. 1982: *Some Concepts and Consequences of the Theory of Government and Binding*. MIT Press, Cambridge, MA.
- Chomsky, N. 1995. *The Minimalist Program*. MIT Press, Cambridge, MA.
- Chomsky, N. 2000. Minimalist Inquiries: The Framework. In: Martin, R., Michaels, D., and Uriagereka, J. (Eds.), *Step by Step. Essay on Minimalist Syntax in Honor of Howard Lasnik*. MIT Press, Cambridge, MA, pp. 89-156.
- Fanselow, G., and C. Féry. 2002. Resolving Conflicts in Grammar. *Linguistische Berichte Sonderheft 11*. Helmut Buske Verlag. Hamburg.
- Grimshaw, J. 1997. Projections, Heads, and Optimality. *Linguistic Inquiry* 28:373-422.
- Grimshaw, J. 2001. *Economy of Structure in OT*. Ms., Rutgers University. Downloadable as paper 434-0601 at [www.roa.rutgers.edu](http://www.roa.rutgers.edu).
- Grimshaw, J. 2005. *How OT Works: Questions and Answers, Descriptions and Explanations*. Talk given at DEAL Conference, Zentrum für allgemeine Sprachwissenschaft, Berlin.
- Jackendoff, R. 1972. *Semantic Interpretation in Generative Grammar*. MIT Press, Cambridge, MA.
- Legendre, G., Grimshaw, J., and S. Vikner. 2001, *Optimality-Theoretic Syntax*. MIT Press, Cambridge, MA.
- McCarthy, J. 2002. *A Thematic Guide to Optimality Theory*. Cambridge Uni-

- versity Press. Cambridge.
- Prince, A. 1997. Endogenous Constraints in Optimality Theory. Talk handout, LSA-Institute, Cornell. Downloadable at <http://ling.rutgers.edu/gamma/talks/insthdt1.pdf>.
- Prince, A. and P. Smolensky. 1993/2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Blackwell Publishers.
- Riggle, J. 2004. Contenders and Learning. In: Schmeiser, B., Chand, V., Keller, A. and A. Rodriguez (eds.) *WCCFL 23 Proceedings*. Cascadia Press, Somerville, MA. pp101-114.
- Samek-Lodovici, V. 1998. Opposite Constraints: Left and Right Focus-alignment in Kanakuru. *Lingua* 104: 111-130.
- Samek-Lodovici, V. 2005. Prosody-Syntax Interaction in the Expression of Focus. *Natural Language & Linguistic Theory* 23: 687-755
- Samek-Lodovici, V. 2006. When right dislocation meets the left-periphery. A Unified analysis of Italian non-final focus. *Lingua* 116: 836-873
- Samek-Lodovici, V. Forthcoming. *Recent Developments in OT-Syntax and Semantics*. Special issue of *Lingua*. Elsevier, Holland.
- Samek-Lodovici, V. and Prince, A. 1999. *Optima*. Technical reports of the Rutgers Center for Cognitive Science RuCCS TR-57. Rutgers University.
- Samek-Lodovici, V. and Prince, A. 2002. *Fundamental Properties of Harmonic Bounding*. Technical reports of the Rutgers Center for Cognitive Science RuCCS TR-71. Rutgers University. (Also available as ROA-785 in the Rutgers Optimality Archives.)
- Smolensky, P., G. Legendre, and B. Tesar. 2006. Optimality Theory: The structure, Use, and Acquisition of Grammatical Knowledge. In: Smolensky, P., and Legendre, G. (Eds.), *The Harmonic Mind*. Vol 1. MIT Press, Cambridge, MA, pp. 453-544.
- Sorace, A. 2000. Gradients in auxiliary selection with intransitive verbs. *Language* 76: 859-890.

- Tesar, B. 1995. *Computational Optimality Theory*. Ph.D. Dissertation. University of Colorado.
- Truckenbrodt, H. 1995. *Phonological Phrases: their Relation to Syntax, Focus, and Prominence*. Ph.D. dissertation, MIT, Cambridge, MA.
- Truckenbrodt, H. 1999. *On the Relation between Syntactic Phrases and Phonological Phrases*. *Linguistic Inquiry* 30, 219–255.
- Williams, E. 2005. “My Theses”. Handout, DEAL Conference, Zentrum für allgemeine Sprachwissenschaft, Berlin.
- Zoll, Cheryl. 1993. Directionless syllabification and ghosts in Yawelmani. Ms. Berkeley, CA. ROA-28.
- Zubizarreta, M. L. 1998. *Prosody, Focus, and Word Order*. MIT Press, Cambridge, MA.

*Dr. Vieri Samek-Lodovici*  
*Department of Italian / Centre for Human Communication*  
*University College London*  
*Gower Street, London WC1E 6BT*

*ucljvsl@ucl.ac.uk*

*<http://www.homepages.ucl.ac.uk/~ucljvsl/>*

# The Simple Generator\*

*Ralf Vogel*

Universität Potsdam

I argue that the shift of explanatory burden from the generator to the evaluator in OT syntax – together with the difficulties that arise when we try to formulate a working theory of the interfaces of syntax – leads to a number of assumptions about syntactic structures in OT which are quite different from those typical of minimalist syntax: formal features, as driving forces behind syntactic movement, are useless, and derivational and representational economy are problematic for both empirical and conceptual reasons. The notion of markedness, central in Optimality Theory, is not fully compatible with the idea of syntactic economy. Even more so, seemingly obvious cases of blocking by structural economy do not seem to result from grammar proper, but reflect (economical) aspects of language use.

*Keywords: OT syntax, generator, markedness, syntax-semantics interface, syntax-phonology interface*

The history of the Chomskyan branch of generative syntax can be seen as an attempt to explain syntactic regularities as much as possible as the result of the rules that govern syntactic construal. Ideally, generatibility and well-formedness converge: every expression that can be generated is well-formed. Grammatical constraints, if they exist, are constraints on syntactic generation, they ideally apply within the derivational process.

This high emphasis on procedural aspects of syntax models has often been challenged, first of all by representationally oriented models like LFG, HPSG, a.o. While I do not want to make a claim related to this issue, I will discuss in

---

\* I would like to thank the audience of the DEAL conference Dec 17-19, 2005, for an extremely helpful discussion, and Hans Broekhuis for comments on an earlier draft of this paper. The work on this paper was partially supported by a grant for the research group “Conflicting Rules in Language and Cognition”, project A3, “Minimality in Optimality Theory”, FOR-375/A3.

this paper how the relation between derivational and representational aspects is handled in Optimality theoretic grammars.

OT makes a distinction between output candidates and optimal outputs. The two sets of generatable and well-formed expressions are non-identical. Prince & Smolensky (1993, 2004) argue that this distinction is already implicit in the history of generative grammar. They observe a shift in focus from the rules for the generation of expressions to well-formedness constraints on output structures. This becomes clear in the following quote:

“As originally conceived, the *RULE* of grammar was to be built from a Structural Description delimiting a class of inputs and a Structural Change specifying the operations that altered the input (e.g. Chomsky 1962). The central thrust of linguistic investigation would therefore be to explicate the system of predicates used to analyze inputs – the possible Structural Descriptions of rules – and to define the operations available for transforming inputs – the possible Structural Changes of rules. This conception has been jolted repeatedly by the discovery that the significant regularities were to be found not in input configurations, nor in the formal details of structure-deforming operations, but rather in the character of the output structures, which ought by rights to be nothing more than epiphenomenal. We can trace a path by which “conditions” on well-formedness start out as peripheral annotations guiding the interpretation of rewrite rules, and, metamorphosing by stages into constraints on output structure, end up as the central object of linguistic study.

As the theory of representation in syntax has ramified, the theory of operations has dwindled in content, even to triviality and, for some, non-existence. [...]”

(Prince & Smolensky 1993, 1; Prince & Smolensky 2004, 1)



The shift in perspective that is formulated here is directly reflected in the architecture of an OT grammar:

(1) **Structure of an OT grammar**

- a.  $\text{Gen}(\text{In}_k) \rightarrow \{\text{Out}_1, \text{Out}_2, \dots\}$
- b.  $\text{H-Eval}(\text{Out}_i, 1 \leq i \leq \infty) \rightarrow \text{Out}_{\text{real}}$

(Prince & Smolensky 1993, 4; Prince & Smolensky 2004, 5)

The grammar derives pairs of underlying forms and output forms ( $\text{input}_i$ ,  $\text{output}_i$ ). Each input form is combined with a candidate set of possible output forms by the candidate generation function Gen. H-Eval compares the candidate outputs on the basis of the set of violable constraints which are ordered in a lexicographic ranking. The optimal candidate is the one that performs best on the constraint hierarchy.

The generator determines what constitutes a possible candidate, i.e., what a linguistic structure is in general. One could also state that Gen consists of inviolable constraints. About the relation of Gen and H-Eval, and their status within the theory, Prince & Smolensky say the following:

“ [...] The function H-eval determines the relative Harmony of the candidates, imposing an order on the entire set. An optimal output is at the top of the harmonic order on the candidate set; by definition, it best satisfies the constraint system. Though Gen has a role to play, the burden of explanation falls principally on the function H-eval, a construction built from well-formedness constraints, and the account of interlinguistic differences is entirely tied to the different ways the constraint-system H-eval can be put together, given UG. [...] ”

Optimality Theory, in common with much recent work, shifts the burden from the theory of operations (Gen) to the the-



ory of well-formedness (H-eval). To the degree that the theory of well-formedness can be put generally, the theory will fulfill the basic goals of generative grammar. To the extent that operation-based theories cannot be so put, they must be rejected. [...]”

(Prince & Smolensky 1993, 5)

This quote makes clear that Optimality theory is conceived as a rival to operation based theories of grammar. The minimalist program as developed by Chomsky (1995), and further extended in later work by Chomsky and his many collaborators, is the most important current theory of this kind in syntax.<sup>1</sup>

Optimality Theory shifts the explanatory burden of a grammar model from the derivational system, the generator, to the system of wellformedness constraints. As a consequence of this, the generator function should be as unconstrained and simple as possible. Let us assume, as a starting point, that we choose a minimalist generator for an OT syntax model. The minimalist generator has at least the following components:

- Merge and Move:
  - substitution
  - adjunction ( $XP, X^0$ )
  - multiple specifiers
- feature checking
- feature strength (alternatively, EPP-features)

<sup>1</sup> Some aspects of minimalism look like candidate competition. A minimalist derivation starts with a list of lexical items, the numeration, initially picks two of them and merges them. From this stage on, there is always a choice for the next derivational step: either a new lexical item from the numeration is merged with the structure, or an element within the structure is moved. This choice between Merge and Move is one motivation for the model of serial optimisation explored by Heck & Müller (2000), which is based on the minimalist architecture.

- interface interpretation

The work of some of these components is taken over by the Eval component of the OT grammar. If the OT generator is designed in a minimalist fashion, then *simplification* should mean that the OT generator gets rid of some of these components. I will try to defend the following claims in this paper:

1. Formal features are the ideal candidate for such a reduction, including the associated mechanisms, feature checking, feature driven movement etc. As a consequence, the notion of ‘interface interpretation’ – the core motivation for features and feature checking –, becomes superfluous, too.
2. Simplicity is not necessarily equal to economy of representation. Markedness considerations suggest that the least marked structures are in balance between *compression* and *redundancy*.
3. A convincing theory of the syntax-prosody mapping imposes particular uniformity requirements on syntactic structures.

## 1 Simplifying the Syntactic Aparatus

### 1.1 Broekhuis 2000 (and others): Eliminate Feature Strength

In early minimalism, movement of *wh*-items is triggered by strong features (alternatively, nowadays, a *wh*-feature on a head with an EPP-feature). The dimension of feature strength (strong vs. weak feature) or the optional presence of EPP-features is not necessary under an OT approach where movement is regulated by the relative rank of the derivational economy constraint STAY. This has been demonstrated, among others, in minimalist work on *wh*-movement (Grimshaw 1997, Ackema & Neeleman 1998, Legendre et al. 1998) and Object Shift (Broekhuis 2000).

The OT generator has the task to generate a set of candidate structures for a given input. In minimalist terms, this means that structures with strong features and structures with weak features are generated in parallel. It is the task of the wellformedness constraints to select the optimal output. The ideal of a minimalist grammar is that one input (or: numeration) can only lead to one single well-formed output. Assuming feature strength (or, nowadays, EPP features) is one way to ensure this.

Broekhuis (2000) argues that one advantage of the OT model lies in the ability to derive what he calls *conditioned feature checking*. In Scandinavian, object shift, the movement of an object noun phrase outside of VP, applies if three conditions are met: (i) the verb has left the verb phrase, (ii) the object is an unstressed pronoun, (iii) no other material c-commanding the object is left within VP.

(2) Object shift in Swedish: (Holmberg 1999)

- a. *Jag kysste henne inte*  
I kissed her not
- b. ??*Jag kysste inte henne*
- c. *Jag kysste inte Marit*  
I kissed not Marit
- d. \**Jag kysste Marit inte*

Broekhuis (2000) follows earlier analyses of this phenomenon in that he assumes that the object pronoun in (2-a) moves to its case position, i.e. in a position where it checks its case feature. An early minimalist analysis would assume here that the case feature either on the noun or on the head AGR-O, which checks the case feature, must be strong in order to evoke this movement.

But then there must be an unchecked strong case feature in (3-a), which should, erroneously, lead to ungrammaticality. It further remains unclear why (3-b) is ill-formed.

- (3) a. *Jag har inte kysst henne*  
       I have not kissed her
- b. \**Jag har henne inte kysst*  
       I have her not kissed

Broekhuis (2000) shows how OT offers a way out: case is unchecked in Swedish in principle, but case movement can be triggered by another factor, here it is the constraint D-PRONOUN which requires definite pronouns to leave VP.

Broekhuis further assumes that the minimal link condition is an inviolable constraint on the generator: there will only be candidates that fulfil the MLC. This explains why (3-b) is ungrammatical: although this structure would fulfil D-PRONOUN, it will not even be generated since the object's movement outside VP violates the MLC if the verbal head has not moved out of VP itself. Broekhuis assumes the constraints CASE, which requires case features to be checked and STAY, which penalises syntactic movement (cf. Grimshaw 1997). The ranking that derives the above observations about object shift is as follows:

- (4) D-PRONOUN  $\gg$  STAY  $\gg$  CASE

The minimalist conception of feature strength is in this account replaced by the relative rank of the constraint that requires feature checking, CASE, and STAY. The high rank of D-PRONOUN leads to “conditioned feature checking”: case movement might apply for a different reason than the checking of the case feature.<sup>2</sup>

Such a reformulation of feature strength as constraint ranking has also been used in various OT accounts of *wh*-movement (cf. Grimshaw 1997, Ackema & Neeleman 1998, Legendre et al. 1998). The general picture that these accounts draw can be sketched as in (5).

<sup>2</sup> Note that Hans Broekhuis has recently revised his earlier account, adopting an approach in terms of “shape conservation”. See the discussion in his excellent paper included in this volume, section 4.2.2.

- (5) Simple economy-of-movement account of *wh*-fronting vs. *wh*-in-situ within OT:
- a. CHECK-WH  $\gg$  STAY yields *wh*-movement.
  - b. STAY  $\gg$  CHECK-WH yields *wh*-in-situ.

One might object that this is hardly more than a reformulation of the minimalist approach. This even holds, e.g., in Ackema & Neeleman's (1998) account of multiple questions, as in (6-a):

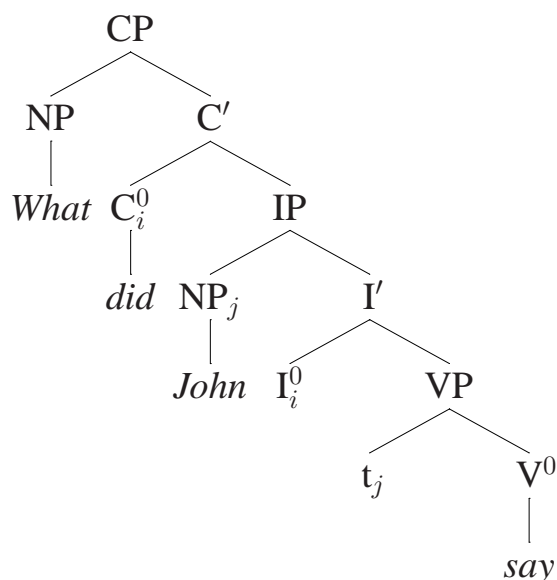
- (6) a. Who bought what?  
 b. What did you buy?

Despite the fact that the *wh*-feature on '*what*' remains unchecked, and would have to be checked in a single question (6-b), (6-a) is grammatical. Traditional generative syntax had to invent complicated devices like *wh*-absorption to explain this. In minimalism, a solution suggests itself that exploits the distinction between the checker and the checkee of a formal feature: if the [+wh] feature on  $C^0$  is strong, while that on the *wh*-phrase is weak, then we expect just one *wh*-phrase to be fronted. The OT approach by Ackema & Neeleman (1998) mimicks this by assuming a 3-constraint system, including STAY, Q-SCOPE (for the *wh*-phrase) and Q-MARKING (for the  $C^0$  head).

## 1.2 Against Economy of Movement as a Violable Constraint

A reformulation of a minimalist analysis that works in OT terms is, of course, a good thing to do. However, it is not a very forceful argument in favour of OT. There are a few further objections to be made. First, it is typical of analyses like these that they silently take over background assumptions. One concern that I have is the question how to rule out a candidate structure like the following one:

(7) What did John say?



A violation of STAY can be avoided by simply inserting the *wh*-phrase directly in [Spec,CP]. This candidate fulfills both Q-MARKING and Q-SCOPE, hence it should be optimal even (wrongly) in in-situ languages.

In minimalism and its predecessors, structure (7) is usually ruled out by interpretive and case requirements: an NP is assigned its  $\Theta$ -role inside VP, and uninterpretable otherwise. Likewise, case is assigned into that position, or another one designated for object case assignment, hence an NP inserted into [Spec, CP] has no case, or its case feature unchecked.

These options are not as straightforwardly applicable in OT. Among most varieties of OT syntax that are on the market, there is consensus at least with respect to one issue: **the input contains an argument structure specification.**

For this reason, an argument against the structure in (7) in terms of a violation of the  $\Theta$ -criterion is much less forceful than in a purely derivational system: omitted merge into  $\Theta$ -position does not lead to a loss of semantic information, if the latter is given in the input.

One principal difference between minimalism (and other purely derivational systems) and OT syntax is the construction of the interfaces between syntax on

the one hand, and semantics and phonology on the other. In the usual generative conception syntactic structures are fed into the semantic and phonological modules, which *interpret* the “instructions” the syntactic structure provides.

An OT conception of the interfaces brings syntactic, semantic and phonological/prosodic representations into *correspondence*. It organises their *mappings*. Semantic and phonological structures are generated independent of the syntactic structure, and they serve as candidates in an OT competition for the optimal syntax $\leftrightarrow$ semantics and syntax $\leftrightarrow$ phonology mappings.

Mapping requirements are typical candidates for violable OT constraints. Examples for constraints on syntax $\leftrightarrow$ semantics mapping are the constraints D-PRONOUN, Q-SCOPE and Q-MARKING, mentioned above. Such constraints can easily come into conflict, and therefore imperfect mappings are expected to be the rule rather than the exception in OT.

Hence, from the logic of an OT model, it would be a mere stipulation to claim that a constraint like  $\Theta$ -MARKING requiring arguments to be inserted in their  $\Theta$ -position was inviolable and part of the generator.

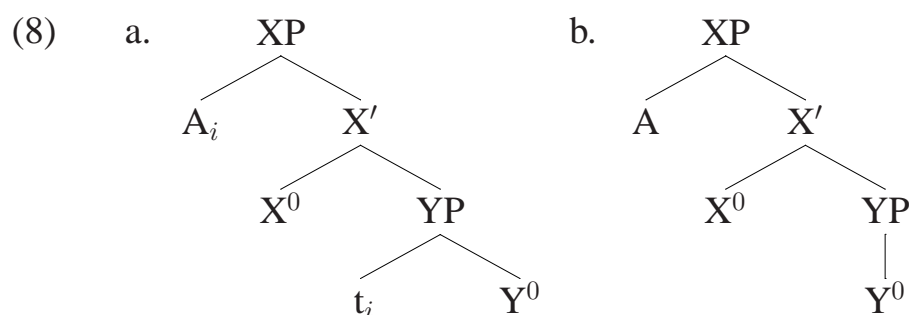
A similar argument can be made with respect to case assignment: an NP might be faced with particular syntactic ordering constraints because it has a particular case, but not necessarily in order to receive case. It might bring its case, being a morphological property, already with it.

It is thus difficult to argue that candidate (7) is ruled out by Gen, as the inviolable principle supposed to hold in Gen can hardly be motivated. Hence, economy of movement cannot help us prevent the candidate in (7) from being optimal in in-situ languages. In other words, *wh*-in-situ does not equal absence of *wh*-movement.

I therefore want to propose that there is no place for constraints like STAY, neither in Eval, nor in Gen. Syntactic movement should be evaluated by its effects *only*. It is welcome if it helps fulfilling highly ranked constraints, and disadvantageous if it leads to their violation. But these constraints should not

be about movement itself, but about the **syntactic consequences** of semantic, morphological and phonological relations among words and constituents, like, for instance, *wh*-phrase placement, syntactic conditions for case licensing and agreement, prosodic structuring etc. The impression that we have of syntax as being an economically designed system should be an emergent by-product of this, if anything.<sup>3</sup>

I want to emphasise that this does not imply the abandoning of syntactic movement per se. At this stage, there is no need to impose such a restriction on the generator. However, one conceptual issue might arise. Given that movement as such is not subject to wellformedness constraints, we might find a situation where two structurally different candidates have an identical constraint violation profile. The case I discussed above could be of this kind, or, more schematically, the following pair of trees:



This situation would be an artefact of the way the generator is defined. It would not be an empirical issue in any sense. We are well-advised to avoid such candidate pairs for conceptual reasons. The question then would be which of the two trees should be given up. Whether we can really do without syntactic movement in the generator is, however, also an empirical issue which lies beyond the scope of this article.

<sup>3</sup> This is very much in line with recent proposals by Grimshaw (2001, 2006) though she takes a different avenue to fulfil this goal.



### 1.3 An alternative account of *wh*-movement in terms of syntax-semantics correspondence

I would like to propose an alternative account of the typology of *wh*-movement that does without economy of movement. It is based on the correspondence between semantics and syntax and can be sketched as in (9).

(9) Alternative OT account of *wh*-movement:

- An object *wh*-phrase as in “What did you say?” has a couple of semantic and morphological properties (*wh*-, case,  $\theta$ -role, a.o.) which are syntactically relevant.
- Assume that case position and  $\Theta$ -position are identical and that the former is dependent on the latter, therefore omissible here.
- Assume further that Q-SCOPE is essentially the demand to have an operator c-command its scope domain.
- To capture the *wh*-movement vs. *wh*-in-situ issue, it is not necessary to refer to economy of movement. Rather, we might postulate that the *wh*-item is in conflict between which of its semantic properties determines its position, [SCOPE] or [ $\Theta$ ]. Assume two constraints, SCOPE-Pos,  $\Theta$ -Pos.
  - a. SCOPE-Pos  $\gg$   $\Theta$ -Pos derives *wh*-movement.
  - b.  $\Theta$ -Pos  $\gg$  SCOPE-Pos derives *wh*-in-situ.

This is an explanation of the typological variation of *wh*-movement in terms of conflicting semantics $\leftrightarrow$ syntax mapping demands (*wh*-scope,  $\Theta$ -role). It might be the conceptually stronger analysis in the sense that it also has something to say about the in-situ position.

Furthermore, I think this kind of approach has an empirical advantage. Note the following problem with Turkish:

[...] It should be noted that, although Turkish is an SOV language, the basic word order is overridden by various other factors. For example, the most unmarked position for a WH-element is to the immediate left of the verb, irrespective of the grammatical relation. The second-best alternative is for the WH-element to be placed in its original position; [...]

(Kornfilt 1997)

- (10) a. *bu kitab-ı kim oku-du?*  
           this book-ACC who read-Past  
       b. *kim bu kitab-ı oku-du?*  
           who this book-ACC read-Past  
           ‘Who read this book?’

As the position left adjacent to the verb is the focus position in Turkish, it is easy to integrate Turkish into our alternative account, assuming that the *wh*-phrase bears focus:

- (11) Constraint Ranking for Turkish:<sup>4</sup>  
       FOCUS-Pos  $\gg$   $\Theta$ -Pos  $\gg$  SCOPE-Pos

An analysis in terms of economy can be extended in the same way, of course. However, Turkish shows that the positioning of *wh*-items is not simply a matter of having *wh*-movement or not. The spirit of the STAY-based analysis is called into question.

The claim being made here is that the surface position of the *wh*-item is *always* determined by some semantic property, no matter which position it is. The *wh*-item bears several semantic properties with conflicting placement re-

<sup>4</sup> This assumes that (10-b) has focus on the direct object – an assumption that has been confirmed to me by Orhan Orgun, p.c.

quirements ( $\Theta$ -role, scope, focus), and the conflict is resolved in the usual OT way.

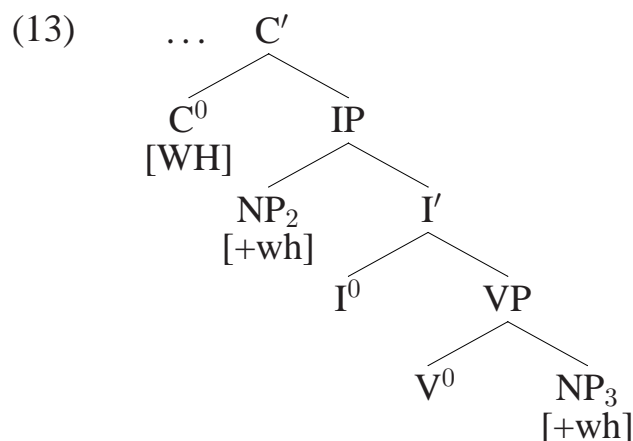
#### 1.4 Reinhart (1995): syntactic economy relativised by syntax-semantics interface needs

In this section, I would like to introduce another application of the OT model in terms of syntax $\leftrightarrow$ semantics correspondence that I illustrated in the previous section. It deals with a problem that has been discussed by Reinhart (1995). She notes the following grammaticality contrast for English:

- (12)    a. \*Bill<sub>1</sub> wonders what<sub>3</sub> who<sub>2</sub> bought.  
           b. Who<sub>1</sub> wonders what<sub>3</sub> who<sub>2</sub> bought?  
           c. Who<sub>1</sub> wonders what<sub>3</sub> Bill<sub>2</sub> bought?

This is a problem for economy of movement, as the order of the *wh*-items in the subordinate clause in (12-b) violates superiority, and hence it should be ruled out for the same reason as (12-a). But, surprisingly, the subordinate clause's subject NP does not induce a superiority violation here, just as in (12-c).

This observation about (12-b) is only correct, as long as the two embedded *wh*-phrases do not compete for the embedded [Spec,CP] position in (12-b). I.e., 'who<sub>2</sub>' has matrix scope. This distinction is difficult to integrate into a minimalist analysis, if [WH] is treated as a purely formal syntactic feature:



The [WH] Comp of the embedded clause should attract the closest [+wh] element, which is  $who_2$  in both (12-a,b). The main problem for the analysis lies in the fact that the [WH] Comp is blind for the semantic scope of the *wh*-elements it attracts.

Reinhart's solution relativises the Minimal Link Condition (MLC) to semantically equivalent syntactic structures. This interpretation of the MLC in terms of competition and blocking is already close to an OT account.

