Die Deutsche Bibliothek – CIP-Einheitsaufnahme

The **role of economy principles in linguistic theorie** /
Chris Wilder ... (ed.). – Berlin : Akad. Verl., 1996
(Studia grammatica ; 40)
ISBN 3-05-002897-1
NE: Wilder, Chris [Hrsg.]; GT

ISSN 0081-6469

Druck und Bindung: Druckhaus „Thomas Müntzer" GmbH, Bad Langensalza

Printed in the Federal Republic of Germany

Introduction

Chris Wilder and Hans-Martin Gärtner

Generative linguistics, it is fair to say, is dominated by developments in syntactic theory; and for the past few years, work in generative syntax has been heavily influenced by what has come to be known as 'the Minimalist Program' (MP). The seminal paper that ushered in the MP was Chomsky's "A minimalist program for linguistic theory" (1993); this was followed by two papers defining a 'second phase'—"Bare Phrase Structure" (Chomsky 1995a) and "Categories and Transformations" (Chapter 4 of Chomsky 1995b).[1] At the core of the MP lies the idea that economy is a central property of the system of language. This is fleshed out in terms of concrete principles of UG that instantiate the overarching economy idea.

Although economy principles have previously figured in explanations in phonology, morphology and lexicon theory (underspecification, elsewhere condition), as well as syntax, it is only with the MP that the role of economy has become a central topic on the linguistic agenda. This is especially true of syntax, but there is also convergence with developments in neighbouring subfields, as this volume illustrates. The papers collected here present differing views on the scope and nature of economy principles in syntax, morphology and the lexicon, offering extensions of and alternatives to the version outlined by Chomsky.

In the following, we set out basic dimensions of the economy concept, and draw attention to some antecedents to the economy principles of MP. We then plot some of the main features of the MP itself. In part II, we introduce the contributions to the volume, relating them to the larger context. (Throughout, we refer to papers in this volume by author's name(s) in small capitals.)

I Background

1 Economy: From Heuristic to Object of Study

In simple, informal terms, an economy principle can be expressed through an injunction like (1). This in turn can be decomposed into a 'last resort' principle, and a 'laziness' ('least effort') principle:

(1) 'Do only what needs to be done.'

(2) a. *Last Resort*: 'Do what must be done.'
b. *Least Effort*: 'Do not do what need not be done.'

(2a) differs from (2b): compliance with (2a) is necessary to attain some goal, while non-compliance with (2b) does not necessarily lead to failure. But under certain conditions—scarcity of resources, a large number of goals to be fulfilled—(2b) can join (2a) as necessary for success. Taken together, these constitute a general economy principle (in practice, Last Resort is usually understood as implying (2b), and Least Effort as implying (2a), so that the two end up describing the same coin by its different sides).

The usual role of (1)/(2) in science is as a heuristic guiding construction of theories commonly referred to as Occam's razor. But such maxims may have different—deeper?—import, reflecting principles underlying the design of the cognitive systems which the theories are about. Throughout, we need to distinguish these two aspects of the MP: [2]

(3) a. *Economy (I):*
A heuristic, a principle governing the way we go about constructing theories (explanations).

b. *Economy (II):*
A property of the object of study, i.e. a constitutive principle in the design of language (UG).

The importance of economy principles in explaining language *use*, and cognitive processes in general, has long been recognised. [3] Such principles have also made regular appearance in models of specific competence domains, cf. the Elsewhere principle in the lexicon, minimality and locality principles in syntax and phonology. The major innovation of the MP has been the elevation of Economy to a fundamental constitutive principle in the design of the language faculty as a whole. [4] This nontrivial step from Economy I to Economy II is much in line with Chomsky's scientific realism:

> A naturalistic approach to linguistic and mental aspects of the world seeks to construct intelligible explanatory theories, taking as "real" what we are led to posit in this quest . . . (1995c:1).

Technically, it is important to recognize the *comparison* (or competion) inherent in the economy concept. A general format for economy principles is a simple injunction: "Minimize X". Minimization is a process of comparison and selection, presupposing (i) a set of alternatives, comparable in terms of bearing some property X, and (ii) a scale that orders these in terms of amount or degree of X: a 'price list'. Comparison of the alternatives determines which is the minimal ('cheapest') one. This then forms the correct choice *by virtue of its properties in comparison with alternative candidates of its set*. In an alternative image, selection can be envisaged as the outcome of competition among alternatives, the 'winner' having scored the most (or least!) points on the scale.

This characterization leaves open parameters—the set of alternatives, the X of the price-list—that can be defined according to the nature of the system being described. It

lends itself to description of *optimization* effects in all kinds of complex systems. The explanatory value of such a description (assuming it is accurate) will depend largely on the plausibility and naturalness of assumptions regarding the sets of alternatives and the X of the price-list, on which appeals to economy feed. [5]

The relevance of economy to a (sub)system entails the presence of a minimal degree of *complexity*, in the sense that alternatives must exist, inside that system. The raison d'etre of an economy principle is selection among alternatives—minimally two. Suppose we have a system *S'* which is *simple*, in that it always generates a singleton output for each goal (no comparable alternatives). If we add an economy principle, *S'* only ever makes a singleton set available to its economy principle. Then of course, each output is a 'born winner' as far as economy is concerned, since it necessarily has less X than all the alternatives (there are none!) in its set. Economy is irrelevant to such a system—if it always applies vacuously, it is not possible to detect it.

Now in theory, *S'* could still exercise the choice between having that output and having none. The latter option would imply failure of the system with respect to the goal at hand. It is crucial that that option is denied to economy, as a matter of principle. If we allow economy to choose no output, then no output will be the universal outcome—economy will prevent the system from fulfilling its goals. To illustrate the point with Occam: a theory with *zero* axioms is no theory at all. [6]

As a heuristic guiding our construction of theories (3a), conceptual economy (Occam's razor) enjoins us: do not multiply your axioms beyond necessity. The theorist seeks to reduce his assumptions to the minimum necessary to account for a body of data. The set of alternatives are the various theories under consideration; the price-list is the scale defined by the number of axioms. Assuming that the theories fall into the same comparison class—i.e. to the extent that their goals are similar, that they share empirical coverage, that their axioms comparable, and so on—then the optimal theory will be the one deploying the least axioms. So much is familiar. So also are the difficulties of putting the principle into practice. How to establish the comparablility of theories? How do we count axioms? And so on. And beyond the practicalities lurks the conceptual issue: why should we seek elegance in our theories anyway? Why should the simplest theory be the correct one? [7]

The phenomena in a particular domain may demand theories in which the economy notion is significant in a different sense. The data may be such that axioms are needed which generate sets of alternatives in the sense described above. The theorist may then decide to include an economy principle of the form "Minimize X", operating over such sets, among his axioms, turning economy into a property of the object under study (Economy II). Of course, there will exist alternative theories which do not include any such principle. But suppose that such alternatives involve *more*, or more *complex* axioms, than the version including such a principle. Then Economy I comes into play—the theory including an economy principle among its axioms is preferred for reasons of conceptual economy.

Natural language—as already indicated above—is such a domain, it seems. Linguistic objects (words, sentences) involve combinations of primitives (features, etc.) that are organized in complex ways. Descriptive goals require the postulation of those primitives and the formulation of axioms (rules, well-formedness conditions) governing possible combinations. In some cases, it has long been claimed that correct

generalization requires the invocation of economy principles, for the sorts of reason just described. One such case is blocking in morphology, interpreted as a reflex of the Elsewhere condition by Kiparsky (1982). [8]

Pairs of V-stems and past tense forms like *move-moved, walk-walked,* motivate the postulation of a V-stem for each pair and an affix (*-ed* [+PAST]) common to all, along with a rule for combining the affix with V-stems. In its simplest formulation, the rule attaches the affix to any V-stem:

(4) /../ ; +V → /...+ed/ ; +V+PAST

A problem arises with irregular past forms in pairs like *go-went, swim-swam* etc—not that an extra [+past] stem (e.g. *went*) or additional past-formation rule (e.g. ablaut) must be assumed, but that 'regular' affixation must be prevented from applying to the V-stem in such pairs (**goed, *swimmed*). The solution lies in assuming that past-formation constitutes a domain in which various alternatives (*-ed*-affixation, ablaut, suppletive forms) are engaged in a competition governed by an economy principle, i.e. the Elsewhere condition.

Suppletive past forms like *went* can be analysed as a rule in the format (1):

(5) /go/ ; +V → /went/ ; +V+PAST

(4) and (5) compete to apply to incoming V-stems; the economy principle orders them on the scale of 'input specificity': a rule with a more specific input condition—i.e. (5)—will win out against a rule with less specific input (4). The output **goed* is blocked, since (5) takes precedence over (4), and gives an output for *go*. If however, the incoming stem does not match a rule's input condition, that rule is no longer in the competition (it can give no output)—e.g. there is no rule relating *walk* with *went*. In that case, the most specific of the remaining rules *which can apply* will win out. This economy principle is the Elsewhere condition, which turns the general rule into a default that applies only where no more specific rule can apply, thus explaining the blocking effect. [9]

It is hard to imagine a simpler alternative account of these facts. Without the Elsewhere condition, we either give up the affixation rule (thereby multiplying lexical entries = 'axioms'), or we have to state the rule in more complex fashion, i.e. with exceptions attached—equivalent to the addition of specific blocking 'axioms' for exceptions. Taking on board the way the Elsewhere account generalizes to numerous similar cases across languages, the argument from conceptual economy for adopting it becomes overwhelming.

Another case for which an economy-style account has been invoked, this time in the syntactic domain, concerns the paradigm (6):

(6) a. John$_j$ V_1 [PRO$_j$ to leave]
b. *John$_j$ V_2 Mary$_k$ [PRO$_j$ to leave]
c. John$_j$ V_2 Mary$_k$ [PRO$_k$ to leave]

The subject argument of the infinitive verb in (6a) is controlled by the subject of the superior verb (*try, hope, intend,* etc.). When the matrix verb takes an object (*persuade, force, ask,* etc.), however, control by the matrix subject is not possible: rather, the controller may only be the object. A control rule based on the pattern (6a)—"the subject of the higher clause controls the infinitive subject PRO"—is falsified by (6b), while a rule based on (6c)—"the object of the higher clause controls..." cannot apply generally. The nature of the problem is familiar: it is the *closest* potential controller that controls. There is object control where this is possible (6c), the possibility of object control blocking subject control in (6b), while there is subject control otherwise (6a). Rosenbaum (1967) proposed just such an account in terms of a 'Minimal Distance Principle'. [10]

Until recently, instantiations of economy in the syntactic theoretical apparatus were fairly isolated. [11] Before considering the emergence of the 'internal economy' as a pervasive factor in syntax in MP, it is useful to consider the role of Economy I (3a)—in rhetoric and in substance—in the the emergence of MP out of its predecessor "Government-Binding" (GB) model.

2 From GB to the Minimalist Program

A recurrent theme in linguistic theory is the tension between descriptive and explanatory adequacy that accompanies the search for the right model of grammar. The facts of natural languages necessitate complex descriptive mechanisms. Conceptual economy dictates that these be generalized and reduced as much as possible. [12] Alongside the incorporation of empirical findings on particular phenomena in individual languages, the elimination of redundancy in theoretical machinery has always been a driving force in the development of the theory.

