2 Economy Conditions in Syntax

CHRIS COLLINS

0 Introduction

An important theme in recent generative grammar is that linguistic opera-
tions, derivations, and representations are subject to economy conditions which
guarantee that they are optimal in some sense (see Chomsky 1998b).

Consider an operation OP applying in a derivation D leading to the repres-
entations (PF, LF) (phonetic form and logical form). Economy considerations
suggest that OP be as small as possible, and be applied in a way that minimizes
search. Given a series of operations that form a derivation D, economy condi-
tions suggest that the length or cost of the derivation must be minimized in
some way. Lastly, economy considerations suggest that the representations
formed in the course of a derivation should be as simple as possible, consist-
ing of a minimal number of syntactic objects, each of which is interpretable (at
either LF or PF).

The main purpose of this chapter is to give an overview of the economy
conditions that have been proposed for syntax and to discuss how the various
conditions are related. I will also present what I consider to be the outstanding
problems and interesting research issues.

This chapter is organized as follows. In section 1, I will discuss Last Resort.
In section 2, I will discuss Minimality. In section 3, I will discuss the Shortest
Derivation Requirement. In section 4, I will discuss timing principles, such as
ASAP (As Soon As Possible) and Procrastinate. In section 5, I will discuss the
issue of local versus global economy conditions. In section 6, I will discuss
economy of representation. Section 7 is the conclusion.

1 Last Resort

The principle of Last Resort is perhaps the most widely used economy condi-
tion. The most intuitive way to state the condition is along the following lines:

(1)   An operation OP may apply only if the derivation would otherwise result
       in an ungrammatical representation (at PF or LF).

A number of immediate questions arise about this formulation. First, are
all operations subject to Last Resort (e.g. Move, Merge), or just some limited
subset (e.g. the spelling out of resumptive elements, *do*-support)? Second, what
kind of ungrammaticality can sanction an operation to apply (a violation of
the Empty Category Principle (ECP), non-convergence, any type of violation)?
I will comment on both of these (largely unresolved issues) in the following
exposition.

This very general formulation describes a number of different phenomena,
investigated by a number of different authors, not all of whom use the term Last
Resort. Consider the case of *do*-support. Chomsky (1991) analyses *do*-support
as resulting from an interpretation of "least effort" such that language specific
operations are more costly than operations that are specified as part of Uni-
versal Grammar (UG). In particular, *do*-support is assumed to be a language
specific operation, and movement is assumed to be an operation specified as
part of UG.

Given these assumptions, consider the following data (I simplify Chomsky's
account for matters of exposition):

(2)   a.   John Infl Agr wrote books
       b.   *John [$_{Infl}$ did] Agr write books
       c.   *John Infl not Agr wrote books
       d.   John [$_{Infl}$ did] not Agr write books

Example (2a) involves the following steps: Infl lowers to Agr, and then Agr
lowers to V, and then V raises back to Infl at LF (via Agr). (2b) is blocked, since
"inserting" *do* in Infl involves a language specific operation that is more costly
than the lowering and raising found in (2a). Example (2c) is disallowed by the
ECP. Once Infl lowers to Agr and Agr lowers to V, V is blocked from raising
back to Infl at LF by negation. Example (2d) is allowed, because even though
lowering and raising is less costly than *do*-support, lowering and raising would
lead to a violation of the ECP. In other words, it is the ungrammaticality of the
derivation leading to (2c) that licenses (2d). In this example, there is a mechan-
ism of *do*-support, which steps in just when the derivation would have been
ungrammatical otherwise.

An alternative account of *do*-support that has recently gained some support
is based on adjacency (Bobaljik 1995, Lasnik 1995b). The basic intuition of this
approach is that in (2) above the inflectional head Infl contains an affix. This
affix may combine morphologically with the following bare verb under adja-
cency, as in (2a). If negation is present (as in (2c)), adjacency is blocked and the
dummy *do* is inserted to support the affix (as in (2d)). A natural way to block
do-support in the sentence (2b) is Last Resort. In other words, *do* can only be
inserted when the derivation would otherwise lead to a stranded affix.

Koopman and Sportiche (1986: 362, 366) claim that the distribution of resumptive pronouns, resumptive verbs, and adjunct extraction particles is determined by "minimalist strategies." They note "It is an often made observation that languages seem to adopt 'minimalist strategies' as unmarked strategies when possible; licensing processes are invoked only when necessary." From this point of view consider the following sentences:

(3) a. yEsO n   dIdO-dIdO suO   la
         how you cut-M-cut-M tree-det WH
         "How did you cut the tree?"
     b. n dI suO   fafa
         I cut tree-det quickly
         "I cut the tree quickly."
         (*dIdOdIdO)

(4) a. alO *(O)   nU mI la
         who he-resum did it   WH
         "Who did it?"
     b. yI   Kofi nU la
         what Kofi did WH
         "What did Kofi do?"

