SYNTAX

CRITICAL CONCEPTS IN LINGUISTICS

Edited by

ROBERT FREIDIN AND HOWARD LASNIK

SYNTAX

Critical Concepts in Linguistics

Edited by
Robert Freidin and
Howard Lasnik

Volume IV

Conditions on derivations and representations (1)

Routledge
Taylor & Francis Group

LONDON AND NEW YORK

Publisher's Note

References within each chapter are as they appear in the original complete work.

SYNTAX

ACKNOWLEDGEMENTS

Disclaimer

CONTENTS

CONTENTS

INTRODUCTION

This volume and the remaining two focus on the general principles of grammar that constrain the application of rules (and derivations more generally) and the representations they generate. Such principles have been the major topic of research in syntax since the late 1960s. The outstanding achievements of this line of research have been the Principles & Parameters framework of generative grammar, which came into focus during the latter half of the 1970s, and the more recent refinement of that framework under the critique of the minimalist program over the past twelve years.

The papers in this volume are concerned with two major subtopics: locality principles, which restrict the syntactic distance over which movement operations may displace constituents, and the related topic of cyclicity, which determines the way transformations may apply within derivations.

The earliest research on locality conditions occurs in Ross's landmark PhD dissertation, *Constraints on Variables in Syntax* (MIT 1967), excerpted in Chapter 56.[1] Ross begins with an examination of Chomsky's A-over-A Principle, first proposed in Chomsky 1964a.[2] He notes that Chomsky 1964a, footnote 10, mentions evidence that shows the A-over-A Principle as formulated is too strong because it rules out acceptable sentences, adding other examples involving pied-piping in the formation of relative clauses, of which (1) is one example.

(1)   the reports, the covers of which the government prescribes the height of the lettering on

The fronted NP *the covers of which* is itself extracted out of a larger NP *the height of the lettering on*, which violates the A-over-A Principle. Ross proceeds to identify further cases of deviant sentences that fall under the A-over-A Principle and to propose alternative principles to handle these cases.

In addition to proposing a Pied Piping Convention to handle the pied-piping cases, Ross formulates four other constraints on movement operations. The Complex NP Constraint (CNPC) prohibits extraction out of clausal constituents of NP. Thus in (2) the CNPC blocks the extraction of the wh-phrase *which politician* from the complex NP, *the article* modified by a relative clause.

(2) *[which politician]$_i$ did John read [$_{NP}$ the article [$_{CP}$ that mentioned t_i]]

The Coordinate Structure Constraint (CSC) prohibits extraction of any constituent that is or is contained within a conjunct of a coordinate structure. In (3) the CSC blocks the extraction of the wh-phrase object of the second VP conjunct. Thus in (3) extraction of the wh-phrase out of the second VP conjunct violates the CSC.[3]

(3) *What will John [$_{VP}$ [$_{VP}$ read the documents] and [$_{VP}$ write t_i]]

The Left Branch Condition (LBC) prohibits the extraction of an NP specifier of a noun, hence the leftmost constituent of an NP, out of that NP. Thus in (4) extraction of the wh-phrase specifier of *books* violates the LBC.[4]

(4) *Whose$_i$ did you read [$_{NP}$ t_i books]

The Sentential Subject Constraint prohibits extraction out of a clause that by itself constitutes the grammatical subject of a larger clause. This constraint excludes the construction of relative clauses in which the relative clause itself contains a sentential subject and the relative pronoun is extracted out of that clausal subject. This prohibits (5), where the relative pronoun *which* originates as the object of *brought* in the subject clause.

(5) *[$_{NP}$ The hat [$_{CP}$ which$_i$ [$_{IP}$ [$_{CP}$ that I brought t_i] seemed strange to the nurse]] was a fedora.

Ross notes that the Sentential Subject Constraint does not appear to be universal because it apparently does not hold for Japanese. In contrast, he designates the CNPC and CSC as universal principles. The status of the LBC is not commented on.

Ross uses these four conditions to define syntactic structures that he calls *islands* (see Ross 1986 for a more precise formulation). Thus islands are syntactic domains from which constituents may not be moved, and these four constraints and others like them have been dubbed 'island constraints.' The topic of syntactic islands has had a long and fruitful history – see especially Chapters 59 and 68.