The move away from construction-centred rule-systems in the late 1970's was accompanied by a shift of focus from rules that derive representations to principles that govern representations. The GB-system (Chomsky 1981) rested on the interaction of the 'principles' of the 'principles-and-parameters' label: Case Filter, Theta Criterion, Binding Conditions, ECP, etc.—essentially well-formedness-conditions on representations. This system attained a degree of deductive depth, especially in the Binding-Case-Theta-A'/A-nexus, which enabled it—extended through the late 1980s with the 'Barriers' model (Chomsky 1986a)—to serve as a research platform for a decade or more (indeed, it still does so).

Nevertheless, transformational grammar has been from its inception a derivational system. Despite its emphasis on 'principles', the derivation (Move-α) plays a central role in the GB-system, too. This is reflected in the 'T-model' architecture that held sway from the late 1970's up to Chomsky (1993), with its four levels of representation at which well-formedness conditions apply (i.e. D-structure, S-structure, Logical Form (LF), and Phonetic Form (PF)).

In GB, principles are formulated in very specific fashion. While concepts of 'minimality' and 'locality' figure in many places, principles such as Subjacency, ECP, and Case Filter appear highly specific—language-faculty-specific and, internally, module- and level-specific—with intricate and seemingly arbitrary internal structure,

not relatable to external concepts. A central role in defining principles was accorded to the government relation, a 'derived primitive' of some complexity, especially in its 'Barriers' incarnation. One aim (only partially realized) of the MP is to reduce this specificity to its 'bare essentials'. In the MP, Chomsky 'goes back to basics'. His approach to the minimal theory of grammar starts from "virtual conceptual necessities"—basic elements that (virtually) anyone would agree on, ones that are incontrovertibly forced on us by the facts.

Sentences, built from words, relate sound and meaning; that is, they are complex representations that are interpreted at LF and at PF—the two independent interfaces with extra-linguistic systems. Thus, we need (i) a lexicon, listing the atomic elements that may enter linguistic representations, and (ii) a way of combining these into (PF,LF) pairs. Furthermore, those representations must encode in some way (iii) the phrase-head distinction (i.e. we need X'-theory), and (iv) the fact that elements within those expressions may be displaced from the position where they are interpreted (i.e. we need movement/chains). Hence, the basic essentials are:

(7) a. units: lexical items (heads), phrases, chains
b. levels: PF and LF
c. a construction/assembly mechanism (forming phrase structures from lexical items)
d. movement (forming chains in trees)

'Minimalism', so the rhetoric goes, dictates that the 'perfect theory' will contain as little more than this as possible. In this sense, the emergence of the MP model was explicitly and fundamentally influenced by Economy I:

> . . . further questions arise, namely those of the Minimalist program. How "perfect" is language? One expects "imperfections" in . . . the lexicon and aspects of . . . the A-P interface. The essential question is whether, or to what extent, these components of the language faculty are the repository of departures from virtual necessity . . . Looking at the same problem from a different perspective, *we seek to determine just how far the evidence really carries us toward attributing specific structure to the language faculty*, requiring every departure from "perfection" to be closely analyzed and well-motivated. (Chomsky 1995b:9; our emphasis)

Taking the challenge seriously, the MP proceeds with wholesale revision of the architecture and core concepts of the GB-model. This extends to levels, principles and relations.

Firstly, if LF and PF are the only externally motivated levels, then DS and SS can be dispensed with, and so should be.[3] The adoption of Generalized Transformations as a tree-construction and lexical insertion mechanism (required anyway), allows D-Structure to be dispensed with. At the same time, the introduction of GT permits a partial unification of structure-building and movement, and a reinterpretation of the cycle as a consequence of the tree-building procedure (more on this in section 4 below).

Secondly, with D-structure eliminated, S-structure is reinterpreted as a 'stage in a derivation' (Spell-Out). This opens the program of reinterpreting other conditions defining SS and DS as independent levels. One aspect consists in showing that

principles such as the Case Filter or Binding conditions can (if possible: must) apply at LF, rather than SS. Another concerns reinterpretation of 'S-structure parameters' governing overt vs. covert application of Move-α (in the sense of Huang's 1982 proposals on *wh*-movement) via conditions on PF and LF.

A third key task is to eliminate government as core relation, replacing it with a locality relation constructed in terms of X'-theoretic relations, independently given (cf. Chomsky's 'Minimal Domain'). Then it becomes necessary to restate Case Theory, Binding Theory, θ-theory, ECP etc—core concepts in the GB-model that relied on government—using these bare minima. The terms for explanations are limited to (i) properties of lexical items themselves, (ii) conditions on LF (or PF) (iii) X'-theoretic locality, (iv) the workings of derivational operations: GT and Move-alpha.

Crucially, MP puts forward the hypothesis that economy-based principles play a major role in the computational system of UG. Principles of the form "minimize X" constrain representations, and the operations that build them. In this sense, we move away from Economy I—economy becomes a property of the object of study, the design of language itself.

It is instructive to see how much of the MP—including economy—was already implicit in the syntactic analyses of the GB framework. Take for example the analysis of *wh*-movement in simple and multiple interrogatives. Adapting May's (1985) or Rizzi's (1991) WH-criterion, assume the level-specific principles in (8), coupled with lexical specifications (9):

(8) a. [+WH] in C must match with [+WH] in an XP in SPEC,CP (at SS)
b. [+WH] in XP must match with [+WH] in C (at LF)

(9) a. C: [+WH] ∅ / [−WH] *that*, etc.
b. D: [+WH] *which*, *who*, *what*, etc. / [−WH] *this*, *the*, *he/she/it*, etc.

These suffice to account for (10): (8a) ensures that movement of *what* applies pre-S-structure in (10a), (8b) then being satisfied at LF. The SS-representation determines the given order of elements in PF, the LF provides a basis for interpretation as indicated.

(10) a. what did John say that he saw
b. 'for which thing *x* , John said he saw *x*'

Assuming that LF-*wh*-raising adjoins a wh-phrase to Spec,CP in case Spec,CP is already occupied, and that adjunction to Spec is sufficient to satisfy (8b), (8)-(9) also nearly suffice to ensure the derivation of (11).

(11) a. who said that he saw what
b. 'for which thing *x*, for which person *y*, *y* said he saw *x*'

The appearance of two *wh*-phrases in (11) however makes available an alternative derivation satisfying (8), in which *what* raises to satisfy (8a) at S-structure, with *who* adjoining to *what* covertly, giving **what did who say that he saw*. To exclude this

derivation, various accounts have been offered (cf. Chomsky's 1973 Superiority Condition, Pesetsky's 1982 ECP-account).

Further, cross-linguistic variation is accommodated by parametrization of the principles (8). Given a division into English-type languages (one *wh*-phrase must raise before SS) and Japanese-type languages (all *wh*-phrases remain in situ at SS), the latter type can be described simply by relocating (8a) at LF in the grammars in question. In sum, an elegant analysis for a complex array of facts. [14]

The movement parameter rests on a principle (8a) that can alternatively apply at SS or LF—an option not available in MP. Chomsky (1993) reinterprets this account in terms of the strong-weak feature distinction (interface-relevant properties of lexical items), interacting with an economy principle ('Procrastinate') that selects among potential derivations.

As well-formedness conditions on representations, matching requirements (8) are interpreted as a reflex of representational economy holding at the interfaces (LF/PF), requiring the removal of formal (later: 'uninterpretable') features introduced into derivations by lexical items (9). The PF-filter is 'turned on' only by the presence of a feature lexically specified as 'strong'. The effect of the SS-parameter is thus obtained by attributing a property 'strength' to features of lexical items (here: the WH-feature of C in English) to which PF-well-formedness is sensitive. [15] This aspect of the reorganisation does little more than restate the GB analysis: an SS filter is removed, by declaring it to hold at PF. [16]

The role of economy in the account is more substantial. (12) represents a 'Japanese' derivation, (13) an 'English' derivation. The filters (8) ensure the LF is the same; the derivation (12) is excluded for English by the SS-Filter. However, while (12) is permitted for Japanese, (13) is not yet excluded.

(12) a. SS: C John said that he saw what
b. LF: what C John said that he saw

(13) a. SS: what C John said that he saw
b. LF: what C John said that he saw

Similarly, nothing that has been said so far prevents (11) from surfacing as **what who said that he saw*. The logic of freeing move-α from level-specific constraints demands an answer to such questions; stipulating that *wh*-movement in Japanese must be delayed is unsatisfactory. Chomsky (1993) proposes that the computational system is 'lazy', it acts to satisfy filters 'as late as possible'. Early (pre-SS) movement generally incurs 'extra cost', which permits the derivations to be ordered. The intuition is expressed as the *Procrastinate* Principle; in the unmarked case, derivation (12a) is cheaper than (13a) (preferred by Procrastinate). In the presence of 'strong' [+WH] in C, however, (12a) leads to 'crash' at PF, leaving (13a) as the cheapest 'convergent' derivation.

The logic of the approach is now familiar—descriptive contingencies indicate alternative derivations; an economy principle selects among costed alternatives, such that cheaper derivations block more expensive derivations, the latter being permitted only when a cheaper derivation is not available.

Chomsky (1993) points out that there is also an apparently simple 'economy' account for superiority effects, in terms of the 'minimal' legitimate derivation blocking all others. The derivation (11) in which (8a) (i.e. strong [+WH] in C) is satisfied by *who* rather than *what*, is more minimal than the alternative in the sense that the path traversed by *who* on raising to C is shorter than the path traversed by *what*. This then becomes a *shortest path* condition on movement. The implementation of this idea, discussed in STERNEFELD's paper, is by no means trivial. HAIDER offers evidence from German that the idea may be simply false.

In the GB-model, derivational mechanisms (e.g. move-α) apply freely to generate a linguistic representation (a quadruple {D,S,P,L}) which independently satisfies level-specific constraints. Research focussed on well-formedness condition on representations, that conspire to constrain movement; but the problem of overgeneration attendant on the conception of 'free rule application' remained.

In the MP, overgeneration is curbed by the elevation of economy ('Last Resort') to an all-constraining principle. Where an operation is not motivated ('triggered'), it is not licensed. The paradigm case illustrating Last Resort is Case-driven NP-movement. In GB-terms, the need for a representation to satisfy the Case Filter at the level of S-Structure acts as a trigger for movement of a NP in a Case-less position into a position where it can be assigned a Case Feature:

(14) a. John was told *t* that it was raining
b. *It was told John that IP

The Last Resort nature of NP-movement is shown by comparison with (15).

(15) a. It seemed to John that it was raining
b. *John seemed to *t* that IP

Even though the subject position is a legitimate target for NP-movement, and even though English permits preposition-stranding under NP-movement, movement of *John* in (15b) is blocked. The NP satisfies Case requirements in its base position (15a), so movement is unnecessary; economy ensures that the derivation leading to (15b) is ungrammatical. [17]

Chomsky (1993) points out that the paradigm motivates a strengthening of Last Resort. A structure like (16) (= (15b) before movement) is ungrammatical as it stands—the clause needs an overt subject (cf. the Extended Projection Principle of the GB model), which is attributed to the need for the (nominative) Case feature in Infl to match with an NP in Spec,IP.

