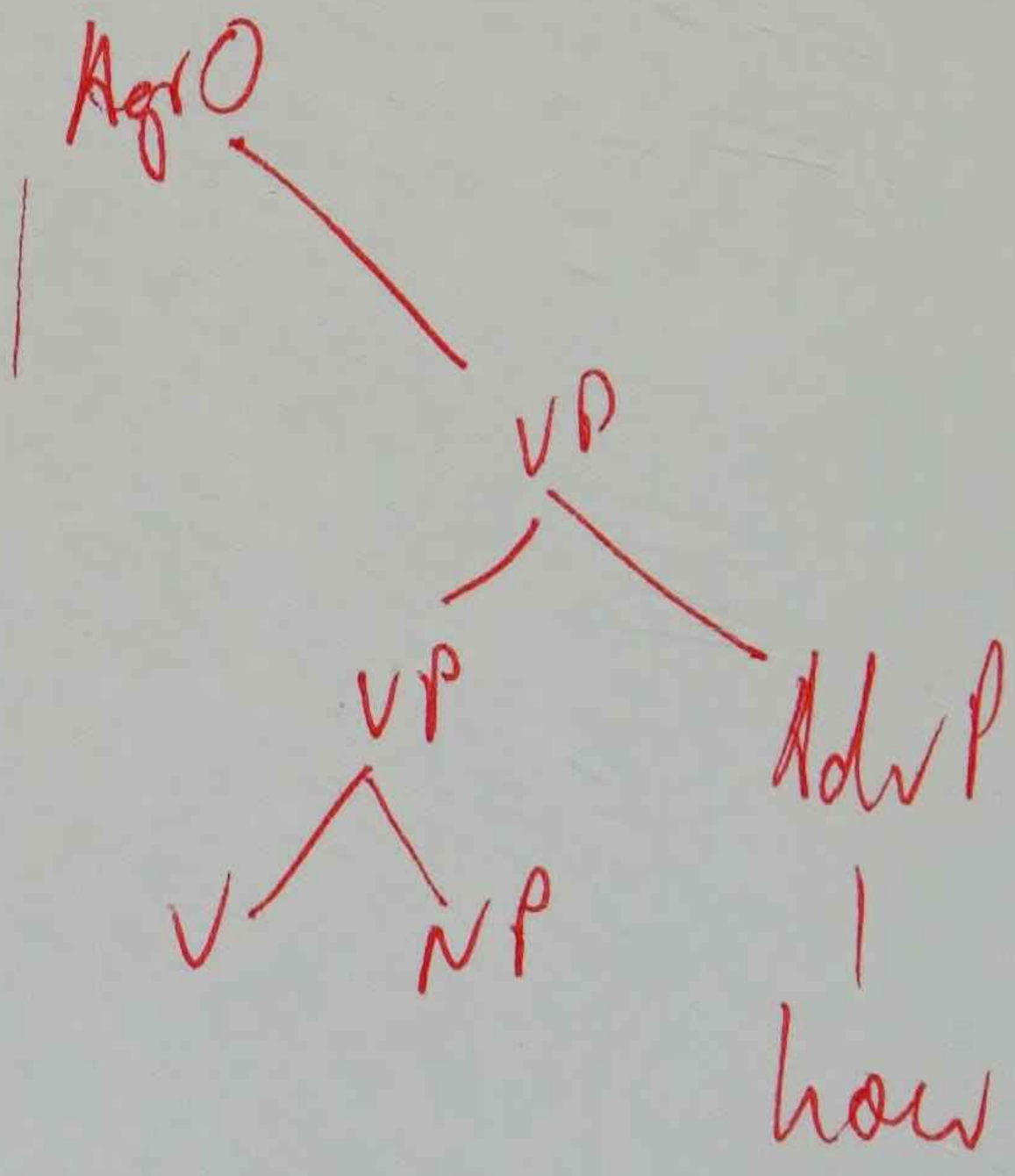
MINIMAL IDEAS

Syntactic studies in the minimalist framework

Edited by
Werner Abraham, Samuel David Epstein,
Höskuldur Thráinsson & C. Jan-Wouter Zwart

John Benjamins Publishing Company


AgrO
VP
VP
V
NP
AdvP
how

MINIMAL IDEAS

LINGUISTIK AKTUELL

This series provides a platform for studies in the syntax, semantics, and pragmatics of the Germanic languages and their historical developments.
The focus of the series is represented by its German title *Linguistik Aktuell* (Linguistics Today).
Texts in the series are in English.

Series Editor

Werner Abraham
Germanistisch Instituut
Rijksuniversiteit Groningen
Oude Kijk in 't Jatstraat 26
9712 EK Groningen
The Netherlands

Advisory Editorial Board

Hans den Besten (University of Amsterdam)
Guglielmo Cinque (University of Venice)
Günther Grewendorf (J.W. Goethe-University, Frankfurt)
Hubert Haider (University of Stuttgart)
Christer Platzack (University of Lund)
Ken Safir (Rutgers University, New Brunswick NJ)
Höskuldur Thráinsson (University of Iceland, Reykjavik)
Lisa deMena Travis (McGill University)
Sten Vikner (University of Stuttgart)

Volume 12

Werner Abraham, Samuel David Epstein, Höskuldur Thráinsson & C. Jan-Wouter Zwart (eds)

Minimal Ideas: Syntactic studies in the minimalist framework

MINIMAL IDEAS

SYNTACTIC STUDIES IN THE MINIMALIST FRAMEWORK

Edited by

WERNER ABRAHAM
SAMUEL DAVID EPSTEIN
HÖSKULDUR THRÁINSSON
C. JAN-WOUTER ZWART

JOHN BENJAMINS PUBLISHING COMPANY
AMSTERDAM/PHILADELPHIA

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences — Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Minimal ideas : syntactic studies in the minimalist framework / edited by Werner Abraham ... [et al.].

p. cm. -- (Linguistik aktuell = Linguistics today, ISSN 0166-0829; v. 12)

Includes bibliographical references and index.

Contents: The minimal links of verb (projection) raising / Marcel den Dikken -- Shortest move and object case checking / K. Scott Ferguson -- Spell-out at the LF interface / Erich Groat & John O'Neil -- The typology of syntactic positions : L-relatedness and the A/Ā-distinction / Liliane Haegeman -- Clause structure, expletives, and verb movement / Dianne Jonas -- Raising quantifiers without quantifier raising / Hisatsugu Kitahara -- Optional movement in the minimalist program / Geoffrey Poole -- Morphology and word order in Germanic languages / Jaume Solà -- On the (non-)universality of functional categories / Höskuldur Thráinsson -- Participles and bare argument structure / Guido vanden Wyngaerd -- "Shortest move" versus "fewest steps" / C. Jan-Wouter Zwart.

1. Minimalist theory (Linguistics) 2. Grammar, Comparative and general--Syntax. 3. Generative grammar. I. Abraham, Werner. II. Series.

P158.28.M56 1996

401--dc20 96-26359

ISBN 90 272 2732 2 (Eur.) / 1-55619-230-4 (US) (Hb; alk. paper) CIP

ISBN 90 272 2733 0 (Eur.) / 1-55619-231-2 (US) (Pb; alk. paper)

© Copyright 1996 - John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. • P.O.Box 75577 • 1070 AN Amsterdam • The Netherlands

John Benjamins North America • P.O.Box 27519 • Philadelphia PA 19118-0519 • USA

Contents

Bibliographical Information vii

List of Contributors ix

Preface xi

Introduction 1–66
Samuel David Epstein, Höskuldur Thráinsson & C. Jan-Wouter Zwart

The Minimal Links of Verb (Projection) Raising 67–96
Marcel den Dikken

Shortest Move and Object Case Checking 97–111
K. Scott Ferguson

Spell-Out at the LF Interface 113–139
Erich Groat & John O'Neil

The Typology of Syntactic Positions: L-Relatedness and the A/$\overline{\text{A}}$-distinction 141–165
Liliane Haegeman

Clause Structure, Expletives and Verb Movement 167–188
Dianne Jonas

Raising Quantifiers without Quantifier Raising 189–198
Hisatsugu Kitahara

Optional Movement in the Minimalist Program 199–216
Geoffrey Poole

Morphology and Word Order in Germanic Languages 217–251
Jaume Solà

On the (Non-)Universality of Functional Categories 253–281
Höskuldur Thráinsson

Participles and Bare Argument Structure 283–304
Guido Vanden Wyngaerd

"Shortest Move" versus "Fewest Steps" 305–327
C. Jan-Wouter Zwart

References 329–346

Index 347–363

Bibliographical information

An earlier version of Kitahara's paper was published in *Harvard Working Papers in Linguistics 1* (editors Susumu Kuno & Höskuldur Thráinsson, Department of Linguistics, Harvard University, 1992). An earlier version of Ferguson's paper was published in *Harvard Working Papers in Linguistics 3* (editors Höskuldur Thráinsson, Samuel David Epstein, and Susumu Kuno, Department of Linguistics, Harvard University, 1993). Earlier versions of the papers by den Dikken, Haegeman, Solà, Vanden Wyngaerd, and C.J.W. Zwart were published in *Groninger Arbeiten zur germanistischen Linguistik 37* (C. Jan-Wouter Zwart, editor, Department of German, University of Groningen, 1994).

Contributors

Werner Abraham
Department of German
University of Groningen
P.O. Box 716
NL-9700 AS Groningen
The Netherlands
abraham@let.rug.nl

Samuel David Epstein
Department of Linguistics
Harvard University
77 Dunster Street
Cambridge, MA 02138
United States of America
epstein@fas.harvard.edu

Höskuldur Thráinsson
Department of Icelandic
University of Iceland
Árnagarði v. Suðurgötu
101 Reykjavík
Iceland
hoski@rhi.hi.is

Liliane Haegeman
Department of Linguistics
University of Geneva
2, rue de Candolle
CH-1211 Geneva 4
Switzerland
haegeman@uni2a.unige.ch

Dianne Jonas
Yale University
P.O. Box 208236
New Haven, CT 06520-8236
United States of America
dianne.jonas@yale.edu

Hisatsugu Kitahara
University of British Columbia
Vancouver, BC
Canada V6T 1Z1
kitahara@interchg.ubc.ca

C. Jan-Wouter Zwart
Department of Linguistics
University of Groningen
P.O. Box 716
NL-9700 AS Groningen
The Netherlands
zwart@let.rug.nl

Marcel den Dikken
Department of Linguistics
Free University of Amsterdam
De Boelelaan 1105
NL-1081 HV Amsterdam
The Netherlands
dikken@let.vu.nl

K. Scott Ferguson
Laboratoire d'Analyse et de
Technologie du Langage
Department of Linguistics
University of Geneva
2, rue de Candolle
CH-1211 Geneva 4
Switzerland
ferguson@latl.unige.ch

Erich Groat
Department of Linguistics
Harvard University
77 Dunster Street
Cambridge, MA 02138
United States of America
groat@fas.harvard.edu

John O'Neil
Department of Linguistics
Harvard University
77 Dunster Street
Cambridge, MA 02138
United States of America
oneil@fas.harvard.edu

Geoffrey Poole
Department of Linguistics
Harvard University
77 Dunster Street
Cambridge, MA 02138
United States of America
poole1@fas.harvard.edu

Jaume Solà
Departement of Spanish Philology
Faculty of Arts
Autonomous University of Barcelona
09193 Bellaterra
Spain
ilfh1@cc.uab.es

Guido Vanden Wyngaerd
Department of Linguistics
Catholic University of Brussels
Vrijheidslaan 17
B-1080 Brussels
Belgium
guidovdw@cc1.kuleuven.ac.be

Preface

The papers in this volume come from two sources. Earlier versions of the papers by Den Dikken, Haegeman, Solà, Vanden Wyngaerd, and Zwart were published in *Groninger Arbeiten zu germanistischen Linguistik* (GAGL) and the authors are all affiliated with European universities (see List of Contributors, and the Bibliographical Information at the end of the Table of Contents). The remaining papers, those by Ferguson, Groat & O'Neil, Jonas, Kitahara, Poole, and Thráinsson, were originally written while their authors were affiliated with the Department of Linguistics at Harvard University and earlier versions of the papers by Kitahara and Ferguson were published in *Harvard Working Papers in Linguistics* (HWPL) (see Bibliographical Information).

All these papers are inspired by the Minimalist Program first outlined in Noam Chomsky's MIT Fall term class lectures of 1991 and in his seminal paper "A Minimalist Program for Linguistic Theory" (Chomsky 1993).

The editors of the present book were involved in editing the working paper volumes mentioned above (Abraham and Zwart for *GAGL,* and Epstein and Thráinsson for *HWPL*) or had worked closely with the authors of the papers. It became clear to us that the papers under discussion were highly similar in spirit. They all sought to develop further some key idea of the Minimalist Program, sometimes in ways deviating from the course taken by Chomsky, but always explicating or exploring the (cross-linguistic) empirical adequacy of the relevant "minimal ideas". For this reason it seemed to us that they should be published in a single volume which could, with a proper introduction, serve as a useful introduction to the Minimalist Program, and, more generally, as excellent text material for a course on modern syntactic theory. We discussed this idea with

the publisher, John Benjamins, who agreed with us. All of the papers have been thoroughly revised and we have added a detailed introduction (written by Epstein, Thráinsson and Zwart) relating the Minimalist Program to earlier versions of generative syntactic theory and putting the present articles into perspective. We hope that the reader agrees that this transatlantic cooperation has been very fruitful.

As the reader will note, most of the articles in this volume are concerned with issues first raised in Chomsky's paper on the economy of derivation and representation (1991) and his first two minimalist papers, namely "A Minimalist Program for Linguistic Theory" (1993, first distributed in 1992) and "Bare Phrase Structure" (1995a, first distributed in 1994). Consequently, each article begins with a quote from one of these papers, which serves to highlight the particular grammatical or theoretical issue that is primarily discussed in the relevant article.

Several articles also address the basic ideas in Richard Kayne's book, *The Antisymmetry of Syntax* (1994, distributed in part in manuscript form in 1993). In many respects, therefore, these articles develop alternatives to ideas proposed in chapter 4, "Categories and Transformations" (Chomsky 1995b) of Chomsky's most recent book, *The Minimalist Program*, without basing their discussion directly on that "version" of the Minimalist Program. Some of the articles contain references to chapter 4 of Chomsky 1995b, however, and comments on similarities and differences between ideas developed in these articles and in chapter 4 of Chomsky 1995b can also be found in the Introduction.

Finally, we would like to thank Kees Vaes of John Benjamins for his interest in this project and his help seeing it through, René Mulder for assistance in preparing the manuscripts, and all the authors for their patient cooperation.

Werner Abraham
Höskuldur Thráinsson
Samuel David Epstein
C. Jan-Wouter Zwart

Introduction

Samuel David Epstein, Höskuldur Thráinsson
& C. Jan-Wouter Zwart

This introduction has two parts. The first part is a general introduction to the Minimalist Program. It opens with a section on the common goals of modern linguistic theory, explaining how the Minimalist Program addresses questions that have guided generative research from its beginnings (section 1.1). The second part is an introduction to the articles in this volume. Apart from the opening section of part one, the two parts are structured alike. They address the following issues: phrase structure theory (section 1.2 and 2.1), movement and syntactic licensing (1.3 and 2.2), economy conditions (1.4 and 2.3), the properties of the structures built and the typology of syntactic positions (section 1.5 and 2.4), parametric variation (1.6 and 2.5), and directionality and word order (1.7 and 2.6). The reader will want to check part 1 of the introduction for a quick survey of the central notions of the Minimalist Program. Part 2 adds critical discussion, puts these notions in (historical) context, and makes the various contributions of the articles in this volume explicit.

1. Introduction to the Minimalist Program

1.1. *Origin and Some Basic Properties*

1.1.1. *Common Goals*

Two of the central goals of linguistic theory are:

(1) a. To provide an explicit **description** of what constitutes knowledge of each natural language.
b. To provide an **explanation** of how it is that a person P can come to know what they do know.

(1a) implies that for every possible natural language we must answer the question "What does a person P, who knows a language L, know?". (1b) means that for each natural language L that is humanly knowable we must explain how it is that a person P can come to know L. More specifically, (1b) requires that we answer two questions:

(2) a. What is the contribution of the organism?
b. What is the contribution of the environment?

As Chomsky (1995a:387) notes:

> The first efforts to address these problems quickly revealed that traditional grammatical and lexical studies do not begin to describe, let alone explain, the most elementary facts about even the best-studied languages ... This is hardly a discovery unique to linguistics. Typically, when questions are more sharply formulated, it is learned that even elementary phenomena had escaped notice, and that intuitive accounts that seemed simple and persuasive are entirely inadequate. If we are satisfied that an apple falls to the ground because that is its natural place, there will be no serious science of mechanics. The same is true if one is satisfied with traditional rules for forming questions, or with the lexical entries in the most elaborate dictionaries, none of which come close to describing simple properties of these linguistic objects.
>
> Recognition of the unsuspected richness and complexity of the phenomena of language created a tension between the goals of descriptive and explanatory adequacy. It was clear that to achieve explanatory adequacy, a theory of the initial state must hold that particular languages are largely known in advance of experience. The options permitted in universal grammar (UG) must be highly restricted; limited experience must suffice to fix them one way or another, yielding a state of the language faculty that determines the varied and complex array of expressions, their sound and meaning, in a uniform and language-independent way. But this goal receded still further into the distance as generative systems were enriched in pursuit of descriptive adequacy, in radically different ways for different languages. The problem was exacerbated by the huge range of phenomena discovered when attempts were made to formulate actual rule systems.

These rule systems formulated took each grammar/language to be

> a rich and intricate system of rules that are, typically, construction-particular and language-particular: the rules forming verb-phrases or passives or relative clauses in English, for example, are specific to *these* constructions in *this* language. (Chomsky 1993:4)

Thus, there arose the essential tension common to all scientific inquiry between descriptive adequacy (empirical coverage) on the one hand and explanatory adequacy on the other, with early theories arguably erring by virtue of being overly descriptive, incorporating myriad language-specific and even construction-specific postulates. Thus,

> In essence, the problem is that there are too many rule systems. Therefore it is hard to explain how children unerringly select one such system rather than another. (Chomsky 1987)

Within the more recent Principles and Parameters approach, it was explicitly recognized that these very serious obstacles to explanatory adequacy are created by the incorporation of construction-specific and language-specific theoretical constructs. An explanatory theory of language, not a theory incorporating postulates specific to English passives or French existential sentences, i.e. specific to the data, was sought. This theory, the theory of universal grammar, was assumed to consist of principles and parameters. The former are fixed and invariant language-independent, construction-independent universal linguistic laws common to all human languages. The parameters, invoked to capture seemingly irreducible and real differences among languages are the vestiges of the language particular rules of the earlier standard theory. There is assumed to be a finite (if not very small) number of parameters, each with a finite number of (perhaps two) values. Thus cross-linguistic variation, as expressed by parameters, is assumed to be highly restricted. A recent even more circumscribed theory of the formal properties of parameters postulates that each parameter is expressible as an irreducible cross-linguistic difference in the value of (a circumscribed set of) morphological features associated with certain lexical categories — the so-called functional syntactic categories such as Inflection/ Agreement, Complementizer, and Determiner.

As Chomsky notes (1995a:387) this theory of Principles and Parameters is consonant with Jespersen's view that it is only "with regard to syntax" that we expect that "there must be something in common to all human speech" — a "universal (or general) grammar", though "no one ever dreamed of a universal morphology".

If the Principles and Parameters theory is on the right track, then "... there is only one computational system [syntax] and one lexicon, apart from this limited [morpho-parametric] kind of variety" (Chomsky 1993: 3).

Thus a unified explanatory theory was sought — counter-intuitively postulating that there is, in effect, only one human language despite the appearance of unlimited diversity. The Minimalist Program — a research program —

represents the most recent and direct outgrowth of this ongoing inquiry.

Perhaps the most distinctive feature of this Program — and that which makes it most exciting and promising — is its explicit programmatic commitment to explanation through the characteristic method of minimization, consistent with Einstein's (1954) perspective that "... the grand aim of all science is to cover the greatest possible number of experimental facts by logical deductions from the smallest number of hypotheses or axioms." As concerns this minimization of the axioms, Nash (1963: 173) rightly notes that "[e]ach 'quality' imputed to a premised entity figures as an additional postulate. Our desire for parsimony of postulates thus evokes a search for theoretical posits having the slenderest possible qualitative endowment".

Consistent with this search for that which we might call "simplicity", Chomsky presents the working hypothesis that

> ... languages are based on simple principles that interact to form often intricate structures... the language faculty is non-redundant, in that particular phenomena are not "overdetermined" by the principles of language (1993: 2) ... the principles outlined are simple and restrictive... (1993: 5)

Once committed to "explanation through minimization" there of course emerge, among others, the entirely formal and empirical questions: "Which aspects of the existing theory should be simplified, which eliminated, and which retained (if any)?" The answers, regardless of the domain of inquiry, are never self-evident, nor the totality of their empirical consequences immediately (if ever) determinable.

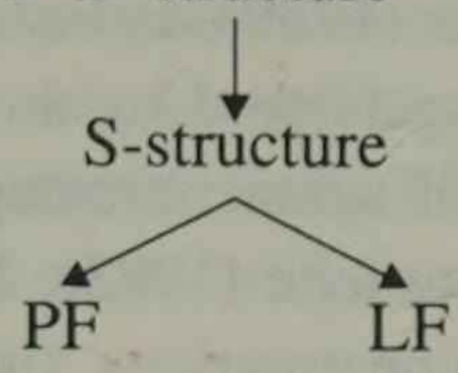
The Minimalist Program is the most recent (programmatic) outgrowth of the theory of generative grammar originally developed in Chomsky (1957, 1965). The immediate predecessor of the Minimalist Program — the so-called Government and Binding Theory (Chomsky 1981, 1982, 1986a, b) — upon informed, studied and insightful scrutiny was argued to exhibit certain suspect formal properties. The Minimalist Program first identifies, through careful and insightful analyses, potentially problematic aspects of the so-called GB theory and then proceeds to explore alternative conceptions of grammar lacking these arguably unattractive properties, eliminating them altogether or replacing them with more natural, less stipulative constraints.

1.1.2. *Minimizing Levels of Representation*

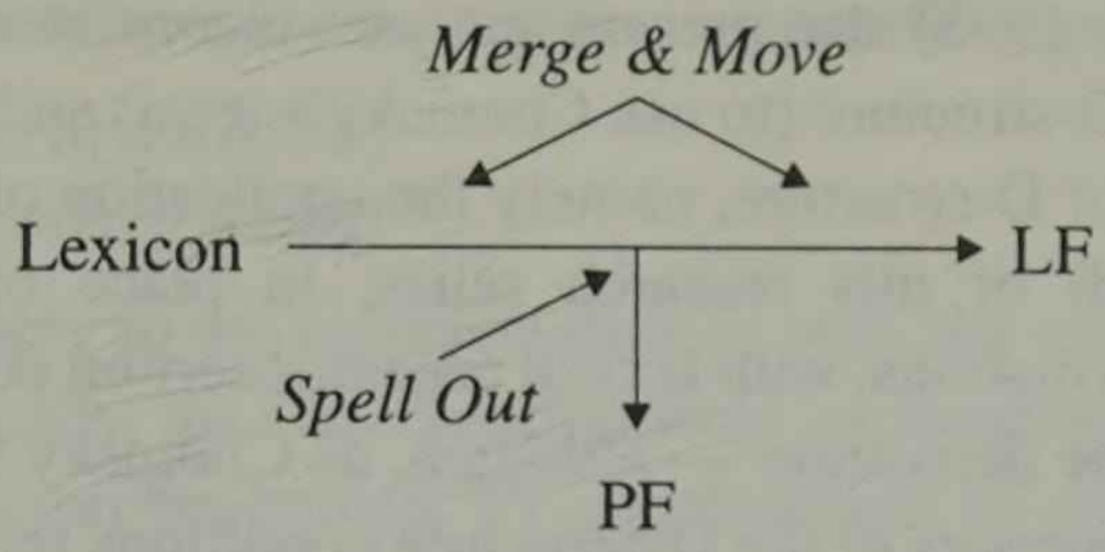
Perhaps the first targets of minimization were levels of representation. Unlike the four-level Government and Binding model of Chomsky (1981),

illustrated in (3a), the Minimalist Program assumes only two levels of representation — one an abstract representation of sound (PF), the other an abstract representation of meaning (LF), each seemingly ineliminable (3b):

(3) a.



b.



The levels of D-structure and S-structure in (3a), which are *internal* representations, in the sense that they feed into no system external to the syntactic component, are eliminated.

In the absence of a level of D-structure representation, there nonetheless remains a Lexicon and two concatenative and formally simple binary transformational (structure building) operations (cf. (3b)). One is *Merge*, a resurrection of Generalized Transformations postulated in Chomsky (1957) and later abandoned in favor of a model in which recursion is performed by (non-transformational) phrase-structure rule application. The other concatenative operation is *Move*. Each of these operations concatenates two and only two categories, forming a third. The rules apply iteratively (subject to the strict cycle condition (Chomsky 1973), reformulated as an "extension condition" (Chomsky 1993: 23)), thereby building constituent structure.

At any point in the process of such iterative binary concatenative rule-application, a (non-transformational) rule/operation called Spell-out can be optionally applied. The result of such rule application is that from the structure thus far derived certain information is split off and sent to the phonological component, while the derivation continues after the split off point *continues*, yielding the logical form (or covert) component. Thus, the split between PF and LF components inherent in the Government and Binding model is retained — the empirical hypothesis that neither component interacts with the other remains. However, D-structure is abandoned, an attractive and seemingly viable hypothesis

especially given an articulated (copy) theory of traces within which all relevant aspects of prior representation are apparently retainable throughout the derivation (cf. Freidin 1978; Koster 1987).

As Chomsky (1993: 21) notes, recent research, including Reinhart (1991), Kroch & Joshi (1985), Kroch (1989), Lebeaux (1988), and Epstein (1991) also raises empirical problems engendered by the postulation of a D-structure level of representation in which all syntactic categories appearing in a derivation D must appear. In addition, Sportiche (1983), Epstein (1987), Larson (1988), and research concerning *tough*-constructions (in particular Lasnik & Uriagereka 1988: 147 and Brody 1993) also present analyses inconsistent with either the all-at-once property of D-structure (to use Chomsky's term) and/or inconsistent with a defining property of D-structure, namely the application of the Theta Criterion at this level.[1] Much of this research relies, in place of D-structure, upon Generalized Transformations, with lexical access/insertion consequently applying over the course of the derivation — although, as Chomsky (1993: 21) notes: "... the empirical consequences of the D-Structure conditions remain to be faced... ."

S-structure as an independent level of representation is, as noted above, also eliminated, being replaced by the optional rule of Spell-out, allowed to apply at any point — splitting the derivation into two tracks exactly as in the Government and Binding-model. This elimination of S-structure was already suggested in Chomsky (1986a) wherein it is observed that S-structure has a unique status within the Government and Binding model in that it is the sole (central) level linking, i.e. having contact with, each of the other three levels — D-structure, PF, and LF. As Chomsky observed, this raises the possibility that its properties may well be deducible from the three requirements that: (i) it must be derivable from D-structure by iterative transformational-rule application and, at the same time (ii) PF representation must be derivable from it by the operations internal to the PF component, and (iii) LF representation must be derivable from it by iterative transformational-(covert)-rule application. As Chomsky (1986a:101) notes:

> properties of S-structure may be reducible to the independent conditions of Full Interpretation holding of PF and LF representation ... given an appropriate account of the ways in which the elements of a structure Σ [i.e. the levels of representation E/T/Z] may be related.

It is this "reduction to independent conditions" that pervades the Minimalist Program. As concerns S-structure representation, the idea is that it is eliminable

---

1. Notice that this defining property is conceptually problematic as well: the Theta Criterion, if it exists, is an interpretive condition, and should therefore apply only at LF.

by appeal to an optional rule of Spell-out which is allowed to apply at any point in the derivation.

The hypothesis is that this level elimination and its concomitant replacement with an optional, unordered and virtually conceptually necessary operation (ultimately yielding two ineliminable representations: one of sound, the other of meaning) may well be empirically adequate. The precise derivational point at which Spell-out applies is, by hypothesis, deducible from independent requirements (Full Interpretation) imposed on the only two (and the natural two) levels of representation: PF and LF.

To take but one illustrative example, consider the representations in (4):

(4) a. [*was arrested he*]
b. [*he was arrested*]

Were Spell-out to apply to the representation in (4a), the claim is that this is not a licit sound, specifically, the Case requirements imposed upon *he*, requirements now interpreted as morphophonetic PF-requirements (S-structure having been eliminated) are not satisfied. If, on the other hand, Spell-out were applied after NP-movement of *he* to subject position, yielding (4b), this representation, submitted to PF scrutiny, is regarded as consisting entirely of legitimate, i.e. legible, PF objects. In particular, *he* is PF-licit by virtue of satisfying Case requirements, requirements which always have been construed as pertaining to (at least) phonetically realized NPs/DPs (cf. Chomsky 1981: 49).

Thus, it is hypothesized that there is neither a D-structure- nor an S-structure-level of representation but rather a Lexicon, a quite simple binary concatenative algorithm yielding, for each expression, a PF (sound) representation and an LF (meaning) representation.

As expected, the properties of these two remaining levels of representation — PF and LF — are themselves also subjected to scrutiny. Thus one must ask the following questions:

(5) a. Why are there two levels of representation (i.e. PF and LF)?
b. Why do they each have the properties they seem to have — i.e. Why are PF and LF representations each constituted of certain objects (not others), and why must the objects be arranged in a particular way?

Chomsky conjectures that the answers lie in the fact that representations at PF and LF (the interface levels) must be read by (legible to) the Articulatory-Perceptual (A–P) and the Conceptual–Intentional (C–I) systems respectively, that

is, the A–P and C–I systems external to the syntactic component proper impose legibility conditions on the interface representations.

For example, a PF representation must, by hypothesis, be linearized in order for it (or some transform of it) to undergo pronunciation, i.e. interpretation by the A–P system. Thus, a fairly fundamental aspect of PF representations — that they contain linearly ordered strings — is arguably deducible from demands imposed from without, i.e. requirements imposed by the articulatory system. Similarly, the following entirely descriptive universal (substantive) filter

(6) *V
+high
+low

may in fact not be a descriptive phonological filter, but might instead be explained by appeal to natural properties of the articulatory systems structurally external to the syntactic component but which nonetheless take as input, interface representations (in this case, PF representations).

These legibility conditions imposed on PF and LF representations by the A–P and C–I systems respectively are known as the Bare-Output Conditions and here again the program looks promising and, importantly, the questions it seeks to address must be posed in any event.

If on the right track, the number and properties of the only two and the natural two representations — those of sound and meaning — are to be deduced from the properties of the external systems A–P and C–I which (in both senses of the word) must read these representations. Thus the Minimalist Program seeks to eliminate, on principled grounds, D–structure representation and S–structure representation while concomitantly hypothesizing a programmatic deduction of the properties of the only two (comparatively uncontroversial and seemingly ineliminable) levels of representations — the interface levels of PF and LF represention.

1.1.3. *Building Representations*

Having discussed the Minimalist reduction of levels of representation, we now must address the question of how representations (PF and LF representations) are generated.

As noted, there is assumed to be a Lexicon — an arguably irreducible component of the grammar expressing what we know when we know the words of a given language. Each lexical entry is assumed to consist of, at least, a set of three sets: a semantic-feature set, a phonological-feature set, and a syntactic-feature set. Again this seems unavoidable — a PF representation containing *cat* /kæt/ must provide instructions to the A–P system specifying that this particular

lexical item (meaning what it does) is pronounced /kæt/, not /tæk/ nor /ækt/. Thus a lexical entry must contain irreducible/unpredictable phonological specifications to be interpreted when PF representations are submitted to the A–P system. The same holds of semantic properties of lexical items — these are tantamount to irreducible atomic instructions (present in LF representations containing lexical items) to the C–I system ("interpret *cat* as ..."). Importantly then, it follows that certain aspects of, or constraints on, lexical representations are deducible from the bare-output conditions — each aspect of a lexical representation (present in a PF or in an LF representation) must be interpretable by A–P and C–I systems, respectively. Hence we have a program for determining the core notion "possible lexical representation".

In addition to phonological features and semantic features a lexical item is assumed to have syntactic features, e.g. categorial features (Noun) and φ-features (including person, number, and gender) identifying its syntactic status and providing instructions to the syntactic combinatorial system itself.

Given a Lexicon, the generative procedure consists of Merge and Move (see 1.2 below). Each is subject to the strict cycle condition (Chomsky 1973) and each concatenates exactly two objects forming a third. For purposes of illustration, we might have a derivation such as:

(7)

a.	1st	*Merge* the_D *and* dog_N *yielding* $[_{DP}\ the_D\ dog_N]$
b.	2nd	*Merge* $arrested_V$ *and* $[_{DP}$ the dog] *yielding* $[_{VP}\ [_V$ arrested] $[_{DP}\ the_D\ dog_N]]$
c.	3rd	*Merge* be_{Infl} *and* $[_{VP}\ [_V$ arrested] $[_{DP}\ the_D\ dog_N]]$ *yielding* $[_{I'}$ be $[_{VP}\ [_V$ arrested] $[_{DP}\ the_D\ dog_N]]]$
d.	4th	*Merge* $[_{DP}$ the dog] *with* $[_{I'}$ be $[_{VP}\ [_V$ arrested] $[_{DPTRACE}\ the_D\ dog_N]]]$ *yielding* $[_{IP}\ [_{DP}$ the dog] $[_{I'}$ be $[_{VP}\ [_V$ arrested] $[_{DPTRACE}\ the_D\ dog_N]]]]$
Ultimately		[the dog **was** arrested $[_{DPTRACE}\ the_D\ dog_N]]]]$ *(via application of agreement)*

The 1st Merge concatenates two lexical items. The 2nd and 3rd Merge each concatenates one lexical (atomic) item and an already constructed (molecular) phrase. The 4th operation, involving Move, takes a phrase within an already constructed tree and concatenates it with that tree. This too is binary — concatenating one member of a tree with the one tree containing it. The strict cycle

condition (Chomsky 1973) prohibits (syntactic) infixation. The strict cycle condition, now called the "extension condition", dictates that no category can be concatenated with a category (already) having a dominating node. Informally, Merge and Move always add a sister to the root node.[2]

In the following sections (sections 1.2-1.6) we will introduce the basic concepts of the Minimalist Program in more detail. In section 1.7, we will present a brief sketch of the antisymmetry hypothesis of Kayne (1994), which complements the Minimalist Program outlined in Chomsky (1993).

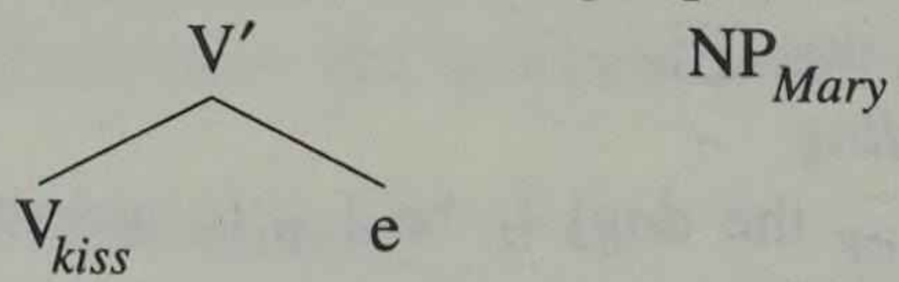
1.2. *Building up Phrase Structure*

As just described, representations are built up in a (strict-cyclic) bottom-up fashion by Generalized Transformation. A Generalized Transformation combines two phrase markers where here we include among the class of phrase markers, lexical items, e.g. [$_N$cat]. Two phrase markers are combined by expanding one (the target phrase marker) so as to include an empty position. This expansion takes place by adding to the target phrase marker a projection of the target phrase marker. This projection is binary branching and has two daughters: the target phrase marker and an empty position. The other phrase marker then substitutes into this empty position. The whole process, illustrated in (8), yields two sister phrase markers in a binary branching subtree.[3]

(8) a. Two independent phrase markers

V_{kiss} NP_{Mary}

b. Expansion of the target phrase marker V

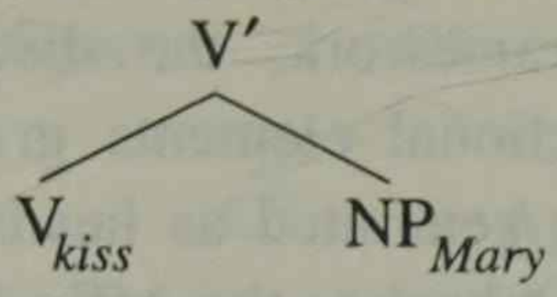


---

2. This condition may itself be deducible from natural and independently motivated conditions (Kitahara 1994, 1995, 1996). Moreover, Merge and Move — each being binary concatenative operations may well be unifiable as a single operation Target-α as proposed by Kitahara (1994, 1995, 1996).

3. Binary branching is a result of this particular formulation of the Generalized Transformation mechanism. The attractiveness of binary branching has been argued for several times in the literature (see e.g. Kayne 1984; E. Hoekstra 1991).

c. Substitution of NP into the empty position in the projection of the target phrase marker



(8a–c) apply internal to a single operation. There is no intermediate representation of the form (8b). The projection of the target phrase marker has the same categorial features as the target phrase marker. The phrase level of the projection of the target phrase marker is determined by the X-bar schema (Chomsky 1986b; Chomsky 1970; Jackendoff 1977), specifying that the ultimate projection of an X (or X^0, or head) will be an XP (or X″, or maximal projection), and that there is an intermediate projection X′ (X-bar) which is the immediate projection of X. This is expressed in the following two rewrite rules:

(9) a. XP → (ZP) X′
b. X′ → (YP) X^0

The order of the elements to the right of the arrows in (9) is irrelevant (but see section 1.7 below). The sister of X^0, YP in (9b), is called *complement;* the sister of X′, ZP in (9a), is called *specifier.*

Because the Generalized Transformation illustrated in (8) combines two independent phrase markers, it is called a *binary* operation. Lexical insertion is a typical binary operation. It is also possible that the empty element created by expanding the target phrase marker is substituted for by an element contained in the target phrase marker. This would be called a *singulary* operation, or Move, e.g. (7d) above.

1.3. *Licensing: Movement and Feature Checking*

A standard distinction exists in linguistic theory between contentful elements and functional elements. Word stems are contentful elements, whereas inflectional morphemes are functional elements. Functional elements express agreement relations between constituents.

In the Minimalist Program, it is assumed that agreement relations are highly local. A maximal projection α agrees with a head β only if α is a specifier of β. Thus in (7) there is a Spec–head relation between [the dog] (singular) and *be*: here singular overt morphological agreement (as well as Past Tense morphology)

(*was*) appears.[4] A head α agrees with a head β only if α is adjoined to β.[5] Moreover, β must be a functional head.

In the Government and Binding framework, the distinction between contentful (or lexical) elements and functional elements gradually took the following shape.[6] Functional elements are generated as heads of independent phrasal projections. This idea has roots going back to the Affix-Hopping analysis of Chomsky (1957, 1965) in which at D–structure agreement morphemes are generated separated from the verb to which they are affixed in the S–structure representation. Thus, for example, the inflectional morphemes tense and agreement are generated separate from the lexical stems. This yields a structure as illustrated in (10) (next page).

In (10), C^0 stands for the complementizer position, T^0 for tense, and $AgrS^0$ and $AgrO^0$ for subject and object agreement morphology, respectively. These functional heads project phrases in accordance with the X-bar schema given in (9) above. AgrOP, TP, AgrSP and CP together constitute the functional domain of a syntactic structure. VP constitutes the lexical domain. The stems, generated in V, have to be united with the inflectional morphemes in the functional heads through a transformational process of adjunction.

In the Minimalist Program, this analysis is maintained in a simplified form. The major difference concerns the content of the lexical and functional heads. In the Minimalist Program, lexical heads are fully inflected forms (stems plus inflectional affixes). These forms carry a feature associated with the inflectional

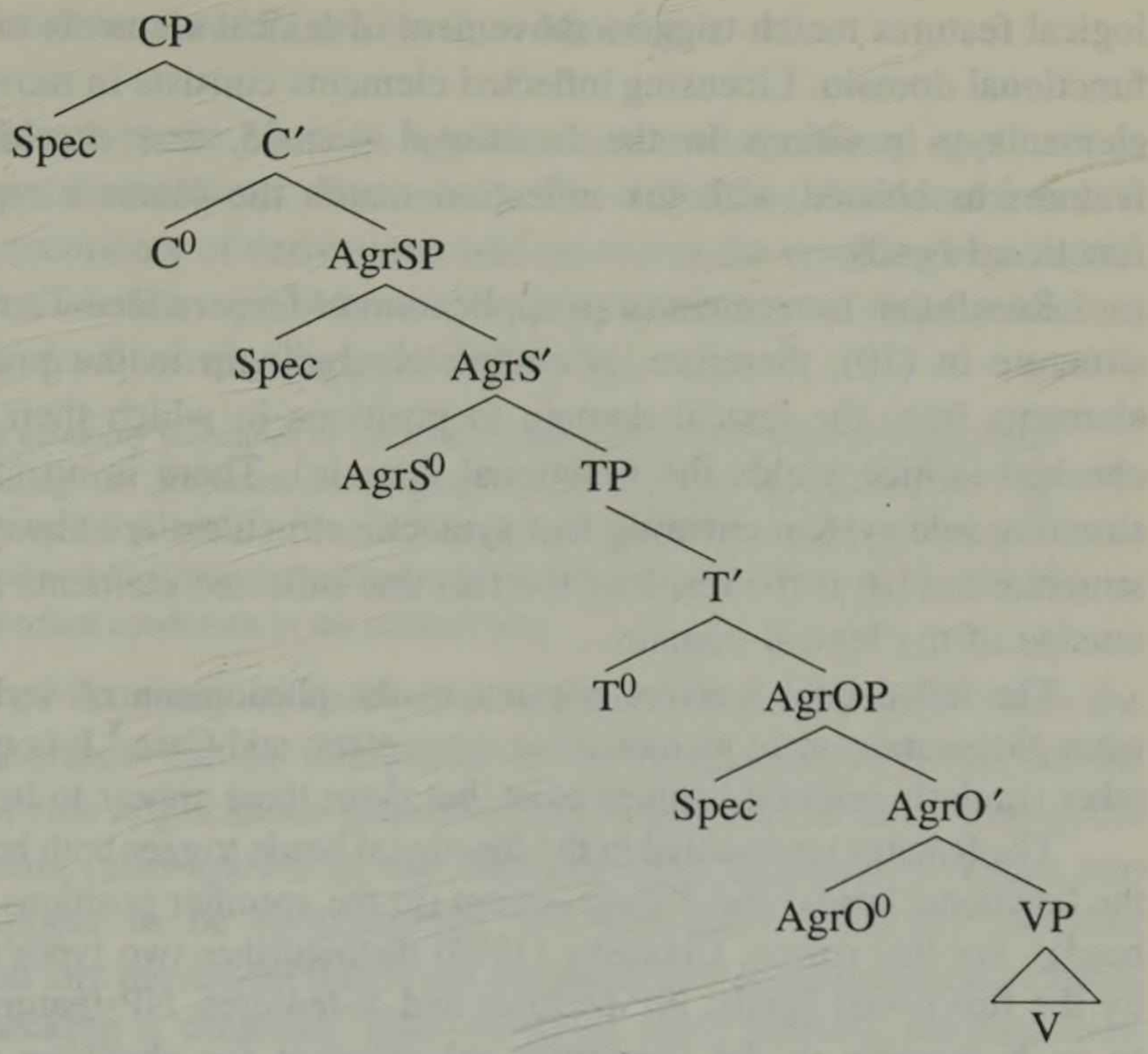
---

4. The Spec–head relation can be instantiated before and after Spell-out. Possibly, in the English example *the dog* and *was* are not in a Spec–head relation in overt syntax, since adverbs may intervene (*the dog probably was not arrested*).

5. The locality requirements are further restricted in Zwart (1992b), where it is argued that an element α agrees with an element β only if α adjoins to β. This implies that in a specifier-head agreement relation, the specifier does not agree with the head, but with the immediate projection of a head. See also Epstein (1994, 1995).

6. The developments in the Government and Binding period are marked by Stowell (1981) and Pesetsky (1982), on the structure of IP, Chomsky (1986b) on the structure of CP, Kayne (1989), Pollock (1989) and Chomsky (1991) on the division of IP into AgrP and TP, and Abney (1987) on the functional domain of noun phrases, DP. See Fukui & Speas (1986) on the relevance of functional projections for parametrization.

(10)



affix. The functional heads likewise consist of features associated with inflectional morphology.[7]

The features associated with the inflectional morphology of lexical categories have to match the features represented in the functional heads. Matching is checked under the same strict locality requirements as is agreement (in fact, agreement is a subcase of feature matching). Thus again returning to (7d), there would be a feature mismatch appearing in the Spec–head relation in e.g. [*the dogs* (plural)] *was* (singular) *arrested*]. Therefore, the requirement that morpho-

---

7. The assumption that abstract features associated with inflectional morphology are of greater syntactic significance than the overt morphology itself is already a crucial part of the Case Theory module of the Government and Binding framework. This Case Theory refers to abstract Case features which are associated with nouns and noun phrases regardless of the morphological manifestation of Case on the nouns and noun phrases (Vergnaud 1982; Chomsky 1981). This theory of abstract Case is subsumed under the Minimalist Program. As a result, the inflectional features associated with Case are assumed to be present on lexical categories, even if there is no overt morphological manifestation of Case on these categories.

logical features match triggers movement of lexical elements to positions in the functional domain. Licensing inflected elements consists in moving the inflected elements to positions in the functional domain, and checking whether the features associated with the inflection match the features represented in the functional heads.

Recall that movement is an application of Generalized Transformation. The structure in (10), therefore, is completely built up in the process of moving elements from the lexical domain to positions in which their features can be checked (which yields the functional domain). There is no top-down phrase-structure rule system ensuring that syntactic structures are always like (10). The structure in (10) is the result of the fact that inflected elements must be licensed outside of the lexical domain.

The inflectional features relevant to the phenomena of verb movement and noun phrase movement include tense, agreement, and Case.[8] It is quite possible that other similarly relevant features exist, but these three appear to be indispensable.[9]

The features represented in the functional heads trigger both head movement (to the functional heads) and XP-movement (to the specifier positions of the functional heads). For this reason, Chomsky (1993) distinguishes two types of features borne by the functional heads: NP-features and V-features. NP-features (also referred to as N-features in the literature) are relevant for checking features of XPs (maximal projections), V-features are relevant for checking features of heads.

---

8. The exact difference between Case and agreement is not very clear in this system. It is assumed that the specifiers of AgrS and AgrO are the positions for checking Nominative and Accusative Case features respectively. This suggests that Case and agreement are identical concepts. However, Chomsky (1993: 29–30) suggests that, while Nominative and Accusative Case features are checked in the specifier positions of AgrS and AgrO, respectively, the features relevant for checking Case do not reside in AgrS and AgrO, but in T and V, respectively (see 1.5). See also the discussion in the papers by FERGUSON and JONAS in the present volume.

9. Iatridou (1990) contends that the approach to inflectional morphology sketched here leads to an "explosion" of functional categories, assuming that every functional category discovered in studying the languages of the world should be present in the grammar of every single language of the world. One way to think of this is to assume that a small number of inflectional features are present in all languages of the world, whereas a larger number may be relevant to specific languages only. What is syntactically universal, however, is the way the presence of inflectional features determines movement and word order. — For further discussion of issues of this kind see the papers by SOLÀ and THRÁINSSON in the present volume.

1.4. *Economy Conditions and Restrictions on Movement*

In direct contrast to the Move-α hypothesis, transformational applications are subject to general conditions of economy. The derivation should take as few steps as possible (economy of derivation), and moreover the resulting representations should have as few symbols as possible (economy of representation) (Chomsky 1991). As Chomsky suggests (1993: 5)

> ... it seems that economy principles of the kind explored in early work play a significant role in accounting for properties of language. With a proper formulation of such principles, it may be possible to move toward the minimalist design: a theory of language that takes a linguistic expression to be nothing other than a formal object that satisfies the interface conditions in the optimal way.

One aspect of economy of derivation is that movement always takes the shortest route. Shortness can be interpreted in two ways, viz. as involving the fewest number of rule applications (Shortest Derivation) and/or as involving the shortest movements (also known as the Minimal Link Condition). These two interpretations appear to be contradictory (cf. Chomsky 1993: 21). Another consequence is that any movement that is not triggered by a requirement of morphological feature checking is excluded. Thus, elements, once licensed, are immobile. This is known as Last Resort (see e.g. Epstein 1992 for recent discussion).

Economy of representation (naturally enough) excludes the presence of uninterpretable material in an interface representation. One instantiation of this is the principle of Full Interpretation (FI).[10]

FI requires that every element of an interface representation must provide a meaningful/legible input to the syntax-external cognitive systems, A–P and C–I. Only these elements are considered to be legitimate objects in an interface representation. Again such requirements (the Bare Output Conditions) seem natural assuming that interface representations are input to the A–P and C–I systems.

---

10. This principle was first introduced in Chomsky (1986a:98), in the context of a discussion of the relation between Case assignment and theta-role assignment. The idea was that noun phrases must be assigned Case at S–structure, because only then would they be visible for theta-role assignment at LF. Since only noun phrases that carry a theta-role are interpretable at the interface of LF and other components of the cognitive system, the principle of Full Interpretation requires Case assignment at S–structure. This concept has been slightly changed in the Minimalist Program. Case checking eliminates features since Case cannot be interpreted at the interfaces. Without Case checking, Full Interpretation (and economy of representation) is violated, since unchecked Case features are uninterpretable.

The syntactic features associated with inflectional morphology are considered to be relevant for syntax only, i.e. they lack, by hypothesis, phonological features and semantic features and therefore are uninterpretable at the interface levels. They play a crucial part in the licensing of inflected elements. However, these features are uninterpretable to components of the cognitive system external to the syntax. In other words, the features associated with inflectional morphology are not legitimate objects at the interface level: they cannot be a part of an interface representation that is to serve as input to other components of the cognitive system.

For this reason, these features must be eliminated during the derivation. It is assumed that matching features on a purely inflectional head are eliminated as soon as they are checked.[11]

Therefore, a minimal number of derivational steps is required to achieve a minimal representation at the interface of the syntactic component and other components of the cognitive system.

Two other economy principles are directly relevant.

First, picture the derivation as a step-wise procedure. At each step, economy of derivation will allow only a minimum of transformational activity. Eventually, movements will have to take place, but economy of derivation dictates that these take place as late in the derivation as possible. This can be formulated as a separate principle, *Procrastinate* (Chomsky 1993: 30).

Second, movement is triggered by the need to license inflected elements — more exactly, by the need to check off and thereby delete the uninterpretable abstract features associated with inflected elements, which, recall, are prohibited at the interface. Elements that are already licensed, or that do not need licensing, are not forced, hence they are not allowed, to move. It follows that such elements can never move, not even in order to assist in the licensing of another element. The economy principle that prohibits such solely altruistic movement is called *Greed* (Chomsky 1993: 33).

1.5. *Phrase Structure and Category Types*

Crucially, rule application is highly restricted within the minimalist system. This stands in direct contrast to the Move-α hypothesis, under which transformational rule application was entirely unconstrained. D–structure is eliminated

---

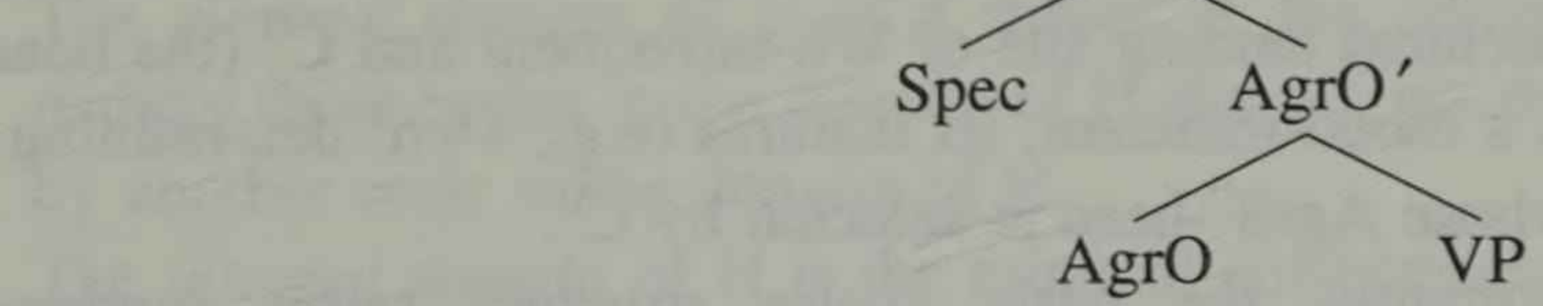
11. This is the current incarnation of recoverable "deletion-under-identity".

and representations are built from the bottom up, step by step (i.e. in a succession of binary rule applications), subject to the strict cycle/extension condition, with Move further restricted/defined by (at least) the derivational economy constraints, Greed and Procrastinate.

The syntactic component is thus, in Chomsky's terms, strongly derivational. Intermediate representations and the formal properties of the recursive procedure generating them are syntactically significant — an empirical hypothesis antithetical to the entirely filter-based (representation-characterizing) Move-α hypothesis.[12] In the Move-α approach, the generative procedure itself is unconstrained (if not undefined). Properties of the output representations (however generated) are the sole object of linguistic significance.

As noted, Chomsky (1993), adapting Pollock's (1989) analysis, assumes the following basic clause structure:[13]

(11) (= Chomsky 1993: example (2), p. 7 (slightly modified))



---

12. For further exploration of this derivational approach see Epstein (1993, 1994, 1995), Epstein, Groat, Kawashima & Kitahara (forthcoming), Groat (1995, in progress), and Kitahara & Kawashima (1994).

13. As Chomsky (1993: 7) notes, this representation omits possible additional structure. As regards the position Spec-TP, see JONAS (1992, 1994, 1995, this volume), Jonas and Bobaljik (1993), Bobaljik & Jonas (1993), Bures (1992, 1993), Bobaljik & Carnie (1992), and THRÁINSSON (1993a, this volume). Also omitted is a possible Negation and/or Affirmation Phrasal Projection (see Pollock 1989; Laka 1990; Zanuttini 1991).

Here Infl of Chomsky (1981) is split into two distinct X-bar-projections – a Subject-Agreement (AgrS) projection and a Tense (T) projection. In addition, an object-agreement (AgrO) phrase appears as sister of T^0, and its head AgrO takes VP as its complement. Agr is assumed to be a collection of φ-features (person, number, and gender) playing a central role in both subject agreement and in object agreement systems.

Chomsky proposes a unified and restricted theory of (abstract) Structural Case-assignment. There is no Exceptional Case-marking nor is the sister to V (the direct object) a Case-marked position. Rather, Structural Case is assigned in only one configuration: Specifier of Agr. However, since the Case borne by the category constituting Spec of Agr does appear to be determined by the featural (Case-assigning) properties of V (for the object) and T (for the subject) and not by Agr itself, it is proposed that T^0 raises (adjoins) to $AgrS^0$, and V raises (adjoins) to $AgrO^0$. These head-movements create complex (molecular) heads which contain the φ-features of Agr as well as the Case-features of V/T adjoined to Agr. An NP (or DP) Specifier of such a complex head (of $[_{Agr}$ T + AgrS] or of $[_{Agr}$ V + AgrO]) bears the Case and agreement features of the complex head. As a result, Case and agreement are also unified — each a manifestation of the Spec-head relation, which along with the head–head relation (V-to-AgrO raising and T-to-AgrS raising) constitute the central X′-theoretic relations involved in inflectional morphology. Notice that the subject and object inflectional systems are also unified — the subject is Spec-Agr(S); (agreement is determined by the φ-features of Agr and Case by the Case-features of Tense adjoined to Agr) and similarly the object is Spec-Agr(O) (with agreement (similarly) determined by AgrO and Case by the V which adjoins to it). Properties of the CP projection remain largely unchanged from previous work (Chomsky 1986a, b) — Spec-CP is the canonical landing site of Wh-movement and C^0 (the head of CP) is construed as a mood-indicator, its features (e.g., +wh) determining the force of the AgrSP whose $AgrS^0$-head is selected by C^0.

Importantly, the basic clause structure raises fundamental questions concerning the proper typology of positions. In Chomsky (1981), NPs occupied two types of positions: A (argument) or $\overline{A}$ (non-argument), defined in relation with (potential) theta-role assignment (Chomsky 1981: 34 ff.; see below, section 2.4). In Chomsky (1993), Chomsky seeks to eliminate the A-vs.-$\overline{A}$ dichotomy and attempts to replace it with a positional taxonomy directly related to the morphological features of lexical items (consistent with the central role played by lexical items and their morphological features within the minimalist framework).

As Chomsky notes, Agr and Tense incorporate (or represent) morphological features of the verb. These features are called "V-features" and their derivational function is to check the Agreement and Tense features of the verb that raises (adjoins) to Agr and Tense (as discussed above) and to check Case and Agreement properties of the NP/DP that raises to their Spec position. Thus, they are mediators that ensure that DP and V are properly related. More generally, V-features, being features of a lexical (L) item are called L-features. Positions, in the minimalist framework, are defined in terms of their relation to L-features. Thus, a position is L-related if it is in a local (X′-theoretic) relation to an L-feature, i.e. if it is the Spec of or the Complement of a head (X^0) bearing an L-feature.

In Chomsky (1993), a fundamental distinction between complement positions and specifier positions is assumed. Complements and specifiers belong to different domains, which are defined on the structures built by Merge and Move. The specifier of an X^0 is part of the checking domain of the X^0 wherein Case and Agreement relations between a head and a DP are established. The complement of an X^0 is part of the internal domain of an X^0 wherein theta-relations between an X^0 and an argumental DP are established.

The definitions of checking domain and internal domain of a head H can be paraphrased as follows:[14]

(12) a. The *domain* of H is the set of nodes dominated by the maximal projection of H, with the exception of H itself and the projections of H.

b. The *complement domain* of H is the part of the domain of H that is (dominated by) the sister of H; the remainder of the domain of H is the *residual domain* of H.