In the OT analysis in terms of semantics $\leftrightarrow$ syntax correspondence presented here, Reinhart's idea can be implemented quite nicely. Assume that there is no formal [WH] feature, no attraction of such features, and no checking. Take the constraints SCOPE-Pos and  $\Theta$ -Pos from above. Consider the following OT competition:

- (14) input: Qxy [ x wonders Qz [y bought z ] ]  
Candidate structures:
- a. \*Who<sub>x</sub> wonders who<sub>y</sub> bought what<sub>z</sub>
  - b. \*Who<sub>y</sub> does who<sub>x</sub> wonder what<sub>z</sub> bought
  - c.  $\checkmark$  Who<sub>x</sub> wonders what<sub>z</sub> who<sub>y</sub> bought

The three candidates are Reinhart's examples in (12). The input specifies the reading where both the matrix subject and the embedded subject have matrix

scope. Which is the optimal syntactic structure for this reading?

(15) OT tableau for (12)-(14):

Qxy [ x wonders Qz [y bought z ] ]	SCOPE	$\Theta$
(14-a)	*(y) *(z)	
(14-b)	*(x)	*(y)*(z)
☞(14-c)	*(y)	*(z)

Structure (14-a) has two violations of SCOPE, because neither the embedded object nor the embedded subject occupy their scope positions. Structure (14-b) has only one violation of SCOPE, because the matrix subject *wh*-phrase remains in situ. However, both the embedded object and the embedded subject occupy their scope positions and therefore incur violations of  $\Theta$ . The candidate in (14-c) exploits the fact that the matrix subject *wh*-phrase simultaneously satisfies both SCOPE and  $\Theta$  in the same syntactic position. This gives this structure the advantage of having one violation of  $\Theta$  less than (14-b), for the embedded *wh*-object, ‘*what*’. The only element that violates SCOPE is the embedded ‘*who*’. This is the optimal candidate.

(14-b) might even be worse under a definition of SCOPE that requires an operator to c-command the elements in its scope domain. This is not met, though  $\text{what}_z$  is in the embedded [Spec,CP], because one element of this domain,  $\text{who}_y$ , has moved higher – such a definition of SCOPE is thus even able to derive *wh*-island effects.<sup>5</sup>

Let me summarise the claims I have made so far:

- Syntactic constraints should formulate placement requirements as consequence of particular semantic, morphological, or (perhaps) phonological properties of syntactic elements.

<sup>5</sup> I carried out this analysis in Vogel (to appear).

- Gen does not contain any checking operations.
- Consequently, the respective features are unnecessary.

So, Gen simply consists of two operations, ...

- Gen = MERGE  $\alpha$  + MOVE  $\alpha$

... plus a restriction on vacuous movement to avoid endless generation of output-equivalent structures.

Given this, it seems that the simple generator is the unconstrained, thus (hopefully not too) powerful generator – still including the possibility of syntactic movement. Note that this is very much in common with the OT principle “Richness of the Base”: no constraints should be imposed on the generation of candidate structures. This (methodological) principle emphasises the role of Eval in OT in the effort to yield explanatory adequacy.

## 2 Syntactic Simplicity and Markedness

Both minimalism and Optimality Theory use meta-principles that are assumed to shape linguistic expressions. In minimalism, this principle is *economy*, both derivational and representational. Optimality Theory relies on the principle of *markedness*. I would like to discuss in this section how these two concepts relate.

OT’s notion of markedness is close to the traditional understanding of this term in traditional linguistics. Typical claims about the differences between marked and unmarked versions of an expression are the following:

- The unmarked expression is typologically more frequent than the marked one.
- When a language has the marked expression, it also has the unmarked expression.

- In languages that have both the marked and the unmarked expression, the contexts in which the marked expression can occur build a proper subset of the contexts in which the unmarked expression can occur.

How does syntactic simplicity correlate with this traditional conception of markedness? Optimality Theory is good at modeling so-called “repair strategies”. In phonology, this repair is typically the *neutralisation* of a marked feature, which happens under particular conditions, as, e.g., in German final devoicing – where syllable-final obstruents lose voice, e.g., /rad/ → [rat]:

- (16) German final devoicing (after Wiese 1996)  
 [+ obstruent] → [– voice] / — ]<sub>σ</sub>

As we will see in the following section, *syntactic* repair strategies are not always the unmarked option, and the unmarked option is not always the structure that is in some sense less complex, i.e., less marked in the original sense. A further issue is the relationship between analytical and synthetic expressions. Sometimes, we use syntactic means in order to fill a ‘morphological gap’. Are these syntactic means therefore less marked? And if so, why is the syntactic route often blocked, when the morphological route is available, and how can this all be integrated in a theory of syntactic markedness?

## 2.1 Optional and Obligatory Complementisers

A nice example for an unmarked-marked pair of two syntactic expressions are the two versions of English subordinate clauses, with and without complementiser, CP vs. IP. Interestingly, ‘*that*’-clauses, i.e. CPs, have to be seen as the unmarked option in the classical sense. The contexts where they are possible are a proper superset of those where the ‘*that*’-less (IP) variant is possible. For

instance, when the clause is fronted, only *that*-clauses are possible, while both forms are legitimate in the final position:

- (17)    a.    I would never say John should leave  
          b.    I would never say that John should leave  
          c.    That John should leave, I would never say  
          d.    \*John should leave, I would never say

The complementiser becomes obligatory with the insertion of an adverbial preceding the subject (cf. Grimshaw 1997):

- (18)    a.    \*She swore/insisted/thought(,) most of the time(,) they accepted this solution.  
          b.    She swore/insisted/thought that(,) most of the time(,) they accepted this solution.

The possibility of complementiser-less clauses is restricted to complements of so-called bridge verbs. Many verbs only allow for a clause with complementiser:

- (19)    a.    I regret that John left  
          b.    \*I regret John left

Considerations about the economy of representation would suggest that the version with the complementiser is the marked option, because it has more structure. This is clearly not the case. We thus conclude that the grammatically unmarked form is not always the shortest (or literally unmarked) form. There is a discrepancy between economy of structure and syntactic markedness.

The complementiser can also be understood as a clausal marker for subordination. I.e., *that*-clauses are literally marked for subordination. This observation seems to stand in opposition to the traditional notion of markedness. However,



this is also a matter of perspective. If we see the two forms as possible variants of English clauses in general, we find that *that*-less clauses can serve as both main and subordinate clauses, while *that*-clauses can only serve as subordinate clauses. *That*-less clauses, from this perspective, have the wider distribution.

All of this suggests that, especially in syntax, expressions are not marked or unmarked as such. They count as (un)marked *for a particular purpose*:

- (i) The unmarked main clause has no complementiser.
- (ii) The unmarked subordinate clause has a complementiser.

We can nevertheless make the following two statements:

- (iii) Subordinate clauses are more marked than main clauses.
- (iv) Clauses with a complementiser are more complex than those without one.

The statements in (i) and (ii) can be interpreted as the result of the interaction of the two markedness tendencies expressed in (iii) and (iv). The latter statements can be reinterpreted as scales:<sup>6</sup>

- (20)    main clause  $\prec$  subordinate clause  
           IP  $\prec$  CP

By using the method of harmonic alignment, as established by Prince & Smolensky (1993, 2004), we can construct two universally fixed sub-rankings of constraints composed by aligning the two scales appropriately:

- (21)    a. \*MainCl/CP  $\gg$  \*MainCl/IP  
           b. \*SubCl/IP  $\gg$  \*SubCl/CP

It is universally more harmonic for a main clause not to have a complementiser, and for a subordinate clause to have one. The interleaving of these two subrankings is open to typological variation. For English, it is crucial that \*MainCl/CP

<sup>6</sup> For ease of representation, I use the labels CP and IP for clauses with and without complementiser.

is ranked higher than all the other constraints, as this is the structure that never occurs.

As this analysis shows, economy of structure does indeed play a role, but perhaps not in a pure way, but only indirectly as part of a constraint subsystem that is derived by harmonic alignment. More complex structures are sometimes preferred, for instance in order to maintain a contrast.

This reminds of Horn's (1984) 'division of pragmatic labour', the observation that unmarked forms tend to be used for unmarked situations and marked forms for marked situations. When a pair of two forms stands in such a relation, the more general form will be blocked by the more specific one in a 'neutral' context. This is not the case with our two sentence types, but the next section will discuss a candidate for such an interaction, English *do*-support.

## 2.2 *Do*-support, Periphrasis, and Markedness

As we saw in the previous section, the decision which of two syntactic structures has to be considered as less marked, is not necessarily decided simply by considering structural complexity. This is also the case with the second example I would like to discuss, English *do*-support. Consider the following examples:

- (22)    a.    John left.  
         b.    \*John did leave.  
         c.    John DID leave.  
         d.    John didn't leave. / \*John left not.  
         e.    Why did John leave? / \*Why left John?

*Do*-support is the periphrastic version of a simple tense form, it alternates with the tense inflection on the verb. A couple of contexts make it obligatory – in (22), we have contrastive *verum focus* (22-c), negation (22-d), and non-subject questions (22-e). Which is the unmarked form, *do*-support or tense inflection?

If we follow the reasoning above, then the unmarked form is the one which is more widely applicable and which occurs especially in difficult environments. This is clearly the case with *do*-support. However, the unmarked expression should also be possible in an unproblematic environment. But as the judgement in (22-b) indicates, this is not the case.

These observations thus do not fit the theory of markedness in syntax that we developed thus far. I see two possible explanations for the oddity of (22-b) which are in line with our theory of syntactic markedness:

- (i) The non-acceptability of (22-b) is not an instance of syntactic illformedness, but due to pragmatic blocking.
- (ii) (22-b) is well-formed, its low acceptability is due to a prescriptive norm within the speech community.

Explanation (i), pragmatic blocking, could rely on the theory of conventional implicatures, as founded by Grice (1975), and further developed, e.g., by Levinson (2000). It can happen that two semantically equivalent forms stand in a scalar opposition. These scales are called Horn-scales after Horn (1984) who was the first to give a systematic account of such phenomena.

The example that Levinson has studied in detail is the English system of pronominal and anaphoric reference. The SELF-anaphora (*himself, herself, itself, myself* etc.) are nowadays the only option for a locally bound pronoun in English. But in Old High English, the simple pronouns *him, her, it* were still possible, i.e., ‘John shaved him’ could mean that John shaved himself. What has changed since then, according to Levinson, is the *conventionalization* of the scale ‘SELF-pronoun – pronoun’. This had the consequence that in contexts where the SELF pronoun is used, the simple pronoun is blocked.

The oddity of (22-b) could be seen as another instance of such a division of pragmatic labour. In general, I would like to propose, the synthetic form is preferred over the periphrastic form:

**Blocking of periphrastic forms** If two forms that differ only in whether they express a feature by a morpheme or by a function word, build a Horn-scale, then the form that uses the morpheme blocks the form that uses the function word.

It is striking that the syntactic structure of (22-b) is not unacceptable per se, but, as we see in (22-c), requires, or induces, an additional semantic feature, *verum focus*. This is in fact a precondition for the building of a Horn-scale: the forms involved in a Horn-scale are wellformed according to core grammatical criteria. Thus, *do*-support is syntactically wellformed, even in (22-b), but because of the division of pragmatic labour, its use induces a semantic contrast – if no such contrast is intended, the use of the dispreferred form is not justified.

While in Standard English the scale ‘*do*-support – morphological tense’ is conventionalized, there exist English dialects which are in a state comparable to Old English in Levinson’s example: they use *do*-support even in neutral environments. This has been reported by Kortmann (2002) for the southwest counties of England<sup>7</sup> where “unstressed *do* [occurs] as simple tense-carrier in affirmative sentences:

*We do breed our own cows. This man what do own this, ...*

*We’ve been up milking at 6 o’clock in the morning, and then we did go on haymaking, ...*”

Among German dialects, this phenomenon is even more widely spread, though also most German speakers will presumably agree that (23) is illformed as a Standard German sentence:

(23) ?\**Maria tut schlafen*  
M. does sleep

<sup>7</sup> Kortmann quotes Wakelin (1986), according to whom this region is mainly constituted by the counties of Cornwall, Devon, Somerset, South Avon, Wiltshire and Dorset, with East Cornwall, Devon and (West) Somerset forming its core.

(23) probably sounds to many Germans like child speech. Smaller Children use this construction quite frequently, just as their dialectal environment does. When children come to primary school, teachers spend much effort on driving them this habit out. So, for Standard German, a sociolinguistic explanation for the low acceptability of (23) seems plausible – it is the result of the exposition to prescriptive pressure at school.

### 2.3 Comparative Adjective Formation

The two versions of comparative adjective formation in English follow a pattern similar to *do*-support: short adjectives are formed with *-er*, those with 3+ syllables are built with *more*. The two options have nearly complementary distribution:

- (24)
- a. easier, \*more easy
  - b. \*intelligenter, more intelligent
  - c. luckier, more lucky

Adjectives with two syllables are somewhat in between. Via a *Google* search, it is possible to find both versions for ‘*lucky*’:

- (25)
- a. [http://www.omgclothing.com/score/36052/Liberals\\_are\\_luckier\\_in\\_love!](http://www.omgclothing.com/score/36052/Liberals_are_luckier_in_love!)
  - b. “How You Can Be More Lucky”  
([http://www.somethingyoushouldknow.net/transcript8\\_13\\_03.htm](http://www.somethingyoushouldknow.net/transcript8_13_03.htm))

Periphrastic comparatives of ‘*easy*’ can be found in coordinated adjectives:

(26) Periphrastic comparative adjectives with a disyllabic adjective:

- a. “But then turn to an open source language, inspired by Unix shell programming, but, oh, so much more easy and powerful.”  
(<http://www.awaretek.com/programming.html>)
- b. “AOSell integrates with America Online software to make researching stocks with AOL more easy and productive.”  
([http://www.softdepia.com/business\\_solutions\\_sub\\_155\\_1.html](http://www.softdepia.com/business_solutions_sub_155_1.html))
- c. “Act for the more easy and speedy recovery of small debts, within the city of Rochester, and the parishes of Strood [etc] and the ville of Sheerness”  
(<http://library.kent.ac.uk/library/special/html/specoll/acts.htm>)

This can even be observed with monosyllabic adjectives:

(27) Periphrastic comparatives with coordinated a monosyllabic adjective:

- a. “Just hope that the script kiddie graphic interface will be more nice and sober in the future.”  
([forum.sysinternals.com/forum\\_posts.asp?TID=7003&PN=1&TPN=57](http://forum.sysinternals.com/forum_posts.asp?TID=7003&PN=1&TPN=57))
- b. “Being the North the poor area, the South the more nice and old area, with medium class all over it and some old rich people also.”  
([geoimages.berkeley.edu/wwp904/html/AYRTON.html](http://geoimages.berkeley.edu/wwp904/html/AYRTON.html))
- c. “I spent around thirty hours or so working on the Everything Engine, trying to refactor it into something a little more nice and usable.”  
([www.oreillynet.com/onlamp/blog/2006/06/refactoring\\_everything\\_retrosp.html](http://www.oreillynet.com/onlamp/blog/2006/06/refactoring_everything_retrosp.html))

Like in the case of *do*-support, periphrasis is an option the system can ‘fall back’ to in a non-trivial syntactic context. Although the expressions “easier and speedier” and “nicer and older” are available, the periphrastic “more easy and speedy” and “more nice and old” are not blocked anymore. Hence, in the context of our discussion about markedness we again notice that the periphrastic form, the ‘*more*’-comparative is the one that is more widely applicable, and, thus, should count as the less marked form, despite its being blocked in the case of small adjectives in unproblematic contexts.

In the absence of a morphological strategy, the periphrastic form is not even blocked in the simple cases. This can be seen with *less*-comparatives:

- (28) a. “That’s less nice. And we hope.”  
 ([www.aquinas.ac.uk/documents/download.asp?nodeid=2631&libraryversionid=1719](http://www.aquinas.ac.uk/documents/download.asp?nodeid=2631&libraryversionid=1719))
- b. “A little less nice and a lot more nasty would have made Shallow Hal twice the film.”  
 ([www.totalfilm.com/cinema\\_reviews/shallow\\_hal](http://www.totalfilm.com/cinema_reviews/shallow_hal))
- c. “I had to make her a bit less nice and a bit more willing to make mistakes and get involved with people.”  
 ([fictionwriting.about.com/od/interviews/a/alixohlin\\_2.htm](http://fictionwriting.about.com/od/interviews/a/alixohlin_2.htm))

This is expected: without a Horn-scale, no pragmatic blocking can apply. If there was a genuinely morpho-phonological or morpho-syntactic constraint ruling out periphrastic comparatives with small adjectives, we would expect this constraint to also apply with the *less*-comparative. ‘*Less nice*’ should be ill-formed. As we see, this is false. The illformedness of ‘*more nice*’ in unproblematic contexts is thus indeed dependent on the existence of a morphological alternative – the two forms build a Horn-scale.

Summarising the discussion in the last two sections, we can state that from a purely formal perspective, periphrastic forms are less marked than synthetic forms, because they are more generally applicable. But whenever we have an alternation between morpheme and function word, and this relation has become conventionalised in the form of a Horn-scale, the less marked periphrastic form is blocked in neutral environments, due to the principle of the ‘division of pragmatic labour’. However, this is an observation about language *use*, not about grammar in the narrow sense.

## 2.4 Agreement with first and Second Person in Relative Clauses

Thus far, the results of our discussion on the relation between markedness and structural simplicity showed that periphrastic forms are the less marked forms, i.e., those forms that are more widely applicable, and the last resort the system can fall back to under difficult circumstances. Thus, richer, more explicit structures are less marked than those which are more condensed.

However, this should not mean that structural richness is less marked in general. One example of a richer, but more marked structure that occurs only as repair form are resumptive pronouns in German relative clauses. German relative pronouns are marked for third person and agree with their head noun in the  $\phi$ -features person, number and gender:

- (29) a. *Der Mann, der da steht ...*  
           the man-3SgMasc the-3SgMasc there stands
- b. *Die Frau, die da steht ...*  
           the woman-3SgFem the-3SgFem there stands

But German lacks relative pronouns in first and second person. Using the third person relative pronoun alone leads to ill-formedness, especially when an appositive relative clause is extraposed (30-a,b). The structure is repaired by in-



serting a resumptive pronoun that bears the missing person features (30-c). This option is ruled out in third person (30-d).

(30) Relative pronoun agreement with first/second person in German:

- a. *\*Ich gehe zu ihr, der sie am besten kennt.*  
 I go to her the-3SgNomMasc her at-the best knows-3Sg  
 ‘I’ll go to her, who (i.e., me) knows her best.’
- b. *\*Ich gehe zu ihr, der sie am besten kenne.*  
 I go to her the-3SgNomMasc her at-the best know-1Sg
- c. *Ich gehe zu ihr, der ich sie am besten*  
 I go to her the-3SgNomMasc I her at-the best  
*kenne.*  
 know-1Sg
- d. *?\*Peter geht zu ihr, der er sie am besten*  
 P. goes to her the-3SgNomMasc he her at-the best  
*kennt.*  
 knows-3Sg
- e. *Peter geht zu ihr, der sie am besten kennt.*  
 P. goes to her the-3SgNomMasc her at-the best knows-3Sg

While (30-a,b) are clearly odd examples, (30-d) sounds first of all ‘archaic’, as if it stemmed from a Shakespeare translation. Nevertheless, leaving the resumptive pronoun out, as in (30-e) is clearly the preferred and fully acceptable option, and this strongly contrasts with (30-a,b).

Using such a resumptive pronoun is totally ruled out in restrictive relative clauses:

- (31) *\*Ich kenne einen Mann, der er Maria kennt*  
 I know a man the-3SgNomMasc he M. knows  
 ‘I know a man who (he) knows Maria’

I conclude that the resumptive pronoun in (30-c) is a repair form that is invoked by agreement requirements. There is an agreement chain starting from the head noun of the relative pronoun, “Ich”, via the relative pronoun to the finite verb of the relative clause. Especially in order to avoid an agreement clash with the finite verb of the relative clause, the resumptive pronoun is required.

(30-c) is the syntactically more complex expression, but in this case it is also the more marked expression. Its occurrence is restricted to cases like (30-c). There is also another important difference: while in all examples that we discussed we are dealing with function words that express a feature that could be expressed by a morpheme, the feature in this latter case is agreement, i.e., a purely formal property of the relative pronoun – of course, one that it is unable to express. In the other cases above, the expressed properties were tense and comparative, i.e., semantically relevant properties.

## 2.5 Summary

Let me briefly sum up the results of this section:

Periphrastic forms where a function word expresses a semantically relevant feature are less marked than their synthetic alternatives, because they have broader application. Their avoidance in unproblematic contexts is due to the division of pragmatic labour. There has been a considerable debate about the integration of these pragmatic aspects into optimality theory, especially in the context of bidirectional OT, see for instance the paper by Blutner (2001), and the collection by Blutner & Zeevat (2004). I sketched a bidirectional model of OT syntax that is able to capture relevant aspects of Horn’s division of pragmatic labour, as they are relevant for syntactic analyses, in (Vogel 2004*a,b*).

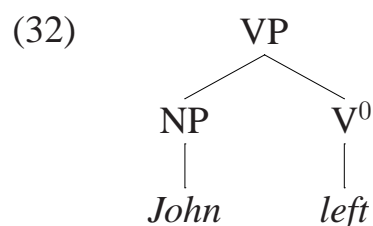
Clitic doubling, as we find it in the preceding section, is used to fulfil agreement requirements. It does not serve a semantic purpose in such a case, has an isolated range of application, and is therefore the marked option.

Structural economy in the strict sense seems to hold if function words are used to express a purely morpho-syntactic property like agreement, but not if they express semantically relevant properties like tense or comparative. Thus, it seems that the unmarked syntactic expressions are typically periphrastic constructions. However, empirically, this can be counterbalanced by the pragmatic constraints governing language use.

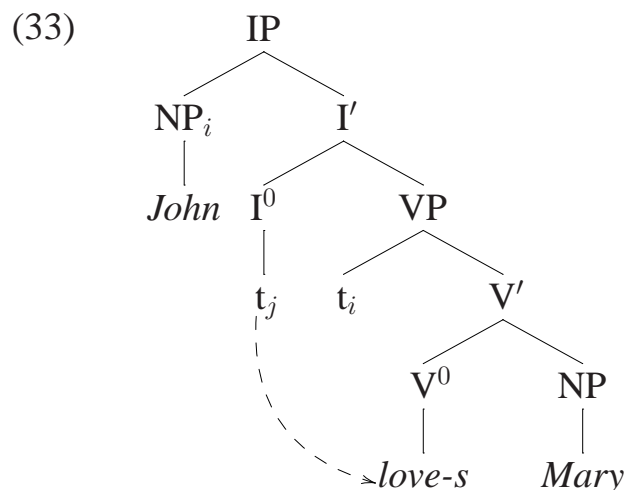
Unmarked syntactic expressions can be seen as standing in a balance between *compression* (synthetic constructions) and *redundancy* (clitic doubling).

### 3 Syntactic Simplicity and Syntax-Phonology Correspondence

A convincing theory of syntax-prosody mapping requires that syntactic stipulation and structural idiosyncracies be reduced to a minimum. One noteworthy problem arises when we apply the theory of syntax-prosody mapping by Truckenbrodt (1999) to Grimshaw's (1997) account of the English verb phrase. In Grimshaw's system, English active clauses with simple tense are analysed as simple VPs:



The standard Chomskian approach of the English main clause assumes that the inflectional affix of a finite verb is base generated under  $I^0$  and then lowered to  $V^0$  (cf. Chomsky 1981):



However, theories of syntax-prosody mapping make crucial use of the distinction between functional and lexical projections (Selkirk 1986, 1996, Truckenbrodt 1995, 1999, a.o.). Of particular importance is the *Lexical Category Condition*:

(34) **Lexical Category Condition (LCC)**

(Hale & Selkirk 1987; Truckenbrodt 1999)

Constraints relating syntactic and prosodic categories apply to lexical syntactic elements and their projections, but not to functional elements and their projections, or to empty syntactic elements and their projections.

The problem that arises when we put the things together is that Grimshaw's (1997) VP analysis for simple English active clauses renders Truckenbrodt's (1995, 1999) and Hale & Selkirk's (1987) otherwise very nice account of prosodic phrasing inapplicable to English.

In (Truckenbrodt 1995), the following two constraints are central for syntax-prosody mapping:

(35) **Wrap-XP** Each lexically headed XP is contained in a phonological phrase.

**Stress-XP** Each lexically headed XP must contain a phrasal stress.

As a consequence of the LCC, the relevant lexical XPs are VP and NP in (36):

- (36)                    ( John )<sub>PPh</sub> ( left )<sub>PPh</sub>                    prosodic structure
- a.    [IP [NP John ] [VP left ] ]                    traditional analysis
- b.    [VP [NP John ] [V<sup>0</sup> left ] ]                    analysis by Grimshaw (1997)

While the prosodic phrasing in (36) fulfils both Wrap-XP and Stress-XP under the syntactic analysis in (36-a), the one in (36-b) violates Wrap-XP for VP. The appropriate prosodic structure for (36-b) would be the wrong ( John left )<sub>PPh</sub> (underlining signals stress). Thus the theory of English syntax-prosody mapping might have to be refined, perhaps in counterintuitive ways. By the way, the same is true, if the structure is atomised into a number of functional projections. Already, if the verbal head leaves VP into some higher functional head, the consequence for prosodic phrasing might result in atomisation:

- (37)    Too many functional projections for syntax-prosody mapping:
- a.    [IP [NP John ] [FuncP loves [VP [NP Mary ] ] ] ]
- b.    ( John )<sub>PPh</sub> ( loves )<sub>PPh</sub> ( Mary )<sub>PPh</sub>
- c.    [IP [NP John ] [FuncP1 loves [FuncP2 [NP Mary ] ... ] ] ]

Certainly, this happens, when both V and the object NP move up to higher functional projections, as in (37-c).

Prosodic phrasing provides indirect evidence for syntactic structure. It is certainly true that a model of the syntax↔prosody mapping works better when syntactic analyses are as coherent, exceptionless, and surface-near as possible. The amount of hidden, ‘invisible’ structure should be reduced to a minimum, but not for the price of a loss of analytical coherence, as the discussion of Grimshaw’s (1997) VP-analysis in this section showed.

Traditional syntax models, generative or not, use maximally general descriptions of sentence patterns, very much like the ‘ $S \rightarrow NP VP$ ’ rule from early generative grammar for English sentences. Because of the generality of such patterns, introducing a different phrase structure rule for unmodified clauses in simple tenses like the one in (32) would mean an unnecessary complication, the introduction of new rule, where the old one was sufficient. The simplicity of such an analysis is only apparent. If our OT syntax model leads us to an analysis like (32), then it is quite likely that our model needs revision. In general, the more variance we introduce into our syntactic analyses, the more we have to take care of in our model of the syntax $\leftrightarrow$ prosody mapping.

#### 4 Summary

The starting point of my discussion was the shift of explanatory burden from Gen to Eval within OT. One consequence of this shift should lie in a simplification of the generator, compared to a purely derivational system like minimalism. I argued that OT’s generator can indeed do without a couple of important ingredients of minimalist theory: features, feature strength, economy of derivation, and also, to a certain extent, economy of representation. Whether it can also do without syntactic movement, is an open issue.

I proposed a correspondence theoretic conception of OT syntax, where derivational economy is reduced and relativised to syntax $\leftrightarrow$ semantics correspondence.

The concept of representational economy is also called into question from an empirical perspective: the syntactic structures that count as unmarked, according to typological and distributional criteria, often are not the ‘shortest’ ones. Syntactically unmarked structures are in balance between compression and redundancy.

Typically, periphrastic constructions are those with the broadest applicability. We further found that situations where periphrastic constructions are ruled

out have two characteristics: we have a syntactically unproblematic context, and a synthetic alternative is available. I argued that these cases should be treated as instances of the pragmatic blocking of the periphrastic form by the synthetic one. However, the (grammatical) well-formedness of the involved expressions is a prerequisite of such pragmatic blocking to apply.

In the final section, I focused on the syntax-phonology interface, especially the mapping between syntactic and prosodic constituents. The syntax↔prosody mapping in English provides another argument against a strict application of representational economy. For the theory of the syntax-phonology interface, it is more important to work with invariant syntactic structures, rather than using the shortest, simplest structure possible.

