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 II

Conditions on derivations and representations (2)

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

VOLUME V CONDITIONS ON DERIVATIONS AND
         REPRESENTATIONS (2)

CONTENTS

INTRODUCTION

Continuing the focus on UG principles begun in the previous volume, this volume covers three distinct but interrelated topics: Case theory, θ-theory, and binding theory. Chapters 69–72 deal with Case theory; Chapters 73–75 with θ-theory; and Chapters 76–83 with binding theory.

One of the most. important breakthroughs in the development of the principles and parameters framework (see Ch. 60, Vol. IV) occurs in a hitherto unpublished letter from Jean-Roger Vergnaud to Noam Chomsky and Howard Lasnik (published here as Ch. 69[1]) regarding the manuscript of their 1977 paper 'Filters and control'. To account for certain properties of infinitival constructions in English, Chomsky and Lasnik had proposed the following surface filter (1):

(1) *[$_\alpha$ NP *to* VP], unless α is adjacent to and in the domain of:
   a. [–N] (adjunct)
   b. [+V] $\varnothing$

(1) applies to indexed NPs – i.e. lexical (phonetically realised) NPs and trace, but not PRO. Vergnaud suggests replacing (1) with a filter concerning the licensing of 'Governed Case' (as opposed to nominative and genitive), which he suggests 'looks very much like a principle of UG'. He assumes that NPs are assigned abstract Case, even when Case has no overt morphological realisation. Vergnaud explores two alternative theories, one in which only phonetically realised NPs are marked for Case and another with includes traces.[2] This constitutes the first proposal of a UG principle to be formulated in terms of Case and government and leads directly to the formulation of the more general Case Filter (Chomsky 1980, see Ch. 70). The focus on concepts of Case and government introduced in Vergnaud's letter opened up extremely fruitful lines of research, still ongoing.

Chapter 70 contains an excerpt from Chomsky's 1980 article 'On binding', which begins with some conceptual and empirical problems with the NP-*to*-VP Filter of Chomsky and Lasnik 1977. The discussion starts with the observation that the obligatory control property of the construction (2) follows from the NP-*to*-VP Filter.

(2) [$_{S'}$ wh-phrase [$_S$ NP *to* VP]]

1

Chomsky then raises as conceptual problems for the filter why PRO is excepted, why the context [–N]__ is excepted under command and adjacency, and that the filter picks out, apparently redundantly, infinitival clause subjects, which are also picked out by the combination of the Specified Subject Condition and the Tensed S Condition (and its descendants). He also raises a pair of technical problems for the filter analysis. Nonetheless, as Chomsky shows, the filter, which is independently motivated by other constructions, automatically accounts for the properties of infinitival relatives without additional apparatus or stipulation – thereby providing a real explanation.

Chomsky considers two alternative approaches to the analysis of infinitival relatives, one involving a theory of deletion and the other, the now famous Case Filter. However, the deletion analysis involves problems that the Case analysis solves in addition to resolving the conceptual and technical problems noted for the NP-*to*-VP Filter analysis. In this excerpt, Chomsky also suggests that the lack of preposition stranding via NP movement follows from a general prohibition against Case conflict (the assignment of more than one Case to an NP).

Chapter 71, the section on Case theory from Chomsky and Lasnik's 1993 article, 'The theory of principles and parameters', provides an overview of the concepts and issues of Case theory over a decade beyond the birth of the theory. A basic assumption is that Case is required for morphological realisation, a requirement that motivates NP movement. However, the morphological realisation approach to Case is problematic. One problem concerns wh-trace, which must be marked for Case but is not morphologically realised (see Lasnik and Freidin 1981 for discussion) Related to this, the Case filter clusters phonetically realised NP, variables and pro (the null pronominal subject (e.g. in Spanish and Italian)), which form an unnatural class in terms of phonetics and morphology.