(13) a. The *minimal domain* of H is a subset of the domain of H, namely those nodes in the domain of H that are not dominated by another node in the domain of H.

b. The *internal domain* of H is the minimal complement domain of H (i.e., the sister of H); the *checking domain* of H is the minimal residual domain of H (i.e., the specifier of H and the adjuncts of (a projection of) H).

According to (13b), the checking domain of a head H is assumed to contain not only the Spec of H (and a head adjoined to H) but also a category adjoined to

---

14. See Chomsky (1993: 10–13) for complete definitions of checking domain and internal domain of a head and of a head movement chain.

H^{max}. Spec is called a "narrowly L-related" position, while a position adjoined to H^{max} is dubbed "broadly L-related". As Chomsky (1993: 28–29) notes:

> A structural position that is narrowly L-related has the basic properties of A-positions; one that is not L-related has the basic properties of $\overline{A}$-positions; in particular [Spec-C], not L-related if C does not contain a V-feature. The status of broadly L-related (adjoined) positions has been debated, particularly in the theory of scrambling.[footnote deleted] For our limited purposes, we may leave the matter open.

Thus, Chomsky intends to replace the A/$\overline{A}$ positional taxonomy with the L-based system recognizing at least the positional types L-related, including broadly and narrowly L-related, and non-L-related.

1.6. *Universality, Parametric Variation, and Morphological Strength*

According to the Minimalist Program, the derivation of a sentence yields two interface representations, a PF and an LF representation, each of which is subject to the principle of Full Interpretation, requiring that an interface representation must consist of legitimate objects only. For PF, the empirical question "What are the legitimate PF objects?" will be answered by an empirically adequate theory of Universal Phonetics. At LF, each legitimate object is assumed to be a chain, including X^0-chains, argument-chains, modifier chains and operator–variable chains.

If an interface representation consists of only legitimate objects the derivation yielding it is said to converge at this level of representation, e.g. "The derivation converges at PF". If a derivation D converges at both PF and at LF, then D is said to converge – period. Failure to converge is called "crashing". A derivation D can crash at PF, crash at LF, or crash at PF and at LF. Importantly, if D converges it is not necessarily the case that each interface representation is interpretable/sensical. Thus, e.g. the well-known cases of nonsense poetry and sentences like *Colorless green ideas sleep furiously* consist entirely of legitimate objects but are not interpretable or are nonsensical.

The other components of the cognitive system that the syntactic component is embedded in are performance systems, concerning (roughly) speech and interpretation. As noted, there are assumed to be two discrete performance systems: articulatory–perceptual and conceptual–intentional (Chomsky 1993: 2). In accordance with this, there are two interface representations (PF and LF), each hypothesized to provide instructions to one of the two performance systems: PF to A–P, LF to C–I.

On the assumption that the human conceptual-intentional performance

system is an invariant property of the species, the interface representation called "LF" must be essentially identical in all languages. By contrast, the interface representation called PF varies from language to language, as can easily be observed. (The point to be made here is actually more subtle. What differs (in a restricted way) cross-linguistically among PF representations is the phoneme order, phoneme inventory, and the language-particular phonetic rules–all presented (in the data) and thereby representing no learnability problem (for humans!). The way the corresponding instructions are interpreted by the articulatory-perceptual performance system is presumably just as universal as the way the LF instructions are interpreted by the conceptual-intentional performance system.)

In the Minimalist Program, it is assumed that the LF interface level is the final stage of a derivation, and that the PF interface level is the reflection of an intermediate stage in the derivation to LF. That is, at a certain point in the derivation, instructions to the articulatory–perceptual system are issued by applying a rule/operation called Spell-out. That part of the derivation preceding Spell-out is called overt syntax, the remainder of the derivation, i.e. that which follows the application of Spell-out is called covert syntax.

The problem of comparative linguistics is to determine the (highly limited) formal respects in which languages can differ, and then to ascertain why language varies in the ways it does. Recall that the principle of Procrastination dictates that movements take place as late in the derivation as possible. This principle, then, has to be violated to some extent in the grammar of certain, perhaps all, languages. The question is why? That is, why isn't "as late as possible" always synonymous with "in the covert syntax"? Equivalently, given Procrastination, why isn't movement always postponed until after Spell-out, in which case not a single language would have overt movement? In other words, why doesn't the A–P system spell out what used to be called D–structure in Government and Binding terms?

The only possible answer to this question within the highly restricted Minimalist framework is that overt movement applies, i.e. Procrastination must be violated, in order to ensure convergence at the PF interface level. In other words, certain objects that would be illegitimate at PF must be eliminated in overt syntax. Adhering to the minimalist assumptions made above, it may be the case that certain inflectional features count as illegitimate objects at PF. These features, then, have to be checked and eliminated in overt syntax, through a process of movement of heads and phrases to positions in the functional domain.

Notice that this view is quite consonant with earlier analyses. For example,

I wonder [Comp ⟨+wh⟩] *John likes what* is the wrong way to say — i.e. the wrong sound/PF in English. Rather, a Comp selected by *wonder* requires a local *wh*-phrase — where locality can be obtained by movement, viz. *I wonder what John likes*. Thus, in English a ⟨+wh⟩ Comp without *wh* is an illegitimate PF object.

The surprising aspect of this mechanism is that the class of illegitimate objects at PF is not universalized but rather varies cross-linguistically. If there were universality, overt syntax would be largely, perhaps completely identical in all languages. As we know, there are very distinct differences in word order between even closely related languages such as English and French (Pollock 1989).

This, then, appears to be the locus of syntactic, i.e. word order parametrization between languages: an inflectional feature may or may not be visible as an illegitimate object at PF. In English, a ⟨+wh⟩ Comp without a *wh*-specifier in its checking domain is an illegitimate PF object. Thus, if in a derivation D we were to Spell-out

(14) *I wonder* $[_{CP\ +wh}\ C^0\ [_{IP}$ *John likes what*]]

the derivation crashes at PF since C^0 lacking a *wh*-specifier is an illegitimate PF object. Categories that are visible as illegitimate objects at PF will have to be eliminated in overt syntax. By the principle of Procrastination those that are not visible at PF will not be eliminated in overt syntax, i.e., "if you don't have to move for convergence, then movement is prohibited." Features that are visible (thus: potentially illegitimate) at PF are called strong; features that are invisible (thus: inoffensive) at PF are called weak.[15]

A minimal assumption is that the strong/weak distinction constitutes the *only* parametric variation among languages. This implies that parametric variation is restricted to certain features (e.g., ⟨+wh⟩) of functional categories (e.g. C^0) (Fukui & Speas 1986). Furthermore it implies that there are no directionality parameters, such as directionality of government.[16] The latter implication is supported empirically by Kayne (1994), who argues that movement is invariably leftward.

---

15. See Koster (1986) for earlier use of this terminology in connection with word order variation.

16. In fact, government has no formal status in the Minimalist Program. For example, Case assignment is reduced to feature checking in a specifier-head configuration (see also section 2.6 below). The consequences for the Empty Category Principle, which incorporates the notion proper government, have not been fully explored.

1.7. *Directionality and Word Order*

The structure building process of iterative (cyclic) generalized transformational application nowhere specifies the linear order of head, complement, specifier, and adjunct. Superficial cross-linguistic examination suggests that languages may differ with respect to the linear order of these elements. In the tradition of generative grammar, the attested variation is described in terms of a parametric option: heads may govern to the left or to the right. A head that governs to the left takes its complement to the left in the initial representation, yielding an OV order at D–structure.

In the minimalist approach, a directionality parameter is no longer available. First, parametric variation must be expressed in terms of the features of functional heads only. A directionality parameter would therefore not suffice to account for the ordering of elements in the lexical domain. Second, government no longer plays a role in the minimalist approach. Therefore, it is unclear whether a directionality parameter could be reduced to properties of such an independently defined grammatical relation. Third, a directionality parameter would be redundant, since much of the word order variation can be accounted for by the interaction of overt and covert movement.

Kayne (1994) presents empirical evidence indicating that movement into the functional domain is invariably leftward. The evidence consists in what we do *not* find, in comparing cross-linguistic movement phenomena. Thus, we can hypothesize from the apparent absence of *wh*-movement to the right that the specifier position of CP is always to the left. Similarly, there appear to be no known cases where verb movement changes a verb–complement order from VO to OV, which suggests that verb movement to the right does not exist. Hence, the functional projections hosting V-features are all assumed to be head initial. Also, the subject precedes the object in almost all languages of the world (Greenberg 1963, Universal 1). Assuming, in connection with this, that AgrSP is hierarchically higher than AgrOP, it also follows that the specifier of AgrSP is situated to the left. Likewise, if the complement of a preposition is extracted, the complement always appears to the left of the preposition, never to its right. Again, this suggests that licensing positions, i.e. specifier positions, under our assumptions, are on the left hand side.

These empirical observations are presented more fully in Kayne (1994), who additionally argues that lexical projections are universally head initial as well. This is an attractive hypothesis, considering the empirical evidence for the universal structure of the projections of the functional domain. However,

empirical evidence in support of this hypothesis is considerably more difficult to obtain, in view of the fact that the observable word order reflects an intermediate stage of the derivation, i.e., one never knows whether the constituents are in a basic position or not.

Kayne (1994) also presents conceptual argumentation in support of the idea that all phrases are head-initial, proposing that asymmetric c-command invariably maps into linear precedence. In order for this mapping to be successful, it must be possible to express the relations between the nodes of a phrase marker that asymmetrically c-command into a set of ordered pairs ⟨x,y⟩ of the terminal (lexical) elements dominated by these nodes. The pairing of two terminal elements x,y thus expresses a relation between x and y. Kayne proposes that the set of ordered pairs of these relations must express a linear ordering, i.e. a total, transitive, and antisymmetric ordering.

Thus, according to this proposal it must be possible to read the relation of each terminal element to all other terminal elements off of the set of ordered pairs. Crucially, these relations must be antisymmetric, i.e., it is excluded that two terminal elements L each other, where L stands for the relevant relation between these two elements.

The axiom that the set of ordered pairs of terminal elements (derived from the set of relations between the nodes of a phrase marker that are in a relation of asymmetric c-command) is a linear ordering of the terminal elements is called the Linear Correspondence Axiom (LCA).

In addition to the LCA, Kayne proposes that the relation expressed by the pairing of terminal elements is a precedence relation. We will refer to this hypothesis as the Extended Linear Correspondence Axiom (ELCA).[17]

Kayne shows that the adoption of the LCA explains many basic facts of phrase structure, such as binary branching and endocentricity. In this respect, the LCA is compatible with the mechanism of generalized transformations as

---

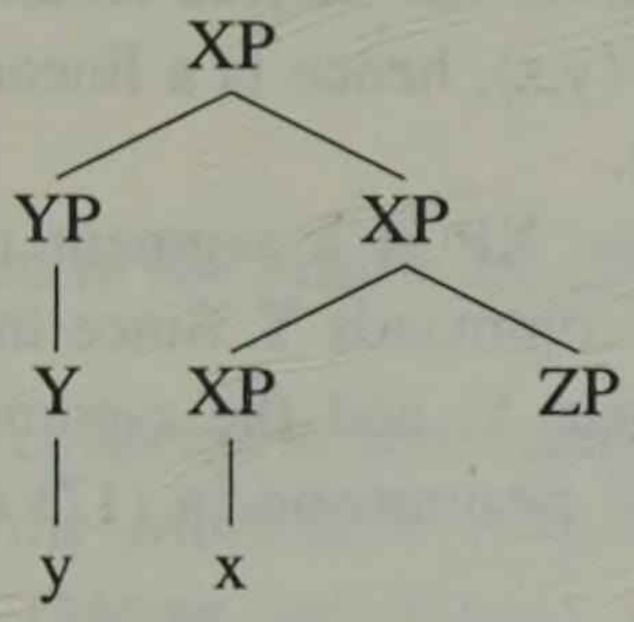
17. Kayne (1994: section 5.3) derives the definition of the relation between the terminal elements in a phrase marker as a precedence relation from the hypothesis that every phrase marker contains a root node dominating all other nodes except itself. On the assumption that this root node also dominates an abstract terminal element *a*, which, as Kayne argues, has to precede all other terminal elements of the phrase marker, it follows that the linear relation between *a* and the other terminal elements is also a precedence relation. Hence, the linear relation between terminal elements must always be a precedence relation. For empirical justification of the Extended LCA, see above.

presented in section 1.2. It follows from the ELCA that adjunction always takes place on the left-hand side.

In some respects, however, the LCA appears to be too restrictive, as Kayne notes. In fact, the LCA excludes adjunction of specifiers and adjuncts. Kayne therefore modifies the definitions entering into the notion c-command in order to allow adjunction of specifiers. He argues, however, that adjunction of adjuncts (i.e. in addition to adjunction of a specifier) is excluded.

To see why adjunction of specifiers is difficult, consider the tree structure in (15), where y and x represent terminal elements:

(15)



Assume the following definition of c-command:

(16) α c-commands β iff every γ that dominates α dominates β.

C-command is asymmetric where, for α c-commanding β, β does not c-command α.

In (15), YP asymmetrically c-commands X and XP asymmetrically c-commands Y. YP dominates the terminal element y, and X dominates the terminal element x. The relation between YP and X therefore can be expressed in the ordered pair of terminal elements ⟨y,x⟩. But since XP dominates x and Y dominates y, the ordered pair of these terminal elements ⟨x,y⟩ is also part of the set of ordered pairs expressing the relations between YP, XP, Y, and X. So now this set contains ⟨y,x⟩ and ⟨x,y⟩. Hence, the relation between x and y (i.e., between a head and its specifier) is not linear, because it is not antisymmetric.

To solve this problem, the pair ⟨x,y⟩ or the pair ⟨y,x⟩ must be excluded. This can be achieved if either YP does not c-command X (eliminating the pair ⟨y,x⟩) or XP does not c-command Y (eliminating ⟨x,y⟩). Kayne proposes to modify the definition of c-command in such a way that XP no longer c-commands Y. This can be done by excluding segments from the definition of

c-command, assuming the higher XP and the lower XP in (15) to be two segments of the same category:[18]

(17) α c-commands β iff
  (i) α and β are not segments, and
  (ii) α excludes β, and
  (iii) every γ dominating α dominates β.

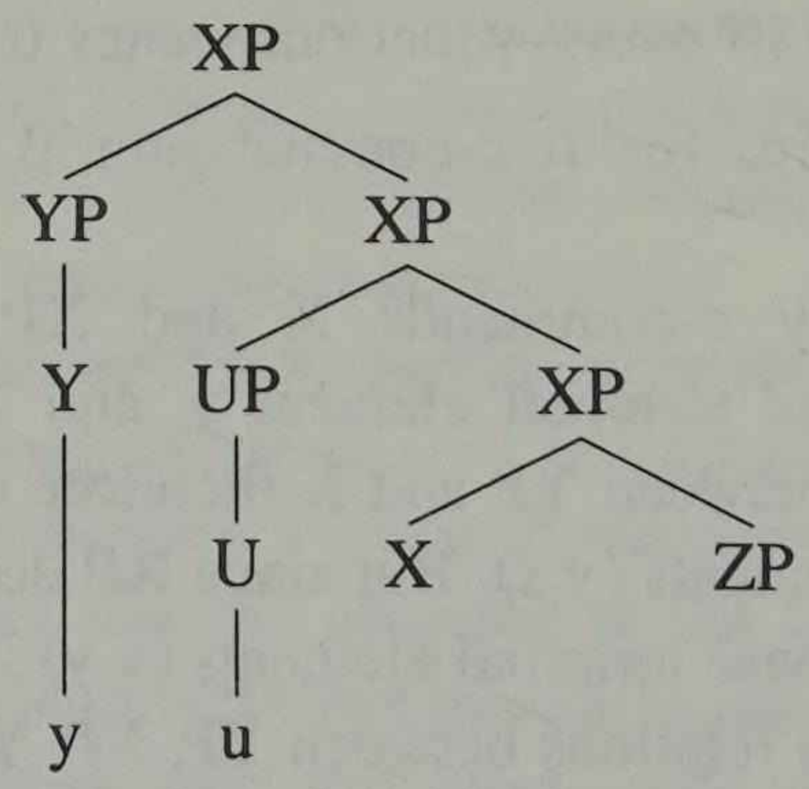
(18) α excludes β if no segment of α dominates β.

In (15), XP is a segment, hence does not c-command Y by clause (i) of the definition of c-command in (17). This gives the desired result that the relation between *x* and *y* in (15) is described by ⟨y,x⟩, hence is a linear relation (hence, following Kayne, a precedence relation).

Notice that the fact that the higher XP is a segment of the lower XP suffices to exclude that the lower XP c-commands Y. Since the higher XP is a segment, the lower XP does not exclude Y, and the c-command relation is barred by clause (ii) of the definition of c-command in (17) (cf. Kayne 1994, note 2 in chapter 3).

To see why (multiple) adjunction is difficult, consider (19):

(19)



In (19), not only the specifier UP, but also the adjunct YP is adjoined to XP. YP asymmetrically c-commands U, yielding ⟨y,u⟩, and UP asymmetrically c-commands Y, yielding ⟨u,y⟩. Hence the relation between the terminal elements y and u is not antisymmetric and therefore (19) is not allowed by the LCA.

Kayne concludes that multiple adjunction (i.e., adjunction of an element in addition to adjunction of a specifier, as in (19)) is universally impossible. It

---

18. Following May (1985: 57), Chomsky (1986b:7). For critical discussion of this segmental theory of adjunction see also Epstein (1987, 1989, 1991, 1993, to appear), Lasnik & Saito (1992), Williams (1988), and Poole (forthcoming).

follows that adjunct elements, such as adverbs, can be present only as specifiers. Thus, for every adjunct there must be a head in the structure creating the required specifier position.

It is important to note that the ELCA does not necessarily follow from the LCA. The ELCA merely interprets linearity in the mathematical sense in one-dimensional terms, yielding precedence. Chomsky (1995a) objects to the way Kayne (1994) employs virtual categories (such as intermediate bar nodes) to determine the relation between terminal elements. He therefore rejects the LCA as proposed by Kayne, while accepting the word order generalizations of Kayne in full.

2. Introduction to the Studies in this Volume

The studies collected in this volume cover a range of issues that have taken center stage in the Minimalist Program: (i) the basic operations that create phrase structure (Merge and Move), (ii) licensing of grammatical categories (movement and feature checking), (iii) economy conditions and restrictions on movement, (iv) the nature of phrase structure and the inventory of functional categories, (v) inflectional morphology and the nature of parametric variation, and (vi) the relation between hierarchical structure and linearization. In the remainder of this introduction, we will present these issues and discuss the contributions that the various papers make in the respective domains.

2.1. *Building up Phrase Structure*

As we have seen in section 1.2, Chomsky (1993) departs from previous analyses, in that the rules or mechanisms deriving phrase structures are considered to operate in a bottom-up fashion. In the tradition of generative grammar, phrase structure is described by a system of rewrite rules, which take a phrase and analyze it into its constituent parts:

(20)
$$\begin{array}{lcll} S & \rightarrow & NP & VP \\ VP & \rightarrow & V & NP \end{array}$$

The rewrite rules can be regarded as the generative counterpart of the immediate constituent analysis of Bloomfield (1933), and of the traditional analysis of a sentence into its parts of speech.

Within generative grammar, there has been a development of category

specific rewrite rules, such as the ones in (20), to the category-neutral rewrite rules of the X-bar Theory, illustrated in (21) (see Chomsky 1970):

(21)
$$\begin{array}{lcll} XP & \rightarrow & YP & X' \\ X' & \rightarrow & X^0 & ZP \end{array}$$

However, this development did not change the top-down character of the phrase structure rule system.

The X-bar Theory holds that phrases are structured according to the category-neutral rewrite rules of the type in (22). The rewrite rules convey two types of information. They express that there is a projection of categorial features running from the head of a phrase to the maximal projection of the phrase (the projection line X^0-X′-XP). Second, they express that the nodes making up the projection line are of different levels (X^0 being the head, X′ the intermediate projection, and XP the maximal projection).

Thus, X-bar Theory basically asks (and anwers) the question: If we have a phrase, what does it consist of? Chomsky (1993) turns the tables, and asks: If we have a syntactic object (a head or a phrase), what do we get when we integrate that object in a syntactic structure?

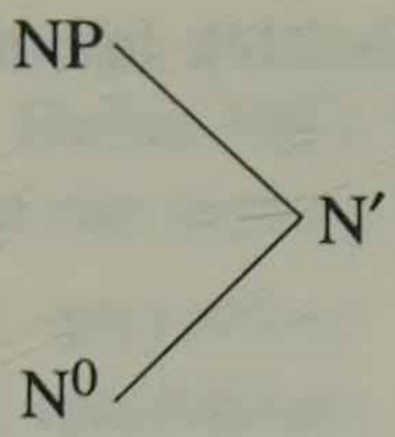
For example, if we want to integrate a head, say V, into a syntactic structure, we need to combine it with something, its complement, say DP, and the combination of V and DP yields a new syntactic object, V′. Hence the mechanisms Merge and Move describe how two syntactic objects are combined, and what syntactic object the combination yields.

In Chomsky (1993), X-bar Theory is not eliminated. The operations Merge and Move are subject to the principles of X-bar Theory:

> The computational system takes representations of a given form and modifies them. Accordingly, UG must provide means to present an array of items from the lexicon in a form accessible to the computational system. We may take this form to be some version of X-bar theory. The concepts of X-bar theory are therefore fundamental. In a minimalist theory, the crucial properties and relations will be stated in the simple and elementary terms of X-bar theory. (Chomsky 1993: 6)

This suggests that the top-down rewrite system and the bottom-up Move/Merge system are just notational variants. However, certain discrepancies do exist. For example, a descriptive system that uses the top-down rewrite system of X-bar theory must introduce the intermediate X′-level, even if no specifier or complement exists. Thus, NP is always rewritten in such a way that the structure in (22) results:

(22)



In the bottom-up system, N′ can only be the result of combining N^0 with a complement. In fact, if no complements and specifiers exist, it may be impossible to tell N^0 and NP apart.

What this example illustrates is that in the bottom-up system, phrase level is a relational concept, concomitant with movement processes, rather than an absolute concept given by rules. Therefore, the shift to a bottom-up system in Chomsky (1993) is more radical than the above quote suggests. It is, however, a perfectly natural shift, given the tendency in generative grammar since the 1970s towards eliminating rules altogether.[19]

The question then remains what Chomsky means by "the concepts of X-bar theory" which are taken to be "fundamental". This must relate to the two types of information that the rules of X-bar theory convey: categorial status and phrase structure level status. We want to maintain that there is a notion of a head that projects, and that the categorial features of the head are also present on the projections of the head. Secondly, we need to have an algorithm that tells us about the phrase structure level of the syntactic object resulting from the operation Merge/Move.

The issue of projection is discussed in much detail in Chomsky (1995a, b). Chomsky argues that if α adjoins to β, only β projects its features (1995b: 243f., 256f.). In other words, γ, the combination of α and β, can be defined as the set of nodes containing α and β, and a bundle of features present on γ, which derive from β, and which are represented as the label of γ:

(23) $\gamma = \{\text{LABEL}\{\alpha,\beta\}\}$

We will pass over the details of the argumentation here, as they are not relevant to the articles in this volume.

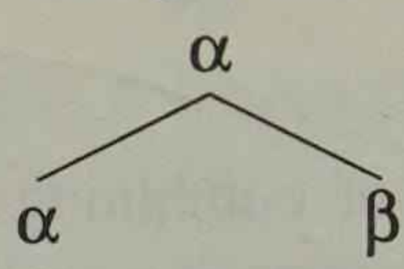
Note incidentally that the question of which node projects is relevant only to the operation Move, as the operation Merge is defined in such a way that

---

19. In Chomsky (1995b:256), the relation of the present theory with the traditional X-bar theory is made explicit. X-bar theory, it is said there, consists of "stipulated properties" which we can no longer make recourse to, "so we hope to show that the conventional assumptions are in fact derivable on principled grounds."

what projects counts as the head (Chomsky 1995b: 244). In other words, Merge combines a head and a nonhead:

(24) Merge



$\alpha = X^0$

$\beta = XP$

With the operation Move, a typical case is where two nonheads are combined (movement to specifier position or adjunction of XP), or (in the case of head movement) two heads. Hence the relation is symmetric, and the question which element projects its features becomes relevant.

The second type of information conveyed by the rewrite rules of X-bar theory concerns the phrase structure level of the various subtrees in the structure. In other words, how can we derive that γ in (24) is X′ or XP, rather than X^0?

Here Chomsky (1995b: 242) resorts to a relational definition of phrase structure, introduced by Muysken (1982). Muysken proposed that phrase structure level is not an inherent property of subtrees, but a relational property, to be expressed in terms of the features ⟨maximal⟩ and ⟨projection⟩. These two features yield the following combinatorial possibilities:

(25) a. ⟨+maximal,+projection⟩ XP
b. ⟨+maximal,–projection⟩ [?]
c. ⟨–maximal,+projection⟩ X′
d. ⟨–maximal,–projection⟩ X^0

As can be seen, these combinatorial possibilities yield the three conventional X-bar levels. ((25b) represents the case of a head that does not project. Possibly clitics are a case in point.) Importantly, however, the node immediately dominating a head is not necessarily X′ in this system. X^0 projects an X′ only if X′ is not the maximal projection of X^0 (i.e., if X′ is immediately dominated by another projection of X^0). If the immediate projection of X^0 is the highest node in the projection line of X^0, its features are ⟨+maximal,+projection⟩, corresponding to XP.

In this system, unlike in the traditional X-bar theory, X^0, X′, and XP are derived notions. In the phrase structure representation proposed in Chomsky (1995a,b), these notions do not occur (see (23)). Notice, however, that the relational definition of phrase structure level based on (25) involves two features only because we want to derive three levels of phrase structure. It has been

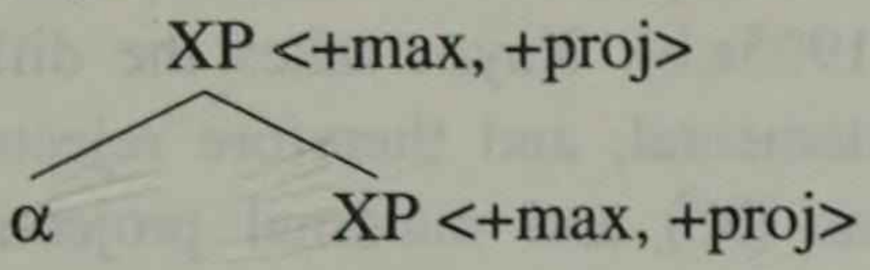
proposed several times that there is no syntactic distinction between XP and X′ (Stuurman 1985; Hellan 1991; E. Hoekstra 1991; Zwart 1992b; Kayne 1994). If this is correct, we can replace (25) by (26):

(26) a. ⟨−projection⟩ X^0
b. ⟨+projection⟩ XP

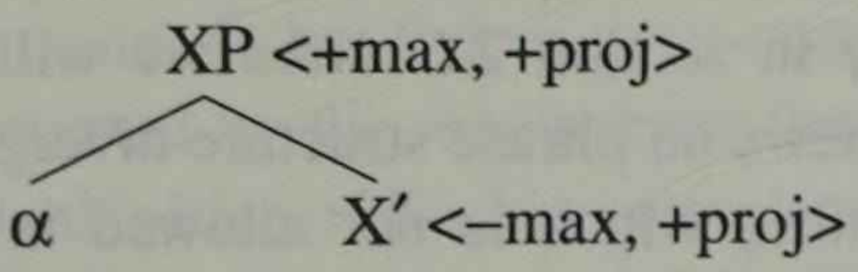
In (26), we seem to have reached a fundamental property of phrase structure: there is a distinction between heads and phrases. It is not clear that (25) reaches the same fundamental level. If not, the system in (25) is just one way of capturing a traditional, essentially stipulative property of X-bar theory, namely that we must distinguish three levels of phrase structure.

The question of the existence of an intermediate X′-level is relevant for at least two issues. First, Chomsky (1995b) maintains that there is a fundamental distinction between substitution and adjunction. Both substitution and adjunction are subcases of the operation Move. Adjunction differs from substitution in that adjunction of α to β yields a syntactic object of the same phrase structure level as β, whereas substitution yields a syntactic object that is one phrase level higher than β:

(27) a. adjunction



b. substitution



This distinction between substitution and adjunction cannot be expressed if phrase structure level is defined in terms of ⟨±projection⟩ only.

Notice that in (27a) the characterization of the lower XP as ⟨+maximal⟩ is not straightforward (as there is a second XP immediately dominating it). Therefore, the two XPs in (27a) are characterized as two "segments" of a single maximal projection (as in May 1985; Chomsky 1986a).

The choice between (25) and (26) is also relevant if we wish to construe an algorithm describing the features of γ, resulting from the combination of α and β (via Merge or Move). As argued in Zwart (1992b, 1993: 25), such an algo-

rithm would be more elegant if we start from the feature system in (26) (i.e., if we do not distinguish XP and X′). Informally speaking, Zwart proposes that γ takes its categorial features from the head (say α), and its phrase structure level from the nonhead (β):

(28) If β^n is combined with α^m, the result is α^n.

(where *m,n* specify phrase structure level, and *α,β* categorial features)

The algorithm in (28) covers the following cases (where Y = α):

(29) a. $XP + Y^0 = YP$
b. $X^0 + Y^0 = Y^0$
c. $XP + YP = YP$
d. $X^0 + YP = Y^0$

Of these, only (29d) does not exist. This, however, follows from the definition of Merge. (29d) involves a head, which can only be combined with a nonhead by projecting (see (24); we could stipulate that *m* in (29) must not be higher than *n*). The other cases are well described by the algorithm in (29), assuming no distinction between maximal projections and intermediate projections.

Kayne (1994) presents a view of phrase structure that differs significantly from that of Chomsky (1993, 1995a,b). Kayne takes the difference between heads and nonheads to be fundamental, and therefore rejects the distinction between intermediate projections (X′) and maximal projections (XP). Like Chomsky, Kayne wants to derive the well-known properties of X-bar theory from more fundamental notions. However, unlike Chomsky, Kayne wants to relate these notions to the linear order of the terminal nodes (see 1.7).

We will return to this issue in section 2.6. Here we will briefly indicate where Chomsky's and Kayne's views on phrase structure diverge. First, Kayne's system requires vacuous projection, which is not allowed in Chomsky's approach. Second, Kayne's system prohibits multiple adjunction, which is not prohibited in the minimalist approach.[20]

As we have seen, Kayne proposes to derive the linear relation between two elements α and β from the hierarchical relation between α and β (the LCA, see

---

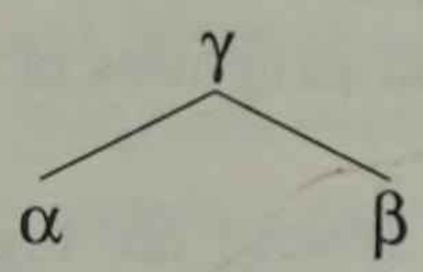
20. Chomsky (1995b:323) argues that adjunction is a severely restricted option in the minimalist approach, but multiple adjunction of specifiers is not. Recall that in Kayne's system, a specifier is technically an adjunct (a sister of XP), so that multiple specifiers in Chomsky's system are multiple adjuncts in Kayne's system.

section 1.7 above). This relation must be asymmetric, and is defined in terms of c-command (cf. Kayne 1994: 16):

(30) α c-commands β iff (i) α and β are not segments, and
(ii) α excludes β, and
(iii) every γ dominating α dominates β

In other words, the relation between α and β is well-defined only if there is a hierarchical relation between α and β (or between the projections of α and β) such that α c-commands β and β does not c-command α. Consider now the case where α and β are sisters:

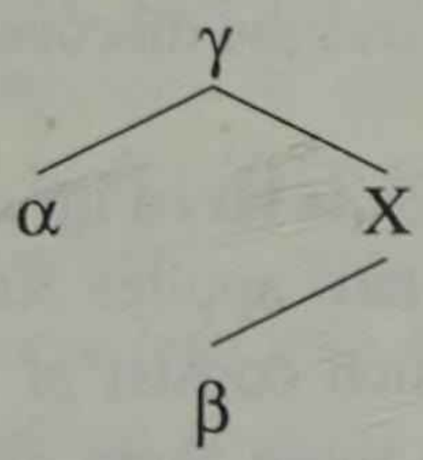
(31)



Here, α c-commands β and β c-commands α. Consequently, there is no hierarchical asymmetry between α and β, and the relation between α and β is not well-defined. Structures like (31) are excluded in Kayne's system.

There are two ways to establish a hierarchical ordering of α and β in (31). The first is to embed β deeper:

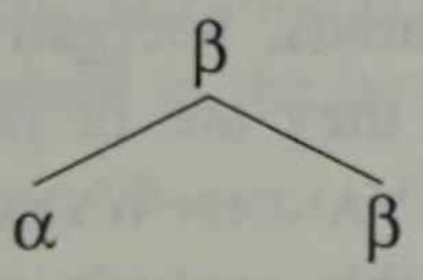
(32)



In (32), α c-commands β, but β does not c-command α. Notice that the presence of X is only motivated by the requirement that α and β be ordered. X does not result from the combination of β with a complement, as in Chomsky's system. Consequently, Chomsky prohibits structures like (32), whereas Kayne requires them.

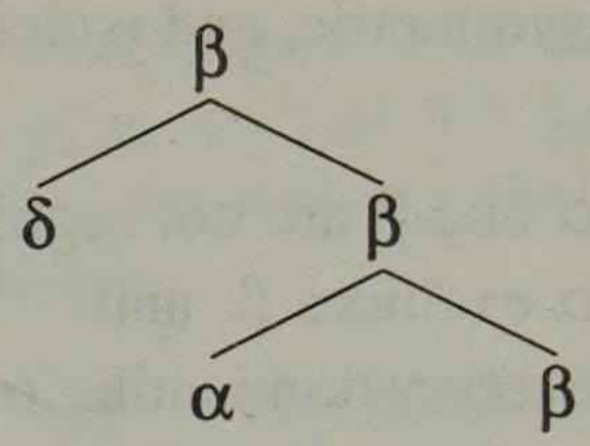
A second way of fixing (31) would be to consider γ a segment of β:

(33)



Being a segment, β does not c-command α by (30i). Notice that adding another segment would yield new problems, as discussed in section 1.7:

(34)



Here the problem is that δ and α c-command each other, as neither is dominated by β. As we have seen in 1.7, Kayne adopts Chomsky's (1986b: 7) modification of the definition of dominate here:

(35) α dominates β only if every segment of α dominates β

Thus, Kayne (1994) disallows multiple adjunction structures of the type in (34), whereas Chomsky (1993, 1995b) permits them.

The phrase structures allowed in Kayne's system are highly regular and maximally simple. Whenever a phrase seems to have two adjuncts adjoined to it (for example, a specifier and an additional adjunct, as in (34)), the second adjunct is assumed to be in the specifier of a higher functional projection. This higher functional projection is present only because the theory of possible phrase structure forces it to be there, not because its head hosts some inflectional feature. Again, this is a considerable departure from traditional thinking (within generative grammar) about phrase structure (see also the discussion of functional projections in sections 2.4–2.5 below).

The question of permissible phrase structures is taken up in this volume in VANDEN WYNGAERD's article. VANDEN WYNGAERD applies Kayne's system to the morphosyntax of Dutch past participles, which consist of a prefix (*ge-*), a verb stem, and an inflectional ending. Treating these three elements as being organized in a phrase structure that meets the requirements set by Kayne (1994), VANDEN WYNGAERD accounts for several features of past participle constructions in Dutch. Among these is the phenomenon that participles that have an infinitive in their complement domain are turned into infinitives in various Continental West Germanic dialects (the Infinitivus Pro Participio effect). VANDEN WYNGAERD also addresses the question of nonbranching projections. If these do not exist, as Chomsky (1995a: 399) holds, unergative verbs can be distinguished from unaccusative verbs only if they are in fact transitive (as argued independently by Hale & Keyser 1993). VANDEN WYNGAERD's analysis of Dutch past participles presents support for this analysis of unergatives as transitives.

The processes of Merge and Move raise a number of other questions which are discussed in the articles in this volume.

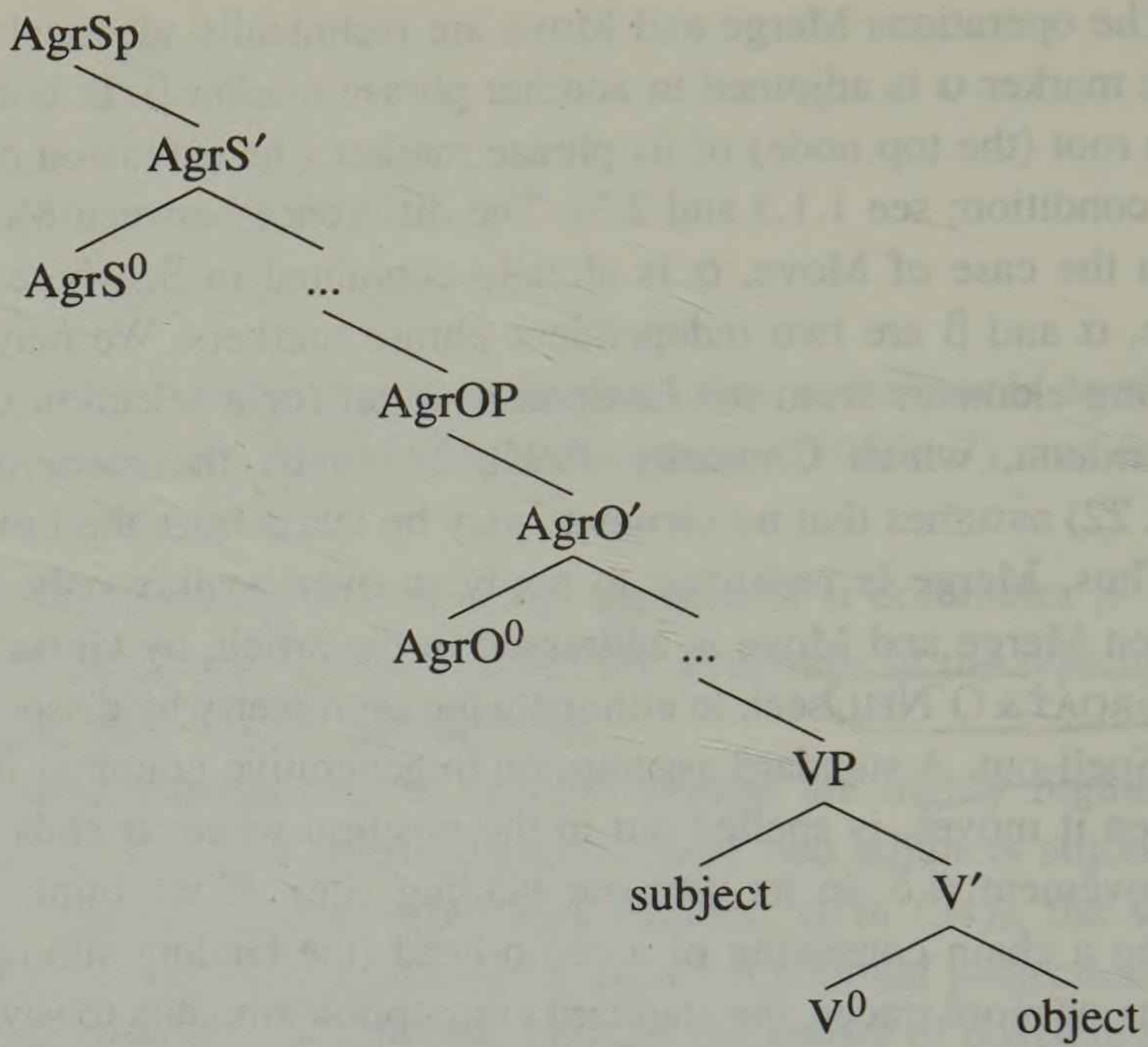
The operations Merge and Move are technically identical: in both cases, a phrase marker α is adjoined to another phrase marker β. In both cases, β has to be the root (the top node) of its phrase marker (the extension condition or strict cycle condition; see 1.1.3 and 2.3). The difference between Merge and Move is that in the case of Move, α is already contained in β, whereas in the case of Merge, α and β are two independent phrase markers. We may think of Merge as taking elements from the Lexicon as input (or a selection of elements from the Lexicon, which Chomsky 1995b: 225 calls the *numeration*). Chomsky (1993: 22) assumes that no elements may be taken from the Lexicon after Spell-out. Thus, Merge is restricted to apply in overt syntax only. This asymmetry between Merge and Move is addressed in the article by GROAT & O'NEIL.

GROAT & O'NEIL seek to eliminate the asymmetry by dissociating movement from Spell-out. A standard assumption in generative grammar is that an element α, when it moves, is spelled out in the position where α ends up as a result of the movement (i.e. in its ultimate landing site). If we think of movement as creating a chain consisting of a chain-head (the landing site of the movement) and one or more traces, the standard assumption amounts to saying that the head of a chain is the element that is spelled out. Chomsky (1993: 34–35) takes a trace of the moved element α to be a copy of α. GROAT & O'NEIL now propose that when α moves, either α or the copy of α (the trace) can be spelled out. In this proposal, languages do not differ with respect to the timing of movement, but with respect to the chain-member (copy) that is spelled out (pronounced). This makes it possible to restrict the operation Move to overt syntax, just like the operation Merge (see also section 2.7 below).

Another question raised in connection with the structure building operations Merge and Move concerns the structure of the functional domain. In traditional X-bar theory, the structure of the functional domain is encoded in the rewrite rules. Thus, since C selects IP, IP must be contained in CP. If Merge operates freely, the ideal case in a minimalist approach, generalizations about the dominance relations among the functional projections must be *derived* in some way.

One such generalization is that the functional projection licensing the object (AgrOP) is contained in the functional projection licensing the subject (AgrSP):

(36)



Chomsky (1993: 18) derives this generalization by imposing a locality condition on the movement operations that take the subject and object from their VP-internal positions to the specifier positions of AgrSP and AgrOP, respectively. These conditions are discussed in section 2.3 below. The question whether these conditions exist, and, if so, how they should be formulated, is addressed in the articles by POOLE and ZWART.

Baker (1988) has proposed that the hierarchical relations within the functional domain mirror the morphological structure of inflected forms. Thus, TP is supposed to be contained in AgrSP when the tense morpheme appears closer to the verb stem than does the agreement morpheme. This *Mirror Principle* is discussed further in section 2.4 below, and plays an important role in the papers by THRÁINSSON and SOLÀ. Discussion of the question of which functional projections must be assumed to exist in the grammar of a given language will also be postponed until section 2.4 below.

A final question regarding the operation Merge concerns the position in which the subject is generated (or, in minimalist terms, the moment in the derivation at which the subject is merged with the existing structure). Chomsky (1993: 8) adopts the by now standard assumption that the subject is generated in the specifier position of VP (merged with V′) (see (36)). In Chomsky (1995b: 331), this assumption is essentially maintained (albeit that VP is now

assumed to contain two VP projections, the subject being generated in the specifier position of the higher VP).[21]

The assumption that the subject is generated in the specifier of the VP is not challenged in the articles in this volume. We mention the issue, however, because this asumption is of crucial importance to several articles, including KITAHARA's (in which the trace of the subject in Spec-VP plays a role in the description of scope asymmetries), and all those articles that address the locality conditions on movement of subjects (FERGUSON, JONAS, POOLE, and ZWART).

Next to the operations Merge and Move, Chomsky (1993: 15) appears to propose a third type of operation forming syntactic objects, *Form Chain.* Form Chain plays a role in successive cyclic movement constructions, such as (37):

(37) *Who*$_i$ *did you say* [$_{CP}$ t_i *that Bill had kissed* t_i]*?*

In (37), the syntactic object created is the chain (*who,t,t*).

Successive cyclic movement typically involves movement out of a local domain (CP in (37)) via intermediate landing sites (Spec-CP in (37)) (Chomsky 1973, 1986b). In the minimalist framework, both the movement to the intermediate landing site and the movement to the final landing site must be triggered by feature checking requirements.

Here, a potential problem arises. It is reasonable to propose that movement to both landing sites is triggered by a requirement to check the *wh*-features of the *wh*-element (*who* in (37)). For the movement to the matrix CP, this is unproblematic. We may assume that the matrix C has ⟨+wh⟩-features, since the clause as a whole has interrogative properties. So we may assume that the *wh*-feature of *who* is checked in the Spec-CP of the matrix clause. But there is reason to believe that the *wh*-feature of *who* is also checked in the Spec-CP of the embedded clause. In many languages the embedded complementizer is affected in some way by the long distance movement, either by special morphology or because a different complementizer is used in that situation. In addition, there are languages in which the *wh*-movement can be partial, i.e., it moves the *wh*-element to the embedded CP only (in that case, an interrogative scope marker appears in the matrix Spec-CP). Both phenomena suggest that checking of the *wh*-features in the embedded CP cannot be excluded.

---

21. For the VP-internal subject hypothesis, see, among others, Kitagawa (1986), Fukui & Speas (1986), Koopman & Sportiche (1991), Burton & Grimshaw (1992), McNally (1992), and references cited there.

The problem, then, is the following. Suppose movement of *who* in (37) to the embedded Spec-CP is triggered by the requirement that the *wh*-feature of *who* be checked. Then, *who*'s *wh*-feature will be checked in the embedded clause, and (assuming that checing invariably entails deletion) there will be no feature left to check in the Spec-CP of the matrix clause. Thus, checking the feature in the embedded Spec-CP will remove the trigger for further movement to the matrix Spec-CP. Suppose on the other hand that movement of *who* to the embedded Spec-CP does not result in checking of the *wh*-features of *who*. This will make it possible for *who* to move on to the matrix Spec-CP, assuming that its features will be checked there. However, now it is unclear what the trigger is for moving *who* to the embedded Spec-CP in the first place, since the movement does not result in feature checking (and should be disallowed by Greed).

Underlying this problem is a conflict between the two economy requirements that the operation Move is subject to. These requirements are that steps in the movement process should be as few and as short as possible. The requirements are conflicting, because the number of steps decreases with the distance of the steps. In (37), *who* could have moved to the matrix Spec-CP in one step, satisfying the fewest steps requirement, but violating the shortest steps requirement.

To solve these problems, Chomsky (1993: 15) introduces the operation Form Chain, which creates (37) in a single step, forming the chain (*who*,*t*′,*t*) in the process. The nature of the operation is admittedly unclear. We will discuss it further below, in section 2.3. In this volume, POOLE and ZWART discuss Form Chain extensively.[22]

2.2. *Licensing: Movement and Feature Checking*

A basic theoretical and descriptive problem in syntax is to account for the fact that only certain syntactic positions are possible argument positions and different types of arguments are licensed in different positions. Thus subjects can occur in certain positions where objects cannot occur and *vice versa*. In the Government-Binding framework it is the task of Case theory to account for facts of this kind. The basic claim is that every overt DP (or NP) must be assigned

---

22. Chomsky (1995b) abandons the Form Chain analysis of (37), and proposes that checking of the *w*h-features of *who* in the embedded Spec-CP does not automatically lead to elimination of these *w*h-features. Not being eliminated, the features provide a trigger for further movement.

(abstract) Case (cf. e.g. the formulation of the Case Filter in Chomsky 1981: 49 and elsewhere, which assumes the basic idea of (abstract) Case usually attributed to Vergnaud (see, for instance, Rouveret & Vergnaud 1980; Vergnaud 1982)). Within the GB framework, then, one consequence of this is that if an argument is base-generated in a position where it cannot be assigned Case, the derivation will be illicit and the string predicted to be ungrammatical unless the argument is moved to a Case position (or passes through such a position in the course of the derivation).

As described in section 1.3 above, phenomena of this sort are handled in a somewhat different fashion in the Minimalist Program. Under the strict lexicalist approach assumed, arguments emerge from the Lexicon with full morphological specification, including Case features. The standard minimalist account is that the Case features (like all other morphosyntactic features) need to be checked in the course of the derivation, either by PF, if they are "strong", or in LF, if they are "weak" (see section 1.6 above and section 2.5 below). Thus Case assignment is replaced by Case checking. Furthermore, while Case assignment in GB theory is typically explained by reference to some notion of government, Chomsky (1993) assumes that Case checking uniformly takes place in a Spec–head relationship between an overt DP (NP) and a functional head (cf. section 1.3 above). More specifically, objects will have their Case checked in a Spec–head relationship with an AgrO head and subjects will have their Case checked in a Spec–head relationship with the functional head AgrS (or possibly T, cf. e.g. Chomsky 1993: 7 and 45 (fn. 11)). This is a very strong claim that raises a number of interesting questions, divisible into at least two groups: questions having to do with universal aspects of the feature checking approach and questions about cross-linguistic (parametric) variation in feature checking. Both types of questions are discussed in considerable detail in this volume. Here we will mention a couple of examples.

First, note that under the standard minimalist account just described, it is assumed that some sort of raising is involved in the Case checking of subjects and objects in all languages. Either the arguments raise overtly to the relevant Spec-AgrP position (or Spec-TP), thereby checking the (strong) Case features by PF, or else the relevant raising and feature checking (checking of weak features) will take place in LF.[23] As discussed in section 1.6 above and section 2.5

---

23. The exact formulation of feature checking is not the same in all of Chomsky's minimalist work, but it is not necessary to go into the details here. The reader can compare the formulations in Chomsky (1993: 27 ff., 1995a and 1995b: 279 ff.). Most of

below, the basic idea is that languages will vary with respect to which movements are overt and which ones covert, but all languages are assumed to have subject and object raising at some level for the purposes of morphological checking. This means, for instance, that an English sentence without overt object shift like *John never read the book* and a corresponding Icelandic sentence with overt object shift (i.e. *Jón las bókina aldrei*, lit. 'Jón read the book never') will have the same LF-representation, a unifying and welcome result given their apparent synonymy. One question is, then, whether we can get some explanatory mileage out of this purported similarity among languages, i.e. whether this way of approaching Case can be independently motivated, i.e. used to explain something else.

This is exactly the point made in KITAHARA's paper in connection with scope interpretation of quantified constructions. KITAHARA argues that core cases of scope interpretation phenomena that have been analyzed by appeal to the LF-rule of Quantifier Raising proposed by May (1977, 1985) are accounted for by the Case checking operations independently needed in the Minimalist Program. KITAHARA assumes (with Chomsky & Lasnik 1993) that each feature checking operation forms a distinct chain and formulates the following Scope Principle:

(38) A quantifier X may take scope over quantifier Y iff X c-commands a member of each chain associated with Y at LF.

Given this, plus standard minimalist assumptions about feature checking, KITAHARA accounts for the fact that (39a) is scope-ambiguous (i.e. *someone* can have scope over *everyone* or vice versa) whereas (39b) is not (*who* must have scope over *everyone*):

(39) a. *Someone saw everyone.*
b. *Who saw everyone?*

No rule of Quantifier Raising (QR) is needed to account for this difference in scope interpretation, nor for the other cases that KITAHARA discusses. This is obviously the kind of result that (if empirically adequate) every theoretical linguist (or more generally, anyone interested in explanation in any domain) dreams about: An analysis that is specifically designed to account for particular phenomena turns out to provide an explanation for apparently independent phenomena elsewhere and provides a solution to a difficult puzzle. Here the

---

the papers under discussion here take the formulations in Chomsky (1993) and/or (1995a) as points of departure, although some do incorporate references to ideas first discussed in Chomsky (1995b).

account is of particular theoretical interest since it follows from a central assumption in the Minimalist Program, something which differentiates it from previous accounts, namely the theory of morphological feature checking. KITAHARA's account raises the important question of whether the widely assumed rule of QR may be completely eliminable from grammar, an attempted reduction or minimization of operation types generally sought within the (aptly named) Minimalist Program. That possibility is raised in his paper, but he leaves the final answer to future research.

Another interesting issue raised by the generalized Spec–head approach to Case checking is how to account for cross-linguistic differences in the licensing of subjects and objects. One aspect of this variation is manifested in the so-called "multiple subject positions" found in some languages but not others. This difference is believed to be responsible for the fact that some languages allow the so-called "Transitive Expletive Construction" while others do not. This is illustrated with examples from the closely related languages Icelandic and Swedish in (40):

(40) a. *...að það hefur einhver stúdent sennilega stolið smjörinu.* (Ice)
that there has some student probably stolen the-butter
'...that some student has probably stolen the butter.'

b. **...att det har någon student antagligen stulit smöret.* (Swe)
that there has some student probably stolen the-butter

If we assume the VP-Internal Subject Hypothesis, it appears that the logical subject *einhver stúdent* 'some student' (the associate of the expletive) has moved out of the VP in (40a) as it precedes the sentential adverb *sennilega* 'probably', which is standardly assumed to be left-adjoined to VP in constructions of this sort. Under the minimalist approach, this means that the subject must be able to check some feature in this position in Icelandic. The corresponding derivation is unacceptable in Swedish, witness the ungrammaticality of (40b). Assuming that the position of the overt expletive, namely the position immediately preceding the finite verb, is the canonical subject position in both Icelandic and Swedish, we can describe the difference between the two languages in terms of the number of subject positions licensed.[24] In that sense Icelandic has an extra subject position which Swedish lacks. The question is how to account for this difference and this is one of the main questions that JONAS discusses in her

---

24. JONAS argues against the possibility that the position of the overt expletive is a topic position rather than a subject position, mentioning arguments given by Ottósson (1989) indicating that the overt expletive is quite different in nature from typical Spec-CP (i.e. Topic) elements.

paper. To have any explanatory power, such an account should link two or more phenomena. Otherwise it would be *ad hoc*.

Following a line of argument also found in work by Jonas & Bobaljik (1993) and others, JONAS maintains that the extra subject position licensed in languages like Icelandic is Spec-TP and the position where the overt expletive occurs is Spec-AgrSP.[25] Her main innovation is to link the licensing of Spec-TP to overt verb movement to T, claiming that a bare T does not project a specifier. She points out that only those Scandinavian languages that show general verb movement in embedded clauses (namely Icelandic and, apparently only dialectally, Faroese) are those which allow the transitive expletive construction. In addition, she maintains that there is evidence for this extra subject position in other constructions in these languages, including raising expletive constructions, but not in languages lacking overt verb movement to T. This difference is illustrated in (41) vs. (42):

(41) a. *Það virðist oftast einhver stúdent*$_i$ *vera* t$_i$ *í herberginu.* (Ice)
there seems usually some student be in the-room
b. *Það virðist oftast vera einhver stúdent í herberginu.*
there seems usually be some student in the-room
'There usually seems to be some student in the room.'

(42) a. **Det verkar vanligen någon student*$_i$ *vara* t$_i$ *i rummet.* (Swe)
there seems usually some student be in the-room
b. *Det verkar vanligen vara någon student i rummet.*
there seems usually be some student in the-room
'There usually seems to be some student in the room.'

Again, the claim is that the extra subject position licensed in Icelandic, but not in Swedish, is Spec-TP.[26] But since the indicated movement of the lexical argument in non-finite clauses like (41a) can hardly be for Case checking

---

25. For a different account of this extra subject position in languages like Icelandic see Chomsky (1995b: 354), where it is argued that the cross-linguistic difference involved is that T in some languages can license two specifiers but only one in others.

26. THRÁINSSON argues in his contribution to this volume that the crucial difference between languages like Icelandic and Swedish does not involve properties of T as such but rather that Swedish has an unsplit IP (i.e., it does not have AgrS and T as separate functional categories) whereas Icelandic has a split IP, namely both AgrSP and TP. Hence the difference in possible subject positions. He wants to relate this to differences in the morphological inventory of the languages. See the discussion in 2.5 below.

purposes, under the generally accepted assumption that non-finite verbs cannot check (Nominative) Case on lexical subjects, JONAS argues that the feature checking involved must concern the Extended Projection Principle (EPP). Her paper thus contains interesting proposals about the nature of feature checking and its relationship to (overt) verb movement.

One interesting prediction that follows from JONAS' analysis is that if a language has a transitive expletive construction (TEC), it should have overt verb movement (assuming that the associate of the expletive (the logical subject) is in Spec-TP in such constructions). Since Dutch for instance has TEC, it should have overt V-to-T. As is well known, this has been a matter of great debate, with the pendulum swinging towards "no-overt-V-to-T" analyses the past five years or so. Thus Reuland (1990) presents arguments against V-to-T in Dutch, and recent work which assumes that IP and VP are left headed in Dutch (see, for instance, the papers discussed in section 2.6 below and references cited there) also contains arguments to the same effect. JONAS' analysis provides an interesting contribution to this debate.

Finally, it should be mentioned here that at least two of the contributors to this volume suggest that Case checking may not take place exclusively in a Spec–head relation. One is THRÁINSSON, who mentions this possibility in connection with his claim that the Mainland Scandinavian languages (which include Swedish but not Icelandic) do not have an AgrSP as a projection separate from TP, and hence possibly not an AgrOP either if AgrSP and AgrOP are "indivisible" (cf. Chomsky 1993: 7). If that is correct, one would seem to be forced to some sort of Case-checking-under-government analysis for Mainland Scandinavian and other languages that do not have an AgrOP. THRÁINSSON does not pursue this point further, however. The other contributor suggesting alternatives to Case checking in a Spec–head relation is FERGUSON, who suggests that N^0 object incorporation is in fact a Case feature checking operation on a par with movement of a DP object to a specifier position, accounting for apparent exemptions from the Case Filter in incorporation structures (cf. Baker 1988). But as FERGUSON's paper is mainly concerned with the proper formulation of economy conditions (the concept "Shortest Move"), we leave further discussion of his paper to the next section of this Introduction.

2.3. *Economy Conditions and Restrictions on Movement*

In section 1.4 we briefly introduced the economy conditions and restrictions on movement of Chomsky (1993). Here, we will discuss these conditions and

restrictions a bit more, as they play a major role in most of the articles in this volume.

In Chomsky (1993), economy of derivation is expressed in two requirements, which we will refer to as the shortest steps requirement and the fewest steps (shortest derivation) requirement.

The shortest steps requirement prohibits movement across a potential landing site. It captures the basic phenomena adduced in support of the Relativized Minimality condition on movement (Rizzi 1990).

The fewest steps requirement entails that a derivation should exhibit as few steps as possible. This is expressed in the principle of Greed: superfluous movement is not allowed. Procrastinate, the principle that movement must be postponed until LF, can also be seen as an implementation of the fewest steps requirement.