On the other hand, prosodic phrases are usually headed by lexical words. Functional categories play a prominent role in syntactic analyses, but they are nearly irrelevant for the syntax↔prosody mapping. The highly abstract syntactic structures that we can frequently observe in the current generative discourse, with a proliferation of elaborated hidden structures and derivations, leads to serious complications for the theory of the syntax-phonology interface. From that perspective, it is much better to follow the opposite route: in our assumptions about syntactic structures, we should avoid abstraction as much as possible, but, see above, use structures with maximal generality.

## Bibliography

- Ackema, Peter & Ad Neeleman (1998). 'WHOT.' *In*: Barbosa et al. (1998), pp. 15–34.
- Barbosa, Pilar, Danny Fox, Paul Hagstrom, Martha McGinnis, & David Pesetsky (eds.) (1998). *Is the best good enough? Optimality and competition in syntax*. MIT Press, Cambridge, Massachusetts.

- Blutner, Reinhard (2001). 'Some aspects of optimality in natural language interpretation.' *Journal of Semantics* 17:189–216.
- Blutner, Reinhard & Henk Zeevat (eds.) (2004). *Optimality Theory and Pragmatics*. Palgrave MacMillan, Houndmills, Basingstoke, Hampshire, England.
- Broekhuis, Hans (2000). 'Against Feature Strength: The Case of Scandinavian Object Shift.' *Natural Language and Linguistic Theory* 18:673–721.
- Chomsky, Noam (1981). *Lectures on Government and Binding*. Foris, Dordrecht.
- (1995). *The Minimalist Program*. MIT Press, Cambridge, Massachusetts.
- Culicover, Peter W. & Ray Jackendoff (2005). *Simpler Syntax*. Oxford University Press, Oxford, New York.
- Grice, H. Paul (1975). 'Logic and Conversation.' In: P. Cole & J. Morgan (eds.), 'Speech Acts,' vol. 3 of *Syntax and Semantics*, pp. 41–58. Academic Press, New York.
- Grimshaw, Jane (1997). 'Projection, Heads, and Optimality.' *Linguistic Inquiry* 28:373–422.
- (2001). 'Economy of Structure in OT.' Rutgers Optimality Archive (ROA). URL: <http://roa.rutgers.edu/view.php3?id=479>
- (2006). 'Chains as Unfaithful Optima.' In: Eric Bakovic, John J. McCarthy, & Junko Ito (eds.), 'Wondering at the Natural Fecundity of Things: Essays in Honor of Alan Prince,' pp. 97–109. BookSurge Publishing, Charleston, South Carolina. Published online by the Linguistics Research Center, University of California at Santa Cruz.  
URL: <http://repositories.cdlib.org/lrc/prince/6/>



- Hale, Kenneth & Elisabeth Selkirk (1987). 'Government and tonal phrasing in Papago.' *Phonology Yearbook* 4:151–183.
- Heck, Fabian & Gereon Müller (2000). 'Successive Cyclicity, Long-Distance Superiority, and Local Optimization.' *Proceedings of WCCFL* 19:218–231.
- Holmberg, Anders (1999). 'Remarks on Holmberg's Generalization.' *Studia Linguistica* 53(1):1–39.
- Horn, Lawrence R. (1984). 'Toward a new taxonomy for pragmatic inference: Q-based and R-based implicatures.' In: D. Schiffrin (ed.), 'Meaning, Form and Use in Context,' pp. 11–42. Georgetown University Press, Washington.
- Kornfilt, Jaklin (1997). *Turkish*. Descriptive Grammmars. Routledge, London.
- Kortmann, Bernd (2002). 'New prospects for the study of English dialect syntax: Impetus from syntactic theory and language typology.' In: Sjeff Barbiers, Leonie Cornips, & Susanne van der Kleij (eds.), 'Syntactic Microvariation,' pp. 185–213. Meertens Instituut, Amsterdam. Online Publication.  
URL: <http://www.meertens.knaw.nl/projecten/sand/synmic/>
- Legendre, Géraldine, Paul Smolensky, & Colin Wilson (1998). 'When is less more? Faithfulness and Minimal Links in WH-Chains.' In: Barbosa et al. (1998), pp. 249–289.
- Levinson, Stephen C. (2000). *Presumptive Meanings. The theory of Generalized Conversational Implicature*. The MIT Press, Cambridge, Massachusetts.
- McCarthy, John J. & Alan Prince (1995). 'Faithfulness and Reduplicative Identity.' In: Jill Beckman, Laura Walsh-Dickie, & Suzanne Urbanczyk (eds.), 'Papers in Optimality Theory,' vol. 18, pp. 249–384. UMass Occasional Papers in Linguistics, Amherst, Massachusetts.

- Prince, Alan & Paul Smolensky (1993). 'Optimality Theory. Constraint Interaction in Generative Grammar.' Rutgers Optimality Archive: ROA-537. Published as Prince and Smolensky 2004.  
URL: <http://roa.rutgers.edu/view.php3?roa=537>
- (2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Blackwell, Oxford.
- Reinhart, Tanya (1995). 'Interface Strategies.' OTS Working Papers, University of Utrecht.
- Selkirk, Elisabeth (1986). 'On derived domains in sentence phonology.' *Phonology Yearbook* 3:371–405.
- (1996). 'The Prosodic Structure of Function Words.' In: James Morgan & Katherine Demuth (eds.), 'Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition,' pp. 187–213. Lawrence Erlbaum Associates, Mahwah, New Jersey.
- Truckenbrodt, Hubert (1995). *Phonological Phrases: Their Relation to Syntax, Focus, and Prominence*. Ph.D. thesis, MIT, Cambridge, Massachusetts.
- (1999). 'On the Relation between Syntactic Phrases and Phonological Phrases.' *Linguistic Inquiry* 30:219–255.
- Vogel, Ralf (2004a). 'Correspondence in OT Syntax and Minimal Link Effects.' In: Arthur Stepanov, Gisbert Fanselow, & Ralf Vogel (eds.), 'Minimality Effects in Syntax,' pp. 401–441. Mouton de Gruyter, Berlin.
- (2004b). 'Remarks on the Architecture of Optimality Theoretic Syntax Grammars.' In: Reinhard Blutner & Henk Zeevat (eds.), 'Optimality Theory and Pragmatics,' pp. 211–227. Palgrave MacMillan, Houndmills, Basingstoke, Hampshire, England.

- (2006). ‘Weak Function Word Shift.’ *Linguistics* 44:1059–1093.
- (to appear). ‘Wh-Islands: A View from Correspondence Theory.’ *In*: Curt Rice (ed.), ‘Modelling Ungrammaticality in Optimality Theory,’ Equinox, London.
- Wakelin, M. F. (1986). *The Southwest of England*. Benjamins, Amsterdam/Philadelphia.
- Wiese, Richard (1996). *The Phonology of German*. Oxford University Press, Oxford.
- Williams, Edwin (2003). *Representation Theory*. MIT Press, Cambridge MA.

*PD Dr. Ralf Vogel*  
*Fakultät für Linguistik und Literaturwissenschaft*  
*Universität Bielefeld*  
*Postfach 10 01 31*  
*D-33501 Bielefeld*

*Ralf.Vogel@Uni-Bielefeld.de*

*<http://www.uni-bielefeld.de/~rvogel>*

# Derivations (MP) and Evaluations (OT)\*

*Hans Broekhuis*

Leiden University (LUCL)

The main claim of this paper is that the minimalist framework and optimality theory adopt more or less the same architecture of grammar: both assume that a generator defines a set  $S$  of potentially well-formed expressions that can be generated on the basis of a given input, and that there is an evaluator that selects the expressions from  $S$  that are actually grammatical in a given language  $L$ . The paper therefore proposes a model of grammar in which the strengths of the two frameworks are combined: more specifically, it is argued that the computational system of human language  $C_{HL}$  from MP creates a set  $S$  of potentially well-formed expressions, and that these are subsequently evaluated in an optimality theoretic fashion.

*Keywords: Minimalist Program, Optimality Theory, Derivation-and-Evaluation model, Object Shift.*

## 1 Introduction

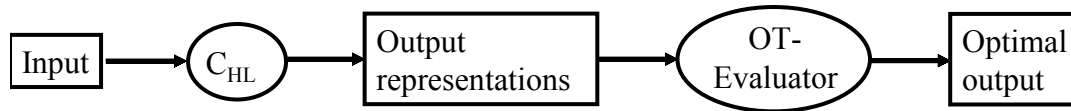
This paper describes and discusses the derivation-and-evaluation model in (1). The central idea underlying this model is that developing an explanatorily and descriptively adequate theory of syntax requires that restrictions be formulated both on the syntactic derivations and the resulting syntactic representations. This is obtained by assuming that the framework combines certain aspects of the minimalist program (MP) and optimality theory (OT). More specifically, it is

---

\* I like to thank the participants of the Workshop *descriptive and explanatory adequacy in linguistics* (DEAL), held from 17 to 19 December in Berlin, for their comments on the oral presentation of this material. This research is supported by the Netherlands Organisation of Scientific Research (NWO), grant 276-70-001, which is hereby gratefully acknowledged.

assumed that representations created by some version of the computational system of human language  $C_{HL}$  from MP are evaluated in an OT-fashion.

**Figure 1: The derivation-and-evaluation (D&E) model**



One reason for seriously investigating the properties of the D&E model in Figure 1 and for being optimistic about its explanatory and descriptive adequacy lies in the insight that whereas MP has been especially successful in formulating a restrictive theory of core grammar, that is, the universal properties of grammar as encoded in  $C_{HL}$ , OT has been very successful in describing the more peripheral, language-specific properties of languages and the variation between languages.<sup>1</sup>

The model in Figure 1 goes against the often tacitly adopted but apparently generally accepted view that MP and OT are *incompatible*, and thus competing, frameworks. In earlier work (Broekhuis and Dekkers 2000; Broekhuis 2000) I have argued, however, that MP and OT are actually *complementary* frameworks, which can therefore be advantageously combined in one overarching theory of grammar: MP is mainly a derivational theory that aims at accounting for the universal properties of language, whereas OT is rather a representational theory that focuses on the language-specific properties of language. This section will take the earlier claim even one step further, and

<sup>1</sup> This paper will use the notion of core and periphery in the sense of Chomsky and Lasnik (1977), without the implication that only the former is part of UG. On the contrary: I will adopt the OT-claim that the constraints that enter the evaluation are part of a universal constraint set CON, and that the only thing that must be acquired by the speaker is the ranking of these constraints. This also implies that the evaluator is part of the ‘core of linguistic investigation’ and that the ‘true periphery’ therefore lies outside the model in Figure 1 and consists of everything that must be learned on an item-to-item or construction-to-construction basis. This will be made explicit in Figure 10 in section 5.

argue that, despite all the differences between them, MP and OT basically assume the same kind of architecture of grammar, which comes very close to the one in Figure 1. The widely held, and in my view erroneous, belief that MP and OT are incompatible theories of grammar seems mainly due to the fact that the proponents of the two frameworks more or less exclusively focus on only one of the two components of the model in Figure 1: most work in MP focuses on properties of  $C_{HL}$ , whereas most work in OT focuses on properties of the OT-evaluator.

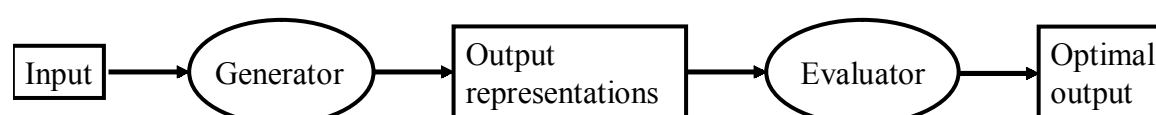
This section is organized as follows. Section 2 will substantiate the claim that MP and OT adopt essentially the same architecture of the grammar, and thus highlights the similarities between MP and OT. Section 3 discusses some differences in the research programs, and argues that these do not inherently follow from the two systems themselves. The discussion in 2 and 3 will lead to the conclusion that it is readily possible to combine MP and OT into a single overarching model of grammar, and that this gives rise to the D&E model in Figure 1. Section 4 will provide a sketch of this model, and briefly illustrate some of its properties. The discussion and claims in this paper are restricted to syntax, but it goes without saying that I believe that the proposal as worked out in section 4 should be extended to other parts of grammar like phonology (see LaCharité & Paradis 2000 for relevant discussion of the role of rules/the generator in OT-phonology).

## **2 Where MP and OT are similar: the architecture of syntax**

This section will argue that most grammars that have been developed during the *principles-and-parameters* (P&P) period of generative grammar assume the architecture in Figure 2, where the Generator and the Evaluator can be held responsible for respectively the universal and language-specific properties of

languages. The essential property of this model is that the generator defines a set  $S$  of potentially well-formed expressions that can be generated on basis of a given input, and that the evaluator selects the expressions from  $S$  that are actually grammatical in a given language  $L$ .

**Figure 2: The architecture of grammar**



The general idea has been very clearly formulated by Chomsky and Lasnik in *Filters and Control* (1977), where they argue that “to attain explanatory adequacy it is in general necessary to restrict the class of possible grammars, whereas the pursuit of descriptive adequacy often seems to require elaborating the mechanisms available and thus extending the class of possible grammars”. In order to solve this tension they propose that “there is a theory of core grammar with highly restricted options, limited expressive power, and a few parameters” next to a more peripheral system of “added properties of grammar”, which “we may think of as the syntactic analogue of irregular verbs”. Core grammar consists of the phrase structure and transformational component (the generator in Figure 2), whereas the more peripheral system consists of language-specific surface filters (the evaluator). Chomsky and Lasnik’s main claim is that the introduction of these filters contributes to the simplification of the transformational rules by bearing “the burden of accounting for constraints which, in the earlier and far richer theory, were expressed in statements of ordering and obligatoriness, as well as all contextual dependencies that cannot be formulated in the narrower framework of core grammar”.

Although the ideas about which aspects of grammar should be considered part of core grammar or part of the periphery have changed over the years (and

no doubt will change in the years to come), the gist of the proposal has survived in the more recent minimalist incarnations of the theory, where core syntax can be more or less equated with  $C_{HL}$ , and the periphery with the interface (or bare output) conditions. The task of reducing core grammar as much as possible has been very successful: the reduction of  $C_{HL}$  to its absolute minimum (internal and external merge) much contributes to the explanatory adequateness of the theory. But, as expected, the contribution of core grammar to descriptive adequacy has diminished accordingly, so that in this respect we have to rely more and more on the interface conditions.

Below, I will attempt to give a necessarily sketchy overview of the ways in which the global architecture in Figure 2 has been given shape in the various proposals that have been put forth over the last thirty years. I will start in section 2.1 with discussing some subsequent proposals within the P&P framework, and show that although the proposed grammars from the earlier period diverge in several respects from the overall structure in Figure 2, the more recent minimalist proposals more and more converge with it. After this I will give a brief discussion of OT in section 2.2, which fits neatly to the global architecture in Figure 2, which is clear from the in fact that some version of it can be found in virtually all introductory texts on OT.

## 2.1 Principles & Parameter Theory

Since Chomsky and Lasnik (1977), the global organization of the different P&P models has had more or less the shape given in Figure 2 above, although in the earlier proposals this is masked by the fact that instead of a fully linear model, a so-called T- or inverse Y-model was assumed, according to which the derivation of the LF- and the PF-representation diverge after a certain point (s-structure or Spell-Out). This property of the early P&P models disappears in the later versions of MP with the introduction of mechanisms like feature movement,

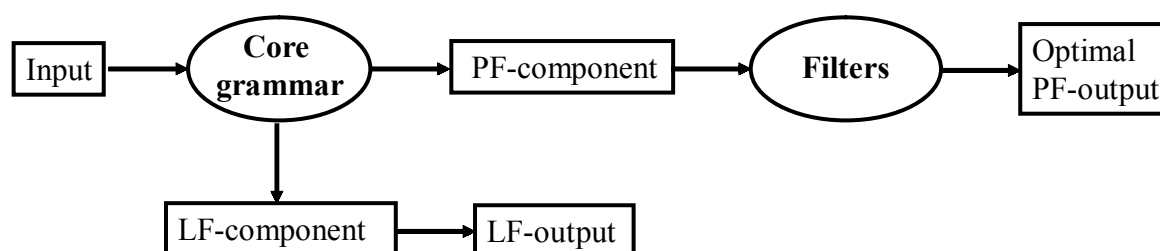


spell out of copies and Agree, which void the need of covert movement. As a result, these later versions fully accord with the essentially linear model in Figure 2.

The answers to the question what is part of the generator and what is part of the evaluator have of course changed over the years. The *that*-trace filter, for example, was originally proposed as a language-specific filter for English, but the Empty Category Principle, which ultimately grew out of it, was rather assumed to be part of core grammar. Furthermore, it is not always easy to determine which ingredients were considered part of generator and which of the evaluator since these were normally not discussed in these terms. It is clear, however, that at least the phrase structure and transformational component have consistently been considered part of the generator in all proposals so far.

In what follows I will compare the various stages of the P&P framework with the global architecture in Figure 2. First consider the model adopted by Chomsky and Lasnik in *Filters and Control*, which is given in Figure 3 below.

**Figure 3: The *Filters and Control* model (Chomsky and Lasnik 1977)**



The input of the system is a set of lexical items. The generator contains a phrase structure and a transformational component. The phrase structure component consists of phrase structure rules constrained by X-bar-theory, which combine the lexical elements from the input into a d-structure representation. The transformational rules are constrained by a set of general conditions and modify the d-structure representation into an s-structure representation, which is

subsequently fed to the LF- and the PF-component of the grammar, where it undergoes further computation. The LF-wing of the grammar contains rules that assign a semantic interpretation to the s-structure representation, for example, rules of construal (binding and control) and quantifier interpretation. The PF-wing of the grammar contains rules that assign a phonetic interpretation to the s-structure representation. Among these phonological rules we find deletion and stylistic rules. The language-specific filters, finally, evaluate the resulting PF-representations: only those representations that pass these filters are acceptable in the language under discussion.

The introduction of a filter component was motivated by the fact that this made a more restrictive formulation of core grammar possible by eliminating ordering statements and language-specific properties from the transformational component of the core grammar. By way of demonstration let us consider the derivation of the relative clauses in (1).

- (1) a. the man who I know  
b. the man that I know  
c. the man I know  
d. \*the man who that I know

The relative pronoun *who* is generated in the regular object position, so that the d-structure of the examples in (1) is as given in (2a). Chomsky and Lasnik further propose that universal grammar (UG) contains a universal principle “Move *wh*-phrase” that requires that relative pronouns (and other *wh*-phrases) be placed to the left of the complementizer, as in the s-structure representation in (2b). The examples in (1) can now be derived by assuming a deletion rule that freely deletes the relative pronoun *who* or the complementizer *that*. The resulting PF-representations are given in (3). Chomsky and Lasnik further assume the language-specific Doubly Filled COMP Filter, which prohibits the

simultaneous realization of the relative pronoun and the complementizer. This excludes representation (3d).

- (2) a. the man [that I know who] (d-structure)  
 b. the man [[COMP who that] I know  $t_{\text{who}}$ ] (s-structure)
- (3) a. the man [[COMP who ~~that~~] I know  $t_{\text{who}}$ ]  
 b. the man [[COMP ~~who~~ that] I know  $t_{\text{who}}$ ]  
 c. the man [[COMP ~~who that~~] I know  $t_{\text{who}}$ ]  
 d. \*the man [[COMP who that] I know  $t_{\text{who}}$ ]

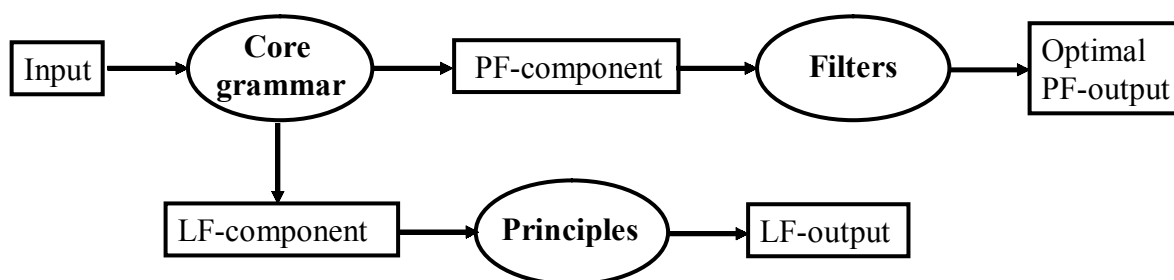
Although the deletion rule is freely applicable in principle, the resulting representation is subject to a recoverability principle, which requires that deleted elements be locally recoverable. This is needed to block deletion of the *wh*-phrase in representations like (4): the recoverability principle in tandem with the Doubly Filled COMP Filter ensures that the examples in (4b-d) are excluded. By the same means, deletion of the preposed PP in relative clauses like (5) is blocked. Deletion of *about which* would violate the recoverability principle because the preposition *about* cannot be recovered locally.

- (4) a. I wonder [who ~~that~~ you met  $t_{\text{who}}$ ]  
 b. \*I wonder [~~who~~ that you met  $t_{\text{who}}$ ]  
 c. \*I wonder [~~who that~~ you met  $t_{\text{who}}$ ]  
 d. \*I wonder [who that you met  $t_{\text{who}}$ ]
- (5) a. the book [about which ~~that~~ he spoke  $t_{\text{about which}}$ ]  
 b. \*the book [~~about which~~ that he spoke  $t_{\text{about which}}$ ]  
 c. \*the book [~~about which that~~ he spoke  $t_{\text{about which}}$ ]  
 d. \*the book [about which that he spoke  $t_{\text{about which}}$ ]

The virtue of Chomsky and Lasnik's proposal of the data above is that by accounting for the language-particular properties of the English constructions by means of the Doubly Filled COMP Filter, we can keep the transformational rule that derives s-structure (2b) maximally simple (Move *wh*-phrase), which makes it possible to attribute this rule to UG.

In the *Government-and-Binding* (Chomsky 1981) and *Barriers* (Chomsky 1986) period, the model of grammar remains essentially the same. The attempts to further reduce the transformational component of the core grammar led to the formulation of the general rule Move  $\alpha$ . As far as the filter component was concerned, it turned out that some of the filters proposed in Chomsky and Lasnik (1977) had a wider application and could be reformulated as more general principles. For example, the so-called *that*-trace filter, which prohibits a trace immediately to the right of the complementizer *that*, was reformulated as/reduced to the Empty Category Principle (ECP), which requires that a trace be properly governed. This change is depicted in Figure 4.

**Figure 4: The LGB/Barriers model (Chomsky 1981/1986)**



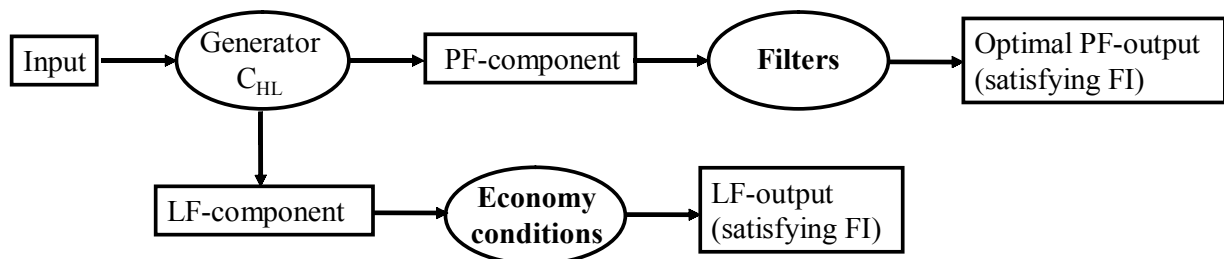
Although the ECP was claimed to be universal, that is, to be part of UG, its function is more or less the same as that of the *that*-trace filter: it excludes structures that have been created by core grammar. Therefore the formulation of the ECP is not a reason to frown with a skeptical eye on the notion of filter: it should rather give us hope that also in the domain of filters a certain degree of explanatory adequacy can be obtained.

In the *Minimalist Program*, as developed by Chomsky since the mid 80's, core grammar seems to have been reduced to its absolute minimum. The computational system of human language  $C_{HL}$ , as it is now called, consists of essentially one merge operation in two guises. External merge combines two independent syntactic objects into a larger syntactic unit, whereas internal merge

takes some element from an existing syntactic object, and merges it to the root of this object, thus deriving the effect of movement. Merge is subject to a number of general conditions. For example, it never involves more than two elements at the same time, which results in binary branching phrase structures. Internal Merge obeys certain locality restriction and is further subject to the Last Resort Condition, which requires that movement be triggered by some uninterpretable/unvalued formal feature. As in Chomsky and Lasnik (1977), descriptive adequacy lies mainly outside the core system: for example, Chomsky (1995:§4.7.3) suggests (rightly or wrongly) that ‘rearrangement’ phenomena like extraposition, right-node raising, VP-adjunction and scrambling are essentially the result of stylistic rules of the phonological component.

Although the notion of filter is not used, MP also heavily relies on the filter component. It seems that this filter component has taken various guises in the various stages in the development of the program. The organization of grammar in Chomsky (1995:ch.3) is more or less as indicated in Figure 5.

**Figure 5: The early MP model (Chomsky 1995:ch.3)**



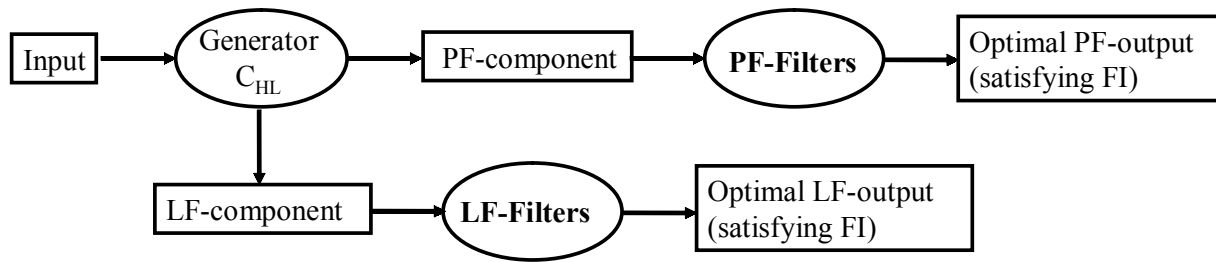
Many of the filters as discussed in Chomsky and Lasnik (1977) have not found an alternative account in MP, but the fact that they are not *discussed* is, of course, no guarantee that they are not *needed*: this motivates the postulation of a set of PF-filters in Figure 5. Furthermore, Chomsky (1995) explicitly assumes that  $C_{HL}$  generates a set of converging (= potentially well-formed) derivations satisfying Full Interpretation, the so-called *reference set*. It is further assumed

that the optimal output is the representation that satisfies a number of global economy conditions best: derivations with a smaller number of derivational steps are preferred (fewest steps), as are derivations with shorter movement chains (shortest steps).

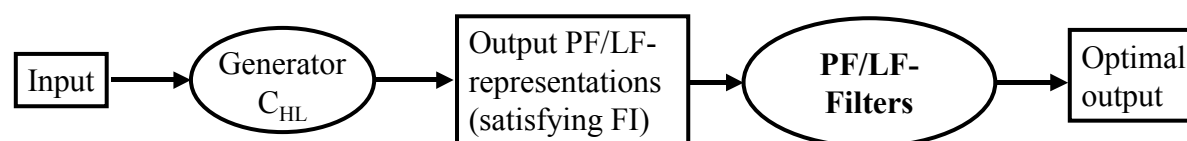
The language  $L$  thus generates three relevant sets of derivations: the set  $D$  of derivations, a subset  $D_C$  of convergent derivations of  $D$ , and a subset  $D_A$  of admissible derivations of  $D$ . FI determines  $D_C$ , and the economy conditions select  $D_A$ . [...]  $D_A$  is a subset of  $D_C$  (Chomsky 1995:220).