As a solution, Chomsky and Lasnik explore linking the Case Filter to θ-theory[3] (see Chs 73–75), where phonetically realised NP, variables and pro are arguments (as opposed to NP-trace). On this view, the Case Filter forms part of the principle of θ-marking. An argument must be visible for θ-role assignment to occur and Case renders it visible. This visibility approach creates a problem for the analysis of PRO, which apparently appears in a Caseless position, but is nonetheless an argument that receives a θ-role. Chomsky and Lasnik propose that PRO has a phonetically null Case and thus behaves just like phonetically realised NPs – Last Resort blocks the movement of Case-marked PRO and visibility forces the movement of Caseless PRO to a Case-marked position.[4]

This section also considers the related phenomena of agreement, where PRO involves both Case and agreement. Chomsky and Lasnik propose that null Case is assigned via the spec-head relation where the head I (*to* or *ing*) has null agreement. They consider generalising all Case assignment

to the spec-head relation, thus eliminating the head-complement relation from Case assignment. This analysis requires a modification in the formulation of Last Resort.

Chapter 72, Freidin and Sprouse's 'Lexical case phenomena' (1991), focuses on the analysis of languages that have rich morphological Case systems (in contrast to English). Such languages (e.g. Russian and Icelandic) manifest two distinct types of morphological Case: configurational (structural) and lexical (a.k.a. quirky or inherent). Configurational Case is assigned solely in terms of syntactic positions, whereas lexical Case is assigned via selection by a specific lexical head (where different heads of the same category may select different lexical Cases). Given that θ-role assignment is also mediated by selection, lexical Case assignment via θ-marking as suggested in the previous chapter seems plausible. In constructions where configurational and lexical Case assignment could be in conflict (e.g. for the object of a verb that assigns lexical Case) the lexical Case requirement must be morphologically realised and thus the configurational Case is morphologically suppressed. This follows from the Principle of Lexical Satisfaction of Freidin and Babby 1984, the ramifications of which are investigated in this article. Lexical Case phenomena establish a distinction between Case assignment and Case licensing. In a clause whose main verb selects a lexically Case-marked subject, that subject must occur in a position that is configurationally licensed for Case. It is necessary but not sufficient that the subject bears the appropriate lexical Case. Thus, a phonetically realised NP must be Case-licensed as well as Case-marked. For configurational Case, licensing and marking appear to be indistinguishable, whereas for lexical Case these are distinct processes, which has consequences for the formulation of the Case Filter.

Chapter 73 contains three sections of Chomsky's *Lectures on Government and Binding* (1981) concerning LF representation and θ-theory (§§2.2 and 2.6) and a related discussion about the base (§2.5). The focus of these sections concerns the mapping of θ-roles on arguments in LF representations. They assume that LF representation is a form of syntactic representation derived via syntactic rules, the null hypothesis. Thus, the LF representation of a sentence like (3a) will be something along the lines of (3b) rather than (3c).

(3) a. John seems to be sad.
     b. John$_i$ [$_{VP}$ seems [$_S$ t$_i$ to [$_{VP}$ be sad]]]
     c. seems (sad (John))

θ-roles are assigned to complements (i.e. a position that satisfies a subcategorisation feature of a head) and also some subject positions. The embedded clause [$_S$ t$_i$ to [$_{VP}$ be sad]] is assigned a θ-role as complement of *seem* and the trace of *John* is assigned another θ-role by the complement predicate *be sad*.

In general, θ-role assignment (θ-marking) is mediated by the θ-Criterion (4).

(4) Each argument bears one and only one θ-role, and each θ-role is assigned to one and only one argument.[5]

Given the θ-Criterion, if subcategorisation entails θ-marking, then (i) there can be no rule of raising to object position given that objects are complements and therefore argument positions[6] and (ii) the variety of lexical items is restricted. The θ-Criterion also suggests that the Projection Principle, the initial formulation of which was restricted to subcategorisation properties of lexical items, be formulated in terms of θ-marking (i.e. the lexical property of selection) such that representations at each syntactic level (D-Structure, S-Structure and LF) are projections of lexical properties.[7] Thus, at D-Structure each argument occupies a θ-position and each θ-position is occupied by an argument. The remainder of §2.2 considers the interrelation of the two factors for determining θ-roles: intrinsic lexical properties and grammatical functions (subject, object, etc.) and spells out the consequences of the θ-Criterion and the Projection Principle.

Section 2.5 focuses on the base component, beginning with a close examination of the categorial component of the lexicon. It compares stipulating properties of the lexicon with deriving these properties from general principles – e.g. using Case theory to account for the non-occurrence of VPs with the order V-S-NP. It addresses the fundamental question of which properties are explicable on general grounds versus which are idiosyncratic to the language in question. As a way of clarifying the strengths of the standard approach, Chomsky compares it to various alternative approaches.