(5) nU Kofi ka   mI nU
     do Kofi Fut-Aux it do
     "Kofi will DO it."

According to Koopman and Sportiche, the reduplication of the verb seen in (3a) is needed to govern the trace of the extracted adjunct properly. If the adjunct has not been extracted, no such morphology is possible. Similarly, in (4a), a subject trace does not obey the ECP, therefore the trace must be spelled out. The trace of an object does obey the ECP, and so no resumptive pronoun is possible. Similar reasoning holds for resumptive verbs, seen in the predicate cleft construction in (5). Since the verbal trace would violate the ECP, an overt copy must be spelled out at the tail of the verbal chain.

Other authors that have a similar analysis of resumptive elements include Rizzi (1990), Ura (1996), Shlonsky (1992), and Pesetsky (1997: 168) on resumptive pronouns. For example, Shlonsky (1992: 443) states that "resumptive pronouns only occur as a Last Resort, when wh-movement fails to yield a grammatical structure."

With respect to the questions asked above, the operation of spelling out the trace of movement (in resumptive pronoun constructions, and in predicate cleft constructions) is considered to be an operation that applies only as a last resort. The principle that triggers the operation is the ECP.

Last Resort like principles are also possible to state in non-derivational frameworks. For example, in the Relational Grammar framework, the Motivated Chomeur Law is a kind of Last Resort principle. This principle states that

Cho arcs arise only when there would otherwise be a violation of Stratal Uniqueness (see Perlmutter and Postal 1983: 99). In particular, one of the consequences of this law is that there can be no spontaneous creation of a chomeur. Last Resort like principles are also used in phonology (for example, the analysis of epenthesis in Prince and Smolensky 1993).

The full range of cases falling under the general form of Last Resort in (1) is not known. But it is without a doubt true that natural language obeys something like (1). In section 5, I will return to the issue of whether the general form of Last Resort in (1) is global or local.

1.1   Last resort and inertness

An implicit assumption of much recent work is that once a syntactic relation is formed it cannot be altered. For example, if X dominates Y and Z, then it is impossible to change the dominance relations, so that later in the derivation X dominates only Y, or X dominates Y, Z, and W (see in particular Collins 1997; see also Watanabe 1995). In addition, if X is the head of a constituent Y, then it is impossible to change the head of Y, so that later in the derivation Z becomes the head of Y. If an empty element ec is a PRO at some step in the derivation, then it cannot become a A′-trace at some later step.

I will state this condition as follows:

(6)   Once a syntactic relation (or syntactic object) is formed, it cannot be changed.

I will call this general principle Inertness or Inalterability. Last Resort and Inertness seem to be two sides of the same coin. Last Resort gives the conditions under which an operation can take place. Inertness gives the conditions under which an operation cannot take place.

For example, consider the discussion of Case assignment in Babby (1987: 96): "a NP is assigned case only once and there are no rules that change one case to another." Babby calls this principle "inertness" and uses it to account for the fact that in Russian the object of a verb that assigns lexical Case (e.g. dative) cannot be passivized (since that would involve the NP being assigned nominative).

This principle has the effect of limiting the number of operations that can affect a particular element (say a NP), and so it seems justified to call it an economy condition. Inertness is obviously related to Chomsky's (1995b) analysis of feature checking, where an uninterpretable feature, once deleted and erased, is inaccessible to further operations. As another example, Chomsky (1998a) postulates that the sisterhood and c-command relations that a label enters into cannot be changed.

As another example, Richards (1998: esp. n. 32) postulates Principle of Minimal Compliance, which may also be a form of Inertness. For example, in his discussion of Subjacency, Richards states "it appears to be true quite generally that in cases involving multiple wh-movement to a single [+wh]

complementizer, only the first moved wh-word will have to obey Subjacency."
It appears that when a structure acquires the status of obeying Subjacency,
that status cannot be changed later in the derivation.

C-command does not seem to be subject to (6). Suppose X c-commands Y,
then Y moves to a position where it c-commands X. In this case, the relation of
c-command has been altered. The force of this counter-example to (6) depends
in part on interpretation of the copy theory of movement, and in part on
whether c-command is even a real syntactic relation.

The topic of Inertness has never been investigated in a systematic way. The
question is what syntactic relations cannot be changed, and which ones can be
changed, and what distinguishes the two sets.