It is not so clear in how far the global economy conditions still play a role in the current formulation of MP. It seems that very soon they lost independent status by being successfully incorporated into the definition of the movement operation. Fewest steps was replaced by Last Resort (Chomsky 1995:280) and shortest steps by the Phase Impenetrability Condition in Chomsky (2001). As a result,  $D_C$  and  $D_A$  can be considered identical and we are left with only two sets of derivations: the set of derivations  $D$  and the set of converging derivations  $D_C$ .

Another important innovation in Chomsky (1995:ch.4, 221) is the introduction of the bare output conditions, which are later normally referred to as the interface conditions. According to Chomsky, these interface conditions are “imposed from the outside” by the performance systems that make use of the representations created by  $C_{HL}$ , and which include (perhaps at most) the articulatory-perceptual and the conceptual-intentional system. Chomsky claims that the interface conditions may be involved in the displacement property of language, and we will see in the discussion of (10/19) below that in later work, he formulates these conditions in the format of a filter on the output of  $C_{HL}$  (Chomsky 2001). So let us provisionally assume that the interface conditions can be formulated as filters on the output of the PF- and the LF-component:

**Figure 6: The later MP model (Chomsky 1995:ch.4)**

As was noted at the beginning of this section, a conspicuous property of the P&P models discussed above is that they differ from the linear model in Figure 2 in that the derivation of the PF- and LF-representations diverge at a certain point in the derivation in order to account for the fact that there can be certain mismatches between linear order and semantic interpretation. Very soon in the development of MP proposals have been put forth to eliminate this property from the grammar. Groat and O’Neil (1996), for example, noted that the copy theory of movement made it possible to account for the discrepancies in PF and LF-representations by assuming that phonology could either spell out the lower or the higher copy in a movement chain (cf. also Bobaljik 2002). Chomsky (1995: chapter 4) noted that economy considerations can account for these discrepancies by assuming that it is more economical to move a syntactic category without its phonological features, pied piping of the phonological features being possible only when there are independent reasons to do so. Finally, the introduction of Agree (feature checking at a distance) in the so-called *Minimalist Inquiry* framework (Chomsky, 2000, and subsequent work) made overt movement totally superfluous from the point of view of core grammar. As a result of this we can assume that the derivation of the LF- and PF-representations proceed in fully parallel fashion. The model of grammar assumed in this framework is therefore as indicated in Figure 7.

**Figure 7: The Minimalist Inquiry model (Chomsky 2000 and later)**

Since Agree makes movement superfluous as far as core grammar is concerned, movement must be forced by external factors, more specifically by the interface conditions imposed on the output representations of  $C_{HL}$ . Actually, the intuition underlying this proposal is much older than the Minimalist Inquiry framework. For example, it has been argued that the motivation for *wh*-movement is that a *wh*-phrase can only be interpreted if it heads an operator-variable chain; cf. e.g. Chomsky (1991:440) and Rizzi (1996). Chomsky (2001) aims at showing that also certain types of A-movement are externally motivated. We will look at this in some detail in what follows.

According to MP, movement of a syntactic object *S* is subject to last resort: it must be triggered by some unchecked or unvalued formal feature of a higher functional head *H* that can be checked or valued by a corresponding feature of *S*. In the earliest proposal it was assumed that these features of *H* come in two forms: weak and strong features. A strong feature on *H* must be checked before the projection of *H* is merged with some higher head; if checking does not take place, the derivation is canceled. A weak feature on *H*, on the other hand, cannot be checked before Spell-Out as a result of the economy condition Procrastinate. This proposal led to a very rigid system in which the question whether a certain movement does or does not apply is mechanically determined by the feature constellation of the functional head *H*. However, it is clear that movement may be sensitive to other factors as well. Consider the case of so-called object shift (OS) in the Icelandic examples in (6).



- 
- (6) a. Jón keypti ekki bókina. *bókina*  $\subset$  focus  
       Jón bought not the book  
       b. Jón keypti bókina<sub>i</sub> ekki t<sub>i</sub> *bókina*  $\subset$  presupposition

The examples in (6) demonstrate that it is possible in Icelandic to move the direct object to the left, across the negative adverb *ekki*. This movement is, however, not obligatory and depends on the information structure of the clause: OS applies only when the object is part of the presupposition (‘old’ information) of the clause; it is excluded when it is part of the focus (‘new’ information) of the clause.

Let us provisionally assume that OS is triggered by the case feature on the light verb  $v^*$  (Vikner 1994; Chomsky 2001): if this case feature were strong, we wrongly expect this movement to be obligatory; if it were weak, we wrongly predict it to be impossible. In order to account for the apparent optionality of OS, we must therefore introduce additional means. One possibility would be to make the strength of the case feature sensitive to the information structure of the clause: only when the object is part of the presupposition of the clause does  $v^*$  have a strong case feature. Apart from being *ad hoc*, this option is not descriptively adequate since OS is never possible in complex tense constructions like (7): OS is excluded irrespective the information structure of the clause, and (7a) is therefore ambiguous.

- (7) a. Jón hefur ekki keypt bókina. ambiguous  
       Jón has not bought the book  
       b. \*Jón hefur bókina ekki keypt t<sub>bókina</sub> —

Another possibility is to follow Holmberg (1999) in claiming that OS is actually not part of core grammar. He proposes that OS is a *phonological* operation that is driven by the interpretation of the object: in the terminology used above, OS is only possible if the object is part of the presupposition of the clause. This is stated in (8a), which paraphrases Chomsky’s (2001:(54a)) summary of

Holmberg's claim. Holmberg (1999:22) accounts for the ungrammaticality of (7b) by postulating the additional restriction on the application of OS in (8b): OS is blocked in (7b) because it would move the object across the main verb.

- (8) a. Object shift is a phonological movement that satisfies condition (8b) and is driven by the semantic interpretation INT of the shifted object:  
       (i) INT: object is part of the presupposition of the clause.  
       (ii) INT': object is part of the focus of the clause.  
       b. Object shift cannot apply across a phonologically visible category asymmetrically c-commanding the object position except adjuncts.

Chomsky (2001:32) argues that Holmberg's proposal is problematic because "displacement rules interspersed in the phonological component should have little semantic effect" (p.15), and he therefore develops a proposal according to which OS takes place in core syntax. The relevant configuration is given in (9), where Obj is the  $\theta$ -position of the object, and XP is a specifier position of  $v^*$  created by OS (note that Chomsky assumes a multiple specifier approach).

- (9) ... [ <sub>$\alpha$</sub>  XP [Subject  $v^*$  [V ... Obj ]]]

Note that (9) is an intermediate stage in the derivation: at some later stage in the derivation the subject is moved into SpecTP; in simple tense constructions the  $v^*+V$  complex is moved to T. Given this, Chomsky (2001:(61)) tries to account for the properties of Icelandic OS in (8) by adopting the assumptions in (10), where INT and INT' are defined as in (8a).

- (10) a.  $v^*$  is assigned an EPP-feature only if that has an effect on outcome.  
       b. The EPP position of  $v^*$  is assigned INT.  
       c. At the phonological border of  $v^*P$ , XP is assigned INT'.

The EPP-feature mentioned in (10a) has the same function as the strong features in the earlier proposals in the sense that it forces movement of some element into a specifier position of the head that it is assigned to. The statement in (10a) must be considered an invariant principle of grammar, which expresses that  $v^*$  is

only assigned an EPP-feature if the resulting movement has some effect on the output representation. According to Chomsky this is only the case when the movement affects the interpretation of the clause, or when it makes A'-movement possible (by placing the object at the phonological edge of the  $v^*P$ -phase). We will see shortly that this leads to a less rigid system in the sense that movement can be made sensitive to factors other than the feature constellation of the attracting head.

Chomsky claims that also (10b) is an invariant principle: in the terminology employed earlier, this claim expresses that an object occupying the position XP in (9) must be construed as being part of the presupposition of the clause. It is important to note that (10b) is only concerned with shifted objects, and leaves open the option that non-shifted objects are ambiguously interpreted as being part of either the focus or the presupposition of the clause. This is needed in order to allow the non-shifted objects in Icelandic examples like (7a) to be interpreted as part of the presupposition of the clause, and, of course, also correctly predicts that the objects in languages like English, which do not have OS of the Icelandic sort, can be part of either the focus or the presupposition of the clause.

Given that (10b) does not restrict the interpretation of non-shifted objects, we need something in addition to account for the fact that OS is obligatory in examples like (6b). This is where (10c) comes in. Let us first consider the notion of phonological border, which is defined as in (11).

- (11) XP is at the phonological border of  $v^*P$ , iff:
- a. XP is a  $v^*P$ -internal position, and;
  - b. XP is not c-commanded by  $v^*P$ -internal phonological material.

The main difference between the examples in (6) and (7) is that in the former the main verb has moved out of  $v^*P$  into T, whereas in (7) it has not and thus occupies a  $v^*P$ -internal position. Example (7a) is therefore correctly predicted to

be ambiguous: since the  $v^*+V$  complex is  $v^*P$ -internal and c-commands the object, clause (10c) does not apply and the object can be interpreted either as part of the focus of the clause (INT') or as part of the presupposition of the clause (INT). Example (7b) is consequently blocked by (10a) because OS has no effect on the outcome as the object can also be assigned the interpretation INT in its base position in (7a). Therefore, in constructions like (7), the EPP-feature can only be assigned to  $v^*$  if it is needed to enable A'-movement. In (6), on the other hand, there is no  $v^*P$ -internal phonological material that c-commands the position Obj. Consequently, if the object occupies this position, (10c) states that it must be assigned INT'. Movement of the object into the XP-position in (9) therefore has an effect on the outcome, and (10a) consequently allows assignment of an EPP-feature to  $v^*$ .

It is important to note that statement (10c) clearly functions as a filter in the sense of Chomsky and Lasnik (1977). First, it is clear that it cannot be considered a condition on the derivation: when we apply it to the intermediate stage in (9), the desired distinction between (6) and (7) could not be made locally (in the sense of Collins 1997), because the verb and the subject are moved out of the  $v^*P$  only at a later stage in the derivation. Chomsky therefore assumes that it applies at the higher phase level (CP). Second, (10c) is a language-specific statement: Icelandic (and the continental Germanic languages) is subject to it, and therefore OS is forced in examples like (6b); the Romance languages, on the other hand, are not subject to it, so that (10a) blocks OS in comparable Romance examples. Thus, statement (10c) has two characteristic properties of the PF-filters proposed Chomsky and Lasnik (1977). It differs from these filters in that it is sensitive both to phonological and to semantic information. But this is, of course, to be expected if filters in one way or another

reflect the fact that the output of  $C_{HL}$  is fed to both the articulatory-perceptual and the conceptual-intentional system.

This subsection has shown that all grammars proposed during the P&P era have the global architecture of grammar indicated in Figure 2, although this was obscured in the early period by the fact that it was assumed that the derivation of the PF- and LF-representation diverge at some point in the derivation. It has been shown that by rejecting this assumption Chomsky's recent Minimalist Inquiry framework fully conforms to the architecture in Figure 2 in that the grammar consists of a generative component that creates representations that are subsequently evaluated by a filter component. The filters place both semantic and phonological constraints on the output of  $C_{HL}$ , which reflects the fact that the representation(s) that pass these filters are subsequently fed to the articulatory-perceptual and the conceptual-intentional system where they undergo further computation in order to receive a phonetic and a semantic interpretation.

## 2.2 Optimality Theory

Optimality theory fits nicely to the global architecture of grammar in Figure 2, which is clear from the fact that it can actually be found in virtually all introductory texts on OT. Nevertheless, it is certainly not easy to describe the substantive contents of each of the components mentioned in the model. The input, for example, depends on the part of grammar we are talking about. For phonology, for example, it is generally assumed that the input consists of underlying phonological representations, which is of course not suitable for syntax. But even if we restrict our attention to syntax, it is clear that there is hardly any consensus on the question what the nature of the input is: in some proposals it is assumed that the input is constituted by a set of lexical elements comparable to the numeration in MP, in other proposals the input is a structured

meaning, and sometime it is even assumed that the input consists of prefabricated syntactic representations (thus leaving open the question how these are created).

Something similar holds for the generator. McCarthy and Prince (1993) assume that the generator consists of linguistic operations subject to “very general considerations of structural well-formedness”. As a rule we only find scattered remarks on the nature of these operations and the restrictions they are subject to: Grimshaw (1997), for example, claims that the generator builds structures in accordance with some version of X-bar-theory. We can therefore conclude that the generator is still largely unanalyzed in optimality theory, certainly where syntax is concerned. Nevertheless, it is crucial that the generator is an overgenerating system. It creates a so-called *candidate set* from which the evaluator selects the optimal candidate(s). It is generally assumed that this candidate set is infinite and contains many candidates that will never surface because they are harmonically bound by some other candidate, where A is harmonically bound by B if A violates at least one constraint on top of the constraints violated by B.

In optimality theory the focus of attention is on the evaluator. It consists of a set of constraints with the properties in (12a-c), which I will more extensively discuss below.

- (12) The optimality theoretic evaluator contains constraints that:
- a. are taken from a universal set of constraints CON;
  - b. are violable, and;
  - c. have a language-specific ranking.

The constraints crucially differ from the language-specific filters assumed in the *principle-and-parameters* theories in that they are generally assumed to be universal, that is, part of UG. It is assumed that there is a universal set of constraints CON from which the constraints that are active in a given language

are taken (normally it is assumed that *all* constraints from CON are active, but that the effects of some constraints are simply not observable). The constraints can nevertheless be used to express language-specific properties due to the two other properties of the constraints: according to (12b) and (12c) languages may differ in the ranking of these constraints, whereby violation of a lower ranked constraint is tolerated in order to satisfy a higher ranked constraint.

The way the OT-evaluator works can readily be demonstrated by means of Pesetsky's (1997;1998) analysis of relative clauses. This will also give me the opportunity to show how the OT-evaluator differs from the filters assumed in the P&P approaches. Consider again the relative clauses from example (1/3) and (5), repeated here as (13) and (14), which were accounted for in *Filters and Control* by taking recourse to the Doubly Filled COMP Filter and the recoverability condition on deletion.

- (13) a. the man [[COMP who ~~that~~] I know  $t_{\text{who}}$ ]  
 b. the man [[COMP ~~who~~ that] I know  $t_{\text{who}}$ ]  
 c. the man [[COMP ~~who that~~] I know  $t_{\text{who}}$ ]  
 d. \*the man [[COMP who that] I know  $t_{\text{who}}$ ]
- (14) a. the book [about which ~~that~~ he spoke  $t_{\text{about which}}$ ]  
 b. \*the book [~~about which~~ that he spoke  $t_{\text{about which}}$ ]  
 c. \*the book [~~about which that~~ he spoke  $t_{\text{about which}}$ ]  
 d. \*the book [about which that he spoke  $t_{\text{about which}}$ ]

When we contrast these examples with the French relative clauses in (15) and (16), we see that English and French differ in that the former allows a wider variety of constructions with a bare relative pronoun than the latter. However, when the relative pronoun is embedded in a PP (or an NP), the two languages behave the same.

- (15) a. \*l'homme [qui<sub>i</sub> ~~que~~ je connais t<sub>i</sub>]  
 b. l'homme [~~qui~~<sub>i</sub> que je connais t<sub>i</sub>]  
 c. \*l'homme [~~qui~~<sub>i</sub> ~~que~~ je connais t<sub>i</sub>]  
 d. \*l'homme [qui<sub>i</sub> que je connais t<sub>i</sub>]
- (16) a. l'homme [avec qui<sub>i</sub> ~~que~~ j'ai dansé t<sub>i</sub>]  
 b. \*l'homme [~~avec qui~~<sub>i</sub> que j'ai dansé t<sub>i</sub>]  
 c. \*l'homme [~~avec qui~~<sub>i</sub> ~~que~~ j'ai dansé t<sub>i</sub>]  
 d. \*l'homme [avec qui<sub>i</sub> que j'ai dansé t<sub>i</sub>]

In order to account for the data in (13) to (16), Pesetsky proposed the constraints in (17), which I slightly simplify here for reasons of exposition. Constraint (17a) is simply the recoverability condition on deletion from Chomsky and Lasnik (1977), constraint (17b) is a constraint that expresses that embedded clauses tend to be introduced by a complementizer, and (17c) is a constraint that expresses that function words (like complementizers) tend to be left unpronounced.


- (17) a. RECOVERABILITY (REC): a syntactic unit with semantic content must be pronounced unless it has a sufficiently local antecedent.  
 b. LEFT EDGE (CP): the first leftmost pronounced word in an embedded CP must be the complementizer.  
 c. TELEGRAPH (TEL): do not pronounce function words.


The ranking of these constraints will determine the optimal output. In order to see this, it is important to note that LE(CP) in (17b) and TEL in (17c) are in conflict with each other: the first wants the complementizer to be pronounced, whereas the latter wants it to be deleted. Such conflicts make it possible to account for variation between languages: when we rank these constraints differently, we get languages with different properties. When we assume that LE(CP) outranks TEL, we get a language in which embedded declarative clauses must be introduced by a complementizer. When we assume that TEL outranks LE(CP), we get a language in which embedded declarative clauses are not introduced by a complementizer. When we assume that the two constraints are in a tie (ranked equally high), we get a language in which embedded declarative



clauses are optionally introduced by a complementizer. The evaluation can be made visible by means of a tableau. Tableau 1 gives the evaluation of embedded declarative clauses with and without a pronounced complementizer in a language with the ranking  $LE(CP) \gg TEL$ .

**Tableau 1: no complementizer deletion in embedded declarative clauses**

	LE(CP)	TEL
.... [ complementizer ....] 		*
.... [ <del>complementizer</del> ....]	*!	

The two asterisks indicate that the constraint in the header of their column is violated. The first candidate, with a pronounced complementizer, violates TEL but this is tolerated because it enables us to satisfy the higher ranked constraint LE(CP). The second candidate, with a deleted complementizer, violates LE(CP), but this is fatal (which is indicated by an exclamation mark) because the first candidate does not violate this constraint. The first candidate is therefore optimal, which is indicated by means of the pointed finger: . The shading of the cells indicates that these cells do not play a role in the evaluation; this convention is mainly for convenience, because it makes it easier to read the tableaux.

Now consider the evaluation of the same candidates in a language with the ranking  $TEL \gg LE(CP)$ , given in Tableau 2. Since TEL is now ranked higher than LE(CP), violation of the former is fatal, so that deletion of the complementizer becomes obligatory.

**Tableau 2: obligatory complementizer deletion in embedded declarative clauses**


	TEL	LE(CP)
.... [ complementizer ....]	*!	
.... [ <del>complementizer</del> ....] 		*

Tableau 3 gives the evaluation of a language in which the two constraints are in a tie  $TEL \diamond LE(CP)$ , which is indicated in the tableau by means of a dashed line. Under this ranking, the rankings  $LE(CP) \gg TEL$  and  $TEL \gg LE(CP)$  are in a sense simultaneously active. Therefore we have to read the tie in both directions: when we read the tie from left to right, the violation of  $LE(CP)$  is fatal (which is indicated by  $>$ ), and the first candidate is optimal; when we read the tableau from right to left, the violation of  $TEL$  is fatal (which is indicated by  $<$ ), and the second candidate is optimal. This correctly predicts that deletion of the complementizer is optional in this case.

**Tableau 3: optional complementizer deletion in embedded declarative clauses**

	LE(CP)	TEL
.... [ complementizer ....] 		$<^*$
.... [ <del>complementizer</del> ....] 	$^*>$	


Let us now return to the difference between English and French with respect the pronunciation of relative clauses. It is clear that English has the tied ranking  $TEL \diamond LE(CP)$ , given that the complementizer is normally optional in embedded declarative clauses. In French, on the other hand, it is clear that  $LE(CP)$  outranks  $TEL$  given that the complementizer is obligatory in embedded declarative clauses. Pesetsky (1997) has shown that this also accounts for the differences between the English and French examples in (13) and (15), in which a bare relative pronoun is preposed. Assume that in both languages the constraint RECOVERABILITY outranks the constraints  $TEL$  and  $LE(CP)$ ; the ranking of the constraints in (17) are then as given in (18).

- (18) a. French:  $REC \gg LE(CP) \gg TEL$   
 b. English:  $REC \gg TEL \diamond LE(CP)$

The evaluation of the French examples in (15) proceeds as in Tableau 4. Since the relative pronoun has a local antecedent it is recoverable after deletion, so that




all candidates satisfy REC. The second candidate is the optimal candidate because it is the only one that does not violate LE(CP); the fact that this candidate violates the lower-ranked constraint TEL is tolerated since this in fact enables the satisfaction of the higher-ranked constraint LE(CP).

**Tableau 4: Relative clauses with preposed relative pronoun**

French	REC	LE(CP)	TEL
l'homme [qui <sub>i</sub> <del>que</del> je connais t <sub>i</sub> ]		*!	
l'homme [ <del>qui</del> <sub>i</sub> que je connais t <sub>i</sub> ] 			*
l'homme [ <del>qui</del> <sub>i</sub> <del>que</del> je connais t <sub>i</sub> ]		*!	
l'homme [qui <sub>i</sub> que je connais t <sub>i</sub> ]		*!	*

The evaluation of the English examples is slightly more complex than that of French due to the fact that LE(CP) and TEL are in a tie: we are therefore dealing with two rankings at the same time: REC >> LE(CP) >> TEL and REC >> TEL >> LE(CP). The first ranking is actually the one we also find in French, and we have seen that this results in selection of the second candidate as optimal. Under the second ranking, violation of TEL is fatal, so that the first and third are selected as optimal. As a result, three out of the four candidates are grammatical in English.


**Tableau 5: Relative clauses with preposed relative pronoun**

English	REC	LE(CP)	TEL
the man [who <sub>i</sub> <del>that</del> I know t <sub>i</sub> ] 		*>	
the man [ <del>who</del> <sub>i</sub> that I know t <sub>i</sub> ] 			<*
the man [ <del>who</del> <sub>i</sub> <del>that</del> I know t <sub>i</sub> ] 		*>	
the man [who <sub>i</sub> that I know t <sub>i</sub> ]		*>	<*

Next consider the evaluation of the French examples in (16), in which a PP containing a relative pronoun is preposed. Since the preposition is not locally recoverable, deletion of it leads to a violation of the highest-ranked constraint REC: this excludes the second and the third candidate. Since the two remaining candidates both violate LE(CP), the lowest ranked constraint TEL gets the final


say by excluding the fourth candidate. Note that this shows that the ranking  $LE(CP) \gg TEL$  does not mean that the complementizer is always realized, but that this may depend on other factors; when the complementizer is preceded by some element that must be realized, TEL forces the complementizer to delete.

**Tableau 6: Relative clauses with preposed PP**

French	REC	LE(CP)	TEL
l'homme [avec qui <sub>i</sub> <del>que</del> j'ai dansé t <sub>i</sub> ] 		*	
l'homme [ <del>avec qui<sub>i</sub></del> que j'ai dansé t <sub>i</sub> ]	*!		*
l'homme [ <del>avec qui<sub>i</sub></del> <del>que</del> j'ai dansé t <sub>i</sub> ]	*!	*	
l'homme [avec qui <sub>i</sub> que j'ai dansé t <sub>i</sub> ]		*	*!

For the English examples in (14) we get the same result as in French: both the second and the third candidate are excluded by REC, and the fourth candidate is excluded because it is harmonically bound by the first candidate: it has a fatal violation of TEL irrespective the question whether we read the tie from left to right or from right to left.

**Tableau 7: Relative clauses with preposed PP**

English	REC	LE(CP)	TEL
the man [who <sub>i</sub> <del>that</del> I know t <sub>i</sub> ] 		*	
the man [ <del>who<sub>i</sub></del> that I know t <sub>i</sub> ]	*!		<*
the man [ <del>who<sub>i</sub></del> <del>that</del> I know t <sub>i</sub> ]	*!	*>	
the man [who <sub>i</sub> that I know t <sub>i</sub> ]		*	*!

The discussion above has shown that that OT fully adheres to the global architecture in Figure 2. The focus of attention is, however, on the evaluator. The OT view on the evaluator seems to be of a more optimistic nature than that of the P&P approaches. The latter consider the evaluator as a more or less random collection of language-specific filters on the output of core grammar. Pesetsky's work has shown, however, that at least some of the filters proposed by Chomsky and Lasnik (1977) can be decomposed into more atomic OT

constraints (see Dekkers, 1999, for more examples). Furthermore, since the OT constraints are claimed to be universal, they make precise predictions about the range of language variation that is allowed: Pesetsky, for example, has shown that his proposal is able to account for the differences between English and French relative clause constructions, and Broekhuis and Dekkers (2000) and Dekkers (1999) have shown that his proposal can be readily extended to relative constructions in Dutch.

### 2.3 Conclusion

This section has argued that the global architecture of grammar is as given in Figure 2, and that the several proposals made within the P&P approach do not differ in this respect from OT-syntax. The two frameworks are similar in assuming that we are dealing both with derivations and with evaluations: a generator creates a potentially multi-membered set of expressions *S*, and an evaluator determines which expressions from *S* are grammatical in a given language *L*. Although this section has mainly focused on the similarities in architecture between the P&P approaches and OT-syntax, it must be noted that there are other similarities between the two frameworks. For example, both MP and OT-syntax adopt some version of Frege's principle of compositionality of meaning by claiming that meaningful elements must be interpreted: in MP it is assumed that interpretable semantic features cannot be deleted and must receive an interpretation (Full Interpretation); the fact that Pesetsky's constraint RECOVERABILITY is universally ranked high expresses more or less the same,<sup>2</sup> as does Grimshaw's (1997) claim that all candidates in a certain candidate set have

---

<sup>2</sup> Given that there are no known cases in which RECOVERABILITY is violated, Broekhuis and Dekkers (2000:421) actually argued that it should actually not be considered a constraint but an inviolable condition on the operation DELETE.

the same meaning. I will not digress on this, however, and continue the discussion by focusing on some differences between the two frameworks.

### 3 Where MP and OT do differ: derivations and evaluations

The previous section has argued that MP and OT assume the same global architecture of grammar. However, there are also obvious differences. This subsection will briefly discuss these and argue that they do not have a principled linguistic motivation, but are the result of a more or less accidental difference in focus of attention between the two approaches: MP is mainly concerned with the universal, derivational aspects of grammar, whereas OT-syntax rather focuses on more language-specific aspects of grammar, or, to put it differently, MP is basically a theory of  $C_{HL}$ , the generator from the model in Figure 2, whereas OT is basically a theory of the evaluator.

This difference between MP and OT is also reflected in the research strategies that the two approaches employ, which in a sense are each other's opposite. Research in MP tends to attribute as many properties of languages to the generator  $C_{HL}$ ; although we have seen in the discussion of Icelandic OS (section 2.1) that MP does allow for filtering devices, researchers seem to take recourse to these as a last resource only. Research in OT, on the other hand, tends to attribute as many properties of languages to the evaluator; although it is generally acknowledged that the generator has certain universal properties, these are hardly ever invoked to account for the data.

Given that MP is a theory of the generator and OT-syntax a theory of the evaluator, it is not surprising that the empirical successes of the two approaches lie in different areas. MP is especially well equipped to account for the universal properties of languages, but there is no generally accepted view on the way we should account for, or even approach, the many ways in which languages may

differ from each other. OT, on the other hand, precisely provides such a general theory of language variation, but since there is no generally accepted theory of the generator, current OT-syntax fails to account for the ‘truly’ universal properties of languages. These differences between MP and OT will be discussed more extensively below.

### 3.1 Universal properties of language (the generator)

Both MP and OT-syntax hold the generator responsible for the invariant properties of language: the generator determines what representations are contained in the output, and hence can take part in the evaluation. The two frameworks differ, however, with respect to the extent that the generator is developed, or invoked in the analysis of the linguistic data.