Based on considerations raised in §2.5, §2.6 returns to an examination of LF representation. Taking θ-roles to be assigned to R-expressions, overt anaphors, pronominals and clauses, Chomsky elaborates how θ-marking works for a variety of clause types (including small clauses (see Ch. 9, Vol. I)). The conclusions Chomsky draws follow from his 'decision to keep to minimal assumptions'. Chomsky goes on to review the consequences of assuming that the θ-Criterion and the Projection Principle hold without exception. He considers the partial overlap between Case theory and the θ-Criterion: 'a verbal element assigns Case to an NP that it governs if and only if it assigns a θ-role to its subject'[8]; and mentions the visibility approach (discussed in Ch. 71) as a possible solution. Chomsky also discusses the problem of wh-phrases, where they are treated as R-expressions when they enter the derivation, but not when they move to Spec-CP where they have no θ-role. That is, a variable (versus a NP-trace) does not transmit a θ-role.

Chapter 74, Belletti and Rizzi's 'Notes on psych-verbs, θ-theory, and binding' (1988) continues the discussion on the mapping of thematic

representations (θ-grids) onto initial syntactic configurations. Belletti and Rizzi are concerned with the situation where the same θ-grid (e.g. (experiencer, theme)) is projected on a variety of syntactic configurations. For example, in (5a) the experiencer occurs as the subject of the clause whereas in (5b) it occurs in the predicate.

(5) a. Gianni teme l'inflazione
         Gianni fears inflation
     b. La musica piace a tutti
         The music pleases to everyone

They argue against an approach that assumes such situations simply reflect lexical idiosyncrasies by constructing an analysis in which θ-marking operations apply uniformly according to syntactic representation.[9] Thus, theme is directly θ-marked by V whereas experiencer is compositionally θ-marked by V' (V + theme). The experiencer can therefore be θ-marked either when inserted directly in subject position as in (5a) or when it occurs in VP as in (5b). When theme shows up as the S-Structure subject of an experiencer verb (in Italian), this results from movement out of VP.[10] In support of this analysis, they show that the theme subject in these constructions shares the properties of derived subjects rather than D-Structure subjects by considering binding of reflexive clitics, the distribution of arbitrary pro subjects, the nonoccurrence of embedded clauses with derived subjects in causative constructions, and the non-occurrence of passivisation in clauses with non-thematic subjects. As further empirical support for their D-Structure analysis, they consider how binding theory applies to constructions containing psych-verbs. Their analysis yields an interesting asymmetry between Principle A and Principles B and C regarding where in the derivation they apply. It also yields a general observation that, when a θ-grid contains both a theme and an experiencer, the experiencer must be merged higher than the theme.

The notion that θ-roles are assigned at a level of D-Structure has been a central part of much of syntactic theory for decades, since the inception of θ-role analysis in Gruber 1965 (see also Jackendoff 1972). However, since the formulation of the minimalist program in Chomsky 1993 (Ch. 90, Vol. VI), a level of D-Structure has been rejected as inconsistent with minimalist assumptions. Chapter 75, Bošković's 'D-Structure, theta-criterion, and movement into theta-positions' (1994), provides an extensive critique of the level of D-Structure[11] and hence the necessity and utility of the Projection Principle and the θ-Criterion. Bošković claims that the two principles are trivial at LF because a linguistic expression will receive a deviant interpretation if they are not satisfied. He discusses *tough*-constructions and double object constructions as problems for the notion of D-Structure as a pure representation of thematic properties. Given a level of D-Structure

along with the Projection Principle and the θ-Criterion, it is easy to account for the ban on movement into θ-positions. Bošković argues that without a level of D-Structure the ban can be derived, but it does not follow from a θ-Criterion now applied to LF. Instead he reanalyses movements into θ-positions as violations of the economy of movement plus the EPP. His strategy is to reduce the θ-Criterion to (6), which differs from that of (4) by replacing 'one and only one' with 'some' in both parts.

(6) a. Every θ-role must be assigned to some argument.
     b. Every argument must be assigned some θ-role.