(16) (*) seemed to John that it was raining

Movement of *John* to the subject would be one possibility for saving (16) from ungrammaticality, and in this sense, would not be a gratuitous operation. But movement is blocked—the NP does not move since it satisfies its own Case requirement in situ. It appears that move-α cannot apply where α does not need to move to satisfy its *own*

requirements. This 'self-serving' property of Last Resort ('Greed'), is held to constrain all applications of move-α:

(17) *Greed*: Move-α must result in satisfaction of requirements of α

Adopting (17) immediately restricts the class of permissible derivations; in particular, it reduces the role of true derivational economy (comparison of possible derivations) in calculating the 'optimal' one. (However, in Chomsky (1995a, 1995b), Greed is revised—as discussed also by FRAMPTON).

This conception also means that optional (unmotivated) movement is not possible. Optimal derivations are 'pre-determined' by the interaction of lexical properties (triggers) with output conditions (FI) and derivational economy principles. For NP-movement, *wh*-movement and V-movement, the move seems successful. But in face of apparent 'optional movement' constructions, such as Scrambling (also other 'stylistic rules' such as extraposition, or the negation data discussed by HAIDER), sticking to this view might involve unacceptable cost in terms of assumptions about lexical properties (triggers). In this volume, 'optional' word order alternations are variously treated as movement triggered by optional morphosyntactic features (GREWENDORF & SABEL, MÜLLER, and STERNEFELD), and as base-generated orders in free variation (HAIDER, REINHART). Reinhart suggests a different role for economy, that of choosing among the outputs at the PF-interface.

We return to economy in syntax and competing derivations in section 6 below.

3 Lexicon, Morphology, and Syntax

The syntactic theory of MP places a heavy burden of explanation on the lexicon—specifically, on properties of lexical items. This reliance on the lexicon has three interlocking sources in MP. Firstly, syntactic operations respond to properties of lexical items. There are no construction-specific rules; only (morphological properties of) lexical items (closed class functional items) can be the source of specific syntactic patterns. Secondly, the hypothesised crosslinguistic invariance of syntax (computational system) means that much of the work in accounting for cross-linguistic syntactic diversity is delegated to the lexicon. Properties of lexical items interact with invariant syntactic principles to produce different syntactic patterns.[18] (The rest of the work must be done by phonology, a second locus of cross-linguistic variation). The third source is the hypothesis of *inclusiveness* (Chomsky 1995b:228), according to which syntactic objects created during the derivation, and in particular interface representations, contain nothing more than the lexical material, arranged in certain ways by the computational system:

> A "perfect language" should meet the condition of inclusiveness: any structure formed by the computation . . . is constituted of elements already present in the lexical items selected . . . no new objects are added in the course of computation apart from rearrangements of lexical properties . . .

I.e. syntax can manipulate (concatenate, copy, delete) lexical items, but may not add new items or features. Phonology forms the major exception to inclusiveness: the phonological derivation must have the capacity to add features and structure (e.g. metrical structure) to lexical material, if any version of underspecification is correct.

Full Interpretation (FI) dictates that interface representations only contain items interpretable there. Given inclusiveness, all such elements must have entered from the lexicon (or have been created from lexical input). While features contained in interface representations (particularly LF) come from the lexicon, not all features from the lexicon are present in the interface. Phonological features are 'stripped' at spell-out; morphological features trigger movement that cause their own elimination. What is left is 'interpretable' at LF.

However, MP has little to say about the structure and organization of the lexicon itself. Lexical items are viewed as simple sets (of sets) of elements of the usual sort: i.e. phonological, morphological, syntactic and semantic features. Except for 'strength', no new lexical properties or entities are assumed for purely syntactic reasons; no additions are made to the special 'syntactic' features (e.g. the [pronominal] and [anaphoric] features dedicated to Binding theory) of the GB system.

On the other hand, neither are there substantive constraints, beyond a general 'minimalist' prerogative, on what one may assume. The logic of Minimalist explanation encourages the addition of new features or properties to the inventory—if there is a movement, there must be a feature to trigger it; likewise, if movement has a certain property, perhaps this is a property of the trigger. This tendency is illustrated in this volume by GREWENDORF & SABEL's Agr-features, and MÜLLER's distinction between optional vs. obligatory features—both are attempts to come to grips with complex movement patterns within the confines of minimalist syntax. Yet the game with features is by no means restricted to syntax: in his analysis of complex inflectional patterns in Georgian and Potawatomi, WUNDERLICH relies on minimal morphological operations, but appeals to unorthodox 'complex' features that denote relative positions in theta- and animacy hierarchies. Rather, it is fair to say that descriptive necessity is the mother of invention here.

Beyond issues of the interaction of lexical items with syntax, the question of the role of economy has independent significance in lexicon theory. Economy has precursors in morphology and lexicon theory; concepts that originated from attempts to clarify the principles governing the format of lexical representations—underspecification, markedness—have a clear affinity to syntactic economy principles and to distinctions (e.g. 'strong' vs. 'weak' features) they respond to. The Elsewhere condition and the blocking effects it accounts for are clearly related to principles like Procrastinate and Shortest Move involved in ordering and selecting derivations in the syntax in MP.

Underspecification forms a major component in hypotheses about the format of lexical entries. The assumption that lexical entries provide only information about marked feature values, with unmarked values being filled in 'later' by (default) rule application—is central in phonology. Similar ideas have been applied to morphological, syntactic (categorization and subcategorization), and semantic information in lexical entries (cf. BIERWISCH, this volume). The choice of underspecification as a descriptive strategy reflects conceptual parsimony (Economy I)—the theorist chooses the minimal format for lexical entries, in the attempt to derive predictable information and

generalizations by maximal appeal to general rules. To the extent that the strategy is successful, it appears that underspecification is a real factor in the organization of lexical knowledge.

Taken as a property of the domain of study, underspecification may then be viewed as reflex of a substantive economy principle governing possible format of lexical entries (Economy II). Thinking of the role of the lexicon as the repository of 'permanent knowledge', underspecification may be explained as a reflex of 'representational economy'. The effect of ensuring maximally underspecified representations is to minimize space needed to store them. Hence, lexical representations may be subject to a specific economy principle (18): [19]

(18) Minimize storage.

A second feature of the organization of lexical information are the blocking effects in derivational and inflectional morphology, which Kiparsky (1982) attributes to the Elsewhere condition ('if a specific form is listed, don't create new ones by regular means'). Viewing word-formation processes as operations (taking a 'derivational view'), the Elsewhere condition can be interpreted as a reflex of a general constraint (19):

(19) Minimize computation.

Both these ideas are implemented in WUNDERLICH's paper.

To the extent that the system of lexical knowledge displays properties not reducible to principles underlying the syntax, the question arises as to the source of such differences. Both the derivational and representational economy principles governing syntax may be exhausted by (19)—assuming 'interpretation' (FI) can be read as 'interface computation'. The lexicon differs from syntax (the 'computational system') in its function as a permanent repository of linguistic knowledge. When applied to the lexical domain, the economy requirement may yield a distinct principle (18). It is tempting to speculate that special principles governing the organization of lexical knowledge related to or even reducible to (18).

There are further questions concerning lexicon-syntax relations which complicate the picture. Some spring from a basic tension between various notions of what the lexicon is. Minimally, the lexicon is simply a store of irreducibly idiosyncratic information: it lists arbitrary associations of form and meaning, plus non-predictable information about those forms and those meanings. All predictable properties of words are then to be derived by rule (or are governed by principles), belonging to 'syntax' in a wide sense—including phonological and interpretive rules/principles, as well as those governing sentence-construction ('syntax proper'). On the other hand, the elements stored in the lexicon are treated as syntactic atoms; the lexicon precedes sentence grammar, i.e. it feeds the syntactic derivation. Serious questions arise as to what properties such syntactic atoms may and must have, beyond the minimal, purely idiosyncratic form-meaning specifications. Words are more than unstructured bundles of features; and their internal structure goes beyond the intrinsic structure that emerges from the onotology of feature-types (phonemes are sequentially ordered; words may be morphologically

complex; predicates have argument structure). To the extent that those properties are non-idiosyncratic, it becomes necessary to assume rules / principles governing the lexicon. Questions then arise concerning the nature of those principles—are they reducible to universal principles that govern 'sentence grammar', are there special universal principles dedicated to the lexicon? These issues are taken up by BIERWISCH.

Related questions concern the status and place of morphology (word-formation and inflectional morphology). Is there a morphological component independent of the lexicon? Does morphology happen before, after, or even during syntax? In the MP, inflection is assumed to happen prior to syntax—in the lexicon (or in the process of forming the 'numeration'—see section 6 below) prior to derivational operations—while room is left both for syntactic word-formation (head-movement in the sense of Baker 1988) and for post-syntactic morphology in the phonological component. The latter is developed by Halle and Marantz (1993) into a model in which the determination of word-forms takes place in an autonomous Morphology component after syntax. In this volume, WUNDERLICH adopts an opposite, 'lexicalist' stance in which all of inflectional morphology precedes syntax, and takes issue with some of Halle & Marantz's assumptions. One issue in this debate concerns the account of syncretisms. These are handled by Halle & Marantz as the effect of 'impoverishment rules', which reduce the information passed to Morphology by syntactically fully specified trees. For Wunderlich, syncretisms reflect feature underspecification in word-forms, which is then inherited by the syntactic structures containing them.

4 Dynamic Conceptions of Phrase Structure

As discussed above, the MP reorganizes the base component of the grammar.[20] D-structure being eliminated, the base rules—called phrase structure rules in earlier models, later generalized as the X-bar format—are left dangling in the air, so to speak. It has become superfluous to generate D-structure as an output that feeds lexical insertion [21] on the one hand and transformational rules which derive surface structure on the other. A different perspective emerges instead. Lexical items can be taken to make up the input of a generalized syntax—the computational system 'C_{HL}'—that maps a set ('array') of such items into interface representations directly.

Factually, this development took place in two phases. The first consists in postulating three operations, *Select*, *Project*, and *Generalized Transformation* (GT). Select takes an item X from the lexicon. Project moulds X according to X-bar structure, that is, X becomes any of (20a-c):[22]

(20) a. X° b. $[_{X'}\ X^{\circ}\ ]$ c. $[_{XP}\ [_{X'}\ X^{\circ}\ ]\ ]$

GT is a more complex operation that targets a structure K, adds the dummy symbol ∅, substitutes a structure K' (K' ≠ K) for ∅, and forms K*, K* subject to X-bar theory. An example is given in (21):

(21) a.	Target K (= X°)	X°
b.	Add ∅	X°∅
c.	Substitute K' (= YP) for ∅	X°YP
d.	Form K*	[X' X° YP]

Given that (21d) seems to employ the operation Project, the relation between GT and Project seems to be amenable to further refinement. The role of Project is taken to be the preservation of certain uniformity properties, such as the requirement that elements in specifier and complement position be of XP status even if made up of a single terminal. We return to some revision of these issues below.

There is a simple argument that something like GT would have been necessary in pre-minimalist syntax to account for the facts under (22) (Lasnik & Uriagereka 1988:147; Frank&Kroch 1995):

(22) a. It is easy to please John.
b. John is easy [$_{CP}$ OP_i [$_{IP}$ PRO to please t_i]].
c. [That economy was difficult to control] is easy to see.