The investigation of the generator ( $C_{HL}$ ) is considered MP’s core business. It has resulted in a sophisticated, restrictive theory on the nature of the generator. It is assumed that  $C_{HL}$  is constituted by a small set of operations that are subject to inviolable conditions that are relatively well understood. Perhaps  $C_{HL}$  can be reduced to a single merge operation, which has two incarnations, external and internal merge. As a result of this, also the output of  $C_{HL}$  is highly restricted; although it can be a non-singleton set, the differences between the members of this set are very limited in nature, and perhaps only involve the number of movements that occurred (cf. the discussion of Icelandic OS in section 2.1). It seems that analyses that do not invoke filtering devices are valued higher in MP than those that do. As a result, research tends to focus on those phenomena that can be successfully approached by means of a derivational account, with a concomitant reduction of the empirical scope of the theory; Chomsky (1995:§4.7.3), for example, suggests that ‘rearrangement’ phenomena like extraposition, right-node raising, VP-adjunction and scrambling are not part of core syntax.

It is generally admitted in OT-syntax that the generator is the locus of the ‘truly’ universal properties of language: for example, Grimshaw (1997) assumes that the structures formed by the generator conform to some version of X-bar-theory, Pesetsky (1998) and Anderson (2000) adopt some version of generative grammar as the generator, and Bresnan (2000) and Sells (2001) argue in favor of some version of Lexical Functional Grammar. The nature of the generator is, however, not a prominent subject of research, which is possibly also related to the fact that the current generation of OT-syntacticians has come from various theoretical frameworks with varying views on the nature of the generator. Furthermore, it is rather exceptional for an OT-researcher to account for some phenomenon by taking recourse to the generator; most research in OT-syntax rather focuses on the variation that can be found than on the universal properties of languages.

Despite the differences in theoretical background (P&P, LFG, etc), it seems that the view on the generator of many (if not most) OT-syntacticians crucially differs from that of the MP-researchers, which becomes especially apparent when we consider the differences in the view on the output of the generator. We have already seen that although MP allows for non-singleton output sets, it is generally taken for granted that this set is very small and that differences between the members of this set are limited in type, perhaps confined to differences in movement. In OT, on the other hand, it is generally maintained that the output of the generator is in principle infinitely large, and that the members of the set may differ in a wide variety of ways. This seems to imply that the generator contains a larger set of operations in OT than is assumed in MP, and that these operations are probably confined in a less strict manner than the operations assumed in MP.

As a result of this different view on the generator, MP and OT tend to provide entirely different explanations for similar phenomena, the former taking



recourse mainly to properties of the generator and the latter to those of the evaluator. This state of affairs seems to strengthen the widely accepted view that we are dealing with two competing and essentially incompatible frameworks. However, it can also be assessed differently, and more positively. Since it is not *a priori* given whether a certain phenomenon belongs to core syntax or to the periphery, it is important to develop alternative analyses that can subsequently be compared and evaluated; the fact that in some domains competing MP- and OT-analyses are available therefore does not mean in itself that we are dealing with competing or conflicting theories.

In fact, there are similar conflicts internally in MP. Take as an example verb second, which has long been considered a prototypical example of a phenomenon that is part of core syntax, and which has played an important role in the development of the theory of functional heads (especially the CP projection) and head movement (verb movement to C). Nevertheless, since  $C_{HL}$  as developed in Chomsky's (1995:§4.10) is no longer able to handle verb second in Icelandic transitive expletive constructions (p.354), Chomsky concluded that it should be considered part of the periphery, as the result of some not further explicated PF-rule (p.368). Taken to its extreme, this proposal may lead to the claim that verb second, like the other 'rearrangement' phenomena mentioned above, is not part of core syntax at all, but essentially a PF-phenomenon; cf. Chomsky (2001:37-8) and especially Boeckx and Stjepanovic (2001), who explicitly argue that head-movement in general is PF phenomenon.

In short, the fact that OT and MP provide competing analyses for the same phenomena does not show that MP and OT should be seen as competing or conflicting theories but should rather be seen as a normal reflex of the fact that it is not *a priori* given whether a certain phenomenon belongs to core syntax or to the periphery. The question which analyses are most feasible is therefore essentially an empirical one.

### 3.2 Variation (the evaluator)

One of the main concerns of both MP and OT is **cross-linguistic variation**. However, the way they approach this problem is entirely different — at least, at first sight. Let us start with discussing the way MP approaches the issue. Language variation is assumed to arise as a result of additional constraints on the application of the otherwise universal generator ( $C_{HL}$ ). The generator can basically perform two operations: external and internal merge. Let us provisionally adopt the standard assumption in MP that external merge is indispensable given that it is needed in order to assemble lexical items into semantically interpretable structures, e.g., by saturating the thematic roles of a given lexical head. Despite the fact that internal merge may have certain semantic implications, it is not essential in the creation of semantically interpretable structures, so that we expect to find language variation in this domain. Note that since MP is mainly concerned with core syntax it also mainly studies differences between languages that are somehow related to movement: variation in other domains is attributed to other modules (like PF), and is generally not discussed any further.

In early MP, the locus of variation between languages is solely attributed to the lexicon. Differences in the displacement property of languages are due to differences in the ‘strength’ property of the morpho-syntactic features that trigger movement: strong features trigger overt movement, whereas the weak features allow covert movement (which is favored by Procrastinate). In the more recent Agree-based theories, which reject the idea of covert movement, the core idea is preserved by assuming that movement only takes place if a functional head  $F$  contains an EPP-feature, which requires that the specifier of  $F$  be present. Under this view, the task of the language learner is to determine whether the functional head  $F$  has a weak or strong feature, or, alternatively, whether it has an EPP-feature, and to store this information in the lexicon.

The scope of OT goes much beyond the displacement property of languages: in principle, all (phonological, syntactic, semantic, pragmatic, etc.) properties can be fruitfully investigated, as long as one can plausibly postulate constraints bearing on the phenomenon in question. As we have already seen variation between languages is attributed to the evaluator in Figure 2, more specifically to the differences in ranking of the otherwise universal constraints. Under this view, the task of the language learner is therefore to determine the constraint ranking (and the lexicon) of the language.

The discussion above seems to reveal another important difference between MP and OT: in the former cross-linguistic variation is solely due to differences in lexical specifications, whereas in the latter it rather due to the ranking of the universal constraints. This is indeed the case when we compare early MP with OT-syntax, but it does no longer hold when we compare the most recent Minimalist Inquiry framework and OT-syntax.

The early MP thesis that the sole locus of cross-linguistic variation is the lexicon runs into severe problems when we consider **variation within a single language**, because it predicts that languages cannot have ‘optional’ movement, by which I refer to movement operations that occur only under well-defined semantic or phonological conditions. One example of this type of movement is Icelandic OS (already discussed in section 2.1), which can only apply when the object is part of the presupposition of the clause (cf. (6)), and when it does not cross the verb (cf. (7)) or other  $v^*P$ -internal material. This kind of optionality cannot arise under the early MP thesis because the postulation of feature strength or an EPP-feature gives rise of to a very rigid system: when a feature is strong/an EPP-feature is present, movement must apply; when a feature is weak/an EPP-feature is not present, movement is blocked by Procrastinate.

This problem has led to proposals according to which in some cases certain features are optionally strong or an EPP-feature is optionally present. In

order to avoid circularity, the choice must be made sensitive to external factors like the semantic and phonological conditions imposed on the pertinent movement, and this is precisely what Chomsky (2001) did in his account of OS in Icelandic in (10), repeated below as (19): as we have seen, the language-specific statement in (19c), in tandem with the universal principles in (19a&b), precisely derives the circumstances under which Icelandic OS applies.

- (19) a.  $v^*$  is assigned an EPP-feature only if that has an effect on outcome.  
 b. The EPP position is assigned INT.  
 c. At the phonological border of  $v^*P$ , XP is assigned INT'.

Chomsky (2001:36) presents clause (19c) as a parameter that distinguishes OS from non-OS languages. French, for example, has verb movement to I, but nevertheless OS does not apply. This can be accounted for by assuming that (19c) does not hold for French. As a result, the interpretation INT can be assigned to the object when it is at the phonological border of  $v^*P$ ; as a result, movement of the object to the EPP-position is not needed and assignment of an EPP-feature to  $v^*$  is consequently blocked by (19a).

It seems, however, that (19c) is unlike the parameters of the earlier P&P framework in that it is not binary, because it is not the case that languages can be straightforwardly divided between OS and non-OS languages. This will become clear when we consider the Danish examples in (20) and (21), taken from Vikner (1994:502). The examples in (20) show that Danish, unlike Icelandic, does not have OS of non-pronominal DPs, whereas the examples in (21) show that it does have OS of weak pronouns.

- (20) a. Hvorfor læste studenterne ikke artiklen?  
           why      read  the students not  the article  
 b. \*Hvorfor læste studentene artiklen<sub>i</sub> ikke  $t_i$ ?

- (21) a. Hvorfor læste studenterne den<sub>i</sub> ikke t<sub>i</sub> ?  
           why      read  the students it     not  
       b. \*Hvorfor læste studenterne ikke den?

This can be accounted for by assuming that clause (19c) must be further refined as in (19c'). This clause correctly expresses (i) that non-pronominal DPs that are part of the presupposition of the clause (= INT) must undergo OS in Icelandic, but not in Danish or the Romance languages, and (ii) that definite pronouns (which are assigned INT by definition) must undergo OS in Icelandic and Danish but not in the Romance languages.<sup>3</sup>

- (19) c'. At the phonological border of  $v^*P$ , XP is assigned INT'  
       (i) XP = DP (Icelandic)  
       (ii) XP = definite pronoun (Danish)  
       (iii) XP =  $\emptyset$  (Romance)

What I want to stress here is that the adoption of language specific statements like (19c) or (19c') is a radical break with the early MP thesis that the sole locus of cross-linguistic variation is the lexicon. Since these statements essentially function as language-specific filters on the output of  $C_{HL}$ , it should be attributed to the evaluator in the model in Figure 2, and not to the lexicon. In fact, it seems that Chomsky's proposal makes it possible to eliminate the EPP-features entirely: when we assume that movement is subject to Last Resort but applies optionally, we could simply replace clause (19a) by the claim that movement is possible only if it has an effect on the outcome. This would make it possible to attribute cross-linguistic language variation *entirely* to the evaluator, just like in OT. In (22) I attempt to rephrase Chomsky's proposal such that reference to the notion of EPP-feature becomes superfluous.

---

<sup>3</sup> For completeness' sake, note that the fact that English does not have OS does not follow from clause (19c'): since English does not have V-to-I movement, objects are never at the phonological border of  $v^*P$  so that (19c') never applies and OS is always blocked by (19a).

- 
- (22) a. Movement is possible only if it has an effect on outcome.  
 b. The derived object position is assigned INT.  
 c. At the phonological border of  $v^*P$ , XP is assigned INT'.  
     (i) XP = DP (Icelandic)  
     (ii) XP = definite pronoun (Danish)  
     (iii) XP =  $\emptyset$  (Romance)

### 3.3 Conclusion

Since we have seen in section 2 that MP and OT assume more or less the same global organization of grammar, we may conclude that the differences in the research strategies of MP and OT are somewhat accidental: as far as I can see, there are no theory-internal reasons for these frameworks to limit their investigation to respectively the generator or the evaluator. The fact that MP and OT occasionally provide alternative analyses for similar data as a result of these differences in research strategy does not follow from insurmountable theoretical differences between the two frameworks either, but simply reflects the fact that it is not *a priori* given whether a certain phenomenon belongs to core syntax or to the periphery.

Early MP and OT-syntax do seem to adopt conflicting views on the nature of variation between languages: the former adopts the thesis that language variation can be reduced to differences in the feature specifications of the lexical elements (feature strength/EPP-features), whereas the latter assumes that language variation is due to the evaluator, that is, to differences in constraint rankings. In Chomsky's current *Minimalist Inquiry* framework, however, the early MP thesis has been dropped: language variation is (also) attributed to parameters like (22c), which essentially function as language-specific filters on the output of  $C_{HL}$ . Current MP and OT therefore both attribute language variation to the evaluator, and the main difference between MP and OT boils

down to the question whether the evaluator takes recourse to output filters or to ranked constraints.

In sum, we may conclude that MP and OT-syntax are actually much more alike than is generally assumed or one would think at first sight. Given the fact that the strengths and weaknesses of the two frameworks are somewhat complementary (MP being especially successful in accounting for the universal, derivational aspects of grammar, and OT-syntax being especially well equipped to account for variation), it is fully justified and useful to investigate whether the strengths of the two frameworks can somehow be combined. This will be the topic of the next section.

#### **4 The derivation-and-evaluation model**

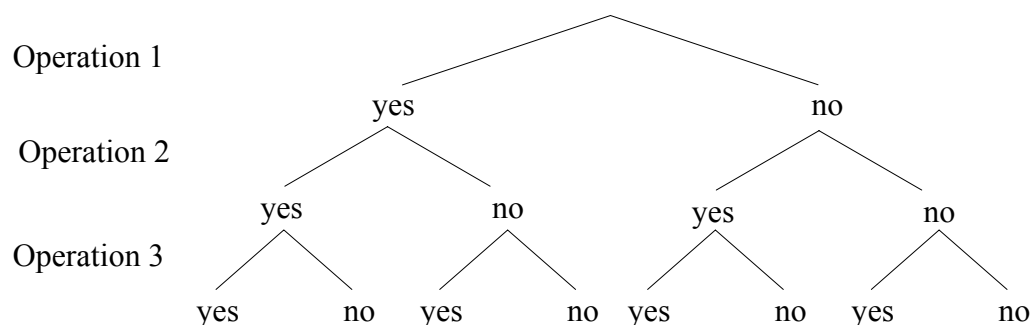
This section sketches the derivation-and-evaluation (D&E) model in Figure 1, in which the strengths of MP and OT are combined. The name of the model underlines the claim that the generator and the evaluator are equally important for providing descriptions and explanations of linguistic phenomena. The D&E model differs from the current versions of OT-syntax in that it adopts a version of  $C_{HL}$  as its generator, and it differs from MP in claiming that the output of  $C_{HL}$  is not evaluated by means of filters but in an optimality-theoretic fashion. Adopting the D&E model makes it necessary to seriously investigate the *interaction* between the generator and the evaluator: after all, when both the generator and the evaluator are to be taken seriously, they are expected to interact in intricate ways so that properties ascribed to the former may have far-reaching consequences on the design of the latter, and vice versa. Section 4.1 and 4.2 will discuss the generator and evaluator, respectively, and compare the D&E assumption with those normally adopted in MP and OT-syntax.

## 4.1 The generator

D&E adopts the standard assumption from MP that the computational system  $C_{HL}$  is universal and consists of operations that are conceptually necessary, such as the two incarnations of the merge operation, and possibly the operation Delete. The latter operation is needed to account for deletion of the phonological features of complementizers and relative pronouns (cf. the discussion of relative clauses above), although it is not *a priori* clear whether Delete should be considered an operation of  $C_{HL}$  or of the phonological component. Furthermore, D&E adopts the claim that these operations are subject to inviolable conditions: movement, for example, must satisfy the Last Resort Condition, according to which movement of a syntactic object S must be triggered by some unchecked or unvalued formal feature of a higher functional head H that can be valued by a corresponding feature of S, and Delete is subject to the recoverability condition (cf. fn. 2).

The main difference between D&E and the ‘standard’ versions of MP is that the former assumes that  $C_{HL}$  is not parameterized: more specifically, it is assumed that there are no strength/EPP-features that may force or block the application of a certain operation, and neither can an operation be blocked by the availability of a more economical option (cf. Broekhuis and Klooster, 2001, who argue that there is no general preference for external over internal Merge). At any point P in the derivation,  $C_{HL}$  may choose at random between applying or not applying the operation(s) that could in principle be performed (= would satisfy Last Resort) at P. Consequently, the number of candidates in the candidate set is therefore at most  $2^n$ , where  $n$  is the number of operations that satisfy Last Resort.



**Figure 8: The construction of the candidate set**

$C_{HL}$  thus defines a candidate set that contains a limited number of candidates, and which is defined by the optional application of the operations Merge and Delete. Of course, the effects of the strength/EPP-features must be mimicked in some way, but we have seen in section 3.2 that the filters introduced in Chomsky (2001) in effect already determine whether certain movements may or may not apply, so that they make the EPP-features superfluous: cf. the discussion above (22). We may therefore conclude that, as far as the generator is concerned, the D&E model in Figure 1 comes very close to the more current versions of MP.

The D&E claim that the generator should be identified with the computational system  $C_{HL}$  from MP breaks radically with the generally adopted OT-claim that the candidate set is infinite; the claim that the operations of the generator, although being subject to a Last Resort Condition, can in principle be optionally applied, results in candidate sets that are very small.<sup>4</sup> By way of illustration, (23a&b) give the maximum size of the candidate sets for derivations with respectively 8 and 16 operations that satisfy Last Resort. Actually, it is

<sup>4</sup> What is maintained, however, is that the candidate set can be assumed to be very similar for all languages: variation may arise but this is mainly the result of differences in the lexicon, such as the availability of certain lexical items, or the (non-)affixal status or the categorial nature of the lexical elements involved in the derivation.

even possible to reduce these numbers much further by adopting some version of phase theory. This is shown in (23a'&b').

- (23) The size of the candidate set:
- a. 8 operations:  $2^8 = 256$
  - a'. 8 operations in 2 phases of 4 operations each:  $2 \times 2^4 = 32$
  - b. 16 operations:  $2^{16} = 65.536$
  - b'. 16 operations in 4 phases of 4 operations each:  $4 \times 2^4 = 64$

I believe that this radical break with the OT-tradition is also advantageous from the OT point of view. First, of course,  $C_{HL}$  can be invoked to provide a non-*ad hoc* account for the truly universal properties of languages, which the OT-evaluator by its very nature is not able to do so. Secondly, since part of the descriptive burden is now placed on the generator, we may hope that this will enable us to considerably reduce *the number of constraints* in the universal constraint set CON. This, in its turn, will result in a dramatic decrease of *the number of constraint rankings*, and, consequently, of the number of possible natural languages. Thirdly, the fact that  $C_{HL}$  does not only limit the candidate set, but also the *type of differences* that can be found among the candidates in this set, which are defined by the application or non-application of the operations of  $C_{HL}$ , suggests that it will be possible to also reduce *the number of constraint types*, and, consequently, also the ways in which natural languages can differ from each other. It goes without saying that all these consequences contribute to considerably enhancing the explanatory adequacy of OT-syntax.<sup>5</sup>

---

<sup>5</sup> The discussion above will make it clear that I disagree with Samek-Lodovici's (this volume) claim that it is an inherent virtue of OT that is more powerful than MP. Given that the grammar should define the notion of possible natural language, reduction of generative power is desirable when it leads to the exclusion of languages that are likely not to be part of the set denoted by this notion.

## 4.2 The evaluator

The previous subsection has briefly mentioned the D&E proposal that the EPP-features should be eliminated by attributing the intended effects of these features to the evaluator. From the point of view of MP, this step seems quite natural since I have already argued in section 3.2 that the filters introduced in Chomsky (2001) actually suffice to determine whether certain movements may or may not apply. A general problem with filters is, however, that they tend to take the form of *ad hoc* stipulations that simply reformulate descriptive generalizations or the description of certain states of affairs in a semi-formal language. Since it is not obvious that this will lead to any deeper insights, the D&E framework adopts the idea that filters should be subject to further investigation, and be derived from more primitive notions of the theory. It further assumes that that this is precisely what OT does: work by Pesetsky (1997;1998) and Dekkers (1999) has already shown that at least some of the filters from Chomsky and Lasnik (1977) may receive a natural explanation in this way, and this section will show that also the language-specific filter in (22c) can be expressed by means of the interaction of a small set of more primitive constraints (cf. Costa 1998 and Broekhuis 2000).

The previous section has also argued that by adopting  $C_{HL}$  as the generator, the OT-evaluator can be considerably simplified: since the inviolable conditions on the operations of the generator carry part of the descriptive burden, we may expect a reduction of the number of constraints that in CON, and since the candidates in the candidate set only differ from each other in a small number of well-defined ways, we may also expect the number of constraint types to be rather small.

In order to get some idea about the syntactic constraints and constraint types that we may expect to arise, I will adopt as my point of departure that the OT-evaluator is a formalization of the so-called interface conditions postulated in MP. If that is indeed so, we expect the syntactic constraints in CON to be

somehow related to three components involved: the computational system  $C_{HL}$ , which creates the relevant syntactic representations in the candidate set, and the two interpretative systems that interpret them: the articulatory-perceptual and the conceptual-intentional component. Let us therefore assume that the syntactic constraints in CON can be divided into the two basic types in (24).

- (24) The syntactic constraints in CON are of two basic types:
- a.  $C_{HL}$  constraints
  - b. Interface (PF and LF) constraints

Before I discuss these constraint types, I want to point out that, in my view, it is not only desirable to restrict the number and kind of constraints, but also to restrict the possible format of the constraints. I will therefore adopt Eisner's (1999) proposal that there are basically two formal types of constraints which should be formulated as positive or negative generic statements (which Eisner refers to as the implication and clash families). Furthermore, I will assume that the formulation of the constraints is simple in the sense that connectives like *and*, *or*, *unless*, etc. cannot be used.

#### 4.2.1 $C_{HL}$ constraints

The D&E framework assumes that the application of the operations of the generation is essentially free. Nevertheless, it is clear that in most languages there are strict restrictions on the application of these operations. A good example of this is OS: languages like Icelandic have it, whereas the Romance languages do not. Given the claim that the generator is universal and cannot be parameterized, it must be the evaluator that penalizes the application of this movement. Therefore, we must postulate a set of clash constraints that favor the non-application of the operations of C, and which I will henceforth refer to as **economy constraints**.

A first example of such an economy constraint is STAY, which I prefer to call \*MOVE in order to highlight the fact that it is a clash constraint. \*MOVE forbids internal merge, and thus militates against superfluous movement steps in the derivation. Assuming this constraint seems uncontroversial: it is assumed in most work in OT-syntax, and it has its MP counterpart in the claim that movement is a costly operation. In the early MP period, this claim has played a crucial role in the formulation of principles like Procrastinate and Fewest Steps, and it has survived in the later period in the form of the proposal that movement is licit only when an EPP-feature is present.

It has been proposed that the economy constraints on movement may take a more specific form. For example, Grimshaw (1997) proposes the constraint No-Lexical-Movement (NOLEXM), which blocks movement of the lexical ( $\theta$ -role assigning) verbs. This constraint is a reformulation of Pollock's (1989) ban of movement of lexical verbs to weak AGR-phrases: English has a weak AGR, and therefore movement of a lexical (but not an auxiliary or a modal) verb is blocked in (26a); French has a strong AGR, and consequently movement of a lexical (as well as an auxiliary or a modal) verb is possible in (26b).

(25) NOLEXM: don't move lexical ( $\theta$ -role assigning) verbs.

- (26) a. John <\*kisses> often <kisses> Mary.  
 b. Jean <embrasse> souvent <\*embrasse> Marie.

Given that the economy constraints block the application of the operations of the generator, we must also introduce means that allow or force the operations of  $C_{HL}$  to apply. Since we have seen that languages differ in their displacement properties, we cannot take recourse to some general property of the conceptual-intentional or the articulatory-perceptual component to force movement. Therefore, we have to postulate constraints that favor movement, so that the relative ranking of these constraints and the economy constraint \*MOVE will

determine whether a certain movement does or does not take place. Of course, we want to restrict the class of constraints that force movement as much as possible. In order to obtain this let us assume that all probes prefer movement of their goal into their local domain (I will use the notion of local domain instead of the notion of checking domain in order to avoid the connotation that movement into the local domain of a head H is required to value the unvalued features). In a sense, this means that we are generalizing the EPP to all unvalued features. The general form of the **EPP constraints** is given in (27), and they force movement of the goal into the local domain of the probe. Consequently, if the goal of probe F is an XP, this constraint forces it to move into a specifier of the head that has F as its sublabel, and if the goal is a head it is adjoined to the head that has F as its sublabel. Potential specific instantiations of the ‘generalized’ EPP constraint are given in (27i-iii). The constraints EPP(case) and EPP( $\phi$ ) require movement of a DP into the specifier of a head containing case or  $\phi$ -features, and EPP(tense) requires head-movement of the finite verb to T.

- (27) EPP(F): probe F attracts its goal.
- (i) EPP(case): an unvalued case-feature attracts its goal.
  - (ii) EPP( $\phi$ ): unvalued  $\phi$ -features attract their goal.
  - (iii) EPP(tense): an unvalued tense feature attracts its goal.
  - (iv) etc.

It is obvious that the number of EPP constraints cannot be larger than the number of unvalued constraints that are postulated in the grammar. It is, however, less clear whether the two numbers are equal. Take *wh*-movement. Watanabe (1991) has argued on empirical grounds that so-called *wh*-in situ languages like Japanese actually have overt *wh*-movement of an empty operator: among other things, this accounts for the fact that also these languages exhibit *wh*-island effects. Chomsky (1995:ch.3) claimed on the basis of Watanabe’s findings that *wh*-features are universally strong. When we abandon covert movement in favor

of Agree, Watanebe's findings suggest that Agree does not suffice to license *wh*-constructions, but that movement must apply. So the question is: Why? Earlier proposals have maintained that *wh*-phrases can be interpreted by the conceptual-intentional component only if it heads an operator-variable chain; cf. e.g. Chomsky (1991:440) and Rizzi (1996). If so, the obligatoriness of *wh*-movement follows immediately from semantic considerations, since any construction in which *wh*-movement does not apply will either crash as a violation of Full Interpretation or, at least, receive an anomalous interpretation. Consequently, the postulation of a constraint like EPP(*wh*) has no effect, so that we may safely assume that it does not exist. If movement of the goals of other [+affect] features like [topic], [focus] or [neg] are similarly forced by semantic considerations, we may also conclude for them that they do not fall under the generalized EPP constraint in (27). This would eliminate a large set of potential constraints from the grammar, and thus considerably reduce the set of possible grammars. Since this issue does not play a prominent role in the present study, I will not pursue this issue any further, and leave it to future research.

Word order variation between languages is accounted for by assuming that the EPP constraints interact in an optimality-theoretic fashion with the economy constraints. Ranking (28a) expresses that probe F (normally) does not trigger movement due to the fact that the EPP constraint is outranked by the economy constraint \*MOVE: this ranking will be called weak, since it is more or less equivalent to assuming that probe F is weak or has no EPP-feature associated with it. Ranking (28b), on the other hand, expresses that probe F (normally) does trigger movement due to the fact that that the EPP constraint outranks the economy constraint \*MOVE: this ranking will be called 'strong', since it is more or less equivalent to assuming that probe F is strong or has an EPP-feature associated with it.

- (28) a. weak ranking: \*MOVE >> EPP(F)  
 b. strong ranking: EPP(F) >> \*MOVE

The choice between the weak and the strong ranking of a certain probe F constitutes one of the ways in which languages can be parameterized. In fact, (28) constitutes a clear example of what one may call a macro-parameter. For example, if we continue to assume that OS is triggered by the case features on  $v^*$ , we may distinguish between languages like Icelandic, which has full OS, and languages that have only partial OS or no OS at all, by the two rankings in (29). Of course, we have seen that OS is much more complicated than this, and this is where the interface (PF and LF) constraints come in.

- (29) a. \*MOVE >> EPP(case): object shift is (normally) blocked.  
 b. EPP(case) >> \*MOVE: object shift (normally) applies.

#### 4.2.2 *The Interface (PF and LF) constraints*

One of the disadvantages of early MP was that the postulation of feature strength or the association of an EPP-feature with certain formal features gave rise to a very rigid system: if a certain formal feature is assumed to be strong or to be associated with an EPP-feature, it is predicted that it invariably triggers movement; if a certain formal feature is assumed to be weak or not to be associated with an EPP-feature, Procrastinate predicts that the pertinent movement is invariably blocked. As we have seen above, Chomsky (2001) has tried to make the system more flexible by making the selection of the EPP-features dependent on semantic and phonological factors. The three statements in (19), repeated here as (25), ultimately have the effect that  $v^*$  is only assigned an EPP-feature (i) when the object is assigned the interpretation INT (= when the object is part of the presupposition of the clause), and (ii) when the object is at the phonological border of  $v^*P$ , that is, when OS does not cross  $v^*P$ -internal material.