We will discuss the shortest steps requirement and the fewest steps requirement in 2.3.1 and 2.3.2. Section 2.3.3 addresses the question of whether optional movement should (or can) be allowed in a minimalist approach, and section 2.3.4 discusses another condition on movement, the extension (or strict cycle) condition.

2.3.1. *Shortest Steps*

The hypothesis that long distance movement proceeds via a succession of local steps is probably one of the more important within generative grammar (Chomsky 1973). Much effort has since been devoted to the exact definition of what constitutes a local domain (see Koster 1978; Chomsky 1981; Kayne 1984; Chomsky 1986b; Rizzi 1990; Cinque 1990; Manzini 1992; and the references cited there).

The concept of minimality, introduced in Chomsky (1986b), relates locality to the domain of a governing head: no head may govern into the domain of another head. Chomsky (1986b) also proposes that head movement enlarges the local domain. This idea re-emerges in the Minimalist Program.

Relativized minimality (Rizzi 1990) introduces another locality condition, one for which the presence of a head is irrelevant. The condition basically states that movement of an element to an α position across another α position is not allowed, where α ranges over $\{X^0, A, \overline{A}\}$. In this approach, nonlocal head movement is not allowed because it skips a head position, nonlocal raising is not allowed because it skips an A-position, and nonlocal *wh*-movement is not allowed because it skips an $\overline{A}$-position. This idea also returns in the Minimalist Program.

In Chomsky (1993), the shortest steps requirement forces movement to the closest potential landing site. What constitutes a potential landing site depends on the type of movement, much like in the relativized minimality approach. Thus, for head movement, the closest potential landing site is the first head up, etc.

Assuming now that subjects are generated inside the VP, and that subjects and objects are licensed in the functional domain associated with V (i.e. outside the VP), a problem arises with respect to the movement of the subject and the object out of the VP. The positions in which subject and object are generated are A-positions (see section 2.4). The landing sites for subject movement (Spec-AgrSP or Spec-TP; we will leave this open here) and object movement (Spec-AgrOP) are also A-positions. Movement of the object across Spec-VP therefore violates the shortest steps requirement:

(43) $[_{AgrSP/TP}$ — $AgrS^0/T^0$ $[_{AgrOP}$ — $AgrO^0$ $[_{VP}$ SUBJ $[_{V'}$ V^0 OBJ]]]]

To solve this problem, Chomsky proposes that head movement of V to AgrO enlarges the local domain to which the shortest steps requirement is sensitive (in this case the local domain is VP). Chomsky's formulation of this proposal is that V-to-AgrO movement makes Spec-VP and Spec-AgrOP *equidistant* from the position of the object (OBJ). Since Spec-VP and Spec-AgrOP are equidistant from OBJ, Spec-VP is not a closer landing site than Spec-AgrOP, and movement of the object to Spec-AgrOP therefore satisfies the shortest steps requirement.

Chomsky's solution to the object movement problem now predicts that "overt object raising will be possible only with overt V-raising" (1993: 18). That there is a connection between overt verb raising and overt object raising has been suggested by Holmberg (1986). In fact, Chomsky's analysis appears to predict what is known in the literature as *Holmberg's Generalization*. Holmberg's Generalization is based on facts from Swedish and Icelandic, where movement of a pronoun (in Swedish) or noun phrase (in Icelandic) is not possible in past participle constructions. If we assume that the participle does not move, it appears that object movement is dependent on verb movement:

(44) a. *Johan köpte den inte* (Swe)
John bought it not

b. **Johan har den inte köpt*
John has it not bought

c. *Johan har inte köpt den*
John has not bought it

However, it is not clear that Holmberg's Generalization has the scope that Chomsky predicts it to have. For instance, Zwart (1993b, 1994b) argues that in the Continental West Germanic languages (Dutch, German, Frisian), where overt object raising takes place, the verb does not move to AgrO in embedded clauses (in main clauses, the verb is always fronted):

(45) a. *...dat Jan het boek niet koopt* (Dutch)
that John the book not buys
b. **...dat Jan het boek koopt niet*
that John the book buys not

There seem to be two ways to align the Continental West Germanic facts with the Scandinavian facts. First, we could hypothesize that the object raising in Continental West Germanic is not an object raising of the type under discussion here (i.e., A-movement), but movement to an $\overline{A}$–position. That there are two types of object movement in Dutch and German has been claimed various times in the literature (see Vanden Wyngaerd 1989a; Mahajan 1990). In this volume, DEN DIKKEN presents argumentation supporting the hypothesis that object shift in Dutch may be $\overline{A}$-movement. (FERGUSON addresses the same point in his footnote 11.) But this does not immediately solve the problem, assuming that the object has to undergo A-movement out of the VP before it can undergo the $\overline{A}$-movement . If Chomsky's prediction is correct, the A-movement would have to be accompanied by overt V-raising, contrary to fact (cf. (45b)). Second, one might assume that the Continental West Germanic languages differ from the Scandinavian languages and English in that AgrO is situated to the right in Continental West Germanic, and to the left in Scandinavian and English. This ties in with the traditional analysis of Dutch and German as head-initial languages (Bach 1962; Koster 1975), but is not compatible with Kayne's phrase structure generalizations (see 1.7 and 2.6).

In this volume, Holmberg's Generalization is discussed in the articles by DEN DIKKEN, FERGUSON, GROAT & O´NEIL, and THRÁINSSON.

DEN DIKKEN stresses the point that in Dutch there are two types of object movement: object raising to Spec-AgrOP, and an $\overline{A}$-movement which he calls "scrambling".

FERGUSON proposes a relaxation of the shortest move requirement of Chomsky, essentially refining the notion of "potential landing site". In FERGUSON's article, not every A-position, for example, is a potential landing site for every A-movement. A given position α is a potential landing site for β only if β's features can be checked in α. This solves the object shift problem immediately: Spec-VP is not a position in which the object's features can be

checked. FERGUSON points out that his reformulation of the shortest move requirement still makes it possible to derive Holmberg's Generalization. The crucial assumption here (also made in Chomsky 1993: 8, and Chomsky 1995b: 352) is that the object is not checked by AgrO but by V. An object can be licensed either by incorporation of N^0 into V (an option not available in Scandinavian languages), or by movement to Spec-AgrOP. But since V has to check the object, the object raising has to be accompanied by verb raising.

GROAT & O'NEIL, while accepting Holmberg's Generalization, reject Chomsky's analysis of object shift, in particular the idea that verb movement makes specifier positions equidistant. Instead, GROAT & O'NEIL link the presence of the object in Spec-AgrOP to the presence in AgrS/T of the verb selecting the object (here, the discussion is restricted to Icelandic, which has fronting of finite verbs in both main and embedded clauses).

Finally, THRÁINSSON discusses Holmberg's Generalization in the light of an analysis that assumes an impoverished functional domain for some languages but not others.

ZWART's article takes an entirely different stand. ZWART takes the incompatibility of the shortest steps requirement and the fewest steps requirement as a starting point, and argues that the shortest steps requirement is superfluous. One of the arguments is that most cases of nonlocal movement that we want to rule out involve movement across a position where certain features must be checked by the moved element.

In defense of the shortest steps requirement, it should be noted here that it is this requirement that derives (part of) the structure of the functional domain. Thus, AgrOP is contained in AgrSP/TP because if AgrSP were contained in AgrOP, the subject would have to move first, and the object would have to cross too many A-positions:

(46) $[_{AgrOP}$ — $AgrO^0$ $[_{AgrSP}$ — $AgrS^0$ $[_{VP}$ SUBJ $[_{V'}$ V^0 OBJ]]]]

According to the definitions in Chomsky (1993: 11f.), Spec-AgrOP cannot be made equidistant to both Spec-AgrSP and Spec-VP. This is because verb raising is a stepwise head-adjunction process, involving in this case a V-to-AgrS step and a AgrS-to-AgrO step. Equidistance is calculated after each step, and because of that, the number of equidistant specifier positions is never greater than 2. Thus, if Merge generated the structure in (46), the object could never be checked. The structure in (43) is therefore postulated.

2.3.2. *Fewest Steps*

In the Principles and Parameters framework (Chomsky 1981), application of the transformational rule Move α was unconstrained. Overgeneration was blocked by principles governing the output of movement (such as the Empty Category Principle).[27]

Perhaps the heart of the minimalist movement theory is the hypothesis that movement can apply only if it is triggered by the requirement that some formal feature be checked. Checking results in elimination of the feature. Since the relevant features are there for syntactic purposes only, failure to eliminate them would lead to a violation of Full Interpretation: the output representation of the syntactic component would contain elements that are not interpretable outside the grammatical component.

An even stronger requirement states that α can move only if α carries a feature that is checked (i.e. eliminated) as a result of the movement. This is the principle of Greed (section 1.4).

In Chomsky (1995b), two major changes regarding movement and feature checking are proposed. First, it is proposed that checking does not automatically result in elimination of the feature. Some features, such as categorial features, need to be interpreted at LF, and therefore cannot be eliminated (even if they must be checked). (The *wh*-features of *wh*-elements are among these interpretable features, as we have seen above.) Second, the trigger for movement has been changed: the trigger used to be an unchecked feature of the element that moves, but now the trigger is an unchecked feature of the head targeted by the element that moves. We could say that a head that has an unchecked feature "looks for" a moveable element that has the corresponding formal feature (e.g., T^0 looks for a noun phrase with nominative Case features), and *attracts* it. The consequence is that an element that has moved and has its features checked (such as a *wh*-element that has moved to Spec-CP in the embedded clause) can still move on to check features of some higher head (the matrix C^0), provided the relevant features are of the interpretable type and have not been eliminated. In this volume, the articles address the movement theory of Chomsky (1993), which incorporates Greed.

---

27. Nevertheless, it was always clear that some connection between movement and "something else" had to be there, otherwise the movement would be an ad hoc element of the analysis. This "something else" could be formal licensing of noun phrases (Case assignment), overt expression of scope (*wh*-movement), linking of a verb to tense (verb raising), etc. In other words, even though movement was unconstrained in principle, there was a strong tendency to avoid arbitrary movement.

Another principle that could be related to the fewest steps requirement is the Procrastinate principle. Movement is covert, unless the feature triggering movement is "strong". As a result, overt syntax is characterized by a well defined, restricted number of movements, with no arbitrary movements allowed.

As has been noted above, the fewest steps requirement and the shortest steps requirement appear to be incompatible. Chomsky (1993b: 15) solves this problem by introducing the operation Form Chain. The articles by ZWART and POOLE are attempts to further define this operation.

ZWART argues that Form Chain should be decomposed into the two familiar structure building operations, Merge and Move. In the particular case of long distance *wh*-movement, he assumes that the intermediate trace is in fact an empty *wh*-element which is merged with the embedded C′ (i.e., base generated in the Spec-CP of the embedded clause). The lexical *wh*-element then moves in one step from its base position to the Spec-CP of the matrix clause. This derivation maximally satisfies the fewest steps requirement, and does not incur the problem that features checked in the intermediate landing site should still be visible to trigger further movement. In this analysis, the lexical *wh*-element in the matrix Spec-CP, the empty *wh*-element in the embedded Spec-CP, and the trace of the lexical *wh*-element are all combined in a chain, which is needed to link the lexical *wh*-element to its theta-position.

POOLE maintains successive cyclic movement as a succession of movement steps, but makes the Form Chain mechanism more explicit. Form Chain operates on the result of the Move operations, and "hooks together" the moved category and its traces. POOLE then proposes that only the Form Chain operation is subject to Greed, while Move is not. This solves the problem of successive cyclic movement containing non-greedy movements (i.e. movement of elements that have their features checked already). Since Move is not subject to Greed, the fewest steps requirement must be thought of as being less crucial to economy of derivation than the shortest steps requirement.

Thus, while ZWART argues against the shortest steps requirement, and wants to give the fewest steps requirement center stage, POOLE argues for the exact opposite.

Greed plays an important role behind the scenes in FERGUSON's article. FERGUSON essentially relates "shortness" to the presence of relevant features (i.e., to Greed). Elements must make the shortest greedy movement, which amounts to saying that the fewest steps requirement is more important than the shortest steps requirement.

The Procrastinate principle is addressed in the article by GROAT & O'NEIL.

Procrastinate expresses an asymmetry between overt and covert syntax, and GROAT & O´NEIL's article can be seen as an attempt to eliminate such asymmetries. GROAT & O´NEIL argue that Procrastinate is a "global economy" principle, a principle comparing possible (convergent) derivations. They seek to reformulate Procrastinate as a purely local constraint that operates upon one derivation at a time. The heart of the proposal is that movement always takes place in overt syntax, but that languages differ as to which of the copies (the head or the tail of the chain) is spelled out. This can be expressed in terms of strong and weak features (as in standard minimalism), but in this approach, we no longer have to say anything about which copy is spelled out in the default case. Procrastinate is precisely such a default rule.

2.3.3. *Optionality*

It seems that in the strictest implementation of the minimalist framework, optional movement is not allowed. In Chomsky (1991: 431), optional movement patterns were considered to be the result of two derivations being equally costly. In the minimalist approach, this would have to be reformulated in terms of features being optionally strong, an unattractive result.

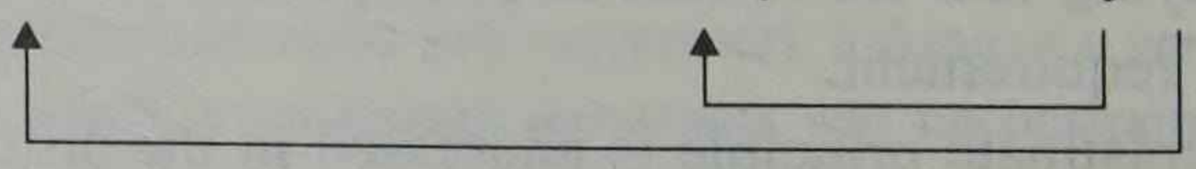
This problem is addressed in POOLE's article. As we have seen in section 2.3.1, POOLE proposes to separate Move and Form Chain, and argues that only Form Chain is subject to economy of derivation. In other words, movement is "cost free". POOLE then proposes that optional movement occurs whenever Move is not accompanied by Form Chain. In those cases, a derivation with movement is not more costly than a derivation without movement. This re-introduces Chomsky's (1991) analysis of optionality as resulting from two derivations that are equally costly.

2.3.4. *Strict Cyclicity*

The strict cycle condition (Chomsky 1973) was introduced into syntactic theory as a rule ordering constraint needed to account for island effects in successive cyclic movement constructions. Consider the *wh*-island construction in (47a), schematically represented in (47b):

(47) a. **When*$_i$ *did you wonder who*$_j$ *Bill kissed* t_j t_i

b. $[_{CP}$ *when*$_i$ C^0 $[_{IP}$ ... V^0 $[_{CP}$ *who*$_j$ C^0 $[_{IP}$ ... t_j t_i]]]]



In (47), *who* occupies the intermediate landing site for *when*, blocking successive cyclic movement. The strict cycle condition is needed to ensure that *who* moves into the Spec-CP of the embedded clause before *when* moves. In other words, we must make sure that the island configuration is not created after the movement it is supposed to block has taken place. (See Freidin 1978 for argumentation that the strict cycle condition is superfluous in this domain if we assume that successive cyclic movement leaves traces.)

In the minimalist framework, the strict cycle condition is reintroduced as a condition on the operations Merge and Move, stating that Move can only target the root of a phrase marker. Movement of *who* in (47) can target the root only if it applies before the embedded CP is combined with the matrix V^0 (hence, also before *when* moves). The strict cycle condition (or "extension condition") ensures that tree structures are built from the bottom up, in an orderly fashion. As illustrated in (47), Chomsky employs the condition in order to complement his analysis of Relativized Minimality phenomena in terms of the shortest move condition. In addition, the extension condition prohibits movement into complement position (which makes other principles prohibiting this, such as the Theta Criterion and the Projection Principle, superfluous in this domain).

Chomsky notes, however, that the extension condition does not operate without exceptions. Head movement does not seem subject to the extension condition. Chomsky also tentatively restricts the extension condition to substitution (i.e., countercyclic adjunction is allowed) in overt syntax (i.e. covert countercyclic substitution is allowed) (1993: 24).

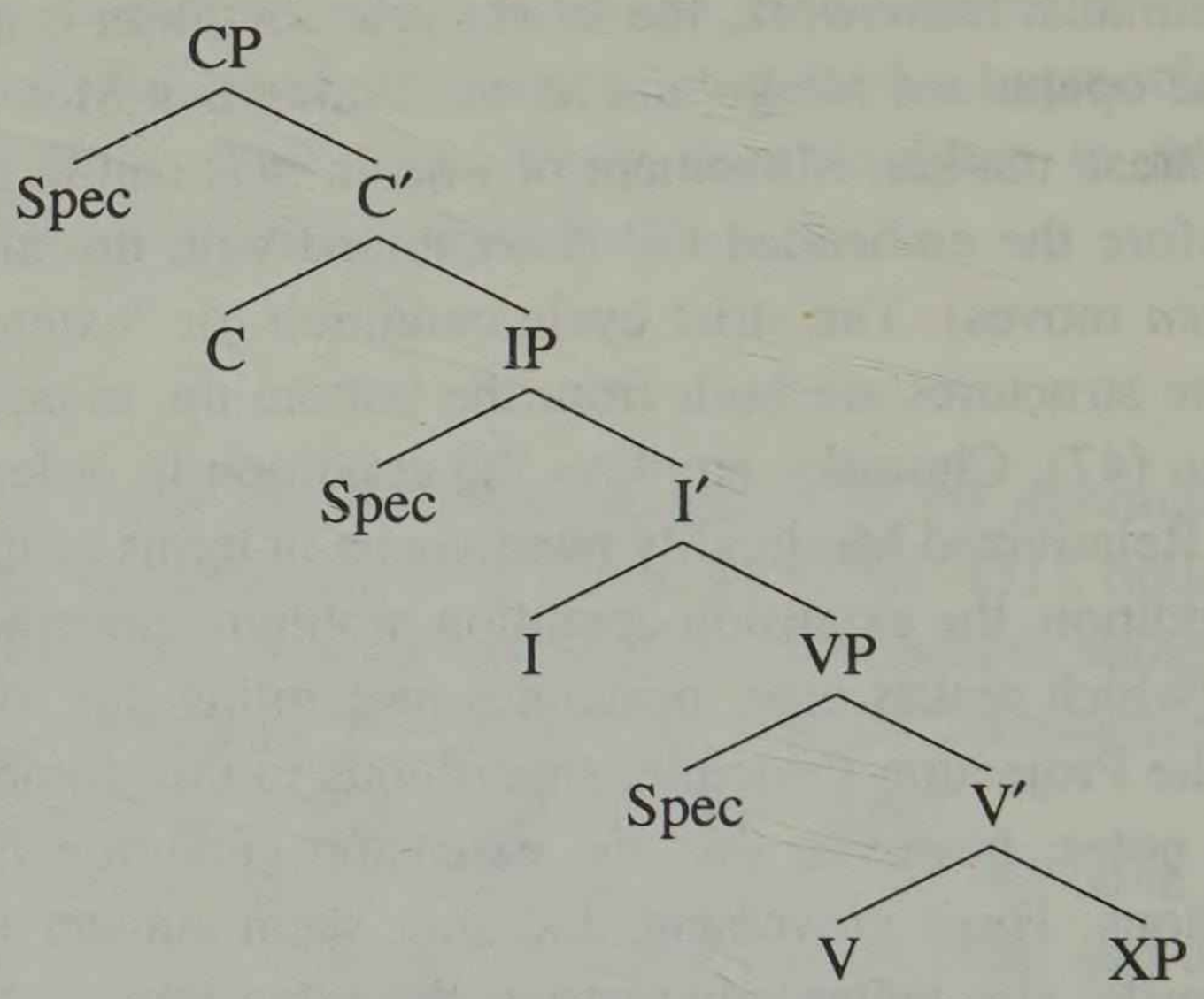
The asymmetry between overt and covert syntax is eliminated in GROAT & O'NEIL's article. In their analysis, all movement is overt, and language variation is a matter of spelling out the higher or the lower copy (the head or the foot of the chain). Hence, there is no covert substitution (for example, movement of the object to Spec-AgrOP in languages without overt object raising, like English).

The remaining exception to the extension condition, adjunction, does not seem to have any substance outside the domain of head movement. (See section 2.1 for the somewhat tenuous distinction between substitution and adjunction in the minimalist framework.) This suggests that head movement has a special status, and that all XP-movement can be constrained by the extension condition. There have been several recent attempts to derive the special status of head movement (Kitahara 1994; Watanabe 1994; Collins 1994). ZWART's paper also addresses the issue, proposing that the extension condition prohibits projection of an XP inside another XP. This would make countercyclic head movement possible.

2.4. *Phrase Structure and Category Types*

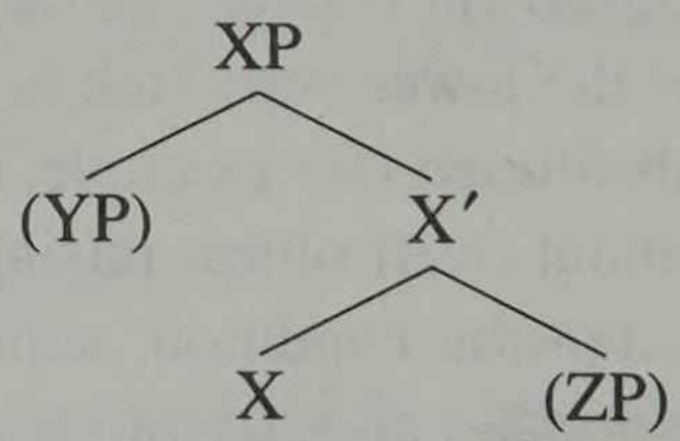
As discussed in sections 1.2 and 1.5 above, the distinction between lexical categories on the one hand and functional categories on the other is a crucial one in most contemporary theories of syntax, although the proposed inventory of functional categories will vary. Consider for instance the structure in (48), which would be typical for, say, Chomsky (1986b) and later work:

(48)



This structure conforms to the X-bar theory outlined in Chomsky (1986b: 3), which states that syntactic projections conform to the so-called X-bar-schema illustrated in (49) (leaving aside adjunction):

(49)



XP in (49) is the maximal projection of the head X, YP is its specifier and ZP its complement. In a structure like (48), the projection VP would be a (maximal) lexical projection of the (lexical) head V, whereas CP and IP are (maximal) functional projections of the (functional) heads C and I. XP in (48) stands for the complement, which could be a maximal projection of various kinds, including CP and NP (or DP).

In the Government-Binding framework, it is assumed that "not all positions are created equal." More specifically, it is assumed that there is an important

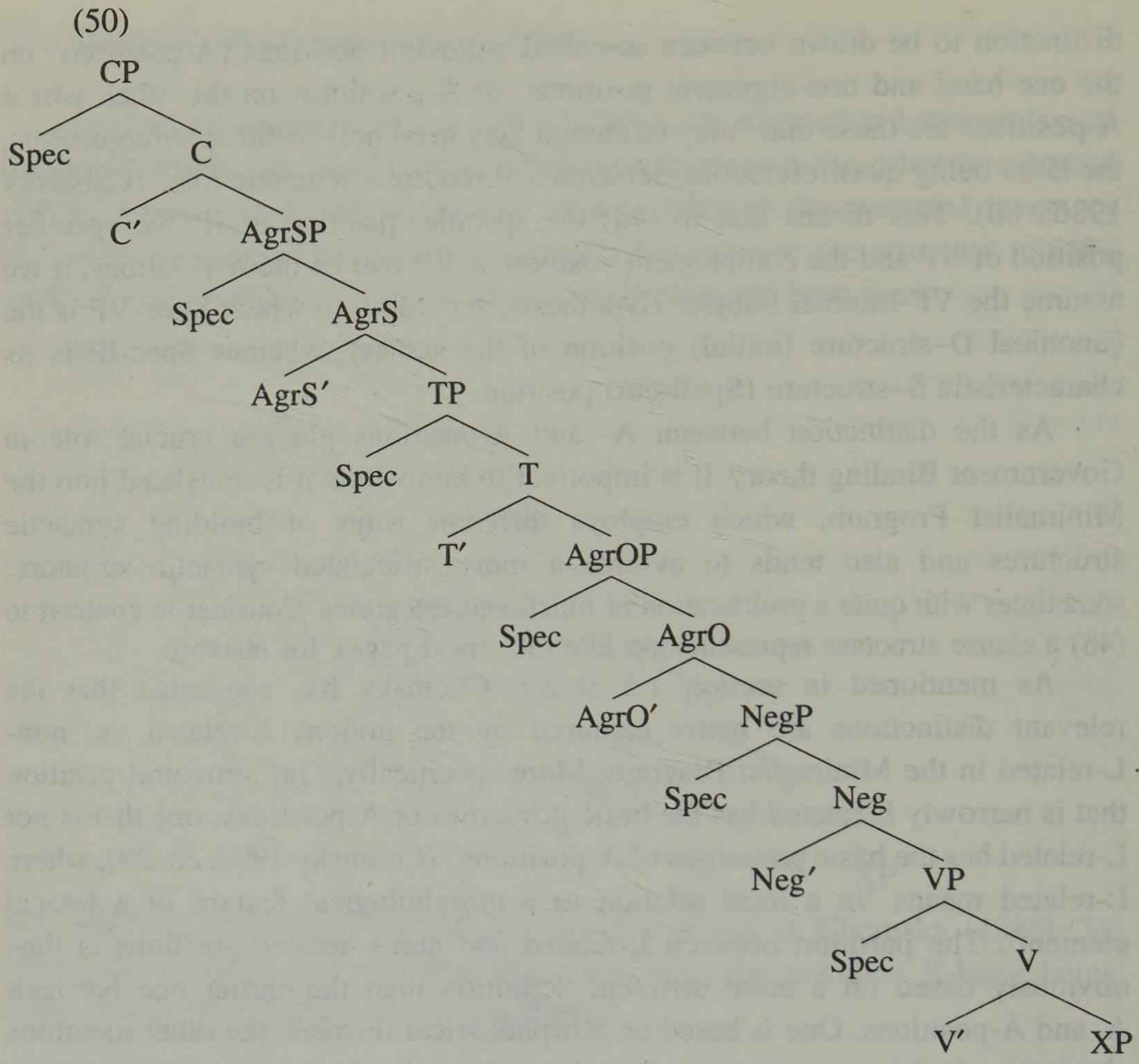
distinction to be drawn between so-called argument positions (A-positions) on the one hand and non-argument positions (or $\overline{\text{A}}$-positions) on the other, where A-positions are those that "may (although they need not) be filled by arguments, the latter being quasireferential elements that require a semantic role" (Chomsky 1986a: 80). This means that in (48) the specifier position of IP, the specifier position of VP and the complement position of VP will be the A-positions, if we assume the VP-Internal Subject Hypothesis, according to which Spec-VP is the canonical D–structure (initial) position of the subject, whereas Spec-IP is its characteristic S–structure (Spell-out) position.

As the distinction between A- and $\overline{\text{A}}$-positions plays a crucial role in Government Binding theory, it is important to know how it is translated into the Minimalist Program, which employs different ways of building syntactic structures and also tends to assume a more articulated syntactic structure, sometimes with quite a proliferation of functional categories. Consider in contrast to (48) a clause structure representation like (50) (next page), for instance.

As mentioned in section 1.5 above, Chomsky has suggested that the relevant distinctions are better captured by the notions L-related vs. non-L-related in the Minimalist Program. More specifically, "[a] structural position that is narrowly L-related has the basic properties of A-positions; one that is not L-related has the basic properties of $\overline{\text{A}}$-positions" (Chomsky 1993: 28–29), where L-related means 'in a local relation to a morphological feature of a lexical element'. The partition between L-related and non-L-related positions is thus obviously based on a quite different definition than the earlier one between A- and $\overline{\text{A}}$-positions. One is based on morphological features, the other mentions elements requiring a semantic (i.e. thematic) role. It is thus an interesting empirical question whether the distinctions made on the basis of these different definitions are empirically equivalent. This is one of the main concerns of HAEGEMAN's paper in this volume.

HAEGEMAN discusses different types of clitics, in both Germanic and Romance, analyzing their similarities and differences. Whereas Romance clitics are argued to be verbal and to interact with verb movement, Germanic clitics are analyzed as non-verbal and move to a functional head that is not L-related in the minimalist sense. HAEGEMAN provides an overview of some of the discussion of the L-related/non-L-related typology of syntactic positions within the Minimalist Program and relates this to her analysis, with the main emphasis on the behavior of West Flemish clitics. She claims that although the distinction between L-related and non-L-related positions is a linguistically significant one, it does not in fact replace the A- vs. $\overline{\text{A}}$-distinction. In particular, she maintains that there

(50)



appear to exist A-positions that are not L-related. This is clearly a central and hence important claim regarding the typology of syntactic positions specified in UG. It is well known that a simple A/$\overline{\text{A}}$- distinction seems insufficient in many cases, e.g. in connection with the analysis of the properties of scrambling (see e.g. the discussions in Webelhuth 1992 and Mahajan 1990) and it would be interesting to see whether HAEGEMAN's proposal can help shed new light on some old A/$\overline{\text{A}}$-problems.

A related issue concerns similarities and differences between specifiers and adjoined phrases. Chomsky (1993: 28) states that the specifier of a head with an L-feature (a morphological feature) is narrowly L-related whereas a category adjoined to the maximal projection of such a head would be broadly L-related. As we saw above, Chomsky suggests that only the first type (i.e. the narrowly L-related position) has the properties of an A-position. This is quite standardly assumed in much recent syntactic work. The nature of the whole issue becomes

somewhat different in a framework where multiple specifiers of a single head are allowed, as assumed, for instance, in some of Chomsky's most recent work (see Chomsky 1995b and references cited there). The multiple specifier analysis provides an alternative to the AgrSP+TP analysis advocated in the papers by JONAS and THRÁINSSON in the present volume in connection with transitive expletive constructions and other "multiple subject" constructions (see the discussions in section 2.2 above and 2.5 below). But, as pointed out by these authors, it is not immediately clear what the multiple specifier analysis entails regarding the syntax of the verb second phenomenon as it appears in transitive expletive constructions, for instance. The reason is that if we assume that the expletive occurs in the higher specifier and its associate in the lower specifier of the same projection (say TP) in such constructions, the appearance of the verb between the two rather than in the head position "below" both (i.e. in T in this case) has no obvious syntactic analysis. Hence the multiple specifier analysis is forced to postulate that the second position of the verb in such instances is a "phonological" phenomenon.

Finally, as a comparison of the structures in (48) and (50) above suggests, work in the Minimalist Program tends to assume a large number of functional categories. The suggestion that the structure of the Inflection Phrase (IP) is more complex than illustrated in (48) and that it should be decomposed into an Agreement Phrase (AgrP) and a Tense Phrase (TP) goes back to Pollock's influential (1989) paper, thus predating minimalism. But functional categories play an even more important role in the Minimalist Program than before. Consequently it is important to ask about their precise role(s) and universality. Do different clause types, for instance, vary with respect to which functional projections they instantiate? Although it sometimes seems to be assumed that the functional categories present in all types of clauses are invariant, it should be pointed out that the common hypothesis that Small Clauses exist (i.e. clauses without an I-projection, see e.g. Stowell 1983), presupposes that clause types may indeed vary in terms of the functional categories they contain. This question is discussed in some detail in THRÁINSSON's paper in this volume, as will become clear in the following section.

2.5. *Universality, Parametric Variation and Morphological Strength*

2.5.1. *Some Basic Ideas*

As discussed in sections 1.1 and 1.6 above, two of the major concerns of modern linguistic theory are arguably the following:

(51) a. What is universally true of all grammars of natural languages — and, by hypothesis, known by humans in advance of any linguistic experience?

b. To what extent can human grammars vary and how does the "learner" come to know the particular variant commonly referred to as "the language the child has learned"?

This emphasis on discovering what is universal and what is not is particularly apparent in the so-called Principles-and-Parameters approach to syntax. The origin and nature of the basic ideas underlying this approach is described in a very accessible fashion in Williams (1987) and Freidin (1991), for instance, and the approach is further explicated by Chomsky & Lasnik (1993).

An important contribution to this discussion of universality and linguistic variation is Pollock's seminal (1989) paper on verb movement and clause structure. As is well known, Pollock's main innovation was to suggest that it is possible to account for certain cross-linguistic differences if one assumes that the categorial inventory includes an Agreement (Agr) projection and a Tense (T) projection rather than just a single Inflection (I) projection as had been previously assumed. Pollock further suggested that the Agr-projection could have different properties cross-linguistically, depending on whether the language in question had a "morphologically "rich" system of agreement inflections" or not (see, e.g., Pollock 1989: 418–419). As Pollock points out, Roberts (1985) had also attempted to explain cross-linguistic syntactic differences in terms of varying degrees of richness of inflection but without assuming a separate Agr-projection. In further development of Pollock's approach to clause structure, Chomsky (1991) and Belletti (1990) also discuss the notion of "morphological strength/richness" employed by Pollock. Inspired by research of this kind, many syntacticians investigating comparative Germanic and Romance syntax have tried to determine how "rich" verbal morphology must be in order to make the relevant inflectional projections "strong". More specifically, what these authors are mainly concerned with is that there seems to be evidence for overt verb movement in some languages but not others and the presence of overt verb movement seems to have some correlation with richness of inflection. More recent research on comparative Germanic and Romance syntax has sought to define the relevant notion of "rich inflection", including work by Roberts (1993), Rohrbacher (1994), Vikner (1995a, 1995b), and others. We shall return to this question below.

Another attempt to account for the relationship between morphology and syntax is embodied in Baker's Mirror Principle (see e.g. Baker 1985, 1988),

which maintains that there is a direct relationship between syntactic derivations (and thus syntactic structure) and (overt) morphology. In a framework where it is assumed that inflectional morphemes, such as agreement and tense morphemes, are functional heads and these inflectional elements are then "picked up" by lexical heads that adjoin to (unite with) the relevant functional head positions, the implementation of the basic idea behind the Mirror Principle is relatively straightforward. It is not as clear, however, what role the Mirror Principle can play in the strictly lexicalist approach embodied in the Minimalist Program.[28]

These are some of the basic issues having to do with the universality/diversity of functional categories and their relation to morphology. Some of the papers in this volume discuss these issues in considerable detail and the following two subsections contain a brief introduction to these discussions.

2.5.2. *Functional Categories, Parametric Variation, and Overt Morphology*

Chomsky's original work on the Minimalist Program (1993) assumes the so-called "Split Infl Hypothesis" of Pollock (1989) which had been further developed in Chomsky's own work (1991) and by Belletti (1990) and others. However, instead of just one Agr-projection, as Pollock had originally suggested, Chomsky (1993), following Chomsky (1991: 434), postulates an AgrS-projection that enters into a checking relationship (Spec–head relationship) with the subject, and an AgrO projection that plays a corresponding role in object checking. Both Chomsky (1991) and Belletti (1990) argue that AgrSP dominates TP and this seems to be quite generally accepted in the pre-minimalist literature that adopts the Split-Infl Hypothesis. Work by Ouhalla (1988, 1991, 1994) and Campbell (1991) is somewhat of an exception in this respect, as they have argued that the hierarchical order of AgrSP and TP may vary and that this variation reflects

---

28. This simplified exposition ignores the fact that languages vary with respect to "overt verb movement". As discussed by Pollock (1989) and Chomsky (1991), for instance, French differs from English in this respect. Hence the idea (which goes back to Chomsky 1957) that a rule of Affix Hopping (or Affix Lowering) might be an appropriate way of accounting for aspects of the morphosyntax of finite verbs in English while a rule of Verb Raising (overt verb movement that is) would be appropriate for French. The strong lexicalist approach to morphosyntax adopted in the Minimalist Program (Chomsky 1993) makes it possible to minimize this difference. Under such an approach the difference between French and English is no longer claimed to be that the former has Verb Raising and the latter Affix Lowering but rather that the former has overt Verb Raising and the latter covert Verb Raising.

differences in morphological structure according to Baker's Mirror Principle. Thus if the subject agreement morpheme is "further away from" the stem of the verb than the tense morpheme is, then the syntactic AgrS projection (AgrSP) will dominate the T projection (TP).[29]

In the Minimalist Program, a strong lexicalist approach to inflectional morphology is postulated in the sense that words are supposed to emerge from the Lexicon fully inflected. Thus there is no affixation of inflectional morphemes via the application of syntactic head-adjunction. Instead, it is assumed that morphological features of lexical elements are checked against matching features of functional heads in the syntactic structure. As pointed out above, it is not immediately obvious how the Mirror Principle could be cast within such a model. In his Minimalist Program paper, Chomsky (1993: 28) makes a brief suggestion about how the spirit of the Mirror Principle could be maintained under the strong lexicalist approach he assumes. This idea is taken up in THRÁINSSON's paper in the present volume and developed further. He suggests that if we assume that morphological features are associated with particular inflectional morphemes (to the extent that such morphemes can be found in the language in question) and if the features of morphemes closer to the stem must be checked first, then we can capture the essence of the Mirror Principle in a lexicalist framework (see Halle & Marantz 1993: 166f. for discussion).

Now it should be noted that the relationship between morphological feature checking and overt morphology assumed by THRÁINSSON is much closer than assumed in Chomsky's Minimalist Program paper (1993). Thus while Chomsky keeps the basic idea advocated by Pollock and many others that overt verb movement is triggered by "strong morphology" (see the references in 2.5.1 above), he argues that the parametric variation should be expressed by assuming strong vs. weak features on functional heads and that these features are abstract in the sense that their value has no direct relationship to overt morphological distinctions. SOLÀ's paper in the present volume is very similar in spirit to THRÁINSSON's. Thus SOLÀ argues that the crucial parameter which determines

---

29. Ouhalla argues that while the agreement marker is further from the stem of the verb than is the tense marker in French and thus AgrSP dominates TP, the agreement marker is closer to the stem of the verb than is the tense marker in Arabic and hence the T projection dominates the AgrS projection (see e.g. Ouhalla 1994: 45. — For a different account see Shlonsky, in press.). Campbell (1991) argues that within a single language, German, TP dominates AgrSP in the present tense, whereas in the past tense, AgrSP dominates TP.

the occurrence of overt head movement in general (and overt verb movement in particular) is "the presence/absence of inflectional morphology for a given functional category on a given word". We will return to this aspect of SOLÀ's paper in section 2.5.3 below.

Attempts to relate functional structure in the syntax to overt morphological structure in such a direct fashion raise the question of the universality of functional categories. As it is well known that languages exhibit considerable variation in overt inflectional morphology, directly relating functional structure to inflectional morphology would seem to predict a great cross-linguistic variation in functional structure. Such a concept of functional structure seems clearly at odds with much recent work in theoretical syntax, which tends to assume (near) universality of functional structures, as THRÁINSSON notes in his present paper. But Pollock's and Chomsky's original ideas about the decomposition of the Infl-projection were actually criticized early on by Iatridou (1990). Iatridou warned against the idea that all languages should be analyzed as having Agr- and T-projections, for instance, although there might be (morphological and syntactic) evidence for them in some languages. THRÁINSSON's paper is close in spirit to Iatridou's, as one of his main claims is that only some languages have a split Infl in the Pollockian sense whereas others simply have a single unsplit IP-projection. THRÁINSSON goes on to discuss how morphological triggers may facilitate acquisition of the value of what he refers to as "The Split IP Parameter". He claims, for instance, that English and Mainland Scandinavian (MSc) have a "fused" IP whereas French and Icelandic have a split IP. This is supposed to explain some syntactic differences between these two types of languages, including the availability of the "extra subject position" in Icelandic as opposed to MSc. THRÁINSSON agrees with JONAS that this extra subject position is Spec-TP and argues that the reason it is not available in MSc is that MSc simply does not have a T-projection separate from the AgrS-projection and hence there is only one such specifier (subject) position present, namely Spec-IP. This is then related to the fact that MSc does not have separate tense and agreement morphemes whereas Icelandic does.[30]

---

30. As discussed in THRÁINSSON's paper, a somewhat similar approach is taken by Bobaljik (1995), although Bobaljik does not want to tie the presence of a separate Agr-projection as directly to overt morphology as THRÁINSSON attempts to do. A different position is taken by Chomsky (1995b: 354), who suggests that there may be no need for a separate Agr-projection at all in any language and the "extra subject position phenomenon" may instead result from the ability of T to project more than one specifier in some languag-

Now, if THRÁINSSON and SOLÀ are right in that there is a rather direct relationship between functional structure and syntactic movement on the one hand and overt morphology on the other, then the question arises whether different clause types within a given language vary with respect to the functional categories they instantiate. If the Agr-projection, for instance, is really an agreement projection, then we might expect it to be absent in non-finite clauses since these typically do not exhibit (overt) agreement. This issue is briefly addressed in THRÁINSSON's paper and also in the paper by JONAS, but it is also related to the question of what triggers overt movement, namely whether there are abstract strong features or whether overt inflectional morphology is solely responsible. This issue will be examined further in the next subsection.

2.5.3. *Morphological Strength and Overt Order*

As mentioned above, Chomsky assumes in his Minimalist Program paper (1993) that (abstract) strong features on functional heads trigger overt syntactic movement. The original idea was that strong features had to be checked before Spell-out since unchecked strong features would be visible but uninterpretable at PF and thus cause the derivation to crash.[31] Weak features, on the other hand, need not be checked in overt syntax, and hence they will not be, since covert movement is more economical than overt movement (cf. the discussion of economy principles in sections 1.4 and 2.3 above).

As already mentioned, SOLÀ rejects the notion of (abstract) strong features and suggests instead that the key to the relevant syntactic variation can be found in overt morphological distinctions. He argues that given the proper morphological analysis, we can account for cross-linguistic differences in functional structure and overt syntactic movement. Most of his arguments are based on comparisons of English and (other) West Germanic languages. His analysis adopts Kayne's (1994) proposals, in particular their adaptation to West Germanic languages along the lines suggested by Zwart (1993b), for instance. This approach delimits cross-linguistic phrase structure variation (see the discussion

---

es but not in others. The basic idea of the relevance of "strong inflection" could be implemented in such an approach by saying that only "strong" T can have this "multiple specifier" property.

31. The account of the relationship between strong features and overt movement is somewhat different in Chomsky's most recent work (1995b), but a detailed description of that alternative is not directly relevant here (see also section 2.3.2).

of directionality in 1.7 above and 2.6 below). The basic idea is that an inflected lexical element moves only if it "contains overt morphology of the target of [the] movement;[32] and there is no covert movement". This is obviously a quite radical departure from the basic idea of Chomsky's Minimalist Program, which assumes both overt and covert movement.

The paper by GROAT & O'NEIL proposes ideas somewhat similar to SOLÀ's. More specifically, GROAT & O'NEIL suggest that there are "no post-Spell-out syntactic operations". In this respect they agree with SOLÀ. Their particular implementation of the idea is different, though. In particular, they actually do assume both overt and covert movement. The difference is not, however, that covert movement applies "after Spell-out" or "in LF", as standardly assumed in the Minimalist Program, but rather that all movement precedes Spell-out and "overt movement is movement which carries phonological features to the head of the chain being created, while covert movement leaves phonological features behind".[33]

This idea of a single-interface model proposed by GROAT & O'NEIL will be discussed further in section 2.7 below. It should be noted here, however, that GROAT & O'NEIL's analysis differs from SOLÀ's in that the former assume that it is feature strength of functional heads that triggers overt movement, rather than overt inflectional morphology of lexical items, as SOLÀ argues.

Finally, an interesting claim made by GROAT & O'NEIL is that the chains

---

32. An interesting question here is whether this implies, say, that verb forms not overtly inflected for tense, like infinitives in Germanic languages, would not be expected to move to T. As has been argued, there appears to be verb movement in control infinitives in Icelandic, for instance (see e.g. the papers by JONAS and THRÁINSSON in this volume and references cited there). This could be taken to suggest that although overt morphology may provide clues to the language acquirer about the nature of functional projections in the target language, their "strength" is determined once and for all in each language and does not vary depending on the inflectional properties of the lexical heads in different clause types (finite and non-finite). This is obviously an interesting and broad topic requiring further research.

33. Interestingly, the idea of moving the relevant features and "leaving phonological material behind" is also proposed in Kitahara (1994) and further developed in Chomsky's most recent work (Chomsky 1995b: 261ff.). The main difference is that Chomsky assumes that this feature movement takes place "after Spell-out", i.e. at LF, and movement which has overt consequences (pied-pipes phonological material) takes place "before Spell-out". GROAT & O'NEIL, on the other hand, assume that Spell-out is always "applied to the same phrase-marker that receives LF-interpretation."

created by overt vs. covert movement have different licensing properties. This is an issue also discussed in the paper by JONAS in the present volume. Recall that JONAS argues that overt verb movement to T licenses Spec-TP as a subject position. Jonas & Bobaljik (1993) had previously argued that having Spec-TP as an available subject position was a prerequisite for object shift in languages like Icelandic, for instance. GROAT & O'NEIL argue, on the other hand, that overt verb raising to T (or T/AgrS in their terminology) makes Spec-AgrOP an available theta position and thus makes it possible to base-generate the object in Spec-AgrOP, creating the illusion of object shift.

2.6. *Directionality and Word Order*

In traditional grammar, in typological research, and in the pre-minimalist stages of generative grammar, a distinction is made between head-final and head-initial languages. It is not entirely correct to equate head final with OV and head initial with VO. A language can be characterized by OV word order without necessarily having head-final structure, if the OV order is the result of object raising. Thus, the VO/OV distinction refers to surface structure, not necessarily to the order of elements in the initial representation (the product of Merge). We can therefore continue to speak of VO, OV, SVO, VSO, etc. languages, without committing ourselves to claims about the basic structure of these languages.

In the minimalist framework, the VO, OV, etc. patterns can be described in terms of the interaction of overt and covert movement. Thus, a language can be characterized as VO if its basic structure is head initial and no overt movements take place, or if the object and the verb move overtly, with the verb ending up to the left of the object, etc. Notice that if we assume that all languages are head initial (or head final, for that matter), the interaction of covert and overt movement is the only factor bringing about the surface OV, VO, etc. distinctions. This is one of the conceptual arguments in Kayne (1994) for assuming that all languages are head initial (see section 1.7).

In the Principles and Parameters framework, it was assumed that the order of head and complement at D-structure was an instance of parametric variation. A language could have one or the other setting of the *headedness parameter*. Kayne (1984) and Koster (1987), among others, derived the headedness parameter from directionality of government: in some languages, the verb governs to the right, in others, to the left. If theta-roles are assigned under government, rightward governing verbs take their complement to the right (yielding head

initial structures), and leftward governing verbs take their complement to the left (yielding head final structures).

This derivation of the headedness parameter is no longer possible in the minimalist framework, in which the government relation is no longer available. To review briefly, government is not needed for theta-role assignment which is now a function of the operation Merge (i.e. of the sisterhood relation), nor is it needed for Case assignment, which is now redefined as feature checking in a Spec–head configuration, nor for the formulation of the PRO-theorem, since PRO is now defined as an element with Null Case, not as an ungoverned element (Chomsky & Lasnik 1993), etc.

Viewed from this perspective, Kayne's conjecture that all languages are head-initial is highly consistent with the minimalist framework of Chomsky (1993).[34]

In this volume, Kayne's generalizations regarding phrase structure are adopted and employed in an analysis of Dutch past participle constructions by VANDEN WYNGAERD (see the discussion in 2.1). VANDEN WYNGAERD adopts the analysis of Dutch, an OV language traditionally regarded as head-final (cf. Koster 1975), as involving head-initial structures only (cf. Zwart 1994a; see Zwart, to appear, for an introductory text on the analysis of Dutch as a head-initial OV language).

The other article concentrating on Dutch, by DEN DIKKEN, also adopts Kayne's assumptions, taking Dutch to be a head-initial language. DEN DIKKEN pursues earlier suggestions by Kaan (1992) and Zwart (1993b) and analyzes West Flemish Verb Projection Raising constructions as involving head-initial structures with multiple VPs, and AgrOPs either dominating the VPs or inserted between two VPs. In Verb Projection Raising constructions, like (52), a sentence final cluster of verbs is broken up by material appearing in between the verbs:

(52) *...da Jan wilt geen vlees eten* (West Flemish)
that John wants no meat eat-INF
'...that John does not want to eat meat.'

In the traditional analysis of Dutch as a head-final language, (52) is derived by moving the verb projection *geen vlees eten* 'eat no meat' to the right, across the higher verb *wilt* 'wants'. DEN DIKKEN points out various inadequacies of this traditional approach, and strengthens the Kaynean analysis by deriving certain

---

34. In Chomsky (1995b: 335f), the LCA is adopted as a linearization device operating between Spell-out and PF. We will not discuss this development here.

scope ambiguities (or the absence thereof) in these Verb Projection Raising constructions. In doing so, DEN DIKKEN employs KITAHARA's analysis of scope ambiguities, also presented in this volume.

The analysis of Dutch and German as head-initial languages is not uncontroversial (see e.g. Abraham 1994 and Sternefeld 1994). In this volume, the issue is relevant to the status of Holmberg's Generalization and to the hypothesis advanced in the article by JONAS regarding the distribution of transitive expletive constructions. Holmberg's Generalization correlates overt object raising and overt verb raising. Assuming that Dutch and German have object raising of the relevant type (i.e. movement to Spec-AgrOP), the hypothesis that Dutch and German are head-initial is not consistent with Holmberg's Generalization: the nonadjacency of the object and the verb indicates that the verb has not raised to AgrO (see section 2.3.1 above). JONAS presents the hypothesis that languages that have transitive expletive constructions also have overt verb raising to T. But if Dutch, a language with transitive expletive constructions (Bennis 1986; Zwart 1992a), has a head-initial TP, the verb should not be allowed to stay to the far right in embedded transitive expletive constructions, contrary to fact:

(53) *...dat er veel mensen een praatje hielden* (Dutch)
that there many people a talk held
'...that many people gave a talk.'

Thus it seems that something has to give here.

2.7. *The Organization of the Grammar*

As discussed above, the Minimalist model of grammar eliminates both D–structure (a level which is presumed to interact with no non-linguistic system) and S–structure, a level which similarly interacts with no non-linguistic system while concomitantly interacting with D–structure, PF and LF (in the standard Government and Binding model of Chomsky 1981). As Chomsky & Lasnik (1993: 511) note, we might think of S–structure, informally, as a "... (presumably unique) solution ...", i.e., the only possible representation successfully mediating between D–structure, PF and LF via transformational application. Thus, in Chomsky (1993) S–structure as a level of representation is eliminated — in its place, there is an optional rule of Spell-out: at any point in a derivation (i.e., after any given transformational rule application) the representation derived may be spelled-out. Doing so sends that representation into two separate, non-interacting components of the grammar: the PF component and the LF (or covert) component. The leading idea is that properties of S–structure representa-

tion may well be deducible. For example, if, in an English derivation, the following representation were spelled-out

(54) [*I wonder* [$_C$0 [*John likes what*]]]
+wh
+strong

the strong ⟨+wh⟩ feature of C^0 would, as a result, appear in PF. However, strong features are illegitimate PF objects, and consequently the derivation crashes at this level. Thus, a PF (not an S–structure) condition is violated — an empirical hypothesis — carrying with it the intuitively appealing and seemingly correct prediction that (54) is not a correct representation of sound/pronunciation.

While S–structure as a level of representation is targeted for deduction (an appealing goal), notice that in certain central respects the Government and Binding model nonetheless persists — even in the wake of D–structure and S–structure elimination. That is, Spell-out splits the derivation into two tracks, PF and LF.

Just as van Riemsdijk & Williams (1981) investigated alternatives to the Government and Binding model of Chomsky (1981) (going back to Chomsky & Lasnik 1977), GROAT & O'NEIL (this volume) explore an alternative to the Minimalist model of grammar. As GROAT & O'NEIL note, the Minimalist model incorporates three quite fundamental, but stipulated, asymmetries concerning the application of pre- vs. post-Spell-out operations:

(55) a. Pre-Spell-out operations are more costly than are post-Spell-out operations (Procrastinate) (Chomsky 1993: 30).
b. The strict cycle condition (the extension condition) applies only to (certain) pre-Spell-out operations, but does not constrain post-Spell-out operations. In particular, LF object-shift in languages like English is apparently countercyclic (non-extending) and hence the extension condition is presumed inapplicable in the covert component (Chomsky 1993: 24).
c. There is lexical access before Spell-out, but not after Spell-out (Chomsky 1993: 22).

As GROAT & O'NEIL note, the formal operations (binary and singulary Transformations, Merge and Move) which apply before Spell-out are, by hypothesis, formally identical to those applying in the covert component. Thus, given the three asymmetries just noted, the model of the Minimalist Program in fact postulates two distinct computational systems that are, except for these asymmetries, identical. A worthy goal and the one sought by GROAT & O'NEIL is the

elimination of these asymmetries, yielding a single computational system.

In order to eliminate pre- vs. post-Spell-out asymmetries, GROAT & O´NEIL hypothesize a model in which it is logically impossible to even state such asymmetries. In their model, a derivation (iterative transformational-rule application) yields a single phrase-marker which feeds both the phonological and the interpretive components — thus their title "Spell-out at the LF interface". Given such a model of grammar, there are simply no post-Spell-out operations and therefore it is impossible to stipulate pre- vs. post-Spell-out asymmetries. But if one and the same phrase-marker is both spelled-out and undergoes interpretation, how can overt operations (triggered by strong features) vs. covert operations (triggered by weak features) be distinguished?

For GROAT & O´NEIL, overt vs. covert movement does not concern the derivational timing of movement (pre- vs. post-Spell-out). Rather, strong feature checking is phonological feature checking, and it therefore requires that movement carry-along the phonological features of the moved category. By contrast, weak feature checking is not phonological feature checking, and it therefore allows phonological features to be left on the trace, i.e., not carried along by movement. Procrastinate (in Chomsky 1993, a global transderivational economy condition) is recast as a local, intraderivational economy condition: "It is more work to carry along phonological features then it is not to: Weak feature checking therefore precludes moving phonological features." (See also Kitahara 1994, 1995, 1996, and Chomsky 1995b: 262.) Notice now that English object shift need not be formulated as a countercyclic LF operation. Instead English object shift is cyclic but since it is a weak (N-)feature checking operation, the now local intraderivational economy condition Procrastinate prohibits moving the phonological features of the shifted English object. Given this analysis, the extension condition (strict cycle condition) need not be stipulated to be inapplicable in LF — i.e., this unexplained pre- vs. post-Spell-out asymmetry is eliminated.

Further support for their model is provided by GROAT & O´NEIL's reanalysis of subjacency effects in Japanese *wh*-movement (Watanabe 1992) and Icelandic object shift.

The Minimal Links of Verb (Projection) Raising

Marcel den Dikken
Vrije Universiteit Amsterdam/HIL

Let us sharpen the notion of "shortest movement" as follows: (..) [i]f α, β are in the same minimal domain, they are equidistant from γ. In particular, two targets of movement are equidistant if they are in the same minimal domain.

— Chomsky 1993: 17–18

Abstract

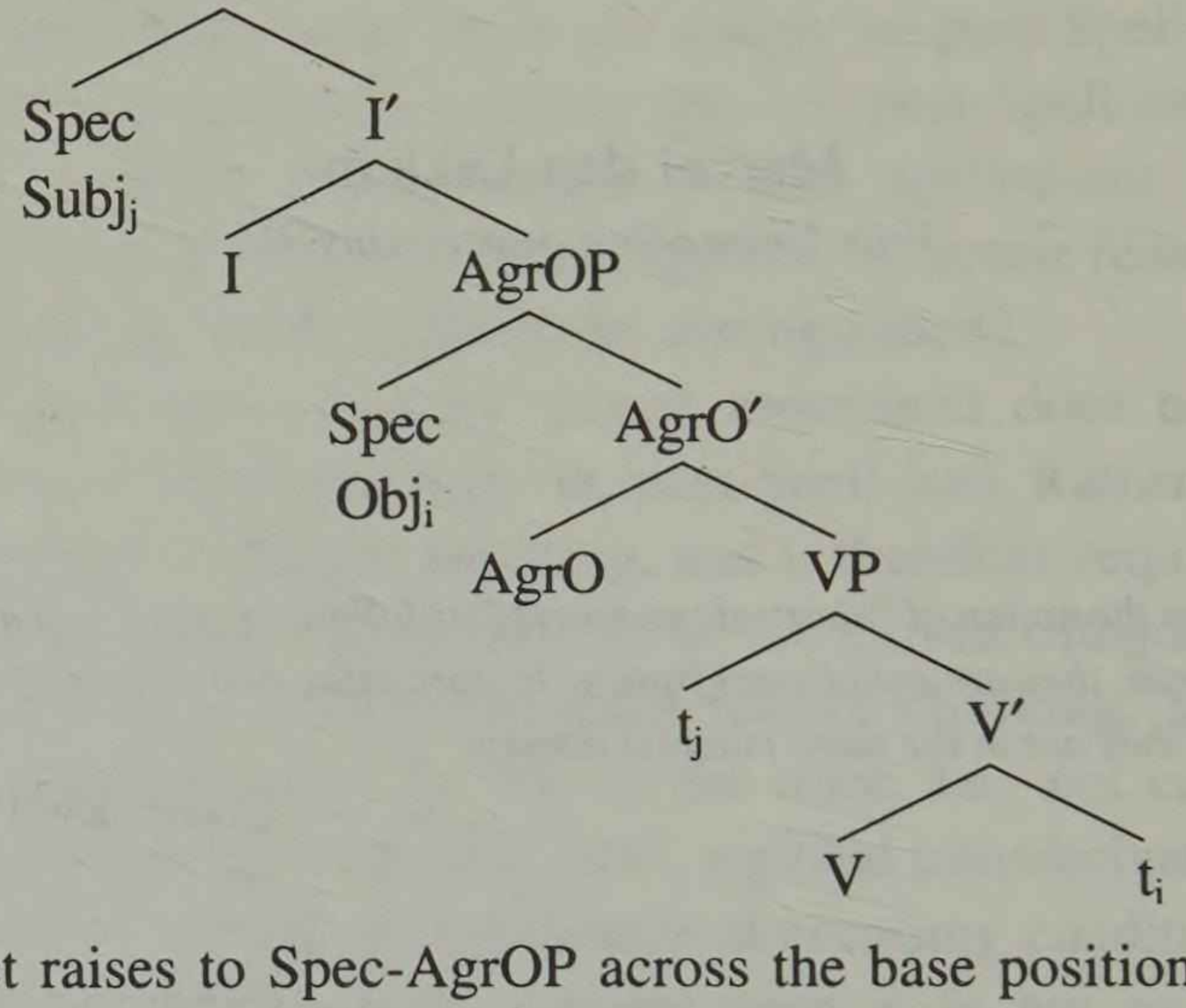
This article exploits the theory of phrase structure and locality of movement outlined in the original formulation of the minimalist program for linguistic theory (Chomsky 1993), with its AgrPs and its head-chain based theory of EQUIDISTANCE, to construct an analysis of Verb Raising and Verb Projection Raising phenomena in the Continental West Germanic languages from a minimalist and antisymmetric (Kayne 1994) perspective. I argue that the essential difference between Verb Raising and Verb Projection Raising constructions is the presence in the latter of a TP in the complement of the matrix verb, a TP whose presence is forced by minimalist locality considerations and whose presence is empirically supported by several properties of VPR constructions, including temporal adverbial modification and cluster-contained subjects in transitive expletive constructions with VPR. The analysis of V(P)R also leads to the conclusion that the Continental West Germanic languages feature a SCRAMBLING operation distinct from Case-driven NP-movement, a conclusion to which the scope facts of double object Verb (Projection) Raising constructions lend empirical support.