1.2   *Last resort and movement*

In the Minimalist Framework (Chomsky 1995b), the discussion of Last Resort
has been mostly focussed on Move, which is a generalization of traditional
movement operations such as constituent questions, clefting, passive, and rais-
ing. The original insight that Move is subject to a Last Resort condition is due
to Chomsky (1986a: 143, 160). He claimed that Move-alpha is applied only
when a failure to apply it would lead to a structure that violates a grammatical
condition such as the Case filter. The main generalization that Chomsky wanted
to explain is what is often called the Chain Condition:

(7)   In an A-chain of the form $(\alpha_1, \ldots, \alpha_n)$, α_1 occupies its unique
      Case position and α_n occupies its unique theta-position. (modified from
      Chomsky 1986a: 137)

In particular, Chomsky was concerned about the part of the generalization
above that states that only the head of a chain (not the tail) is in a Case posi-
tion. The basic reasoning is that if a NP has had Case assigned to it, then it
would not have to move any more to satisfy the Case Filter.

Chomsky (1993, 1995b: 200, 201) picks up the notion of Last Resort again,
defining it as follows: "a step in a derivation is legitimate only if it is necessary
for convergence." Chomsky refines Last Resort to make it self-serving. This is
the condition of Greed (self-serving Last Resort), stated below:

(8)   Move α applies to an element α only if the morphological properties of α
      itself are not otherwise satisfied.

In this regard, consider the following sentence:

(9)   *John seems that [$_{TP}$ t is nice]
      "John seems to be nice."

The question is why *John* could not raise from the embedded subject position
to the matrix subject position. According to Chomsky, this is blocked by Greed,
since at the time of the raising, the Case feature of *John* is already checked.

The condition of self-serving Last Resort or Greed has been argued against most extensively by Lasnik (1995a) and Collins (1995, 1997). Lasnik argues that in a large number of cases no independent condition of Greed is needed. For example, Lasnik points out that (9) is ruled out by an independent principle that states that if T has an unchecked Case assigning feature, then the derivation crashes. Since the Case feature of *John* is checked in the embedded clause, it no longer can check the Case assigning feature of the matrix T. So no principle of Greed is needed to block (9).

Collins (1995, 1997) points out that successive cyclic movement, ECM, locative inversion, and quotative inversion all seem to demand a weakening of the condition in (8). First consider successive cyclic movement:

(10)   a.   [$_{IP}$ John to be [$_{SC}$ t nice]]
       b.   [$_{IP}$ John seems [$_{IP}$ t to be [$_{SC}$ t nice]]]

A structure involving successive cyclic movement is illustrated in (10b). In order to form this structure it is necessary, at some point in the derivation, to raise *John* to the intermediate Spec IP position (to satisfy what has traditionally been called the EPP, or extended projection principle). The problem is that this movement is not licensed by the condition in (8), since the morphological properties of *John* would have been satisfied if *John* had raised directly to the matrix Spec IP position (without passing through the embedded Spec IP). This would give rise to the following representation:

(11)   [$_{IP}$ John seems [$_{IP}$ to be [$_{SC}$ t nice]]]

This representation violates the EPP, but that is not relevant to the movement of *John*. In addition to this criticism, Collins (1997) also argues that (8) is inherently global (see section 5 below), because it refers to alternate derivations (this is what the word "otherwise" means).

Lastly, Collins (1997) argues that quotative inversion and locative inversion do not satisfy (8). Consider the following:

(12)   a.   "I am so happy", thought John
       b.   down the hill rolled John

Collins argues that the quotation in (12a) undergoes movement to the specifier of TP, but that this movement does not result in the checking of Case (since the Case feature of the quotation has already been checked). The movement is purely the result of the need to check the EPP feature of the matrix T.

Given these considerations, Collins (1995, 1997), Lasnik (1995a), and Chomsky (1995b, 1998a) all argue that (8) should be weakened to a condition like the following:

(13)   Move raises α to the checking domain of a head H with a feature F only
       if the feature F of H enters into a checking relation with a feature F of α.

(14)   F1 and F2 enter a checking relation iff F2 is in the checking domain of F1 and F1 is deleted (F2 may also be deleted).

The crucial aspect of this condition is that checking relation is defined *asymmetrically*. What this means is that movement of a constituent α does not have to result in the features of α being checked. We can illustrate this with the example in (10a). Assume, with Chomsky (1995b), that the infinitival *to* checks its EPP feature against the D feature of *John*. In that case, the movement in (10a) satisfies (13), since a checking relation has been established. Note that the checking relation is asymmetric, since the D feature of *John* has not been deleted.

   This account of Last Resort as it affects movement answers the questions under (1) above in the following way. First, all movement operations are subject to Last Resort. Second, only morphological feature checking can trigger movement, nothing else (such as the ECP, or another syntactic condition).

   Is the narrow form of Last Resort in (13) reducible to the general form of Last Resort in (1)? This is far from clear, and the issue has not been addressed systematically.