- (30) a.  $v^*$  is assigned an EPP-feature only if that has an effect on outcome.  
 b. The EPP position is assigned INT.  
 c. At the phonological border of  $v^*P$ , XP is assigned INT'.

The statement in (30c) is assumed to be a parameter: OS languages have it, whereas non-OS languages do not. Further, we have seen that the introduction of (30c) makes the postulation of EPP-features superfluous, and that we can simply replace (30a) by the assumption that movement is optional in principle, as in (31a). Finally, we have seen that the parameter in (30c) does not suffice, since some languages like Danish have limited OS with definite pronouns. This means that (30c) must be further refined as in (31c).

- (31) a. Movement is possible only if it has an effect on outcome.  
 b. The derived object position is assigned INT.  
 c. At the phonological border of  $v^*P$ , XP is assigned INT'.  
     (i)  $XP = DP$  (Icelandic)  
     (ii)  $XP = \text{definite pronoun}$  (Danish)  
     (iii)  $XP = \emptyset$  (Romance)

Macro-parameters in the format of (28) introduce the same kind of flexibility as filters like (31c). Although movement is normally blocked under the weak ranking in (28a), movement can be forced provided that there is some higher ranked constraint A that favors this movement (cf. (32a)); in the terminology of Chomsky (1995:ch.3), one might say that constraint A overrules 'Procrastinate'. Similarly, although movement is normally forced under the strong ranking in (28b), it can be blocked if there is some higher ranked constraint B that disfavors it (cf. (32b)); in other words, constraint B overrules 'Strength'.

- (32) a.  $A \gg *MOVE \gg EPP(F)$ :  
         if A favors movement, 'Procrastinate' is overruled.  
 b.  $B \gg EPP(F) \gg *MOVE$ :  
         if B disfavor movement, 'Strength' is overruled.

The claim that I want to make here is that it is the function of the interface constraints to overrule macro-parameters of the type in (28). I will illustrate this more specifically for the macro-parameter in (29).

We have seen that Danish has OS of a more limited type: although lexical DPs do not shift, definite pronouns do; cf. (20) and (21). This can be accounted for by assuming that Danish has the weak ranking in (29a), but that this weak ranking is overruled by a constraint that requires definite pronouns to be *vP*-external. The claim that there is a restriction of this sort on the placement of pronouns is not new: Diesing (1997:380), for example, claims that definite pronouns are variables that due to their definiteness cannot remain within the nuclear scope of the clause (VP), and Vogel (to appear) a.o. has argued that weak pronouns must leave the VP for phonological reasons. Let us assume that something of the sort is indeed the case, and postulate the clash constraint D-PRONOUN in (33a), which requires that definite pronouns be *vP*-external. The fact that Danish has OS with definite pronouns only can now be accounted by assuming the ranking in (33b), as is shown by the evaluations of the examples in (20) and (21) in Tableaux 8 and 9.

- (33) a. D-pronoun:  $*[_{vP} \dots \text{pron}_{[+def]} \dots]$ .  
 b. Danish: D-PRONOUN >> \*MOVE >> EPP(case)

**Tableau 8: Danish (no object shift of lexical DPs)**

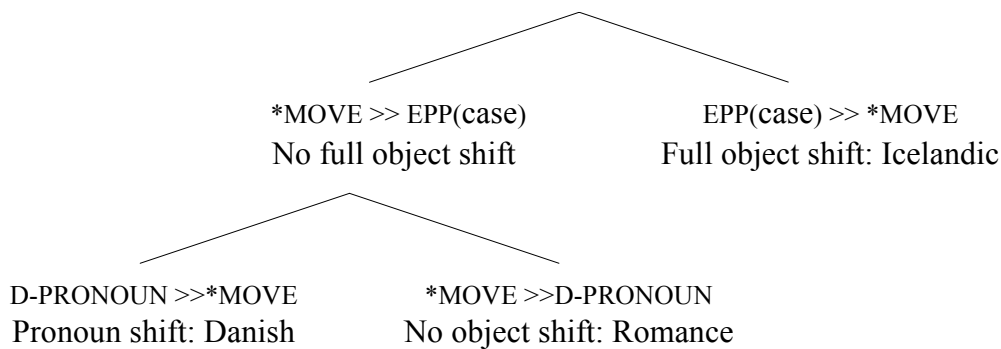
	D-PRONOUN	*MOVE	EPP(case)
Hvorfor læste studenterne ikke artiklen? ➡			*
Hvorfor læste studentene artiklen <sub>i</sub> ikke <i>t<sub>i</sub></i>		*!	

**Tableau 9: Danish (pronoun shift)**

	D-PRONOUN	*MOVE	EPP(case)
Hvorfor læste studenterne ikke den	*!		*
Hvorfor læste studenterne den <sub>i</sub> ikke <i>t<sub>i</sub></i> ➡		*	

The ranking  $D\text{-PRONOUN} \gg *MOVE$  can again be seen as a macro-parameter which subdivides the languages that do not have full OS into languages that do and languages that do not allow pronoun shift. This shows that the constraints we have introduced so far can successfully account for the division postulated by the clauses in (31ci-iii). Observe that the ranking of  $D\text{-PRONOUN}$  and  $*MOVE$  is immaterial for the full OS languages, since movement of the pronoun is already forced by the strong ranking of  $EPP(case)$ .


**Figure 9: Macro-parameterization of languages with respect to object shift**




Also the semantic conditions on the application of objects shift in Icelandic can be taken care of by means of an interface constraint. As we have seen in (6) above OS is normally obligatory in Icelandic, but blocked when the object is part of the focus (new information) of the clause. When we adopt the constraint  $ALIGNFOCUS$  in (34a) from Costa (1998) and rank it above  $EPP(case)$ , we will derive the desired result. The ranking in (34b) correctly predicts that all object DPs must undergo OS, unless they are part of the focus of the clause: OS of a non-presuppositional object across some phonetically realized constituent is excluded. The evaluation of the two examples in (6) is given in the tableaux below.

- (34) a. ALIGNFOCUS: The prosodically unmarked focus is the rightmost constituent in its clause.<sup>6</sup>  
 b. Icelandic: ALIGNFOCUS >> EPP(case) >> \*MOVE

**Tableau 10: Icelandic (object in not focus)**

	AF	EPP(case)	*MOVE
Jón keypti ekki bókina		*!	
Jón keypti bókina <sub>i</sub> ekki t <sub>i</sub> 			*

**Tableau 11: Icelandic (object in focus)**

	AF	EPP(case)	*MOVE
Jón keypti ekki bókina 		*	
Jón keypti bókina <sub>i</sub> ekki t <sub>i</sub>	*!		*

The discussion above has shown that recourse can be taken to the interface constraint D-PRONOUN in (33a) to account for the fact that some languages that do not have full OS do have pronoun shift. By taking recourse to the interface constraint ALIGNFOCUS in (34a), on the other hand, we are able to account for the fact that OS is sometimes blocked in languages that normally do have it. By introducing these constraints, we account for almost the same range of data as (31); the only thing that we have not capture yet is that OS cannot cross  $v^*P$ -internal material; cf. the Icelandic example in (7). In order to account for this, we may take recourse to a number of PF constraints involving linearization. Since these constraints effectively require that the underlying order of heads and arguments be maintained in the surface realization, I will refer to these as ‘shape conservation constraints’. Two examples are given in (35).

<sup>6</sup> Note that the notion of *prosodically unmarked focus* in (34a) refers to the new information of the clause and stands in opposition to the notion of *presupposition*, and should not be confused with the notion of exhaustive or contrastive focus.

- (35) **Shape conservation (PF) constraints** (do not change the base order):
- a. Relativized Minimality (RELMIN): X-movement retains the relative order of elements in X-positions, where X = A, A' or H.
  - b. HEAD-COMPL: a head precedes all terminals dominated by its complement.


The constraint RELMIN in (35a) is of course a direct descendant of the most influential ‘shape conservation’ principle from the earlier P&P period is Rizzi’s (1990) Relativized Minimality, but reinterprets it as a constraint on the output of the generator. Although this will not be illustrated here, this constraint plays a role in prohibiting OS of a direct object across an indirect object (earlier proposals that assume similar constraints/principles are e.g. Williams 2002 and Müller 2000/2001). When we adopt Kayne’s (1994) conjecture that all languages have the underlying the head-complement order, also the constraint HEAD-COMPL in (35b) can be construed as a shape conservation constraint.<sup>7</sup> HEAD-COMPL disfavors OS across the main verb because this would result in a surface order that differs from the underlying order. Consequently, by assuming that HEAD-COMPL outranks EPP(case) in Icelandic, OS will be blocked in examples like (7); the evaluation of these examples is given in Tableau 12. Note that the relative ranking of HEAD-COMPL and AF cannot be determined on the basis of the present set of data, since OS in (7) will be blocked irrespective the question whether the object belongs to the focus of the clause, that is, irrespective the question whether the star between parentheses is present or not.

---

<sup>7</sup> HEAD-COMPL must not be confused with the alignment constraint HEAD-LEFT that can be found in much recent OT-work (e.g. Grimshaw 1997), which also requires a head to precedes its complement, but competes with its counterpart HEAD-RIGHT, which does not feature in my proposal. Alignment constraints play a prominent role in OT-syntax, and have generated a lot of new insights. They have been employed e.g. by Legendre (2000) to account for the linearization of the clitics in the Bulgarian clitic cluster, by Anderson (2000) to account for verb second and other second position phenomena, and by Sells (2001) for describing Swedish object shift. These alignment constraints differ from HEAD-COMPL in that they do not take recourse to an underlying word order but express certain word order generalizations directly.

(36) Icelandic: HEAD-COMPL  $\langle \rangle$  ALIGNFOCUS  $\gg$  EPP(case)  $\gg$  \*MOVE

**Tableau 12: Icelandic (no object shift in complex tense constructions)**

	HEAD-COMPL	AF	EPP(case)	*MOVE
Jón hefur ekki keypt bókina 			*	
Jón keypti bókina <sub>i</sub> ekki t <sub>i</sub>	*!	(*)		*

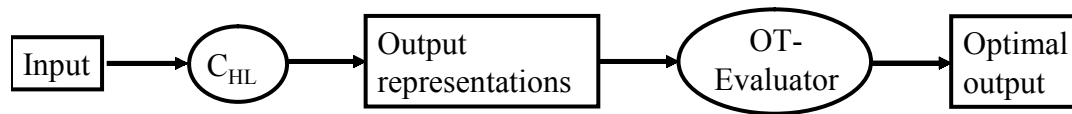
By introducing the Interface constraints D-PRONOUN, ALIGNFOCUS and HEAD-COMPL the present proposal accounts for the same range of facts as the set of statements in (31). There are, however, a number of reasons to prefer the present constraint approach to an approach that takes recourse to filters. First, filters tend to take the form of *ad hoc* stipulations that simply reformulate descriptive generalizations or the description of certain states of affairs in a semi-formal language, and it is not obvious that this will lead to any deeper insights. The constraint approach, on the other hand, derives these generalizations from more primitive notions of the theory. Secondly, the constraint approach (but not the filter approach) makes very precise predictions about what types of languages are possible. The postulation of HEAD-COMPL, for example, predicts that there are also languages in which EPP(case) outranks HEAD-COMPL, and which therefore allow OS across the verb (Dutch and German are of this type). Thirdly, the constraint approach (but not the filter approach) provides us with a general format for approaching other word order phenomena.

## 5 Summary

This paper has provided an updated version of the derivation-and-evaluation (D&E) framework originally proposed in Broekhuis and Dekkers (2000) and Broekhuis (2000). The leading idea of the framework is that, in order to arrive at a descriptively and explanatory adequate theory, restrictions must be placed both on the syntactic derivation and the resulting syntactic representations. This has

been given shape by assuming a framework in which aspects of the minimalist program (MP) and optimality theory (OT) are combined. More specifically, it was claimed that representations created by some version of the computational system of human language  $C_{HL}$  from MP are evaluated in an optimality theoretic fashion, as indicated in Figure 1, repeated below.

**Figure 1: The derivation-and-evaluation (D&E) model**



In MP and OT-syntax the explanatory burden is normally placed on the generator and the evaluator, respectively. By placing the explanatory burden on both systems, these systems cannot be developed independently in the D&E framework: properties ascribed to the one may have far-reaching effects on the format of the other. The following table summarizes the central claims of D&E, and compares these to those normally adopted in MP and OT-syntax.

Derivation-and-evaluation model		MP	OT
I. The generator is some version of $C_{HL}$		+	—
(a)	all operations are subject to inviolable conditions	+	?
(b)	all operations are subject Last Resort	+	—
(c)	the generator is autonomous and operations apply at random; there are no EPP-features.	—	+
II. The evaluator consists of a ranked set of syntactic constraints		—	+
(a)	the syntactic constraints are taken from a universal set CON	d.n.a.	+
(b)	the number of syntactic constraints in CON is small	d.n.a.	—
(c)	the number of syntactic constraint types in CON is small	d.n.a.	—
III. The input and output			
(a)	the lexical items from the input are selected directly from the lexicon without the intervention of a numeration	—	?
(b)	all candidate in the candidate set share the same meaning	+	+

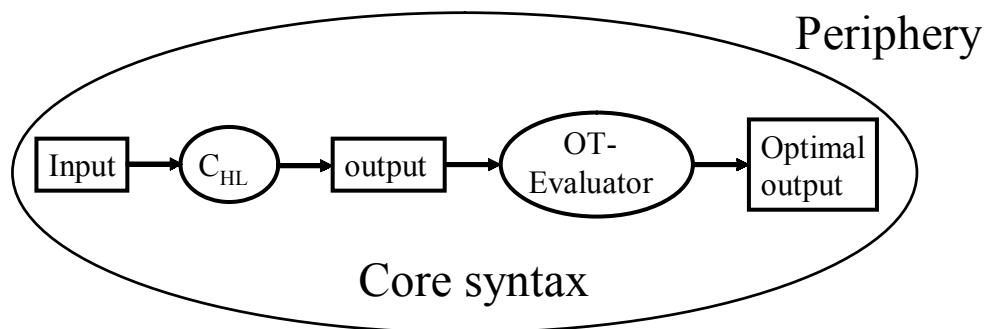
It seems to me that MP and D&E have a descriptive apparatus of more or less the same size, and are facing a similar task in that they both have to identify the features that may trigger movement. The frameworks mainly differ in that they provide different answers to the question what determines whether the movements that are allowed by the Last Resort Condition actually do take place in a given language *L*. In MP it is commonly assumed that movement is forced by the presence of an EPP-feature, and since certain movements, like Icelandic OS, only apply under certain well-defined conditions, the question is raised what determines the distribution of the EPP-features. Chomsky (2001) claims that the distribution of these EPP-features is determined by certain ‘parameters’ that take the form of language-specific output filters. In D&E the answer takes the form of an optimality-theoretic evaluation, as indicated in (28) and (32).

D&E differs from OT-syntax in that the former postulates the computation system  $C_{HL}$  from MP as the generator. As a result of this, many imaginable derivations are blocked by the inviolable conditions on the operations of  $C_{HL}$ , so that the number of candidates in the candidate set is very restricted, and the candidates in this set can differ in well-defined manners only. This has led to the conjecture that there are not only a limited number of syntactic constraints, but also a limited number of constraint types. In order to establish these types, I have assumed that the evaluator is actually a hypothesis about the interface condition postulated in MP, and, consequently that the constraints fall into the two main classes in (24). The  $C_{HL}$  constraints can be further subdivided into two families of constraints, viz. the economy constraints that disfavor the operation of  $C_{HL}$  to apply, and the EPP constraints that favor them: the ranking of these constraints determine whether a certain operation normally does or does not take place. The interface (PF and LF) constraints seem to be more varied in nature, and it is still an open (empirical) question how many there actually are.



In a sense, the D&E framework directly descends from the Chomsky and Lasnik's (1977) *Filters and Control* in postulating two independent systems for generating and evaluating syntactic structures. Chomsky and Lasnik left open the option that the periphery (the evaluative component) uses "much richer resources, perhaps resources as rich as contemplated in the earlier theories of TG", but our hope should be that this will turn out not to be the case, and that also the rules of the periphery will be largely determined by our genetic endowment, that is, by the innate and thus universal constraint set CON. From the D&E perspective, Chomsky and Lasnik's use of the notions 'core' and 'periphery' for respectively the generator and the evaluator is therefore misleading: the generator and the evaluator constitute core syntax together, and the periphery rather lies outside these systems, and should refer to everything that must be learned on an item-to-item or construction-to-construction basis.

**Figure 10: Core and periphery in syntax**



Actually, at some places, Chomsky and Lasnik seem to have had something like this in mind as well, given that they "think of theory of grammar T as consisting of two parts: a universal grammar UG that determines the class of potential grammars and the way they operate, and a system of evaluation that ranks potential grammars in terms of 'optionality' or 'simplicity'" (Chomsky and Lasnik 1977:44). This seems a very apt description of the D&E framework.

## References

- ANDERSON, STEPHEN R.: 2000, Towards an optimal account of second-position phenomena, in Joost Dekkers, Frank Van der Leeuw and Jeroen Van de Weijer (eds.), *Optimality Theory: phonology, syntax and acquisition*, Oxford University Press, Oxford.
- BOBALJIK, JONATHAN: 2002, A-chains at the PF-interface: copies and 'covert' movement, *Natural Language & Linguistic Theory* **20**, 197-267.
- BOECKX, CEDRIC and SANDRA STJEPANOVIC: 2001, Head-ing toward PF, *Linguistic Inquiry* **32**, 345-55.
- BRESNAN, JOAN: 2000, Optimal Syntax, in Joost Dekkers, Frank Van der Leeuw and Jeroen Van de Weijer (eds.), *Optimality Theory: phonology, syntax and acquisition*, Oxford University Press, Oxford.
- BROEKHUIS, HANS: 2000, Against feature strength: the case of Scandinavian object shift, *Natural Language and Linguistic Theory* **18**, 673-721.
- BROEKHUIS, HANS and JOOST DEKKERS: 2000, The minimalist program and optimality theory: derivations and evaluations, in Joost Dekkers, Frank van der Leeuw and Jeroen van de Weijer (eds.), *Optimality Theory: phonology, syntax and acquisition*, Oxford University Press, Oxford/New York, 386-422.
- BROEKHUIS, HANS and WIM KLOOSTER: 2001, On Merge and Move/Attract, in Marc van Oostendorp and Elena Anagnostopoulou (eds.), *Progress in Grammar. Articles at the 20th Anniversary of the Comparison of Grammatical Models Group in Tilburg*, Electronic publication, <http://www.meertens.knaw.nl/books/progressinggrammar>.
- CHOMSKY, NOAM: 1981, *Lectures on government and binding*, Foris, Dordrecht.
- CHOMSKY, NOAM: 1986, *Barriers*, MIT Press, Cambridge, Mass.
- CHOMSKY, NOAM: 1991, Some notes on economy of derivation and representation, in Robert Freidin (ed.), *Principles and parameters in comparative syntax*, MIT Press, Cambridge (Mass.).
- CHOMSKY, NOAM: 1995, *The minimalist program*, MIT Press, Cambridge (Mass.).
- CHOMSKY, NOAM: 2000, Minimalist inquiries: the framework, in Roger Martin, David Michaels and Juan Uriagereka (eds.), *Step by step. Essays on minimalist syntax in honor of Howard Lasnik*, MIT Press, Cambridge (Mass.), 89-155.
- CHOMSKY, NOAM: 2001, Beyond explanatory adequacy, *MIT Occasional Papers in Linguistics* **20**, Cambridge (Mass.).

- CHOMSKY, NOAM: 2001, Derivation by phase, in Michael Kenstowicz (ed.), *Ken Hale. A life in Language*, MIT Press, Cambridge (Mass.), 1-52.
- CHOMSKY, NOAM and HOWARD LASNIK: 1977, Filters and control, *Linguistic inquiry* **8**, 425-504.
- COLLINS, CHRIS: 1997, *Local economy*, MIT Press, Cambridge (Mass.)/London.
- COSTA, JOÃO: 1998, *Word order variation. A constraint-based approach*, LOT Dissertations 14. University of Leiden: Ph.D.-thesis.
- DEKKERS, JOOST: 1999, *Derivations & evaluations. On the syntax of subjects and complementizers*, HIL dissertations 21: Phd.-thesis.
- DIESING, MOLLY: 1997, Yiddish VP order and the typology of object movement in Germanic, *Natural Language and Linguistic Theory* **17**, 369-427.
- EISNER, JASON. (1999). *Doing OT in a straitjacket*. Handout of a talk given at UCLA on 14 June 1999. Available at <http://www.cs.jhu.edu/~jason/resume.html#pubs>.
- GRIMSHAW, JANE: 1997, Projection, heads and optimality, *Linguistic Inquiry* **28**, 373-422.
- GROAT, ERICH and JOHN O'NEIL: 1996, Spell-Out at the LF interface, in Werner Abraham, Samuel David Epstein and Höskuldur Thráinsson (eds.), *Minimal Ideas. Syntactic studies in the minimalist framework*, John Benjamins, Amsterdam/Philadelphia, 113-39.
- HOLMBERG, ANDERS: 1999, Remarks on Holmberg's generalization, *Studia Linguistica* **53**, 1-39.
- KAYNE, RICHARD S.: 1994, *The antisymmetry of syntax*, MIT Press, Cambridge (Mass.).
- LACHARITÉ, DARLENE and CAROLE PARADIS: 2000, Derivational residue: hidden rules in Optimality Theory, in Joost Dekkers, Frank van der Leeuw and Jeroen van de Weijer (eds.), *Optimality Theory: phonology, syntax and acquisition*, Oxford University Press, Oxford/New York, 386-422.
- LEGENDRE, GÉRALDINE: 2000, Morphological and prosodic alignment of Bulgarian clitics, in Joost Dekkers, Frank Van der Leeuw and Jeroen Van de Weijer (eds.), *Optimality Teor, Phonology, Syntax and acquisition*, Oxford University Press, oxford, 423-62.
- MCCARTHY, JOHN J. and ALLEN PRINCE: 1993, Prosodic Morphology: Constraint Interaction and Satisfaction, University of Massachusetts, Amherst/Rutgers University, <http://roa.rutgers.edu/files/482-1201/482-1201-MCCARTHY-0-1.PDF>.
- MÜLLER, GEREON: 2000, Shape conservation and remnant movement, in A. Hirotani, N. Hall Coetzee and J.-Y. Kim (eds.), *Proceedings of NELS 30*, GLSA, Amherst (Mass.), 525-39.
- MÜLLER, GEREON: 2001, Order preservation, parallel movement, and the emergence of the unmarked, in Géraldine Legendre, Jane Grimshaw and

- Sten Vikner (eds.), *Optimality-theoretic syntax*, MIT Press/MITWPL, Cambridge (Mass.)/London, 113-42.
- PESETSKY, DAVID: 1997, Optimality theory and syntax: movement and pronunciation, in Diana Archangeli and Terence Langendoen (eds.), *Optimality theory*, Blackwell, Malden/Oxford.
- PESETSKY, DAVID: 1998, Some optimality principles of sentence pronunciation, in Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis and David Pesetsky (eds.), *Is the best good enough?*, MIT Press/MITWPL, Cambridge (Mass.)/London, 337-83.
- POLLOCK, JEAN-YVES: 1989, Verb movement, Universal Grammar and the structure of IP, *Linguistic Inquiry* **20**, 365-424.
- RIZZI, LUIGI: 1990, *Relativized Minimality*, vol. 16, MIT Press, Cambridge (Mass.)/London.
- RIZZI, LUIGI: 1996, Residual Verb Second and the Wh Criterion, in Adriana Belletti and Luigi Rizzi (eds.), *Parameters and functional heads. Essays in comparative syntax*, Oxford University Press, Oxford/New York.
- SAMEK-LODOVICI, VIERI: this volume, Optimality Theory and the Minimalist Program.
- SELLS, PETER: 2001, *Structure Alignment and optimality in Swedish*, CSLI Publications, Stanford.
- VIKNER, STEN: 1994, Scandinavian Object Shift and West Germanic Scrambling, in Norbert Corver and Henk van Riemsdijk (ed.), *Studies on Scrambling. Movement and Non-movement Approaches to free Word-Order Phenomena*, Mouton de Gruyter, Berlin/New York, 487-517.
- VOGEL, RALF: to appear, Weak Function Word Shift, *Linguistics*.
- WATANABE, AKIRA: 1991, *Wh-in situ, Subjacency and Chain-formation*. MIT: PhD dissertation.
- WILLIAMS, EDWIN: 2002, *Representation theory*, MIT Press, Cambridge, Mass.

*Hans Broekhuis*  
*Leiden University Center for Linguistics (LUCL)*  
*Leiden University*  
*P.O. Box 9515*  
*2300 RA Leiden*  
*The Netherlands*

*hans.broekhuis@uvt.nl*

*[http://fdlwww.uvt.nl/~broekhui/home/home\\_page.html](http://fdlwww.uvt.nl/~broekhui/home/home_page.html)*

# An Optimality-Theoretic Analysis of Scandinavian Object Shift and Remnant VP-Topicalisation

*Eva Engels and Sten Vikner*

University of Aarhus, Denmark

Holmberg (1997, 1999) assumes that Holmberg's generalisation (HG) is derivational, prohibiting Object Shift (OS) across an intervening non-adverbial element at any point in the derivation. Counterexamples to this hypothesis are given in Fox & Pesetsky (2005) which show that remnant VP-topicalisations are possible in Scandinavian as long as the VP-internal order relations are maintained. Extending the empirical basis concerning remnant VP-topicalisations, we argue that HG and the restrictions on object stranding result from the same, more general condition on order preservation. Considering this condition to be violable and to interact with various constraints on movement in an Optimality-theoretic fashion, we suggest an account for various asymmetries in the interaction between remnant VP-topicalisations and both OS and other movement operations (especially subject raising) as to their order preserving characteristics and stranding abilities.

*Keywords: Object Shift, VP-topicalisation, Order preservation.*

## 1 Introduction

In the Scandinavian languages, a pronominal object may move from its base position to a position to the left of a sentential adverbial. This movement operation is called **Object Shift** (OS).<sup>1</sup>

---

<sup>1</sup> Icelandic differs from the Mainland Scandinavian languages in that not only pronominal objects but also full DPs may undergo OS (Vikner 2005: 394).

- 
- (1) Da a. \*Hvorfor læste Peter aldrig \_\_\_\_ den?  
           why read Peter never it  
       b. Hvorfor læste Peter den aldrig \_\_\_\_ \_\_\_\_?

A defining characteristic of OS is that it depends on verb movement. OS is only possible if the main verb moves itself. In other words, the pronominal object cannot undergo OS if the main verb remains within VP, as e.g. in clauses with a non-finite main verb, (2), or in embedded clauses in the Mainland Scandinavian languages (MSc, i.e. Danish, Norwegian, and Swedish), cf. (3).<sup>2</sup>

- (2) Da a. Hvorfor havde Peter aldrig læst den?  
           why had Peter never read it  
       b. \*Hvorfor havde Peter den aldrig læst \_\_\_\_?  
 (3) Da a. Jeg spurgte hvorfor Peter aldrig læste den.  
           I asked why Peter never read it  
       b. \*Jeg spurgte hvorfor Peter den aldrig læste \_\_\_\_.

- 
- (i) Ic a. Af hverju las Pétur aldrei \_\_\_\_ þessa bók?  
           why read Pétur never this book  
       b. Af hverju las Pétur þessa bók aldrei \_\_\_\_ \_\_\_\_?  
 (ii) Da a. Hvorfor læste Peter aldrig \_\_\_\_ bogen?  
           why read Peter never book-the  
       b. \*Hvorfor læste Peter bogen aldrei \_\_\_\_ \_\_\_\_?

<sup>2</sup> Icelandic differs from MSc in that finite verb movement ( $V^\circ$ -to- $I^\circ$ -to- $C^\circ$  movement) and, consequently, OS is restricted to main clauses in MSc, (1) vs. (3), while finite verb movement ( $V^\circ$ -to- $I^\circ$  movement) and OS also take place in embedded clauses in Icelandic, (ii); cf. (Vikner 2005: 394/6).