1. Introduction

The minimalist program of linguistic theory incorporates an economy constraint that prescribes "keeping links minimal ('Shortest Movement' Condition)" (Chomsky 1993: 15). In the strictest of cases this minimal link constraint would

prohibit any movement across a potential landing-site. But such a strict interpretation of the condition would be untenable, for it would prevent the grammar from deriving any simple transitive sentence, whose minimalist structure reads as in (1):[1]

(1) *Transitive sentences according to Chomsky (1991, 1993)*



In (1) the object raises to Spec-AgrOP across the base position of the subject, and the subject moves to Spec-IP crossing the landing-site of the object, Spec-AgrOP. Neither movement is compatible with a strict interpretation of the minimal link constraint.[2] It appears, then, that its demands should be relaxed,

---

1. For ease of initial exposition, I have conflated Chomsky's (1991, 1993) AgrSP and TP into a single projection IP. The distinction between AgrSP and TP will not play a role in this introductory section; later in the article, TP will come to play a central role. The point in the derivation at which NP-movement of the subject and object takes place will not be immediately relevant either; I will address this issue on the next two pages.

2. For the case of object movement to Spec-AgrOP across the subject's base position, one might wonder, though, whether the minimal link condition has anything to say about it, given that Spec-VP is not a position in which the object could ever check a feature (Spec-VP not being in the checking domain of a functional head, and feature checking taking place only in the checking domains of functional heads), hence is not a potential landing site (given Greed, the economy condition that says that each and every movement step must be triggered by the need to check a feature of the moved constituent). Chomsky assumes that it has. But the issue is not crucial to the present article. In the analysis of Verb (Projection) Raising constructions to be developed in what follows, the key cases of "shortest movement" involve raising of the subject across Spec-AgrOP.

and that the theory should allow specifier-skipping movement of the type depicted in (1) under certain conditions. In particular, such movement is legitimate if the position skipped and the position landed in are equally far removed, or EQUIDISTANT, from the position that the moved constituent starts out from. And two positions are equidistant if they are members of the same minimal domain (Chomsky 1993: 17–18). In each grammatical case of apparently non-local movement, we should hence identify some minimal domain that contains the intervener and the landing site as two of its members.

In the minimalist domain theory of Chomsky (1993), such a minimal domain can only be formed by the application of head movement — particularly, movement of the head whose specifier is skipped to the immediately higher head position whose specifier serves as the landing site of movement. To take the case of movement of the object to Spec-AgrOP across Spec-VP in (1), we should raise V to AgrO, creating the head-movement chain (V, *t*), whose minimal domain comprises the nodes Spec-AgrOP, Spec-VP and the object's base position.[3] And in the case of subject movement crossing the object's landing site, there should be domain-extending head movement of AgrO to I, rendering Spec-IP and Spec-AgrOP equidistant from the extraction position of the subject. The minimalist locality theory thus predicts a strong correlation between Spec-skipping NP-movement and domain-extending head movement — the former will always be contingent on facilitating movement of the head

---

3. Let us quickly run through the way in which minimal domains are computed. The minimal domain of a chain (X_1, X_n) — where $n=0$ (in the case of a trivial, one-member chain) or 2 (head-chains in practice having maximally two members on Chomsky's assumptions; see Chomsky 1993: fnn. 17, 20) — is a subset of the nodes contained in the domain of that chain. The two notions are defined in (i). In the particular case of the two-member chain (V, *t*) encountered in the main text example, the domain of the chain is the set of nodes constituted by Spec-AgrOP, Spec-VP, object-*t* *and all nodes dominated by them*; the *minimal* domain of this chain reduces this rather heterogeneous bunch of nodes to just Spec-AgrOP, Spec-VP and object-*t*, shedding the italicized addition in the preceding clause.

(i) a. The *domain* of a chain (X_1, X_n) is the set of nodes contained in the maximal projection of the head of the chain (X_1) that are distinct from and do not contain any X.

(i) b. The *minimal domain* of the chain (X_1, X_n) is the largest subset of the domain of the chain such that none of its members is dominated by a member of the domain of the chain.

whose specifier is crossed to the higher head whose specifier is targeted by NP-movement.

From Chomsky's (1993) derivationalist perspective, the correlation between Spec-skipping and domain-extending head movement is even tied to *the point in the derivation* at which movement takes place, such that if Spec-skipping NP-movement obtains in the overt syntax (i.e. prior to SPELL-OUT), facilitating head movement must also take place overtly. Chomsky suggests that in the domain of object movement to Spec-AgrOP the facts of OBJECT SHIFT in the Scandinavian languages bear out this level-wise correlation, which has come to be known by the name of Holmberg's Generalization, after its discoverer (see Holmberg 1986: 165 ff.).[4] However, as especially Zwart (1994b) has argued at some length, the validity of the strict correlation between overt Object Shift and overt movement of the Case-assigning verb expressed by Holmberg's Generalization is dubious, even in Icelandic, which seems just about the only language for which it might truly hold. While I hence will not be assuming a direct correspondence between *the point in the derivation* at which Object Shift takes place and *the point in the derivation* at which domain-extending head movement obtains, I continue to assume a contingency of specifier-skipping NP-movement and facilitating head movement, taking a more representationalist view — the landing site of movement and the position skipped must ultimately be equidistant in the LF output representation. I will exploit the head-movement based theory of equidistance in an analysis of Verb Raising and Verb Projection Raising phenomena in the Continental West Germanic languages (especially West Flemish), approaching their syntax from a minimalist and antisymmetric (Kayne 1994) perspective and showing that an analysis of Verb (Projection) Raising (henceforth, V(P)R) built on the premise that phrase structures are uniformly of the type Specifier–Head–Complement is not only theoretically feasible but also empirically highly accurate in that it makes some precise and correct predictions about what is and is not possible in V(P)R constructions.

The paper is structured as follows. Section 2 introduces the phenomenon of Verb (Projection) Raising and briefly browses through the substantial body of literature devoted to it. In Section 3 I then present an analysis of V(P)R built on the minimalist locality theory of Chomsky (1993), showing that this theory ensures on purely structural grounds that VPR constructions involve TP comple-

---

4. Curiously, Chomsky (1993: 18) attributes it to Vikner (1990), and suggests that its scope extends across "the Germanic languages". Both the attribution and the suggestion are inaccurate; see Zwart (1993b, 1994b) for discussion.

mentation. The empirical correctness of this theoretical prediction is demonstrated in Section 4, which *en passant* shows that, apart from Case-driven NP-movement of the object to Spec-AgrOP, the Continental West Germanic languages also feature an independent scrambling operation. This conclusion is confirmed in Section 5 by a discussion of the scopal properties of double object V(P)R constructions. After a discussion of VPR in transitive expletive constructions in Section 6, Section 7 closes the paper.

2. Verb (Projection) Raising

The West Flemish example pair in (2) and (3) illustrates the basic patterns of simple VR and VPR constructions, respectively:

(2) *da Jan dienen boek wilt kuopen* (Verb Raising; VR)
that Jan that book wants buy

(3) *da Jan wilt dienen boek kuopen*
that Jan wants that book buy (Verb Projection Raising; VPR)
both: 'that Jan wants to buy that book'

VR constructions are characterized by a clustering of verbs at the end of the clause, their VPR variants differing in that some non-verbal material (*dienen boek* 'that book' in (3)) breaks up the clause-final sequence of verbs.

The OV–based perspective on the syntax of the Continental West Germanic languages has brought forth a wealth of accounts of V(P)R constructions such as (2) and (3), all starting out from the underlying order *da Jan dienen boek kuopen wilt*. Setting aside Haegeman & Van Riemsdijk's (1986) reanalysis approach and Broekhuis' (1993) analysis in terms of PF inversion, we can say that the VR pattern in (2) used to be quite generally analyzed in terms of right-adjunction of the main verb *kuopen* to the higher auxiliary verb *wilt*, and that VPR was taken to involve rightward movement of the raised VP, either to an adjunction position to the higher verbal head or to an adjunction position to some maximal projection dominating the higher verb (see Evers 1975; Rutten 1991; Vanden Wyngaerd 1989b; Den Dikken 1989; Haegeman 1992; among many others). In one way or another, all of these OV–based analyses treat V(P)R in terms of operations that are not obviously compatible with the restrictive theory of syntax that has been developed over the last decade or so. The problems are particularly severe in the case of VPR. An analysis of this phenomenon involving overt adjunction of the embedded verb's projection to the

higher verb violates the restrictions on adjunction imposed in both Chomsky (1986b: 6) and Chomsky (1995a: §5), and a VPR analysis featuring rightward adjunction of the lower verb's projection to some higher maximal projection is questionable in the light of the minimalist program, since it is difficult to conceive of rightward VP–movement as being *triggered* by feature-checking considerations.

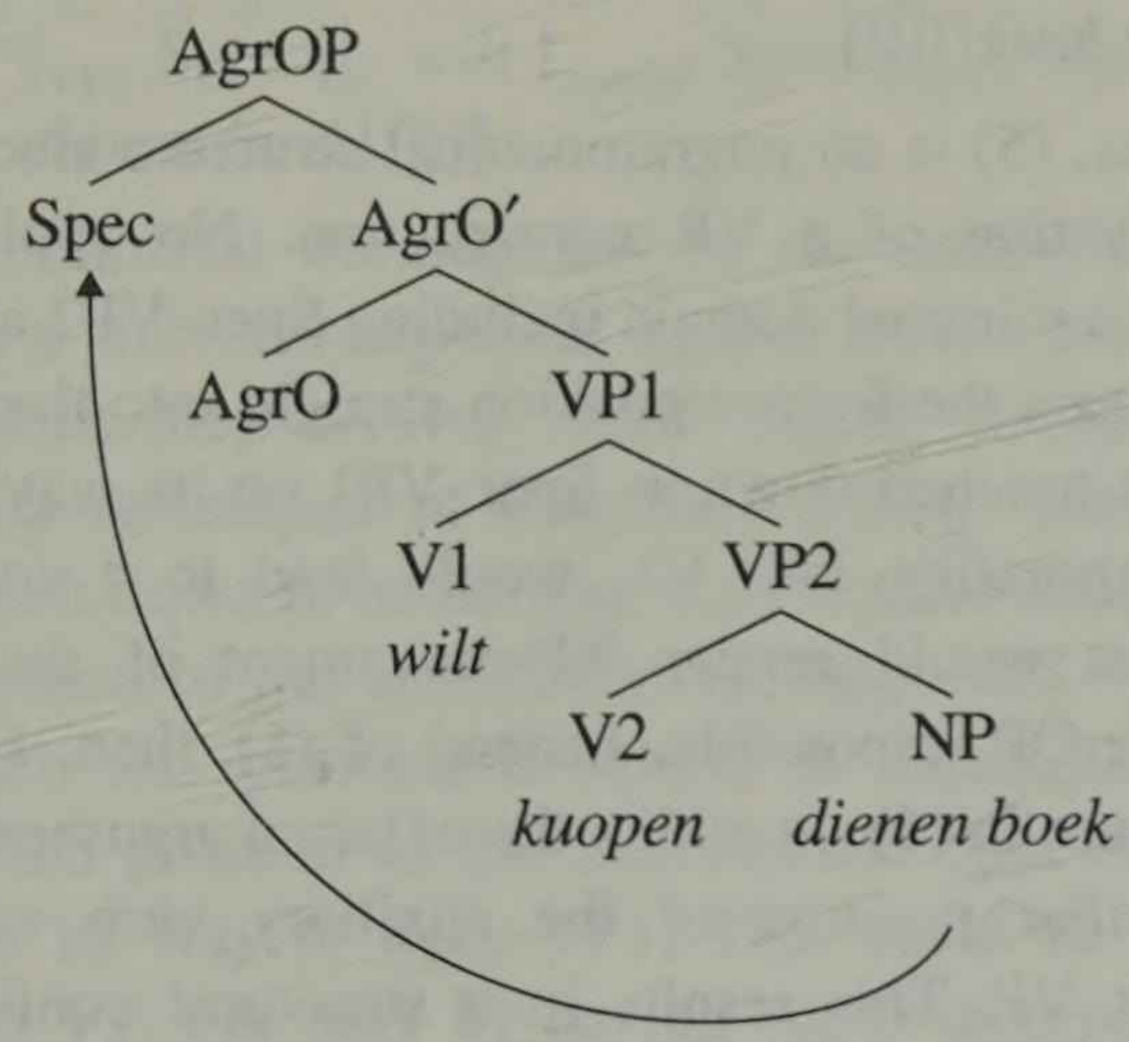
Moreover, all OV–based accounts make (implicit or explicit) reference to a directionality parameter of sorts which makes West Flemish differ from a language like English with regard to the direction in which V governs its complement — V precedes its complement in the underlying structure of English, but follows it in the base structure of West Flemish. Such a directionality parameter is unstatable in a minimalist theory. We could not phrase it in terms of Case assignment, since Case is no longer taken to be assigned under government but instead is checked in the checking domain of designated Case-checking heads. And we cannot state it in terms of θ-role assignment either, since the minimalist theory takes the restrictive tack of confining all parametric variation to featural properties of non-substantive (functional) heads. Recent theorizing hence leaves us with two options concerning the relative order of complements and heads — it could either be entirely free, or it is rigidly fixed by some overall theoretical constraint. Kayne (1994) has provided us with reasons for believing that the latter option is the right one, developing a theory of phrase structure according to which it is uniformly of the basic form Specifier–Head–Complement. Such a theory rules out in principle any underlying OV order. And moreover, it also eradicates all rightward movement operations, thereby discrediting all the basic ingredients of the traditional analysis of V(P)R phenomena.

Viewed from an antisymmetric VO perspective (see Zwart 1993b; 1994a for detailed discussion), the surface order of the verbs in the VR and VPR constructions in (2) and (3) reflects their initial order. With respect to verb sequencing, then, the antisymmetric approach certainly seems simpler, *prima facie* at least. The placement of the object, hence the difference between (2) and (3), is now going to give us some trouble. How does it come to pass that the object in the VR construction ends up in a position outside the verbal cluster, while in the VPR example in (3) the two verbs embrace the object? In the next section I will answer this question by devising an analysis of VR and VPR constructions anchored in the original version of the minimalist theory, outlined in Chomsky (1993).

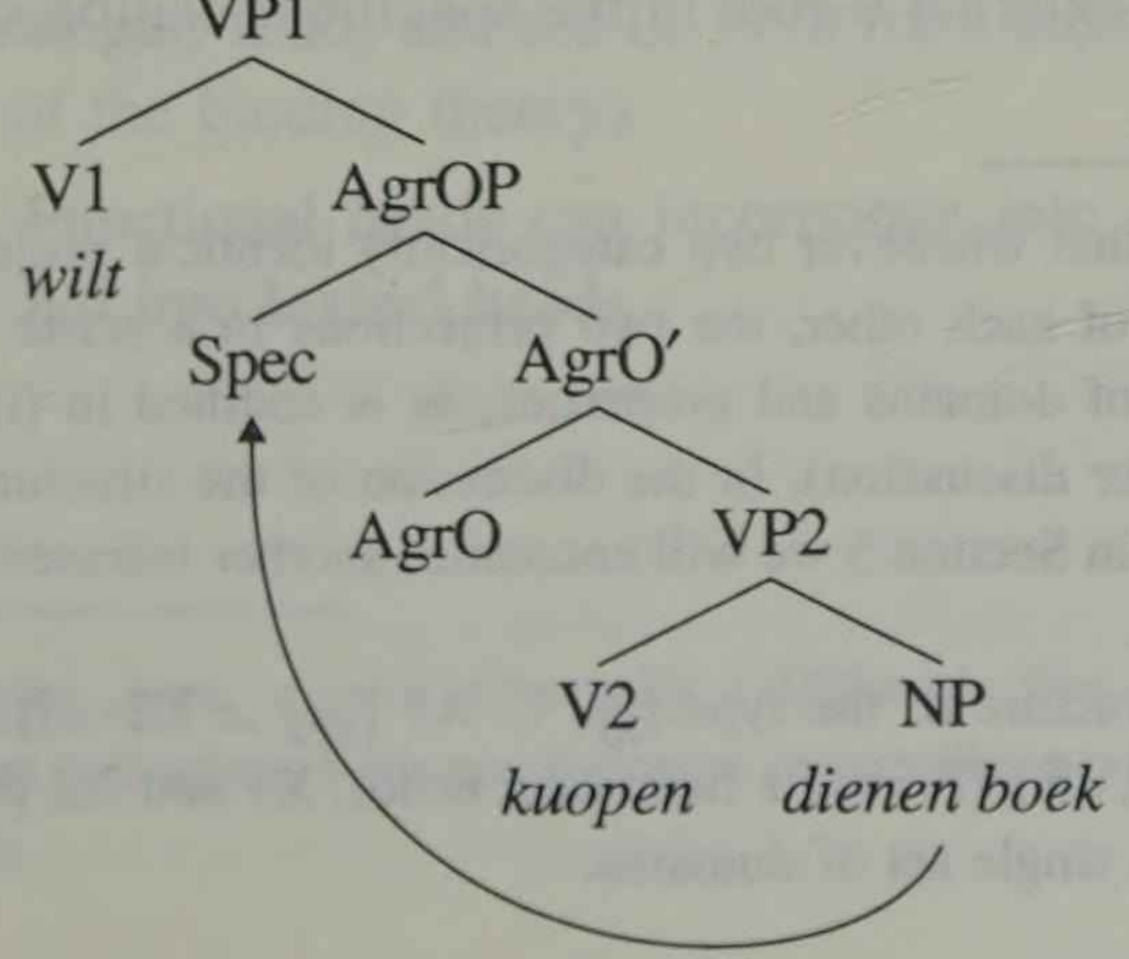
3. A Minimalist Analysis of Verb (Projection) Raising

The earliest attempts at coming to grips with VR and VPR in the minimalist and antisymmetric era, represented by the work of Kaan (1992) and Zwart (1993b), postulated a simple and minimal difference between transitive VR and VPR constructions like (2) and (3). While in the latter the AgrOP responsible for the checking of the object's features is generated in the complement of the auxiliary verb, in VR constructions AgrOP finds itself above the auxiliary verb. This is illustrated in the trees in (4a, b), which faithfully reflect Kaan's and Zwart's proposals. AgrO's strong Case feature forces overt syntactic object movement to its specifier, which makes (4a) yield the VR pattern while (4b) produces VPR.

(4) a. *VR à la Kaan (1992), Zwart (1993b)*



b. *VPR à la Kaan (1992), Zwart (1993b)*



The structures in (4) are simplifications of the full structure of VR and VPR constructions in at least one immediately obvious respect — they are mute on the base position of the subject and the movement operation(s) that it has to undergo to reach its feature-checking position. Where could the subject be base-generated in (4a, b)? In answering this question, let us first of all make the general assumption that the subject is base-generated within VP, as is standard (see Kuroda 1988; Fukui & Speas 1986; Koopman & Sportiche 1991 and much recent work, for discussion of the VP-internal subject hypothesis). With this said, now consider the VR structure in (4a). As a natural initial hypothesis we might assume that the subject, which after all is the external argument of V2, originates in the specifier position of VP2. Suppose, then, we adopted the structure in (5):

(5) *[$_{\text{AgrOP}}$ Spec [$_{\text{AgrO'}}$ AgrO [$_{\text{VP1}}$ Spec [$_{\text{V'}}$ *wilt* [$_{\text{VP2}}$ *Jan* [$_{\text{V'}}$ *kuopen dienen boek*]]]]]]

As the asterisk indicates, (5) is an ungrammatical structure since it cannot lead to a well-formed derivation of a VR construction. No application of head movement could create a minimal domain including Spec-VP2 and Spec-AgrOP, so object movement across the former position straight into the latter is illegitimate. And if the object touched down in Spec-VP1 on its way to Spec-AgrOP (which, with V2 incorporating into V1, would lead to a converging object raising derivation), that would render NP-movement of the subject across Spec-VP1 and Spec-AgrOP impossible. Instead of (5), then, I will assign VR constructions the structure in (6), in which the external argument of V2 is base-generated in the specifier position of the auxiliary verb whose projection immediately dominates VP. This results in a structural configuration that is typical of "clause union" — whenever two VPs are stacked immediately on top of each other, as in (6), and the higher verb has no external θ-role to assign, the lower verb deploys its external θ-role in the specifier position of the outer VP.[5]

---

5. The general idea is that whenever two categorically identical projections are stacked immediately on top of each other, the two projections in a sense "merge" into one, sharing a single set of domains and properties, as is codified in (i) (see Den Dikken 1995a: Section 5.2 for discussion). In the discussion of the structure of double object V(P)R constructions in Section 5 we will encounter another instance of domain sharing in stacked structures.

(i) In a structure of the type [$_{\text{XP1}}$ ... X1 [$_{\text{XP2}}$ ... X2 ...]], where X1 = X2 ∈ {A,N,P,V,F} (F = some functional head), X1 and X2 project together and share a single set of domains.

(6) $[_{AgrOP}$ Spec $[_{AgrO'}$ AgrO $[_{VP1}$ *Jan* $[_{V'}$ *wilt* $[_{VP2}$ *kuopen dienen boek*]]]]]

With V1 (which has access to V2's objective-Case feature in a stacked structure such as (6)) raising to AgrO, movement of the object to Spec-AgrOP is unproblematic; and subject movement out of Spec-VP1 across Spec-AgrOP is possible in the usual way, contingent on regular AgrO raising to I.

For VR constructions, then, the Kaan/Zwart approach in (4a) is essentially correct. For VPR, on the other hand, it is not. Since in (4b) the two verbal projections are not stacked immediately on top of each other, the external argument of V2 cannot be projected onto Spec-VP1 this time. Instead, it is assigned to the subject in Spec-VP2. Now let us see what happens if we just stick in the base position of the subject and leave everything else in (4b) unchanged, as depicted in (7):

(7) *$[_{VP1}$ Spec $[_{V'}$ *wilt* $[_{AgrOP}$ Spec $[_{AgrO'}$ AgrO $[_{VP2}$ *Jan* $[_{V'}$ *kuopen dienen boek*]]]]]]

While this is a minimal development of the simplified structure in (4b), it presents us with a technical argument against an analysis of VPR constructions according to which they involve "bare" AgrOP complementation, given the theory of locality developed in Chomsky (1993). The first position that the subject could touch down in on its way to its feature-checking position would be the specifier position of the auxiliary's VP, Spec-VP1.[6] Now, in order that the subject can reach this position it should be equally far removed from the subject's extraction position as the intermediate Spec-AgrOP position that the subject skips on its way to Spec-VP1. This could be possible only if AgrO incorporates into V1. For only by having AgrO incorporate into V1 could we create a minimal domain that contains both Spec-VP1 and Spec-AgrOP. But arguably, Agr, like functional heads in general, cannot incorporate into a lexical head (cf. also Branigan 1995; and see Li 1990 for a possible theoretical account of this in terms of the binding theory):

(8) Functional heads can incorporate into other functional heads, *not* into lexical heads.

---

6. I am assuming here, *contra* Chomsky (1995a, b), that densely successive-cyclic NP-movement operations with touch-downs in specifier positions of lexical projections are legitimate.

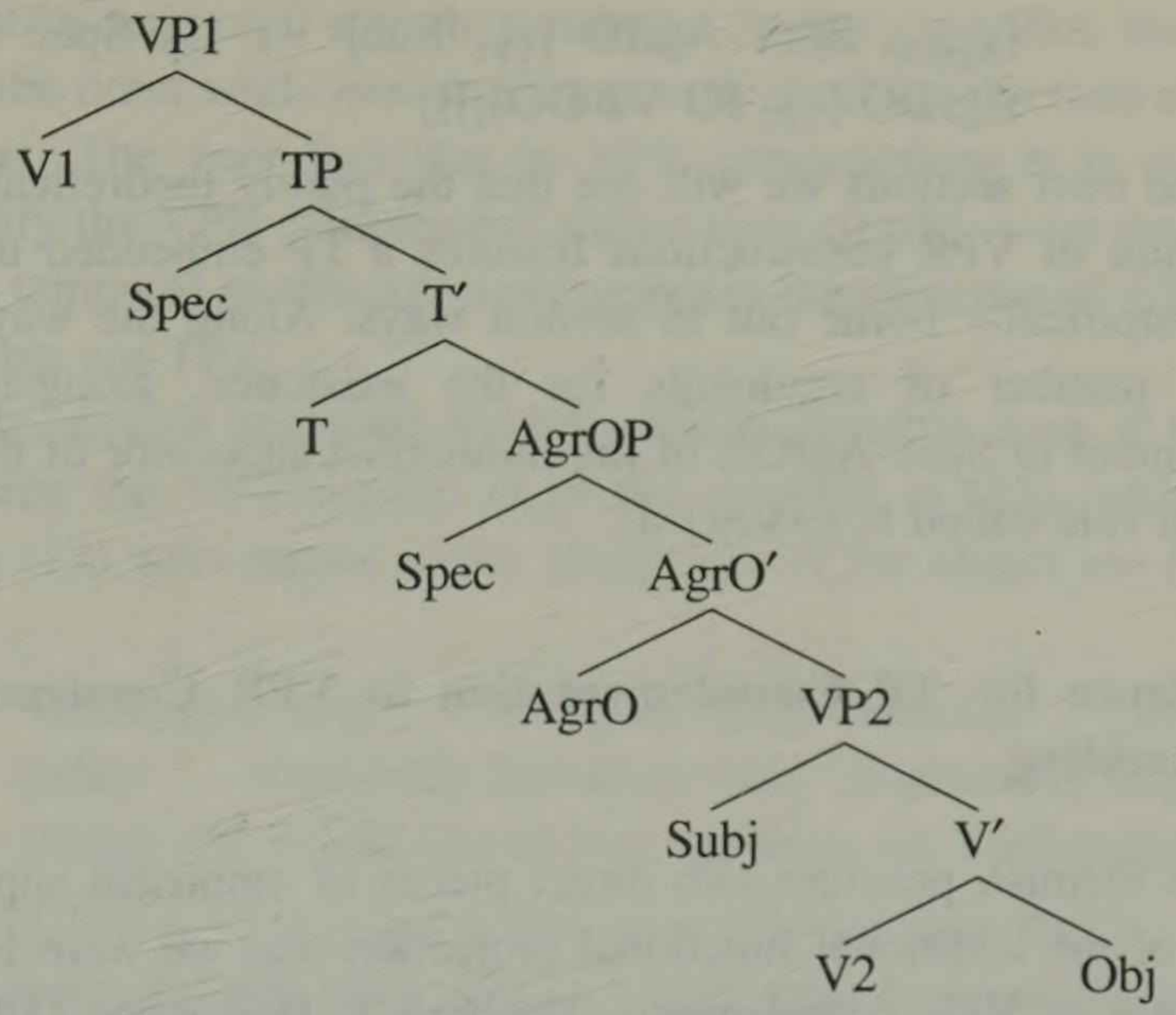
But with AgrO incorporation into V1 excluded, a grammatical derivation of a VPR construction built on (7) is impossible.[7]

The original minimalist locality theory thus gives occasion to an elaboration of the structure of VPR constructions. We need an additional projection in V1's complement in whose specifier position the subject can make an intermediate touch-down on its way up, and whose head is capable of incorporating the AgrO in its complement. Especially this latter consideration leads us to the conclusion that this ancillary projection must be a *functional* projection — for in (8) I stated, following Li's (1990) theory of head movement possibilities, that functional heads (like AgrO) can incorporate into other functional heads only. What could the label of this functional projection be? In the functional structure of clauses, AgrOP is immediately dominated by a projection of T. So let us hypothesize that the structure of VPR constructions must be adorned with an additional TP between V1 and AgrOP in (4b). This leads us to adopt the extended structure in (9), which unlike the "bare" AgrOP complementation structure in (7) caters for a well-formed derivation of VPR constructions. In particular, the subject can now raise to Spec-TP on its way up (the trigger for which will be discussed in Section 6). Though this movement skips Spec-AgrOP, this is legitimate if AgrO raises to T, rendering Spec-TP and Spec-AgrOP equidistant from the subject's extraction position. Unlike AgrO incorporation into V1 (which one would have to resort to in (7)), functional head movement of AgrO to T is of course in perfect conformity with (8).

---

7. In Den Dikken (1994a) I presented a structure of the type in (6) as one of the structural representations of transitive perfective constructions, with V1 being the auxiliary of the perfect, and I had AgrO incorporate into the auxiliary in the course of the derivation. There I said that the auxiliary *be* (which changes into *have* whenever the AgrO incorporating into it possesses a Case feature) is the only verbal head capable of incorporating Agr. From the perspective on *be* developed in more recent work (see Den Dikken 1995c) it now follows why *be* is unique in this respect — *be* is the lexical realization of a *functional* (not a lexical) head.

(9)



So we see that the locality theory of Chomsky (1993) effectively *forces* the inclusion of a TP in the complement of the higher verb in VPR constructions.[8] On the basis of a very similar train of thought, Collins & Thráinsson (1993) come to the conclusion that in a Larsonian type of approach to the structure of double-object constructions coupled with the assumption that both the direct and the indirect object check their Case features in specifier positions of AgrOPs, there is a purely theoretical need to postulate a TP in the complement of the higher (abstract causative) verb. The structure that Collins & Thráinsson assign to a double-object construction is depicted in (10). In this structure the direct object can move to Spec-AgrDOP skipping the indirect object's base position if V2 undergoes domain-extending head movement to AgrDO; the indirect object can move across the direct object's landing site under movement to Spec-TP if AgrDO incorporates into T; and with T incorporating into V1, IO can finally travel on to Spec-AgrIOP skipping the base position of the subject.[9]

---

8. The structure in (6) continues to be accurate for VR constructions; see also the discussion of (14) and note 10, below.

9. This last step continues to be problematic if incorporation of a functional head (here, T) into a lexical head (V1) is prohibited, as I have assumed (see (8)). Though bringing it up serves a purpose in the present context, I will not take over Collins & Thráinsson's (1993) analysis of double-object constructions; see also Section 5 and note 13, below.

(10) $[_{AgrIOP}$ Spec AgrIO $[_{VP1}$ Subj V1 $[_{TP}$ Spec T $[_{AgrDOP}$ Spec AgrDO $[_{VP2}$ IO V2 DO]]]]]

In the next sections we will see that the purely theoretical prediction that the structure of VPR constructions features a TP embedded under the higher verb is empirically borne out in several ways. Along the way we will come across a number of arguments for the existence, alongside Case-driven NP-movement to Spec-AgrOP, of the minimalist successor of the classic object movement rule called SCRAMBLING.

4. Evidence for TP Complementation in VPR Constructions, and for Scrambling

West Flemish provides two direct pieces of empirical support for the TP labelling of the additional functional projection that we were led to include in the structure of VPR complements. Haeberli & Haegeman (1992: fn. 5) agree with me that the structure of VPR complements must be larger than "bare VP", showing that it can contain a projection of the negative head, NegP. The evidence they present concerns the fact that the example in (11) has a so-called negative concord reading (in contrast with a double negation reading). On Haegeman's assumptions this is indicative of the presence of NegP.

(11) *da Valère durft tegen niemand nieks nie zeggen*
that Valère dares against no one nothing not say
'that Valère dares not to tell anything to anyone'

There is a close interrelationship between NegP and T. As Zanuttini (1991) argues, NegP is generable only in the presence of Tense. Saying this much (and referring the reader to Haegeman 1995a for more detailed discussion of the syntax of negation), we can take facts of the type in (11) as initial corroboration of our conclusion that the complement of VPR verbs is a TP.

Even more direct evidence for the TP status of the ancillary functional projection between V1 and AgrOP in (9) comes from the fact that it is possible to modify the VPR complement with a temporal adverbial *independently* of the matrix clause. This is shown in (12) (Haegeman, p.c.):

(12) ***Vandoage*** *zou-ze* $[_{V'}$ *willen* $[_{TP}$ $[_{AgrOP}$ *eur kleed* ***oensdag*** *kuopen*]]].
today would-she want her dress Wednesday buy
'today she would like to buy her dress on Wednesday'

In (12) the matrix temporal adverb *vandoage* 'today' specifies the time of wanting to buy the dress while *oensdag* temporally modifies the time of actually buying the dress. The very fact that in VPR constructions it is possible to temporally modify the VPR complement shows very clearly — on the plausible assumption that temporal modification presupposes the presence of a TP — that VPR complements are TPs.

Interestingly, double temporal modification is possible even if the object finds itself outside the VPR cluster, as in the example in (13), which differs minimally from (12) with regard to the placement of the object *eur kleed* 'her dress':

(13) *Vandoage zou-ze eur kleed willen oensdag kuopen.*
today would-she her dress want Wednesday buy
'today she would like to buy her dress on Wednesday'

The example in (13) is a VPR construction, as is evident from the fact that non-verbal material (here, the adverbial *oensdag*) intervenes between the modal and the main verb. But the object is not part of the VPR cluster. Does this mean that the AgrOP licensing the object's features is situated above the modal's projection, as in (14)?

(14) *$Subj_j$ ... [$_{AgrOP}$ Obj_i [$_{AgrO'}$ AgrO [$_{VP1}$ t_j [$_{V'}$ V1 [$_{TP}$ t_j [$_{T'}$ T [$_{VP2}$ t_j [$_{V'}$ V2 t_i]]]]]]]]

It turns out that the answer to this question must be negative — (14) gives rise to an ungrammatical derivation, since the object cannot reach Spec-AgrOP crossing both the base position of the subject and Spec-TP, where the subject makes an intermediate touch-down; and if the object touched down in Spec-TP on its way to Spec-AgrOP, it would trap the subject in its base position, since in that case the subject would never be able to cross both Spec-TP and Spec-AgrOP. In general, no application of (non-skipping) head movement can ever render more than two positions equidistant from a given extraction site.[10]

---

10. The parenthesis "non-skipping" in the previous statement is inconsequential in the case at hand; it might play a role, however, in contexts in which a feature-relativized minimality condition of the type argued for in Ferguson & Groat (1994) and Den Dikken & Hoekstra (1995) (see the discussion underneath (16), below) rules in head-skipping head movement. See Den Dikken & Broekhuis (1993: Section 2.7–2.8) for early discussion of these issues. Also see the discussion of (22a, b), below, for cases in which three specifiers are equidistant, as a result of the fact that two of them are specifiers of categorically identical projections immediately stacked on top of each other.

If the object cannot raise to an outside AgrOP in (13) for structural reasons, it must be the case that, in (13) as in (12), the object checks its features internal to the complement of the modal auxiliary, and that it ends up outside the VPR cluster via a movement operation other than Case-driven Object Shift to Spec-AgrOP, as depicted in (15):

(15) Subj_j ... Obj_i $[_{VP1}$ t_j $[_{V'}$ V1 $[_{TP}$ t_j $[_{T'}$ T $[_{AgrOP}$ t_i $[_{AgrO'}$ AgrO $[_{VP2}$ t_j $[_{V'}$ V2 $t_i]]]]]]]]$

Movement from Spec-AgrOP to a VP1–external position in (15) is what is traditionally referred to in the literature on Dutch, German and their dialects as scrambling (see the recent collection of articles in Corver & Van Riemsdijk 1994, and the references cited therein). What (13) suggests, then, is that, even in the AgrOP era, Dutch and German objects continue to have the option of undergoing a clause-internal movement operation that is distinct from Case-driven movement to Spec-AgrOP. And since scrambling is not Case-driven, the fact that it crosses two positions included in the chain formed by Case-driven subject movement (Spec-TP and Spec-VP1) incurs no locality problems, given a feature-relativized approach to minimality (see note 10 and the discussion of (16), below).

The existence of scrambling will receive interesting confirmation from the domain of quantifier scope interactions in V(P)R constructions, which are discussed in Section 5. Before turning to these facts, however, let us dwell some more on the difficult question of what the status of scrambling could be in a minimalist theory. Since minimalism reduces all movement to the need to check features, some sort of feature checking must be going on in scrambling. And given that features can only be checked in the checking domains of functional heads, this in turn implies that scrambling must target some Spec-FP position, the nature of "F" remaining to be determined. It is clearly a head different from the one in whose specifier topicalized constituents land. This can be ascertained particularly clearly from an inspection of the contrast in (16), noted in Müller (1993):

What is said in the main text concerning the ill-formedness of (14) carries over to VR constructions — no TP can be present in the complement of the higher verb whenever the AgrOP licensing the embedded verb's object is generated in the matrix functional domain, for reasons of locality of movement.

(16) a. **daß* [*zu lesen* t_i]$_j$ *keiner das Buch*$_i$ *versucht* t_j *hat*
that to read no one the book tried has
b. [*zu lesen* t_i]$_j$ *hat keiner das Buch*$_i$ *versucht* t_j
to read has no one the book tried
'(that) no one has tried to read the book'

Scrambling of *das Buch* 'the book' interferes with scrambling of the remnant infinitival clause *zu lesen*, as in (16a), but it does not obstruct remnant topicalization (Den Besten & Webelhuth 1990) of the infinitival clause, as the grammaticality of (16b) shows.

Basing himself on Müller's (1993) original account, Kitahara (1994: 84–87) presents a minimalist analysis of the pair in (16), built on a feature-relativized minimality approach (cf. also Ferguson & Groat 1994 and Den Dikken & Hoekstra 1995). The idea is that a constituent bearing a certain feature may not move across a c-commanding element bearing the same feature unless the feature-checking position and the feature-identical position crossed in the course of movement are in the same minimal domain. In the case of (16a), the scrambled infinitival phrase moves across the scrambled direct object, and no minimal domain can be constructed that contains the scrambled object and the landing site of the infinitival clause, hence a violation of locality theory ensues. In (16b), on the other hand, the infinitival phrase undergoes topicalization, which is triggered by a different feature, hence is not hampered in any way by scrambling of the object.

The discussion of (16) strongly underscores the fact that objects can undergo scrambling. For clearly, if all the object got involved in (16a) were Case-feature checking movement to Spec-AgrOP, there would be no reason why it should block scrambling of the infinitival phrase across it, given that the infinitival clause certainly does not undergo Case-checking movement. Of course this last point in itself already suggests that not all scrambling can be reduced to Case-driven movement to Spec-AgrOP. But the account of (16a) also makes the more specific point that structurally Case-marked NPs can undergo not just movement to Spec-AgrOP but also scrambling.

5. Verb (Projection) Raising, Scope and Scrambling

As announced in the previous section, interesting confirmation for the existence of scrambling in West Flemish can be found in the domain of

quantifier scope interactions. Of central concern in this section will be the double-object triplet in (17):

(17) a. *dan ze twee studenten vier boeken wilden geven* (2 >/< 4)
that they two students four books wanted give

b. *dan ze wilden twee studenten vier boeken geven* (2 >/?< 4)
that they wanted two students four books give

c *dan ze twee studenten wilden vier boeken geven* (2 >/*< 4)
that they two students wanted four books give
'that they wanted to give two students four books'

While (17a) is ambiguous between a reading in which precisely four books are at stake (wide scope for *vier boeken*; 2<4) and one in which the total number of books is eight (narrow scope for *vier boeken*; 2>4), and (17b) is similarly ambiguous (though apparently somewhat less readily), the "object split" example in (17c) is not ambiguous — it only has the narrow-scope reading for *vier boeken*.[11] The scopal rigidity of (17c) and the ambiguity of the other two sentences can be made to follow from the minimalist structures of the sentences in (17), given a theory of scope interactions that makes crucial reference to Case-driven movement to Spec-AgrP (and does away with Quantifier Raising; see May 1985). Before returning to the paradigm in (17) I will first of all outline such a theory of scope — the one presented in Kitahara (1992).[12]

5.1. *Scope and Feature-driven Movement*

Kitahara (1992) develops a theory of scope interactions built on Aoun & Li's (1991) Scope Principle (which I have stated as in (18)) and his own hypothesis in (19):

(18) *Scope Principle* (Aoun & Li 1991)
X has scope over Y if X c-commands a member of the chain containing Y.

(19) *Chain Formation* (Kitahara 1992: 56)
Each feature-checking operation creates a *distinct* chain.

---

11. I thank Liliane Haegeman for providing me with these subtle scope judgements.

12. Other recent accounts of quantifier scope interactions similarly exploiting movement to AgrP are presented by Hornstein (1995) and Pica & Snyder (1994). I have not seen the former in any detail yet; the latter does not seem to capture the facts in (17).

Adopting Chomsky's (1993) checking approach to movement, he accounts for the scope contrast in (20) with the aid of the structural representations in (21):

(20) a. *Someone loves everyone.* (*someone >/< everyone*)
b. *Who loves everyone?* (*who >/*< everyone*)

(21) a. $[_{IP}$ *someone*$_i$ I $[_{AgrOP}$ *everyone*$_j$ AgrO $[_{VP}$ t_i *loves* $t_j]]]$
b. $[_{CP}$ *who*$_i$ C $[_{IP}$ t_i^x I $[_{AgrOP}$ *everyone*$_j$ $[_{VP}$ t_i^y *loves* $t_j]]]]$

Scopal ambiguity arises in (21a) due to the fact that *someone* c-commands *everyone*, and *everyone* c-commands a member of the chain containing *someone* (viz. the trace t_i of *someone*). In (21b), on the other hand, while *who* c-commands *everyone* and can accordingly take scope over the universal QP, *everyone* cannot have scope over *who*. This is so because *everyone* does not c-command a member of *the chain containing who* — *everyone* does c-command t_i^y, but this trace is a member of the chain (t_i^x, t_i^y), a chain that does not contain *who*; *who* is a member of the chain (*who*, t_i^x), which is a separate chain (since it involves a different feature-checking operation — viz. checking of the *wh*-feature) of which no member is c-commanded by *everyone*. The combination of (18) and (19) thus correctly leads us to conclude that the only reading that (21b) can yield is one in which *who* takes scope over the universal quantifier.

5.2. *Scope Interactions in Double-Object V(P)R Constructions*

With this in mind, let us now return to the paradigm in (17) and see how the QR-less scope theory of Kitahara (1992) can make sense of it. Let us take up the discussion of the examples in (17) one by one, starting with the VR case in (17a), which has both objects outside the verbal cluster. This example can be analyzed as in (22a), with both AgrOPs outside the projection of V1 and no TP in V1's complement (see the discussion of (14) and fn. 10, above):

(22a)

$[_{AgrIOP}$ Spec $[_{AgrIO'}$ AgrIO $[_{AgrDOP}$ Spec $[_{AgrDO'}$ AgrDO $[_{VP1}$ Subj $[_{V'}$ V1 $[_{VP2}$ V2 IO DO]]]]]]]

In this structure, both objects undergo simple Case-checking NP-movement to their respective Spec-AgrOP positions.[13] And in so doing, they obtain ambigu-

13. The indirect object can skip Spec-AgrDOP and the subject can move across both Spec-AgrOPs given that the two AgrOPs are stacked immediately on top of each other and hence share a single set of domains (see note 5), and that V1 raises to AgrDO,

ous scope over each other — the Case-checking chain of the indirect object intersects with that of the direct object. Thus, (22a) correctly ensures scopal ambiguity for (17a). And a highly similar story applies to (17b), the VPR construction with both objects contained within the cluster. For this example the theory provides the structure in (22b), in which once again the Case-checking chains of the direct and indirect objects intersect, resulting in ambiguous scope relationships.

(22b)

$[_{VP1}$ V1 $[_{TP}$ Spec $[_{T'}$ T $[_{AgrIOP}$

Spec $[_{AgrIO'}$ AgrIO $[_{AgrDOP}$ Spec $[_{AgrDO'}$ AgrDO $[_{VP2}$ Subj $[_{V'}$ V IO DO]]]]]]]]]

Finally, consider the scopally rigid "object split" example in (17c). Suppose we were to assign this sentence the structure in (22c), which differs from the one in (22b) only with respect to the positioning of the AgrIOP — outside the projection of V1 in (22c), and inside it in (22b). If that was all there is to it, then nothing would lead us to expect a difference between (17c) and (17a,b) with respect to scopal ambiguity. After all, with the two objects both undergoing nothing more than simple Case-checking movement to Spec-AgrOP positions and with their movement chains crossing in the familiar way, the minimalist scope theory outlined in Section 5.1 would predict (17c) to be just as ambiguous as (17a ,b), contrary to fact.

(22c)

*$[_{AgrIOP}$ Spec $[_{AgrIO'}$ AgrIO $[_{VP1}$ V1 $[_{TP}$ Spec $[_{T'}$ T $[_{AgrDOP}$ Spec $[_{AgrDO'}$ AgrDO $[_{VP2}$ Subj $[_{V'}$ V IO DO]]]]]]]]]

So this cannot be all there is to it. And in fact it is fairly easy to see that (22c) cannot be correct. If the derivation of (17c) were to involve just Case-driven NP-movement of the various NPs, as in (22c), movement of the indirect object to Spec-AgrIOP across the base position of the subject, the landing site

---

thereby rendering Spec-AgrIOP, Spec-AgrDOP and Spec-VP1 equidistant. This provides a solution to the locality problem posed by double-object constructions that serves as an alternative to that proposed in Collins & Thráinsson (1993), mentioned above. Unlike Collins & Thráinsson's solution, it allows us to restrict movement of functional heads in such a way that they are only allowed to move to other functional heads, not to lexical heads (see (8)) — a restriction (going back to Li 1990) that played a central role in forcing the inclusion of a TP in the structure of VPR complements.

of the direct object, and Spec-TP would irrevocably violate the minimalist locality theory, there being no minimal domain that could ever comprise all four positions involved.

This leads me to conclude that the derivation of (17c) does *not* in fact involve NP-movement of the indirect object to a Spec-AgrIOP position outside the verbal cluster. Instead, the indirect object in (17c) is *scrambled* outside the verbal cluster. So what happens in the derivation of (17c) is that the direct and indirect objects both move to Spec-AgrOP positions in the complement of the auxiliary verb, and that subsequently the indirect object takes an additional, non-NP-movement step to a position to the left of V1 — scrambling. The derivation that ensues is summarized in (23):

(23)

IO_i [$_{VP1}$ V1 [$_{TP}$ $Subj_k$ [$_{T'}$ T [$_{AgrIOP}$ t_i^x [$_{AgrIO'}$ AgrIO [$_{AgrDOP}$ DO_j AgrDO′ AgrDO [$_{VP2}$ t_k [$_{V'}$ VP t_i^y t_j]]]]]]]]]

The thing to bear in mind about the derivation in (23) is that the step that the indirect object takes to end up to the left of V1 in the surface order of (17c) involves a feature-checking operation *distinct* from the one that the direct object undergoes.

With this said, the fact that in (17c) *vier boeken* 'four books' can only have narrow scope with respect to *twee studenten* 'two students' now naturally follows from the Scope Principle (18) and Kitahara's (1992) Chain Formation condition (19). While the shifted direct object (*vier boeken*) in (23) c-commands t_i^y, this trace is not a member of the chain that physically contains the indirect object *twee studenten*; instead it is a member of the chain (t_i^y, t_i^x), of which the indirect object is not physically a part. No member of the chain of which does contain the indirect object — the scrambling chain (IO_i, t_i^x), involving a *distinct* feature-checking operation — is c-commanded by the direct object at any point in the derivation. The Scope Principle and (19) thus correctly guarantee that the direct object cannot take scope over the indirect object in (17c). As pointed out before, in the other two constructions both the direct and the indirect object move only once, in a "crossing paths" fashion. As a result, the Scope Principle predicts that there will be scopal ambiguity in these examples, which is in agreement with the facts.

The scopal properties of the paradigm in (17) thus follow from a minimalist analysis of the word order of VR and VPR constructions, in combination with the Scope Principle (based on Aoun & Li's 1991 work) and Kitahara's (1992) Chain Formation condition in (19), both of which are independently supported.

An additional important result of this discussion is that (17c) confirms our earlier conclusion that the grammar of West Flemish features a scrambling operation that is featurally distinct from Case-driven NP-movement to Spec-AgrOP.

6. VPR and Cluster-contained Subjects in Transitive Expletive Constructions

After this interlude on scope interactions in V(P)R constructions, let us return to the main theme of the paper — the argument for TP complementation in VPR constructions. In this section I will present a fairly detailed investigation of VPR in West Flemish TRANSITIVE EXPLETIVE CONSTRUCTIONS (TECs), which will be seen to bring forth yet another argument for the inclusion of a TP in VPR complements and which moreover will reveal that — just as it should, on minimalist assumptions — the subject moves to the specifier position of the embedded TP for feature-checking purposes, thereby clearing up one of the issues remaining after the discussion in Section 3.

TECs are transitive sentences whose lexical subject surfaces in some position lower than the root Spec-AgrSP position, this position being occupied by an expletive. The analysis of TECs was first addressed from a minimalist perspective by Zwart (1992a), who focused on Dutch examples. More recently, TECs have grown into a minimalist's pet due to the detailed attention that Chomsky (1995a, b) has given to Jonas & Bobaljik's (1993) study of Icelandic TECs. My discussion of TECs in this section will be limited to West Flemish examples, of which the ones in (24) illustrate the basic TEC pattern in simple, non–V(P)R constructions:[14]

(24) a. *dan der vee studenten dienen boek kuopen*
that there many students that book buy

---

14. All examples of West Flemish TECs quoted in this section are due to Liliane Haegeman (p.c.). The paradigms of TECs reported here sketch a pretty much exhaustive picture of the possibilities that the combination of TECs and VPR offers. Whatever the ultimate value of the theoretical perspectives on these examples to be presented in what follows, it is to be hoped that the discussion in this section will contribute to a widening of the empirical scope of minimalist discussions of transitive expletive constructions.

b. *dan der dienen boek vee studenten kuopen*
that there that book many students buy
'that many students buy that book'

In (24a) the subject, *vee studenten* 'many students', has demonstrably raised across the landing site of the object (which, at the lowest, is Spec-AgrOP). But apparently the subject does not have to so raise in West Flemish TECs, as the grammaticality of (24b), in which the indefinite subject shows up to the right of the shifted object, bears out.

There is of course an immediate question posed by the pair in (24): where is the subject? Continuing to assume the Agr-based theory of phrase structure, I will follow Chomsky (1995a) in taking the expletive *der* to occupy Spec-AgrSP in (24a, b). With the object shifted to a position at least as high as Spec-AgrOP, the subject in (24a) must have moved overtly to Spec-TP. There must be a trigger for this instantiation of NP-movement. Arguably, then, the subject checks some strong feature in Spec-TP in (24a). Let us assume, as seems plausible, that the feature in question is the strong nominative Case-feature of T. Now consider (24b). Here, the subject in all likelihood occupies its base position.[15] But if the subject has a reason to move overtly in (24a), how come it can abstain from overt movement in (24b)? The only possible answer to this question, from a minimalist perspective, is that the subject is allowed to stay put

---

15. The likelihood of the subject's being in its base position in (24b) is enhanced by two observations. One is that the DO–Subj order displayed by this example is much more acceptable than the same order in a non-expletive construction like (i), in which the subject is at least as high as Spec-TP and the object is scrambled across it — cross-subject scrambling is very marginal in West Flemish; but since no degradation of this sort is noticeable in (24b), cross-subject object scrambling cannot be involved here. This latter point is confirmed by the fact that in (ii) the adverb *morgen* 'tomorrow' finds itself to the left of the object, which hence has not scrambled across it. If the placement of *morgen* in (ii) shows that no object scrambling has taken place in this example, it must be the case that the object is sitting in its Case-checking position, Spec-AgrOP; for the subject there would then basically remain exactly one position — Spec-VP, its base position.

(i) ??*dan dienen boek al de studenten gekocht een*
that that book all the students bought have
'that all the students bought that book'

(ii) *dan der morgen dienen boek vee studenten kuopen*
that there tomorrow that book many students buy
'that many students will buy that book tomorrow'

in Spec-VP because T's strong Case-feature is checked by some other element in (24b). The obvious candidate for this role is the expletive *der*, which may in principle be freely inserted at any point in the derivation.[16] In (24b), then, *der* is base-inserted in Spec-TP, checking T's nominative-Case feature there, and raises to Spec-AgrSP to satisfy the Extended Projection Principle (which requires every clause to have a subject).[17] The analyses of (24a, b) that these suggestions give rise to are made explicit in (25):

(25) a. $[_{AgrSP}$ *der* AgrS $[_{TP}$ *vee studenten*$_i$ T $[_{AgrOP}$ *dienen boek*$_j$ AgrO $[_{VP}$ t_i *kuopen* $t_j]]]]$

---

16. In Den Dikken (1995b) I make a case for the view that *there*-expletives are inverted small-clause predicates, hence not truly expletive elements. But for the sake of argument (so as to closely align the discussion of West Flemish transitive expletive constructions with Chomsky's account of their Icelandic counterparts), I will here assume that *transitive* expletive constructions of the West Flemish, Dutch and Icelandic type receive a Chomskyan analysis, which takes the expletive to be a dummy element. (This does not imply, however, that I necessarily agree with Chomsky that the expletive needs to be "replaced" at LF via adjunction of the associate.)

17. Chomsky (1995a: 431) rules out insertion of the expletive in the matrix Spec-TP and subsequent expletive raising to Spec-AgrSP, on the assumption that the expletive, though possessing a Case-feature, lacks the agreement features that could force its movement to Spec-AgrSP. Chomsky's suggestion is built on the premise that there exist no alternations between TECs with the structure in (i-a) and those with the structure in (i-b); in order to prevent (i-a) from being blocked by (i-b), Chomsky invents a way of making (i-b) crash. West Flemish shows, however, that alternations between (i-a,b) do exist, and that (i-b) is not ungrammatical.

(i) a. $[_{AgrSP}$ EXPL AgrS $[_{TP}$ Subj T ... $[_{VP}$ ...]]]

b. $[_{AgrSP}$ EXPL$_i$ AgrS $[_{TP}$ t_i T ... $[_{VP}$ Subj ...]]]

I believe that, even though expletives lack agreement features that could trigger their movement to Spec-AgrSP, expletive movement to this position could still be forced by whatever causes the effects of the Extended Projection Principle; for the sake of concreteness, let us call it an "EPP feature", which is strong.

Chomsky's (1995a) suggestion concerning the (counterfactual) ill-formedness of (i-b) is eliminated in Chomsky (1995b), where it is assumed that expletives can check neither Case nor agreement features. I will continue to assume that expletives check Case. In general I will by and large refrain from discussing Chomsky's (1995b) account of TECs in terms of a strong "multiple specifier" feature on T; I will also largely ignore questions about economy (if it is cheaper to insert an expletive than to move a constituent into a newly created position, how could there ever be a choice between the two, as, apparently, in the case of Spec-TP in (i)?).

b. $[_{AgrSP}$ *der*$_k$ AgrS $[_{TP}$ t_k T $[_{AgrOP}$ *dienen boek*$_j$ AgrO $[_{VP}$ *vee studenten kuopen* $t_j]]]]$

With these analyses of (24a,b) in place, we are now sufficiently equipped to face the intricacies of V(P)R in transitive expletive constructions. Consider the paradigm in (26), the last three examples of which display one of the most startling facts about VPR — the ability of the subject (which is underlined throughout (26)) of a transitive construction to appear *within* the VPR complement in TECs:

(26) a. VR
dan der meer studenten dienen boek zoun moeten kuopen
that there more students that book would must buy

b. VR
dan der dienen boek meer studenten zoun moeten kuopen
that there that book more students would must buy

c. Split VPR
dan der meer studenten zoun moeten dienen boek kuopen
that there more students would must that book buy

d. Split VPR
dan der dienen boek zoun moeten meer studenten kuopen
that there that book would must more students buy

e. Total VPR
dan der zoun moeten meer studenten dienen boek kuopen
that there would must more students that book buy

f. Total VPR
dan der zoun moeten dienen boek meer studenten kuopen
that there would must that book more students buy
'that more students should buy that book'

The paradigm in (26) is organized according to ascending degree of "cluster containment" — the examples in (26a, b) instantiate simple VR, those in (26c, d) are VPR constructions with one of the two arguments (the object in (26c) and the subject in (26d)) inside the VPR cluster and one outside it, and the sentences in (26e, f) exemplify full VPR, both the subject and the object being contained in the VPR cluster. Though of these examples, the last three are going to be of particular interest, let me briefly show that the analyses of the sentences in (26a–c) essentially run parallel to those of (24a, b). The VR examples in (26a–c) all feature a stacked VP structure in which the complement of the auxiliary *zou* 'would' is the bare VP of the auxiliary *moeten* 'must', which in its turn takes

the bare VP projected by the main verb *kuopen* 'buy' as its complement. In agreement with the analysis of VR constructions outlined in Section 3, the base position of the subject in (26a–c) is the specifier of the highest head (*zou*) of the stacked VP structure. In (26b) this Spec-VP position is the one that the subject surfaces in. Overt object movement to (at least) Spec-AgrOP, which finds itself above the stacked VPs, results in the object–subject order exhibited by this example, just as in (24/25b). In the examples in (26a, c), on the other hand, the subject raises to Spec-TP in overt syntax, to check T's strong nominative Case-feature, which would otherwise remain unchecked in these two constructions, in which the expletive is base-inserted in Spec-AgrSP, as in (24/25a).