Ross's discussion of island phenomena primarily concerns wh-movement and extraposition, the two rules discussed in the first chapter of his dissertation. His formulations of the structural description of these rules contain a variable between the launching site and the landing site. The constraints he formulates provide the necessary restrictions on what structures the variable could legitimately cover so that the rule could apply. By the mid 1970s, Ross's approach to the formulation of movement rules had been generalised to all movement rules. His work was the first to demonstrate

that the existence of such general constraints on movement operations made it possible to state movement transformations in a more general form. Ultimately rules would be stated as simple elementary operations (adjunction and substitution) without reference to the various specific contexts in which they applied in different languages. Thus formulating UG constraints that regulated the behavior of movement rules allowed for the formulation of the rules themselves as part of UG rather than as highly articulated, language specific rules containing UG elements.

Chomsky's 'Conditions on transformations' (1973, first draft 1970), excerpted in Chapter 57, constitutes a major extension of the work begun in Ross 1967. First, the conditions that Chomsky formulates are not limited to extraction operations, but are stated generally for any rules that 'involves' two constants X and Y. Second, there is explicit discussion of how these conditions cover NP movements – e.g. passive and raising[5] and generalise to rules of interpretation (e.g. a rule of disjoint reference for pronouns (called RI)). Thus consider Chomsky's formulation of the Tensed S Condition (TSC)[6]:

(6)   The Tensed-S Condition:
      No rule can involve X, Y in the structure
      $\ldots X \ldots [_\alpha \ldots Y \ldots] \ldots$
      where α is a tensed sentence.

This formulation covers insertion[7] as well as extraction and also interpretation. It mentions general structural considerations rather than specific constructions like coordinate structures and complex NPs.

In addition to the TSC, the Specified Subject Condition (SSC) is postulated to account for restrictions on rules involving a nonfinite clause boundary – i.e. blocking a rule from involving X and Y when Y is contained in an infinitival clause and is not the subject of the clause. Both the TSC and SSC apply when X and Y occupy A-positions. Given the assumption that these conditions apply to all rules, complicated formulations are needed to allow for 'long distance' wh-movement. Thus the paper explicates how COMP (now Spec-CP) provides an escape hatch from these conditions for wh-movement. NP movement cannot make use of this because of a COMP-to-COMP condition on movement such that once a phrase is moved into COMP it can only move to another COMP position (see Ch. 79, Vol. V on how this restriction is derived from binding theory). Using wh-movement phenomena, the paper also clarifies the notion of the syntactic cycle by proposing a more rigorously constrained notion of cyclicity, the Strict Cycle Condition (SCC), which prohibits operations from applying within cyclic subdomains of the cyclic domain currently being processed (see Ch. 66 for detailed discussion). To motivate the SCC, the paper employs sentences involving multiple wh-phrase questions – constructions that are

characterised as wh-island violations where a clause containing a clause initial wh-phrase is considered to be a syntactic island (see Ch. 59 for further discussion). The paper also postulates the Superiority Condition to account for the fact that in a multiple wh-phrase question containing only bare wh-pronouns, only the highest wh-phrase can move to COMP. The Superiority Condition supplements the SCC in that it covers the case where only one clause is involved and therefore cannot constitute a wh-island violation. The paper also postulates the Subjacency Condition, also formulated in terms of cyclic domains – i.e. that no movement rule can involve X, Y where the number of cyclic domains (bounding categories) in which X and Y are constituents differs by more than one such domain (see Chs 36, Vol. II, 59 and 60 for further discussion of which syntactic domains are relevant to the application of the Subjacency Condition). It follows from the Sub-jacency Condition that wh-movement must be successive cyclic through c-commanding COMP positions. As Chomsky notes, this condition covers many examples that are also covered by the CNPC.

Each of the conditions postulated in Chomsky 1973 opened up a line of research that is still being actively pursued in current work.

The successive cyclic application of movement transformations was not universally accepted when it was first proposed. Postal's 1972 article 'On some rules that are not successive cyclic' (Ch. 58) offers two arguments against the proposal that wh-movement in both questions and relative clauses and the movement that topicalises a constituent by moving it to the front of a clause (what he calls U-rules) cannot be successive cyclic because that analysis creates insurmountable empirical problems. The first argument concerns the possibility that a wh-phrase object of prepositional phrase can optionally be moved with its preposition. Postal argues that this allows a derivation to pied-pipe the preposition on one cycle and strand it on the next, thereby generating a deviant structure, as in (7).

