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

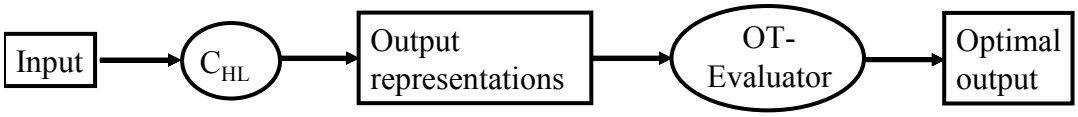
This paper describes and discusses the derivation-and-evaluation model in (1). The central idea underlying this model is that developing an explanatorily and descriptively adequate theory of syntax requires that restrictions be formulated both on the syntactic derivations and the resulting syntactic 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.

Linguistics in Potsdam 25 (2006): 137–193
Hans Broekhuis and Ralf vogel (eds.):
Optimality Theory and Minimalism: a Possible Convergence?

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.[1]

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

[1] 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.

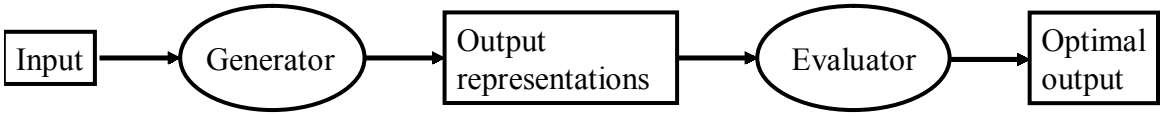
This section is organized as follows. Section 2 will substantiate the claim that MP and OT adopt essentially the same architecture of the grammar, and thus highlights the similarities between MP and OT. Section 3 discusses some differences in the research programs, and argues that these do not inherently 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 a 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

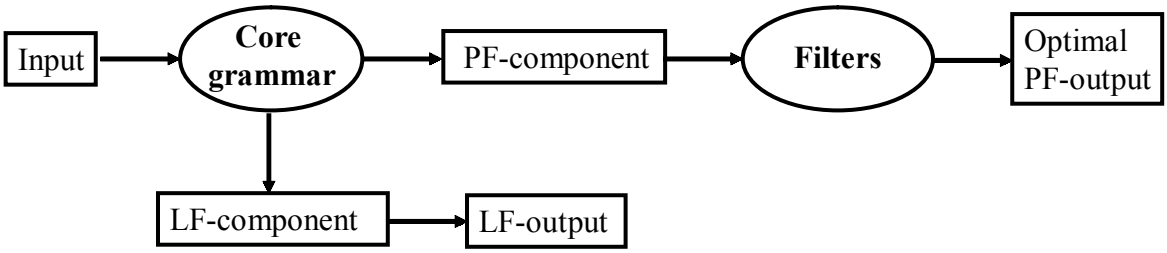
Since Chomsky and Lasnik (1977), the global organization of the different P&P models has had more or less the shape given in Figure 2 above, although in the earlier proposals this is masked by the fact that instead of a fully linear model, a so-called T- or inverse Y-model was assumed, according to which 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_{who}] (s-structure)

(3) a. the man [[$_{COMP}$ who ~~that~~] I know t_{who}]
b. the man [[$_{COMP}$ ~~who~~ that] I know t_{who}]
c. the man [[$_{COMP}$ ~~who that~~] I know t_{who}]
d. *the man [[$_{COMP}$ who that] I know t_{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.

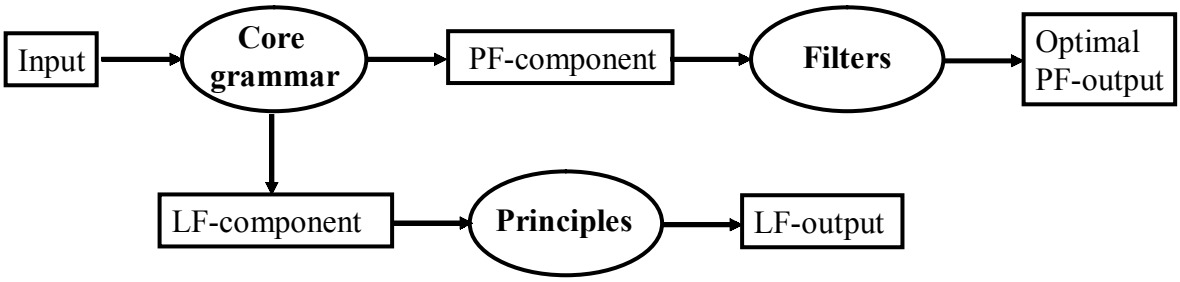
(4) a. I wonder [who ~~that~~ you met t_{who}]
b. *I wonder [~~who~~ that you met t_{who}]
c. *I wonder [~~who that~~ you met t_{who}]
d. *I wonder [who that you met t_{who}]