Now let us turn to the examples in (26d–f), which all feature the subject in some lower position, contained in the VPR complement. Of these three examples, (26d) is derived from either (26e) or (26f) via scrambling of the object from out of the VPR complement into a position in the matrix functional domain (see the discussion of (13) in Section 4). This example is hence not going to show us anything specific with respect to the position of the subject which the sentences in (26e, f) could not also show, which is why I will concentrate on (26e, f) in the remainder of this section. Within the VPR complement of these examples, the object moves overtly to Spec-AgrOP, given that it ends up to the left of the main verb *kuopen*. For the position of *meer studenten*, this means that in (26f) this subject NP in all likelihood occupies Spec-VP. But in (26e) it cannot be in its base position. Nor can it have moved to the matrix Spec-TP, the highest position that the subject can reach in the overt syntax of transitive expletive constructions. For (26e), then, we are in need of some intermediate landing site for the subject. This landing site is readily supplied by the analysis of VPR constructions outlined in Section 3, and supported with empirical evidence in Section 4 — the specifier position of the TP embedded under the modal can serve as such.

The presence of a TP in the VPR complement is vindicated for (26e) by the possibility of temporally modifying the VPR complement with the aid of an adverb like *morgen* 'tomorrow', as shown in (27a). For the example in (26f), which has the subject surfacing all the way to the right of the cluster-contained object, similar evidence for the presence of an embedded TP is available, as (27b) shows.

(27) a. *dan der zoun moeten (morgen) meer studenten (morgen) dienen boek kuopen*
b. *dan der zoun moeten (morgen) dienen boek (morgen) meer studenten kuopen*

The examples in (27) each pose a question of their own. In (27a) the subject has moved overtly to Spec-TP, so the question is what could have triggered this movement. For VPR constructions in which the subject moves all the way out of the VPR complement, one might assume that Form Chain readily deploys an intermediate trace in the embedded Spec-TP to comply with the demands of the minimal link condition, and that there need not be any particular TP-related feature-checking motive involved in such cases. But clearly, for (26e) and (27a) no such strategy will work. The subject overtly moves no further than the embedded Spec-TP, and must have some intrinsic motive for undertaking this journey. Following Chomsky & Lasnik (1993), I will assume that non-finite T possesses a strong "null Case" feature. I will further assume, in the wake of a suggestion made by Ferguson & Groat (1994: 25) with reference to the analysis of to-infinitival ECM constructions, that the NP moving to non-finite Spec-TP in transitive expletive VPR constructions of the type in (26e)/(27a) undergoes this movement operation in order to check T's null Case feature. This is what Ferguson & Groat refer to as "short-sighted" movement to the closest Case-checker.

Let us assume that this sufficiently answers the trigger question arising in connection with (27a). Next consider the case of (27b). The subject has not moved overtly in this example, as is particularly evident from its placement to the right of the temporal adverb *morgen* 'tomorrow'. At the same time it is clear, from the very presence of *morgen* within the VPR complement, that there must be a TP present in the complement of the modal in this example — and along with it, the strong feature that the subject checks in (27a) will be present in the VPR complement as well. So here the question is how this strong feature is checked in (27b). Since the subject evidently does not check it, the only other candidate is the expletive *der*, which ends up in the matrix Spec-AgrSP at the end of the overt-syntactic derivation. I conclude, then, that in (27b) (and (26f)), the expletive is inserted in the specifier position of the *embedded* TP, checking this T's strong feature, and subsequently raises, via the root Spec-TP, to its final destination (which it must reach in order for the Extended Projection Principle's demands to be complied with).

Thus, (27a) and (27b) differ minimally with respect to the position in which the expletive is base-inserted — in the root Spec-TP in the former, and the embedded Spec-TP in the latter. Since the embedded Spec-TP is otherwise unoccupied, the subject must overtly move there in (27a) to check T's strong ("null Case") feature; in (27b), on the other hand, there is no motive for subject movement, hence the subject stays put in Spec-VP. As was to be expected, then,

the pair in (27) is essentially on a par with our earlier pair in (24), the difference being that the TP into whose specifier position the subject must or must not move is the *root* TP in the latter and an embedded, infinitival TP in the former.

The discussion in this section has shown that in VPR constructions the subject is base-generated within the VPR complement, that it can move overtly to the specifier position of the TP in the VPR complement to check T's strong feature (which I have suggested is a "null Case" feature), and that T's strong feature can also be checked by insertion in the embedded Spec-TP of an expletive, which subsequently undergoes raising into the matrix clause to comply with the Extended Projection Principle. This clearly only scratches the surface of the extremely murky morass of problems associated with transitive expletive constructions. In particular, I have not addressed the question of how it can be the case that all of (26a–f) happily coexist.[18] This is an issue which our current understanding of both the theory and the intricacies of transitive expletive constructions does not seem to allow us to clear up at this stage of research. In any case, though the suggestions made here have been somewhat tentative and deserve further thought, the investigation of transitive expletive VPR constructions in this section has certainly underpinned our earlier conclusion that VPR complements include a TP, and has shown that movement of the subject to the specifier position of this embedded TP is a minimalistically well-behaved case of feature-driven movement.

7. Concluding Remarks

In this article I have presented a minimalist analysis of Verb (Projection) Raising constructions built on an antisymmetric VO structure (cf. Kayne 1994), and centered around the structures in (6) (for VR) and (9) (for VPR). The basic difference between VR and VPR constructions comes down to the absence or presence of an additional TP below the auxiliary verb. This difference is both empirically motivated, especially by the facts of temporal modification discussed in Section 4, and structurally forced, by the Agr-based minimalist locality

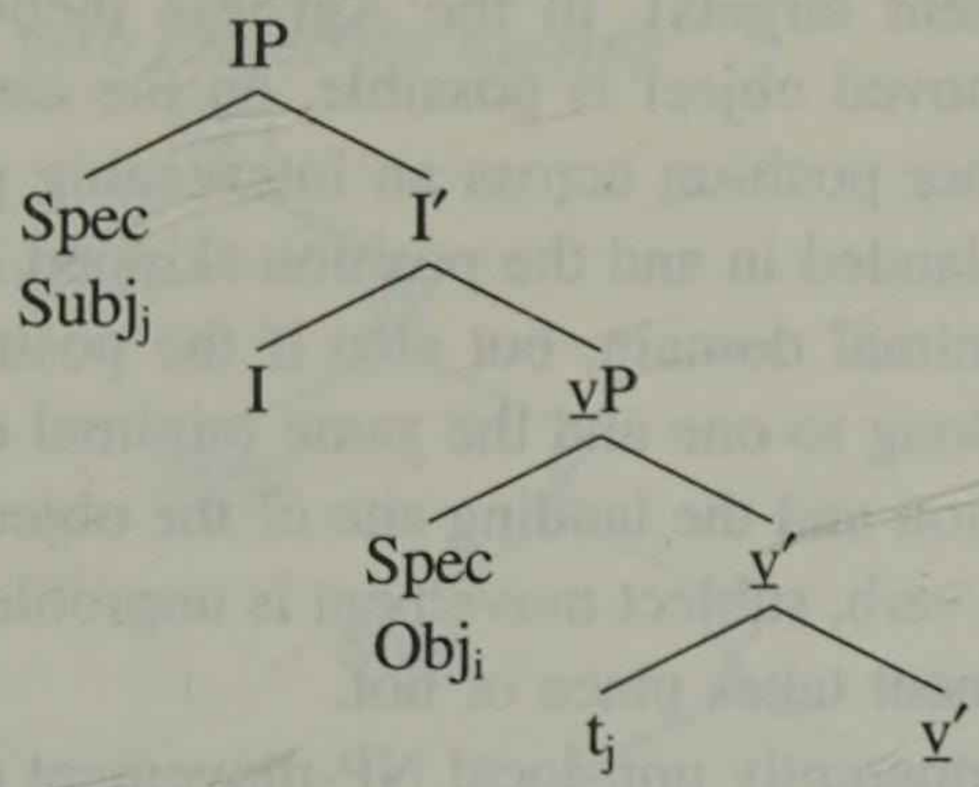
---

18. Chomsky (1995b: Section 10.2) makes an attempt at coming to grips with some of the variation in surface patterns in transitive expletive constructions (or multiple subject/specifier constructions in general), based on his Agr-less theory of phrase structure in which, in an example like (24a) above, both the expletive and the subject occupy multiple specifier positions of a single TP. I will not discuss this.

theory, which attributes a central role to domain-extending head movement. Interesting side-effects of this analysis of V(P)R constructions are the arguments that it presents for the existence of a scrambling operation distinct from Case-driven object movement to Spec-AgrOP, and the light that it sheds on the account of transitive expletive constructions, discussed in Section 6.

In my account of V(P)R constructions I have consistently made use of the Agr-based theory of feature checking of Chomsky (1991, 1993). Recently, Chomsky (1995b: Section 10) has suggested that we might be able to do without AgrPs. In the particular case of the object, we might assume that its features are checked in the "outer specifier" position of the projection of a "light verb" (see (28)).

(28) *Transitive sentences according to Chomsky (1995b: Section 10)*



This light verb projection is inspired by Larson's (1988) work on "VP shells", and more directly has as its root the work of Hale & Keyser (1993) on the representation and projection of argument structure. Following the spirit (though not the letter) of this work, Chomsky (1995b: Section 6) assumes that all transitive clauses feature a layered VP structure, the base verb combining with its object to form a VP, and this VP subsequently combining with a light verb, v. The bearer of the external θ-role is generated in the (inner) specifier position of this light verb projection, vP in (28), "the agent role being understood as the interpretation assigned to the v–VP configuration" (Chomsky 1995b: 315). And not only does the light verb take on a crucial role in the assignment of the external θ-role, it also takes over AgrO's function of checking the object's Case-features.

Comparing (28) with (1), above, the reader will quickly realize that, as Chomsky (1995b: 355) puts it, "to accommodate the change from an Agr-based to a multiple-Spec theory, we have to simplify the notions of equidistance and closeness that entered into the definition of Attract/Move." While in the

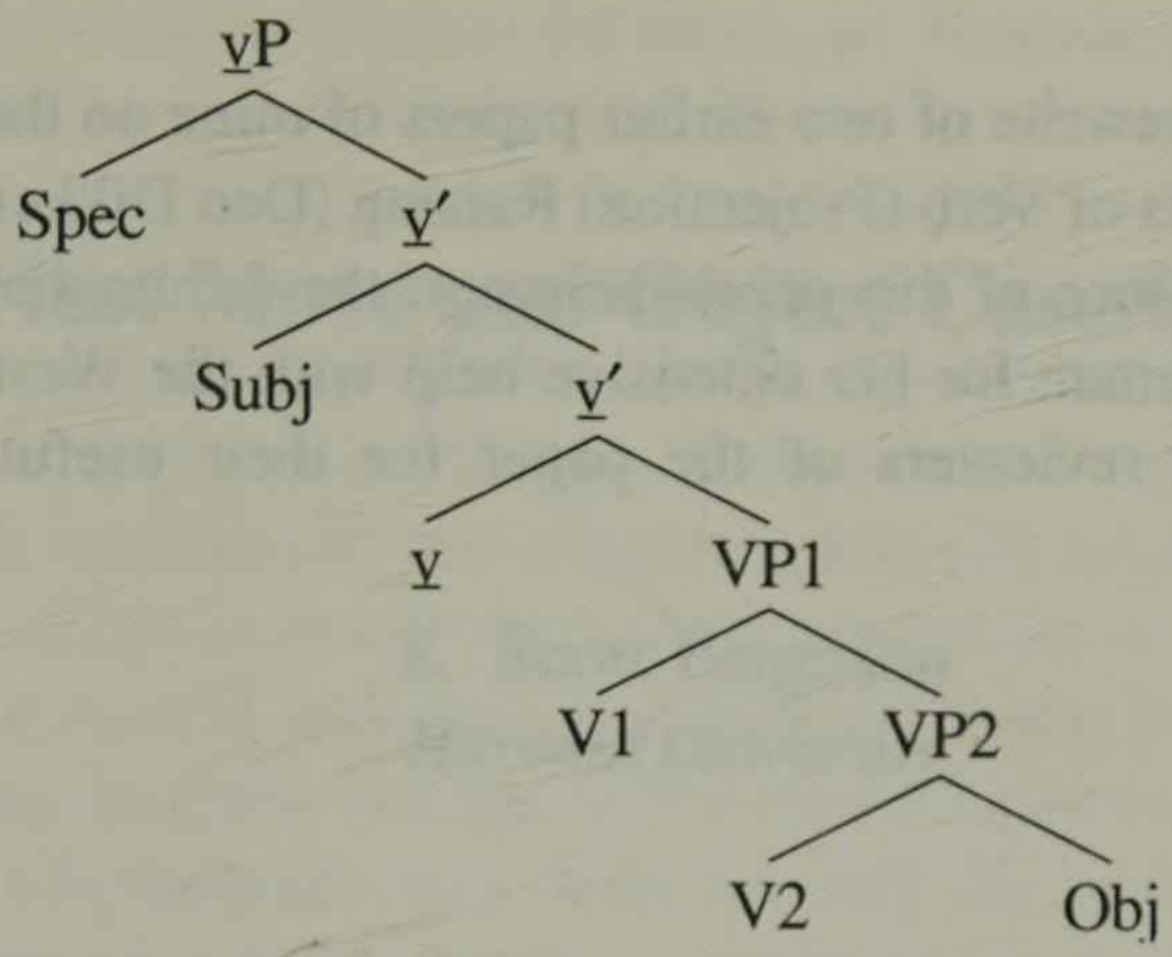
structure in (1) the object can raise to Spec-AgrOP only if there is domain-extending head movement of the verb from V to AgrO (see Section 1 for discussion), in (28) movement of the object to the "outer specifier" of the light verb projection cannot be contingent on movement of the verb. In fact, no facilitating head movement is necessary, since the object's landing site and the base position of the subject are members of the minimal domain of the trivial chain of the light verb even in the absence of movement. Thus the original dependency of L-feature-driven object movement on head movement of the verb vanishes in the "bare theory".

So does the contingency on head movement of subject movement to Spec-IP across the object's landing site. While in Chomsky (1993) this is possible only if AgrO raises to Infl (which would render Spec-AgrOP and Spec-IP equidistant movement targets), in the Agr-less theory "unmediated" subject raising across the moved object is possible, on the assumption that an element can move to a higher position across an intervening potential landing site not only if the position landed in and the position skipped in the movement process are in the same minimal domain, but also if the position started from and the position skipped belong to one and the same minimal domain. Since in (28) the subject's base position and the landing site of the object are both in the minimal domain of the light verb, subject movement is unproblematic regardless of whether any head movement takes place or not.

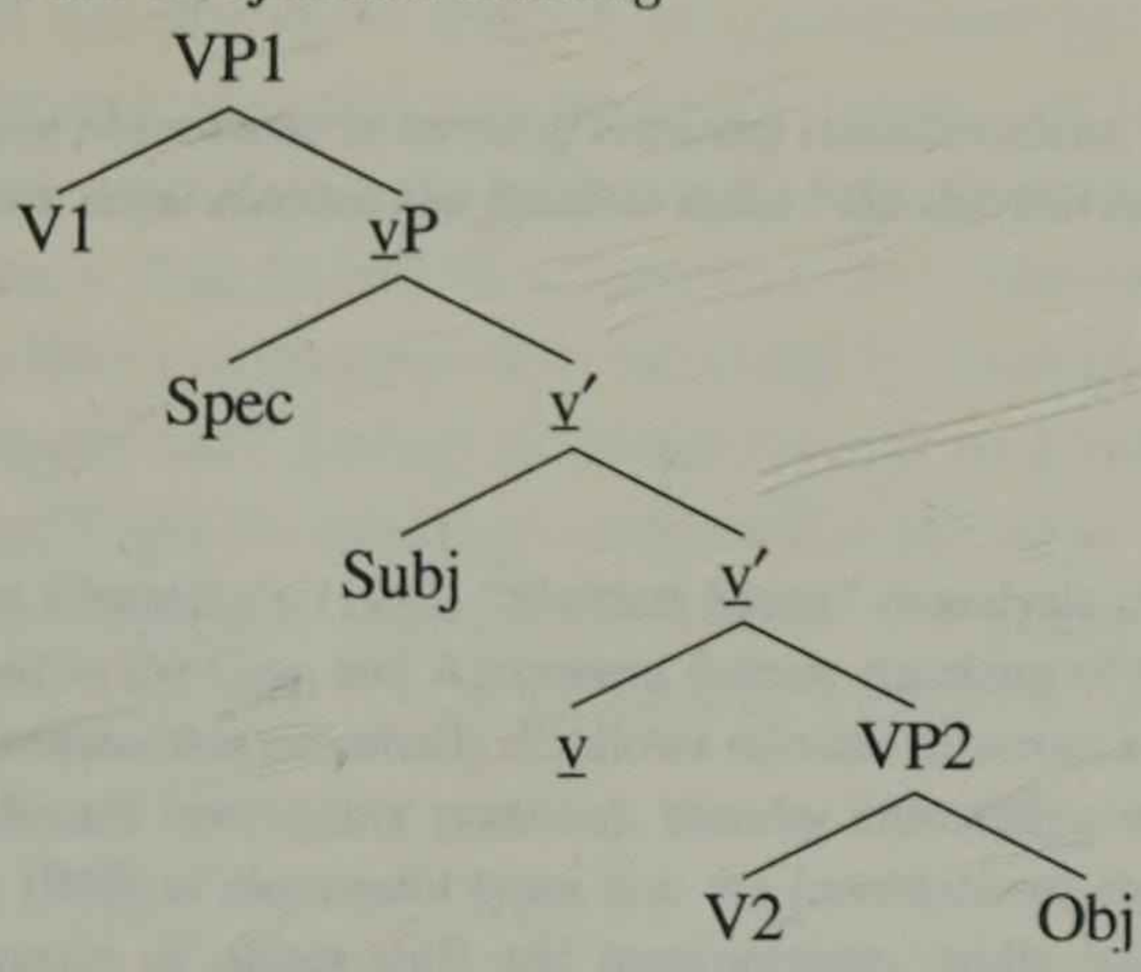
The two instances of apparently non-local NP-movement obtaining in (28) are hence legitimate automatically, without there being any need for domain-extending head movement. While this renders the minimalist locality story quite a bit simpler, it entails not only that "we have lost Holmberg's generalization and other effects of verb-raising on extension of chains" (Chomsky 1995b: 358), which may not be particularly harmful (see Section 1), but also that the theory will fail to *guarantee* the presence of a TP in VPR complements. Let us see why this is so.

Adopting the Agr-less "bare theory", we could perfectly legitimately assign VPR constructions a structure of the type depicted in (30) (with (29) being the "bare theory" representation of VR constructions, which does not differ significantly from the one given in (6), above).

(29) *Verb Raising*



(30) *Verb Projection Raising*



In (30) the object can move to the "outer specifier" of the light verb projection embedded under V1 (the auxiliary) without any trouble; and the subject unproblematically raises from its base position across the object's landing site. No additional projection between V1 and the VPR complement is necessary in order for the derivation of a VPR construction to converge. In this regard the "bare theory" differs non-trivially from the Agr-based theory in which my analysis of VPR was couched. For recall that the Agr-based locality theory *forces* the presence of a TP in the complement of the higher verb in VPR constructions, by structural considerations of locality of movement. To the extent, then, that there is evidence that VPR complements *systematically* include a TP, there is a motive for preferring the original Agr-based theory of Chomsky (1993) to its multiple-specifier successor.

Acknowledgements

This article is a total rewrite of two earlier papers of mine on the minimalist and antisymmetric analysis of Verb (Projection) Raising (Den Dikken 1994b, 1995a), focusing on the relevance of this phenomenon to the minimalist locality theory. I thank Liliane Haegeman for her extensive help with the West Flemish data. I also thank the three reviewers of the paper for their useful comments and suggestions.

Shortest Move and Object Case Checking

K. Scott Ferguson
Harvard University

Looking at these phenomena in terms of economy considerations, it is clear that in all of the "bad" cases, some element has failed to make "the shortest move".

— Chomsky (1993: 14)

Abstract

This paper examines Chomsky's (1993) "Shortest Move" reanalysis of Rizzi's (1990) Relativized Minimality as applied to the Case and Agreement feature checking of the direct object. I propose a Shortest Move requirement that essentially disallows movement across a relevant intervening checker (as opposed to a relevant intervening position), thereby eliminating equidistance stipulations and relativization (Rizzi 1990) of movement types (i.e. the [non-natural] class: A, $\overline{A}$, X^0). As a result, I obtain a unified account of object shift and incorporation, while partially deducing Holmberg's (1986) Generalization. Disjunctive formulations of the Case Filter and the Visibility Condition are also eliminated.

1. Introduction

This paper presents a reformulation of Chomsky's (1993) "Shortest Move" account of Rizzi's (1990) Relativized Minimality effects and considers its consequences regarding the checking of the syntactic Case and Agreement features of the direct object. Examining Superiority and Relativized Minimality effects (e.g., superraising, head-movement constraint violations, and *wh*-islands), Chomsky (1993: 14–15) states that the common property of all such violations is that "some element has failed to make 'the shortest move'." Regarding Relativized Minimality in particular, Chomsky (1993: 14–15, see also Frampton 1991)

proposes that in each of the deviant cases "the moved element has 'skipped' a position it could have reached by a shorter move, had that position not been filled." While pursuing the basic spirit of Chomsky's analysis, this paper investigates the intuition that an intervening CHECKER is a more natural choice for an intervening element than is an intervening POSITION in a "Shortest Move" account, especially given the driving morphological forces underlying feature checking movement (see Last Resort/Greed, Chomsky 1993: 31–34).

Section 2 outlines the relevant principles and assumptions of the minimalist framework, including an articulated I(nflection) P(hrase) and a checking theory of Case. Section 3 first reviews Chomsky's "Shortest Move" account and then presents a Shortest Move requirement that essentially disallows movement across an intervening checker of the relevant type. This Shortest Move Requirement eliminates the need for both an appeal to EQUIDISTANCE as well as stipulated relativizations of types of movements and types of positions. Section 4 shows that the proposed Shortest Move Requirement permits a unified account of object shift and object incorporation whereby each type of movement is analyzed as being an instance of movement to the closest Case checker. This analysis permits the deduction of that part of Holmberg's (1986) Generalization expressing the contingency of Object Shift on (prior) Verb Raising. Finally, the analysis eliminates the need for disjunctive versions of the Case Filter and the Visibility Condition whereby incorporated nominals are stipulated to be exempt from Case requirements. Section 5 concludes.

2. The Minimalist Framework

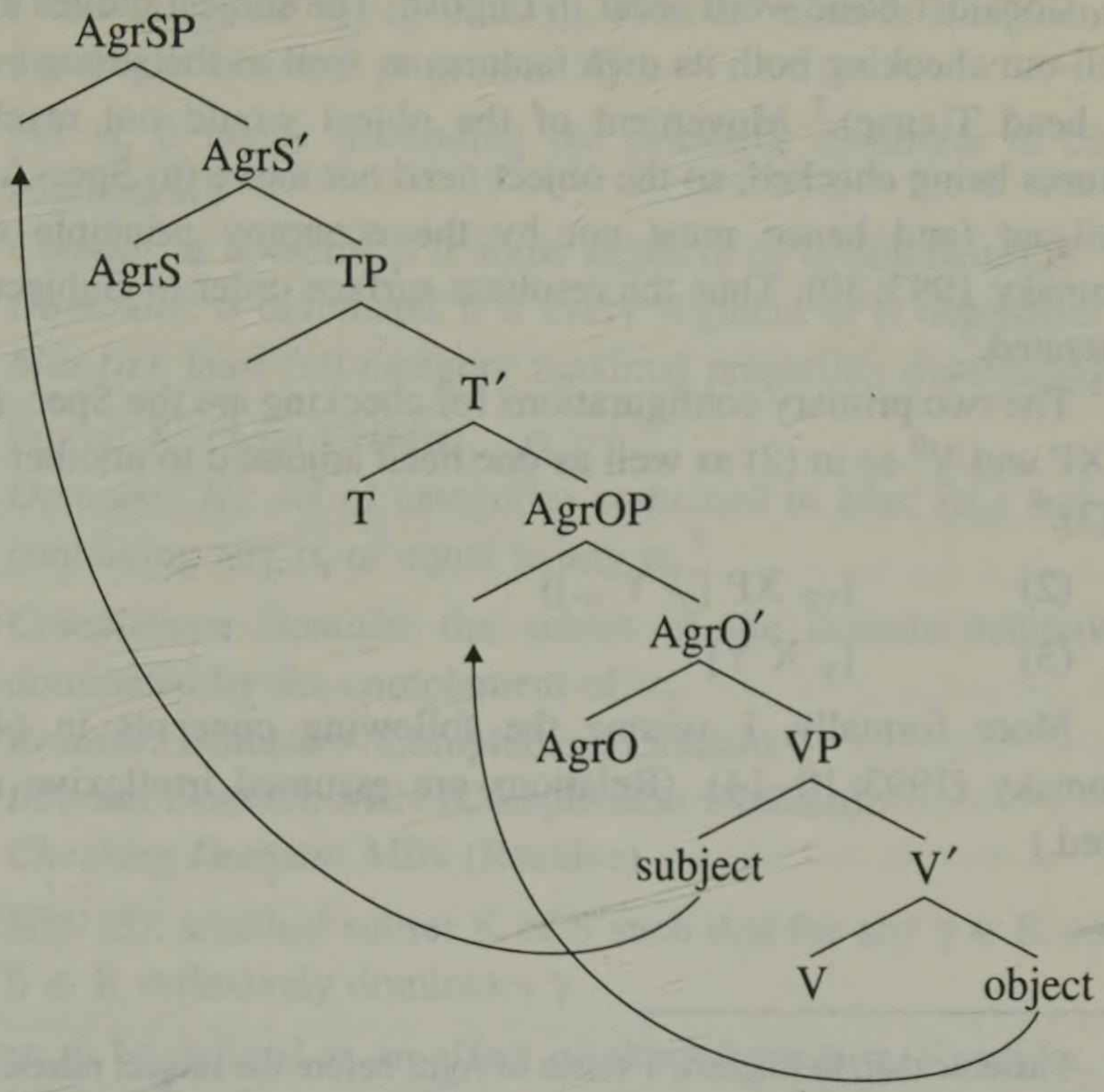
This paper generally assumes the Minimalist Program as outlined in Chomsky (1993), supplemented in places with more explicit proposals from Chomsky & Lasnik (1993). Key assumptions include the (Predicate) Internal Subject Hypothesis and an articulated IP structure (i.e. AgrSP–TP–AgrOP–VP, see also Chomsky 1991 and Pollock 1989). The subject and the object each move from a VP-internal "base" position to a Spec-AgrP position for Case and Agreement feature checking. In particular, the subject moves to Spec-AgrS, the object to Spec-AgrO.[1] Thus, (1) illustrates the basic clausal structure (apart

---

1. I generally exclude from further discussion the checking of the subject's Case and Agreement features. T(ense) presumably checks the Case of the subject — at least in nominative–accusative languages (Chomsky 1993: 9).

from CP) and the movements of the subject and object to their respective checking positions.

(1)



All movement is assumed to be motivated by the need to check features. In particular, a category moves only to check its own features (see Greed, Chomsky 1993: 33). Feature checking is possible only when an XP or an X^0 is in the checking domain (Chomsky 1993: 10–14) of a Y^0 and the XP or X^0 matches features with the Y^0. Visible unchecked features are not interpretable at the interface levels — Logical Form (LF) and Phonetic Form (PF). Uninterpretable or illegitimate objects (e.g. a starred trace/copy [*t]) cause the derivation to "crash" at the interface. All unchecked features are visible at the LF interface, hence both subject and object must raise (checking features on each as well as the relevant functional heads) by LF. Strong unchecked features are also visible at the PF interface; weak unchecked features are not visible at the PF interface. Hence, only the strong features must be checked in the overt syntactic component (i.e. by SPELL-OUT [cf. S-structure, Chomsky 1981], or at least PF if

checking is possible in the phonological component) — weak features need not be checked until LF.

Consider basic word order in English. The subject moves to Spec-AgrSP by Spell-out checking both its own features as well as the strong ⟨+NOM⟩ feature of the head T(ense).[2] Movement of the object would not result in any strong features being checked, so the object need not move (to Spec-AgrOP) until after Spell-out (and hence must not by the economy principle of Procrastinate; Chomsky 1993: 30). Thus the resultant surface order of Subject–Verb–Object is generated.[3]

The two primary configurations for checking are the Spec–Head relation for an XP and Y^0 as in (2) as well as one head adjoined to another for two heads as in (3).[4]

(2) $[_{YP}\ XP\ [_{Y'}\ Y\ ...]]$

(3) $[_{Y}\ X\ Y]$

More formally, I assume the following concepts in (4) adapted from Chomsky (1993: 10–14). (Relations are assumed irreflexive unless otherwise stated.)

---

2. I assume that, in English, T raises to AgrS before the subject raises. Should T not raise, Spec-TP arguably becomes a closer checking site (see Chomsky 1993, footnote 11). Jonas & Bobaljik (1993) argue that the subject does move overtly to Spec-TP for checking purposes in certain Icelandic constructions (e.g. Transitive Expletives).

3. But see Johnson (1991) for an analysis of English with verb raising and object movement before Spell-out.

4. Other checking domain positions for a phrase ZP with respect to a head Y^0 in the structure (i) include ZP adjoined to YP, to XP, or even possibly to WP.

(i) $[_{YP}\ XP\ [_{Y'}\ Y\ WP]]$

Chomsky (1993: 10) assumes that "there is no adjunction to complement" position. Taking "complement" in the X′-theoretic sense of "sister to a head" implies that adjunction to a maximal projection should only be possible to the root CP or maximal projections in Specifier or adjoined positions. Compare this with an arguably weaker restriction on adjunction from Chomsky (1986b: 6):

(ii) Adjunction is possible only to a maximal projection (hence, X″) that is a non-argument. [Chomsky's (6)]

If only lexical (not functional) heads take arguments, adjunction of a phrase UP to the complement WP of a functional head Y should be both licit and sufficient for UP to enter into Y's checking domain. I leave this latter possibility open.

(4) A segment-category distinction for adjunction structures (see also May 1985, Chomsky 1986b).
(Thus in $[_{XP}$ ZP $[_{XP}$ $[_{X'}$ X ...]]], each XP node is a segment of the single maximal projection XP.)

For α, β nodes (extending the primitive *dominate* to cover adjunction):
Contain: α contains β if some segment of α dominates β.
Dominate: α dominates β if every segment of α dominates β.
Max (α): least full-category maximal projection dominating α.

For an X^0-chain CH=(α_1,...,α_n):
Domain(CH): set of categories contained in Max (α_1) and not containing any α_i or equal to any α_i.[5]
Complement Domain: the subset of the domain reflexively dominated by the complement of α_1.
Residue: Domain – Complement Domain.
Internal Domain: MIN (Complement Domain).
Checking Domain: MIN (Residue).

MIN (S): smallest subset K of S such that for any $\gamma \in$ S, some $\beta \in$ K reflexively dominates γ.

Domains are taken to be defined as an effect of chain-formation. Consider, for example, a head X^0. When inserted from the lexicon (i.e. via generalized transformation), X^0 has domains defined. Suppose that X^0 adjoins to Y^0 yielding $[[_Y$ X Y] ... $t_X]$. Only the chain (X, t_X) has domains defined. Suppose further that the Y amalgam then moves to incorporate into a head Z^0. Then the new chain (Y, t_Y) has domains defined. However, the now positionally-altered head chain (X, t_X) does not have new domains defined.

3. Shortest Move

This section summarizes Chomsky's Shortest Move account of Relativized Minimality and then presents an arguably more natural formulation of the Shortest Move Requirement that blocks movement across an intervening checker

---

5. Chomsky (1993: 13) says "set of nodes ..."; however, his probable meaning "set of categories" seems clear from the text.

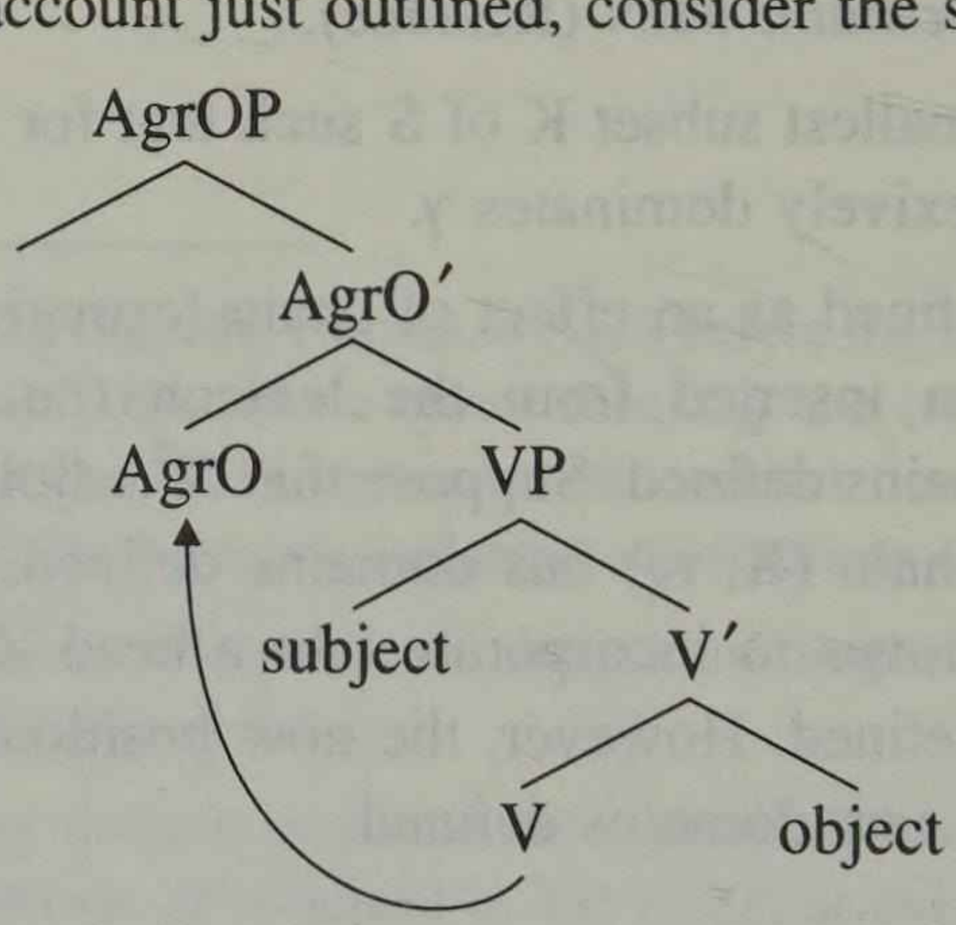
as opposed to an intervening filled position. This Shortest Move Requirement eliminates equidistance and the relativization of various types of movement.

3.1. *Shortest Move*

Chomsky postulates a Shortest Move account of Relativized Minimality effects (superraising, head-movement constraint, *wh*-islands). Chomsky's (1993: 14–15, see also Frampton 1991) account states that the common property of Relativized Minimality phenomena is that "the moved element has 'skipped' a position it could have reached by a shorter move, had that position not been filled". Following Rizzi (1990), Chomsky assumes that relevant positions and chains come in three varieties: A, $\overline{\text{A}}$, and head. Thus when a category moves to form an A-chain, it cannot skip a filled A-position; when forming an $\overline{\text{A}}$-chain it can not skip a filled $\overline{\text{A}}$-position; when forming a head chain it can not skip a head position. Note that this relativization is stipulative and non-unified; it is unclear why there are these three particular movement types, the postulation of which presupposes a non-unified theory of movement.

Given the account just outlined, consider the structure in (5):

(5)



Informally, Spec-VP (a filled A-position) is expected to block movement to Spec-AgrOP (also an A-position). Thus, in order to allow the object to skip the VP-internal subject position when raising to Spec-AgrOP, Chomsky (1993: 17–18) stipulates that "two targets of movement are equidistant [from the source of movement, K.S.F.] if they are in the same minimal domain." So in (5), provided that V first adjoins to AgrO, the resulting chain (V, t_V) will include in its minimal domain both Spec-AgrOP and Spec-VP, hence raising to Spec-AgrOP is permitted.

To sum up, the minimalist account of shortest movement relies on (what

may be a descriptively correct) stipulation of equidistance, as well as a stipulated trinary relativization (A, $\overline{A}$, X^0) over movement and position types.[6]

3.2. *The Shortest Move Requirement*

This section presents a Shortest Move requirement that captures more directly the idea that movement is motivated by the need to check features (cf. O'Neil 1992):[7]

(6) *Shortest Move Requirement (SMR)*:
Given a category α moving to check feature β:
α may not skip moving to the checking domain of the closest c-commanding Y^0 capable of checking feature β.

(7) *C-command*: α c-commands β iff the first (branching) node dominating α dominates β.

Following Chomsky & Lasnik (1993: 546), failing to meet the Shortest Move Requirement leaves behind a starred trace/copy (*t), an illegitimate object at LF.

3.3. *The Naturalness of the Shortest Move Requirement*

The feature-checking based account of Shortest Move given in the Shortest Move Requirement is composed of deeper principles of universal grammar (UG). The notions of Greed, checking, c-command, and Least Effort combine to yield the SMR. Greed specifies that a category α must be undergoing movement in order to check some feature β. If α moves to check some feature β, then α must move to the checking domain of a head capable of checking feature β. Drawing on Kayne (1994), I will assume that syntactic relations are possible only in the configuration of c-command. Finally, I argue that Least Effort forces movement to the closest such position, given that local operations are more computationally tractable (cf. Chomsky 1993: 33–34), all other things aside. Given the strong independent motivation for each of these primitives, the Shortest Move Requirement is a natural principle.

---

6. Jonas & Bobaljik (1993: 63) give a clear, explicit version of Chomsky's account nicely illustrating the stipulativeness of these particular facets of his account. See also Branigan (1992) for proposals similar to Chomsky's.

7. I leave it as an open question whether the "feature β" should be generalized to the "feature set β".

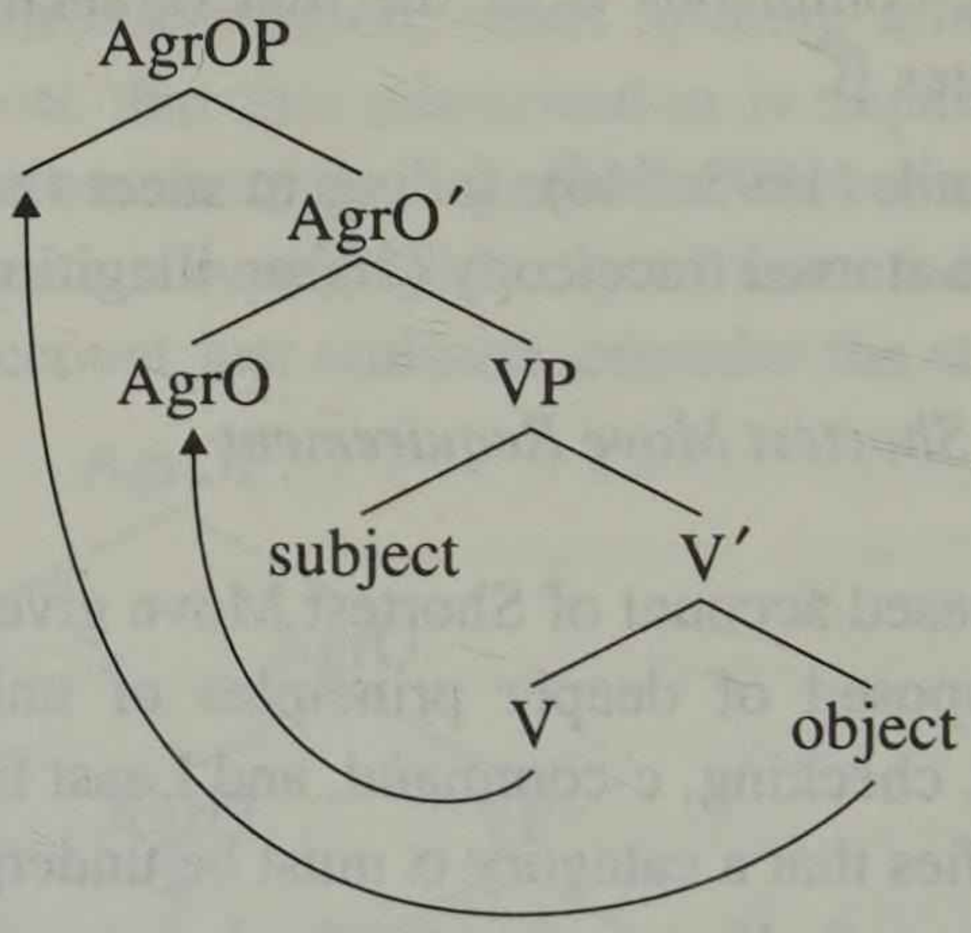
4. Object Case Checking

By object Case checking, I refer to the checking of the syntactic Case and Agreement features of the direct object. Object Case and Agreement are checked by V and AgrO. The V supplies the required accusative Case feature while the AgrO provides the necessary Agreement features (e.g. person, number, gender). I assume that even in the absence of overt morphology, the abstract features that underlie overt morphology are syntactically present.

4.1. *Object Shift*

Object Shift refers to the movement of the direct object to Spec-AgrOP as illustrated (along with V-raising) in (8).

(8)



An Icelandic example (which also shows further V-raising) is given in (9).

(9) *margir stúdentar*$_i$ *lásu*$_j$ [$_{AgrO}$ *bókina*$_k$ [$_{VP}$ *ekki* [$_{VP}$ t_i t_j t_k]]
many students read book-the not
'Many students did not read the book.'

In pre-AgrO frameworks, Holmberg (1986) and Vikner (1991) each argue for the overt A-movement of objects in Icelandic. Moreover, analyses are presented more specifically for movement of objects to Spec-AgrOP by Jonas & Bobaljik (1993) for Icelandic, Branigan (1992) for English and French, Déprez (1992) for Scandinavian, Mahajan (1990) for Hindi, and Vanden Wyngaerd (1989a) and Zwart (1993b) for Dutch.

Suppose as in (8) that the verb raises, adjoining to AgrO to form a complex

head [$_{\text{AgrO}}$ V AgrO]. From the object's perspective, the closest c-commanding element capable of checking its Case is the (V, t_V) chain. Spec-AgrO is included in the verb chain's checking domain. Hence the object may move to Spec-AgrOP, meeting the Shortest Move Requirement and getting its Case and Agreement features checked. Although Spec-VP is an intervening A-position, given V-raising, it is not in the checking domain of the V chain, the closest Case checker (to the unshifted object). Hence, the need for equidistance does not arise. In sum, object shift is movement of a nominal to the checking domain of the closest Case checker.

4.2. *Object Incorporation*

In this subsection, I argue that OBJECT INCORPORATION (just like Object Shift) is movement of a nominal to the checking domain of the closest Case checker.

4.2.1. *Standard Cases of Noun Incorporation*

Following Baker (1988), NOUN INCORPORATION (NI), is analyzed as the adjunction of a noun to another head. Consider the standard cases of noun incorporation into V (see Baker 1988: 81–87) and noun incorporation into P (see Baker 1988: 90) illustrated in (10) and (11):[8,9]

---

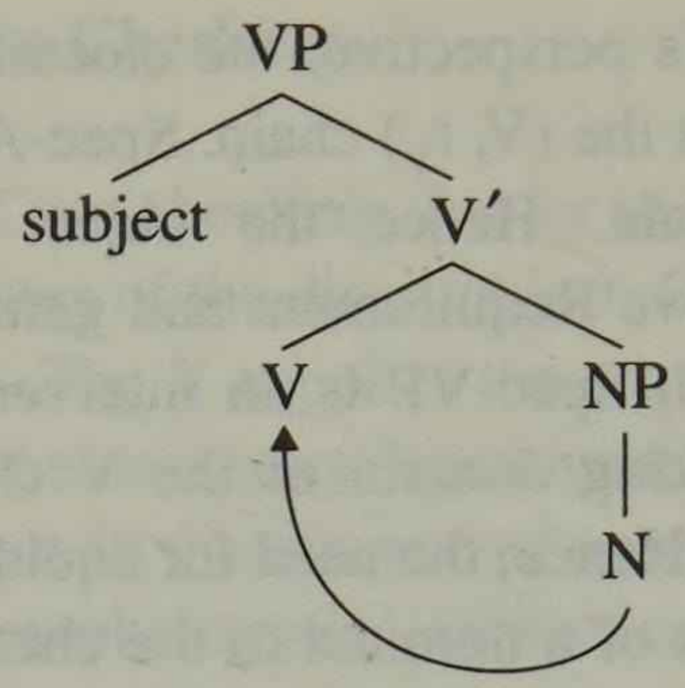
8. Baker (1988: 87–90) reports that an unaccusative subject (which is an underlying object by hypothesis) can also incorporate into a verb as in (i).

(i) *Ka.**hsahe²t**.ahi.hw.i*
3NEUT.**bean**.spill.CAUS.ASP
'The beans spilled.' (Onondaga: Woodbury 1975; Baker's: 35b)

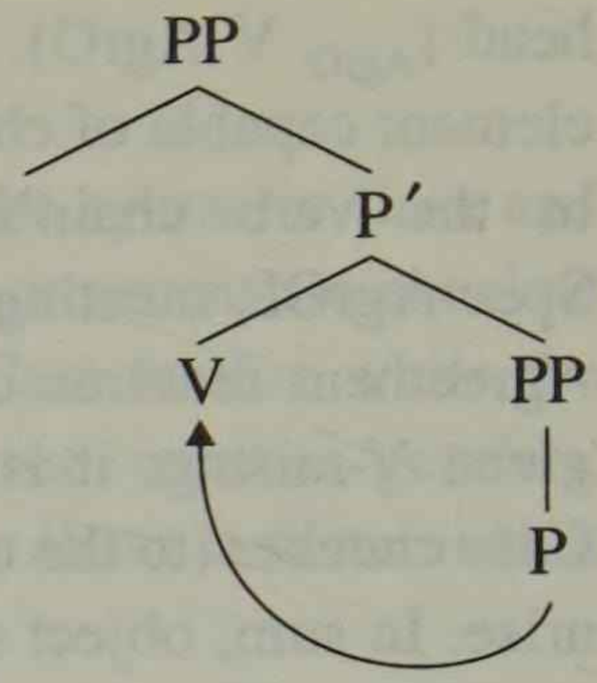
Unaccusative subjects presumably check Case with Tense as other subjects do. If all noun incorporation is driven by Case checking, then either the unaccusative subject incorporates into the verbal complex after the verb picks up Tense, or the verb somehow appears to be capable of checking Case when it is not. See Ferguson & Groat (1995) for similar cases. Regardless, the SMR in no way prohibits the noun from incorporating into the verb.

9. In the trees illustrated in (10) and (11) and elsewhere, I show a N^0 category dominated by an NP node. Following Lasnik & Kupin (1977), I assume non-branching projection such as the NP to be essentially notational (i.e., the structure doesn't really exist). Given Chomsky's (1994) bare theory of phrase structure, we can assume that the noun

(10)



(11)



An example of NI into V is given in (12), an example of NI into P in (13).[10]

(12) *Yao.wir.aʔa ye.**nuhs**.nuhweʔ.s*
PRE.baby.SUF 3FEM-SUBJ/3NEUT.**house**.like.ASP
'The baby house-likes' (Mohawk: Postal 1962; Baker's: 14b)

(13) *o.**'hoñt**.ako'*
PRE.**bush**.in
'into the bushes' (Mohawk: Hewitt 1903; Baker's: 43a)

In standard accounts, the lexical categories V^0 and P^0 have the capacity to assign Case features to their nominal complements. Since nouns typically incorporate into V and P, and movement is assumed to be motivated only by the need to check features, I argue for the null hypothesis, namely, that incorporation is motivated by need for feature-checking, in particular Case-feature checking. Thus, the complement N moves to the checking domain of the closest c-commanding head capable of checking Case (i.e. it adjoins to V or to P).

4.2.2. *Noun Incorporation over a Determiner*

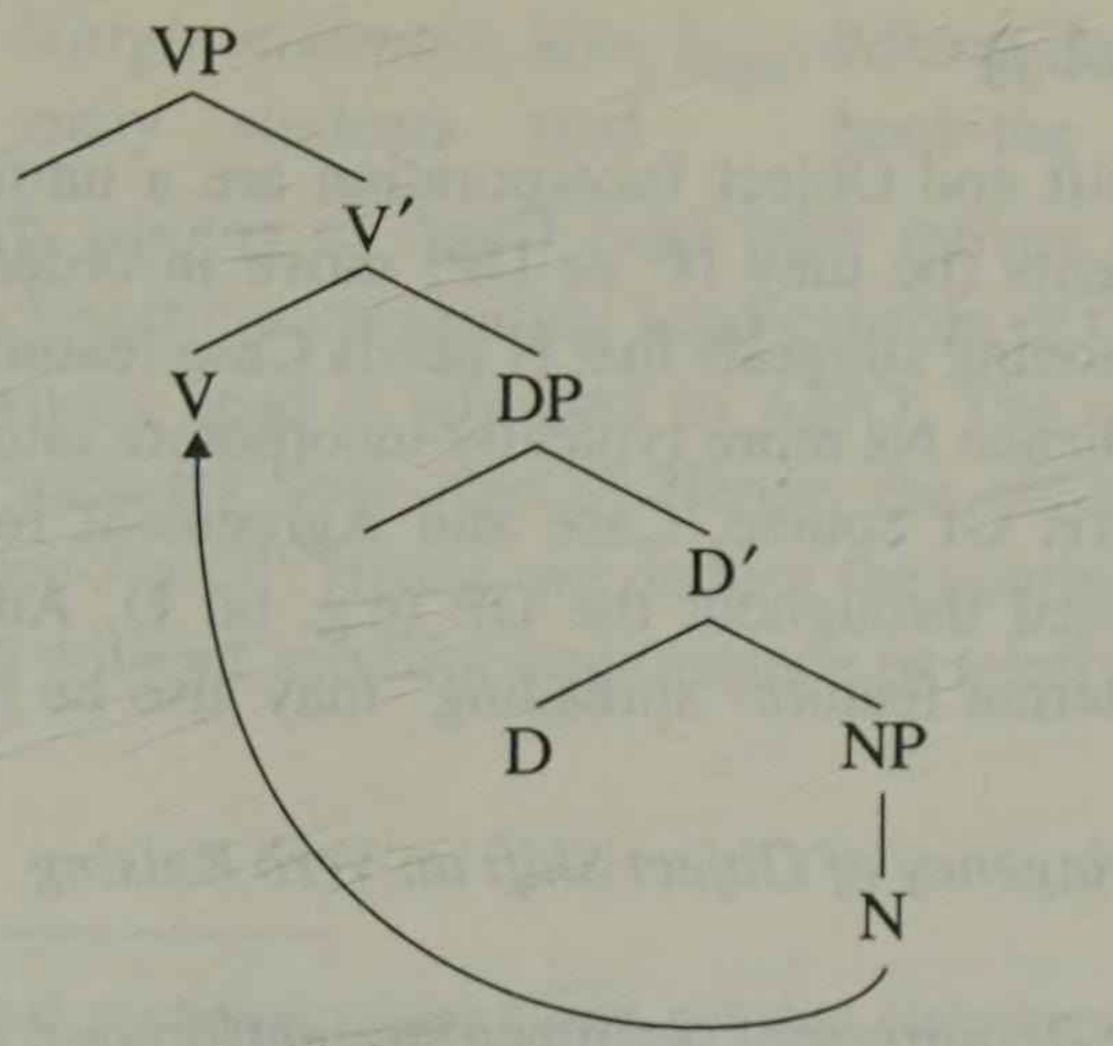
The Shortest Move Requirement also yields an account of a problematic "stranding" case discussed in Baker & Hale (1990) (see also Baker 1988: 92–105). There they examine cases of NI where a noun incorporates into a verb apparently "skipping" over the Determiner as illustrated in (14).

---

is categorially both N^0 and NP without any (non-branching) projection. Note that I also generally abstract away from linear ordering.

10. Here and throughout I give the examples as transliterated in the original references.

(14)



An example is given in (15).

(15) $[_{DP}$ *yede* $t_N]$ *a.**seuan***$_N$*.mũ.ban*
that 2SG-SUBJ/A.**man**.see.PAST
'You saw that man.'
(Southern Tiwa: Allen, Gardiner & Frantz 1984; Baker & Hale's: 7)

Abney (1987) reanalyzes such noun phrases as Determiner Phrases (DPs) headed by a D^0 having a NP complement. Given such an analysis, Baker & Hale note that the N skips the D head to incorporate into the verb, apparently violating Relativized Minimality (i.e., a head undergoing movement may not skip over another head position). In order to permit this violation, Baker & Hale (1990: 292) propose in essence that Relativized Minimality be further relativized between lexical and functional heads, that is, a functional head may skip a lexical head position when moving and a lexical head may skip a functional head position when moving. Since D^0 is a functional head, the lexical head N^0 may skip it and incorporate into V^0. Note that Baker & Hale's "Relativized Relativity" requires an increasingly non-unified *four-way* relativization (A, $\overline{A}$, lexical X^0, functional X^0) over movement and position types.

These cases can be accounted for without "Relativized Relativity". Given the hypothesis that Ns incorporate to check Case features, notice that the first head capable of checking such a feature is the V head — not the intervening D, hence incorporating into V (albeit, over D) meets the Shortest Move Requirement (as does shifting the object over the non-checking domain [albeit A] position; Spec-VP). Note that NI out of PP is still blocked as P typically checks Case (for examples, see Baker 1988: 388–392).

4.2.3. *Object Case Checking*

In sum, Object Shift and Object Incorporation are a unified phenomenon whereby nominal elements (be they N^0 or DP) move in order to check Case. This general line of reasoning suggests that N needs Case features and DP needs Agreement features — hence Ns more typically incorporate into V or P whereas DPs move to Spec-AgrP. Of course, Case and Agreement features are often morphologically expressed throughout the DP (e.g. on D, Adjective, and N), hence additional DP-internal feature-"spreading" may also be required.

4.3. *Deducing the Contingency of Object Shift on Verb Raising*

The Shortest Move Requirement permits a straightforward deduction of part of Holmberg's (1986) Generalization. Regarding the Scandinavian languages, Holmberg finds that object shift is not possible without V raising. Consider the Icelandic example in (16) which shows that object shift is not possible if the verb has not raised (presumably because the Tense position is already filled by the Aux *hafa* or its trace).

(16) *margir stúdentar*$_i$ *hafa* [$_{AgrO}$ *bókina*$_k$ [$_{VP}$ *ekki* [$_{VP}$ t_i *lesið* t_k]]
many students have book-the not read
'Many students have not read the book.'

Suppose the verb *lesið* stays *in situ,* that is, does not adjoin to AgrO. From the object's perspective, the closest c-commanding head capable of checking its Case feature is the verb in its base position.

Recall that UG provides two options for the checking of an object's Case feature. First, a DP object can shift directly to Spec-AgrOP. This option is blocked in (16) as Spec-AgrOP is not in the checking domain of the closest checker, the unraised V. Second, an N^0 object can undergo head movement and incorporate into the verb.

As for (16), productive noun incorporation is not available in Icelandic, hence the object can not incorporate to check its Case. Note that the non-availability of noun incorporation in Icelandic is irrelevant to the SMR. Object incorporation into the unraised verb is the only move which meets the Shortest Move Requirement. Since this kind of incorporation is unavailable, object movement for Case is blocked in the absence of verb raising.

Consider again the "good" object shift case (which has verb raising), repeated from (9) as (17).

(17) *Margir stúdentar*$_i$ *lásu*$_j$ [$_{AgrO}$ *bókina*$_k$ [$_{VP}$ *ekki* [$_{VP}$ t_i t_j t_k]]
many students read book-the not

Suppose the verb adjoins to AgrO (and later moves on). From the object's perspective, the closest c-commanding head capable of Case feature checking is the verb chain whose head is adjoined to AgrO. The checking domain of the verb chain now includes Spec-AgrOP. Hence, the object can undergo full DP object shift to Spec-AgrOP. Hence, we deduce the contingency of object shift on prior verb raising without need for equidistance or relativization over movement types.[11]

---

11. Zwart (personal communication) points out that Holmberg's Generalization does not appear to obtain in Continental West Germanic scrambling. Recalling the deduction of the contingency of object shift on (prior) verb raising, it is clear that demonstrating that Holmberg's Generalization fails to hold requires (as a necessary condition) that at least two points be established:
(i) Scrambling in Dutch and German is indeed movement to Spec-AgrOP.
(ii) Scrambled phrases indeed move to check Case.
Even granting (i), as Vanden Wyngaerd (1989a) and Zwart (1993b) argue for Dutch, it is not clear how tenable (ii) is. First, as Zwart (1993b: Chapter 6, Section 2.2.2.) discusses in detail, scrambled NPs in Dutch do appear to license parasitic gaps, a property more typical of $\overline{A}$-elements than of NPs shifted for Case. Second, as Vikner (1991: 287) notes, PPs may be scrambled in German:
(iii) *Ich habe nicht für das Buch bezahlt.*
(iv) *Ich habe für das Buch nicht bezahlt.*
I have (for the book) not (for the book) paid
By standard assumption, PPs do not require Case, hence this kind of Scrambling must be driven by some other (morphological?) requirement. Furthermore in Icelandic, where movement to Spec-AgrOP is arguably driven by Case checking (i.e. there is no general "scrambling"), movement of a PP to Spec-AgrOP is disallowed.
(v) *Ég keypti ekki bókina.*
(vi) *Ég keypti bókina ekki.*
I bought (the-book) not (the-book)
'I did not buy the book.'
(vii) *Ég borgaði ekki fyrir bókina.*
(viii) **Ég borgaði fyrir bókina ekki.*
I paid (for-the book) not (for-the book)
'I did not pay for the book.'
To conclude, if Scrambling is not driven by Case checking, then the deduction of the contingency of Object Shift on Verb Raising from the SMR is not necessarily disconfirmed by the Dutch and German data. Moreover, my (partial) deduction of

4.4. *Object Checking and Case Theory*

The analysis of object checking presented here eliminates the need for disjunctive versions of the Case Filter and the Visibility Condition. In (18), versions of the Case Filter and the Visibility Condition are given from Chomsky (1986a):

(18) Case Filter: Every phonetically realized NP must be assigned (abstract) Case (Chomsky 1986a: 74)

Visibility Condition: An element is *visible* for θ-marking only if it is assigned Case (Chomsky 1986a: 94)

In (19), disjunctive versions of the same filters are given:

(19) A noun phrase whose head noun is incorporated does not need to receive Case, even though it is phonologically overt [overt NP needs Case OR overt NP needs no case iff N^0 incorporates, K.S.F.] (Baker 1988: 105)

In order for an argument to be visible for θ-role assignment at LF, it must either

a. be assigned Case, OR

b. have its head morphologically united with an X^0

(Baker, Johnson, and Roberts 1989: 239)

...the requirement that NPs be Case-marked is actually an instance of a more general requirement that nominals be associated with a Case feature. This association takes place in one of two ways: either by means of assignment of a feature from a head to the nominal, OR by means of incorporation of the nominal into the head bearing the Case feature

(Rizzi & Roberts 1989: 5)

Under the present analysis both unincorporated and incorporated nominals receive a unified account. A nominal — be it N^0 or DP — checks its Case feature while in the checking domain of an X^0 Case checker. Whether it moves overtly or not depends on the strength of the features checked. Visibility reduces to whether a nominal has its features checked by the interface levels. There is thus no need to invoke disjunctive filters.