(7)   *$[_\alpha$ who do you think $[_\alpha [_{PP}$ to t] (that) John gave a new car t']]^8

Thus (7) constitutes the empirical evidence for the 'preposition dangle' argument against successive cyclic movement.[9] The second argument concerns the derivation of constructions like (8), which involve both wh-movement and so-called *tough*-movement.[10]

(8) $[_\alpha$ what did you think $[_\alpha$ would be difficult $[_\alpha$ for Melvin to find]]]?

Postal claims that successive cyclic movement on the lowest clause would bleed the application of *tough*-movement on the intermediate clause. Exactly what the problem here is is not made clear – in part because no explicit analysis of the putatively problematic derivation is given. Adopting an alternative analysis of *tough*-constructions where the subject of

the *tough* predicate is externally merged and the predicate's complement clause contains a covert object (e.g. Chomsky 1977 or Chomsky 1981) would avoid the problem Postal is citing. Postal's argument is one example of the weakness of empirical arguments against general principles (e.g. successive cyclic movement imposed by Subjacency). The crucial constructions are usually susceptible to alternative analyses that do not create the problem.

Rizzi's 1980 article 'Violations of the *wh*-island constraint and the subjacency condition' (Ch. 59) is particularly concerned with the way in which the Subjacency Condition provides an account for wh-island phenomena. Rizzi notes that in Italian it appears that clauses with a wh-phrase in Spec-CP do not always function as wh-islands with respect to wh-movement. The paper gives a detailed analysis of the different cases, showing that rather than providing counter-evidence to the Subjacency Condition, the data of Italian supports such a condition if the bounding categories for Italian include CP (S-bar) rather than IP (S). In footnote 25 the paper gives one of the first demonstrations of how a UG principle could be parameterised. Thus the Subjacency Condition, restricting to one the number of bounding categories that could separate a launching site from its immediate landing site, still holds. However, what the bounding categories are for various languages can differ within a restricted range.

The study of linguistic variation in terms of theory of general principles of grammar in conjunction with parameters (beginning in the mid 1970s) led directly to the Principles & Parameters framework (P&P), which has been the standard framework for generative grammar ever since. Chomsky's 1981 paper 'Principles and parameters in syntactic theory' (Ch. 60) provides a condensed but clear explication of P&P[11] as it had developed at the end of the 1970s.

The paper begins with a sketch of the standard framework and then discusses how recent developments, starting with Ross's 'very important work' (see Ch. 56), leads to the formulation of Move α as a single transformation for all displacements. It then demonstrates how the new empirical problems created by this formulation can be solved at the level of UG via general principles – in particular, Subjacency, the Empty Category Principle (ECP), and the Case Filter.

The paper investigates a new formulation of Subjacency where both S and S-bar (i.e., IP and CP) constitute bounding categories.[12] Chomsky demonstrates that Subjacency applies to movement but not rules of construal with the contrast in (9), where the movement in (9a) results in deviance whereas the binding in (9b) is acceptable.

(9) a. *who_i did the man expect that [$_{NP}$ pictures of t_i] would be on sale
     b. The men_i expected that [$_{NP}$ pictures of each $other_i$] would be on sale.

The paper also discusses parametric variation for Subjacency, based in part on Rizzi's discussion in the previous chapter.

Chomsky goes on to discuss that extraction of the subject of a finite clause from a wh-island is a stronger violation (i.e. more deviant) than an extraction of a constituent from the predicate in a wh-island. What accounts for this difference is that the former violate the ECP, formulated as a requirement that a trace must be lexically governed, as well as Subjacency (for English). While the ECP rules out such examples in French and English, the corresponding examples in Italian are not deviant. Chomsky discusses Rizzi's solution which proposes free subject inversion in Italian (but not in English or French), hence another parametric difference.[13]

The notion 'government' links the ECP and the Case Filter, the requirement that all phonetically realised NPs and wh-trace must have Case (see Chs 69–72, Vol. V). Case is assigned in most instances under government by a lexical head (e.g. V or P). The Case Filter is inviolable, but associated with a parameter – formulated in this chapter as S-bar deletion, thus allowing a V to non-finite complement clause by a matrix V. The S-bar deletion option is triggered by certain specific lexical items in certain languages.