   The version of Last Resort in (13) takes the point of view of the head H that contains a feature F. This feature needs to be checked in order for movement operation to take place. It is in part for this reason that Chomsky (1995b: 297) reinterprets Last Resort as part of the definition of Attract (see section 2 for a similar remark about the Minimal Link Condition). Chomsky (1998a) also assumes that the narrow form of Last Resort in (13) is a part of the definition of applying an operation. In order to apply some operation OP (Agree or Merge), a probe or selector must be found, and then satisfied.

2   Minimality

Minimality states that given a choice between two comparable operations, the smallest is chosen. This way of stating Minimality is more general than the statements usually found in the literature (see Rizzi 1990, Chomsky 1995b). Looked at this way, Minimality bears a strong resemblance to the Shortest Derivation Requirement (see below) that says that the number of operations in a derivation should be minimized. The intuition behind both conditions is that a grammar tends to minimize whatever can be ranked along some scale: length of movements, steps in a derivation, or the number of violations of some condition (at least in OT syntax/phonology).

   Consider Chomsky's (1995b: 296) Minimal Link Condition:

(15)   Minimal Link Condition (Chomsky 1995b: 296)
       α can raise to a target K only if there is no operation (satisfying Last Resort) Move β targeting K, where β is closer to K.

To illustrate, consider the following example:

(16)   *John$_i$ seems that it was told t$_i$ that Mary left

In this example, the DP *John* raises from the embedded clause to target the matrix T'. In the embedded clause, it checks the Case and EPP feature of T. The problem is that the DP *it* is closer to the matrix T' than *John*, and so it blocks movement of *John* by the MLC in (15).

Chomsky's MLC (in (15)) takes the point of view of the head H containing the feature F that needs to be satisfied. The feature F' closest to F is the one that can enter into a checking relation with F. This is natural, as Chomsky points out, if the MLC is built into the definition of Attract (just as Last Resort was in section 1). Chomsky (1998a) develops this line of reasoning even further, proposing that the MLC is just one reflection of a general constraint that the search needed to apply an operation is minimized.

There are many other cases of Minimality conditions in syntax, which have just begun to be investigated. For example, Nakamura (1994, 1997) claims that extraction in Tagalog and Bantu applicative constructions is mediated by a type of Minimal Link Condition. Collins (1997) claims that binary branching in syntax is a result of Minimality. Richards (1997) claims that the order of *wh*-phrases in Slavic multiple *wh*-movement is the result of Minimality. Takahashi (1994a) attempts to derive the CED from the Minimal Link Condition. There are also several Minimality like conditions in phonology, including the EDGEMOST condition of Prince and Smolensky (1993).

One of the most exciting areas of syntax is the search for examples where Minimality like conditions play a role, and to unify these conditions. At this point in the search for phenomena, the largest possible net should be cast.

Since Rizzi (in this *Handbook*) takes up the issue of Relativized Minimality, I will not elaborate any further on these notions.

3   Shortest Derivation Requirement

The principle that the number of steps in a derivation be minimized is often called the Shortest Derivation Requirement (SDR), and it is stated below:

(17)   Minimize the number of operations necessary for convergence.

Note that there must be some way in which derivations are comparable in order for (17) to be useful. If any derivation could be compared to any other derivation, the zero derivation (with no operations) would always win. The set of comparable derivations is often called the reference set.

There are a number of different proposals as to how to define the reference set. Under one account, if derivations D1 and D2 are both convergent, and start from the same Numeration, then they are comparable (see Chomsky 1995b). Under another account, if D1 and D2 are both convergent, and they both lead to the same interpretations, then they are comparable (see Fox 1995).

Other researchers that have employed the Shortest Derivation Requirement are Kitahara (1997), Chomsky (1991, 1995b), and Epstein (1992). In particular,

Collins (1994a) uses the Shortest Derivation Requirement to rule out cases of chain interleaving.

Of all the economy conditions, the Shortest Derivation Requirement is the one with the least intuitive appeal. For example, consider a derivation with 18 steps. By the Shortest Derivation Requirement this derivation will be blocked by any comparable derivation (having the same Numeration) with 17 or fewer steps. This appears to be a case of the grammar counting, in a way that has long been held to be impossible. What grammars seem to be able to do well is to verify whether some simple condition holds of a representation or an operation (e.g. does the representation R satisfy condition A? Does Move X result in a feature being checked?).

To illustrate what is at issue, consider the following example that has been employed by Chomsky (1995b: 357) and Kitahara (1997: 19) to illustrate the Shortest Derivation Requirement. The following sentence illustrates object shift in Icelandic:

(18) a. Í gær   las Jón bækurnar ekki.
          yesterday read John the books not
          "Yesterday, John did not read the books."
      b. *Í gær   las bækurnar Jón ekki.
          yesterday read the books John not
          "Yesterday, John did not read the books."