Holmberg's Generalization may help to explain why the Dutch and German scrambling appear not to obey Holmberg's Generalization; that is, it may well be the case that such movement is in fact not driven by Case checking.

5. Conclusion

The Shortest Move Requirement permits a unified analysis of object Case checking, without either equidistance or relativization over movement and position types. Hence, (DP) object shift and (N^0) object incorporation are each analyzed as movement of a nominal to a closest c-commanding head capable of checking a needed feature. In this paper, I have neglected a number of important effects which ought to be captured by any Shortest Move account. First, the obligatorily crossed paths of the subject and object as illustrated in (1) are not deduced (see Jonas & Bobaljik 1993: 63–66). Second, given nothing but the SMR, heads might be expected to move freely "over" other heads with no relevant features — to the extent that the Head Movement Constraint holds, this "skipping" must be constrained (see Ouhalla 1989 and Zwart, this volume). Finally the question of $\overline{A}$-movement and its interactions must be addressed. Ferguson & Groat (1995) and Ferguson (1996) investigate in detail some of these additional consequences.

Acknowledgements

This paper is based on Ferguson (1993). Thanks are due Mark C. Baker, Jonathan Bobaljik, Noam Chomsky, Erich M. Groat, Dianne Jonas, Colin Phillips, and Geoff Poole as well as Samuel D. Epstein's Spring 1993 syntax seminar for discussion and access to work-in-progress. Special thanks to Werner Abraham, Naomi Bolotin, and especially Samuel D. Epstein, Steve Peter, Höskuldur Thráinsson, and C. Jan-Wouter Zwart for invaluable comments and discussion on earlier drafts of this paper. Uncited Icelandic judgements are due to Höskuldur Thráinsson. Needless to say, the views expressed here, along with any errors, are my own.

PF interface cond

TNS

Case,

strong features : phonological features
= now are the only ones that move

(problem (gerund) with null operator mvt)

	[= movt	phonol.]
there	×	
PRO	✓	⊘
Peter	✓	✓

moves as well

Spell-Out at the LF Interface
Achieving a Unified Syntactic Computational System in the Minimalist Framework

Erich Groat & John O'Neil
Harvard University

Languages differ with respect to where Spell-Out applies in the course of the derivation to LF. (Are wh*–phrases moved or* in situ*? Is the language French-style with overt verb raising or English-style with LF verb raising?)*

— Chomsky (1993: 24)

Abstract

In Chomsky's (1993) "A Minimalist Program for Linguistic Theory" (MPLT), syntactic derivations yield two generally distinct output phrase-markers that are relevant to the linguistic system: the LF phrase-marker, which receives semantic interpretation, and a phrase-marker which is Spelled-out to the phonological/morphological systems. Furthermore, the rule SPELL-OUT may be applied freely, i.e. at any point in the derivation of an LF phrase-marker, subject to well-formedness conditions indirectly imposed on the phrase-marker by PF-interface conditions (for example, the obligatory non-appearance of uninterpretable strong features) and principles of economy (for example, the global economy principle Procrastinate). However, the system thus formulated introduces small but significant asymmetries into the syntactic computational system: pre-Spell-out operations are stipulated to be more costly than those occurring after Spell-out, strict cyclicity is enforced only before Spell-out, and the lexicon is inaccessible after Spell-out.

Our paper attempts to eliminate these asymmetries by postulating that Spell-out does not occur freely, but rather must always be applied to the same phrase-marker that receives LF-interpretation, i.e. there are no post-Spell-out syntactic operations. The potential for pre-/post-Spell-out asymmetries is eliminated, resulting in a uniform computational system without stipulated (and hence undesirable) asymmetries. We provide further support for our model on a conceptual basis, through an analysis of the strict cycle condition in which we *derive* the Extension Condition for substitution operations from principles of X-bar theory, giving the strong result that there could be no countercyclic substitution at all, even after Spell-out — a result that is not problematic if there are no syntactic operations after Spell-out.

In our system, the difference between overt and covert movement is found on the level of

chain-formation: overt movement is movement which carries phonological features to the head of the chain being created, while covert movement leaves phonological material behind. This approach allows us to construe the Procrastinate principle as a purely local economy constraint. Analyses of subjacency effects in Japanese null operator movement (Watanabe 1992) and Icelandic Object Shift provide empirical support for the analysis.

1. Introduction

In this paper we propose an alternative to the model of grammar developed in Chomsky's (1993) *Minimalist Program for Linguistic Theory* (MPLT). We hope to show that a model in which a single phrase marker serves as the interface to the phonological and interpretive components is a viable model, with a number of conceptual advantages over Chomsky's MPLT approach.

In Chomsky's model, a phrase marker is created and transformed through a series of binary and singularly transformations (MPLT: 22–24). At any point in the derivation, a phrase marker K_i can be directed to the PF component, via the operation Spell-out, which marks the beginning of phonological operations on the phrase-marker. Conditions of the phonological component force phrase-markers to adhere to certain constraints at Spell-out so that they are interpretable by the PF component. For example, one such condition is that no strong features may remain in the phrase-marker at Spell-out, since these features are visible but uninterpretable in the PF component. Phrase-markers that are fed to PF via Spell-out that fail to meet the conditions for phonological operations are said to be illegitimate at PF. After Spell-out, however, the derivation proceeds from the phrase-marker K_i "covertly," yielding a phrase-marker K_j for the interpretive interface, LF. Conditions on legitimacy at LF will again restrict the class of convergent derivations. For example, any unchecked syntactic features (strong or weak) are uninterpretable at LF, so syntactic features must be checked by LF or the representation is illegitimate at LF. Phrase-markers that fail to meet these interface conditions are illegitimate. Derivations creating illegitimate representation at PF or LF are said to "Crash" at those interface levels.

In our model, the derivation proceeds to create a *single* phrase-marker K_f, which is both spelled out to the PF interface, *and* interpreted by the LF interpretive component. Thus, after Spell-out, no further syntactic operations are applied to the phrase-marker. This single phrase-marker must therefore meet all the requirements of *both* the phonological and interpretive components, or it is an illegitimate representation, and the derivation crashes. The difference between "overt" and "covert" movement of a given category is not expressed in the

timing of movement with respect to Spell-out, but rather with respect to where a category is pronounced. Thus, in the simple case of a category forming a single chain in the syntax, what is traditionally called "overt" movement amounts to pronouncing the *head* of the chain, while "covert" movement entails pronouncing the *tail* of the chain. We postulate that chain-formation, always being a reflex of feature-checking, copies phonological material to the head of the chain only when the features to being checked are strong.

The primary motivation for this proposal is conceptual. We mainly seek to eliminate certain asymmetries extant in the pre- and post-Spell-out computational components of the grammar in MPLT, which we argue result in two separate syntactic computational systems. These include (1) the inapplicability of the Strict Cycle condition after Spell-out (MPLT: 22–23; also see the EXTENSION REQUIREMENT below), (2) the lack of access to the lexicon after Spell-out, and (3) the difference in relative cost of otherwise identical operations when performed before vs. after Spell-out (post Spell-out operations being cheaper) — the economy principle known as PROCRASTINATE. By postulating a single interface phrase-marker for both Spell-out and LF interpretation, the very possibility of these asymmetries is eliminated. By eliminating these asymmetries, we argue that a simpler model of greater explanatory power is demonstrated.

As additional motivation for this model, we show that the structure building apparatus of GENERALIZED TRANSFORMATIONS (GT) postulated in MPLT, when examined closely, turns out to enforce strict cyclicity for all substitution operations; it is not necessary to stipulate an Extension Requirement on substitution operations, since the structure-building apparatus of GT itself forces all substitution to be cyclic. This is a welcome result, since it renders the stipulation of an Extension Requirement unnecessary. However, the type of cyclicity we derive is stronger than the version of strict cyclicity imposed in MPLT: strict cyclicity would necessarily apply throughout the derivation, after Spell-out as well as before Spell-out. This contrasts with the MPLT approach, in which the Extension Requirement is stipulated to hold only for overt operations, enforcing cyclicity only before Spell-out. Our derivation of the Extension Requirement, within the framework of MPLT, forces us to conclude that there are no countercyclic substitution operations, even after Spell-out. But MPLT crucially relies on countercyclic substitution operations, such as English object shift to Spec-AgrOP for Case-checking at LF. A solution to this problem comes for free in our model, since *all* transformations occur before reaching the single PF/LF phrase-marker. Thus we may maintain our derivation of strict cyclicity from independent and elementary principles of structure-building already extant in the theory,

eliminating a stipulation advanced in MPLT.

We also show that Procrastinate in our model takes the form of a local, intraderivational economy condition, not a global, transderivational one which compares derivations. In MPLT, Procrastinate stipulates that LF operations are cheaper than identical operations performed before Spell-out. Two convergent derivations that are identical but for the timing of an operation with respect to Spell-out are compared, and the cheaper derivation excludes the more costly one. Procrastinate is thus a transderivational principle. In our single level model, Procrastination-effects are the result of a purely local economy consideration: do phonological features need to be moved or not? The stipulation that a post-Spell-out operation is cheaper than the equivalent pre-Spell-out operation is replaced in our model by the stipulation that it is cheaper not to move phonetic features for any given movement operation; therefore they will not move unless forced to move by considerations of Greed. Along these lines, Procrastinate is reformulated as a purely local, intraderivational economy condition, arguably more consistent with the derivational nature of grammatical principles in the minimalist program.

In the next section, we example the conceptual arguments for this framework, as outlined in this introduction. In section 3, we examine two empirical arguments for this framework: the properties of Japanese *wh*-movement as described by Watamabe (1992), and Icelandic object shift. We then conclude.

2. Conceptual Arguments

In this section we review some important aspects of the approach to a model of grammar advanced in MPLT. We will concentrate on certain stipulations and asymmetries necessary in this model. We then introduce our model and show how these asymmetries and stipulations can be either eliminated or deduced.

2.1. *MPLT Model*

There are three points of concern for us in the Minimalist program: (1) the stipulation of an Extension Requirement which constrains only overt substitution operations, (2) three pre-/post-Spell-out asymmetries of the model: a) the Extension Requirement, b) access to the lexicon, and c) Economy (Procrastinate), and (3) the status of Procrastinate as a transderivational economy principle.

2.1.1. *The Extension Requirement: Stipulating the Strict Cycle Condition*

MPLT proposes the following formalization of the Strict Cycle Condition, which we refer to here as the "Extension Requirement" (MPLT: 22–23):

> (1) Suppose we restrict substitution operations still further, requiring that ∅ be *external* to the targeted phrase-marker K. Thus GT [Generalized Transformations] and Move α extend K to K*, which includes K as a proper part. For example, we can target K = V′, add ∅ to form $[_{\beta}$ ∅ V′], and then either raise α from within V′ to replace ∅ or insert another phrase marker K^1 for ∅. In either case, the result must satisfy X-bar theory, which means that the element replacing ∅ must be a maximal projection YP, the specifier of the new phrase-marker V″ = β.

This stipulation rules out countercyclic substitution operations, which would invariably involve creating a phrase-marker that did not include the original phrase-marker as a "proper part" (as defined in the above quote). Note, on the other hand, that head-movement, which adjoins a head X to a head Y, can never result in extending the phrase marker. For example, an Agreement phrase before adjunction of V to Agr might have the form $[_{Agr'}$ Agr $[_{VP}$ ... V...]]; adjunction of V to Agr would result in the structure $[_{Agr'}$ $[_{Agr}$ V Agr] $[_{VP}$ ... t_V...]]. The phrase-marker has not been "extended"; it does not include the old phrase marker as a "proper part".

With respect to the adjunction of a phrase to a phrase, which requires a clear definition of the notion "include" and "proper part" (see Poole 1994), the question arises whether these too might be subject to the Extension Requirement. MPLT assumes that countercyclic adjunction of phrases is possible (p. 24). The Extension Requirement is therefore defined to apply only to substitution operations. Additionally, covert substitution operations such as English object shift to Spec-AgrOP (necessarily occurring after movement of the subject to Spec-AgrSP, which c-commands the object's landing site Spec-AgrOP) suggest that covert operations must not be subject to the Extension Requirement. Chomsky therefore also exempts covert operations from the condition: "The extension requirement holds only for substitution in overt syntax." (MPLT: 24)

Thus we see that the Extension Requirement as formulated provides the desired result that all overt substitution is cyclic, but at the cost of a number of stipulations: (1) The Extension Requirement itself must be stipulated; (2) Adjunction operations must be stipulated to be exempt from the condition, and (3) the condition is stipulated not to hold covertly, i.e. after Spell-out. This

issues have been addressed at length in Kitahara (1993, 1994, 1995), who proposes a means to derive Strict Cyclicity by means of Economy considerations. In Section 2.2 we will show that it is possible to derive these stipulations from structure-building considerations under a single-level model.

2.1.2. *Asymmetries of MPLT Pre-/Post Spell-out*

There exist a number of asymmetries between the pre- and post-Spell-out components of the syntax in the minimalist model. The first we have already seen in Section 2.1.1, namely that the Extension Requirement on substitution operations does not hold after Spell-out.

A second asymmetry involves the inaccessibility of the lexicon after Spell-out. Chomsky motivates the stipulated lack of access to the lexicon in LF and in the phonological component in a natural way: addition of lexical items to the representation in LF would result in a derivation whose interpretation involved lexical items that did not enter the phonology at all, even if they have overt phonological form: "...we must ensure, for example, that *John left* does not mean 'They wondered whether John left before finishing his work'..." (MPLT: 22). A parallel argument holds for banning access to the lexicon in the phonological component: strings of words receiving PF but not LF interpretation must not be permitted (though see Marantz & Halle [1993] for arguments in favor of lexical access after Spell-out). Thus we have a second asymmetry in the syntactic computational system: pre-Spell-out computation allows access to the lexicon, while post-Spell-out computation in the LF component disallows it.

The third asymmetry involves the question of Economy. A given operation of Form–chain incurs a different cost depending on whether it occurs before or after Spell-out: pre-Spell-out operations are stipulated to be more costly than post-Spell-out operations. "The system tries to reach PF 'as fast as possible,' minimizing overt syntax" (MPLT: 30–31). Thus English, which does not require overt main verb raising to check strong inflectional features, disallows overt main verb raising in general, since by Economy LF raising is cheaper, and cheaper derivations are preferred over more costly ones. French, on the other hand, requires overt main verb raising for convergence, since French is hypothesized to have strong verbal Agr/Tense features that are uninterpretable at the PF interface and must be checked and deleted before Spell-out (MPLT: 30). Thus the more costly overt operation is allowed for the sake of convergence. This economy principle is known as Procrastinate in MPLT, since it requires operations to be performed after Spell-out if possible.

To summarize, then, three asymmetries exist between overt and covert syntax: (1) the Extension Requirement holds only before Spell-out; (2) access to the lexicon is barred after Spell-out, and (3) pre-Spell-out operations are more costly than identical post-Spell-out operations under the Economy metric.

Why should we want to eliminate these asymmetries? After all, it is clear that not all components of the grammar have the same computational character: the phonological component involves a variety of operations not found in the syntax (for example the establishment of phonological and intonational phrases, phonological rules operating on and between words, etc.). However, it is clear that the computational character of the PF component is radically different from that of the syntax, as is the computational character of the semantic (post-LF) component. The differences are not unexpected. On the other hand, pre- and post-Spell-out operations are exactly the same: both operate on phrase-markers through binary and singularly transformations (and perhaps deletion), both create chains through the operation FORM-CHAIN; both are constrained by the Greed principle. To the extent that the asymmetries between them must be stipulated and are not derived from independent principles, we effectively end up with *two* computational systems that are nearly identical. A model in which we had only one uniform computational system would therefore be a simplification of this model — as long as no additional stipulations needed to be made.

2.1.3. *Procrastinate as a Global Principle*

The third issue concerns again the question of the Economy principle Procrastinate. As it is formulated in MPLT, the principle is transderivational in nature: convergent derivations that result in the same LF representation are compared, and the cheapest derivation blocks other derivations. Thus, English derivations involving overt and covert Verb Raising are compared, both resulting in legitimate LF and PF representations. The cheaper derivation will always be that in which Verb Raising occurs after Spell-out; thus overt Verb Raising is always blocked.[1]

---

1. We will not address the question of the overt raising of auxiliary verbs such as *have* and *be* here. MPLT (p. 31) contains the suggestion that auxiliaries are "semantically vacuous," and are invisible to LF rules; thus if they are to raise they must raise overtly, otherwise they cannot raise and the derivation will crash with unchecked features at LF. (See also Epstein [1993] for discussion.) Note, however, that such an approach introduces yet another asymmetry in the syntactic computational system; some

However, one of the conceptual guidelines of the minimalist framework is the notion that computation takes place *locally* and *derivationally*. For example, the notion of Minimal Domain is (crucially) defined for a given category within a given phrase marker at a given point in a derivation. (See Chomsky: 1993: 12; see also Jonas & Bobaljik 1993 and Ferguson & Groat 1995 for extensive work on the derivational nature of domain extension.) Greed, the principle that a category moves only to check its own features and not for the sake of the legitimacy of other categories in the representation, also restricts the possibilities for movement of a category on a completely local scale: does a particular instance of Form-chain, at a particular point in the derivation, contribute to the legitimacy of a category or not?

With respect to the definition of a category's domain and the Greed principle, the computational determinacy of the system is established derivationally. Thus, English Verb Raising to a C^0 which lacks verb features would be a violation of Greed, even though it is hard to see why such movement would result in an illegitimate representation (as we see V-to-C raising in many Germanic languages, for example, presumably motivated by some checking requirement absent in English). But with respect to Procrastinate, no principle applies *within* the derivation to determine whether a movement is overt or covert: overt main Verb Raising in English is not blocked within the course of a derivation, rather the whole derivation is blocked by comparing the derivation with the cheaper derivation which is identical but involves covert Verb Raising.

While transderivational economy conditions are not in themselves problematic, they are inconsistent with the minimalist philosophy that the computational system has an *intraderivational* nature. We will show in Section 2.2 that Procrastinate can be recast as a consequence of Greed considerations, so that whether an operation results in overt or covert movement is forced by local considerations of Greed and Economy, rather than by a transderivational economy metric that compares derivations.

---

categories are visible both before and after Spell-out, while others are visible only before Spell-out. It is also unclear whether auxiliaries are truly "vacuous" and act as pure "place-holders", since they result in feature checking of Tense, hardly a semantically vacuous set of features.

2.2. *The Single Level Model*

We have seen that the MPLT model of the grammar introduces (1) certain stipulations concerning structure-building to instantiate a Strict-Cycle Condition, (2) three asymmetries between the otherwise identical pre- and post-Spell-out syntactic computational systems, and (3) a transderivational economy metric required for the Procrastinate principle.

We shall now propose a different model of the grammar, in which there is no syntax after Spell-out: the same phrase-marker that is "spelled-out" is fed to the LF interface. We argue that minimal considerations of structure-building (outlined in Section 2.2.1) allow us to derive the Extension Requirement on substitution, but in a way which would not allow countercyclic substitution after Spell-out. We call this condition the "Strong Cycle". Following this idea, we propose that there are no post-Spell-out syntactic operations at all, and suggest a means of determining the position where categories are pronounced that does not rely on the timing of Spell-out within the derivation. This model trivially eliminates the asymmetries we discussed in Section 2.1.2, resulting in a single, uniform syntactic computational system. Finally, a natural consequence of this model results in a principle of Procrastinate that is completely local in character and does not compare convergent derivations.

2.2.1. *Structure-Building Considerations Enforce "Strong Cycle"*

First let us review the way phrase-markers are constructed in MPLT. The mechanism used is the Generalized Transformation (GT), which is either binary (inserting one phrase-marker into another) or singularly (moving, or more precisely copying, a category and inserting it elsewhere within the same phrase-marker). GT works as follows (from Chomsky 1993):

(2) a. A category X is targeted and projected in a manner consistent with X-bar theory;
b. a sister ∅ to X is created;
c. a category is inserted into ∅ (a separate phrase-marker in the case of binary GT, a subtree of the phrase-marker in the case of singularly GT) which yields an X-bar consistent structure.

The result of this operation is the creation of a new phrase marker, which in the case of overt substitution must properly include the target X, the root node of the previous phrase-marker (by the Extension Requirement). Overt substitution must always target the root node of the phrase-marker.

It is also possible to target a category X, project it, and *not* introduce a ∅ into which some category is inserted as sister of X. For example, consider an Agr′, whose Specifier need not yet be filled. This Agr′ is to be inserted as complement of T^0. To maintain an X-bar-consistent structure, Agr′ must first undergo non-branching projection to AgrP before insertion into a projection of T. Strictly, steps (2b) and (2c) constitute an actual operation GT, while (2a) may be performed without the insertion of a category by steps (2b) and (2c). (See again Kitahara 1993, 1994, 1995 for discussion.)

Putting adjunction briefly aside, let us see what happens when the targeted category is not the root node of the phrase-marker. Consider the case of an $AgrOP_1$ which is a complement to T, as in (3).

(3) $[_{T'}\ T^0\ [_{AgrOP1}\ [_{AgrO'}\ AgrO^0\ [_{VP}\ \ldots\ ]]]]$

Now assume that a direct object is to be moved to Spec-AgrOP for Case-checking at LF where the Extension Requirement does not hold, by stipulation. This would be accomplished by targeting AgrO′, projecting a new $AgrOP_2$, creating a sister ∅ to AgrO′, and inserting the direct object DP into ∅. The result would be the structure in (4):

(4) $[_{AgrOP2}\ DP\ [_{AgrO'}\ AgrO^0\ [_{VP}\ \ldots\ ]]]$

But what relation does this new $AgrOP_2$ have with the rest of the phrase-marker? It is not identical with the $AgrOP_1$ that originally immediately dominated AgrO′ in (3), nor could it be construed as a segment of $AgrOP_1$, since we are assuming substitution, not adjunction. But this means that AgrO′ is immediately dominated by both $AgrOP_1$, its original dominator, and the newly created $AgrOP_2$. This is clearly not an X-bar consistent structure; in fact it does not even fall under the definition of a tree structure (among other things, there is more than one node which is not dominated by any other node). Thus using GT to countercyclicly substitute a phrase into a Specifier is not possible, since again by definition the result must be X-bar consistent.

It would not be difficult to rectify this situation by stipulating that if a category X^1 that is already immediately dominated by some category Y is projected to form X^2, then Y immediately dominates X^2 instead of X^1. The resulting structure, following countercyclic substitution in (4), is then as in (5):

(5) $[_{T'}\ T^0\ [_{AgrOP1}\ [_{AgrOP2}\ DP\ [_{AgrO'}\ AgrO^0\ [_{VP}\ \ldots\ ]]]]]$

This structure is consistent with X-bar theory. However, this would require an additional stipulation concerning structure-building.

Note that the problem does not arise in the case of countercyclic adjunc-

tion, since the targeted category X projects a segment of X; immediate dominance relations with the *category* X remain undisturbed. For instance, assume that an AdvP is adjoined to AgrOP in (3). The output structure is given in (6).

(6) $[_{T'}\ T^0\ [_{AgrOP1b}\ AdvP\ [_{AgrOP1a}\ [_{AgrO'}\ AgrO^0\ [_{VP}\ \ldots\ ]]]]]$

T′ immediately dominates the *segment* AgrOP1b (but not AgrOP1a), and AgrOP1b immediately dominates both AdvP and the *segment* AgrOP1a. However, since AgrOP1a and AgrOP1b are segments of the same category, the immediate domination relation *with respect to that category* is the desired one: T′ (category) immediately dominates the *category* AgrOP1. Thus it is in fact only countercyclic substitution that requires additional stipulation in order for GT to allow the operation.

Instead of stipulating a way out of the problem of countercyclic substitution, let us assume that it is in fact disallowed by considerations of structure-building assumed in MPLT. We then derive what we call the STRONG CYCLE, in which countercyclic substitution is impossible. This comes very close to the Extension Requirement, which stipulates the impossibility of overt substitution that does not target the root node (since the phrase-marker would not be "extended," failing to properly contain the original phrase-marker). But there is an important difference: the Strong Cycle bans *all* countercyclic substitution, not just overt substitution. Phenomena like LF object shift, or covert *wh*-movement in *wh-in situ* languages would be a problem under the Strong Cycle.

Still, the fact that we can derive a version of the Strict Cycle condition from independent considerations of structure building — instead of stipulating an Extension Requirement, as well as some sort of patchwork for countercyclic LF substitution — seems to be reason enough to pursue this approach. In the next section we will introduce our way of handling the problem of covert substitution.

2.2.2. *Introduction to the Single-Level Model*

In our model, a derivation proceeds until all features, weak and strong, have been checked, yielding a single "final" phrase-marker K_f, which is the object of the interpretive mechanism and of the phonological component. In other words, Spell-out and LF interpretation take the same K_f as their input. All syntactic operations have taken place before interpretation and before PF; there are no post-Spell-out syntactic operations.

This means, however, that all movement, including "covert" object shift or Verb Raising in English, have already taken place before Spell-out. How is it, then, that English verbs and direct objects are pronounced *in situ*? The answer

lies in a reformulation of the definition of strong vs. weak features and the nature of chain-formation. In our model, when a category is moved, a copy of the category is left as its trace, but unlike in the MPLT approach, this copy is not automatically marked (either in the syntactic component or in the PF component) as phonologically null.[2] Rather, we postulate the following two principles:

(7) a. Strong features may be checked only in a checking relation with a node specified for phonological features.

b. Moving Phonological features to the head of a chain is more costly than leaving them in the tail of the chain.

These two stipulations replace two similar ones made in MPLT: first, that strong unchecked features are visible PF objects and are uninterpretable (forcing pre-Spell-out checking and deletion of strong features), and second, that pre-Spell-out movement is more costly than post-Spell-out movement.

The stipulations in (7) appear at least as natural as those they are intended to replace. Strong features are seen as akin to affixes; they require some phonological material as a host in order to be licensed. Weak features lack this character and are licensed purely by virtue of a syntactic relation with the category, without the need for a phonological host. The idea that copying phonological features of a category is more costly than not doing so also seems natural; fewer features need be copied in the latter case. While MPLT must

---

2. It is not clear by precisely what mechanism copies of categories are to be marked as phonologically null. One possibility is the pure stipulation that Form-chain automatically marks the tail of the chain as unpronounced, in which case our reformulation simply specifies in greater detail which end of the chain ends up with phonological features.

In contrast, on pp. 48–50, Chomsky seems to suggest that the PF component may be responsible for marking copies of categories as phonologically null through a process of "deletion through identity" such as that found in VP-deletion contexts; i.e. if a category has been pronounced already, don't pronounce it again. A complication arises for this approach in the case of X^0-movement in a head-final language: raising of X^0 would be leftwards, so the tail of the X^0-chain would *follow* the head. Would the *head* of this X^0-chain then be pronounced — being the first copy with respect to linear order — instead of the tail, as in a head-initial language? This is not implausible, but it shows that much more needs to be said within the MPLT framework about the question of where a category is to be pronounced given the choice of a number of copies of a category distributed throughout various chains. In answering this question by enriching the mechanism Form-chain, we are not in fact enriching the theory, which must specify *somewhere* where a category is to be pronounced.

stipulate some means to ensure that the tail of a chain is marked in some way as not pronounced (perhaps in the PF component), we instead enrichen the principle Form-chain with respect to the COPY THEORY: forming a chain results in copying all syntactic features of the category moved, but does not copy the category's phonological matrix: it either moves it to the new position or fails to move it. This enrichment of Form-chain replaces whatever mechanism is stipulated to mark the tail of a chain as phonologically null.[3]

Consider the case of English Verb Raising. By hypothesis, Agr and T do not bear strong verb features. Under a minimalist account, this would mean that a main verb would not raise to check Tense and Agreement until after Spell-out, by Procrastinate. Under our account, the verb does in fact raise: first to AgrO, which then raises to T and AgrS,[4] but the phonological material is not copied, since the features are weak. In fact, the phonological material *must* not be copied, because for each move for feature checking, the most economical will be the move that does not copy the phonological features of the verb. Thus, the verb chain to AgrO is extant in the phrase-marker that is spelled out, but the tail of the chain is pronounced, in the verb's base position.

In contrast, a French verb will move its phonological material to T and

---

3. A question arises concerning the raising of phonologically null categories, such as PRO or null operators. There are two possibilities, neither of which we will explore in detail here. First, it may be the case that these categories simply cannot check strong features, since they have no phonological content. In this case, non-finite clauses with PRO subjects must not have strong subject Case features; i.e. the Null Case features of T in non-finite clauses must be weak. This might be a defining characteristic of non-finte T, with weak N-features being associated with Null Case. Similarly, constructions involving null-operator movement cannot involve strong *wh*-features in the CP to which the null-operator moves, since these features could not be checked by a phonologically null element; again, this might be a defining characteristic of CPs in which null operators are permitted.

   A second possibility is that PRO and null operators *do* have a phonological matrix associated with them that allows them to check strong features, but that these features are "weak" in a similar sense to the weak character of some features on functional categories in MPLT: they are *invisible* to the PF component, just as weak features were stipulated to be invisible to the PF component.

4. For concreteness we assume the analysis in MPLT (pp. 29–30) and previous work cited there that T adjoins to AgrS before V-raising; the verb adjoins to the complex head $[_{AgrS}$ T AgrS]. The point is that our account of this analysis differs only in its theoretical implementation.

AgrS, since these categories bear strong V-features in that language. The option to leave the phonological material *in situ* does not exist, since the strong features would not be checked without the phonological material, and chain formation without phonological feature movement would violate Greed.

We see then that the effects of Procrastinate may be gained through a completely local economy principle combined with Greed. At any point in a derivation where a chain is formed, whether phonological material is transferred or not is determined locally, at that point in the derivation: they must be copied if the features to be checked are strong, and must not be if they are weak, since moving them would be more costly at that point in the derivation. No trans-derivational comparison of convergent derivations is needed.

2.2.3. *Elimination of Asymmetries*

This model automatically eliminates the three pre- vs. post-Spell-out asymmetries that were enumerated in Section 2.1.2 above.

The first asymmetry concerned the lack of an Extension Requirement on covert substitution. By deriving the Strong Cycle from structure-building considerations, we can do without a stipulated Extension Requirement, at the cost of disallowing countercyclic substitution at any point in the derivation, including after Spell-out. This lead us to a model in which there are no post-Spell-out operations. But if there are no post-Spell-out operations, no asymmetries between pre- and post-Spell-out operations exist. Similarly, the question of lack of access to the lexicon after Spell-out need not be stipulated, since there is no syntax after Spell-out. Though in this case the stipulated asymmetry seemed natural, the otherwise identical character of pre- and post-Spell-out computation suggests that this asymmetry should be derivable from independent principles; under our model the asymmetry of lexical access does not arise at all. Finally, the asymmetry of the economy of identical operations before and after Spell-out (Procrastinate) is also eliminated: all identical operations (whether overt or covert) incur the same cost, if any, regardless of when they occur. Under our system, checking strong and weak features are not identical operations, since checking the former requires copying phonological features of a category, and the latter does not; the relative costs are determined by the number of features that are moved.

We thus arrive at a uniform computational system for syntactic operations that functions in the same way across the entire derivation.

3. Empirical Consequences

In this section we will look at some empirical evidence that supports our model, and examine some interesting predictions it makes concerning object shift. Section 3.1 concerns *wh*-movement in Japanese, where we take the evidence of Watanabe (1992) as an argument for overt *wh*-movement which does not involve copying of phonological features. In Section 3.2 we consider Icelandic object shift, and argue that differing chain structures formed in the process of checking strong and weak verb features result in different complement domains for the verb.

3.1. *Japanese:* wh-*movement and Subjacency*

In this section we consider the work of Watanabe (1992) and consider how we might recast his analysis in our framework.

To explain examples of apparent LF subjacency effects, Watanabe (1992) argues that Japanese has *wh*-movement *before* S-Structure of a "pure *wh*-operator" which is base-generated in Spec-DP. (We will not cover all of the data discussed in Watanabe (1992), but rather pick out the simplest examples that suggest that his analysis can be recast within our theory.) This operator is phonologically null, thus no moved category is actually observed. Since this movement occurs before S-structure, it is sensitive to subjacency effects, as Watanabe (1992) demonstrates in great detail. As in English, only one phrase is moved to Spec-CP; others remain *in situ* and are licensed by an LF process of absorption. At LF, the DP out of which the null operator moved adjoins to the operator in Spec-CP.

We propose that what Watanabe considers to be null-operator movement from Spec-DP can be analyzed instead as full-phrasal movement of that DP. Japanese, we argue, has weak *wh*-features in C^0, not strong features as found in English, so *wh*-movement does not result in copying of phonological features to Spec-CP. Thus, the full *wh*-phrase moves — and subjacency effects are expected — but only the tail of the *wh*-chain is pronounced, giving the surface appearance of *wh-in situ.*

Consider the *wh*-island violations in (8) and their Japanese counterparts in (9). (The Japanese sentences and judgements, here and throughout, are from Watanabe 1992.)

(8) a. **What do you remember where we bought?*
(= (1a) in Watanabe 1992)

b. **Who do you like books that criticize?* (= (2a))

(9) a. [$_{CP}$ **Op** *John.wa* [$_{DP}$ $\mathbf{t_{Op}}$ **nani.o** *katta hito*].*o*
TOP what.ACC bought person.ACC
sagasite iru no]? (= (5a))
look.for Q
'What is John looking for the person who bought?'

b. ?[$_{CP}$ **Op** *John.wa* [$_{DP}$ t_{Op} *Mary.ga* ***nani.o*** *katta ka dooka*]
TOP NOM what.ACC bought whether
sirigatte iru no]? (= (5b))
want.know Q
'What does John want to know whether Mary bought?'

The illicit sentences in (8) are examples of extraction from a *wh*-island (8a) and from a relative clause (8b) in English. In each case two bounding nodes are crossed: two IPs in (8), and an IP and an NP in (8b). In contrast, the Japanese equivalent to (8b) is (9a), which is grammatical. It appears at first glance that Japanese *wh*-movement, not being overt, is not susceptible to Subjacency.

However, as Watanabe points out, there is evidence that in (9a–b), it is not the *wh*-phrase *nani-o* that moves to Spec-CP at LF, but the entire complex DP [$_{DP}$ *nani-o katta hito*]-*o,* which is pied-piped to Spec-CP at LF. Thus, the *overt* movement of a null operator is from the Specifier of this complex DP to Spec-CP, not from the Specifier of *nani-o.* Evidence for this is found in examples (10a–c):

(10) a. *[$_{CP}$ **Op** *Mary.wa* [$_{DP}$ *John.ni* [*ittai* [$_{DP}$ $\mathbf{t_{Op}}$ ***nani.o***]] *ageta*
TOP DAT the.hell what.ACC gave
hito].*ni* *atta no?*]
person.DAT met Q
'What the hell did Mary meet the person who gave to John?'
(= Watanabe's (6a))

b. [$_{CP}$ **Op** *Mary.wa* [$_{DP}$ *John.ni* [$_{DP}$ $\mathbf{t_{Op}}$ *nani.o* *ageta*
TOP DAT what.ACC gave
hito].*o* *atta no*? (= (6b))
person.DAT met Q
'What did Mary meet the person who gave to John?'

c. [$_{CP}$ **Op** *Mary.wa* [*ittai* [$_{DP}$ $\mathbf{t_{Op}}$ [*John.ni* *nani.o* *ageta*
TOP the.hell DAT what.ACC gave
hito]]].*ni* *atta no?*]
person.DAT met Q
'What did Mary meet the person who gave to John the hell?'

Following Pesetsky (1987), Watanabe (1992) assumes an analysis of the word *ittai,* roughly translated as "the hell," as a particle that attaches to the left of a phrase undergoing LF *wh*-movement.[5] If *ittai* is attached to the left of *nani-o,* as in (10a), then it is *nani-o* that must raise at LF; thus the *overtly* moved null operator originates in the Specifier of *nani-o* as shown. Thus it appears that the null operator must cross *three* bounding nodes in (10a): the DP *nani-o* to which *ittai* is attached, the complex DP, and the IP. In (10b) and (10c), on the other hand, the null operator crosses only two bounding nodes: the complex DP (to which *ittai* is adjoined in (10c)), and IP. The former results in ungrammaticality, but not the latter, though it appears that the null operator is crossing one bounding node too many for the correct prediction.

To reduce the number of bounding nodes the null operator must cross in each case, Watanabe (1992: 54–58) argues that the null operator may adjoin to the DP from which it originates. The argument goes as follows: the traces of the null operator–variable chain formed by overt movement can be deleted at LF, since the chain is replaced by the operator–variable chain of the entire DP adjoining the operator in Spec-CP at LF. The null-operator chain itself plays no semantic role, and may be deleted. For example, in (10c) (from Pesetsky 1987), the null operator in Spec-CP is adjoined to by the DP from which it originated, i.e. the complex DP, so all traces of the null operator included by the complex DP may be deleted. Thus, since LF interpretation does not find any null operator trace adjoined to the argument DP, Chomsky's (1986b) ban on adjunction to arguments is circumvented, and the overt adjunction does not interfere at LF with semantic interpretation of the DP.[6] But by adjoining to the DP from whose Spec-the null operator originates, the null operator need cross only one bounding node at a time.

Thus, in (10b) and (10c), the null operator may adjoin to the complex DP, crossing one bounding node, and then continue raising to Spec-CP, again crossing only one bounding node, IP. The lack of a subjacency effect is correctly predicted. However, in (10a), the position of *ittai* shows that only

---

5. One potential problem with this account is the exact status of *ittai,* which we cannot presume to be adjoined to DP, since this would be in violation of the ban on adjunction to arguments he assumes following Chomsky (1986b). Where it is "attached" we presently leave unanswered.

6. Watanabe (1992) in fact extends the ban on adjunction to arguments to a more general ban on adjunction to *any* category with a semantic role, including adjuncts (though not predicates).

nani-o raises at LF; thus the null-operator originates in the Spec-of the DP *nani-o*. The null operator may adjoin to this DP, since this DP raises at LF to adjoin to then null operator in Spec-CP. However, it still must cross two bounding nodes: the complex DP, and the matrix IP. Thus a strong subjacency effect is predicted for (10a).[7]

Under our analysis, the situation is more straightforward. Consider sentence (9a). Following Watanabe (1992), we assume that pied-piping of complex DPs is possible in Japanese. All we need to say is that the entire phrase is pied-piped to Spec-CP before Spell-out, but that Japanese has *weak wh*-features, so no phonological features are moved. The movement itself is constrained by subjacency — as when we force movement of the *wh*-phrase inside the complex DP by prefixing *ittai* in example (9c) — but it is not visible simply because no phonological features are moved. No movement of null operators (out of DPs

---

7. A potential problem exists for the analysis predicted by Watanabe (1992) for (10b–c). His argument is that the operator–variable chain formed by null-operator movement is "replaced" by the operator chain formed by LF movement of the DP from which the null operator originated through a process of deletion. Now, the full S-structure for (10a) would be as in (i):

(i) *[$_{CP}$ **Op** *Mary-wa* [$_{DP}$ *John-ni* [$_{DP}$ **t′$_{Op}$** [*ittai* [$_{DP}$ **t$_{Op}$ nani-o**]]] *ageta hito*]-*ni atta no?*

LF raising of *ittai nani-o* would result in the deletion of the traces of the null operator (though the sentence is still ungrammatical, due to the Subjacency violation in null operator raising from t′$_{Op}$ to Spec CP). Note that the traces that must be deleted are not just the variable trace, but crucially also the adjoined trace, which would otherwise interfere with LF theta-role assignment (see also the previous note). Thus all null operator traces must be deleted.

However, if this is the case, then why could we not avoid the subjacency violation in (i)/(10a) by adjoining the null operator to the complex DP, as in (ii)?

(ii) *[$_{CP}$ **Op** *Mary-wa* [$_{DP}$ **t″$_{Op}$** [$_{DP}$ *John-ni* [$_{DP}$ **t′$_{Op}$** [*ittai* [$_{DP}$ **t$_{Op}$ nani-o**]]] *ageta hito*]]-*ni atta no?*

Each link of the null-operator chain crosses only one bounding node in (ii). If all traces of the null operator are deleted at LF, since the chain is semantically "replaced" by the chain formed by *nani-o* at LF, then no adjoined trace on the complex DP *John-ni ittai nani-o ageta hito-ni* will interfere with theta role assignment of the complex DP. An additional stipulation is needed guaranteeing that t″$_{Op}$ is not deleted at LF — something of the form, "only those traces of the null operator that are included in the phrase adjoining to it at LF are deleted" — to rule out this possible derivation. This complicates the analysis considerably.

which later adjoin to them at LF) need be hypothesized. Thus the structure at Spell-out under our model of (9a), for example, is as in (11):

(11)

[$_{CP}$ [$_{DP}$ *nani.o katta hito*].*o* [$_{IP}$ *John.wa* [$_{DP}$ *nani.o katta hito*].*o sagasite iru no*]]

The complex DP *nani-o katta hito-o* is moved to Spec-CP, but its phonological features are left behind. The underlined copy is the pronounced copy.

This analysis carries over straightforwardly to other cases of overt *wh*-movement as analyzed in Watanabe (1992). In each case under our analysis, it is not a null operator that moves from Spec-DP, but the entire *wh*-phrase DP, leaving a copy with phonological features behind. Thus the parametric difference between English and Japanese lies in the feature strength of *wh*-features borne by C: in English these features are strong, and we see *wh*-movement that carries the phonological features; in Japanese they are weak, so we do not see the movement directly, but rather indirectly, through Subjacency effects.

3.2. *Object Shift*

In this section we make a tentative suggestion concerning object shift in Icelandic. We argue that the conditions under which objects appear to shift to their Case position in Spec-AgrOP are in fact the conditions for base-generation of direct objects in Spec-AgrOP. We contrast the case of Icelandic with that of English, to show that the option for base-generation of objects in Spec-AgrOP does not exist in English, as a natural consequence of the different verb chain structures formed by the need to check strong vs. weak V-features.

3.2.1. *Icelandic: Optional Base-generation of Object in Spec-AgrOP*

Consider the following sentences from Holmberg (1986), with bracketing added:

(12) a. *Jón keypti ekki* [$_{VP}$ t_V ***bókina***].
John bought not book.the
'John did not buy the book.'

b. *Jón keypti* [$_{AgrOP}$ ***bókina*** *ekki* [$_{VP}$ t_V]].
John bought book.the not
'John did not buy the book.'

c. *Jón hefur ekki* [VP *keypt* ***bókina***].
John has not bought book.the
'John has not bought the book.'

d. **Jón hefur* [$_{AgrOP}$ ***bókina*** *ekki* [$_{VP}$ *keypt*]].

Let us assume, following Holmberg (1986), that negation in Icelandic is adjoined to VP. In (12a), then, we have raising of the verb out of VP to T/AgrS where it checks its inflectional features, while the object apparently remains *in situ* in VP. In (12b), however, we find the object to the left of negation, suggesting (again following Jonas & Bobaljik 1993) that it is in its Case-position, Spec-AgrOP. Under our analysis, the object must in fact have shifted to Spec-AgrOP in (12a) as well as (12b), since all movement occurs before Spell-out in our model. But we do not see the object there. Under our model, we therefore assume that AgrO bears weak N-features in (12a), but strong N-features in AgrO in (12b). The movement of the object to Spec-AgrOP is thus visible in (12b), since it carries the phonological features to Spec-AgrOP to check these strong N-features, while the movement is not visible in (12a), since the phonological material is left *in situ.* Thus, if the optional appearance of the direct object in Spec-AgrOP is the result of optionally moving the direct object from within VP, we must conclude that there are two types of AgrO in Icelandic: one with strong N-features, one with weak N-features.[8]

Now observe (12c–d). In (12c), we find the main verb appearing within the VP. An auxiliary verb has moved to T/AgrS to check inflectional features; the main verb is not so inflected and thus must not raise to T/AgrS, by Greed. But in this situation, we see that the object may not appear in Spec-AgrOP, as shown by the ungrammaticality of (12d). The four examples in (12) constitute a classic example of Holmberg's Generalization that overt Object Shift is contingent on Verb Raising to T/AgrS.

Given our account of the optionality of object shift in (12a–b), in which AgrO's N-features were hypothesized to be optionally strong, we would appear to have a problem: what would preclude using an AgrO with *strong* N-features in (12d), resulting in Case-chain formation of the object to Spec-AgrOP which

---

8. Note that this is true also of an account of optional Object shift in a strictly minimalist model; Procrastinate would prohibit shift of the object to Spec-AgrO before Spell-out unless AgrO bore strong N-features; thus optional movement must imply optionally strong features.

copies the phonological material to check these strong features, incorrectly yielding (12d)?

Jonas & Bobaljik (1993) address this issue directly within the framework of MPLT, and give arguments based on Chomsky's notion of Domain Extension: overt object shift creates a minimality barrier for subject raising, and this barrier can be overcome only by Verb Raising from $AgrO^0$ to T^0. This extends the minimal domain of the verb, rendering Spec-AgrOP and Spec-TP *equidistant,* so that subject raising over Spec-AgrOP does not constitute raising over a closer potential landing site. Since participles do not raise to T^0, Spec-AgrOP and Spec-TP/Spec-AgrSP are never equidistant, to a subject raisng to either of these positions would always be crossing the "closer" Spec-AgrOP where the object resides. However, we would like to suggest another way to explain this phenomenon.

Above, we concluded that AgrO had optionally strong N-features; thus we saw optional object shift in (12a–b). But this caused a problem for (12c–d) which must be resolved. Let us begin instead with a different assumption: AgrO *always* has weak N-features in Icelandic. This is conceivably more desirable than assuming two possible values for a parameter of feature strength. But the appearance of overt object shift cannot now be due to the optionality of N-feature strength on AgrO; rather, it must be due to some other means through which the object might end up with its phonological features in Spec-AgrOP. What follows is a close analysis of the structure of the chains formed in the process of V-raising to $AgrO^0$ and T^0.

First note that the following sentence is ungrammatical in Icelandic:

(13) **Jón hefur* [$_{AgrOP}$ ***keypt*** *ekki* [$_{VP}$ $\mathbf{t_V}$ *bókina*]].
John has bought not book.the
'John has not bought the book.'

It is not appear possible for an uninflected participle to raise overtly out of VP. However, we must assume that it does raise to check its Agr features against $AgrO^0$; we must therefore conclude that AgrO bears weak V-features in Icelandic (unlike T and AgrS), or else the verb's phonological material would be moved to $AgrO^0$, incorrectly generating (13).[9]

Now consider the derivation of (12a), in which the inflected verb *keypti* has

---

9. Our analysis predicts nothing new concerning auxiliaries and the auxiliary in (13) in particular. We assume that the auxiliary in (13) originates above AgrOP; whether it moves to T^0 from a separate head position, or whether it is base-generated in T^0 is not relevant to our argument, as long as the auxiliary has the property of rendering movement of the main verb to T^0 unnecessary.

moved visibly to $AgrS^0$, via T^0. The derivation proceeds in the following way: first, the verb raises to $AgrO^0$ from its base position in VP. Phonological material is obligatorily not transferred, since AgrO, as we concluded above, has weak V-features. We thus have the following structure:

(14) $[_{AgrO'}\ [_{AgrO^0}$ ***keypti*** $AgrO^0]$ *ekki* $[_{VP}$ *Jón* ***keypti*** *bókina*]]

(As before, the underline indicates where the phonological material of the category is found.) The object must now shift to Spec-AgrOP, before additional structure is created dominating AgrO′, by the Strong Cycle. But we hypothesize here that AgrO's N-features are always weak; thus the tail of the object Case-chain is pronounced, and we have the following structure:

(15) $[_{AgrOP}$ ***bókina*** $[_{AgrO^0}$ *keypti* $AgrO^0]$ *ekki* $[_{VP}$ *Jón keypti* ***bókina***]]

Now we draw T^0 from the lexicon, project it to T′, and insert the above AgrOP into its complement position by GT:

(16) $[_{T'}$ $\boldsymbol{T^0}$ $[_{AgrOP}$ *bókina* $[_{AgrO^0}$ *keypti* $AgrO^0]$ *ekki* $[_{VP}$ *Jón keypti bókina*]]]

Now, following again assumptions made in Jonas & Bobaljik (1993), T has strong V-features, so the inflected verb checks its tense features by adjoining to T^0 and copying phonological material. Under Jonas and Bobaljik's analysis, checking of T's features is accomplished by raising the $AgrO^0$ head containing the V^0 to T^0. But under our analysis this would fail to check the strong features of T, since there are no phonological features in $AgrO^0$: the verb's phonological features are *still in the verb's base position in VP.* Thus our analysis requires that a V^0-chain is formed to T^0, instead of an $AgrO^0$-chain, to T^0 as shown below:

(17)

$[_{T'}\ [_{T^0}$ ***keypti*** $\mathbf{T^0}][_{AgrOP}$ *Bókina* $[_{AgrO^0}$ *keypti* $AgrO^0]$ *ekki* $[_{VP}$ *Jón* ***keypti*** *bókina*]]]

This is the crucial point in the analysis: because V-raising to Agr must not copy phonological features, they are left behind in the verb's base position; thus when they are needed to check strong features in T, a chain must be formed from the position *from which those phonological features can be transferred,* i.e. from the verb's base position in VP. We thus have a V-chain from V's base position to T^0.[10,11]

---

10. The example in (17) appears to create a violation of the Head Movement Constraint; see footnote 13 for the resolution of this and related issues.

Now, by the definition of Internal Domain in MPLT (Chomsky 1993: 11–17), Spec-AgrOP is in the Internal Domain of the V-chain in (17).[12] In MPLT, it is suggested that the Internal Domain is the domain in which internal arguments of verbs receive their internal theta-role at LF (MPLT: 12). It suffices, then, that a direct object be in the Internal Domain of the verb by LF to receive its theta-role. Why could we not *base-generate* a direct object in Spec-AgrOP in a sentence like the above, where it would be in the internal domain of the verb by LF, after V-raising to T^0, to receive its theta-role? There seems to be no principled reason why this should not be allowed.

This is the possibility we see realized in (12b): the object has not *shifted* to Spec-AgrOP from its base position as the complement of V, but rather is simply *base-generated* by binary GT into Spec-AgrOP, where its Case and Agr features are checked, and where it receives its internal theta-role from the V-chain at LF. The structure for (12b) is thus as follows (after Subject Raising, with phonological features, to Spec-AgrSP):

(18)

Jón ... [T′ [T^0 *keypti* T^0][$_{AgrOP}$ ***bókina*** [$_{AgrO^0}$ *keypti* $AgrO^0$] *ekki* [$_{VP}$ *Jón keypti*]]]

By virtue of the structure of the V-chain forced by the interaction of weak-feature-checking in AgrO and strong-feature-checking in T, Spec-AgrOP is a theta-position, being in the verb's Internal Domain. Thus Icelandic "object shift" is not really shift at all, but rather optional base-generation of the object in Spec-AgrOP.

---

11. This reasoning implies that there can be no movement to a strong position *via* a weak position. We predict, then, that in the case of English *wh*-movement of an object DP, the Case chain to Spec-AgrOP (where weak N-features are checked) and the *wh*-chain to Spec CP (where strong *wh*-features are checked) *both* have their tails in VP. Empirical arguments are difficult to construct (but cf. Kitahara's [1993] argument that the tail of an object *wh*-chain is in Spec-AgrOP, not complement of V).

12. The Domain of the V-chain is the set of nodes contained in Max(V) (=T′) that are distinct from and do not contain V, that is {Spec-AgrOP, Spec-VP, Complement of V} and whatever these categories dominate; the Complement Domain is the subset of the Domain reflexively dominated by in the complement AgrOP of the construction (the same set in this case), and the Internal Domain is the minimal Complement Domain, or exactly the set {Spec-AgrOP, Spec-VP, Complement of V}.

Note that no V-to-T chain is formed in examples (12c–d), since the verb is not inflected for tense. Thus Spec-AgrOP is never in the Internal Domain of the theta-marking main verb. The object could not be base-generated in Spec-AgrOP, since it would not be in the Internal Domain of the verb by LF, and would not receive its internal theta-role. Instead, it must be base-generated as complement of V, and raise to Spec-AgrOP to check N-features there. This raising will not result in movement of phonological features of the object, since by hypothesis AgrO always has weak N-features, as in (19):

(19) *Jón* ... [$_{T'}$[$_{T^0}$ *hefur* T^0][$_{AgrOP}$ *bókina* [$_{AgrO^0}$ *keypt* $AgrO^0$] *ekki* [$_{VP}$ *Jón keypt bókina*]]]

Object "Shift" (i.e. base-generation in Spec-AgrOP) is correctly predicted to be impossible in sentences without Verb Raising to T/AgrS.[13]

3.2.2. *English: No Optional Base-Generation of Object in Spec-AgrOP*

The question arises whether the posibility of base-generating objects in Spec-AgrOP might overgenerate. Consider an English finite clause without an auxiliary, where the verb must check features against $AgrO^0$, T^0, and $AgrS^0$. Following MPLT and Pollock (1989), we assume the V-features of all of these functional heads to be weak; thus we do not see movement of the verb: phono-

---

13. Two issues arise with respect to possible minimality violations. First, in sentence (16), we hypothesize Object Shift has occurred to Spec-AgrOP, but is not visible since AgrO has weak N-features. However, the (invisible) object in Spec-AgrOP would present a minimality barrier for subsequent Subject Raising out of VP to Spec-AgrSP (or Spec-TP) for Case-checking: it occupies an intervening A-position between Spec-VP and Spec-AgrSP/Spec-TP. Since no further Verb Raising occurs past $AgrO^0$ (the verb being an uninflected participle), there would seem to be no way of resolving this problem through an Equidistance account like Jonas & Bobaljik's (1993) analysis. However, a theory of Shortest Move such as that proposed in Ferguson (this volume) resolves this problem; the closest visible Case-checking category for the subject is T, and intervening categories (like the object in Spec-AgrOP) are therefore irrelevant to the Shortest Move Principle. We thus adopt Ferguson's principle of Shortest Move in this paper.

The other potential problem involves the apparent HMC violation in (17) and (18), where formation of the verb chain crosses the $AgrO^0$ head. However, again following Ferguson (this volume), this poses no problem: for the verb to form a tense-chain to check its tense features, it suffices that there are no intervening heads capable of checking tense features. $AgrO^0$ does not check tense-features, so the movement is unproblematic.

logical features are left behind. Let us take the following simple sentence as an example:

(20) *John saw Mary.*

By Spell-out, the structure will be as follows (for simplicity, we do not show the full copy of the $AgrO^0$ trace with verb adjoined, but represent it in the usual way with a trace "t_{AgrO^0}"):[14]

(21)

$[_{AgrSP}$ *John* $[_{AgrS}$ $[_{AgrO}$ *saw* $AgrO^0]$ $[_{AgrS}$ T^0 $AgrS^0]]$ $[_{TP}$ t_{T^0} $[_{AgrOP}$ *Mary* t_{AgrO}

$[_{VP}$ *John saw Mary*$]]]]$

We see that the only category whose phonological features are copied in movement is the subject DP *John;* this is because only the Case features of T are strong in English, requiring phonological material to be checked. The object, *Mary,* is pronounced in complement position of VP, since no phonological material is moved to Spec-AgrOP (AgrOP has weak N-features in English).

Note that we have indicated the chain from AgrO to AgrS; that is the movement of the verb in AgrO to T/AgrS to check tense and agreement features. Since this chain is an AgrO chain, not a V-chain, Spec-AgrOP is not in the Internal Domain of the verb. Thus there would be no possibility of base-generating the object in Spec-AgrOP. But why couldn't the chain have a structure analogous to (18), with a V-chain to $AgrS^0/T^0$ having a tail in VP, as in (22)? Then we could base-generate the object in Spec-AgrOP, since that object would be in the internal domain of the V-chain to AgrS/T:

(22)

$[_{AgrSP}$ *John* $[_{AgrS}$ *saw* $[_{AgrS}$ T^0 $AgrS^0]]$ $[_{TP}$ t_{T^0} $[_{AgrOP}$ *Mary* $[_{AgrO}$ *saw* AgrO$]$

$[_{VP}$ *John saw*$]]]]$

*'John Mary saw.'

In this structure, the category moving to AgrS/T is V^0, not $AgrO^0$. The difference between (22) and the Icelandic sentence in (18) is that phonological

---

14. We also assume, following Chomsky (1993), that T^0 adjoins to $AgrS^0$ before V-raising, as in note 3.

material is not transferred in (22), since the features being checked in T^0 are weak. But we still have a V-chain to T/AgrS, instead of an AgrO-chain, so the Internal Domain of the verb includes Spec-AgrOP.

Note, however, that verb movement to T^0 from V^0, instead of from $AgrO^0$, was *obligatory* in Icelandic Verb Raising, since the base V^0-position was where the phonological features were, and phonological features needed to be moved to check strong T-features. Here, no such restriction applies, since no phonological features need to be checked; thus either of the structures (21) or (22) should be possible. Yet we know that *John Mary saw* is quite ungrammatical: how do we rule out the derivation shown in (22)?

The answer lies simply in derivational economy: given two options for the verb to satisfy its tense and agreement feature checking, the option chosen will be the most economical. With respect to chain length, clearly the $AgrO^0$ chain of (21) is shorter than the V^0-chain of (21), which must additionally cross V′ and VP. Thus the shortest chain that can be formed to satisfy tense and agreement checking for the verb is the chain shown in (21), and the chain in (22) is excluded. (Cf. Kitahara [1993] for a similar use of a "Shortest Chain Requirement" in deducing Superiority effects, though using a transderivational economy metric.)