This chapter also reformulates opacity conditions on proper binding (i.e. the Nominative Island Condition of Chomsky 1980 and the SSC) in terms of government. It discusses how these conditions apply to traces, lexical anaphors, and PRO.[14] It also covers how opacity conditions apply for disjoint reference in the construal of pronouns and referential expressions. The reformulation of binding conditions in terms of government yields the PRO Theorem, from which it follows that PRO cannot occur in a lexically governed position.

Lasnik and Saito's 1984 paper 'On the nature of proper government' (Ch. 61) refines the ECP analysis that makes the Move α proposal viable. This chapter is primarily concerned with the definition of 'proper government', the level(s) of representation at which the ECP applies, and its relationship with other principles. Lasnik and Saito conclude that (in agreement with Chomsky 1981) there are two subcases of proper government: lexical government (whose core case is the head-complement relation) and antecedent government (where a trace is coindexed by a c-commanding antecedent but not separated from the antecedent by an intervening barrier (NP or CP)).

They propose that the ECP is a filter that applies to LF representations. The ECP filters out representations containing a trace in an A-position that is not marked by a feature $[+\gamma]$. The γ feature is assigned obligatorily at S-Structure and LF under certain conditions, $[+\gamma]$ being assigned to a trace that is properly governed. γ-assignment applies after the optional rule of Affect α, a formulation that subsumes Move α and Delete α. Furthermore, movement optionally leaves a trace. The apparent obligatoriness of Affect α in certain situations follows from independent principles of grammar,

including the ECP, the Projection Principle, and the prohibition against vacuous quantification. Subjacency applies only in the derivation from D-Structure to S-Structure, but not from S-Structure to LF.

This analysis demonstrates how the distribution of wh-trace both in S-Structure and LF follows from the optimally general formulation of transformations as Affect α in conjunction with general principles along with parameters.

The notion of proper government is also exploited in the Constraint on Extraction Domains (CED) of Huang 1982, which provides a general account for the impossibility of extraction out of constituents properly contained in the subject of a finite clause, an adjunct phrase, or a topic. Huang notes that all three contexts share a common property: they are not properly governed. The CED prohibits extraction out of constituents that are not properly governed. Given that both the CED and the ECP are formulated in terms of proper government, the question arises as to whether there is some way of collapsing the two constraints into a single constraint.

Browning's 1989 paper 'ECP≠CED' (Ch. 62) addresses this issue and provides a negative answer using intricate data to check various assumptions. The contrast in (10) constitutes an empirical focal point (examples (7a–b) in the paper).

(10) a. *[t_i dormire], Gianni$_i$ sembra.
              to sleep Gianni seems
      b. [PRO$_i$ dormire], Gianni$_i$ vorrebbe.
              to sleep Gianni would wish

Adopting Lasnik and Saito's γ-marking approach to the ECP this paper argues that NP trace in Italian must be subject to the ECP, thus the trace in (10a) but not the PRO in (10b) violates the ECP. The analysis assumes that the CED is a condition on derivations, not representations, therefore the point in the derivation at which *Gianni* raises will not violate the CED.

The proper characterisation of government domains is central to the correct formulation of the ECP. Rizzi's 1990 monograph *Relativised Minimality* is concerned with the role of minimality in the definition of these domains. Given the two kinds of government, head government and antecedent government, there arises a question about how they interact to determine the appropriate government domain for the ECP. In the first chapter of the monograph 'Opacity effects on adjunct variables' (Ch. 63), Rizzi contrasts two approaches to minimality: rigid minimality, which allows head government to block antecedent government, and relativised minimality, under which only governors of the same kind can block government via minimality. He argues that under relativised minimality as formulated in (11) the ECP provides a unified treatment of the phenomena identified in (12).

(11) Relativised Minimality (RM): X α-governs Y if no Z, Z a typical potential α-governor for Y, intervenes between X and Y.

(12) a. no adjunct extraction out of wh-islands
     b. VP initial adverbial QP selectively blocks extraction from certain VP internal elements
     c. adverbial elements cannot be extracted from the scope of negative operators (inner islands, Ross 1984)

RM covers wh-islands (A-bar chains), super raising (A chains), and head movement (X° chains). The chapter contrasts two formulations of the ECP, both involving antecedent government but differing on whether the second type is lexical government or theta government. It extends the discussion of RM to pseudo opacity constructions and other inner islands.