Bošković claims that (5) 'holds by virtue of interpretation'. Other restrictions on θ-structures (e.g. the ban on movement to a θ-position) must follow from other principles. The third part of the paper deals with empirical problems concerning the θ-Criterion and a level of D-Structure. In particular, Bošković argues that movement to a θ-position is viable for certain constructions – specifically, those involving Romance restructuring verbs and English root modals. Thus, the ban on movement to a θ-position is only partial; it applies when such movements violate economy conditions on movement or the EPP, which the derivations he proposes for Romance restructuring constructions and English root modals do not.[12]

Like θ-theory, binding theory constitutes another module of UG that is intrinsically linked to issues of interpretation. It is concerned with the syntax and interpretation of anaphoric elements (e.g. pronouns, reflexive pronouns, reciprocal expressions). The early work on this topic, from Lees and Klima 1963 up to (but not including) Lasnik's 1976 article 'Remarks on coreference' (Ch. 76) focused on formulating coreference rules that would specify under what syntactic conditions a given NP could be construed as the antecedent of an anaphoric expression. Lasnik's article begins with a discussion of transformational versus interpretive approaches to the syntax and interpretation of pronouns. Given that pronouns must exist independently in the lexicon, how the grammar avoids generating (7), where the pronoun and the name happen to bear the same referential index, becomes a central problem.

(7) *He$_i$ finally realised that Oscar$_i$ is unpopular.

Lasnik solves the problem by postulating a noncoreference rule (8).

(8) If NP$_1$ precedes and commands NP$_2$ and NP$_2$ is not a pronoun, then NP$_i$ and NP$_2$ are noncoreferential.[13]

He then argues that (9a–c) are descriptively equivalent theories:

(9) a. noncoreference rule
 b. noncoreference rule plus a coreference rule
 c. noncoreference rule plus a pronominalisation transformation.

Both the coreference rule and the pronominalisation transformation are then superfluous. Thus, coreference between a pronoun and a potential antecedent can hold when this is not blocked by the noncoreference rule. However, there is no rule of grammar that specifically marks a pronoun as coreferential with some antecedent.

Lasnik makes two modifications to the noncoreference rule. One concerns constructions like (10), where NP_1 (= *Mary's*) commands NP_2 (= *Mary*) but noncoreference is not obligatory.

(10) Mary's friends gave Mary a going away present.

Lasnik replaces the relation command with 'kommand', where a node α kommands a node β if the first cyclic node dominating α dominates β (the cyclic nodes being NP and S). Thus *Mary's* commands *Mary* in (10), but does not kommand it. The second concerns the anaphoric relation between plural pronouns and a split antecedent, as in (11) versus (12).

(11) Jill told Sydelle that they could leave in ten minutes.

(12) They assume that Bob will talk to Tom.

On one interpretation of (11), *Jill* and *Sydelle* can be construed as the split antecedent of *they*. Not only is there no interpretation of (12) where *they* takes both *Bob* and *Tom* as a split antecedent but, moreover, there is no interpretation where just one name is construed as part of the (split) antecedent of the pronoun. The issue here is not coreference but rather the more restrictive relation of non-intersecting reference. Thus, NP_1 (= *they*) must be disjoint in reference from NP_2 (= either *Bob* or *Tom*).

Reinhart's 1981 article 'Definite NP anaphora and c-command domains' (Ch. 77) is also concerned with domain restrictions involved in the interpretation of anaphora (see also Reinhart 1976). It begins with a discussion of some empirical problems with the 'precede and command' formulation of the noncoreference rule – in particular, 'backward pronominalisation' cases like (13), where the pronoun precedes its antecedent, and blocked 'forward pronominalisation' cases like (14), where the pronoun follows the putative antecedent.

(13) a. Near him_i, Dan_i saw a snake.
 b. The gossip about him_i upset Ben_i's wife.

(14) *Near Dan$_i$, he$_i$ saw a snake.

The 'precede and command' formulation rules out the possible interpretations in (13) and allows the impossible interpretation in (14).[14] Reinhart's solution to these problems is to restrict the domain statement further. Instead of command (or kommand), she proposes the relation of constituent-command (c-command): a node α c-commands a node β iff the branching node most immediately dominating α dominates β.[15] Given the c-command requirement, which is defined purely on hierarchical structure, the linear order restriction 'precedes' becomes superfluous. Pronominal anaphora in sentences involving preposed PPs provides a major focus of her analysis (see §4). As Reinhart notes, c-command gives a natural definition of syntactic domain in terms of hierarchical structure without reference to linear order – in effect, a single phrase. The final section of the article (§5) briefly surveys some unsolved problems.