The question is why the inverted structure in (18b) is unacceptable. Given the clause structure in Chomsky (1995b: 352), where there are no Agr projections, sentence (18a) would have the following structure (I leave out V-movement and V/2-effects for convenience of exposition):

(19)
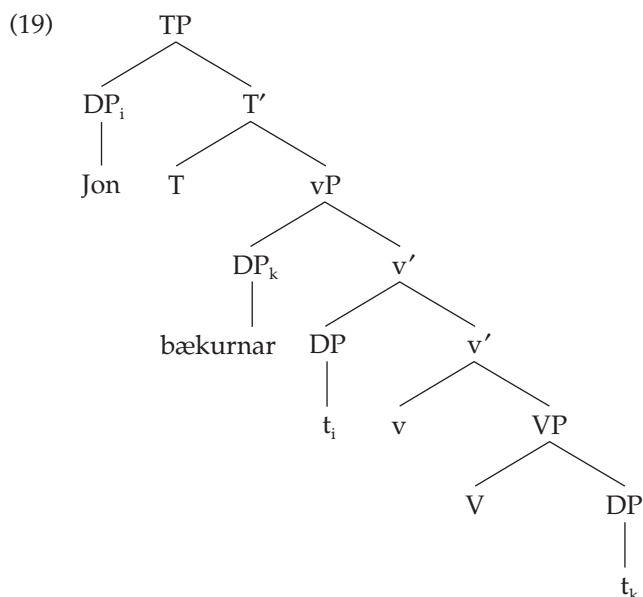
The derivation of this structure is as follows. The DP *bækurnar* "the books" has raised over negation (which is adjoined to v′), to the outer specifier of vP. In this position, *bækurnar* checks the accusative Case of v, and the EPP feature of v (the feature which Chomsky (1995b) assumes drives overt object shift). Second, the subject raises from the inner specifier of vP to the specifier of TP. In this position, the DP *Jón* checks the nominative Case and the EPP feature of T.

But there is a second possible derivation, the one which leads to (18b). Suppose after the object raises to Spec vP (checking the Case and EPP of v), it then raises to Spec TP (satisfying only the EPP feature of T). At this point, the only features left to check are the Case and phi-features of T and the Case feature of the DP *Jón*. Finally, the FF(*Jón*) (Case, D, phi features) raises covertly and adjoins to T at LF. Therefore, there is a good derivation of (18b).

Chomsky (1995b) proposes that this derivation is ruled out in the following way. Consider again the two relevant derivations:

(20)   Non-inverted derivation: (= 18a)
    a.   the DP *bækurnar* moves to Spec vP
    b.   the DP *Jón* moves to Spec TP

(21)   Inverted derivation: (= 18b)
    a.   the DP *bækurnar* moves to Spec vP
    b.   the DP *bækurnar* moves to Spec TP
    c.   the FF(*Jón*) moves to adjoin to T

Since the inverted derivation has one more step than the non-inverted derivation, and both converge, the inverted derivation is blocked by the Shortest Derivation Requirement.

Although inversion in (18b) is blocked by the Shortest Derivation Requirement, Collins (1997) argues extensively that the Shortest Derivation Requirement makes the wrong empirical predictions about inversion phenomena in English, in particular quotative inversion and locative inversion. Therefore, it is worthwhile asking whether there is not a different explanation of the unacceptability of (18b).

Chomsky (1998a) proposes that if the Case feature of a DP is checked, the DP becomes inactive and cannot undergo any further movement or participate in any further agreement relation. Such a condition would immediately rule out (21) with no further stipulation (a point made by Chomsky 1998a: 45).

In conclusion, we see that even examples that apparently show that the Shortest Derivation Requirement is needed can be explained in different ways.

One question that we have not addressed is the relation of the Shortest Derivation Requirement to Last Resort, stated in (1), and repeated below:

(22)   An operation OP may apply only if the derivation would otherwise result in an ungrammatical representation (at PF or LF).

Clearly the two principles are related. In many cases, both principles prefer a shorter derivation over a longer derivation. However, the two principles have a slightly different emphasis. We could imagine a scenario where the number of steps in derivation D1 is less than or equal to the number of steps in derivation D2, but D2 is chosen over D1 on the basis of the general form of Last Resort (1). This situation arises in the analysis of *do*-support as a last resort operation (see (2)). Also, the principle of Procrastinate below has the effect of making overt operations last resort operations, and choosing a derivation D2 over D1 even if they have exactly the same number of steps.

4   Timing Principles

Timing principles are principles that regulate when an operation can and must apply in a derivation. The two principles that I will discuss are Procrastinate and ASAP (As Soon As Possible).

4.1   Procrastinate

Perhaps the most famous timing principle is Chomsky's (1993) Procrastinate, which can be stated as follows:

(23)   Covert movement is less costly than overt movement.