The following Chapter (64), the second chapter of Brody's monograph *Lexico-Logical Form* entitled 'Multiple Wh-/Neg-relations and subjacency', is based on the radical premise that there are no movement derivations, hence all constraints on Move α, including RM, are to be construed as conditions on representations (i.e. on chains). This construal conflicts with the analysis of Subjacency as a condition on derivations discussed in Lasnik and Saito (Ch. 61). The bulk of the chapter is devoted to showing that the representational approach to Subjacency[15] is viable and that the derivational approach is problematic.

In addition to Subjacency and the ECP, the syntactic cycle provides an additional general constraint on how rules may apply in derivations. The basic notion of the cycle is that rules apply to the smallest syntactic domains first and then to successively larger domains. The SCC (Chomsky 1973, §3 (Ch. 57)) prohibits the application of transformational operations internal to a cyclic subdomain of the current cycle.

The original formulation and arguments for a syntactic cycle are given in Chapter 65, Chomsky's third chapter of *Aspects of the Theory of Syntax* (1965), 'Deep structures and grammatical transformations'. In the original formulation of transformational grammar (Chomsky 1955), subordinate and coordinate sentential structures are derived transformationally via generalised transformations that combine multiple phrase markers. This approach allowed for the possibility that an embedding operation could be dependent on the prior application of a transformation to the clause in which another clause would be embedded. However no empirical evidence supported this possibility. To account for this situation, Chomsky introduces sentential recursion in the phrase structure rules, which now replace generalised transformations, and proposes the syntactic cycle as a general constraint on how transformations process the generalised phrase markers created by the new phrase structure rules, those allowing recursion of sentential structures.

The syntactic cycle as formulated in the SCC focuses on the constraint against applying a transformation within a cyclic subdomain of the current cycle. From this constraint it follows that transformational derivations must proceed bottom-up from the smallest to the largest cyclic domains. Freidin's 1978 article 'Cyclicity and the theory of grammar' (Ch. 66) investigates the empirical motivation for the SCC from the perspective of trace theory. The paper demonstrates how the empirical effects of the SCC can be derived from other independently motivated principles of grammar. For example, the SCC rules out the following wh-island violation.

(13)   *What did you wonder to whom John gave?

One derivation of (13) violates no known movement constraints and yet yields deviant output. Thus the SCC appeared to be necessary to rule out this derivation. Under trace theory, however, this derivation yields the same output as a derivation that violates some other constraints on movement (e.g. the Subjacency Condition). While the SCC can only be interpreted as a condition on derivations, Subjacency can also be interpreted as a condition on representations – more precisely, a locality condition on trace binding. Under this interpretation, the SCC becomes superfluous. Freidin 1978 generalises this result to other cases, not considered in Chomsky 1973, whose derivations also violate the SCC. This chapter demonstrates how most wh-movement cases could be handled by Subjacency,[16] whereas NP-movement cases involve other conditions.[17] By deriving the empirical results of the principle of the syntactic cycle from other independently motivated general principles of grammar, this chapter demonstrates how the cyclicity of the computational system is built into the general architecture of UG.[18] This result also raises the issue of the derivational vs. representational interpretation of general principles and provides an argument for the latter interpretation, given that some violations of the SCC also violate Subjacency interpreted as a condition on representations.

The intrinsic connection between Subjacency and the syntactic cycle also figures prominently in the following chapter, Torrego's 1984 paper 'On inversion in Spanish and some of its effects'. It follows from Subjacency that wh-movement must be successive cyclic. This chapter provides empirical evidence for successive cyclic wh-movement in Spanish concerning subject inversion, where the subject shows up in PF to the right of VP. Such inversion occurs obligatorily in a clause in which wh-movement occurs. According to Torrego, such inversion is also obligatory in embedded clauses through which a wh-phrase moves successive cyclically. The wh-phrase in Spec-CP triggers the obligatory inversion, even when the wh-phrase moves to a higher position. Inversion is however optional in the clause from which the wh-phrase first moves. This provides evidence that CP rather than IP is the relevant bounding category for Subjacency in Spanish – cf.