(22a) shows that the matrix predicate assigns only one θ-role. Thus, the NP *John* in (22b) does not occupy a θ-position and can therefore not have been present at D-structure—(22b) corresponding to the standardly assumed analysis of tough-movement constructions at S-structure. To assume that *John* has been inserted transformationally further implies that even complex structures, illustrated in (22c), must be allowed to be generated in parallel. Hence, an operation that combines such structures—traditionally called a generalized transformation—is required.

In a second phase of refining the base-component of UG, the X-bar format is fully eliminated as an independent mediator between lexicon and syntax. Lexical items now belong to the class of syntactic objects which are directly accessible to a modified form of GT called *Merge*. Merge is a binary operation that takes two syntactic objects, α and β, and creates a labeled set from them:

(23) Merge (α, β) → {α, {α, β } }

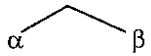
Labelling is achieved by taking the set of the constituent syntactic objects to be a member of another set, which contains the label as an additional element. The label is a copy of (the head of) α or β, α in (23) for the sake of concreteness. If α is the label, it can be said that the entire structure is a projection of α. Thus, Merge incorporates Project, a development foreshadowed by step (23d) of GT above.

In order to render the objects created by Merge commensurable with traditional notions of syntactic tree structures Chomsky defines *terms*, corresponding to nodes (1995b:247):

(24) For any structure K
a. K is a term of K.
b. If L is a term of K, then members of the members of L are terms of K.

Accordingly, { α, { α, β } }, the product of merging α and β, consists of three terms: (i) { α, { α, β } }, (ii) α, and (iii) β. These could be mapped into a standard tree, as in (25): [23]

(25) {α, {α, β } }



Crucially, labels do not constitute terms. Bar-level status is no longer considered an inherent property of syntactic objects but is computed contextually:

> A category that does not project any further is a maximal projection XP, and one that is not a projection at all is a minimal projection X^{min}, any other is an X' . . . (ibid.:242)

This definition puts an end to questions such as whether a category needs a specifier in order for it to count as maximal, issues that led to sometimes quite artificial debate. At the same time, generalizations that concern X-bar status can still be expressed.

More generally, the development sketched above is driven by two forces. The feature bundles called lexical items are elevated to central agents in syntax, allowing a greater focus on features per se and acknowledging the priority of this substantive core of language (cf. section 3 above).

> I . . . will therefore keep to the minimalist assumption that phrase structure representation is "bare", excluding anything beyond lexical features and objects constructed from them. (ibid.:245)

Another guideline for modeling the computational system is stressed by FRAMPTON, who observes:

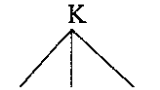
> We are certainly not at the point where we can pretend to be modeling mental computations at the level of algorithm. But feature-driven syntax may offer the possibility of taking some steps in this direction. [24]

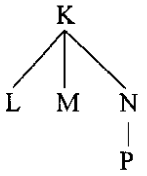
Given the central role of Merge, it can be asked whether there are any consequences for core notions of syntax like the concept of constituent, and the dominance, precedence, and c-command relations.

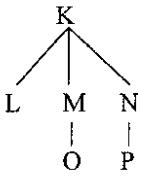
In fact, Kayne (1994) showed how to derive the precedence relation from (asymmetric) c-command with the help of a *Linear Correspondence Axiom* (LCA), together with the requirement that terminals be totally ordered: [25]

(26) *LCA*
If a node α asymmetrically c-commands a node β, then everything reflexively dominated by α precedes everything reflexively dominated by β.

It is possible to derive a large part of X-bar theory from the LCA. Thus it follows that syntactic structures are at most binary branching, i.e. the number of nodes immediately dominated by another cannot exceed two. Take the structures in (27):

(27) a. 

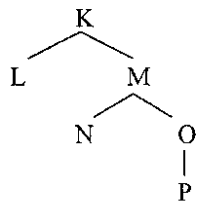
b. 

c. 

In (27a), trivially, L, M and N cannot be linearly ordered because none asymmetrically c-commands any of the others. Adding P in (27b) makes L and M asymmetrically c-command and thus precede P; but L and M remain unordered with respect to each other. Now, adding O in (27c), although it yields a precedence ordering for L and O, undoes the ordering between M and P, because N now asymmetrically c-commands O, so that P is required to precede O, which results in a contradiction. Obviously, adding further elements will not remedy the situation. Binary branching is the limit for syntactic structures, a fact that is axiomatic in the theory of Merge.

Sidestepping a number of technical issues about specifiers and intermediate projections, we may wonder to what extent the LCA is compatible with Chomsky's minimalist enterprise. The most natural interpretation seems to be to regard LCA as a PF-principle. (Chomsky 1995b:334ff). Under that view, the sister relations created by Merge are unordered; precedence only comes to be defined by an operation governed by LCA in the PF-wing of the grammar.

The LCA is possibly a candidate for interpretation as an economy principle. There are 2^n choices for linearizing the terminals of a binary tree with n branching nodes (without crossing the branches); the terminals of (28a) have the four possibilities shown:

(28) a. 

b. L-N-P , L-P-N , N-P-L , P-N-L

Regarding these as 'competitors', LCA serves to distinguish one as correct. The question that arises is—why is linearization based on *this* correspondence (c-command

maps to precedence), and not some other? Is there a natural interpretation of the LCA-mapping as defining the most 'economical' route from a hierarchical arrangement to a linear sequence of terminals? URIAGEREKA suggests such an interpretation, under which the LCA falls out as a consequence of more basic assumptions.

Further reduction has been envisaged by Epstein (1995), who observed that c-command itself is closely related to the operation Merge. Given that a node α c-commands a node β iff β is the sister node of α or dominated by the sister of α, and given further that sisterhood is what results from Merge, c-command can be conceived of as dependent on Merge. Although a precise formulation of this insight is far from trivial and in spite of the fact that head movement adjunction does not easily yield itself to Epstein's analysis—the c-comand domain of X° adjoining to Y° is standardly assumed to comprise Y' and whatever Y' dominates although X° merges only with Y°, not Y'—Epstein's theory highlights another important current in the MP: a *dynamic* view on syntax.

One might ask why syntax does not make use of just any relation definable over fully-fledged tree structures. It appears to be the case that somehow the *window* of syntactic processes is smaller. Relations that are not established dynamically, that is, by the operations that determine the course of a derivation, play no role. [26]

A dynamic way of handling syntactic information had already served as an empirical motivation for adopting GT during the first developmental phase under consideration here. Lebeaux (1988) used GT to derive the contrast in (29).

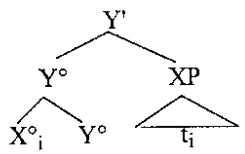
(29) a. * $[_{CP1}\ [_{DP}$ Which $[_{NP}$ claim $[_{CP2}$ that John$_i$ was asleep $]]]_k$
was he$_i$ willing to discuss t_k] ?

b. $[_{CP1}\ [_{DP}$ Which $[_{NP}\ [_{NP}$ claim] $[_{CP2}$ that John$_i$ made $]]]_k$
was he$_i$ willing to discuss t_k] ?

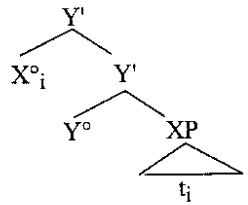
It was stipulated that adjuncts, unlike arguments, could be inserted into the structure after movement has taken place. This presupposes again that syntactic structures can be generated in parallel and that GT is available. Consequently, t_k does not contain a copy of CP2 in (29b)—as opposed to (29a)—and no principle C violation can arise.[27]

Chomsky (1993) captured these facts by exempting adjunction from the general requirement on pre-Spell Out operations to extend their target. Thus, K' in (21) above has to become the sister of the root node of K, a new root immediately dominating both K and K' as a result. Likewise, the transformation Move is subject to this so called "extension" requirement on overt operations, which therefore follow the principle of strict cyclicity. Again adjunction is the exception for obvious reasons. X°-movement being adjunction to a c-commanding head Y° cannot target the root node of the structure as illustrated in (30).

(30) a.



b. *



Indeed, cyclicity is an important ingredient of locality requirements on movement, and finds its way into the discussions by HAIDER, MÜLLER, STERNEFELD and URIAGEREKA.

Kitahara (1995) offers a mechanism sensitive to a fewest steps metric to derive the cyclicity property from X-bar theoretic distinctions not available to Chomsky's conception of bare phrase structure. This is discussed by HAIDER, who adds some critical remarks on the concept of 'cyclicity' with respect to the architecture of UG.

It has to be mentioned here that the account of (29) appears to be lost under the second phase of reforming phrase structure:

> With regard to Merge, there is nothing to say. It satisfies the extension condition for elementary reasons already discussed. (Chomsky 1995b:328)

The account of (29), however, requires Merge to target a non-root node, namely, NP in (29b), which is already dominated by DP and CP1 at that point in the derivation, i.e. after *wh*-movement has taken place. So, unsurprisingly, no mention is made of the data in (29).

5 Adjunction

The issue of adjunction stands out as defying full integration with minimalist principles. It is an open question whether the cost incurred by the technical complications which adjunction necessitates might not actually make it desirable to dispense with adjunction altogether. Technically, adjoining β to α gives rise to a segment-category distinction.

(31) Merge (α, β) → {<α,α> {α, β } }

Although α projects, the result is not a new category—instead, a two segment category is formed. The special status of an adjunction projection of α, called the upper segment, is indicated by its label <α,α>. The definition of terms in (24) applies equally to (31), yielding three terms: (i) { <α,α> { α, β } }, (ii) α, and (iii) β. These could in principle also be interpreted as nodes (but cf. Chomsky 1995b:247). A category must then be allowed to comprise a set of terms, containing { <α,α> { α, β } } and α, in our example. It remains to be seen whether allowing sets of terms to be functioning elements lies beyond what Chomsky calls

> . . . a fairly disciplined minimalist approach. Thus, with sufficiently rich formal devices (say, set theory), counterparts to any object (nodes, bars, indices, etc.) can readily be constructed from features. There is no essential difference, then, between admitting new kinds of objects and allowing richer use of formal devices; we assume that these (basically equivalent) options are permitted only when forced by empirical properties of language. (ibid.:381, fn. 7)

Definitions have to be formulated carefully. The relational definition of X-bar status (see above) applies to categories only. Dominance and c-command are restricted to terms (ibid.: 339). If it were not for X°-adjunction, crucially employed in head-movement, no further complication should arise.[28] With the notion "disconnected" taken to hold for two terms if neither dominates the other, the following (standard) definition of c-command suffices.

(32) X c-commands Y if (a) every Z that dominates X dominates Y and (b) X and Y are disconnected. (ibid.:339)

In head-movement adjunction (30a), however, the moved X° is standardly taken to c-command its trace. This has sometimes been ensured by index percolation, which transfers the c-command properties of Y° to X°.

Yet, c-command properties are not defined for a two segment category that restricts dominance and c-command to terms. Adding such definitions for categories as well, it is possible to make the necessary distinctions for an adjunct to c-command *outside* the category it is adjoined to (cf. ibid. 339). Thus X° c-commands its trace in (30a). However, complications arise in other areas of the phrase structure. If adjuncts are allowed to attach to intermediate projections (ibid. 330) they will c-command the specifier of the same projection.