According to this principle, overt movement is not allowed, unless the derivation would otherwise crash. A derivation crashes at some interface level (either PF or LF) if there is some feature present that is not interpretable at that level. For example, if a Case feature is present at LF, the derivation crashes. If a strong feature is present at PF, the derivation crashes (Chomsky 1993).

As pointed out above, Procrastinate can be interpreted as an instance of Last Resort condition (see (1)). An overt operation OP can apply only if otherwise the derivation would crash.

The principal motivation for Procrastinate is to describe the differences between languages as to whether movement occurs overtly or not. Consider the following two examples illustrating the difference between French and English:

(24)   a.   *John kisses often Mary
         b.   John often kisses Mary

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

Assuming that the adverb occupies some position (adjoined or specifier) between T and V, this paradigm can be described in the following way. In English, the verb cannot raise to T. In French, the verb must raise to T.

Chomsky (1993: 30) proposes to account for this paradigm in terms of feature strength. In French, the V feature of T is strong, and in English the V feature of T is weak. Chomsky makes the assumption that strong features are visible at PF, and weak features are not. He further assumes that strong features are not legitimate PF objects (unlike the feature [labial], or some other phonological feature). Therefore, if the strong V feature of T in French is not checked (or deleted), it will cause the derivation to crash. Thus the verb raises in French to ensure convergence.

Since the V feature of T is weak English, the verb does not have to raise. Chomsky proposed in addition that the verb is forced to stay in situ by Procrastinate, given in (23). Therefore, Procrastinate rules out (24a).

Chomsky also assumes that Procrastinate selects among the set of convergent derivations (1994: 428). This is necessary, since otherwise overt movement of the verb in (25a) would be blocked by the unacceptable (25b). Since the derivation leading to (25b) does not converge (since the strong V feature of T has not been checked), it does not compete with the derivation leading to (25a).

Procrastinate postulates a fundamental difference between overt and covert movements: overt movements are more costly than covert movements. Chomsky (1995b: 262) proposes that there is a more principled distinction that can be made between overt and covert movement.

Consider the following example. If a *wh*-movement takes place in order for the *wh*-feature of the *wh*-phrase to enter into a checking relation with the Q feature of Comp, then it is strange that the whole DP (including phonetic and semantic features) raises to Spec CP:

(26) $[_{CP} [_{DP}$ which book$]_i [_{C'}$ did John read $t_i]]$

Chomsky (1995b) suggests that the Q feature of Comp attracts only the *wh*-feature of the DP, but the rest of the DP (the phonetic and semantic features) is pied-piped, or carried along. The reason for this pied-piping is that separating the formal features from a word before Spell-Out would cause the derivation to crash.

To regulate this kind of generalized pied-piping, Chomsky postulates the following economy condition (see also Watanabe 1992a: 57, for similar remarks). I will call this condition the Weight Condition, since it is equivalent to saying that what is moved should be as small or as light as possible:

(27)   F carries along just enough material for convergence.

In the covert component (after Spell-Out), there is nothing to prevent a feature from being separated from its lexical item. Therefore, according to (27) only bare features (and not whole categories) can move in the covert component.

Chomsky (1998a) develops this line of thought even further, and proposes that there is no covert movement at all. Rather, there are agreement relations that are established (e.g. between a Q Comp and a *wh*-feature), some of which are accompanied by the overt movement of some constituent.

4.2 ASAP

As opposed to Procrastinate, there seems to be another condition in grammar which, following Yang (1997), we can call ASAP (see Pesetsky 1989 for a related condition). This is stated below:

(28)   If it is possible for an operation to apply, then it must apply. (See Collins 1999.)

The first question to ask is whether ASAP is actually an economy condition. In fact, ASAP looks like an anti-economy condition, the opposite of Procrastinate. But in one important sense ASAP is an economy condition, since it allows simplification in the computation needed to decide whether a particular operation applies. Suppose that OP can apply at a particular step S in the derivation. If UG permitted a choice between applying or not applying OP at step S, a decision would have to be made. Information that allowed the decision to be made would have to be found. It would be considerably simpler to adopt (28). Condition (28) is very similar to Chomsky's (1999) condition "Maximize matching effects."

Perhaps the most obvious instantiation of this principle has to do with checking. ASAP states that if a checking relation can be established, it must (see especially Chomsky 1995b: 280). Consider the following example:

(29)   [$_{IP}$ John$_i$ seems that [$_{IP}$ t$_i$ is nice]]

Chomsky (1995b: 284) and Lasnik (1995a) claim that this sentence is unacceptable for the following reason. Suppose that the NP John has a Case feature, and that finite Infl has a Case feature. At some point in the derivation, the following representation will be formed:

(30)   [$_{I'}$ seems that [$_{IP}$ John$_i$ is nice]]

At this step, the Case feature of *John* and the Case assigning feature of the embedded finite Infl enter into a checking relation, and they are both deleted. Because of this, *John* does not have a Case feature, and therefore, the Case feature of the matrix Infl can never be checked, even if *John* moves to the matrix Spec IP position as in (29).