Thus we find that in English, the internal domain of the verb is *not* extended by V^0-raising, since it is $AgrO^0$ that actually raises to T/AgrS. Icelandic, on the other hand, requires the formation of a V^0-chain to T^0, in order that phonological features can be moved to T^0. Thus the Internal Domain of the verb *is* extended, and the direct object has the option of being base-generated in Spec-AgrOP, where it can receive its theta-role at LF. If the verb does not raise past $AgrO^0$, as in the case of participle constructions (12c–d), the Internal Domain of the verb is not extended to include Spec-AgrOP, and thus "object shift" will be impossible.

In sum, then, we see that different chain structures ensue by virtue of whether or not phonological material needs to be moved to T/AgrS when verb movement occurs. Our analysis predicts these different structures, which in turn provide an account of Holmberg's Generalization and the optionality of Icelandic "object shift" without recourse to variable N-feature strengths for AgrO in Icelandic.

4. Summary

We have suggested a model of human grammar, based essentially on the minimalist approach, in which the syntactic computational system is uniform, without pre-/post-Spell-out asymmetries. This approach has the conceptual advantages of allowing us to deduce a Strong Cycle from simple notions of structure-building, eliminating the asymmetries in favor of a single, uniform computational system, and recasting a transderivational economy constraint (Procrastinate) in derivationally local terms.

We then saw that this approach simplifies Watanabe's (1992) account of subjacency effects in Japanese *wh*-movement, allowing us to dispense with his null-operator movement. Instead, full *wh*-phrase movement, exactly as in English but without movement of phonological features, accounts for subjacency effects in Japanese. We also saw that this approach makes certain predictions about the chain structure for verbs in a language like Icelandic with overt V-raising to T/AgrS, suggesting that object shift in such a language is simply a reflex of the option of base-generating objects in Spec-AgrOP, where the predicted V-chain structure in Icelandic makes this a theta-position at LF. This option is precluded in English, which due to the weak inflectional features of T/AgrS has different V-chain structures that do not make Spec-AgrOP a theta-position.

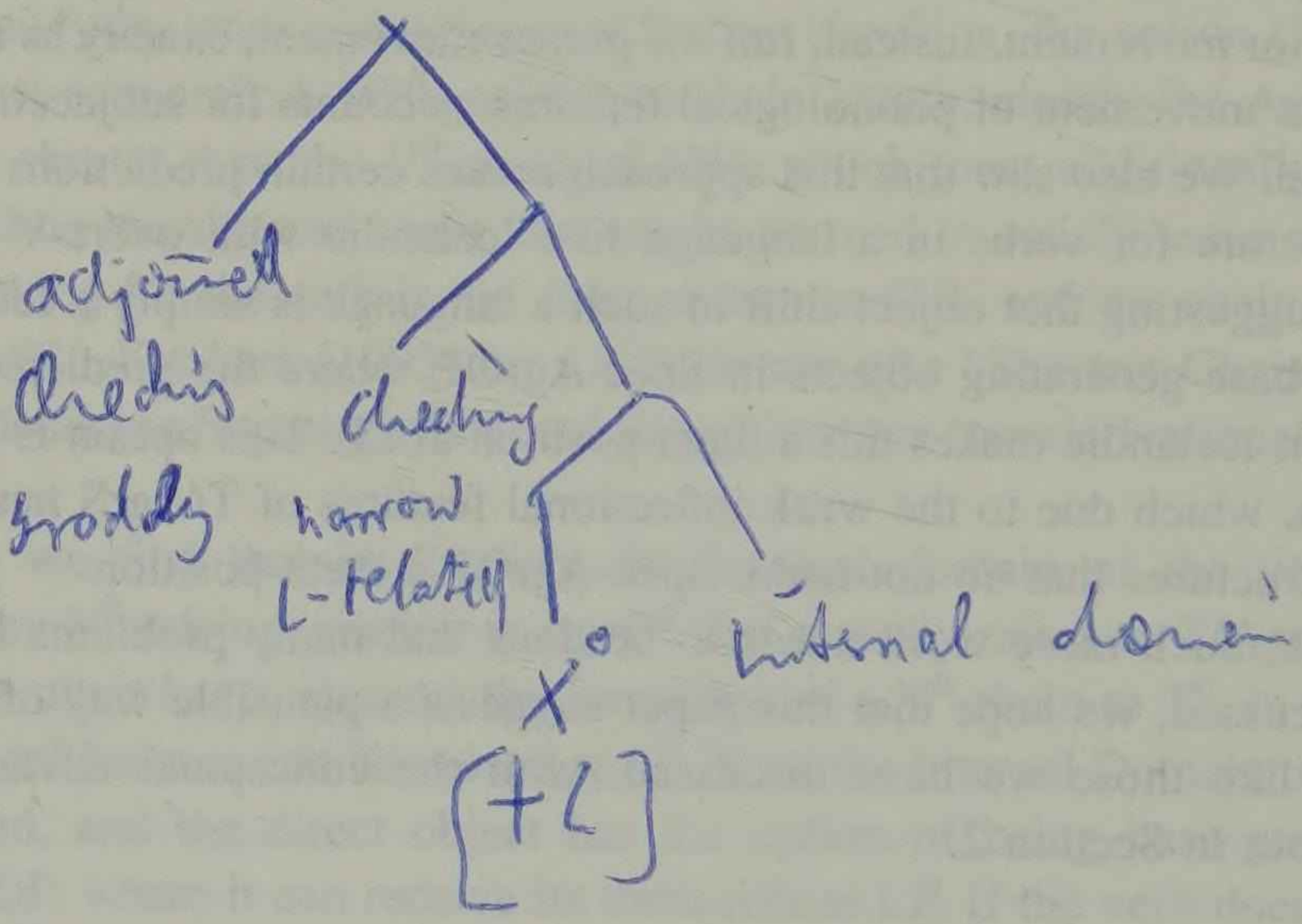
While much more work needs to be done and many problems have been left undiscussed, we hope that this paper suggests a plausible way of attacking problems like those we have discussed, with the conceptual advantages we sketched out in Section 2.

Acknowledgements

The authors would like very much to thank Noam Chomsky, Samuel D. Epstein, Scott Ferguson, Hisatsugu Kitahara, David Pesetsky, and Höskuldur Thráinsson for helpful discussion, and a plethora of crucial suggestions. All remaining errors are the responsibility of the authors and their extended families.

delimitation more to AgrOP

V vs L features



two types of AgrOP
check case at Spec AgrOP
check R-feature at higher projection
⇓
(φ-feature)

The Typology of Syntactic Positions
L-Relatedness and the A/$\overline{\text{A}}$-distinction

Liliane Haegeman
University of Geneva

> *The functional elements Tense and Agr therefore incorporate features of the verb. Let us call these features* V-features*: the function of the V-features of an inflectional element I is to check the morphological properties of the Verb selected from the lexicon. More generally, let us call such features of a lexical item L* L-features*. Keeping to the X-bar-theoretic notions, we say that a position is* L-related *if it is in a local relation to an L-feature, that is, in the internal domain or checking domain of a head with an L-feature. Furthermore, the checking domain can be subdivided into two categories: non-adjoined Spec and adjoined. Let us call these positions* narrowly *and* broadly *L-related, respectively. A structural position that is narrowly L-related has the basic properties of A-positions; one that is not L-related has the basic properties of $\bar{A}$-positions, in particular, Spec-C, not L-related if C does not contain a V-feature.*
>
> — *Chomsky (1993: 28–29)*

Abstract

In the traditional Government and Binding framework the distinction between XP-positions is expressed in terms of the A/$\overline{\text{A}}$-opposition. Chomsky (1993: 28–29) proposes that the contrast between A-positions and $\overline{\text{A}}$-positions be replaced by a contrast in terms of ± L-relatedness. I present evidence to suggest that the A/$\overline{\text{A}}$-distinction ought to be maintained independently of the notion of L-relatedness. The empirical focus of the paper is the distribution of the "scrambled" object DPs and that of object clitics *ze* 'them/her' and *t* 'it' in West Flemish.

In addition to the familiar projections such as CP, AgrSP, TP, NegP, AgrOP, this paper argues that the WF clause contains functional projections high in the clausal domain whose heads host a feature R, a *phi* feature. Two types of leftward DP-movement should be distinguished: (i) short, Case-driven, DP-movement to the specifier of AgrOP; and (ii) DP-movement to the specifier of the higher projections for the checking of the R-feature ("Zone 1").

Definite DPs first move to Spec-AgrOP and, depending on the presence of the R-feature, they move to the higher domain. The object clitic first undergoes DP-movement to Spec-AgrOP, for Case licensing. Then it moves as DP to the specifier of the functional R-projection. Finally, the clitic-head of the DP in the specifier of the R-projection is extracted and adjoins to the head of the R-projection.

The functional R-projections in Zone 1 are not V-related. R is an N-feature, it is not a V-feature. As a result, finite V^0-movement to C^0 will by-pass the functional heads in Zone 1. On the other hand, since ⟨R⟩ is an N-feature, the specifiers of the R-projections are A-positions. The functional projections in Zone 1 are then evidence for the dissociation of the A/$\overline{\text{A}}$-distinction and V-relatedness.

1. Introduction

1.1. *Aim and Scope of the Paper*

This paper focuses on the typology of syntactic positions. In the traditional Government and Binding framework the main distinction between XP-positions is expressed in terms of the A/$\overline{\text{A}}$-distinction. As the above quotation from Chomsky (1993) shows, the distinction in terms of L-relatedness (± L-related) replaces the A/$\overline{\text{A}}$-distinction in the Minimalist Program.

In this paper I present some empirical evidence to suggest that the A/$\overline{\text{A}}$-distinction ought to be maintained independently of the notion of L-relatedness. Specificallly, I will show that there are A-positions which appear not to be L-related. The empirical focus of the paper is the distribution of clitics in West Flemish (WF).

1.2. *Organization of the Paper*

The paper is organized as follows. The remainder of Section 1 contains a summary of the standard Government and Binding assumptions concerning the notions of A- and $\overline{\text{A}}$-positions, as well as a brief discussion of the notion L-relatedness in Minimalist terms. Section 2 is a discussion of the distribution of argument DPs in WF. In Section 3 I discuss the distribution of clitics in WF. Section 4 is the conclusion.

1.3. *The Government and Binding Tradition: A-positions vs. $\bar{A}$-positions*

In this section I briefly go over the discussion of A-positions and $\overline{\text{A}}$-positions developed in Haegeman (1995a). This section elaborates and extends the analysis of Rizzi (1991b).

1.3.1. *A-positions*

In the Government and Binding literature (Chomsky 1981 and subsequent work) A-positions are the nuclear positions in the clause to which Grammatical Functions (GF) such as subject or object are assigned in traditional grammar. $\overline{\text{A}}$-positions are peripheral positions which are not associated with a GF. Spec-IP, for instance, is an A-position: it is the canonical subject position; Spec-CP, on the other hand, is a prototypical $\overline{\text{A}}$-position. A-positions are relevant for binding, the Binding Theory being a theory of A-binding.

In the earlier versions of the theory, A-positions were defined as potential thematic positions, i.e. positions to which thematic roles could be assigned. Since the canonical subject position, Spec-IP, was assumed to be the base position of the subject in transitive and intransitive sentences, it was a potential thematic position, hence it was an A-position.

Following work by, among others, Sportiche (1988), Koopman & Sportiche (1991), etc. it is now generally assumed that the canonical subject position, i.e. Spec-IP or Spec-AgrSP in the split Infl tradition (Pollock 1989; Belletti 1990), is *not* the thematic position of the subject. The thematic position of the subject is VP-internal; the subject DP moves to Spec-AgrSP in order to check agreement features. Still, the arguments for the A-status of the canonical subject position Spec-IP remain valid. I repeat one such argument here (cf. Haegeman 1995a, Chap. 5). Consider (1):

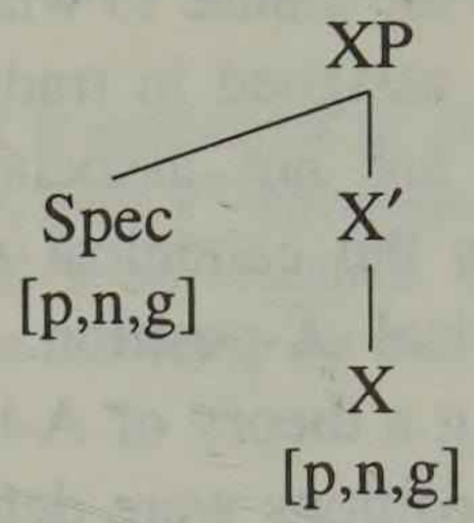
(1) *He*$_i$ *is* [$_{VP}$ t_i *going to bed*] [*because John*$_{j/*i}$ *is sick*].

Coreference between the subject of the matrix clause, *he,* and that of the embedded reason clause, *John*, is excluded, a typical Principle C effect. If the reason clause *because John is sick* is adjoined higher than VP, then we cannot impute the disjoint reference effect to the role played by the VP-internal trace of the subject, t_i, since t_i will be too low in the structure to c-command the DP *John*. We conclude that it is the DP in the canonical subject position of the matrix clause, *he*$_i$, which is responsible for the Principle C effect in (1): *John*$_i$ is bound illicitly by the pronoun *he*$_i$. This means that Spec-AgrSP, the position occupied by the subject *he*, is an A-position.

Integrating the proposals that the canonical subject position, Spec-AgrSP, is *not* a theta position — the base position of the subject being VP-internal — and that Spec-AgrSP *is* an A-position, Rizzi (1991b) modifies the definition of A-positions and proposes that A-positions are (i) thematic positions or (ii) specifiers in construal with an Agr head (see also Vanden Wyngaerd 1989a). Making the second component of the definition more precise, let us say that

what is relevant for the A-status of a (non-thematic) position is that it be a specifier position of an Agr projection, and that the DP in the specifier position share *phi* features (*p*erson, *n*umber, *g*ender) with the head:

(2)



1.3.2. *Ā-positions*

In the Government and Binding tradition A-positions and $\overline{\text{A}}$-positions were originally considered to be in complementary distribution. There was no positive definition for $\overline{\text{A}}$-positions: $\overline{\text{A}}$-positions were those (XP)-positions which did not qualify as A-positions. In Haegeman (1995a) I develop an account in which $\overline{\text{A}}$-positions, as well as A-positions, are defined in terms of their intrinsic properties. I will briefly sum up the basic ideas here.

Let us consider the prototypical $\overline{\text{A}}$-positions: Spec-CP and Spec-NegP. The $\overline{\text{A}}$-status of Spec-CP is based on the observation that it characteristically hosts operators (such as *wh*-phrases). Spec-NegP is assumed to be occupied by an (negative) operator. Let us therefore identify $\overline{\text{A}}$-positions as those which have operator features. If both Spec-CP and Spec-NegP are $\overline{\text{A}}$-specifiers, the blocking effect of Spec-NegP on *wh*-movement of adjuncts (cf. Rizzi 1990) is expected.

I assume that *wh*-operators are subject to the *Wh*-Criterion (Rizzi 1991a) and that negative operators are subject to the Neg-Criterion (Haegeman 1995a, Chap. 2),

(3) *Wh*-criterion (May 1985; Rizzi 1991a: 2)

a. A *wh*-operator must be in a Spec–head configuration with an X^0 with the feature ⟨*wh*⟩;

b. An X^0 with the feature ⟨*wh*⟩ must be in a Spec–head configuration with a *wh*-operator.

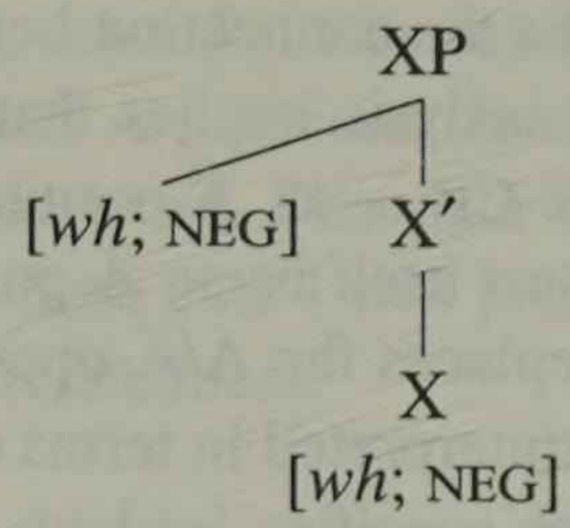
(4) Neg-criterion (Haegeman 1995a)

a. A Neg-operator must be in a Spec–head configuration with an X^0 with the feature ⟨Neg⟩;

b. An X^0 with the feature ⟨Neg⟩ must be in a Spec–head configuration with a Neg-operator.

Wh-operators and negative operators have to attain a specifier-head relation with a head with the matching operator feature. By analogy with the definition of A-positions above, we can then say that a constituent occupies an $\overline{\text{A}}$-position if it occupies a specifier position in which it shares operator features with a head, where operator features include *wh,* Neg.[1]

(5)



We now arrive at a symmetric definition of A and $\overline{\text{A}}$-positions, in which both the configurational relations (specifier-head) and the content of the positions (feature sharing) play a part. A-specifiers agree with their heads in terms of *phi* features (person, number, gender, perhaps case); $\overline{\text{A}}$-positions agree with their heads in terms of operator features (Neg, *wh*) (cf. Agouraki 1993 for similar ideas).[2]

1.4. *L-relatedness and A-positions*

As illustrated by the quotation introducing this article, Chomsky (1993) fully assimilates the contrast between L-related positions and non-L-related positions to the A/$\overline{\text{A}}$-distinction. This view would presumably lead us to expect that only L-related positions can block chains of L-related positions, and that only non-L related positions can block non-L-related chains. Consider (6) and (7):

(6) a. *Why did* [$_{\text{IP}}$ *they say t* [$_{\text{CP}}$ *that* [$_{\text{IP}}$ *he had fired John*]]]?
b. *Why did* [$_{\text{IP}}$ *they say* [$_{\text{CP}}$ *that* [$_{\text{IP}}$ *he had fired John t*]]]?

(7) a. *Why did* [$_{\text{IP}}$ *they* [$_{\text{NegP}}$ *not say t* [$_{\text{CP}}$ *that* [$_{\text{IP}}$ *he had fired John*]]]]?
b. **Why did* [$_{\text{IP}}$ *they* [$_{\text{NegP}}$ *not say* [$_{\text{CP}}$ *that* [$_{\text{IP}}$ *he had fired John t*]]]]?

---

1. The status of Spec-TP is not clear. Jonas & Bobaljik (1993) consider it to be an A-position. Rizzi & Roberts (1989) consider it to be an $\overline{\text{A}}$-position.

2. In Haegeman (1995a, Chapter 5) I argue that given the definitions developed here we also expect that there are mixed A/$\overline{\text{A}}$-positions, i.e. positions which qualify both as A– and as $\overline{\text{A}}$-positions. See also Rizzi (1991b). This point is tangential to the present discussion.

In (6a) *why* is construed locally, with the matrix clause; in (6b) it has long construal, i.e., it is chain-connected to a trace in the embedded clause. In (7a) the matrix clause is negated; *why* has local construal and the sentence is grammatical. We see in (7b) that when the matrix clause is negated long construal of the adjunct is no longer possible. Following Rizzi (1990) we assume that the intervening *not* in Spec-NegP blocks the connection between the moved adjunct *why* and its trace in (7b). Such an analysis implies that Spec-NegP is a position of the same kind as Spec-CP. Spec-CP is an $\overline{A}$-position. If Spec-NegP blocks movement to an $\overline{A}$-position, then it must itself be an $\overline{A}$-position.

In a framework where L-relatedness replaces the A/$\overline{A}$-opposition, Spec-CP is non-L-related. If the A/$\overline{A}$-distinction is reinterpreted in terms of L-relatedness, the syntactic properties of Neg^0 and of its specifier lead us to contradictory proposals. On the one hand, in (7b) Spec-NegP blocks adjunct-movement to a non-L-related position. This should mean that Spec-NegP is not L-related, hence that the associated head, Neg^0, is not L-related. But on the other hand, V^0-movement to AgrS typically cannot skip Neg^0. Rather, V^0 and the negative head form one complex head, which, containing V^0, is L-related (see Belletti 1990). The specifier of NegP will be in the checking domain of the derived V^0+Neg^0 head, an L-related head. This should mean that Spec-NegP is L-related.

In order to account for the blocking effect of Spec-NegP on the movement of the adjunct to Spec-CP, one option is to maintain the contrast between A-positions and $\overline{A}$-positions, in addition to the distinction in terms of L-relatedness. The head Neg^0 could then be said to be L-related but since Neg^0 also contains an operator feature ⟨Neg⟩, Spec-NegP will be identified as an $\overline{A}$-position. If this account is on the right track we cannot equate A-positions with L-related positions and $\overline{A}$-positions with non-L-related positions; rather than two types of positions we end up with four types:

(8) *The typology of positions*

	L-related	non-L-related
A	Spec-AgrSP	
+N	SpecAgrOP	
	?Spec-TP	
	(Chomsky 1993; Jonas & Bobaljik 1993)	
$\overline{A}$ −N	Spec-NegP	Spec-CP

The interdependence of the A/$\overline{\text{A}}$-distinction and the ± L-relatedness opposition can be expressed in terms of N-features and V-features.[3] Heads are A-heads if they carry N-features; heads are $\overline{\text{A}}$-heads if they carry operator features. Heads are L-related (or, more specifically, V-related) if they carry V-features. Heads carrying N-features (person/number/gender) but no V-features would be A-heads which are not L-related. In the classification in (8) there are no examples of such non L-related A-positions. In the remainder of the paper I will argue that such positions do exist. The data I will be using will be drawn from West Flemish. Given these data I will propose that A/$\overline{\text{A}}$-distinction is independent of the notion of L-relatedness.

2. Distribution of Arguments: "Scrambling"

The crucial evidence for the existence of non-L-related A-positions will be derived from the distribution of lexical DP complements and clitic complements in West Flemish (WF). I discuss lexical DP complements in Section 2 and clitic complements in Section 3.

2.1. *Some Preliminary Notes on the Syntax of West Flemish*

Like Standard Dutch, and like German, WF is a so-called Verb Second (V2) language: in matrix clauses the finite verb ends up in second position. For the purposes of this paper I assume that the finite V^0 moves to C^0 giving rise to the V2 effect. I will not enter into the details of this analysis, specifically I do not want to decide between an analysis where V^0 always moves to a specific position in C^0 and an analysis in which the V2 phenomenon is not unified (Zwart 1993b; Haegeman 1995b). Nor do I want to examine the nature of the V2 constraint. I hope to return to this issue in future work.

Non-finite verbs and finite verbs of non-root clauses appear in sentence-final position in WF. In the traditional analysis this is taken as evidence that the functional projections of the IP system, AgrP and TP, are head final. In seminal work on Dutch syntax Zwart (1993b; 1994a) adopts a universal base hypothesis (cf. Kayne 1994) which proposes that the specifier–head order and the head-complement order are universally fixed: specifiers precede their heads and heads precede their complements. Following this line the functional projections of the

---

3. I thank an anonymous reviewer for clarifying this point.

IP system of Germanic are head-initial, and the sentence-final position of the verb in the contexts mentioned is due, on the one hand, to the fact that the V^0 does not move to the highest functional head before Spell-out, and, on the other hand, to a generalized leftward movement of the complements and adjuncts.[4]

2.2. *The Distribution of Complements*

2.2.1. *Three Zones for Arguments*

Consider the distribution of the direct object *dienen boek* in (9). Following my own work on negation (Haegeman 1995a) I assume that Spec-NegP in WF is occupied by *nie*.[5] In the examples in (9) we see that the object DP may follow (9a) or precede (9b) the sentence adverbial *verzekerst*, but it must precede the negation marker *nie* (9c).

(9) a. *da Valère verzekerst dienen boek*$_i$ | *nie an Marie* t_i *gegeven eet*
that Valère probably that book | not to Marie given has
'that Valère probably did not give that book to Mary.'

---

4. It is not obvious if inflected verbs undergo any movement at all in embedded clauses. Zwart (1993b) assumes that the inflected V does not move to AgrS in the syntax. Koopman (1995) argues that the inflected verb in Dutch fails to move to T or to Agr. Kayne (1994) proposes that the inflected V^0 does move to AgrS. Kayne argues for V-movement given the observation that Dutch verbs inflect for person. In an analysis of Infinitivus pro Participio constructions in WF I propose that inflected verbs in embedded clauses may move to T (Haegeman 1995c).

5. In earlier work (Haegeman 1995a) I have assumed that *en* is the head of NegP:
(i) a. *da Valère dienen boek nie en-kent*
that Valère that book not *en* knows
If NegP is head initial and if *nie* is a specifier of NegP this is not tenable, since *nie* can be separated from the head *en* by intervening material:
(i) b. *da Valère dienen boek nie an Marie getoogd en-eet*
that Valère that book not to Marie shown *en* has
'that Valère de not show that book to Marie.'
I hope to return to this issue in later work. Let me simply point out that the distribution of negative quantifiers suggests strongly that *nie* occupies a fixed position in the clause. *Nie* seems to function as the marker of sentential negation, much along the lines of French *pas*. If we continue to assume that *nie* is in Spec-NegP the position of *en* will have to be reinterpreted. Thanks to Jan-Wouter Zwart for discussing this point extensively with me. See also Haegeman (1995c).

b. *da Valère dienen boek*$_i$ *verzekerst nie an Marie* t$_i$ *gegeven eet*
that Valère that book probably not to Marie given has
'that Valère probably did not give that book to Marie.'

c. **da Valère verzekerst nie an Marie dienen boek gegeven eet*
that Valère probably not to Marie that book given has

An analysis of the distribution of definite argument DPs suggests that between the sentence-initial complementizer and the sentence-final finite V there are (at least) three zones in which object arguments appear in WF, which I schematically represent in (10):[6]

(10) Three zones for objects

1	Subj ↔ adverb		DPs
2		adverb ↔ *nie*	DPs
3		*nie* ↔ V	*DPs

In a framework where it is assumed that SVO is the universal base structure, Zone 3 can be equated with the Predicate Phrase as discussed by Zwart (1994a) and Koster (1995). Zone 2 is the domain immediately to the left of the specifier of NegP, *nie*. Definite object DPs either follow the sentential adverbial, and occupy Zone 2 (9a), or they precede a sentence adverbial, i.e. they appear in Zone 1 (9b). They cannot appear in Zone 3 (9c). Leftward movement into Zone 1 and Zone 2 is often referred to as scrambling. There is a vast literature on scrambling in Dutch and German and for reasons of space I do not go into it here. The reader is referred to Vanden Wyngaerd (1989a), Johnson (1991), de Hoop (1992), Frank *et al.* (1992), Zwart (1993b; 1994a), Haeberli (1994).

2.2.2. *A-movement*

In Haegeman (1993a) I argue that the leftward movement of the definite DP arguments is A-movement. I will recapitulate the arguments here. The first argument for the assumption that the leftward DP-movement is A-movement is the observed asymmetry between DP arguments and PP arguments. While DP-movement to the left of *nie* is obligatory, PP-movement is optional, This is illustrated in (11). When realized as a PP the indirect object can remain lower than negation, as shown in (11a); a parallel DP argument must move to precede the negation (11b). Similarly, in (12a) the verb *helpen* 'help' may select a PP

---

6. I do not discuss indefinite arguments here. See de Hoop (1992) for relevant discussion.

complement, which may remain lower than negation. When a DP-complement is selected by *helpen*, then the complement must precede negation (12b).

(11) a. *da Valère dienen boek nie *(an) Marie gegeven eet*
that Valère that book not to Marie given has
'that Valère did not give that book to Marie'

b. *da Valère Marie dienen boek nie gegeven eet*
that Valère Marie that book not given has
'that Valère did not give Marie that book'

(12) a. *da Valère nie *(me) Marie gholpen eet*
that Valère not with Marie helped has
'that Valère did not help Marie'

b. *da Valère Marie nie gholpen eet*
that Valère Marie not helped has
'that Valère did not help Marie'

In Haegeman (1993a) I have shown that there is a rigid ordering constraint on the sequencing of argument DPs in WF, which applies both in Zone 1 and in Zone 2. (13) summarizes the data. The definite subject (Subj) always precedes the indirect object (IObj), which in turn precedes the direct object (DObj).

(13) a. *da Valère Marie dienen boek verzekerst nie toogt*
that Valère Marie that book probably not shows
C Subj IObj DObj Adv *nie* V
'that Valère probably does not show that book to Marie'

b. *da Valère Marie verzekerst dienen boek nie toogt*
that Valère Marie probably that book not shows
C Subj IObj Adv DObj *nie* V
'that Valère probably does not show that book to Marie'

c. *da Valère verzekerst Marie dienen boek nie toogt*
that Valère probably Marie that book not shows
C Subj Adv IObj DObj *nie* V
'that Valère probably does not show that book to Marie'

The alternative orders are ungrammatical. Note specifically that the IObj can never follow the DObj:

WF=
always: S iO dO
DP DP DP
(iO
DP

(14) a. **da Valère dienen boek Marie verzekerst nie toogt*[7]
that Valère that book Marie probably not shows
b. **da Valère dienen boek verzekerst Marie nie toogt*
that Valère that book probably Marie not shows
c. **da Valère verzekerst dienen boek Marie nie toogt*
that Valère probably that book Marie not shows

The ordering constraint which imposes the sequencing Subj–IObj–DObj applies to DP arguments both in Zone 1 and Zone 2. Again there turns out to be an asymmetry between DP arguments and PP arguments. The ordering constraint only applies to DP arguments; IObj PPs, unlike IObj DPs, can follow the direct object, both in Zone 1 and in Zone 2:

(15) a. *da Valère dienen boek an Marie verzekerst nie gegeven eet*
that Valère that book to Marie probably not given has
'that Valère probably did not give that book to Marie'
b. *da Valère dienen boek verzekerst an Marie nie gegeven eet*
that Valère that book probably to Marie not given has
'that Valère probably did not give that book to Marie'
c. *da Valère verzekerst dienen boek an Marie nie gegeven eet*
that Valère probably that book to Marie not given has
'that Valère probably did not give that book to Marie'

One interpretation of the DP/PP asymmetry is to argue that the leftward DP-movement to Zone 1 and Zone 2 is A-movement, i.e. movement to the specifier of a functional head with Agr features. PPs undergo optional $\overline{A}$-movement. I tentatively assume that PPs adjoin to functional projections in Zone 2 or in Zone 1.

A third argument for interpreting the leftward movement of the DPs to Zone 1 and Zone 2 in terms of A-movement is that it contrasts sharply with uncontroversial $\overline{A}$-movement. In sentences with $\overline{A}$-movement of a DP complement the rigid ordering constraint does not apply: the DObj may move to Spec-CP and precede the IObj.

---

7. An anonymous reviewer points out that sequences such as text example (14a) improve in Dutch when a particle is associated with the verb:
(i) *??dat Jan dat boek Marie teruggeeft*
that Jan that book Marie back-gives
'that Jan returns that book to Marie.'
No such improvement is found in WF.

(16) [$_{CP}$ *Dienen boek* [$_{C0}$ *ee* [$_{IP}$ *Valère Marie gegeven*]]]
that book has Valère Marie given
'That book, Valère gave to Marie.'

A final argument in support of interpreting the leftward DP-movement as A-movement is that it does not licence parasitic gaps. In (17a) the direct object has moved to Zone 1 and the parasitic gap in the adjoined clause is not licensed. In (17b) movement is only to Zone 2, again parasitic gaps are not licensed.

(17) a. **da Valère die brieven verzekerst* [*zonder* *(*ze*) *te*
that Valère those letters probably without them to
overlezen] *ipgestierd een*
re-read sent has
b. **da Valère verzekerst die brieven ipgestierd een* [*zonder* *(*ze*) *te*
overlezen]

Uncontroversial instances of $\overline{A}$-movement, such as movement to Spec-CP, (marginally) license parasitic gaps in WF: in (18a) the preposed *wh*-phrase *wavuonen brief* licenses a parasitic gap in the adjunct clause; in (18b) the topicalized DP *dienen brief* licenses a parasitic gap.

(18) a. ?*Wavuonen brief ee Valère zonder te overlezen ipgestierd?*
which letter has Valère without to reread sent
'Which letter did Valère mail without rereading?'
b. ?*Dienen brief ee Valère zonder te overlezen ipgestierd*
this letter has Valère without to re-read sent
'This letter, Valère sent without rereading it.'

2.2.3. *A-movement and Feature Checking*

I propose that DP arguments undergo obligatory movement to Zone 2 and what seems to be optional movement to Zone 1. Concerning the obligatory movement of the DPs to Zone 2, Vanden Wyngaerd (1989a) proposes that this is Case-driven A-movement.[8] I propose that the movement to Zone 1 is not optional, but that it is triggered by the presence of a feature ⟨R⟩, associated with

8. Observe, though, that on a universal SVO account, Object Shift in WF (and in other Westgermanic languages) does not interact with V-movement. There is no obvious way in which the movement of both object DP to Spec-AgrOP and of the subject DP to Spec-AgrSP is made possible by the leftward movement of the finite V to AgrO and AgrS. I will not address this issue here.

the functional projections in Zone 1. Given the discussion in Section 2.2.2 I assume that movement to Zone 1 is also A-movement, and that R is a *phi* feature. One interpretation of this would be to relate R to the specificity interpretation of DPs (cf. Cecchetto 1994).

As a first approximation one might propose that when indirect object and direct object both move to Zone 1 they check the R-feature with respect to one functional head bearing R. Conceivably the direct object could occupy the specifier of the projection and be narrowly related to the relevant head, while the indirect object might be adjoined to the functional projection (cf. Ura 1993). However, we will see in Section 3 that the distribution of clitics in WF suggests that Zone 1 must consist of a number of discrete functional projections.

On the basis of the distribution of clitics we will also see that the functional heads in Zone 1 cannot be landing sites for V^0-movement, hence the heads of these projections do not carry V-features. This means that the functional projections whose specifiers host the DPs in Zone 1, and whose heads I will argue host clitics, correspond to the missing projections in (8). Their heads carry N-features but no V-features: they are non-L-related heads with A-specifiers.

3. The Distribution of Clitics in WF

3.1. *Clitics Move to Zone 1*

I have suggested that two domains for leftward DP-movement are to be demarcated: Zone 1, to the left of the adverbial, and Zone 2, to the left of sentential negation. Definite object DPs must appear to the left of the negative marker *nie*; they need not precede the adverbial. Object clitics, on the other hand, obligatorily move to Zone 1: they must precede adverbials. This is shown in (19). The clitic must not follow the adverb *misschien* 'perhaps' (19a), it may appear either to the right (19b) or to the left (19c) of the subject DP *Valère*.

(19) a. **da Valère misschien ze kent*
that Valère perhaps them knows

b. *da Valère ze misschien kent*
that Valère them perhaps knows
'that Valère perhaps knows them'

c. *da ze Valère misschien kent*
that them Valère perhaps knows
'that Valère perhaps knows them.'

I will interpret the ungrammaticality of (19a) in terms of the structure outlined: I assume that clitics cannot remain in Zone 2 and must move to Zone 1.[9]

The assumption that clitics move to Zone 1 has further consequences for the structure of this domain. If the WF clitics are heads then their distribution is evidence that Zone 1 must be composed of discrete functional projections whose heads host the clitics. In the remainder of this section I will describe the distribution of clitic elements — *t* 'it', *ze* 'them', 'her' and *der* 'some' — in WF in more detail.[10] On the basis of a comparison with the syntax of lexical DPs I will provide evidence that they are indeed heads.

3.2. *The Categorial Status of Clitics: The Case of Romance*

In this section I consider the categorial status of clitics. First I will briefly go over the properties of Romance clitics. This will be the basis for the discussion of the WF clitics in Section 3.3.

3.2.1. *Romance Clitics as Heads*

The traditional assumption is that the landing site for cliticization in Romance is a V-related functional head (Belletti 1993; Kayne 1989c, 1991; Rizzi 1994). In finite clauses in French, for instance, the clitics move as high as the finite V; in (20a) the clitic *les* 'them' and the finite auxiliary *ai* 'have' are hosted by AgrS, as can be seen from the fact that they precede *pas*, the specifier

---

9. Observe also that it would not be obvious how to account for the distribution of clitics in (19) if one were to adopt an approach in which Zone 1 and Zone 2 are collapsed into one domain.

10. For more detailed description of WF clitics and a comparison with other pronominal elements see also Haegeman (1993a, 1993b, 1993c). I consider *t, ze* and *der* as clitics. Their behavior is different from that of unstressed *men* 'me' and *jen* 'you', which are arguably weak pronouns in the sense of Cardinaletti & Starke (1993). I assume that the analysis in terms of clitics carries over to Dutch *t* 'it', *er* 'some' and *ze* 'them' and to German *es* 'it'. For recent discussion of Germanic clitics cf. among others, Cardinaletti (1992a, 1992b), Holmberg (1991), Joseffson (1992).

of NegP. In (20b) the clitic *les* 'them' moves to C^0 along with the inflected auxiliary *as* 'have'.

(20) a. *Je ne les ai pas invités à la fête*
I *ne* them have not invited to the party
'I did not invite them to the party.'

b. *Pourquoi ne les as-tu pas invités à la fête?*
why *ne* them have you not invited to the party
'Why did you not invite them to the party?'

Evidence that the verb and the clitic form a syntactic unit is that the clitic is carried along by the Aux-to-C movement in (20b). Under standard approaches (20c) will be excluded by the Head Movement Constraint (HMC) (Travis 1984) which reduces to the ECP.

(20) c. **Pourquoi n'as-tu les pas invités à la fête?*
why *ne* have you them not invited to the party

The example in (21), due to Kayne (1989c: 240), illustrates a free-standing clitic. This construction is not grammatical in contemporary French but it existed in earlier stages of French and is also found in dialects (cf. Kayne 1989c: 240 for discussion, Lehmann in preparation). The clitic *les* has moved to AgrS, while the infinitival verb *faire* need not check its Agr features in the syntax and remains in a lower position.

(21) **Jean a promis de les bien faire*
Jean has promised for/to them well do
'Jean has promised to do them well.'

As (21) shows, the clitic and V^0 are not always adjacent in Romance. Observe, however, that the pattern in (22) where the verb precedes the clitic and is separated from it by intervening maximal projections is not attested in Romance (see example (20c)).

(22) $*V^0_i$ XP cl t_i

As mentioned, the ungrammaticality of (22) follows from the HMC: we assume that the clitic is hosted by a V-related functional head, i.e. a functional head with V-features. In order to reach the position to the left of the clitic in (22) V^0 would have to skip the relevant head, hence it would fail to check its features in the head which hosts the clitic. Moreover, in the representation in (22) the functional head hosting the clitic will block the connection between V^0 and its trace. I assume that excorporation of the head is not possible.

3.2.2. *Clitic-movement: DP-movement and Head-movement*

Given the past-participle agreement in examples such as (23) where the past participle *vus* 'seen' agrees with the direct object clitic for gender and number features, it is standardly assumed that the clitic first moves as a maximal projection, triggering agreement with the participial head (cf. Kayne 1989a).[11]

(23) *Je les ai vus*
I them have seen+PL
'I have seen them.'

Then the head of the DP, the clitic, is extracted and adjoins to the higher functional head, which has both N-features and V-features. The presence of V-features on the head which hosts the clitic means that V^0 must also check its (head) features. V^0 and the clitic will merge into a complex head and subsequently they will move together.

3.3. *West Flemish Clitics*

In Section 3.2 I have summarized the view on clitic-movement in Romance (Kayne 1989c) which proposes that Romance clitics first move as DPs and then their head is extracted and adjoins to a V-related functional head. I now turn to the distribution of WF clitics. I will propose that like Romance clitics, WF clitics are syntactic heads (3.3.1). Like their Romance counterparts they undergo a two step movement: first the DP containing the clitic moves to a specifier of a functional head and then the clitic is extracted and adjoins to a head (Section 3.3.2).

3.3.1. *WF Clitics as Heads*

As already has been discussed briefly in Section 3.1, object clitics in WF do not have the same distribution as lexical DPs. The difference in distribution between lexical DPs and clitic arguments suggests that the latter be treated as X^0 elements while the former are maximal projections. I will compare the distribution of lexical DPs and of clitics more closely in this section.

---

11. In Sportiche's (1992) account clitics are base generated as functional heads of so called Clitic Projections and a non-overt DP, *pro*, moves to their specifier. It is the movement of *pro*, i.e. a DP, which triggers participial agreement. I will not examine this alternative here. For some problems with this account see Haegeman (1994).

We have seen that WF clitics must move higher than lexical DPs. As already seen in (19) lexical DP arguments can remain in Zone 2 and may move to the higher Zone 1; clitics must not remain in Zone 2. This is also illustrated in (24).

(24) a. *da Valère misschien die boeken/*ze kent*
that Valère perhaps those books them knows
'that Valère perhaps knows those books'
b. *da Valère die boeken/ze misschien kent*
that Valère those books them perhaps knows
'that Valère perhaps knows those books/them'

A second contrast between clitics and lexical DPs concerns the relative ordering of the arguments. Recall that in WF, there is a rigid ordering constraint on lexical DP arguments, imposing the sequence Subj–IObj–DObj. Clitic arguments have additional options. While a lexical IObj DP cannot precede the subject, the IObj clitic *ze* can precede the subject DP.

(25) a. **da Marie Valère die boeken gegeven eet*
that Marie (IObj) Valère (Subj) those books given has
b. *da ze Valère die boeken gegeven eet*
that her Valère those books given has
'that Valère has given her those books.'

The examples in (25b) show that the indirect object clitic *ze* 'her' can undergo additional leftward movement which is not available to the lexical indirect object DP *Marie* in (25a).

Lexical direct objects cannot precede the indirect object or the subject:

(26) a. **da Valère die boeken Marie gegeven eet*
that Valère those books Marie given has
b. **da die boeken Valère Marie gegeven eet*
that those books Valère Marie given has

Direct object clitics may precede both the indirect object and the subject:

(27) a. *da Valère ze Marie gegeven eet*
that Valère them Marie given has
'that Valère gave them to Marie.'
b. *da ze Valère Marie gegeven eet*
that them Valère Marie given has
'that Valère gave them to Marie.'

Again the leftward movement past the indirect object and the subject is available only for object clitics and is unavailable for lexical DPs.

One way of accounting for the distribution of *ze* in (25b) and in (27) is to assume that the leftward movement which distinguishes the clitic argument from the lexical DP argument be interpreted as X^0-movement. X^0-movement is not available to lexical DPs. Leftward X^0-movement does not interact with the DP positions.

If we consider elements like *ze* as clitics, we are led to the conclusion that WF clitics (and Germanic clitics in more general terms) cannot be "verbal clitics": they are not hosted by a V-related functional head. The independence of clitics and V^0 in WF is confirmed by the distribution of object clitics in root clauses: the clitic is not carried along by the V-to-C movement.

(28) a. *Woarom geeft Valère Marie ze?*
why gives Valère Marie them
'Why does Valère give them to Marie?'

b. *Woarom geef-je gie ze dienen boek?*
why give-you you her that book
'Why do you give her that book?'

The examples in (28) contrast with French where V-movement to C takes the clitic along (29a). The sentence in (29b), where the clitic is stranded after V-to-C movement, is ungrammatical.

(29) a. *Pourquoi les donnes-tu à Marie?*
why them give-you to Marie
'Why do you give them to Marie?'

b. **Pourquoi donnes-tu les à Marie?*
why give-you them to Marie

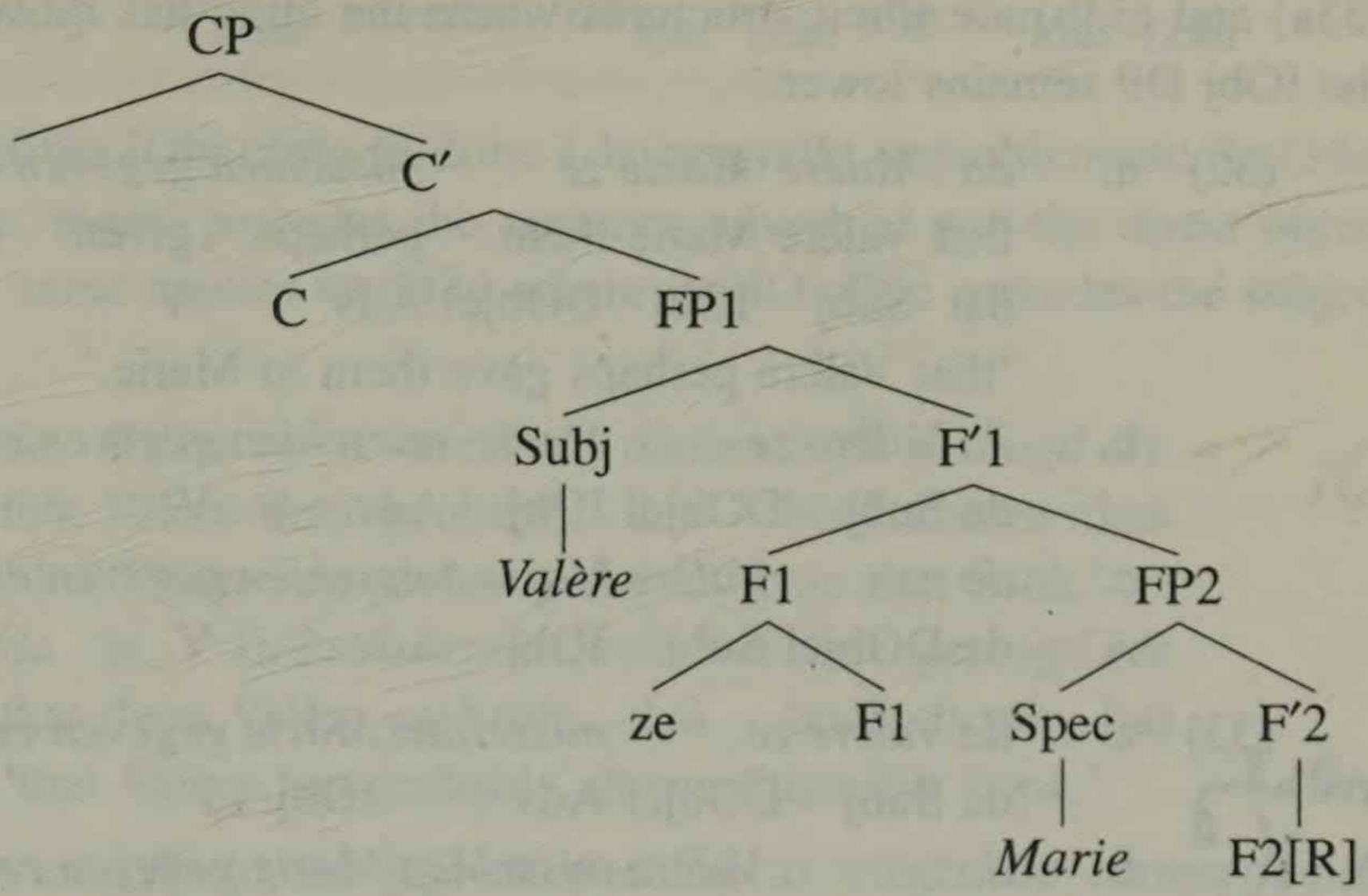
The question arises how to account for the fact that the WF (i.e. Germanic) clitic is not hosted by a verb in general and that V-movement strands the clitics, leading to a representation as in (30), without apparent HMC violation (cf. for contrast (22)).

(30) OK Ger V^0_i XP cl t_i

I propose that WF clitics are like Romance clitics in that they adjoin to a functional head. As a first approximation, to be modified below, (31a) would have the partial structure in (31b).

(31) a. *dat Valère ze Marie geeft*
that Valère them Marie gives
'that Valère gives them to Marie.'

b.



The indirect-object DP *Marie* occupies the specifier of a functional projection, FP2, whose head hosts *phi* features. The subject DP *Valère* occupies the specifier of the highest functional projection, FP1. The direct-object clitic *ze* left-adjoins to the head of FP1, F1.

Romance clitics are hosted by a V-related functional head carrying both N-features and V- features. Since V-movement does not take along the clitic in WF, we conclude that the functional head which hosts the WF clitics does not carry V-features. As a consequence, WF V-movement to C will not interact with clitic-movement.

3.3.2. *WF Clitic-movement: DP-movement and Head-movement*

In the previous section we have seen that WF clitics undergo leftward movement which is unavailable to lexical DPs. We propose that this movement be interpreted in terms of X^0-movement. In this section we will see that to some extent, though, clitic movement *is* dependent of lexical DP-movement. The observed interdependence of clitic movement and DP-movement suggests that clitics do not undergo head-to-head movement directly from the argument position. Rather, following Kayne's (1989c) analysis for French, we will conclude that the WF clitics are DPs which first move as DPs and then undergo head-to-head movement.

In (32) and (33) the movement of the direct object clitic to a position to the left of the adverbial is shown to be dependent on the movement of indirect object DP. Examples (32a)–(32c) illustrate the grammatical patterns. Examples

(33a) and (33b) are illicit structures where the clitic has moved to Zone 1 but the IObj DP remains lower:

(32) a. *da Valère Marie ze misschien gegeven eet*
that Valère Marie them perhaps given has
da Subj IObj DObjcl Adv V
'that Valère perhaps gave them to Marie.'

b. *da Valère ze Marie misschien gegeven eet*
da Subj DObjcl IObj Adv V

c. *da ze Valère Marie misschien gegeven eet*
da DObjcl Subj IObj Adv V

(33) a. **da Valère ze misschien Marie gegeven eet*
da Subj DObjcl Adv IObj V

b. **da ze Valère misschien Marie gegeven eet*
da DObjcl Subj Adv IObj V

Clitic movement to Zone 1 does not depend on PP-movement. In (34) the clitic DObj has moved to Zone 1 and the IObj, realized as a PP, remains lower:

(34) a. *da Valère ze misschien an Marie gegeven eet*
that Valère them perhaps to Marie given has
'that Valère perhaps has given them to Marie'

b. *da ze Valère misschien an Marie gegeven eet*
that them Valère perhaps to Marie given has

As a descriptive generalization we can say that the DObj clitic can only cross the lexical IObj DP once both IObj and DObj have reached Zone 1. This pattern will follow if the clitic first moves as a DP to Zone 1, and undergoes head-to-head-movement once it has reached Zone 1. I therefore propose that WF clitics move leftward as DPs via Zone 2 to reach Zone 1. Both the movement to Zone 2 and that to Zone 1 respect the conditions of DP-movement described above, specifically the ordering of IObj–DObj has to be respected. In the illicit examples in (33) where a DObj clitic precedes an adverbial and the lexical IObj DP follows the adverbial, the clitic would have to move to Zone 1 crossing the IObj DP. If we assume that the movement of the clitic from Zone 2 to Zone 1 is XP-movement, the crossing of the IObj DP by the clitic will result in a representation which violates the rigid ordering constraint on arguments discussed above. (35) would be the intermediate representation for the examples in (33), where the clitic has moved as DP to Zone 1. The sequence DP_{DOBJ}–DP_{IOBJ} is ungrammatical.

(35) *C Subj DP_{DOBJ} adverb DP_{IOBJ} t_{DOBJ} *niet* V t_{IOBJ} t_{DOBJ}

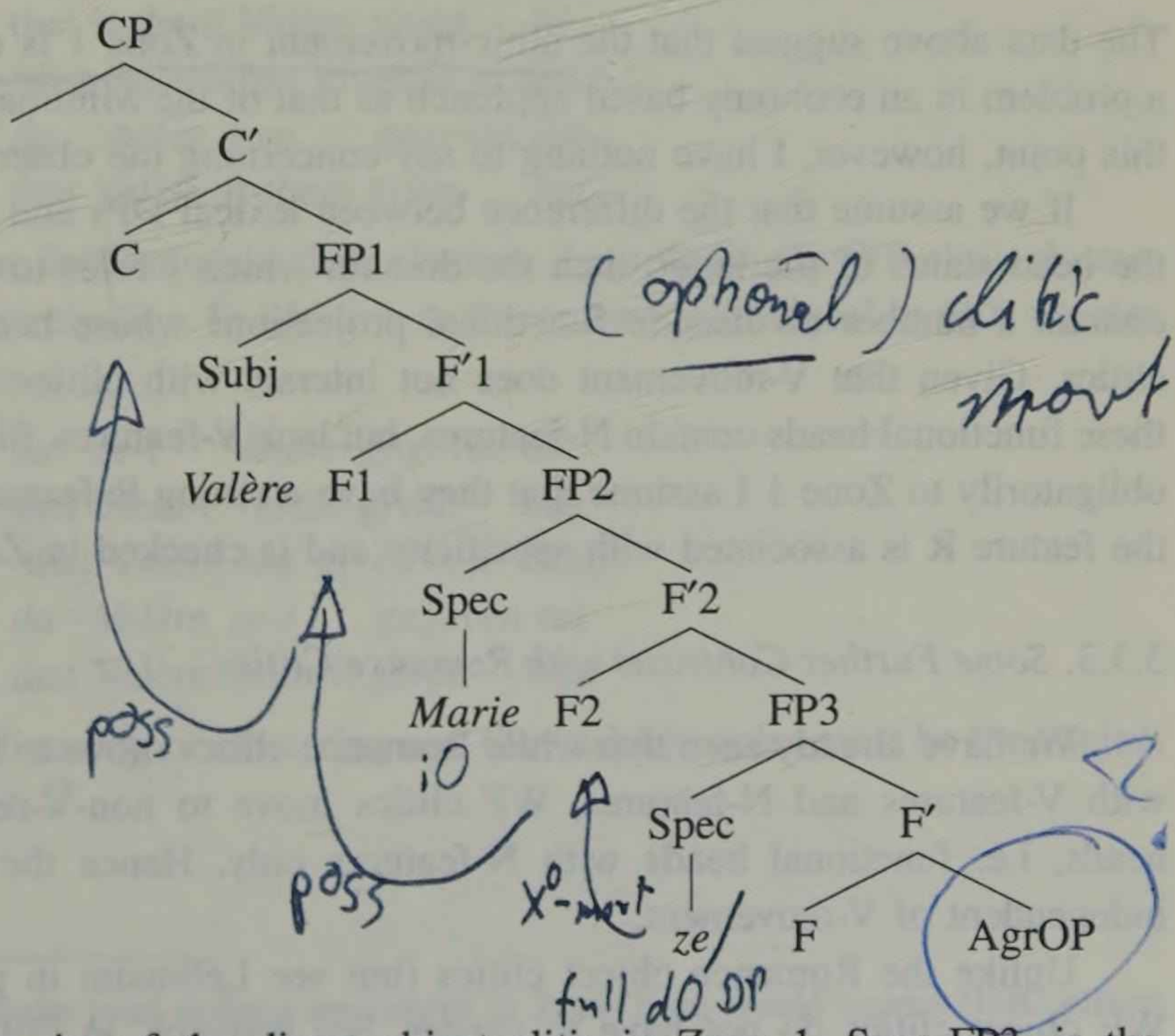
The movement of the IObj clitic to Zone 1 is generally unproblematic. In (36a) the IObj clitic *ze* 'them' precedes the sentence adverbial and the direct object DP follows. The same applies to (36b) where the IO clitic precedes the subject DP.

(36) a. *da Valère ze verzekerst dienen boek getoogd eet*
that Valère them probably that book shown has
'that Valère has probably shown them that book.'

b. *da ze Valère verzekerst dienen boek getoogd eet*
that them Valère probably that book shown has
'that Valère has probably shown them that book.'

WF clitics move as DPs to the highest position to which the corresponding lexical DP would move. A DObj clitic, for instance, will move to Spec-AgrOP, i.e. into Zone 2; then it will move to the A-position in Zone 1 which the corresponding lexical DObj DP would occupy (37). Clitics cannot remain in the Spec-AgrOP in Zone 2.

(37)



The landing site of the direct object clitic in Zone 1, Spec-FP3, is the landing site of a direct object lexical DP. It is also the launching site for

head-to-head-movement. We assume that the clitic is the head of the DP. This head, i.e. the clitic, may undergo head-to-head movement and adjoin to F2, the head whose specifier hosts the IObj. In this case there will be no reordering:

(38) a. *da Valère Marie ze gegeven eet*
that Valère Marie them given has
'that Valère has given them to Marie.'

Alternatively, the clitic may undergo further head-to-head movement and adjoin to F1, the head whose specifier hosts the subject. In this case the direct object clitic reorders with respect to the indirect object lexical DP.

(38) b. *da Valère ze Marie gegeven eet*
that Valère them Marie given has
'that Valère has given them to Marie.'

When the clitic precedes the subject I assume it moves further to C:

(38) c. *da ze Valère Marie gegeven eet*
that them Valère Marie given has
'that Valère has given them to Marie.'

The data above suggest that the clitic-movement in Zone 1 is optional. This is a problem in an economy based approach as that of the Minimalist Program. At this point, however, I have nothing to say concerning the observed optionality.

If we assume that the difference between lexical DPs and clitics relates to the head status of the latter, then the domain which I refer to as Zone 1 must contain a number of discrete functional projections whose heads can host the clitics. Given that V-movement does not interact with clitics we assume that these functional heads contain N-features, but lack V-features. Since clitics move obligatorily to Zone 1 I assume that they have a strong R-feature R. Recall that the feature R is associated with specificity and is checked in Zone 1.

3.3.3. *Some Further Contrasts with Romance Clitics*

We have already seen that while Romance clitics move to functional heads with V-features and N-features, WF clitics move to non-V-related functional heads, i.e. functional heads with N-features only. Hence the WF clitics are independent of V-movement.

Unlike the Romance object clitics (but see Lehmann in preparation), the WF object clitics do not have to cluster. For instance, in (39) the IObj clitic precedes the subject, but the DObj clitic follows it.

(39) a. *da ze Valère z' a gegeven eet*
that her Valère them already given has
'that Valère has already given them to her.'

Consider also (39b) where the perception verb *zien* 'see' takes a non-finite complement.

(39) b. *da Valère ze Marie* t *gisteren nie zien geven eet*
that Valère them Marie it yesterday not see give has
'that Valère did not see Marie give it to them yesterday.'

In (39b), the subject of the complement clause, *Marie*, has moved to the matrix domain. This is seen from the fact that it precedes the matrix negation (see also Haegeman 1993a). In (39b) I assume *ze* is hosted by the head whose specifier contains the subject *Valère*; *t* is hosted by the head whose specifier hosts the embedded subject *Marie*. From the split clitic patterns in (39) we conclude that Zone 1 must be composed of discrete projections whose heads can host clitics.