(33) $[_{XP}$ YP $[_{X'}$ ADJ $[_{X'}$ X° ZP]]]

As soon as both YP and ADJ are complex, they cannot be ordered with respect to each other—assuming the LCA is adopted (see above)—and no precedence ordering can be assigned to them. Similar effects—plus potential binding theoretic consequences—will arise from multiple XP-adjunction to YP, as has often been proposed for scrambling:

(34) $[_{YP}$ XP_i $[_{YP}$ ZP_k $[_{YP}$ . . . t_i . . . t_k . . .]]]

In (34), XP and ZP cannot be ordered with respect to each other.

Even with these issues in mind, the considerable success of using adjunction in the area of movement theory and locality cannot easily be overlooked (see GREWENDORF & SABEL for further arguments that specifiers and adjunction sites should be kept apart.) Still, radical departures have been proposed. Thus, the innovative work of Alexiadou (1995) reanalyzes all adjuncts as specifiers of separate (functional) projections. Zwart's (1993) pioneering study of Germanic SOV-languages opens the possibility that what has been called scrambling is actually movement into functional specifiers, AGRoP most prominently. Even X°-movement can be conceived of as creating its own landing sites rather than adjoining to a target (Ackema, Neeleman & Weerman 1993). Moreover, much of the data associated with adjunction might pertain to a stylistic component of the language faculty dealing with marked phenomena and thus belong

outside the computational system as modelled by minimalist principles (Chomsky 1995b:324f). See also REINHART, who presents an integrative theory.

6 Competing Derivations, *there*-Sentences and FI

The introduction of economy principles into the computational system requires a reappraisal of *global* conditions.[29] In order to be fully aware of the options available, one has to look closely at the reasoning that transformed "guidelines [which] have a kind of 'least effort' flavor to them" into "actual principles of language" (Chomsky 1991:418). A central case in point is the analysis of X°-movement in English and French. Thus, the following contrast has been explained by counting the number of operations required to derive the structures:

(35) a. Jean embrasse souvent Marie.
John kisses often Mary
b. *Jean souvent embrasse Marie.

Simplifying matters radically,[30] we can say that (35a) blocks (35b) because the derivation of (35a) employs one step less than the derivation of (35b). Both involve a V°-to-I° movement, overt in (35a) but covert in (35b), the surface structure of (35b) having arisen from an additional, and therefore crucially superfluous I°-to-V° lowering operation.

It is immediately clear—although it was not discussed at the time when (35) was analyzed—that blocking effects cannot just be a global numerical property. Thus, (36a), which arguably involves fewer derivational steps, certainly does not block (36b).

(36) a. Jean dort.
'John sleeps'
b. Je sais qu'il dort.
'I know that he sleeps'

Trivial though this point may appear at first sight, it is essential that, for transderivational economy principles to apply properly, *candidate or reference sets* be determined among which economy principles are supposed to select the optimal competitor (recall the discussion in section 1 above). This central task is confronted by STERNEFELD. The same issue underlies much of the debate in MÜLLER's and REINHART's papers.

Chomsky's response to this question has been to introduce the concept of a *numeration*, i.e. the array of lexical items selected from the lexicon to be mapped into interface representations (see above). As a first approximation, economy conditions apply relative to such numerations.

A second domain in which 'least effort' plays a role is the insertion of dummy elements. (non-emphatic) *do*-insertion in (37b), which otherwise has to occur in negation and interrogative contexts must not block (37a):

(37) a. John wrote books.
b. *John did write books.

(37b)—again we simplify radically—involves only *do*-insertion to pick up inflectional material in I°, whereas (37a) requires overt I°-to-V° lowering plus V°-to-I° raising at LF. To prevent a blocking effect in this case, language particular rules like *do*-insertion are assumed to be costlier than the universal movement rules operative in (37a). *Do*-insertion is only a last resort to salvage otherwise ungrammatical derivations. Another line of dissociating (37a) and (37b) would be to include *do* in the numeration for (37b). Numerations would thus differ and no competition among the derivations can arise. The illformedness of (37b), of course, would have to be explained independently (cf. Wilder & Cavar 1994 for a proposal).

Further refinement is needed to analyze *there*-insertion, a construction that has been at the core of much of minimalist syntax. Here it has been assumed that *there* will always be part of the numeration of the sentences in which it appears. Thus, (38a) and (38b) do not compete:

(38) a. A unicorn is in the garden.
b. There is a unicorn in the garden.

Now, to account for the distribution in (39), Chomsky (1995b:344ff) assumes that economy conditions apply in a more local fashion.

(39) a. There$_i$ seems t_i to be someone in the room.
b. *There seems someone$_i$ to be t_i in the room.

Given the strictly bottom-up procedure of generating syntactic structures (see above), choice points arise at which the cheapest option possessing a well-formed continuation (of whatever complexity) must be selected (cf. FRAMPTON). As soon as (40) has been arrived at, the specifier of IP must be filled:

(40) [$_{I'}$ to be someone in the garden]

Since *there* is part of the numeration, we have two options: either insert *there* or move *someone*.

Independently, it is assumed that the operation Merge is cheaper than Move, Move being subject to Procrastinate. Arguably it is also desirable to empty the numeration as fast as possible to reach the end of the derivation. Consequently, *there* has to be inserted. When it comes to filling Spec,IP of the matrix clause in (41), no option arises.

(41) [$_{I'}$ seems [there to be someone in the garden]]

There, being the only element visible for the movement rule, will be attracted to that position. The result is that (39a) blocks (39b) as desired.

Yet, things turn out to be more complicated than that. The reasoning employed to analyze the data in (39) predicts that (42a) should block (42b).

(42) a. [A rumour [$_{CP1}$ that there was a unicorn in the garden]]$_i$ was t_i in the air.
b. There was [a rumour [$_{CP1}$ that a unicorn$_i$ was t_i in the garden]] in the air.

Both, however, are fully grammatical. The logic of embedding requires CP1 to be completed first. Filling IP-Spec in CP1 gives rise to the first relevant choice point (43a). Again, *there* must be inserted, yielding (43b):

(43) a. [$_{I'}$ was a unicorn in the garden]
b. [$_{IP}$ there was a unicorn in the garden]

When it comes to filling the IP-Spec of the matrix clause, once more no option is left. *There* having been taken out of the numeration, movement of the complex NP is all that can be done, turning (44a) into (44b) (= 42a).

(44) a. [$_{I'}$ was [a rumour [$_{CP1}$ that there was a unicorn ...]] in the air]
b. [A rumour [$_{CP1}$ that there was a unicorn ...]]$_i$ was t_i in the air.

The reverse ordering of Move and Merge operations required to derive (42b) is not available under the assumptions made.[31]

FRAMPTON offers a cross-linguistic analysis of *there*-constructions that does not run into the problem just described. Taking up the idea exploited for *do*-insertion above, *there* is not part of the numeration but can be inserted at stages where this is required as a last resort.

There-constructions also originally provided a core argument for the principle of representational economy (FI). The central facts of the construction—(i) verbal agreement is determined by the NP associate of *there* (45); (ii) *there* and its NP associate stand in the same relation to each other as links of an A-chain; and (iii) *there* occupies what used to be called a Case position—can be accounted for by assuming that *there* is replaced at LF by its associate via A-movement, the associate being in need of Case licensing.[32]

(45) a. There is a unicorn in the garden.
b. There were some unicorns in the garden.

Replacing *there* would then independently satisfy the principle of *Full Interpretation* (FI) that requires that interface representations contain no symbols which cannot be assigned an interpretation, i.e. symbols may only transport articulatory-perceptual instructions at PF, and conceptual-intentional instructions at LF. Thus, pleonastic elements, being semantically vacuous (Chomsky 1995b:287), do not qualify as legitimate LF-objects.[33]

While the MP incorporates conditions on representations (FI), much of the burden of constraining derivational possibilities lies with specific 'derivational economy' principles, governing applicability and timing of derivational operations (Procrastinate, Shortest Path, Greed, etc.). In this, it diverges from the GB paradigm—research in the 1980's focussed on the formulation of principles governing *representations*, to constrain

movement. A natural tendency was to view the model as essentially equivalent to (or as a notational variant of) a purely representational model, with chains—a representational notion—replacing the derivational operation of Move-α (however, as Chomsky has argued, there may be real differences between the two).

A representational approach has been argued to be preferred in that it incorporates "radical derivational economy"—derivational operations are minimized, i.e. there are none (cf. HAIDER's paper; also Brody 1995); but that amounts simply to the claim there is no derivational economy principle at all (given the point about zero output made in section 1). The point is far more that proponents of representational models have to show how the insights are to be captured without appeal to derivational economy.

Even among the papers which adopt a derivational syntax, the tendency is to rein in the scope of true derivational economy (transderivational comparison). Thus STERNEFELD argues that the Shortest Path condition should be formulated as a derivational condition (an absolute restriction on movement), rather than a transderivational condition (comparing potential derivations). FRAMPTON's approach also is "motivated partly by . . . the intuition that the role of optimality in the theory should be minimized". This matches the direction taken in Chomsky (1995b:Ch.4). There, aspects of both 'derivational economy' (e.g. the Minimal Link Condition) and FI (the requirement for successful feature matching) are reanalyzed as conditions on the movement operation itself, such that 'violations' lead directly to termination (cancellation) of the derivation, rather than further computation to convergence, with comparison among convergent derivations limited now to Procrastinate. Considering this trend—what John Frampton (p.c.) has aptly called "the withering away of economy"—one might be drawn to conclude that HAIDER's remarks are not too far off the mark:

> If the UG-facilitated choice of a core grammar L is determined by economy principles on the derivational complexity, this does not necessarily mean that the core grammar of L contains economy principles.

II The Papers

The first six contributions are concerned with the role of economy principles in the domain of syntax proper.

FRAMPTON takes issue in his paper "Expletive Insertion" with problems of computational complexity that arise from the assumption of economy conditions that compare alternative derivations. The paradigm example is (46)—in Chomsky's analysis, the derivation leading to (46a) blocks the one leading to (46b) (cf. section 6 of part I above):

(46) a. There seemed to be someone in the room.
b. *There seemed someone to be in the room.

His main proposals are (i) that the associate nominal in this construction is not a DP but an NP, so that (46a) and (46b) do not involve identical numerations, hence do not stand

in a blocking relation; (ii) —returning to an earlier idea—that expletives are not included in the numeration prior to the syntactic derivation, but are inserted in the course of the derivation; and (iii) that Economy conditions include not only *Avoid Movement* (Chomsky's *Procrastinate*, i.e. merge < move) but also *Last Resort Insertion*—which renders expletive insertion only possible if neither merge nor move yield a convergent derivation (i.e. merge < move < insert).

Empirical coverage includes the 'dative subject' constructions of Icelandic—which necessitate a broader concept of Case features—and some notoriously difficult 'exceptions' to the definiteness effect in *there*-constructions, which are analyzed on a par with Possessor Raising.