(5) a. the book [about which ~~that~~ he spoke $t_{about\ which}$]
b. *the book [~~about which~~ that he spoke $t_{about\ which}$]
c. *the book [~~about which that~~ he spoke $t_{about\ which}$]
d. *the book [about which that he spoke $t_{about\ which}$]

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 α. 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)



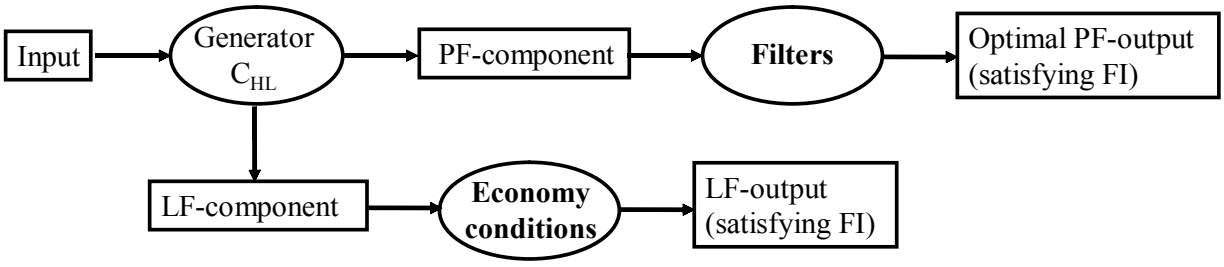
Although the ECP was claimed to be universal, that is, to be part of UG, its function is more or less the same as that of the *that*-trace filter: it excludes structures that have been created by core grammar. Therefore the formulation of the ECP is not a reason to frown with a skeptical eye on the notion 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

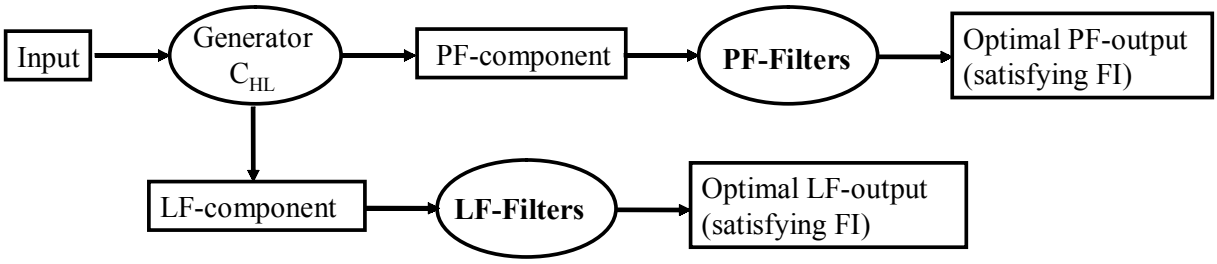
that the optimal output is the representation that satisfies a number of global economy conditions best: derivations with a smaller number of derivational steps are preferred (fewest steps), as are derivations with shorter movement chains (shortest steps).

> 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.

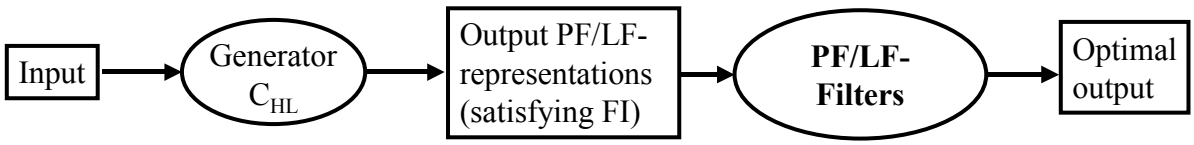
Another important innovation in Chomsky (1995:ch.4, 221) is the introduction of the bare output conditions, which are later normally referred to as the interface conditions. According to Chomsky, these interface conditions are "imposed from the outside" by the performance systems that make use of the 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* ⊂ focus
Jón bought not the book
b. Jón keypti bókina$_i$ ekki t_i *bókina* ⊂ 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_{bokina} —

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 θ-position of the object, and XP is a specifier position of *v** created by OS (note that Chomsky assumes a multiple specifier approach).