- (i) Ic a. \*Af hverju las Pétur aldrei \_\_\_\_ hana?  
           why read Pétur never it  
       b. Af hverju las Pétur hana aldrei \_\_\_\_ \_\_\_\_?  
 (ii) Ic a. \*Ég spurði af hverju Pétur læsi aldrei \_\_\_\_ hana.  
           I asked why Pétur read never it  
       b. Ég spurði af hverju Pétur læsi hana aldrei \_\_\_\_ \_\_\_\_.

The observation that the object only moves if the main verb has moved forms the basis of what is called Holmberg's generalisation (Holmberg 1986: 165, 1999: 15).

- (4) **Holmberg's Generalisation (HG)** (Holmberg 1999: 15)  
Object Shift cannot *apply* across a phonologically visible category asymmetrically c-commanding the object position except adjuncts.

HG does not specifically refer to main verbs but to any intervening non-adverbial element. As shown in (5), a verbal particle precedes an object in Swedish, and OS cannot take place across the particle, (6). However, OS is possible if the verbal particle has moved itself, cf. (7).<sup>3</sup>

- (5) Sw a. Jag har inte skrivit upp det.  
I have not written up it  
b. \*Jag har inte skrivit det upp.
- (6) Sw a. Jag skrev inte upp det.  
I wrote not up it  
b. \*Jag skrev det inte upp \_\_\_\_.
- (7) Sw a. UT kastade dom mej inte \_\_\_\_ (bara ned för trappan).  
out threw they me not (only down the stairs)  
b. (Ja, ja, jag ska mata din katt, men)  
(All right, I will feed your cat but)  
IN släpper jag den inte \_\_\_\_.  
in let I it not (Holmberg 1997: 209)

<sup>3</sup> In Danish, Norwegian, and Icelandic, a pronominal object precedes a verbal particle, (i), and, consequently, OS may take place, (ii).

- (i) Da a. \*Jeg har ikke skrevet op det.  
I have not written up it  
b. Jeg har ikke skrevet det op.
- (ii) Da a. \*Jeg skrev ikke det op.  
I wrote not it up  
b. Jeg skrev det ikke \_\_\_\_ op.

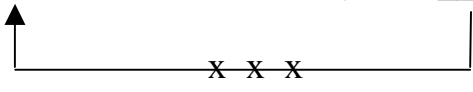






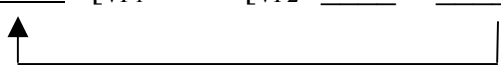
- (i) Sw    a. Mannen såg                 inte den.  
                man-the saw              not it  
                b. Mannen såg den inte . (Josefsson 2003: 201)



(13) Remnant VP-topicalisation? Holmberg (1997, 1999): NO!

- a. [CP har [IP jag [VP<sub>1</sub> inte [VP<sub>2</sub> kysst henne]]]]]
- b. [CP har [IP jag henne [VP<sub>1</sub> inte [VP<sub>2</sub> kysst \_\_\_\_]]]]]
- 
  
**violation of HG!!!**
- c. [CP [VP<sub>2</sub> Kysst \_\_\_\_] har [IP jag henne [VP<sub>1</sub> inte \_\_\_\_]]]
- 

(14) V°-topicalisation? Holmberg (1997, 1999): YES!

- a. [CP har [IP jag [VP<sub>1</sub> inte [VP<sub>2</sub> kysst henne]]]]]
- b. [CP [V° Kysst] har [IP jag [VP<sub>1</sub> inte [VP<sub>2</sub> \_\_\_\_ henne]]]]]
- 
- c. [CP [V° Kysst] har [IP jag henne [VP<sub>1</sub> inte [VP<sub>2</sub> \_\_\_\_ \_\_\_\_]]]]]
- 

Note that OS in the V°-topicalisation analysis is countercyclic: It targets a lower position than the previous movement of V°, which is why Holmberg (1997, 1999) has to assume that OS does not take place in syntax proper but in a special part of the grammar, Stylistic Syntax, where Chomsky's (1993: 22) Extension Condition does not hold. Moreover, the V°-topicalisation analysis involves movement of an X° to an XP-position.

Furthermore, if V°-topicalisation would be possible, we would expect the sentences in (15)b/(16)b to be acceptable, contrary to fact.

- (15) Da a. Jeg har ikke smidt den ud.  
           I have not thrown it out  
       b. \*Smidt har jeg den ikke \_\_\_\_ ud.
- (16) Da a. Jeg har ikke stillet det på bordet.  
           I have not put it on table-the  
       b. \*Stillet har jeg det ikke \_\_\_\_ på bordet.



main clause and the pronominal object undergoes OS, maintaining their relative order  $V < O$ .

- (18) Da a. Jeg kyssede hende ikke \_\_\_\_ \_\_\_\_.  
           I     kissed   her     not  
       b. VP: [<sub>VP</sub> V O]  
             Ordering:        **$V < O$**   
       c. CP: [<sub>CP</sub> S V [<sub>IP</sub> t<sub>S</sub> O Adv [<sub>VP</sub> t<sub>V</sub> t<sub>O</sub>]]]  
             Ordering:        **$S < V$**                 **$V < O$**   
                                    **$V < O$**   
                                    **$O < Adv$**   
                                    **$Adv < VP \rightarrow \emptyset$**

By contrast, OS is impossible in an embedded clause as (19) in MSc. The ordering statements produced at Spell-out of CP, (19)c, contradict the statement " $V < O$ " established at Spell-out of VP, (19)b: Given that at CP, the object precedes the adverb (" $O < Adv$ ") which in turn precedes the verb (" $Adv < V$ "), the object must precede the verb - in contrast to their relative order at VP.

- (19) Da a. \*... at jeg hende ikke kyssede \_\_\_\_.  
           that I   her     not   kissed  
       b. VP: [<sub>VP</sub> V O]  
             Ordering:        **$V < O$**   
       c. CP: [<sub>CP</sub> Comp [<sub>IP</sub> S O Adv [<sub>VP</sub> V t<sub>O</sub>]]]  
             Ordering:        **$C < S$**                 **$V < O$**   
                                    **$S < O$**   
                                    **$O < Adv$**   
                                    **$Adv < VP \rightarrow Adv < V$**

Hence, Fox & Pesetsky (2005) derives HG from ordering contradictions. OS cannot take place if it results in ordering statements at CP that contradict those established at Spell-out of VP. Correspondingly, the asymmetry between stranding of an IO and stranding of a DO by remnant VP-topicalisation illustrated in (17) above is expected by order preservation. Stranding of an IO, but not stranding of a DO gives rise to contradictory ordering statements at the

various Spell-out domains: At VP, "IO<DO" is established, which is maintained at the Spell-out of CP in (17)a but not in (17)b.

Note that Fox & Pesetsky (2005) predicts that movement operations that do not obey HG have to proceed successive cyclically: The underlined constituents in (20) have to move through the edge of VP prior to linearisation of the VP domain to prevent ordering contradictions at the Spell-out of CP. These movement operations comprise various instances of A-movement and A-bar-movement operations, such as Scandinavian Negative Shift (see Christensen 2005), *wh*-movement, topicalisation, and subject raising.

- (20) Da    a. Måske       har       han       ingen bøger læst \_\_\_\_\_.  
               probably    has       he       no books    read \_\_\_\_\_  
               b. Hvad       har       du                        læst \_\_\_\_\_?  
                       what       have    you                        read \_\_\_\_\_  
               c. Bøgerne    har       jeg                        læst \_\_\_\_\_.  
                       books-the    have    I                        read \_\_\_\_\_  
               d. Måske       blev       bøgerne                læst \_\_\_\_\_.  
                       perhaps       were    books-the                read \_\_\_\_\_

Hence, the crucial difference between the various movement operations in (20) and OS is that the former may - and indeed must – go through the edge of VP, but as Fox & Pesetsky (2003) states, in their analysis OS cannot involve movement to the edge of VP.

### 3 An Optimality-theoretic approach to object shift and remnant VP-topicalisation

#### 3.1 Object shift and order preservation

Although there are a number of OT analyses of OS, the ones we are familiar with, e.g. Broekhuis (2000) or Vogel (2004), predate Fox & Pesetsky (2005) and do not consider remnant VP-topicalisation at all. As far as we can tell, these analyses would not be able to account for it.

Building on the insights of Fox & Pesetsky (2005), we consider HG to be a linear restriction. The condition on order preservation is expressed by the constraint in (21)a that requires base order precedence relations among non-adverbial elements to be maintained at the final representation; cf. Déprez (1994), Müller (2001), Sells (2001), and Williams (2003). Pronominal OS is taken to be triggered by the constraint SHIFTPRONOUN in (21)b:<sup>7</sup>

- (21) a. ORDER PRESERVATION (ORDPRES):  
       If the foot of the chain of some non-adverbial element  $\alpha$  precedes the foot of the chain of some element  $\beta$ , the head of the chain of  $\alpha$  also precedes the head of the chain of  $\beta$ .  
       b. SHIFTPRONOUN (SHIFTPRON):  
       A weak pronoun precedes and c-commands the lowest VP (of the same clause) that contains all other VPs and all VP-adjoined adverbials.

SHIFTPRON requires movement of a pronoun to a position at the left edge of VP; ORDPRES penalizes this movement if it results in the reversal of the order of non-adverbial elements. Hence, the ranking ORDPRES >> SHIFTPRON captures HG: The violation of ORDPRES blocks OS across an intervening non-adverbial element such as the *in situ* main verb in Tableau 1. However, if the main verb moves itself to a position to the left of the target position of OS, OS is possible since the base order precedence relation between the verb and its object are maintained in accordance with ORDPRES; compare Tableau 2.<sup>8</sup>

<sup>7</sup> As mentioned in footnote 1, OS may also apply to full DPs in Icelandic. Vikner & Engels (2006) considers full DP Shift to be triggered by a more general constraint SHIFT that requires a non-focused constituent to precede and c-command the lowest VP (of the same clause). The contrast between Icelandic and MSc in the applicability of OS to full DPs depends on the relative ranking between SHIFT and STAY, see (31) below.

<sup>8</sup> In contrast to HG in (4), ORDPRES is not restricted to OS; rather, the constraint penalizes any kind of movement that changes the order of elements. The fact that OS contrasts with other types of movement operations, such as the ones mentioned in (20) above, in that the latter ones do not have to preserve the base order can be captured by differences in the ranking of ORDPRES relative to the constraints that require the corresponding movements.

**Tableau 1: Blocking of OS by intervening verb**

Da	ORDPRES	SHIFTPRON	ex.
☞ a Sub Aux Adv <u>V</u> <u>Pron-Obj</u>		*	(2)a
b Sub Aux <u>Pron-Obj</u> Adv <u>V</u> <u>t<sub>O</sub></u>	*!		(2)b

**Tableau 2: Object Shift**

Da	ORDPRES	SHIFTPRON	ex.
a Sub <u>V</u> Adv <u>t<sub>v</sub></u> <u>Pron-Obj</u>		*!	(1)a
☞ b Sub <u>V</u> <u>Pron-Obj</u> Adv <u>t<sub>v</sub></u> <u>t<sub>O</sub></u>			(1)b

The ranking  $\text{ORDPRES} \gg \text{SHIFTPRON}$  does not only predict that OS is blocked by intervening non-adverbial elements, it also accounts for the fact that multiple OS has to maintain the order of elements.

- (22) Da a. \*Jeg gav ikke hende det.  
           I gave not her it  
       b. \*Jeg gav hende ikke det.  
       c. \*Jeg gav det hende ikke \_\_\_\_\_.  
       d. Jeg gav hende det ikke \_\_\_\_\_.

For example,  $\text{ORDPRES}$  is outranked by the constraint  $\text{WHSPEC}$  that requires *wh*-movement to Spec,CP ( $\text{WHSPEC} \gg \text{ORDPRES}$ ), predicting that unlike OS, *wh*-movement is not blocked by an intervening verb, verbal particle, or object; compare (2), (6), and (8), respectively.

- (i) Da a. Hvad har Peter læst \_\_\_\_?  
           what has Peter read  
       Sw b. Vad smutsade Kalle ner \_\_\_\_?  
           what dirtied Kalle down  
       Sw c. Vad gav Kalle Elsa \_\_\_\_?  
           what gave Kalle Elsa

((i)b,c from Bobaljik 2002: 236)



**Tableau 3: Multiple OS**

Da	ORDPRES	SHIFTPRON	ex.
a Sub V Adv <u>Pron-IO</u> <u>Pron-DO</u>		*!*	(22)a
b Sub V <u>Pron-IO</u> Adv t <sub>IO</sub> <u>Pron-DO</u>		*!	(22)b
c Sub V <u>Pron-DO</u> <u>Pron-IO</u> Adv t <sub>IO</sub> t <sub>DO</sub>	*!		(22)c
☞ d Sub V <u>Pron-IO</u> <u>Pron-DO</u> Adv t <sub>IO</sub> t <sub>DO</sub>			(22)d

### 3.2 Asymmetry I: Stranding of IO vs. stranding of DO

As shown in the preceding section, the ranking  $\text{ORDPRES} \gg \text{SHIFTPRON}$  captures the fact that OS is blocked by an intervening non-adverbial element, predicting that OS is dependent on movement of the main verb. However for OS to be possible, the main verb does not necessarily have to undergo  $V^\circ$ -to- $I^\circ$ -to- $C^\circ$  movement as in Tableau 2. What is crucial is that the main verb moves to a position in front of the target position of OS, such that their relative order is preserved. This can also be achieved by placing a non-finite verb in topic position, as illustrated in Tableau 4.

**Tableau 4: OS with remnant VP-topicalisation**

Da	ORDPRES	SHIFTPRON	ex.
a <u>V</u> Aux Sub Adv <u>Pron-Obj</u>		*!	(11)
☞ b <u>V</u> Aux Sub <u>Pron-Obj</u> Adv			(10)

We propose that in this case the pronominal object undergoes OS prior to remnant VP-topicalisation. In Holmberg's (1997, 1999) approach such remnant VP-topicalisation is ruled out by the assumption that HG is derivational, i.e. that it cannot be violated at any point in the derivation, compare (13) above. The OT constraint  $\text{ORDPRES}$ , by contrast, is representational: Constraint violations are computed based on the final structure of the candidates. Hence, although the



(25) TOPIC: Elements with a [+topic] feature occur in Spec,CP.

**Tableau 5: VP-topicalisation, taking along both IO and DO**

Da/Sw	Topic: V & Pron-IO & Pron-DO	TOPIC	ORDPRES	SHIFTPRON	ex.
☞ a	[ <sub>VP</sub> V <u>Pron-IO</u> <u>Pron-DO</u> ] Aux Sub Adv t <sub>VP</sub>			**	(24)a
b	[ <sub>VP</sub> V <u>Pron-IO</u> t <sub>DO</sub> ] Aux Sub <u>Pron-DO</u> Adv t <sub>VP</sub>	*!		*	(23)a
c	[ <sub>VP</sub> V t <sub>IO</sub> <u>Pron-DO</u> ] Aux Sub <u>Pron-IO</u> Adv t <sub>VP</sub>	*!	*	*	(23)b
d	[ <sub>VP</sub> V t <sub>IO</sub> t <sub>DO</sub> ] Aux Sub <u>Pron-IO</u> <u>Pron-DO</u> Adv t <sub>VP</sub>	*!*			(24)b

**Tableau 6: Remnant VP-topicalisation, stranding both IO and DO**

Da/Sw	Topic: V	TOPIC	ORDPRES	SHIFTPRON	ex.
a	[ <sub>VP</sub> V <u>Pron-IO</u> <u>Pron-DO</u> ] Aux Sub Adv t <sub>VP</sub>			*!*	(24)a
b	[ <sub>VP</sub> V <u>Pron-IO</u> t <sub>DO</sub> ] Aux Sub <u>Pron-DO</u> Adv t <sub>VP</sub>			*!	(23)a
c	[ <sub>VP</sub> V t <sub>IO</sub> <u>Pron-DO</u> ] Aux Sub <u>Pron-IO</u> Adv t <sub>VP</sub>		*!	*	(23)b
☞ d	[ <sub>VP</sub> V t <sub>IO</sub> t <sub>DO</sub> ] Aux Sub <u>Pron-IO</u> <u>Pron-DO</u> Adv t <sub>VP</sub>				(24)b

As Tableau 5 and Tableau 6 show, SHIFTPRON favors stranding of a pronoun which is, however, only possible if the pronoun is not marked [+topic]. The asymmetry between stranding of a DO and stranding of an IO is expected by the ranking ORDPRES >> SHIFTPRON. OS of a DO maintains the ordering relations in remnant VP-topicalisations, satisfying ORDPRES (see Tableau 7).<sup>10</sup> In contrast, remnant VP-topicalisation does not re-establish the base order relations if the IO is stranded. Consequently, the violation of ORDPRES rules out stranding of the

<sup>10</sup> Note that it is crucial for the remnant VP-topicalisation constructions that OrdPres refers to precedence rather than c-command relations: While the precedence relations are maintained in (23)a, the c-command relations are not - neither the verb nor the IO c-commands the shifted DO.

IO in OS position, compare Tableau 8 below. Instead, the IO has to be taken along by VP-topicalisation, giving rise to neutralization: Despite the different input specifications with regard to topichood, the same candidate (namely, candidate a) arises as output in Tableau 5 and Tableau 8. (But note that stranding of the IO is possible if it does not result in a violation of ORDPRES, namely if both objects are stranded as in (24)b.)

**Tableau 7: Remnant VP-topicalisation that strands DO**

Da/Sw Topic: V & Pron-IO	TOPIC	ORDPRES	SHIFTPRON	ex.
a [VP V <u>Pron-IO</u> <u>Pron-DO</u> ] Aux Sub Adv t <sub>VP</sub>			**!	(24)a
☞ b [VP V <u>Pron-IO</u> t <sub>DO</sub> ] Aux Sub <u>Pron-DO</u> Adv t <sub>VP</sub>			*	(23)a
c [VP V t <sub>IO</sub> <u>Pron-DO</u> ] Aux Sub <u>Pron-IO</u> Adv t <sub>VP</sub>	*!	*	*	(23)b
d [VP V t <sub>IO</sub> t <sub>DO</sub> ] Aux Sub <u>Pron-IO</u> <u>Pron-DO</u> Adv t <sub>VP</sub>	*!			(24)b

**Tableau 8: No remnant VP-topicalisation that strands IO**

Da/Sw Topic: V & Pron-DO	TOPIC	ORDPRES	SHIFTPRON	ex.
☞ a [VP V <u>Pron-IO</u> <u>Pron-DO</u> ] Aux Sub Adv t <sub>VP</sub>			**	(24)a
b [VP V <u>Pron-IO</u> t <sub>DO</sub> ] Aux Sub <u>Pron-DO</u> Adv t <sub>VP</sub>	*!		*	(23)a
c [VP V t <sub>IO</sub> <u>Pron-DO</u> ] Aux Sub <u>Pron-IO</u> Adv t <sub>VP</sub>		*!	*	(23)b
d [VP V t <sub>IO</sub> t <sub>DO</sub> ] Aux Sub <u>Pron-IO</u> <u>Pron-DO</u> Adv t <sub>VP</sub>	*!			(24)b

Similarly, the unacceptable sentence in (12)d, repeated here as (26)c, is ruled out by the violation of ORDPRES. These data led Holmberg (1997, 1999) to assume that remnant VP-topicalisation is not possible.

- (Holmberg 1997: 206)

Sw	Topic: V & V	TOPIC	ORDPRES	SHIFTPRON	ex.
☞ a	[ <sub>VP</sub> <b>V</b> <u>Pron</u> <b>V</b> ] Aux Sub Adv			*	(26)b
b	[ <sub>VP</sub> <b>V</b> t <sub>Pron</sub> <b>V</b> ] Aux Sub <u>Pron</u> Adv		*!		(26)c

(27) Da	a.	[ <sub>VP</sub> Smidt <u>den</u> ud]	har	jeg	ikke.
		thrown it out	have	I	not
	b.	*[ <sub>VP</sub> Smidt ____ ud]	har	jeg	<u>den</u> ikke.
(28) Da	a.	[ <sub>VP</sub> Stillet <u>det</u> på bordet]	har	jeg	ikke.
		put it on table-the	have	I	not
	b.	*[ <sub>VP</sub> Stillet på bordet]	har	jeg	<u>det</u> ikke.

(i) Sw Bo ska han — i Malmö, men han ska jobba i Köpenhamn.  
live will he in Malmö but he will work in Copenhagen  
(Holmberg 1999: 12)

- (29) Da a. \*<sub>[VP]</sub> Smidt den \_\_\_ har jeg ikke ud.  
           thrown it have I not out  
       b. \*<sub>[VP]</sub> Smidt \_\_\_ \_\_\_ har jeg den ikke ud.
- (30) Da a. \*<sub>[VP]</sub> Stillet det \_\_\_ har jeg ikke på bordet.  
           put it have I not on table-the  
       b. \*<sub>[VP]</sub> Stillet \_\_\_ \_\_\_ har jeg det ikke på bordet.

As argued above, stranding of a pronominal object is triggered by SHIFTPRON, requiring a VP-external position for the pronoun. Elements for which movement is not independently required by some constraint cannot be stranded by remnant VP-topicalisation due to the constraint STAY.<sup>12</sup>

<sup>12</sup> Note that not just pronominal objects may be left behind when the verb occurs in clause-initial position, but - according to Holmberg (1999: 10) - also epithetic DPs may be stranded. "V-Topicalization requires narrow contrastive focus on V, and is therefore most natural when other VP-constituents are 'defocused', in which case they are most naturally referred to by pronouns. [...] In terms of information structure a pronominal epithet is equal to a pronoun, but with regard to Object Shift, they behave like full DPs, i.e. they are generally not shifted in MSc. [...] [(i)a] featuring a pronominal epithet as object should be as well-formed as [(10)a], which indeed it is" (Holmberg 1999: 10).

- (i) Sw a. Sett har jag inte den idioten, ...  
           seen have I not that idiot  
       b. \*Sett har jag den idioten inte \_\_\_\_\_, ...  
           (... men jag har talat med honom på telefon).  
           but I have talked with him on phone (Holmberg 1999: 11)

Stranding of full DPs is unexpected in our approach as they cannot undergo OS (in MSc) and the remnant VP therefore is not expected to be a constituent. This goes not only for epithets as in (i) but also for focused non-epithetic DPs which can be stranded too, (ii).

- (ii) Da Kysset har jeg ikke MARIE, men SOPHIE.  
           kissed have I not Marie but Sophie

At the first glance, the fact that full DPs can be stranded although they cannot undergo OS would seem to support Holmberg's (1999) claim that these sentences involve V°-topicalisation rather than remnant VP-topicalisation. However, stranding of a full DP is also possible in clear instances of remnant VP-topicalisation as the one in (iii) where the DO of a double object verb is left behind.

Tableau 10: No stranding of other VP-internal right-peripheral constituents

Da	Topic: V & Pron-Obj	TOPIC	ORDPRES	SHIFTPRON	STAY	ex.
☞ a	[ <sub>VP</sub> <b>V Pron-Obj</b> <u>PP</u> ] Aux Sub Adv					(28)a
b	[ <sub>VP</sub> <b>V Pron-Obj</b> t <sub>PP</sub> ] Aux Sub Adv <u>PP</u>				*!	(30)a

(32) Da a. \*Jeg kan ikke uden briller [VP læse den].  
I can not without glasses read it  
b. Jeg kan ikke [VP læse den]  
uden briller.  
c. [VP Læse den] kan jeg ikke \_\_\_\_\_  
uden briller

From the discussion in the previous section, we might expect that all that matters is that the stranded object originally occupied a right-peripheral position in the topicalised remnant VP, because then all orderings are preserved. However, not all objects on the right edge may be left behind during VP-topicalisation: The object of an infinitival clause cannot be stranded by remnant topicalisation of the main clause VP although it is the rightmost element within that VP.

(iii)Sw Har du verkligen lånat Per din gamla dator?  
 ('Have you really lent Per your old computer?')  
 Lånat honom har jag inte det gamla skitet, jag har GETT honom det.  
 lent him have I not the old crap I have given him it  
 (Gunlög Josefsson, p.c.)





(36) Sw  $[\text{VP } \textbf{kastat} \text{ thrown } [\text{PrTP } t_o' [\text{PrTP } \textbf{bort} \text{ out } t_o]]]$  = (34)b

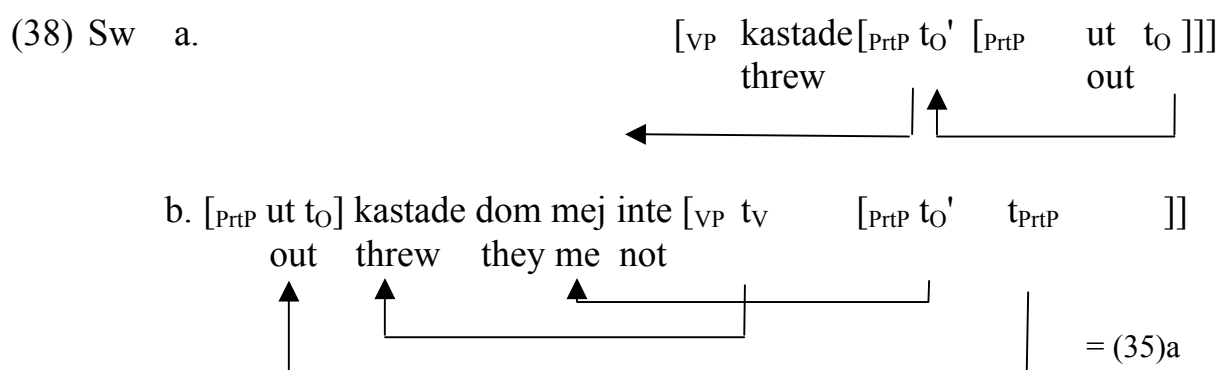
(37) SW    [VP t<sub>DO'</sub> [VP **gett** [VP t<sub>V</sub> **henne** t<sub>DO</sub>]]] = (23)a  
              given                  her

<sup>13</sup> Similarly, remnant topicalisation of the main VP in (33)b would have to take along the intermediate trace: The two VPs do not have the same head such that OS would have to involve adjunction to the embedded VP.

a. [<sub>CP</sub> [<sub>VP</sub> set [<sub>IP</sub> ham [<sub>VP</sub> fotografere hende]]] har [<sub>IP</sub> jeg [<sub>VP</sub> ikke t<sub>VP</sub>]]]  
seen him photograph her have I not  
b. \*[[<sub>CP</sub> [<sub>VP</sub> set [<sub>IP</sub> ham [<sub>VP</sub> fotograferer t<sub>O</sub>]]]] har [<sub>IP</sub> jeg hende [<sub>VP</sub> ikke t<sub>VP</sub>]]]

Spec,CP in both (36) and (37)) may be licensed in a different way, e.g. simply by being in a thematic position.<sup>14</sup>

The difference between (34) and (35) is now that in (35), only the PrtP is topicalised (the verb is also moved, but by a different movement, V°-to-I°-to-C° movement) and so there does not have to be an intermediate trace inside VP, and, therefore it is possible for remnant VP-topicalisation to take place without an intermediate trace occurring in Spec,CP.



To sum up, remnant VP-topicalisation may strand an object in OS position as long as the precedence relations are maintained (ORDPRES >> SHIFTPRON) and its base position is not too deeply embedded (i.e. the topicalised VP does not

<sup>14</sup> Under the assumption that the intermediate step has to target the minimal XP of the selecting/theta-assigning head (excluding any adjuncts to XP), the ungrammaticality of (i)c follows: The intermediate trace of the object is adjoined to the inner segment of VP such that topicalisation of the outermost segment of VP necessarily takes this trace along. In contrast, remnant topicalisation of the innermost VP (excluding the intermediate trace) is marginally acceptable.