The paper tracks several important revisions of the MP (Chomsky 1993) that are adopted in Chapter 4 of Chomsky (1995b). A central innovation is the distinction between 'interpretable' ('intrinsic') and 'non-interpretable' ('formal') features. The latter—e.g. Case—retain the property of needing to be checked and eliminated, while the former—e.g. ϕ-features on nominals, which are clearly able to influence semantic interpretation—are assumed to survive at the LF-interface. Intrinsic features may therefore enter checking relations, but need not do so, and do not delete if they do. Hence, unlike formal features, intrinsic features may enter into more than one checking relation. This entails the abandonment of the strong version of the Greed principle ("move α only to satisfy properties of α") in favor of a weaker version that permits movement of α triggered solely by properties of a head at the target position. The shift in perspective from the properties of α to properties of the target position as determining movement, leads to its reformulation as *Attract* α.

The theoretical treatment of Scrambling phenomena found in German, and in several East Asian languages, forms a central theme in the next four papers. Scrambling poses a challenge for MP in at least three respects: (i) it shows strict locality effects; (ii) it is apparently optional; (iii) it is not obviously morphologically triggered. The status of scrambling with respect to the A/A'-distinction is also controversial. These issues are all addressed in GREWENDORF & SABEL's paper "*Wh*-scrambling in the Minimalist framework" (= G&S).

Most accounts treat Scrambling (e.g. in German and Japanese) as movement to an adjoined position. Compared with earlier models (e.g. Chomsky 1986a), the role of phrasal adjunction diminishes in the MP, which concentrates on Spec-to-Spec and head-movement. But the account of long movement as iterated short steps, enforced by locality (*Minimal Link Condition*), is retained. Locality should apply to movement to adjoined positions just as it does to movement to specifiers, permitting iterated local steps. G&S propose a radical alternative, the *Constraint on Adjunction*:

(47) α may undergo movement and adjunction only *once* in a derivation.

The paper explores an empirical domain with special relevance to (47)—the complex interaction of scrambling and *wh*-movement in Japanese. Scrambling in Japanese (and Korean—cf. MÜLLER's and STERNEFELD's papers) is far more liberal than in German: while scrambling is a finite-clause internal phenomenon in German, phrases can 'scramble' out of finite clauses in Japanese ('long distance scrambling'). Also, *wh*-phrases may scramble in Japanese, but not in German. Additional complexities in

Japanese include the possibility for a *wh*-phrase to scramble to a position outside of its scope domain, and the so-called 'anti-superiority effect' in multiple questions.

In response to the 'optionality' and 'triggering' questions, G&S develop a distributionally intricate system of Σ- (= scrambling) and *wh*-features, associated with functional heads, which then drive overt scrambling and covert *wh*-movement. Further differences are derived by appeal to properties of those features, and of the functional heads which host them.

G&S argue that in Japanese, short scrambling is A-movement, while both short scrambling in German and long scrambling in Japanese are instances of A'-movement. The contrast with respect to short scrambling is related to the (non-)availability of 'extra' specifier positions as landing sites for scrambling (exploiting the possibility for multiple specifiers in the system of Chomsky 1995a). Assuming that in German, specifier positions are not available as landing sites for scrambling, scrambled phrases can only adjoin; the strictly local nature of German scrambling then follows from (47). The properties of short scrambling scrambling in Japanese indicate that additional A-specifier positions are available, while the possibility of long scrambling is due to the availability of extra A'-specifiers, enabling the effects of (47) to be avoided. Fully within the premises of MP, the differences between German and Japanese must be parametric, deriving from properties of functional items: G&S propose that AGR is the element involved.

The MP reopens the case for transderivational constraints on syntax; alongside Procrastinate, these include the principles of 'Fewest Steps' and 'Shortest Path' as conditions on syntactic operations. STERNEFELD formulates the notion more precisely as follows:

(48) *Global Economy Condition*
Given two derivations D1 and D2 in the same reference set RS, D1 is preferred over D2 if and only if D1 fares better than D2 with respect to some metrical measure M.

His article "Comparing Reference Sets" tackles the main issue of syntactic globality, namely, how to single out the right competitors, derivations in the case at hand. Concentrating on superiority effects (which motivate the Shortest Path condition), and scrambling phenomena, Sternefeld subjects various formulations of the central notion of reference set to close scrutiny.

Careful investigation of the Shortest Path approach to superiority phenomena reveals that neither numerations alone, nor a combination of numeration plus LF-output, furnish a sufficient basis for RSs. Consider (49)-(52):

(49) Whom$_j$ did John persuade t_j to visit whom?
(50) * Whom$_j$ did John persuade whom to visit t_j ?
(51) Who wonders who$_j$ t_j bought what?
(52) Who wonders what$_j$ who bought t_j ?

Competition on the basis of identical numerations alone—although adequate for (49) vs. (50)—would improperly predict (51) to block (52). Adding a requirement for

identical LF-output to the definition of RSs would correctly place (51) and (52) in different RSs—but would fail on the contrast (49)-(50) in turn.

An alternative definition of RSs which invokes identity of (truth-conditional) meaning instead of identity of LF-output is considered and rejected. This would run into problems with such 'stylistic' movements as scrambling to IP in German, or long distance scrambling in Korean—these displacements lack truth-conditional effects, and should thus be blocked by their counterparts (competitor derivations) displaying shorter scrambling.

Given that superiority effects are absent (at least) in German (cf. also HAIDER's paper), Sternefeld considers parametrizing the defining property of RSs. Reviewing the facts relating to cyclicity and Fewest Steps, he concludes that even for English, no fully satisfactory definition of RSs emerges.

The remainder of the paper is devoted to replacing transderivational constraints by derivational constraints, which enables the notion of RS to be dispensed with entirely. On the assumption that scope-taking elements possess scope indices, superiority can be assimilated to Crossover effects, and the contrast between English and German can be made to fall out directly.

Sternefeld concludes that, given the computational complexity of transderivational constraints as well as the doubtful status of the concept of meaning in the treatment of syntactic anomaly, a rival account in terms of purely derivational constraints should be preferred.

MÜLLER's "Optional Movement and the Interaction of Economy Constraints" develops a approach to 'optional' movement within the MP that permits an account of attested instances (53), while continuing to exclude unattested instances (54):

(53) a. Partial *wh*-movement in Iraqi Arabic and German
b. Intermediate steps in *wh*-movement targeting [–WH]-COMP positions

(54) a. Topicalization of *wh*-phrases
b. Scrambling of *wh*-phrases
c. Covert *wh*-movement from a [+WH]-COMP

This is done on the basis of the central definitions (55)-(57):

(55) *Reference Set*
Two derivations D1 and D2 are in the same reference set iff they yield the same LF output and converge at LF and PF.

(56) *Fewest Steps*
If two derivations D1 and D2 are in the same reference set and D1 involves fewer checking operations than D2, then D1 is to be preferred over D2.

(57) *Last Resort*
Move raises α to a position ß only if ß is a typical checking position for an unchecked morphological feature of α.

Given certain idealizations concerning intermediate traces in the LF-output and assuming movement to be subject to the *Minimal Link Condition* (MLC), the following account emerges.

(53b) is ruled in, since intermediate steps do not count as checking operations; Movement obeys Last Resort (a weakening of the original version of *Greed*) since COMP is a typical checking position for *wh*-elements. Skipping intermediate COMPs on the other hand does not involve fewer checking operations, while it violates MLC. Although surface facts are somewhat tricky for partial *wh*-movement (53a), the account is fully analogous to (53b).

Topicalization of a *wh*-phrase (54a) must be followed by *wh*-movement at LF, and requires two checking operations, [Spec,Top] not being a typical checking position of *wh*-elements. The same LF-output— modulo intermediate traces—can be arrived at by direct *wh*-movement into the [+WH]-COMP, requiring only one checking operation. *Wh*-topicalization is thus banned. Scrambling of *wh*-phrases (54b) is ruled out analogously.

The case in (54c) concerns the absence of a wide scope reading for *where* in (58) (cf. Epstein 1992):

(58) Who wonders where we bought what?
(59) Who wonders what we bought where?

The derivation permitting a matrix scope reading of *where* and an embedded scope reading of *what* in (59) will block the derivation leading to the same reading for (58), because it involves fewer checking operations (assuming that *where* in (58) could bear two checkable [WH]-features).

Müller's account (54a)-(54c) constitutes a derivational version of the Principle of Unambiguous Binding (Müller&Sternefeld 1993), which rules out simultaneous binding of traces from different types of positions. The theory is accordingly extended to cases covered by that earlier account, including the ban on 'Super-Raising', 'Chain Interleaving', and long distance Scrambling in German; and the analysis of *wh*-island asymmetries.

REINHART's paper "Interface Economy and Markedness" offers an intriguing perspective on the link between word order, stress patterns and Economy. The core claims are (i) the assignment of phrasal stress patterns is regulated by Economy; (ii) alternative word orders can enter into that Economy calculation. Following Cinque (1993), neutral phrasal stress is assumed to be calculated from syntactic structure by algorithm. 'Special' or marked patterns are assigned by 'discourse grammar', overriding the output of grammar, to meet discourse needs that cannot be expressed by the neutral pattern. The choice between 'costfree' neutral stress and 'costly' marked stress is governed by economy.

Reinhart is concerned with the empirical content of 'markedness'—how can a pattern be 'marked' when it is neutral in that context for that sentence? Markedness effects may be only indirectly observable: where a language has a word order option that permits neutral stress in a certain context, that option must be used, while the same context may force special stress in a language that does not have that word order pattern. This idea is

explored by contrasting Dutch, which has object-scrambling, with English, which does not.

In Dutch, neutral stress falls on the object in 'unscrambled order', but when the object is scrambled (in pre-adverb position), it falls on the verb:

(60) a. dat ik altijd het krant leze
that I always the paper read
b. dat ik het krant altijd leze

Reinhart argues that there is no single determinant of scrambling to be found in properties of the object (contra e.g. de Hoop 1992). Rather, the common determinant (where there is one) is to be found in discourse properties of the clause as a whole. The choice between scrambled and non-scrambled orders is made at PF, where neutral stress determines overlapping but distinct 'appropriate contexts' for each.

For the purposes of the economy calculation, both word orders are free; but special (shifted) stress is more costly than neutral stress. In a context that demands the use of a scrambled order in Dutch (60b), English can only use the basic order, and stress shift applies (61a). Crucially, non-scrambled order with stress shift cannot be used in the same context in Dutch (unlike (61a), (61b) permits only contrastive focus on the verb):

(61) a. I always read the newpaper (= 60b)
b. dat ik altijd het krant leze (≠ 60b)

In Dutch, the use of the scrambled order with neutral stress blocks the use of the basic order with stress shift, as dictated by economy.

The proposal has consequences for the treatment of scrambling: not only is it 'optional', but it also is 'free' as far as syntax is concerned. This is not expected under a movement analysis, assuming that movement is a costly operation. Indeed, Reinhart favours the 'non-movement' view in which 'basic' and 'scrambled' orders merely reflect options for the base-generation of objects.

In his paper "Formal and substantive elegance in the minimalist system", URIAGEREKA addresses both external and internal aspects of economy: the research strategy of eliminating redundancy, as well as the exploitation of economy as a principle within the system. Minimalism, he points out, directly invites the suspicion that linguistic phenomena do not directly instantiate principles of the underlying system. The search for deeper explanations involves continual reassessment of principles—including the economy principles of MP— and the generalizations that gave rise to them, as theorems of more abstract principles, such that the former become 'emergent properties' of the system (hence the subtitle: "on the emergence of some linguistic forms").