(9) ... [$_{\alpha}$ 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 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 [[$_{\text{COMP}}$ who ~~that~~] I know t_{who}]
b. the man [[$_{\text{COMP}}$ ~~who~~ that] I know t_{who}]
c. the man [[$_{\text{COMP}}$ ~~who that~~] I know t_{who}]
d. *the man [[$_{\text{COMP}}$ who that] I know t_{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$_i$ ~~que~~ je connais t_i]
b. l'homme [~~qui~~$_i$ que je connais t_i]
c. *l'homme [~~qui~~$_i$ ~~que~~ je connais t_i]
d. *l'homme [qui$_i$ que je connais t_i]

(16) a. l'homme [avec qui$_i$ ~~que~~ j'ai dansé t_i]
b. *l'homme [~~avec qui~~$_i$ que j'ai dansé t_i]
c. *l'homme [~~avec qui~~$_i$ ~~que~~ j'ai dansé t_i]
d. *l'homme [avec qui$_i$ que j'ai dansé t_i]

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) >> TEL.

Tableau 1: no complementizer deletion in embedded declarative clauses

	LE(CP)	TEL
.... [complementizer ....] ☞		*
.... [~~complementizer~~ ....]	*!	

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 >> 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 ....]	*!	
.... [~~complementizer~~ ....] ☞		*

Tableau 3 gives the evaluation of a language in which the two constraints are in a tie TEL <> LE(CP), which is indicated in the tableau by means of a dashed line. Under this ranking, the rankings LE(CP) >> TEL and TEL >> 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 ....]	☞		<*
.... [~~complementizer~~ ....]	☞	*>	

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 <> 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 >> LE(CP) >> TEL
b. English: REC >> TEL <> 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$_i$ ~~que~~ je connais t_i]		*!	
l'homme [~~qui~~$_i$ que je connais t_i] ☞			*
l'homme [~~qui~~$_i$ ~~que~~ je connais t_i]		*!	
l'homme [qui$_i$ que je connais t_i]		*!	*

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$_i$ ~~that~~ I know t_i] ☞		*>	
the man [~~who~~$_i$ that I know t_i] ☞			<*
the man [~~who~~$_i$ ~~that~~ I know t_i] ☞		*>	
the man [who$_i$ that I know t_i]		*>	<*

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) >> 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$_i$ ~~que~~ j'ai dansé t_i] ☞		*	
l'homme [~~avec qui~~$_i$ que j'ai dansé t_i]	*!		*
l'homme [~~avec qui~~$_i$ ~~que~~ j'ai dansé t_i]	*!	*	
l'homme [avec qui$_i$ que j'ai dansé t_i]		*	*!

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$_i$ ~~that~~ I know t_i] ☞		*	
the man [~~who~~$_i$ that I know t_i]	*!		<*
the man [~~who~~$_i$ ~~that~~ I know t_i]	*!	*>	
the man [who$_i$ that I know t_i]		*	*!

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,[2] as does Grimshaw's (1997) claim that all candidates in a certain candidate set have

[2] 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 the 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$_i$ ikke t_i?

(21) a. Hvorfor læste studenterne den$_i$ ikke t_i ?
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.[3]

(19) c′. At the phonological border of *v**P, XP is assigned INT′
(i) XP = DP (Icelandic)
(ii) XP = definite pronoun (Danish)
(iii) XP = ∅ (Romance)

What I want to stress here is that the adoption of language specific statements like (19c) or (19c′) is a radical breaks 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.

[3] 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 = ∅ (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

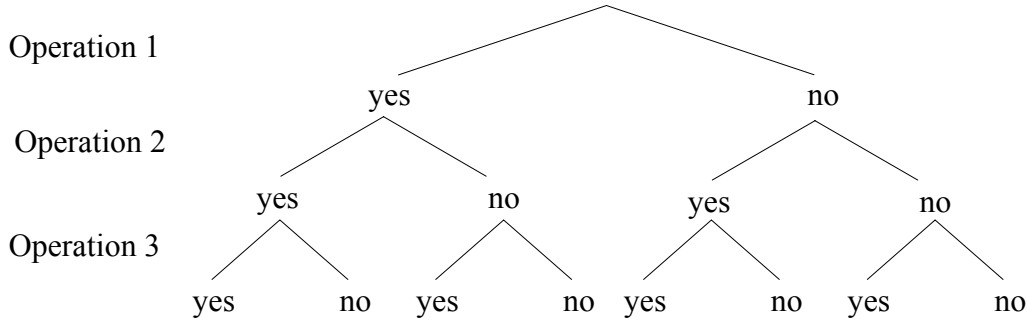
This section sketches the derivation-and-evaluation (D&E) model in Figure 1, in which the strengths of MP and OT are combined. The name of the model underlines the claim that the generator and the evaluator are equally important for providing descriptions and explanations of linguistic phenomena. The 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.[4] 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

[4] 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.[5]

[5] 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 (θ-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 (θ-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(φ) require movement of a DP into the specifier of a head containing case or φ-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(φ): unvalued φ -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 = definite pronoun (Danish)
(iii) XP = ∅ (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 >> *MOVE >> EPP(F):
if A favors movement, 'Procrastinate' is overruled.
b. B >> EPP(F) >> *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 *v*P-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 *v*P-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: *[$_{vP}$ ... pron$_{[+def]}$ ...].
b. Danish: D-PRONOUN >> *MOVE >> EPP(case)