  Chapter 78, Freidin and Lasnik's 'Disjoint reference and *wh*-trace' (1981), is primarily concerned with the question of how conditions on binding apply to traces, at the time thought to be empty categories that are bound to a lexical antecedent – in particular, the traces of wh-phrases. Chomsky 1973 and 1976 establish a parallelism between lexical bound anaphors (reflexive pronouns and reciprocal expressions) and traces left by NP movement: the binding of both is constrained by the TSC and SSC of Chomsky 1973 (see Ch. 57, Vol. IV). In Chomsky 1980, this analysis is formalised by treating the two conditions as rules that affect indexing of NPs. Applying these principles/rules to bound anaphors and NP-traces will yield an NP with an ill-formed index. The same principles also apply to pronouns, although in this case they modify the index so that the disjoint reference interpretation does not occur. However, in the case of R-expressions (e.g. names) neither principle/rule applies. Chomsky 1980 demonstrates that a wh-trace in VP behaves like an R-expression (see also Chomsky 1976), but assumes that this does not hold for wh-trace in the subject of a finite clause.[16] Freidin and Lasnik demonstrate that, with respect to binding, wh-trace in finite clause subject position also manifests the same behavior as an R-expression. Just as the pronoun cannot bind the name in (15b), the pronoun cannot bind the wh-trace in (15a).

(15) a. Who$_i$ did she say t_i had just arrived?
      b. She said Mary had just arrived.

Taking the wh-trace, a variable at LF, to be the empty category analogue of an R-expression, the disjoint reference rule predicts the non-anaphoric interpretation of the pronoun in both examples. Freidin and Lasnik show how May's 1979 derivation of the Comp-to-Comp constraint (Chomsky 1973 – see Ch. 57, Vol. IV) generalises to the case involving wh-movement out of

finite clause subject position. Their analysis yields a simplification in the analysis of binding and indexing. This result has consequences for the analysis of the Chomsky and Lasnik (1977) *that-trace filter and provides another argument that Subjacency is to be interpreted as a condition on representations (see Freidin 1978 (Ch. 66, Vol. IV)).[17]

The attempt to integrate concepts from distinct modules of UG has led to much insightful work in syntactic theory. One case in point is the formulation of the Nominative Island Condition (NIC) in Chomsky 1980.

(16)   Nominative Island Condition: A nominative anaphor cannot be free in S'.

The NIC provides a reformulation of the TSC that does not overlap with the SSC. Rouveret and Vergnaud 1980 (excerpted in Ch. 79) generalises this approach to meld the NIC and SSC into a single constraint, the Case Island Condition (CIC).

(17)   Case Island Condition:
       An anaphor x must have an antecedent in the minimal binding domain containing the element that assigns its Case to x.

The binding domain for nonull anaphors is S', whereas null anaphors (NP traces) have VP and PP as well as S' for binding domains. The CIC accounts for the following contrast.

(18)   a.   the books$_i$ were sold t_i to Mary by Peter
       b.   *the books$_i$ sold t_i to Mary by Peter

In (18b) the verb *sold* assigns accusative Case to the NP-trace and therefore VP is the minimal binding domain for the trace, so (18b) violates the CIC. In contrast, the passive predicate in (18a) does not assign Case to the trace, so the trace binding does not violate the CIC. Interestingly, the CIC prevents some movements between θ-positions, independently of the θ-Criterion.

In *Lectures on Government and Binding* (1981), Chomsky takes the integration of concepts one step further. In §3.2.3, 'The theory of binding' (excerpted in Ch. 80) he observes that because government plays a central role in Case theory, Case theory forms a subpart of the theory of government. Moreover, given the connection between Case and binding – witness Rouveret and Vergnaud's generalisation of the NIC as the CIC in the previous chapter, it is reasonable to try to develop the theory of binding in terms of the core concept of government, thereby constructing a more principled theory, which it is hoped can avoid the conceptual and empirical problems of the previous one.[18]