The paper contains two major proposals, one concerning the LCA, the other concerning Condition B of the Binding theory. First, a model is outlined in which the LCA emerges as a theorem, given reasonable assumptions concerning external necessities ('bare output conditions') and economy. Building on Epstein's (1995) proposal that c-command is a direct reflex of Merger, the key new feature is that spell-out—including linearization—is interspersed with merger operations, hence scattered through the derivation ('multiple spell-out'). There are far-reaching implications to be

worked out; it is certainly a natural extension of the strictly derivational approach to the architecture of grammar pursued by Chomsky, Epstein and others.

Against this background, Uriagereka approaches two seemingly dissociated aspects of grammar: local obviation effects, and the role of Case in the minimalist system. Among the aspects of grammar that are sensitive to locality, these two are conspicuous. Case is mysterious—the only clear instance of features that are always 'uninterpretable', with solely system-internal function and no role to play at the interface. Condition B stands out as a local 'negative' condition, blocking binding relations for pronouns, which unlike R-expressions, are not generally required to have distinct reference from c-commanding expressions.

In Chomsky's (1995b) system, following movement to functional heads, the interpretable features of subject and object(s) end up in one checking domain, Case having been eliminated. When two such arguments are pronouns with the same (interpretable) ϕ-features, the question arises of how the two feature-sets are distinguished *at all* at the interface. Uriagereka's proposal is that Case has this function: to 'mark' the feature-sets of arguments within a single local domain as distinct from one another. A pronoun's Case will mark it as distinct from other feature sets in its domain. The local obviation property now follows with one additional, reasonable assumption: feature sets formally marked as distinct in the grammar receive distinct interpretations at the interface.

HAIDER's paper "Economy in Syntax is Projective" defends a view on syntax which diverges from mainstream models. Taking strings of terminals as its starting point, the task of syntax is to assign hierarchical structures to such strings. Constraints are essentially constraints on representations; including constraints open to interpretation as being economy-based, which are responsible for banning string-vacuous chains, empty projections and non-branching structure. Haider presents arguments against (current versions of) derivational economy principles, based on head movement constructions in English that would have to be taken to arise from optional movement in a derivational model. The effect of optional movement is attributed in the 'projective' model to underspecification of the items involved. Further empirical difficulties for the derivational model are set out in an insightful analysis of superiority phenomena in German and English.

Conceptually, projective economy is claimed to best correspond to UG's role as a "grammar co-processor"; this is Haider's own reconstruction of UG, primarily motivated by its being better designed to cope with questions of parameter fixing (cf. Haider 1991). Haider's string-to-structure economy further purports to be a natural extension of language processing, where time can be considered a limited resource. As the following passage makes clear, these proposals are founded on an unorthodox view of the competence-performance distinction, and hence of the domain of linguistic theory:

> The problem addressed in the Chomskyan question "What is the structure of the grammar?" is directly connected with the question "How is the grammar put to use?". The grammar is to provide optimal data structures for actual usage. This implies that UG is the system of cognitive routines that guarantee this result, that is, grammars that determine optimal data structures for actual usage.

The final two papers move beyond the domain of syntax proper, to explore the role of economy principles in morphology and the lexicon.

In "Lexical Information from a Minimalist Point of View", BIERWISCH explores aspects of the tension between the conception of the lexicon as nothing more than a 'repository of idiosyncrasy', on the one hand; and as a 'system of entries' defining the atoms that enter the syntactic derivation, on the other. He concludes that "the assumption that lexical information is subject to economizing principles that properly belong to the Lexical System should be taken as an indispensible ingredient of the minimalist perspective".

Both individual lexical entries (LEs) themselves, and the lexical system (LS) as a whole, are shown to be in fact finely structured. This structure is governed by principles that apply to LS (but not to syntax proper); and these principles may be invariant (universal)—e.g. the principles determining hierarchy of argument positions; or emerge in language-specific variants—e.g. the rules determining the Case-feature in the LE of a Case-assigning item.

It is argued that "the notational convention of of disjunctive ordering, independently motivated for phonology" should be "extended to lexical information in other domains, where they also seem to apply naturally". On this basis, the effects of economizing principles (underspecification, elsewhere) can be properly exploited in accounting for the actual patterning that occurs. Just as the phonological information stored in a LE is underspecified with respect to the information associated with that entry at later stages of the derivation and at the A-P interface, so is the semantic information in a LE underspecified with respect to the information necessary to compute its contrbution to interpretation at the C-I interface. This is illustrated with subtle but general interpretive properties of change-of-state verbs and their 'unaccusative' cognates. The use of elsewhere is illustrated a.o. with Case-assignment by German prepositions. By exploiting default rules, Case-features need not be specified in lexical entries, except where a lexical specification is necessary to override the default. As in the case of affixation, closer examination of the interaction of subregularities with the Elsewhere condition reveals non-trivial effects. While the default Case for P is Dative, the accusative assigned by directional variants of prepositions such as *auf* (on, onto) must be analyzed as following from a more specific default rule. This in turn means that the Dative assigned by directional prepositions such as *zu* (to) must be marked (lexically specified), despite the existence of the Dative default.

WUNDERLICH's "A Minimalist Model of Inflectional Morphology" presents a general framework for accounting for the distribution and properties of word-forms in paradigms. *Minimalist Morphology* (MM) is based on (i) a maximally simple rule of combination; (ii) specific filters that are plausibly grounded in other systems; (iii) general principles of economy governing the selection of optimal word forms.

The sources for word-forms are sets of stems and affixes, both maximally underspecified. Free combination of affixes and stems (i) represents the minimal assumption about the 'form-generation' component. Independent principles (ii) that determine the general shape of legitimate output forms include the *Affix Order* constraint (forbidding e.g. the attachment of an Aspect-affix to a V-stem already bearing a Tense affix), imposed by external, semantic factors. The potentially massive overgeneration that is the price of assuming free combination is further constrained by a few simple, natural economy principles (iii) (cf. the strategy for curbing move-α in MP). The key mechanism is the paradigm: only those forms that are legitimated in a paradigm will enter into the syntax (hence this is a lexicalist approach—word forms are determined prior to syntax). The dimensions of a paradigm are defined by the inherent features of the most specific candidate form; what paradigm slot(s) a given form eventually occupies is determined by a set of interacting economy principles. These set up a competition among the candidate forms generated to fill the slots opened in a paradigm, by ranking the forms via the specificity of their inherent feature specifications, and inserting them accordingly.

The model is not only economical in the assumptions it makes (e.g. no rules other than free combination), it also relies crucially on both derivational and representational economy conditions, based on specificity, simplicity, number of operations. One of its attractive features is that it dispenses with a range of more powerful rules / mechanisms assumed by other authors to handle word-form selection. To the extent that it can stand up to the empirical challenge, MM thus seems to gain an intrinsic edge over competing models. It is with this in mind that the model is tested in case studies on two challenging verb inflection systems—Georgian and Potawatomi.

Notes

1 Chomsky (1993) was circulated as Chomsky (1992), and reprinted as chapter 3 of Chomsky (1995b). We treat Chapter 4 of Chomsky (1995b) as the 'definitive' version of the 'second phase'. It is important to be aware that, although much of the "Bare Phrase Structure" paper (first circulated as Chomsky 1994) was incorporated into "Categories and Transformations", some proposals from the former were abandoned in the latter.

2 Cf. Chomsky's (1995b:8-9) 'internal' and 'external' notions of 'simplicity', the former having "resurfaced [in the minimalist program] in the form of economy considerations that select among derivations"; the latter "operative as always . . .".

3 Prime examples are parsing and pragmatics. The central role played by 'Minimal Attachment' and similar principles in theories of sentence processing is reviewed in Frazier & Clifton (1996:Ch.1). In pragmatics, see especially Sperber and Wilson's Relevance Theory of communication and cognition, according to which "human cognitive processes [including utterance interpretation] . . . are geared to achieving the greatest possible cognitive effect for the smallest possible processing effort" (Sperber & Wilson 1986:vii).

4 Economy is now writ large in phonology in the guise of Optimality Theory, following Prince & Smolensky (1993). Work is also being done on OT-approaches to syntax, e.g. Grimshaw (1995). Since phonology is not a topic of this volume, and none of the papers discuss optimality-theoretic approaches to syntax, OT is not discussed here.

5 Cf. discussion in Shoemaker (1991). There are terminological traps to be aware of. The use of terms like 'best' / 'optimal' in their everyday (nontechnical) sense amounts to an *interpretation* of the ranking in terms of X, imposed from the outside. The (quantitative) ranking itself must not be confused with how it may be qualitatively evaluated. In similar vein, there is no question that the description of a system in terms of economy as discussed here involves attribution of 'conscious intentions' and the like, either to that system or to the 'competing' alternatives it generates—rather, such a description represents an analysis at a purely mechanistic level, e.g. an abstract characterization of some algorithm.

6 The case where 'no output' can be legitimately selected by an economy principle is a different one: 'no output' must be among a set of alternatives which the system makes available for selection, i.e. the

system is not *S'*. Moreover, 'no output' represents an option for achieving the goal at hand, rather than failure with respect to that goal. Cf section 6 below.

7 A caveat is in order where elegance comes into the picture. It is by no means clear that elegance and economy considerations lead to the same results. The topic of simplicity in scientific theories goes beyond the scope of this introduction—an early serious treatment can be found in Goodman (1977).

8 On blocking effects in derivational morphology, see Aronoff (1976:42-45); for example, while the affix *-ity* applies to adjectives in *-ous* to derive abstract nominals (*curious - curiosity*; *porous - porosity*), the existence of a corresponding unaffixed N-form blocks *-ity*-affixation to the N+*ous* adjective. Thus, *glory* blocks affixation of *-ity* to *glorious* (**gloriosity*). Kiparsky (1982) applied the Elsewhere condition, originally deployed to impose disjunctive ordering on phonological rules, to do the same with derivational and inflectional morphological processes, thereby deriving blocking effects in morphology.

9 This illustration simplifies Kiparsky's (1982) formulation, while preserving the key intuition. Incorporating ablaut and other cases requires more complexity; cf. Halle & Marantz (1993) for comprehensive treatment of English past-formation in this spirit.

10 An account of exceptional 'subject control' with *promise* consistent with the Minimal Distance Principle is found in Larson (1990).

11 Leading instances in the 1980's are also the immediate precursors to the MP: (i) the Last Resort condition for Case-driven NP-movement, discussed in Chomsky (1986b:137), and extended in Chomsky (1991); and (ii) the minimality condition on government of Chomsky (1986b), which developed into the theory of Relativized Minimality (Rizzi 1990).

12 This aspect of the tensiond presupposes simplicity as a criterion for explanatory adequacy. There is another aspect of the same tension, induced on the one hand by the facts of language variation, pulling us towards a diversity of grammars; and on the other, by the facts of language acquisition, leading us to expect a single grammar. These two aspects are strictly speaking logically independent, though progress on each seems to lead in the same direction (cf. Chomsky 1995b:4-5).