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

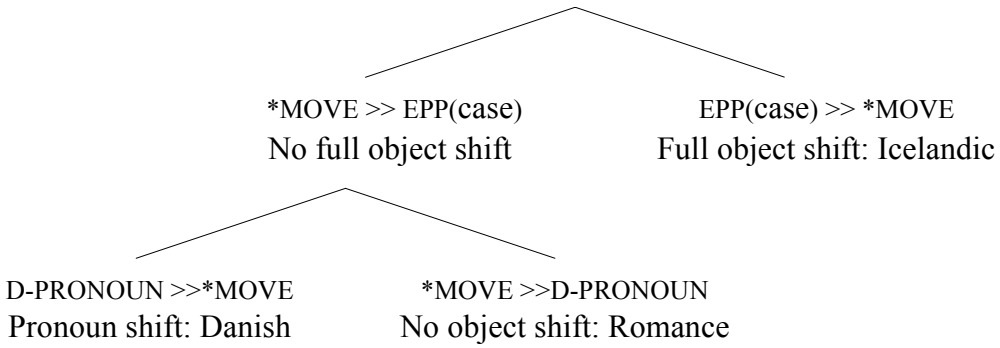
	D-PRONOUN	*MOVE	EPP(case)
Hvorfor læste studenterne ikke artiklen? ☞			*
Hvorfor læste studentene artiklen$_i$ ikke t_i		*!	

Tableau 9: Danish (pronoun shift)

	D-PRONOUN	*MOVE	EPP(case)
Hvorfor læste studenterne ikke den	*!		*
Hvorfor læste studenterne den$_i$ ikke t_i ☞		*	

The ranking D-PRONOUN >> *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-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.[6]
b. Icelandic: ALIGNFOCUS >> EPP(case) >> *MOVE

Tableau 10: Icelandic (object in not focus)

	AF	EPP(case)	*MOVE
Jón keypti ekki bókina		*!	
Jón keypti bókina$_i$ ekki t_i ☞			*

Tableau 11: Icelandic (object in focus)

	AF	EPP(case)	*MOVE
Jón keypti ekki bókina ☞		*	
Jón keypti bókina$_i$ ekki t_i	*!		*

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).

[6] 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.[7] 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.

[7] 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 <> ALIGNFOCUS >> EPP(case) >> *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$_i$ ekki t_i	*!	(*)		*

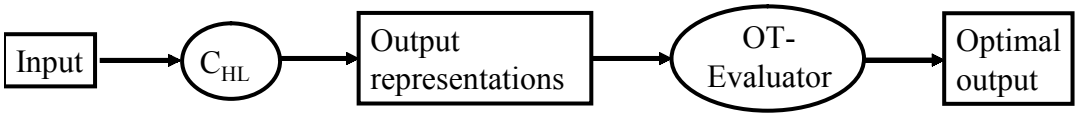
By introducing the Interface constraints D-PRONOUN, ALIGNFOCUS and HEAD-COMPL the present proposal accounts for the same range of facts as the set of statements in (31). There are, however, a number of reasons to prefer the present constraint approach to an approach that takes recourse to filters. First, 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.

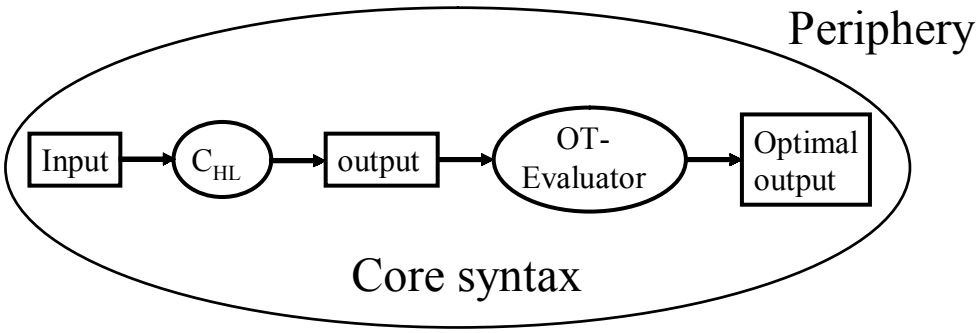
Derivation-and-evaluation model		MP	OT
I. The generator is some version of C_{HL}		+	—
(a)	all operations are subject to inviolable conditions	+	?
(b)	all operations are subject Last Resort	+	—
(c)	the generator is autonomous and operations apply at random; there are no EPP-features.	—	+
II. The evaluator 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/progressingrammar.

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: PhD.-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