This account crucially relies on the assumption that the checking of nominative Case in the embedded clause in (30) must be symmetric (see section 1.2 above for a discussion of asymmetric feature checking). Symmetric feature checking is forced by ASAP.

It is possible that ASAP is related to the "Disjunctive Principle" of Gruber (1973: 440). Gruber discusses the use of kinship terms in =Hoan (a Khoisan language of Botswana) and notes that the use of specific terms for a kinship relation always takes precedence over general terms. As Gruber states: "The disjunctive principle operates in the lexical attachment process to delimit the environment of applicability of a lexical item by that of another which is an inclusive subcategorization of it."

To see what the issues are, consider a similar example closer to home (roughly following Halle and Marantz 1993). We use the suppletive form *went* to the exclusion of the analytic form *go-ed*. Suppose at some point in the derivation we form $[V_{go} + Past]$, where V_{go} is a set of semantic and syntactic features that is spelled out as "go" in infinitival contexts. In order to spell out $[V_{go} + Past]$ we have a choice. The first choice is to spell out V_{go} as "go" and then Past as "-ed." The second choice is to spell out $[V_{go} + Past]$ as "went." The second option is preferred on the basis of Gruber's Disjunction Principle. It is also preferred on the basis of ASAP (or similarly maximize matching effects), since the first option does not spell out Past as quickly as possible. Note that the second option may also be preferred since it uses fewer morphemes (see section 6 below for a related principle).

The above explanation assumes that there is no economy condition that minimizes the number of features spelled out (or features checked) at a particular step in the derivation.

All three principles (ASAP, Disjunction Principle, maximize matching effects) are essentially economy principles. It is unclear for the moment what the empirical scope of each of the principles is, or whether any can be eliminated.

5   Global versus Local Economy

One issue that is becoming increasingly important in linguistic theory is the distinction between global and local theories. This distinction cross-cuts the derivational/representational distinction. Basically, a theory is a local theory if the decision about whether or not an operation OP may apply to a representation R is made only on the basis of information available in R (and not some other representation). Given this definition, Procrastinate is global. Consider again the example in (24). In order to determine whether V-movement must take place in English, it is necessary to see what happens at the PF interface. If not doing V-movement causes the derivation to crash, the V-movement is allowed.

A similar definition can be made for representational theories. We can say that a representational theory of syntax is local if the evaluation of whether a predicate P holds of R is restricted to information in R. For example, OT Syntax/Phonology is definitely global, since the evaluation of the predicate "R is optimal" crucially necessitates comparing R to other representations.

Procrastinate and the Shortest Derivation Requirement are global principles. Is the general form of Last Resort in (1) a global principle? The question is

whether the application of some operation needs to make reference to global information.

Consider again the issue of resumptive verbs in a predicate cleft construction (see (5) above). I claimed (following Koopman and Sportiche 1986) that spelling out the tail of the verbal chain is a last resort operation triggered by a potential violation of the ECP. Does the spelling out of the trace of the moved verb make reference to global information? This answer is not entirely clear.

Many researchers are now adopting the point of view that global conditions should be avoided (see Collins 1997, Chomsky 1998a, Poole 1995, Ura 1996, Yang 1997).

6 Economy of Representation

In most of the analyses described above, an operation or derivation is guaranteed to be minimal by an economy condition. We can now ask whether there is a condition on representations that they must be minimal in some way. Chomsky (1986a) proposes the condition of Full Interpretation that requires that each element in a PF or LF representation must be licensed. For example, Chomsky states that "at the PF level each phonetic element must be licensed by some physical interpretation." Chomsky (1995b: 219) articulates Full Interpretation further, stating that a representation meets Full Interpretation if it consists of legitimate objects that can be interpreted (at the interface levels).

Are there further economy conditions on representations over and above Full Interpretation? A good example is the treatment of finite and infinitival clauses in Bošković (1997b). Bošković gives the following condition on representations:

(31) **The Minimal Structure Principle**
      Provided that lexical requirements of relevant elements are satisfied, if two representations have the same lexical structure and serve the same function, then the representation that has fewer projections is to be chosen as the syntactic representation serving that function.

Bošković claims that the MSP can be used to rule out certain infinitival relative clauses. Consider the following example:

(32) a. *The man [$_{CP}$ Op$_i$ Comp [$_{IP}$ t$_i$ likes John]]
      b. *The man [$_{IP}$ Op$_i$ [$_{IP}$ t$_i$ likes John]]
      c. The man [$_{CP}$ Op$_i$ that [$_{IP}$ t$_i$ likes John]]

Bošković argues that the MSP chooses representation (32b) over (32a), since (32a) and (32b) have the same lexical items and (32b) has fewer projections. Note that (32c) is not blocked by (32b), since (32c) involves an additional lexical item "that" (which is nominal in nature) and is therefore not comparable with the other structures.