13 Koster (1987) had already forcefully argued that among the three genuinely syntactic strata one would be sufficient to handle the data—S-structure on his proposal.

14 There are of course many complications; for early discussion, cf. Lasnik & Saito (1984). These issues are taken up in this volume by STERNEFELD, GREWENDORF & SABEL, and MÜLLER.

15 Interface well-formedness is Full Interpretation (FI) (see below). The idea is that the morphosyntactic features that need to be matched (or assigned) are introduced into derivations as parts of lexical items, but are not interpretable at the LF interface, hence not tolerated there. Strong features are additionally not tolerated at PF. The matching requirements (8), recast as checking theory (matching and deletion of features, governed by X'-theoretic locality) still serve to 'motivate' movement; movement and checking are now interpreted as the means used by the computation to remove disturbing features. The checking theory of Chomsky (1993) is overhauled and refined in Chomsky (1995b:Ch.4).

16 The cost is extension of the 'lookahead' requirement: a successful computation must determine whether well-formedness at a non-syntactic level (PF) necessitates application of an operation. 'Strength' is reinterpreted in Chomsky (1995b:233) as affecting not PF but the compuational operations themselves—the introduction of a strong feature from the lexicon into a derivation requires immediate satisfaction (checking). See URIAGEREKA for some speculations on the nature of 'strength'.

17 The role of the pleonastic subject in such paradigms is discussed in section 6 below; this also forms the topic of FRAMPTON's paper.

18 Note the conceptual closeness to the architecture of categorial grammar where lexical categories (types) largely determine syntactic combination.

19 Given that the 'storage capacity' of the human brain is vastly underoccupied, there is no a priori reason to expect that minimization of storage space is an externally dictated necessity.

20 Of course, different approaches to phrase structure exist within the generative paradigm. See Baltin & Kroch (1989) and Brody (1995b).

21 This process is characterized as "an operation, call it *Satisfy*, which selects an array of items from the lexicon and presents it in a format satisfying the conditions of X-bar theory." (Chomsky 1993)

22 Project can also be considered a relation holding between X and X° in a, X°—being [$_{X°}$ X] strictly speaking—between X° and X' in (20b), and between X' and XP in (20c). Additionally, this relation is transitively closed.

23 This is an informal translation indeed. The 'ontology' of minimalist syntax has yet to be spelled out in a transparent way. For a promising attempt see Wartena (1994).

24 For an interesting discussion of these difficult and controversial matters see Stabler (1983).

25 (22) is adapted from Kayne's originally more elegant formulation:

(i) *LCA* d(A) is a linear ordering of T (Kayne 1994:6)

A is the set of ordered pairs <x,y> of nodes, such that x asymmetrically c-commands y. T is the set of terminals.

26 Incidentally, there is considerable similarity to the compositionality principles advocated in Montague semantics and categorial grammar. Thus, the sisterhood relation is the canonical configuration for function application.

27 See Safir (1986) for similar considerations and Brody (1995) for a critical review of the reconstruction facts.

28 Recall that X°-movement adjunction also constitutes an exception to Epstein's (1995) theory.

29 Globality here concerns transderivational conditions. Lakoff (1970) deals with derivational conditions only. The latter type of globality has since been incorporated into syntax by means of trace theory, feature percolation, and checking mechanisms, as well as conditions on representations.

30 See Grewendorf (1990) and Lasnik (1994) for detailed discussion, including mention of some recalcitrant problems with this analysis.

31 See Lasnik (1995) for some discussion of the same paradigm and an alternative analysis of *there*-constructions.

32 The analysis of ECM contexts in this regard is non-trivial.

(i) I believe there to be a unicorn in the garden.

33 Cf the general discussion of the interpretability of features by Chomsky (1995b:276ff). The analysis of *there*-constructions has seen a lot of technical refinement, far too detailed and controversial to be covered here. Again, we refer the reader to FRAMPTON , Lasnik (1995) and Chomsky (1995b, sections 4.5.3, 4.9, and 4.10.3). The tendency, however, seems to be to dissociate the properties of *there* and its NP associate, such that the role of FI diminishes. Note, incidentally, that applying the logic of *there*-constructions to *do*-insertion would require the English main verb to replace *do* at LF for reasons of FI. In structures like (i)

(i) John does not write books

The crossing of NEG° by V° at LF would, however, result in exactly the kind of ECP violation that the entire analysis was supposed to avoid.

References

Ackema, Peter, Ad Neeleman and Fred Weerman. 1993. Deriving Functional Projections. *OTS Working Papers*. TL-93-001.

Alexiadou, Artemis. 1994. Issues in the Syntax of Adverbs. Doctoral dissertation. University of Potsdam.

Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge, Mass.: MIT Press.

Baltin, Mark and Anthony Kroch eds. 1989. *Aternative Conceptions of Phrase Structure*. Chicago: Chicago University Press.

Baker, Mark. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: Chicago University Press.
Brody, Michael. 1995a. *Lexico-Logical Form*. Cambridge, Mass.: MIT Press.
Brody, Michael. 1995b. Towards Perfect Syntax. *Working Papers in the Theory of Grammar, Vol. 2, No.4.* Budapest.
Chomsky, Noam. 1973. Conditions on Transformations. In: S. Anderson & P. Kiparsky (eds.) *A Festschrift for Morris Halle*. Holt, Reinhart & Winston: New York. 236-286.
Chomsky, Noam. 1981.*Lectures on Government and Binding*. Foris: Dordrecht.
Chomsky, Noam. 1986a. *Barriers*. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 1986b. *Knowledge of Language: Its Nature, Origins and Use*. New York: Praeger
Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In *Principles and parameters in comparative grammar*, ed. R. Freidin, 417-454. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 1992. A Minimalist Program for Linguistic Theory. *MIT Occasional Papers in Linguistic Theory*, 1. (= Chomsky 1993)
Chomsky, Noam 1993. A Minimalist Program for Linguistic Theory. In Hale and Keyser, ed. 1-51.
Chomsky, Noam. 1994. Bare Phrase Structure. *MIT Occasional Papers in Linguistic Theory*, 5. (= Chomsky 1995a)
Chomsky, Noam. 1995a. Bare Phrase Structure. In *Government-Binding Theory and the Minimalist Program*, ed G. Webelhuth, 383-439. Oxford: Blackwell.
Chomsky, Noam. 1995b. *The Minimalist Program*. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 1995c. Language and Nature. *Mind* 104:1-61.
Epstein, Samuel D. 1992. Derivational Constraints on A'-chain Formation. *Linguistic Inquiry* 23: 235-259.
Epstein, Samuel D. 1995. Un-Principled Syntax and the Derivation of Syntactic Relations. Ms., Harvard University.
Frank, Robert and Anthony Kroch. 1995. Generalized Transformations and the Theory of Grammar. *Studia Linguistica* 49:103-151.
Frazier, Lynn & Charles Clifton. 1996. *Construal*. Cambridge, Mass.: MIT Press.
Goodman, Nelson. 1977. *The Structure of Appearance*. Dordrecht: Reidel.
Grewendorf, Günther. 1990.Verbbewegung und Negation im Deutschen. *Groninger Arbeiten zur Germanistischen Linguistik* 30:57-125.
Grewendorf, Günther. 1995. *Sprache als Organ - Sprache als Lebensform*. Frankfurt/Main: Suhrkamp.
Grimshaw, Jane 1995. Projection, Heads, and Optimality. ms. Rutgers University.
Haider, Hubert. 1991. Die menschliche Sprachfähigkeit - exaptiv und kognitiv opak. Kognitionswissenschaft 2:11-26.
Hale, Kenneth and Samuel Keyser, ed. 1993 *The View From Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Cambridge, Mass.: MIT Press
Halle, Morris and Alec Marantz, 1993. Distributed Morphology and the Pieces of Inflection. In Hale and Keyser, ed., 116-176.
Hoop, Helen de. 1992 Case Configuration and NP Interpretation. PhD dissertation, Groningen.
Huang, C.-T. James. 1982. Logical Relations in Chinese and the Theory of Grammar. PhD., MIT.
Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.
Kiparsky, Paul. 1982. Lexical morphology and phonology. In *Linguistics in the Morning Calm*, ed. the Linguistic Society of Korea, 3-91. Seoul: Hanshin.
Kitahara, Hisatsugu. 1995. Target α: Deducing Strict Cyclicity from Derivational Economy. *Linguistic Inquiry* 26:47-77.
Koster, Jan. 1987. *Domains and Dynasties*. Dordrecht: Foris.
Lakoff, George. 1970. Global Rules. *Language* 46:627-639.
Larson, Richard. 1991. *Promise* and the Theory of Control. *Linguistic Inquiry* 22: 103-139
Lasnik, Howard. 1994. Verbal Morphology: 'Syntactic Structures' Meets the Minimalist Program. Ms., University of Connecticut, Storrs.
Lasnik, Howard. 1995. Case and Expletives Revisited. *Linguistic Inquiry* 26:615-633.

Lasnik Howard & Mamoru Saito. 1984. On the Nature of Proper Government. *Linguistic Inquiry* 15. 235-289.
Lasnik, Howard & Juan Uriagereka. 1988. *A Course in GB Syntax*. Cambridge, Mass.: MIT Press.
Lebeaux, David. 1988. Language Acquisition and the Form of Grammar. Doctoral dissertation, University of Massachusetts, Amherst.
May, Robert. 1985. *Logical Form: its Structure and Derivation*. Cambridge, Mass.: MIT Press.
Müller, Gereon & Wolfgang Sternefeld 1993 Improper Movement and Unambiguous Binding. *Linguistic Inquiry* 24:461-507.
Pesetsky, David. 1982. Paths and Categories. PhD., MIT.
Prince, Alan and Paul Smolensky. 1993. Optimality Theory: Constraint interaction in generative grammar. Technical Report of the Rutgers Center for Cognitive Science, Rutgers University.
Rizzi, Luigi. 1990. *Relativized Miminality*. Cambridge, Mass.: MIT Press.
Rizzi, Luigi. 1991. Residual Verb Second and the Wh-Criterion. Technical Reports in Formal and Computational Linguistics 2., Université de Genève.
Rosenbaum, Peter. 1967. *The Grammar of English Predicate Complement Constructions*. Cambridge, Mass.: MIT Press.
Safir, Ken. 1986. Relative Clauses in a Theory of Binding and Levels. *Linguistic Inquiry* 17.4:663-689.
Schoemaker, Paul. 1991. The Quest for Optimality: A Positive Heuristic of Science?. *Behavioral and the Brain Sciences* 14:205-245.
Sperber, Dan & Deirdre Wilson. 1986. *Relevance: Communication and Cognition*. Oxford: Basil Blackwell.
Stabler, Edward. 1983. How are Grammars Represented? *Behavioral and the Brain Sciences* 6:391-421.
Wartena, Christian. 1994. Spelling Out the Minimalist Program. Doctoral dissertation, University of Nijmegen.
Wilder, Chris & Damir Cavar. 1994. Word Order Variation, Verb Movement and Economy Principles. *Studia Linguistica* 48:46-86.
Zwart, Jan-Wouter. 1993. Dutch Syntax: A Minimalist Approach. Doctoral dissertation, University of Groningen.