- (i) Da. a. Han har måske nok [VP omhyggeligt læst den]  
 he has possibly well carefully read it  
 men har han forstået den?  
 but has he understood it?
- b. ?[VP Omhyggeligt [VP læst den]] har han måske nok, ...
- c. \*[VP Omhyggeligt [VP tO' [VP læst tO]]] har han den måske nok, ...
- d. ??[VP Læst tO] har han den ikke [VP omhyggeligt [VP tO' tVP]],  
 read has he it not carefully,  
 kun ret overfladisk.  
 only rather superficially

contain an intermediate trace). Consequently, only an object that is right-peripheral in VP may be left behind, giving rise to the asymmetry between stranding of an IO and stranding of a DO.

### 3.4 Asymmetry II: Object shift vs. subject raising

Apart from the asymmetry between stranding of a DO and stranding of an IO discussed in section 3.2, there is an asymmetry between remnant VP-topicalisation leaving behind an argument in OS position and remnant VP-topicalisation leaving behind an argument in subject position. This indicates that a non-peripheral trace in the topicalised VP is not a problem as such. The base order of elements does not have to be maintained by remnant VP-topicalisation if the element that has left VP occurs in subject position (as in passives), while the order cannot be changed if it occurs in OS position, see (39)b/(41)b vs. (40)b/(42)b.

- (39) Da a. [<sub>VP</sub> Smidt den ud] har jeg ikke.  
           thrown it out have I not  
       b. \*[<sub>VP</sub> Smidt \_\_\_\_ ud] har jeg den ikke.
- (40) Da a. \*[<sub>VP</sub> Smidt den ud] blev ikke.  
           thrown it out was not  
       b. [<sub>VP</sub> Smidt \_\_\_\_ ud] blev den ikke.
- (41) Da a. [<sub>VP</sub> Stillet det på bordet] har jeg ikke.  
           put it on table-the have I not  
       b. \*[<sub>VP</sub> Stillet \_\_\_\_ på bordet] har j eg det ikke.
- (42) Da a. \*[<sub>VP</sub> Stillet det på bordet] blev ikke.  
           put it on table-the was not  
       b. [<sub>VP</sub> Stillet \_\_\_\_ på bordet] blev det ikke.

This contrast is accounted for if ORDPRES is outranked by the constraint that triggers subject movement to Spec,IP, cf. Grimshaw & Samek-Lodovici (1995), Samek-Lodovici (1996, 1998).<sup>15</sup>

- (43) SUBJECT: The highest A-specifier is structurally realized.  
(Samek-Lodovici 1998: 4)

**Tableau 11: No object stranding in Danish particle verb constructions**

Da	Topic: V & Prt	SUBJECT	ORD PRES	SHIFT PRON	STAY	ex.
☞ a	[ <sub>VP</sub> V <u>Pron-Obj</u> Prt] Aux Sub Adv			*		(39)a
b	[ <sub>VP</sub> V <sub>t<sub>O</sub></sub> Prt] Aux Sub <u>Pron-Obj</u> Adv		*!		*	(39)b

**Tableau 12: Subject stranding in Danish particle verb constructions**

Da	Topic: V & Prt	SUBJECT	ORD PRES	SHIFT PRON	STAY	ex.
a	[ <sub>VP</sub> V <u>Pron-Sub</u> Prt] Aux e Adv	*!		*		(40)a
☞ b	[ <sub>VP</sub> V <sub>t<sub>S</sub></sub> Prt] Aux <u>Pron-Sub</u> Adv		*		*	(40)b


The ranking SUBJECT >> ORDPRES is supported by the fact that movement to subject position does not depend on verb movement; i.e. subject movement may cross an intervening (unaccusative) verb.

- (44) Da a. Derfor har Elsa ikke \_\_\_\_\_ ringet.  
          therefore has Elsa not \_\_\_\_\_ called  
      b. Derfor er Elsa ikke kommet \_\_\_\_\_.  
          therefore is Elsa not \_\_\_\_\_ come

<sup>15</sup> Under the assumption that all extraction out of VP has to proceed via adjunction to the minimal XP containing the selecting/theta-assigning head (see section 3.3), the grammaticality of (40)b suggests that the prohibition against intermediate traces in Spec,CP also is a violable constraint (which is outranked by SUBJECT). *Den* ('it') in (40)b originates in the complement position of the particle and it would thus have to adjoin to PrtP before moving on to the subject position. Consequently, the topicalised VP includes an intermediate trace of the subject.

(i) Da [<sub>VP</sub> Smidt [<sub>PrtP</sub> t' [<sub>PrtP</sub> ud t]]] blev den ikke.

**Tableau 13: Subject raising**


Da	SUBJECT	ORDPRES	SHIFTPRON	STAY	ex.
a e Aux V <u>DP</u>	*!				-
 b <u>DP</u> Aux V t <sub>DP</sub>		*		*	(44)b

At the same time, ORDPRES predicts that in double object constructions the IO rather than the DO is promoted to subject in passives, as borne out in e.g. Danish and English.<sup>16</sup>

- (45) Da a. Derfor har jeg ikke givet Elsa bogen.  
           therefore have I not given Elsa book-the  
       b. Derfor blev Elsa ikke givet \_\_\_\_ bogen.  
           therefore was Elsa not given \_\_\_\_ book-the  
       c. \*Derfor blev bogen ikke givet Elsa \_\_\_\_.

- (46) En a. I did not give Elsa the book.  
       b. Elsa was not given \_\_\_\_ the book.  
       c. \*The book was not given Elsa \_\_\_\_.

**Tableau 14: Promotion to subject in passive double object constructions**

Da	SUBJECT	ORDPRES	SHIFTPRON	STAY	ex.
a e Aux V <u>DP</u> <u>DP</u>	*!				-
 b <u>DP</u> Aux V t <sub>DP</sub> <u>DP</u>		*		*	(45)b
c <u>DP</u> Aux V <u>DP</u> t <sub>DP</sub>		**!		*	(45)c


<sup>16</sup> However, promotion of the DO to subject in passive double object constructions is possible in Swedish and Norwegian.

- (i) No a. Marie gav ham den.  
           Maria gave him it  
       b. Han ble gitt \_\_\_\_ den.  
           he was given it  
       c. Den ble gitt ham \_\_\_\_.  
           it was given him

As expected by ORDPRES, promotion of the DO to the subject of a passive is possible if the recipient is expressed by a PP because in that case the DO precedes the PP in the base order.

- (47) Da    a. Derfor    har    jeg    ikke givet    bogen    til Elsa.  
               therefore have I    not given    book-the    to Elsa  
               b. Derfor    blev    bogen    ikke givet    \_\_\_\_\_    til Elsa.  
               therefore was    book-the    not given    \_\_\_\_\_    to Elsa  
               c. \*Derfor    blev    Elsa    ikke givet    bogen    til \_\_\_\_.
- (48) En    a. I    did not give    the book    to Elsa.  
               b. The book    was not given    \_\_\_\_\_    to Elsa.  
               c. \*Elsa    was not given    the book    to \_\_\_\_.

**Tableau 15: Promotion to subject in passive DP PP constructions**

Da	SUBJECT	ORDPRES	SHIFTPRON	STAY	ex.
a    e Aux V <u>DP</u> <u>PP</u>	*!				-
 b <u>DP</u> Aux V t <sub>DP</sub> <u>PP</u>		*		*	(47)b
c <u>DP</u> Aux V <u>DP</u> [P t <sub>DP</sub> ]		***!*		*	(47)c

Hence, the asymmetry between a subject and an object not moving along in remnant VP-topicalisations is accounted for by the difference in the ranking of SUBJECT and SHIFTPRON relative to ORDPRES: SUBJECT >> ORDPRES >> SHIFTPRON predicts that OS but not subject raising is blocked whenever it would result in a reversal of the order relations.

### 3.5 Asymmetry III: Remnant VP-topicalisation out of a main clause vs. an embedded clause

A third asymmetry in the availability of remnant VP-topicalisation concerns the depth of embedding of the topicalised VP, namely whether the remnant VP is topicalised out of a main clause or out of an embedded clause.

As shown in (49), a full VP may be topicalised from both main clauses and embedded clauses.

- (49) Da a. [<sub>VP</sub> Set ham] har jeg ikke, ...  
           seen him have I not  
           ... hvis jeg skal være ærlig, men jeg har talt i telefon med ham.  
           if I should be totally honest but I have spoken in phone with him
- b. [<sub>VP</sub> Set ham] tror jeg ikke at du har, ...  
           seen him believe I not that you have  
           ... men du kan måske nok have talt i telefon med ham.  
           but you may perhaps well have spoken in phone with him

Topicalisation of a remnant VP, by contrast, is only possible out of a main clause, (50)a, not out of an embedded clause in Danish: The stranded object may neither follow the finite auxiliary (in its base position), (50)b, nor may it precede it, (50)c:

- (50) Da a. [<sub>VP</sub> Set \_\_\_\_] har jeg ham ikke, ...  
           seen have I him not  
           ... hvis jeg skal være ærlig, men jeg har talt i telefon med ham.  
           if I should be totally honest but I have spoken on phone-the  
           with him
- b. \* [<sub>VP</sub> Set \_\_\_\_] tror jeg ikke at du [<sub>V°</sub> har ] ham, ...  
           seen believe I not that you have him
- c. \* [<sub>VP</sub> Set \_\_\_\_] tror jeg ikke at du ham [<sub>V°</sub> har ], ...  
           seen believe I not that you him have  
           ... men du kan måske nok have talt i telefon med ham.  
           but you may perhaps well have spoken in phone with him

This asymmetry is expected under the present analysis: As discussed in section 3.2, stranding of an object involves OS (it is motivated by SHIFTPRON); constituents whose movement out of VP is not independently triggered by some constraint cannot be stranded by remnant VP-topicalisation due to STAY. Targeting a position to the left of the base position of the finite verb, OS is only available if the verb has itself left its base position (ORDPRES >> SHIFTPRON) which it does not in embedded clauses in MSc; cf. (3) above. Hence, the difference between main clauses and embedded clauses in finite verb movement

is crucial for the asymmetry of remnant VP-topicalisation out of main clauses vs. out of embedded clauses.

**Tableau 16: Remnant VP-topicalisation out of a main clause**

Da	Topic: V	ORDPRES	SHIFTPRON	STAY	ex.
a	[ <sub>VP</sub> V <u>Pron-Obj</u> ] Aux Sub Adv t <sub>VP</sub>		*!		(49)a
b	[ <sub>VP</sub> V t <sub>O</sub> ] Aux Sub Adv <u>Pron-Obj</u> t <sub>VP</sub>		*!	*	-
☞ c	[ <sub>VP</sub> V t <sub>O</sub> ] Aux Sub <u>Pron-Obj</u> Adv t <sub>VP</sub>			*	(50)a

**Tableau 17: No remnant VP-topicalisation out of an embedded clause**

Da	Topic: V	ORDPRES	SHIFTPRON	STAY	ex.
☞ a	[ <sub>VP</sub> V <u>Pron-Obj</u> ] V Sub Adv Comp Sub [ <sub>VP</sub> Aux t <sub>VP</sub> ]		*		(49)b
b	[ <sub>VP</sub> V t <sub>O</sub> ] V Sub Adv Comp Sub [ <sub>VP</sub> Aux <u>Pron-Obj</u> t <sub>VP</sub> ]		*	*!	(50)b
c	[ <sub>VP</sub> V t <sub>O</sub> ] V Sub Adv Comp Sub <u>Pron-Obj</u> [ <sub>VP</sub> Aux t <sub>VP</sub> ]	*!		*	(50)c

Note that remnant VP-topicalisation from embedded clauses is possible in passives, i.e. if the element left behind occurs in subject position. This follows from SUBJECT being ranked higher than ORDPRES, as in Tableau 13 above.

- (51) Da a. [<sub>VP</sub> Set \_\_\_\_] blev han ikke, ...  
           seen               was he not  
       b. [<sub>VP</sub> Set \_\_\_\_] tror jeg ikke at han blev, ...  
           seen               think I not that he was  
                               ... men der var nok mange der hørte ham.  
                               but there were probably many who heard him

The hypothesis that object stranding has to involve OS seems to be supported by the fact that Icelandic (which has V<sup>o</sup>-to-I<sup>o</sup> movement and, consequently, also OS in embedded clauses, cf. footnote 2), marginally permits a stranded object in



(52) Ic    ??[<sub>VP</sub> Kysst \_\_\_\_] hélt ég ekki að þú [<sub>i</sub> hefðir] hana oft, ...  
kissed think I not that you have her often  
... bara haldið í höndina á henni.  
only held in hand.the on her  
(Gunnar Hrafn Hrafnbjargarson, p.c.)

However, Fox & Pesetsky (2005) have presented data from double object constructions that clearly show that remnant VP-topicalisation is possible, as long as it does not involve a reversal of the base order of elements, and suggesting that HG is representational. We have collected more data that corroborate Fox & Pesetsky's observation and we agree with them in the assumption that HG is to be accounted for in terms of order preservation. Their approach builds on the assumption that Spell-out applies at various points in the derivation (in particular, at VP and at CP) and that the information about the linearisation of the material of a newly constructed Spell-out domain must not contradict the cumulated information of previous applications of Spell-out. In this way, Fox & Pesetsky (2005) predict that OS differs radically from other types of (A- and A-bar-) movement that can result in a reversal of the order of elements, such as e.g. *wh*-movement or subject raising, in that the latter have to

proceed successive cyclically through the left edge of VP while this is impossible for OS.

In contrast, in our OT approach, order preservation is required by a violable constraint. This means that it is the ranking of the ORDERPRESERVATION constraint relative to the constraints that motivate the various types of movement which accounts for the contrast as to whether or not a certain movement operation has to be order preserving. Hence, OS does not receive a special treatment in our approach; the properties distinguishing it from other movement types result from constraint interaction.

The linear conception of HG as expressed by the constraint ORDPRES and its dominance over the constraint that triggers OS, SHIFTPRON, predicts that only pronominal objects that originate in a right-peripheral position within VP might be left behind in OS position during remnant VP-topicalisation, accounting for the asymmetry in stranding of an IO and stranding of a DO observed by Fox & Pesetsky (2005). However, depth of embedding also plays a role for whether or not an object may have undergone OS out of a topicalised VP: The remnant VP in Spec,CP may not include an intermediate trace of a shifted object. Moreover, we presented new data that showed that subject raising does not underly either of these restrictions, and this may be accounted for by a different ranking of SUBJECT and SHIFTPRON relative to the corresponding prohibitions (including ORDPRES).

Finally, the asymmetry between main and embedded clauses as to the applicability of remnant VP-topicalisation in MSc illustrates that object stranding has to involve OS. Object stranding is only possible in sentences in which finite verb movement has taken place, something that would be expected if any object left behind during remnant VP-topicalisation would have to undergo OS (and that as always, OS has to respect order preservation).

## Appendix 1: Structure Preservation

There are native speakers of Danish whose intuitions do not agree with the acceptability judgments given above. Rather than to subject remnant VP-topicalisation to a linear restriction, permitting stranding of an object in OS position as long as it does not change the base order of elements (cf. (23) and (24) above), these speakers do not allow for object stranding during remnant VP-topicalisation at all. Topicalisation of a full VP, in contrast, is judged acceptable.

- (53) Da a. [<sub>VP</sub> Givet hende den] har jeg ikke.  
           given her it have I not  
       b. \* [<sub>VP</sub> Givet \_\_\_\_\_ ] har jeg hende den ikke.  
       c. \* [<sub>VP</sub> Givet hende \_\_\_\_\_ ] har jeg \_\_\_\_\_ den ikke.  
       d. \* [<sub>VP</sub> Givet \_\_\_\_\_ den] har jeg hende ikke.

The pattern in (53) can be accounted for if in addition to order preservation, a constraint on structure preservation is considered to restrict OS (cf. Déprez 1994, Müller 2001, Sells 2001, and Williams 2003).

- (54)       STRUCTURE PRESERVATION (STRUCPRES):  
           If the foot of the chain of some non-adverbial element  $\alpha$  c-commands the foot of the chain of some element  $\beta$ , the head of the chain of  $\alpha$  also c-commands the head of the chain of  $\beta$ .

In other words, where ORDPRES says "preserve the sequence", STRUCPRES says "preserve the c-command relationships".

Like ORDPRES, the constraint STRUCPRES and its dominance over SHIFTPRON predicts that OS cannot cross an intervening non-adverbial element: For example, OS across a verb *in situ* as in (55)b changes the c-command relation between the verb and the shifted object.

- (55) Da a. Jeg spurgte hvorfor Peter aldrig læste den.  
           I asked why Peter never read it  
       b. \*Jeg spurgte hvorfor Peter den aldrig læste \_\_\_\_.

In contrast to ORDPRES, however, STRUCPRES (>> SHIFTPRON) rules out stranding of an object during VP-topicalisation. While the linear relations between the verb and the objects are maintained in (53)b,c above, their structural relations are not: The verb (and IO) in Spec,CP is too deeply embedded to c-command the stranded (IO and) DO. Consequently, STRUCPRES >> SHIFTPRON rules out stranding of an object during remnant VP-topicalisation while permitting topicalisation of a full VP.

**Tableau 18: No remnant VP-topicalisation**

Da	Topic: V	TOPIC	STRUCPRES	SHIFTPRON	ex.
☞ a	[ <sub>VP</sub> V <u>Pron-IO</u> <u>Pron-DO</u> ] Aux Sub Adv t <sub>VP</sub>			*!*	(53)a
b	[ <sub>VP</sub> V t <sub>IO</sub> t <sub>DO</sub> ] Aux Sub <u>Pron-IO</u> <u>Pron-DO</u> Adv t <sub>VP</sub>		*!*		(53)b
c	[ <sub>VP</sub> V <u>Pron-IO</u> t <sub>DO</sub> ] Aux Sub <u>Pron-DO</u> Adv t <sub>VP</sub>		*!*	*	(53)c
d	[ <sub>VP</sub> V t <sub>IO</sub> <u>Pron-DO</u> ] Aux Sub <u>Pron-IO</u> Adv t <sub>VP</sub>		*!*	*	(53)d

Hence, variation between speakers as to the strandability of objects during VP-topicalisation may be accounted for by a contrast in the ranking of two very similar constraints, one requiring order preservation, the other structure preservation.

## Appendix 2: Remnant VP-topicalisation in German and Dutch

As observed by Fox & Pesetsky (2005) for Swedish, there is also an asymmetry between stranding IO and stranding DO by remnant VP-topicalisation in

German. However, it goes in the opposite direction: Stranding of the IO is preferred over stranding of the DO though the contrast is not as sharp as in Swedish/Danish, (56)c being marginal but not ungrammatical. In other words, changing the base order of the objects as in (56)b is preferable to keeping the base order as in (56)c under remnant VP-topicalisation in German. Furthermore, note that remnant VP-topicalisation in an OV language necessarily reverses the order between the verb and the stranded object.

(56) Ge

- a. [VP Dem Mann das Buch gegeben] hat sie schon gestern.  
       the man    the book    given    has she    already yesterday
- b. [VP \_\_\_\_\_ Das Buch gegeben] hat sie dem Mann schon gestern.
- c. ??[VP Dem Mann \_\_\_\_\_ gegeben] hat sie das Buch schon gestern.
- d. [VP \_\_\_\_\_ \_\_\_\_\_ Gegeben] hat sie dem Mann das Buch schon gestern.

That German allows order reversal in remnant VP-topicalisations is not surprising. The fact that German scrambling of pronominal and non-pronominal elements may change the order of arguments as in (57) requires ORDPRES to be outranked by both SHIFTPRON and the more general constraint SHIFT (see footnote 7) which are taken to trigger scrambling in the continental West Germanic languages as well (see Vikner & Engels 2006). Consequently, it is expected that an argument may be stranded (satisfying SHIFT or SHIFTPRON) although stranding changes the base order precedence relations (violating ORDPRES). (Further research is needed concerning the marginal status of (56)c.)

(57) Ge

- a. Sie hat schon gestern dem Mann das Buch gegeben.  
 she has already yesterday the man the book given
- b. Sie hat das Buch schon gestern dem Mann \_\_\_\_\_ gegeben.
- c. Sie hat das Buch dem Mann schon gestern \_\_\_\_\_ gegeben.
- d. Sie hat es schon gestern dem Mann \_\_\_\_\_ gegeben.  
 she has it already yesterday the man given
- e. Sie hat es dem Mann schon gestern \_\_\_\_\_ gegeben.

In contrast to German, however, scrambling of full DPs has to maintain the order of arguments in Dutch, (58), indicating that the constraint ORDPRES outranks SHIFT.

- (58) Du a. Ze heeft gisteren de man het boek gegeven.  
 she has yesterday the man the book given
- b. Ze heeft de man gisteren \_\_\_\_\_ het boek gegeven.
- c. Ze heeft de man het boek gisteren \_\_\_\_\_ gegeben.
- d. \*Ze heeft \_\_\_\_\_ het boek gisteren de man \_\_\_\_\_ gegeben.
- e. \*Ze heeft het boek de man gisteren \_\_\_\_\_ gegeben.

However, as pointed out to us by Hans Broekhuis (p.c.), the ranking ORDPRES >> SHIFT predicts that remnant topicalisation is not possible at all in Dutch, contrary to fact. Although as an OV-language Dutch necessarily reverses the order of topicalised verb and stranded object in remnant VP-topicalisations, stranding of the IO and stranding of both IO and DO during remnant VP topicalisation is acceptable; stranding of the DO, in contrast, is ungrammatical (59).

- (59) Du a. [<sub>VP</sub> De man het boek gegeven] heeft ze gisteren al.  
 the man the book given has she yesterday already
- b. [<sub>VP</sub> \_\_\_\_\_ Het boek gegeven] heeft ze de man gisteren al.
- c. \*[<sub>VP</sub> De man \_\_\_\_\_ gegeven] heeft ze \_\_\_\_\_ het boek gisteren al.
- d. [<sub>VP</sub> \_\_\_\_\_ \_\_\_\_\_ Gegeven] heeft ze de man het boek gisteren al.  
 (Hans Broekhuis, p.c.)

It is interesting to note that the scrambling operation that precedes VP-topicalisation does not violate HG in (59)b, but only in (59)c (compare (58)b,c above). Maybe the fact that (59)b is grammatical even though it violates ORDPRES says something about ORDPRES being a repair strategy in case HG is violated.

Remember that in the Scandinavian languages, stranding of an object during remnant VP-topicalisation necessarily involves a violation of HG and it is only possible if the base order of elements is maintained, as required by ORDPRES (cf. sections 2.2 and 3.2, respectively). In other words, remnant VP-topicalisation may give rise to a repair effect in these languages, re-establishing the base order relations. In contrast, remnant VP-topicalisation does not restore the base order relations in OV-languages. The violation of HG in (58)c cannot be repaired by remnant VP-topicalisation, (59)c. However, the derivation of (59)b does not violate HG and, consequently, no repair strategy is needed. The fact that remnant VP-topicalisation reverses the order of elements would seem to be irrelevant.

## References

- Bobaljik, Jonathan. 2002. A-chains at the PF-Interface. Copies and Covert Movement. *Natural Language and Linguistic Theory* 20, 197-267.
- Broekhuis, Hans. 2000. Against feature strength: the case of Scandinavian Object Shift. *Natural Language and Linguistic Theory* 18, 673-721.
- Christensen, Ken Ramshøj. 2005. *Interfaces: Negation – Syntax – Brain*. PhD dissertation, University of Aarhus, Denmark.
- Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20: Essays in linguistics in honour of Sylvain Bromberger, Kenneth Hale & Samuel J. Keyser* (eds.), pp. 1–52. Cambridge, Massachusetts: MIT Press.
- Chomsky, Noam. 2000. Minimalist inquiries. In *Step by step: essays on minimalist syntax in honor of Howard Lasnik*, Roger Martin, David Michaels and Juan Uriagereka (eds.), 89-156. Cambridge, Mass.: MIT Press.

- Chomsky, Noam. 2001. Beyond explanatory adequacy. *MIT Working Papers in Linguistics* 20, 1-28.
- Déprez, Viviane. 1994. Parameters of Object Movement. In *Studies on Scrambling*, Norbert Corver and Henk van Riemsdijk (eds.), 101–152. Berlin: Mouton de Gruyter.
- Erteschik-Shir, Nomi. 2001. P-syntactic motivation for movement: imperfect alignment in Object Shift. *Working Papers in Scandinavian Syntax* 68, 49–73.
- Fox, Danny & Pesetsky, David. 2003. *Cyclic Linearization and the Typology of Movement*. Ms., MIT. (available at [http://web.mit.edu/linguistics/www/fox/July\\_19\\_handout.pdf](http://web.mit.edu/linguistics/www/fox/July_19_handout.pdf))
- Fox, Danny & Pesetsky, David. 2005. Cyclic Linearization of Syntactic Structure. *Theoretical Linguistics* 31, 1-45.
- Grimshaw, Jane & Samek-Lodovici, Vieri. 1995. Optimal Subjects. *University of Massachusetts Occasional Papers* 18, 589-606.
- Holmberg, Anders. 1986. *Word Order and Syntactic Features in the Scandinavian Languages and English*. Ph.D., University of Stockholm.
- Holmberg, Anders. 1997. The true nature of Holmberg's generalization. *NELS* 27, 203-217.
- Holmberg, Anders. 1999. Remarks on Holmberg's Generalization. *Studia Linguistica* 53, 1-39.
- Johnson, Kyle & Vikner, Sten. 1994. The Position of the Verb in Scandinavian Infinitives. *Working Papers in Scandinavian Syntax* 53, 61-84.
- Jónsson, Jóhannes Gísli. 1996. *Clausal Architecture and Case in Icelandic*. Ph.D., University of Massachusetts, Amherst.
- Josefsson, Gunlög. 2003. Four myths about Object Shift in Swedish – and the Truth ... In *Grammar in Focus: Festschrift for Christer Platzack 18 November 2003*, vol. 2, Lars-Olof Delsing, Cecilia Falk, Gunlög Josefsson, and Halldór Ármann Sigurðsson (eds.), pp. 199–207. Lund: Department of Scandinavian Languages, Lund University.
- Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19, 335-391.
- Müller, Gereon. 2001. Order Preservation, Parallel Movement, and the Emergence of the Unmarked. In *Optimality-Theoretic Syntax*, Géraldine Legendre, Jane Grimshaw, and Sten Vikner (eds.), pp. 279-313. Cambridge, Mass: MIT Press.
- Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, Mass.: MIT Press.
- Samek-Lodovici, Vieri. 1996. *Constraints on Subjects: An Optimality Theoretic Analysis*. Ph.D. Dissertation. Rutgers University, New Brunswick, NJ.



- Samek-Lodovici, Vieri. 1998. OT-interactions between Focus and Canonical Word Order: Deriving the Crosslinguistic Typology of Structural Contrastive Focus. Ms. University College London.
- Sells, Peter. 2001. *Structure, Alignment and Optimality in Swedish*. Stanford: CSLI Publications.
- Vikner, Sten. 2005. Object Shift. In *The Blackwell Companion to Syntax*, Henk van Riemsdijk & Martin Everaert (eds.), pp. 392-436. Oxford: Blackwell.
- Vikner, Sten & Engels, Eva. 2006. *The Germanic Languages and the SOV/SVO difference, Part VII: Object Shift and Scrambling – An Optimality Theoretic approach*. DGfS/GLOW Summer School: Micro- & Macrovariation, University of Stuttgart, Aug. 14 - Sept. 2, 2006. Ms., University of Aarhus. (available at <http://www.hum.au.dk/engelsk/engsv/papers/stuttgart/vikner-stgt7.pdf>)
- Vogel, Ralf. 2004. Weak Function Word Shift. To appear in *Linguistics*. (available at <http://roa.rutgers.edu> , paper #663.)
- Williams, Edwin. 2003. *Representation theory*. Cambridge, Mass.: MIT Press.

*Eva Engels / Sten Vikner*  
*Institut for Sprog, Litteratur & Kultur*  
*Aarhus Universitet*  
*DK-8000 Århus C, Denmark*

[eva.engels@hum.au.dk](mailto:eva.engels@hum.au.dk)  
[sten.vikner@hum.au.dk](mailto:sten.vikner@hum.au.dk)

<http://www.hum.au.dk/home-uk.htm>