In addition, (32b) is ruled out by an economy condition that prohibits movement operations that are too short (at least one category must be crossed; a single segment of an adjunction structure does not count). Bošković calls this the "ban on superfluous steps."

Bošković notes that his MSP has an element of globality (since it compares two representations). To avoid this problem, Bošković gives a derivational account of the MSP. This leads us to the natural hypothesis that there are no economy conditions that compare two different representations.

There still may be economy conditions on representations, other than Full Interpretation? One possibility is the economy condition proposed by Emonds (1994: 162): "The most economic realization of a given deep structure minimizes insertions of free morphemes ('Use as few words as possible')." Emonds calls this "Economy of Derivation," but I think it is naturally construed as an economy condition governing the spelling out of representations. As such, it is not part of the syntactic computation (from the Numeration to LF), but rather a part of Spell-Out. Adopting a framework like that of Distributed Morphology and Gruber (1973), we can say that at Spell-Out, the lexical items and syntactic structures are provided with actual phonetic content.

There is a wide array of data that seems to support such a principle. For example, Emonds claims that his economy condition favors French *au* "to the" over *à le*. Emonds also claims that this economy condition chooses *bigger* over *more big*. There is an obvious similarity between Emonds's constraint and the Fewest Morphemes constraint of section 4.2, which I will highlight by calling Emonds's constraint the Fewest Free Morphemes constraint. Emonds's constraint might explain the fact that in pro-drop languages, expletive pronouns are obligatorily absent. Related conditions include Chomsky's (1981: 65) Avoid Pronoun principle, and Koopman's (1983: 175) Avoid Phonetics.

In addition, Emonds's constraint might be at the root of the generalization that resumptive elements (verbs or pronouns) are normally avoided. More generally, consider the chain (X, Y). Suppose that the head of the chain and not the tail is pronounced (e.g. John$_i$ was seen t$_i$). We can characterize this in one of two ways. We can say that the phonetic content of the tail of the chain was deleted at Spell-Out. Alternatively, we can say that the phonetic content of the tail of the chain was never provided at Spell-Out. Either way, spelling out the tail of the chain (as a resumptive pronoun, or full copy in predicate cleft) involves an extra word, and so is prohibited by Emonds's constraint.

One of the issues in finding economy conditions on representations is deciding how much of syntactic theory is derivational and representational (see Lasnik, this volume). The above discussion has assumed for the most part that syntactic structures are formed through a derivation, involving minimally some type of mechanism to build phrase structure (Merge) and some type of mechanism to capture syntactic dependencies (Move, Attract, Agree). The economy principles seem to be natural in this kind of framework.

However, the issue is not as simple as it appears at first. As mentioned above, representational theories often have economy like conditions. The

reason for this is that extra steps in a derivation usually correspond to extra symbols (or extra violations) in a representation. As a consequence, economy conditions on derivations can sometimes be restated as economy conditions on representations (e.g. minimize the number of symbols, or minimize the number of violations).

However, there are other reasons to adopt a derivational theory. For example, consider the c-command asymmetry involved in chains (the head of the chain c-commands the tail). This seems like a fairly robust empirical generalization and follows without stipulation on the derivational approach (see, for example, Collins 1997 or Chomsky 1998a). In fact, the derivational approach to syntax provides structures with a natural asymmetry (some parts are built before others), which is reflected in the many asymmetrical properties of chains.

7   Conclusion

I have tried to give an overview of economy conditions in syntax in this chapter. I have discussed Last Resort, Inertness, Minimality, the Shortest Derivation Requirement, Procrastinate, the Weight Condition, ASAP, maximize matching effects, Fewest Morphemes, the Minimal Structure Principle, the ban on superfluous steps, and Fewest Free Morphemes.

As representations are simplified to a greater degree, it will become more and more necessary to articulate the theory of economy conditions. For example, Collins (1999) attempts to eliminate labels on phrasal constituents (such as NP and VP). This simplification is only possible when the economy conditions are articulated in a certain way.

One point that has not been made in the preceding discussion is the importance of optionality for research into economy conditions. Suppose that at a particular step in the derivation either OP1 or OP2 can apply. Then it follows that OP1 and OP2 have the same cost associated with them. Therefore, optionality provides a direct probe into measuring the cost of operations (see Chomsky 1991).

The whole range of economy conditions and the interrelations between them have just begun to be charted. What can be said with certainty is that our understanding of economy at this point is minimal.

NOTE