Though WF has the option of split clitics, clustering is also possible, and this in several positions, as illustrated in (40):

(40) a. *da t-ze Valère gegeven eet*
that it-them Valère given has
'that Valère has given it to them.'

b. *da Valère t-ze gegeven eet*
that Valère it-them given has

Unlike is the case for Romance clitic clusters, the order in the WF clitic clusters seems to be relatively free. In addition to the order *t-ze* 'it-them' in (40) we also have the order *ze-t*:

(41) a. *da ze-t Valère gegeven eet*
that them-t Valère given has
'that Valère has given it to them.'

b. *da Valère ze-t gegeven eet*
that Valère them-it given has

The apparent optionality of clustering and the variable order must be the subject of future research.[12]

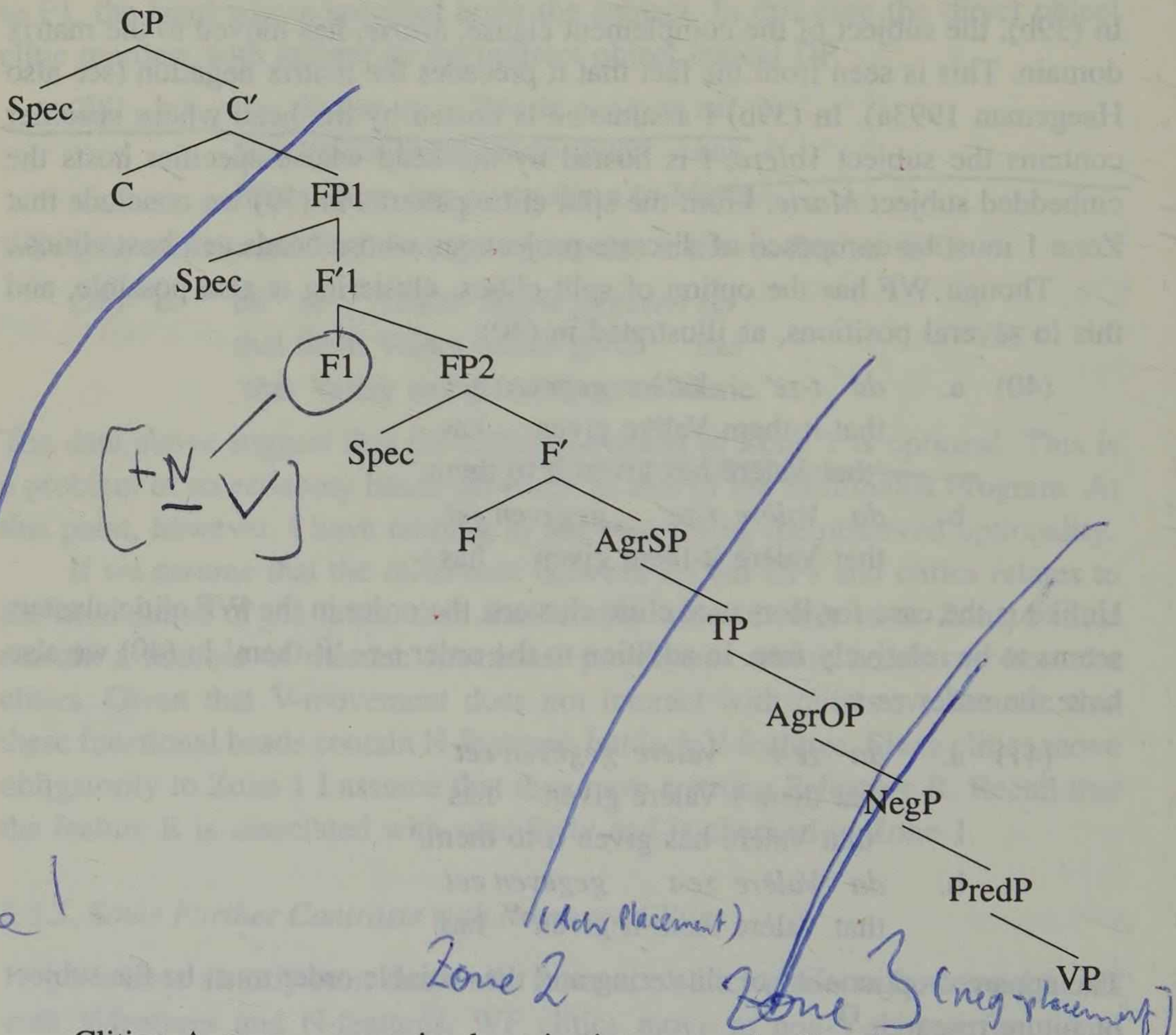
---

12. If clitics undergo head-to-head-movement in Zone 1 we would expect HMC effects when more than one clitic moves. This is indeed the case, as shown in detail in Haegeman (in press).

3.4. *Zone 1*

In addition to the familiar projections such as CP, AgrSP, TP, NegP, AgrOP, this paper argues that the WF clause contains functional projections which dominate AgrSP and whose heads host a feature R. R is assumed to be a *phi* feature. DPs with an R feature must move to the specifier of the relevant functional projections to check the feature.

(42)



Clitics A-move as maximal projections, exactly like the corresponding lexical DP arguments. A direct object DP whose head is a clitic will first undergo DP-movement to Spec-AgrOP, where its case is licensed. Then the DP will move to the specifier of the functional projection containing the R-feature. The latter movement is identical to that of a lexical direct object with R-feature.

The specifier of the R-projection which is the landing site of the direct object DP then serves as the launching site for head-movement. The clitic, which is the head of the DP, is extracted and undergoes leftward head-to-head movement.

The functional R-projections in Zone 1 are not V-related. R is an N-feature, it is not a V-feature. As a result, finite V^0-movement to C^0 will by-pass the functional heads in Zone 1 since the latter are not V-related. On the other hand, because ⟨R⟩ is a N-feature, the specifiers of the R-projections are A-positions. The functional projections in Zone 1 are then evidence for the dissociation of the A/$\overline{\text{A}}$-distinction and L-relatedness.

4. Conclusion

The theoretical focus of this paper is the proposed reinterpretation of the A/$\overline{\text{A}}$-distinction in terms of L-relatedness. I have tried to show that the A/$\overline{\text{A}}$-distinction should be maintained, independently of the notion of L-relatedness. I have provided empirical arguments for postulating a functional projection whose head is non-L-related and whose specifier is an A-position.

The empirical focus of the paper is the contrast between Germanic clitics and Romance clitics. Unlike their Romance counterparts, Germanic clitics are nonverbal clitics; they move to a non-L-related functional head. Romance clitics are verbal clitics which move to an L-related (V-related) functional head. As a consequence, only Romance clitics interact with V-movement.

Acknowledgements

Most of the contents of this paper has been the basis of my DES course at the University of Geneva. Part of the material of this paper was presented at the CNRS comparative syntax round table in Paris (December 1993) , at a seminar at the Autonoma in Barcelona and at the 3ième Cycle conference at Neuchâtel (May 1994). I thank the students of the seminars and the participants of these various meetings for their comments on previous versions of this paper. Special thanks are due to Carlo Cecchetto, Jan-Wouter Zwart and three anonymous reviewers for their comments. Needless to say, I am solely responsible for remaining errors.

Spec, $\overline{TP}$ can be present, and
is where V^0 overtly raises to T^0.

Spec, TP is a case checking position
for subj. in finite clauses

Spec, $\overline{TP}$ for indefs
Spec, AgrSP for defs.

Clause Structure, Expletives and Verb Movement

Dianne Jonas
Yale University

We now regard both agreement and structural Case as manifestations of the Spec-head relation (NP, Agr). But Case properties depend on characteristics of T and the V head of VP. We therefore assume that T raises to Agr_S, forming $[_{Agr}$ T Agr], and V raises to AgrO, forming $[_{Agr}$ V Agr]; the complex includes the phi-features of Agr and the Case feature provided by T,V. [fn 11: I overlook here the possibility of NP-raising to [Spec, T] for Case assignment, then to [Spec, Agr_S] for agreement. This may well be a real option...]

— Chomsky (1993: 7–8)

Abstract

The above quote suggests that there may be two distinct positions in which the Case feature of the subject may be checked: Spec-AgrSP, following raising of tense to AgrS, or Spec-TP. In this paper I examine data from Scandinavian, and English to some extent, and argue that both options are made available by the theory but they are not available in all languages. I present an analysis which examines the interaction between verb movement and subject Case-checking and claim that Spec-TP is only available in languages with overt verb raising to T. In those languages Spec-TP is available as a Case checking position for the subject in finite clauses and also as an intermediate position in non-finite clauses (complements of raising verbs and ECM [Accusative with Inifinitive] verbs) where the subject can check the EPP feature of a non-finite clause but not check its Case. In this sense the paper suggests a refinement of checking theory with respect to subjects.

1. Introduction

It is usually assumed that there is a clear contrast between languages which permit TRANSITIVE EXPLETIVE CONSTRUCTIONS (TECs), such as Icelandic and

Faroese,[1] and those which do not, such as the Mainland Scandinavian languages. This contrast is illustrated in (1a) and (1b).

(1) a. *Það hafa margir jólasveinar borðað búðinginn.* Ice.
there have many Christmas-trolls eaten the pudding
'Many Christmas trolls have eaten the pudding.'

b. **Det har många män ätit puddingen.* Sw.
there have many men eaten the pudding

Jonas & Bobaljik (1993, henceforth J&B) and Jonas (1994) have argued that the position of the indefinite subject in the Icelandic example (1a) is Spec-TP in an articulated INFL structure. One prediction of the Minimalist Program as outlined in Chomsky (1993, henceforth *MPLT*) is that Spec-TP is an overt position for the subject only if the finite verb has raised overtly and adjoined to T^o. *MPLT* predicts that verb movement to T is necessary for the subject to move to Spec-TP following V-to-AgrO and object shift to Spec-AgrOP.[2] The subject can then raise to Spec-TP but only if there is movement of the $[_{AgrO}$ V AgrO] complex to T. This is entailed by the principle of EQUIDISTANCE which will be discussed in Section 4.1.

To put it differently, the claim made here is that TECs should only be possible in languages where Spec-TP is available as a subject position and, furthermore, these should be the languages where there is overt verb movement to T. But if this is correct, we would expect to find other evidence for Spec-TP as a subject position in TEC languages. Second, we would expect to find evidence for overt V-to-T in these languages. With respect to Spec-TP as an "extra" subject position, it is argued in Jonas (1994, 1995) that Spec-TP is the structural position of the indefinite subject in the Icelandic and Faroese finite and non-finite clauses in (2) (while definite subjects typically end up in Spec-AgrSP).

(2) a. TECs

Það byggðu $[_{TP}$ *margir Íslendingar*$_i$ $[_{VP}$ t_i *hús í Þórshöfn*]] Ice.

Tað bygdu $[_{TP}$ *nógvir íslendingar*$_i$ $[_{VP}$ t_i *hús í Havn*]] Far.

there built many Icelanders houses in Tórshavn

---

1. Not all speakers of Faroese accept TECs, see Jonas (1995) for discussion of this.
2. I assume that overt shift of full NP-objects in Icelandic is movement to the Spec–AgrOP position (see, among others, Vanden Wyngaerd 1989a; Déprez 1989; Chomsky 1993; Bures 1993; Jonas & Bobaljik 1993).

b. Passive Expletive Constructions

Það hafa [$_{TP}$ *nokkrar kökur*$_i$ [*verið* [*bakaðar* t$_i$
Tað hava [$_{TP}$ *nakrar lagkøkur*$_i$ [*verið* [*bakaðar* t$_i$
there have some cakes been baked
fyrir veisluna]]]
til veitsluna]]] Ice.
for the party Far.

c. Raising Expletive Constructions

Það virðast [$_{TP}$ *margir menn*$_i$ [$_{VP}$ *vera* t$_i$ *í herberginu*]] Ice.
Tað tykast [$_{TP}$ *nógvir menn*$_i$ [$_{VP}$ *vera* t$_i$ *á kamarinum*]] Far.
there seem many men be.INF in the room

These constructions (with the subject in Spec-TP) are not possible in Mainland Scandinavian (MSc) as shown in (3).[3]

(3) a. TECs

**Det angriper en bjørn en sau.*
there attacks a bear a sheep (No., Åfarli 1992: 105)

b. Passive Expletive Constructions

**Det har noen kaker blitt bakt for selskapet.* No.
there have some cakes been baked for the party

c. Raising Expletive Constructions

**Det verkar en man vara i rummet.* Sw.
there seems a man be.INF in the room

The goal of this paper is, as stated above, to account for data of the sort exemplified in (1–3). The proposal here is that the availability of Spec-TP in a TEC language is dependent upon the presence of overt movement of the finite verb. The verb movement that is concerned here is V-to-T (possibly as part of V-to-AgrS) rather than verb movement to "second" position, whatever that position is argued to be, in main clauses.[4]

---

3. In the following sections, I will use examples from Norwegian, Swedish and Danish rather indiscriminantly to illustrate MSc. I am not aware of any differences between these languages that are relevant to the points I make in this paper so no particular significance should be attached to my particular choice of language for any given example.

4. By this I am not subscribing to the common claim (see e.g. Vikner 1994, 1995a and references cited there) that "second position" is necessarily always the "same" position (like C, for instance) as I argue elsewhere (Jonas 1995). This point is also argued for

The paper is organized as follows: in Section 2, I give evidence from Icelandic supporting an articulated Infl structure. In Section 3, I discuss the verb movement facts in a subset of the Scandinavian languages. In Section 4, I discuss subject positions in various expletive constructions in Scandinavian and the interaction of verb movement, and subject licensing and Case-checking. Section 5 presents the conclusion.

2. Clause Structure

Direct evidence supporting an articulated Infl structure (as opposed to, say, a non-split IP of the pre-Pollockian kind; see Pollock 1989, which Thráinsson, this volume, claims is appropriate for English and MSc) is provided by Icelandic expletive constructions. Icelandic allows overt full NP OBJECT SHIFT (NP-OS) to Spec-AgrOP. When NP-OS occurs in a TEC, the indefinite subject must normally raise overtly, as shown by the contrast between (4a) and (4b):[5]

(4) a. ?*[$_{AgrSP}$ *Það* *málaði*$_j$ [$_{AgrOP}$ *húsið*$_k$ [$_{VP}$ *stundum* [$_{VP}$ *útlendingur*
there painted the house sometimes foreigner
Ice.
t$_j$ t$_k$ *rautt*]]]]
red

b. [$_{AgrSP}$ *Það* *málaði*$_j$ [$_{TP}$ *útlendingur*$_i$ [$_{AgrOP}$ *húsið*$_k$ [$_{VP}$
there painted foreigner the house
stundum t$_j$ t$_k$[$_{VP}$ t$_i$ *rautt*]]]]]
sometimes red
'A foreigner sometimes painted the house red.'

---

by Zwart (1993b), for instance. The only point I want to make here is that verb movement in main clauses (and main-clause-like clauses) may in some instances at least be triggered by other factors than verb movement in embedded clauses and I am not concerned with these other factors here. Rather, I am mostly interested in the verb movement that is independent of these special main clause forces.

5. As discussed in Jonas 1995, Chapter 2, quantified indefinite subjects appear to be able to remain in Spec-VP, witness examples of the following type:
(i) *Það* *máluðu húsið* *aldrei neinir stúdentar rautt*
there painted the house never any students red
Detailed discussion of the conditions under which subjects can remain in Spec-VP is outside the scope of this paper, but see also note 6.

In (4a), the finite verb has raised and adjoined to AgrS (via AgrO° and T°) and the object has moved overtly to Spec-AgrOP to the left of the VP adjoined adverb *stundum* 'sometimes'. Example (4b) shows the acceptable structure where the subject has moved out of the VP to Spec-TP.

As indicated in (4), I am assuming here that the expletive occupies Spec-AgrSP rather than, say, Spec-CP. The arguments for this include the fact that the expletive is very different in nature from typical Spec-CP elements, as pointed out by Ottósson, for instance (1989). In addition, I argue that the expletive has a special checking role to fulfill in Spec-AgrSP, namely that it checks a definiteness feature (a D-feature, see Section 4.4 below and Jonas 1995, Chapter 2). Furthermore I make the standard assumption here that *stundum* 'sometimes' is of the class of "medial adverbs" that are adjoined to VP. Although adverbs are notoriously difficult to pin down, it is clear that *stundum* 'sometimes', *ekki* 'not', *aldrei* 'never' and other adverbs of this class cannot occur VP-finally as shown in sentences with auxiliary verbs which do not allow object shift (Holmberg's Generalization, cf. Holmberg 1986):

(5) **Strákarnir hafa málað húsið ekki/aldrei/stundum.* Ice.
the boys have painted the house not never sometimes

We see therefore that *ekki* 'not', *aldrei* 'never', *stundum* 'sometimes' and other adverbs of the same kind cannot be adjoined lower than to VP. Thus if a shifted object precedes these adverbs, as in (4), it must have shifted out of the VP, presumably to Spec-AgrOP. These are the crucial points about adverbial position that need to be made. It is irrelevant for my purposes whether these adverbs can sometimes adjoin to a higher projection such as TP, for instance.[6]

---

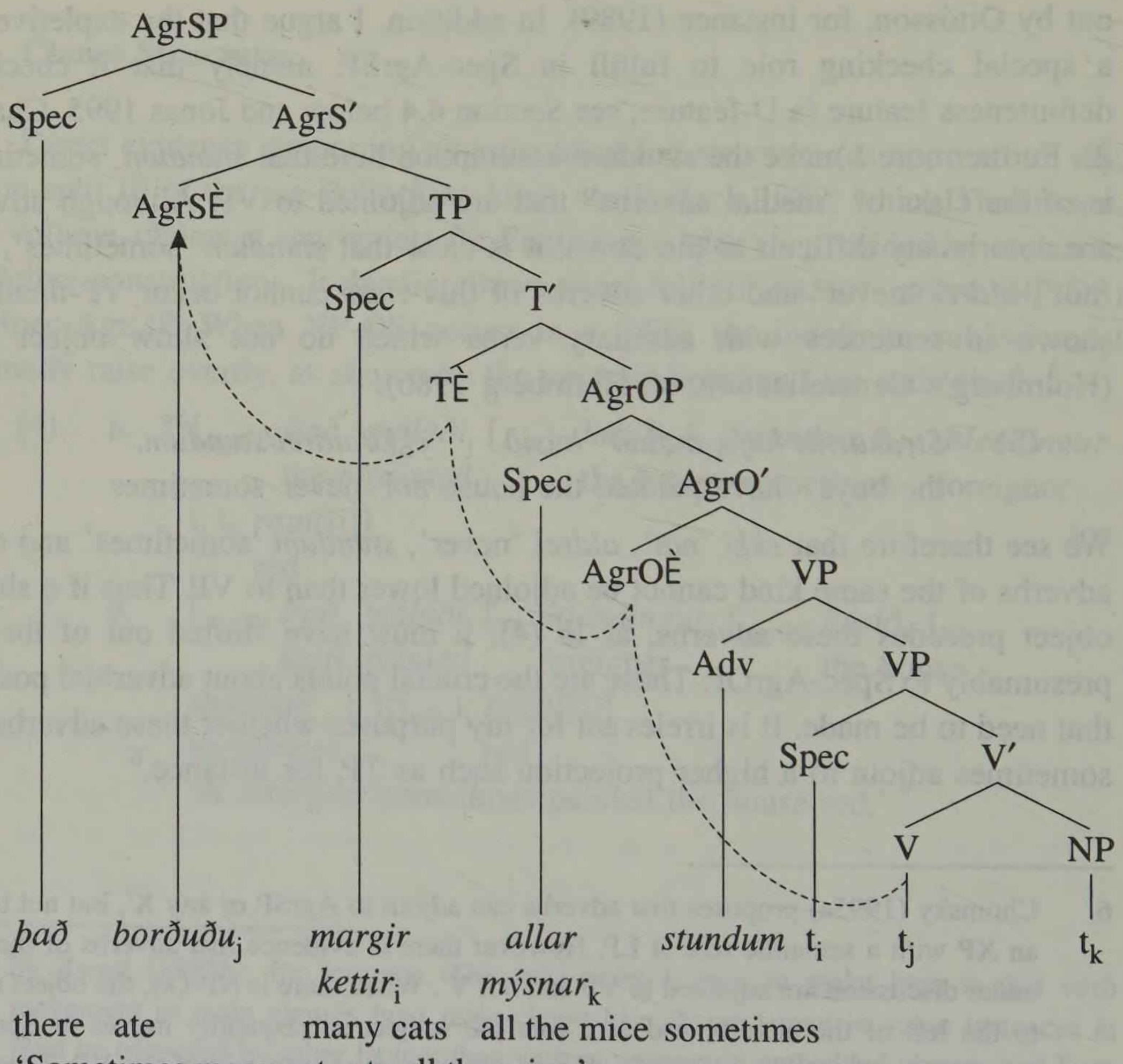
6. Chomsky (1995a) proposes that adverbs can adjoin to AgrSP or any X′, but not to VP, an XP with a semantic role at LF. However there is evidence that adverbs of the type under discussion are adjoined to VP and not V′. When there is NP-OS, the object raises to the left of the adverb, and as a rule the subject subsequently moves to Spec-TP overtly. But there is a class of subjects which may remain within the VP. These are quantified indefinite subjects such as *einhver/neinn (stúdent)* 'some/any (student)', *margir (stúdentar)* 'many (students)', etc. (see note 3). If the adverb were adjoined to V′, then we would expect the order in (i), not the order in (ii):

(i) **[Það *luku*$_i$ [$_{AgrO}$ *verkefninu*$_j$ [$_{VP}$ *margir* [$_{V'}$ *stundum* [$_{V'}$ t_i t_j]]]. Ice.
there finished the assignment many sometimes

(ii) [*Það luku*$_i$ [$_{AgrO}$ *verkefninu*$_j$ [$_{VP}$ *stundum* [$_{VP}$ *margir* [$_{V'}$ t_i t_j]]].
there finished the assignment sometimes many

When there is NP-OS in constructions of the type illustrated in (4), the minimalist principles of EQUIDISTANCE, SHORTEST MOVE and the STRICT CYCLE require that the VP-internal subject move no further than Spec-TP, the first available checking position (see J&B, Jonas 1994, 1995 for extensive discussion of this. See also Ferguson & Groat 1994 and Ferguson, this volume; for an alternative formulation of Shortest Move to that given in J&B.) The example in (6) below shows the structure assumed here for a TEC.

(6)



there ate many cats all the mice sometimes
'Sometimes many cats ate all the mice.'

The fact that (ii) is the order we get thus constitutes an argument against V′-adjunction of these adverbs. See also Vikner (1994: 140) for other empirical arguments against V' adjunction of medial adverbs in Icelandic. See also Jonas (1995) for further implications of examples of this sort.

In (6), the subject *margir kettir* checks its Case feature in Spec-TP against the complex head $[_T\ [_{Agr}\ V_{fin}\ Agr]\ T]$, the latter resulting from verb movement to AgrO followed by AgrO movement to T. The subject presumably raises to Spec-AgrSP at LF to replace the expletive and check (plural) agreement features. Thus TECs such as (6) exemplify the dissociation of Case and agreement checking postulated by Chomsky (1993) (see the above quotation) and discussed further in Chomsky (1995a). This dissociation will play a crucial role in the analysis presented here, and will be discussed in Section 4 of this paper.

3. Scandinavian Verb Movement

3.1. *Overview*

In this section I give a brief overview of the Scandinavian verb movement facts relevant to the analysis presented here. Following Vikner (1994, 1995a) I take the diagnostics for verb movement to T (V-to-T , Vikner's "independent" V-to-I) to include whether or not verb movement is allowed in the complements of so-called non-bridge verbs.[7] It is clear that Vikner does not assume an articulated Infl structure in his analysis, but rather proposes that, of the Scandinavian languages, only Icelandic has "independent" V-to-I. Vikner's "independent" V-to-I is verb movement to I that is not part of the V-to-I-to-C that he proposes for all verb movement in V2 clauses. The presence or absence of verb movement in the complements of non-bridge verbs in the Scandinavian languages divides them into two groups: Insular Scandinavian (ISc, Icelandic and [some dialects of] Faroese; see note 6), which allow verb movement in this context, and MSc (Danish, Norwegian and Swedish), which do not. Icelandic has obligatory verb movement under non-bridge verbs as shown in (7).

(7) a. *Það kom á óvart að María* ***les***$_i$ $[_{VP}$ *ekki* $[_{VP}$ $\mathbf{t}_i$ *bækur*]] Ice.
it was unexpected that Maria reads not books

b. **Það kom á óvart að María* $[_{VP}$ *ekki* $[_{VP}$ ***les*** *bækur*]]
it was unexpected that Maria not reads books

In Danish however, like other MSc languages, verb movement under non-bridge verbs is not possible as shown in (8).

---

7. See Vikner (1994) for an extensive discussion of the complements of bridge and non-bridge verbs in Scandinavian and other verb movement facts in Scandinavian embedded clauses.

(8) a. **Peter tvivler på at Marie* ***ryger***$_i$ [$_{VP}$ *ofte* [$_{VP}$ $\mathbf{t}_i$ *disse cigarer*]]
Peter doubts on that Marie smokes often these cigars
(Da.; Vikner 1994: 123)

b. *Peter tvivler på at Marie* [$_{VP}$ *ofte* [$_{VP}$ ***ryger*** *disse cigarer*]]
Peter doubts on that Marie often smokes these cigars

Vikner (1991) and Rohrbacher (1994) argue that Faroese patterns with Danish, lacking verb movement in non-bridge verb complements. However, for many speakers, verb movement is possible in such complements.[8]

(9) a. *Tað kom óvart, at Maria* ***lesur***$_i$ [$_{VP}$ *ikki* [$_{VP}$ $\mathbf{t}_i$ *bøkur*]] Far.
it was unexpected that Mary reads not books

b. *Tað kom óvart, at Maria* [$_{VP}$ *ikki* [$_{VP}$ ***lesur*** *bøkur*]]
it was unexpected that Mary not reads books

Translating Vikner's analysis into a minimalist framework which assumes an articulated Infl structure, Vikner's "independent V-to-I" can be recast either as V-to-T or V-to-T-to-AgrS (see Thráinsson 1994: 151). Presumably in finite clauses it is V-to-AgrS (via AgrO and T).[9] This differs from Vikner's analysis of MSc V2 clauses in general in which he always assumes the verb moves to a position outside of IP and that the subject moves to Spec-CP when no other XP is fronted (see also Vikner & Schwartz 1995). Zwart (1993b), adopting Chomsky's (1993) approach (and after Travis 1991), argues that within the minimalist framework, this further movement of the verb from AgrS to a higher "second" position is not motivated by some requirement of feature checking.

There are instances in Icelandic where it can be argued that there is just V-to-T in certain non-finite complement clauses, if Thráinsson (1993a) is correct in claiming that there is no AgrS projection in (at least some of) these non-finite complement clauses, since there is no subject agreement to be checked. The following examples illustrate this non-finite verb movement in Icelandic. (10a) shows that verb movement is possible in a raising expletive construction, (10b) illustrates verb movement in a non-finite ECM complement clause.[10]

---

8. There is apparent dialectal variation with respect to verb movement in embedded clauses in Faroese, but there are speakers who readily accept verb movement in these embedded clauses (see Jonas 1995).

9. However, see Næss (in prep.) for arguments that the verb in MSc moves just to T and not to AgrS.

10. The structure of these clauses will be discussed in Section 4.3.

(10) a. *Það virðast* [$_{TP}$ *margir menn*$_j$ *vera*$_i$ [$_{VP}$ *alltaf* [$_{VP}$ t_i t_j
there seem many men be.INF always
í herberginu]]] Ice.
in the room

b. *Jón telur* [*í barnaskap sínum* [*marga stúdenta lesa*$_i$
John believesin his foolishness many students read.INF
[$_{VP}$ *alltaf* [$_{VP}$ t_i *námsbækur*]]]]
always textbooks

MSc does not have movement of the non-finite verb in these contexts in contrast to the Icelandic examples above as shown in (11a, b).

(11) a. **Bengt ansåg* [$_{AgrS}$ *problemet*$_j$ *vara*$_i$ [$_{VP}$ *alltid*
Bengt considered the problem be.INF always
[$_{VP}$ t_j t_i *möjligt att lösa*]]] Sw.
possible to solve

b. **Det verkar* [$_{TP}$ *vara*$_i$ [$_{VP}$ *alltid* [$_{VP}$ t_i *en man i rummet*]]]
there seems be.INF always a man in the room

It is clear that Icelandic has movement of both finite and non-finite verbs in constructions where verb movement is impossible in MSc.[11] One way of capturing this contrast is to state it in terms of the V-feature strength of the heads of the various functional projections. Thus, Icelandic always has a strong V feature for both T and AgrS, MSc has a weak V feature for T.[12]

3.2. *Verb Movement and the Extended Projection Principle*

There are two mechanisms by which T can raise and adjoin to AgrS: one is the independent raising of T (functional head movement independent of any overt movement-and-adjunction of a verb), the other is as a secondary result of overt verb raising. In the latter Case the complex head [$_T$ [$_{Agr}$ V Agr] T] raises

---

11. Thráinsson (1993a) argues that verb movement is not found in all types of non-finite complements in Icelandic. I am not concerned with that difference here. The only point which is important in the present context is that movement of non-finite verbs exists in Icelandic. See also the discussion in Section 4.3.

12. See Zwart (1993b: 208 ff.) for an account of the asymmetry of verb movement in main and embedded clauses in MSc. Under this account, abstract AgrS-to-C movement blocks verb movement in embedded clauses. However, the verb movement of concern in this paper is V-to-T, and V-to-AgrS or higher is not of direct relevance here.

to T. There may also be a third "mixed" mechanism where there is overt verb raising to AgrS, but T raises independently. Each of these will now be discussed in turn.

Chomsky (1993: 7) proposes that in a non-verb-raising language such as English, tense raises independently to AgrS forming [$_{AgrS}$ T AgrS]. The resultant complex head checks both agreement and Case features of the DP, which raises to its specifier position. The DP checks the strong N-features of tense in Spec–AgrSP thus satisfying the EPP.[13] Tense in English has weak V–features, no overt verb raising with adjunction to the head of T occurs. In English, the subject moves overtly and checks its Case against the AgrS complex containing T, but not V. In this paper I propose that in MSc, T raises independently, whether or not there is overt verb movement. Something more needs to be said because, as is well known, MSc has verb movement to second position in main clauses. As discussed above, verb raising in embedded clauses is restricted to a subset of complement clauses, namely those of bridge verbs which have main clause-like properties (Vikner 1994, 1995a). Zwart (1993b: 207–208) argues that verb movement in subject initial MSc main clauses is movement directly to AgrS "to assist in the elimination of the N-feature of AgrS", and further to C in V2 clauses with a fronted XP.[14] This means that in both instances discussed so far, English with covert verb raising, and MSc with main clause verb movement to AgrS or higher, the head of Tense is the trace of T as a result of Tense raising independently, not a complex head formed by adjunction of the complex head [$_{AgrO}$ V AgrO] to T as it is in Icelandic. Unlike both English and MSc, Icelandic and Faroese (to some extent) both have overt V-to-T. When overt V tot T occurs, T does not raise independently. Adjunction of the complex head [$_{AgrO}$ V AgrO] to T forms the complex head [$_{T}$ [$_{Agr}$ V Agr] T]. As we saw above in the discussion of the verb movement facts, the V may be either finite or non-finite in this case. This will play a major role in the discussion of subject Case-checking proposed in the following section.

The EPP may be satisfied in one of two positions: the subject in Spec-TP

---

13. Chomsky (1993: 31) proposes that the EPP reduces to a morphological property of Tense in terms of its N-features. I argue elsewhere (Jonas 1995) that the so-called EPP phenomenon is not uniformly reducible to checking by an element in Spec-TP or Spec-AgrSP but that this depends on the clause structure involved and has interesting consequences. This is discussed in some detail in Section 4 of this paper.

14. I assume the verb moves first to AgrO to form the complex head [$_{AgrO}$ V AgrO] against which the object will check its Case at LF (see also Epstein 1993: 49).

checks the relevant feature of T if T does not raise independently and the complex head $[_T [_{Agr}$ V Agr] T] is formed as a result of AgrO movement to T. If T raises independently, the subject in Spec-AgrSP checks the N-feature of T. Bare T does not project a specifier and once T raises independently, nothing can be checked in Spec-TP. Hence Spec-TP in this instance cannot be an intermediate position to which the subject may move overtly. As we will see later, in Icelandic Spec-TP can be an intermediate position for the subject when there is movement of a non-finite verb to T. I will assume that in such cases the subject checks an EPP feature on T of the non-finite clause (see the discussion in 3.4 below).

Overt or covert verb movement can thus result in three different configurations of T, each of which has distinct licensing properties:

(12) a. Trace of a "bare" T following independent functional head movement of T to Agr. "Bare" T to which the verb does not raise overtly does not project a specifier.

b. The complex head $[_T [_{AgrO} V_{fin}$ AgrO] T] projects a specifier. The subject may move overtly to Spec-TP and check its Case feature.

c. The complex head $[_T [_{AgrO} V_{non\text{-}fin}$ AgrO] T] projects a specifier. A non-finite T cannot check the Case of a subject, but a subject in the specifier position of a non-finite T checks an N-feature (an EPP feature) of non-finite T.

In the next section I will propose that each of these configurations of T plays a different role in the Case-checking and/or licensing of NPs in the Spec-TP position.

To summarize so far, the hypothesis proposed in this paper, namely that languages which have Spec-TP as a Case- (and/or EPP) checking position, are those which have overt V-to-T, is supported empirically in Scandinavian.[15] The

---

15. Since Dutch has TECs, as shown by Zwart (1992a), I need to claim that it has overt verb movement to T. Such a claim appears to be incompatible with analyses of Dutch as a uniformly head-initial language, such as the ones proposed by Zwart (1993a) and some of the papers in this volume that adopt Kayne's (1994) general approach. Zwart (1992a) argues that TECs in Dutch are made possible by overt T-to-AgrS raising, which he maintains does not happen in English nor in Mainland Scandinavian, contrary to the assumptions made here and in Chomsky 1993, for instance. Other researchers have argued that there is no TP (or IP) in Dutch, either because it is fused with the VP (see, e.g. Reuland 1990; Kosmeijer 1993 and references cited there) or does not exist at all (cf. van Gelderen 1993). While I cannot go further into the TP/no-TP controversy in Dutch, I should like to point out that my analysis would seem to be compatible with

Insular Scandinavian languages, Icelandic and Faroese are those which exhibit TECs, where Spec-TP is available, and these languages also have V-to-T. In the next section of this paper I will make explicit the interaction of verb movement and subject Case-checking within the minimalist framework and discuss in detail the licensing of Spec-TP as a subject position.

4. Subject Positions in Expletive Constructions and Infinitival Complements

4.1. *TECs*

Some previous accounts (Ottósson 1989; Vikner 1991; Bures 1993) propose that the indefinite subject in TECs remains in VP-internal position at surface structure.[16] However, Jonas (1992) and J&B have argued that the subject in Icelandic TECs typically occupies Spec-TP overtly and Zwart (1992a) argues the same for Dutch. This is borne out empirically when we consider example sentences with VP-adjoined manner adverbs and NP-OS. In such Cases a bare indefinite subject precedes both the shifted object and the VP-adjoined manner adverb:

(13) a. $[_{AgrSP}$ *Það málaði* $[_{TP}$ *stúdent*$_i$ $[_{AgrOP}$ *húsið* $[_{VP}$ *vandlega*
there painted student the house carefully
$[_{VP}$ t_i *rautt*]]]] Ice.
red
'A student carefully painted the house red.'

b. *$[_{AgrSP}$ *Það málaði* $[_{AgrOP}$ *húsið* $[_{VP}$ *vandlega* $[_{VP}$ *stúdent*
there painted the house carefully student
rautt]]]]
red

In Faroese, which allows TECs, but not overt NP-OS, the following obtains: in a TEC the subject cannot remain in VP-internal position following the VP–adjoined manner adverb *dúgliga* analogous to (13):

---

earlier analyses of Dutch, which assumed (string-vacuous) raising of a VP-final V to an IP-final I (or T in my system), as illustrated in Haegeman (1991: Chapter 11), for instance.

16. Bures (1993) argues that the subject is in Spec-VP only when the object does not raise overtly, if overt object shift occurs the subject is forced to raise to Spec-TP overtly.

(14) a. *Tað bygdu* [$_{TP}$ *nakrir íslendingar* [$_{VP}$ *dúgliga* [$_{VP}$ *hús í Havn*]]]
there built some Icelanders efficiently houses in Tórshavn Far.

b. **Tað bygdu* [$_{VP}$ *dúgliga* [$_{VP}$ *nakrir íslendingar hús í Havn*]]
there built efficiently some Icelanders houses in Tórshavn

Expletive constructions in MSc appear to be restricted to verbs lacking an external argument (see Platzack 1983; Maling 1988 and Vikner 1991, 1995 for an extensive overview of these constructions in Scandinavian). Thus transitive verbs do not appear in expletive constructions in MSc. As there is no such restriction on these verbs in expletive constructions in Icelandic and Faroese, I take this as a starting point for assuming that Spec-TP is not licensed and therefore cannot be a checking position in MSc. However, we need to examine other constructions to compare MSc with Icelandic and Faroese.

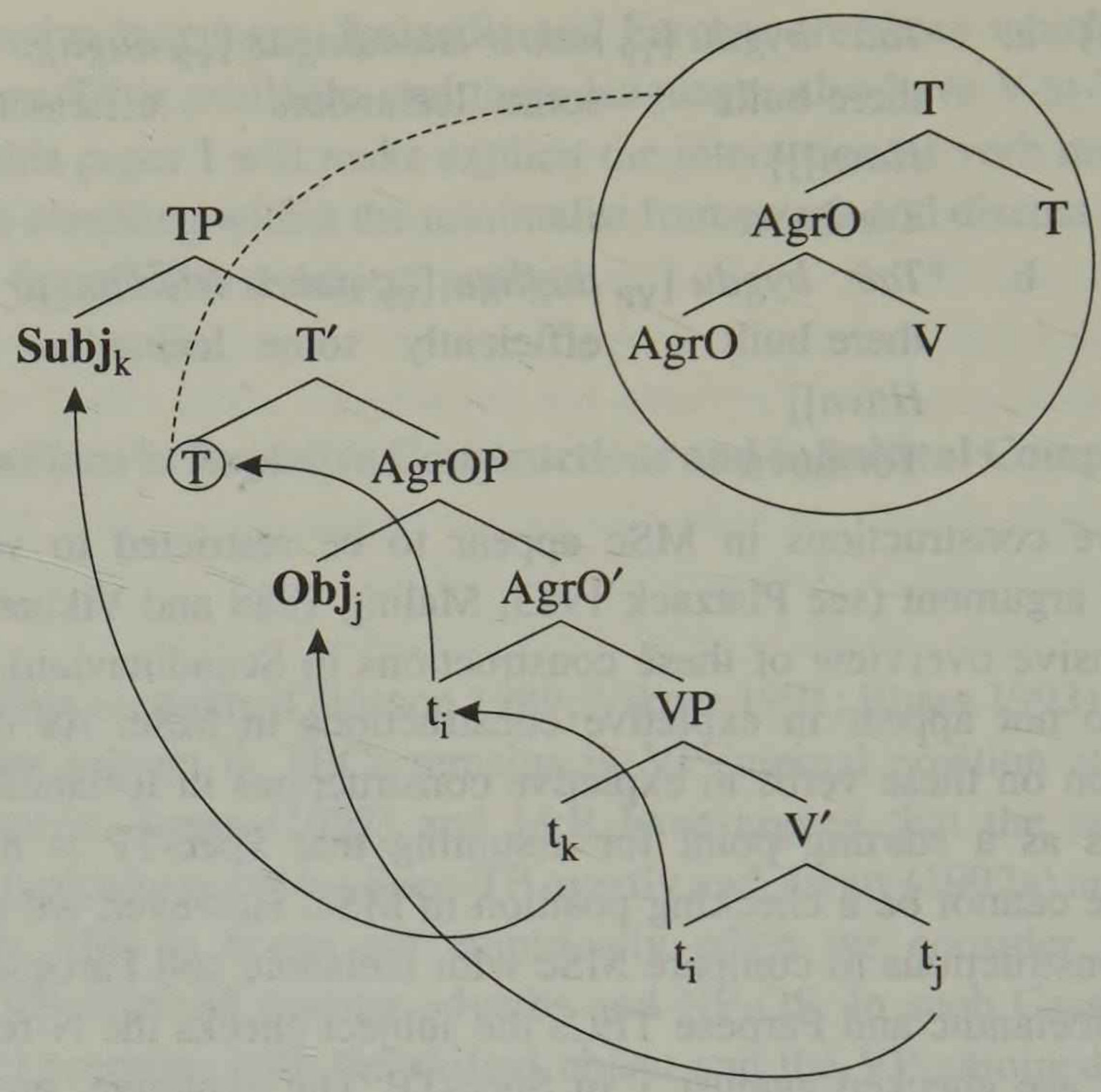
In Icelandic and Faroese TECs the subject checks the N-features of T, and has its Case checked against T in Spec-TP. The expletive, on the other hand, checks a D-feature in Spec-AgrSP (as the expletive in languages like Icelandic and Faroese, for instance, is arguably just a SPELL-OUT of a D-feature).[17] Bures (1993) and J&B argue that overt object shift, overt movement of the object to Spec-AgrOP, requires that Spec-TP be an available A-position. In particular, J&B argue that the interaction of two principles of *MPLT* entail this. These principles are Shortest Movement and Equidistance. The version of Shortest Move developed by Ferguson and Groat (1994) requires that the subject move to the first available checking position.[18] Once the verb has raised, and if the object shifts overtly, the subject can move only as far as Spec-TP by the Equidistance principle. This is illustrated below in (15):

---

17. I assume here that the EPP phenomenon is sometimes the reflection of a D-feature checking requirement, sometimes the reflection of an N-feature checking requirement (see Chomsky 1995b; Jonas 1995). A D-feature is contained in the head of AgrS and requires a DP, a specific NP or an expletive in Spec-AgrSP to be checked. T on the other hand, contains no D-feature but an N-feature and hence allows an indefinite NP in Spec-TP and cannot be checked by an expletive, which is the lexical expression of a D-feature and has no N-feature under my analysis.

18. This is a somewhat different notion of Shortest Move from the one assumed by Chomsky 1993. See the paper by Ferguson in this volume for further discussion.

(15)



The verb moves to AgrO and forming the chain $\{[_{AgrO}$ AgrO V], $t_v\}$. This chain formation renders Spec-VP and Spec-AgrOP equidistant from the position of the object. Thus, following overt verb raising, the object may then move overtly to Spec-AgrOP past the subject in Spec-VP. The subject cannot then raise across Spec-AgrOP directly to Spec-AgrSP because the two specifiers are not equidistant from the base position of the subject, Spec-VP but may move only as far as Spec-TP.

We have thus seen some empirical confirmation already of the entailment of *MPLT* that the subject will move Spec-TP in expletive constructions with overt NP-OS, such as in the Icelandic example in (13). Further, there is no evidence that the subject occupies a different structural position in TECs when there is no NP-OS, as in the Faroese example in (14) or in Icelandic TECs without Object Shift.[19] We now turn to other types of evidence for the availability vs. non-availability of Spec-TP in Scandinavian.

does this imply that obj has to move prior to subject?

19. See, however, note 5.

4.2. *Subject Positions in Expletive Passives in Scandinavian*

In passive expletive constructions there is a clear contrast between Icelandic and MSc as to the positions the subject may occupy overtly. There is a position following the finite auxiliary that the subject may move to overtly in Icelandic that is impossible in MSc. The subject of a passive expletive in MSc may occur only in its base position in MSc or in clause-initial position in Spec-AgrSP as shown in (16).

(16) a. **Det har* ***noen kaker*** *blitt bakt for selskapet.* No.
there have some cakes been baked for the party
b. *Det har blitt bakt* ***noen kaker*** *for selskapet.*
there have beenbaked some cakes for the party
c. ***Noen kaker*** *har blitt bakt for selskapet.* (T. Taraldsen, p.c.)
some cakes have been baked for the party

In Icelandic on the other hand, the subject may occur following the finite auxiliary and if this position is Spec-TP as argued in Jonas & Bobaljik (1993) and Jonas (1994) then this is exactly the contrast that we would expect given that we have argued here that Spec-TP is not available in MSc.

(17) *Það hafa* [$_{TP}$ *nokkrar kökur*$_i$ [*verið* [*bakaðar* t$_i$ *fyrir veisluna*]]]. Ice.
there have some cakes been baked for the party

We can thus assume that the V-features of T are uniformly strong in Icelandic, attracting a finite main verb or a finite auxiliary to T as the closest element with verbal features, forming a complex head with a full set of features. T then projects a specifier and the subject may raise overtly to Spec-TP, its Case checking position. In MSc, the finite verb or auxiliary raises in subject initial main clauses to the [$_{AgrS}$ [T AgrS]] complex head formed following independent functional head movement of T to AgrS. The subject either raises to Spec-AgrSP or an expletive is inserted in this position. The expletive in this position, as in TECs, checks only the D-feature of AgrS and the subject raises at LF to check agreement.

4.3. *Infinitival Constructions in Scandinavian*

As illustrated in Section 3.1, Icelandic allows verb movement in non-finite complement clauses. This gives independent evidence for V-to-T alone distinct from further movement to AgrS. Under the hypothesis presented here, this predicts that there may be evidence for involvement of Spec-TP in Icelandic

infinitival complements, that is Spec-TP as a separate position from Spec-AgrSP. I will argue for this analysis and show that it accounts for word order differences in infinitival complements between Icelandic and MSc. Some Icelandic examples are given in (18).[20]

(18) a. *Sigríður lofaði* [*að* PRO ***senda***$_i$ [$_{VP}$ ***stundum*** [$_{VP}$ t_i
Sigrid promised to send sometimes
tölvupóst]]] Ice.
e-mail

b. *Það virðast* [*margir menn*$_j$ ***vera***$_i$ [$_{VP}$ *alltaf* [$_{VP}$ t_i t_j
there seem many men be.INF always
í herberginu]]]
in the room.

c. *Jón telur* [*marga stúdenta*$_j$ ***lesa***$_i$ [$_{VP}$ *alltaf*
John believes many students read.INF always
[$_{VP}$ t_j t_i *námsbækur*]]]
textbooks

This is an interesting result and indicates that the V-features of T can be strong independent of finiteness in Icelandic.

Given the absence of V-to-T in finite clauses in MSc, on the other hand, we expect that non-finite verb movement should be impossible in the MSc languages. The V-features of T should be uniformly weak and we would expect T to raise independently of the verb movement chain to AgrS. This, too, is borne out by the facts. MSc has no non-finite verb movement. The following Swedish examples can serve as an illustration:

(19) a. **Anders lovade* [*att* PRO$_i$ *dricka*$_j$ [$_{VP}$ *inte* [$_{VP}$ t_j t_i mjölken]]]
Anders promised to drink not the milk
Sw.

b. *Anders lovade* [*att* PRO$_i$ [$_{VP}$ *inte* [$_{VP}$ t_j *dricka mjölken*]]]
Anders promised to not drink the milk

---

20. This verb movement is clearly obligatory in PRO-infinitivals, just as in finite complements (see, for example, Hornstein 1990; Sigurðsson 1991; Thráinsson 1993a). The situation is less clear in raising complements and ECM complements and this calls for an explanation. I will not try to offer one here. The only point I want to make is that verbs move in non-finite complements in Icelandic but they do not in MSc. Hence my analysis predicts that Spec-TP should be available in non-finite complements in Icelandic whereas it should not in MSc. This prediction is borne out.

(20) a. **Det verkar* [*vara*$_i$ [$_{VP}$ *inte* [$_{VP}$ t$_i$ *en man i rummet*]]] Sw.
there seems be-INF not a man in the room
b. *Det verkar* [[$_{VP}$ *inte* [$_{VP}$ *vara en man i rummet*]]]
there seems not be.INF a man in the room

(21) a. **Bengt ansåg* [$_{AgrSP}$ *problemet*$_j$ *vara*$_i$ [$_{VP}$ *inte* [$_{VP}$ t$_j$ t$_i$
Bengt considered the problem be.INF not
möjligt att lösa]]] Sw.
possible to solve
b. **Bengt ansåg* [$_{AgrSP}$ *problemet*$_j$ [$_{VP}$ *inte* [$_{VP}$ t$_j$ *vara möjligt*
Bengt considered the problem not be.INF possible
att lösa]]]
to solve

To summarize so far, in this section I have shown that Icelandic has V-to-T in non-finite complement clauses whereas MSc does not. I will now argue that this licenses Spec-TP in Icelandic non-finite complement clauses but not in MSc.

First, consider the well known contrast illustrated in (22) between Icelandic on the one hand and Swedish and English on the other:

(22) a. *Það virðist* [$_{TP}$ *einhver*$_j$ *vera* t$_j$ *í herberginu*] Ice.
there seems somebody be.INF in the room
b. **Det verkar* [$_{AgrSP}$ *någon vara i rummet*] Sw.
there seems somebody be.INF in the room
c. **There seems* [$_{AgrSP}$ *somebody to be in the room*]

As illustrated in (22a), I am assuming (following Thráinsson 1993b) that the raising complement in Icelandic is a TP. As shown in (10a) and (18b) above, V-to-T is possible in these complements in Icelandic. Hence the T projects a specifier under my analysis and Spec-TP is licensed as an intermediate subject position. My claim is that this position can be occupied by an (indefinite) NP as which checks the EPP-feature of the non-finite complement, but it must eventually raise to Spec-TP in the finite clause to check its Case feature.[21] As Spec-

---

21. Actually, the sentence in (22a) is structurally ambiguous since *einhver* could arguably be in the matrix Spec-TP. In a sentence such as (i), on the other hand, *einhver* appears to be unambiguously in the embedded TP:
(i) *Það virðast* [$_{VP}$ *ekki* [$_{VP}$ [$_{TP}$ *margir stúdentar vera* [$_{VP}$ *alltaf* [$_{VP}$ *hér*]]]]] Ice.
there seem not many students be-INF always here
'It does not seem that many students are always here.'

TP is not available in MSc or English, this is not an option in those languages.

The contrast observed in (22) suggests that there is a structural difference between raising complements in Icelandic on the one hand and English and MSc on the other. As indicated by the bracketing, I assume here that raising complements in English and MSc are AgrSPs. This implies that the embedded T raises to the embedded AgrS, as it normally does in these languages. That means under the analysis presented here that the embedded AgrS has a D-feature to check and this D-feature can be checked by an expletive inserted in the embedded Spec-AgrSP (and then raised to the matrix Spec-AgrSP to check the D-feature of the matrix AgrS). This blocks the raising of an indefinite embedded subject to this position in MSc and English for economy reasons. In Icelandic however, which I assume has Spec-TP as the highest position in raising complements, the expletive (as a pure D-feature) cannot be inserted in Spec-TP as the non-finite T has an N-feature but no D-feature to check. Thus expletive insertion cannot block raising of the indefinite subject to check the N-feature of T.[22] But whatever the explanation, it is clear that the only grammatical version of these sentences in Swedish and English is the one with the indefinite NP following the verb *be*, i.e. the VP-internal position. This version is also acceptable in Icelandic as illustrated in (23):[23]

(23) a. *Það virðist* [$_{TP}$ *vera stelpa í herberginu*] Ice.
there seems be-INF girl in the room

b. *Det verkar* [*vara en flicka i rummet*] Sw.
there seems be-INF a girl in the room

c. *There seems* [*to be a girl in the room*]

In (i) the negation is above the matrix VP and the adverb *alltaf* 'always' is above the embedded VP, as explained before, so both the finite matrix verb and the non-finite embedded verb have moved, at least to T. Thus Spec-TP is available for the lexical subject in the embedded clause (see Jonas 1995, Chapter 3, for further discussion).

22. Another possibility is that the raising complements are "smaller" in English and MSc than they are in Icelandic, possibly just VPs as originally suggested by Bresnan (1982), for instance. However, I assume here that they are AgrSPs.

23. This is also the preferred version in Icelandic if the NP is a bare indefinite. For an interesting discussion of the relationship between definiteness (or specificity) and different subject positions see Vangsnes (1995). See also the discussion in Jonas (1995).

Further interesting evidence for structural differences between Icelandic on the one hand and MSc and English on the other comes from ECM complements. In these complements there is an embedded subject position that cannot be a Case-checking position if one assumes that the subject of the complement clause has its structural Case checked in the Spec-AgrOP position of the matrix clause (see, for example, Chomsky & Lasnik 1993; Jonas 1995; among others). Now recall that it is argued here that the expletive checks a D-feature in Spec-AgrSP. This means that if the position under discussion is Spec-AgrSP in Swedish and English, as claimed above, but Spec-TP in Icelandic,[24] this analysis predicts that expletives should be possible in this position in Swedish and English but not in Icelandic. This is so because Spec-TP is not a position to check a D–feature, which is what the expletive checks under this analysis, whereas Spec–AgrSP is. This prediction is borne out as shown in (24).

(24) a. **Ég tel* [$_{TP}$ *það vera marga stúdenta í þessum bekk*] Ice.
I believe there be.INF many students in this class

b. *Jag anser* [$_{AgrSP}$ *det vara många studenter i denna klass*] Sw.
I consider there be.INF many students in this class

c. *I belive* [$_{AgrSP}$ *there to be many students in this class*]

We see, therefore, that the availability vs. non-availability of Spec-TP is responsible for a number of seemingly disparate word order differences between languages such as Icelandic on the one hand and MSc and English on the other.

4.4. *Subject Checking at LF*

In J&B it was argued, contra Bures (1992, 1993), that Spec-TP is not available at LF. However, when a wider range of data from expletive constructions is considered, we see that the Spec-TP must be an available subject Case-checking position at LF. The data that was not considered in J&B or Bures are

---

24. This difference could follow from some sort of a principle of Economy of Representation (see e.g. Grimshaw 1993 for such ideas): TP is all that is needed for a non-finite clause in Icelandic, since T in Icelandic will project a specifier once the verb has moved to T, so TP is what you will get when you do not "need" an AgrS. In MSc and English, on the other hand, one cannot have anything "smaller" than AgrSP since T always has to raise to AgrS. TP is not an option for a non-finite clause in these languages as a bare T does not project a specifier.

expletive constructions in which the indefinite (heavy) subject occurs in final position.[25] Icelandic allows both transitive and intransitive verbs to occur in these constructions as illustrated below.

(25) a. *Það veiddi allar mýsnar* [*bröndóttur köttur með langt skott*]
there caught all the mice a striped cat with a long tail
Ice.

b. *Það reyktu af kappi* [*margir stúdentar frá Akureyri*]
there smoked intensively many students from Akureyri

c. *Það komu í partíið* [*margir stúdentar frá Akureyri*]
there came to the party many students from Akureyri

In contrast in MSc, only the analogue of the Icelandic (25c) is acceptable (see Maling 1988, Vikner 1991, Platzack 1983).

(26) a. **Det åt en pudding en man med en stor röd hatt.*
there ate a pudding a man in a big red hat
Sw.

b. **Det rökte ivrigt flera studenter från Stockholm.*
there smoked eagerly several students from Stockholm

c. *Det kom till festen många nya studenter från Island.*
there came to the party many new students from Iceland

Thus the lexical restrictions on verbs which may occur in these expletive constructions with "extraposed" subjects are the same that hold for TECs without "extraposed" subjects. Icelandic allows all verbs in these constructions, MSc constructions of this type are restricted to unaccusative and "true" intransitive verbs (see Maling 1988).

Bures (1992, 1993) and J&B argue that it is an entailment of *MPLT* that a language which allows NP-OS must have Spec-TP as an available A-position for subsequent raising of the subject. This led Bures (1993) to propose that within Germanic: those languages which allow NP-OS are exactly those which exhibit TECs. Bures (1993) argues on the basis of this proposed correlation that TECs are the LF analog of overt object shift. He argues that if the object shifts

---

25. Chomsky (1995a) suggests that these constructions may be the result of movement of the subject in the PF component, extraposition in the usual sense as rightward movement not being available in the minimalist framework. I take no stand in this paper on the surface position of the subject or the movement analysis. What concerns us here is how the Case of the subject is checked at LF, since the subject is not in its Case-checking position at Spell-out.

overtly, the subject in a TEC must move overtly to Spec-TP, if the object raises covertly the subject may raise covertly as well, but no further than Spec-TP. He gives this as evidence for the cycle at LF. I adopt this latter proposal here, that the cycle holds at LF, but suggest that relevant data for this argument are the "extraposition" structures shown above in (25)–(26). The entailment of the theory, that Spec-TP must be an available position for subject raising when a language has overt NP-OS, however, is unidirectional. There is nothing in the theory that requires a language in which Spec-TP is an available A-position to allow NP-OS. Spec-TP could potentially be licensed independently of the presence of NP-OS. Therefore the correlation may not be the equivalency between Spec-TP and NP-OS proposed by Bures. Rather, while a language which has NP-OS must have Spec-TP as an available A-position, the converse is not entailed by the theory. That is, a language may, in theory, have Spec-TP as an available A-position, but not exhibit NP-OS. Faroese (or at least one variety of it, see Jonas 1994, 1995) has Spec-TP as an available A-position as it allows TECs, but there are no speakers of Faroese for whom NP-OS is acceptable. This confirms the empirical correctness of the absence of this converse entailment.

MSc does not exhibit the expletives in (26) with transitive or unergative verbs because the same requirement holds at LF and overtly: once the object has raised, Spec-TP must be available as a Case-checking position for subsequent raising of a subject that has not already been raised to its checking position overtly, as in the expletive constructions with "extraposed" subjects.

4.5. *Concluding Remarks*

In this paper I have attempted to define the roles of the specifier positions Spec-AgrSP and Spec-TP in terms of checking theory. I have argued that while Spec-AgrSP is available as a checking position in all the languages considered here, Spec-TP is only available in those languages that exhibit overt V-to-T, such as Icelandic and (to some extent) Faroese. In languages with independent T-to-AgrS raising, such as the Mainland Scandinavian languages and English, only Spec-AgrSP is available as a checking position for subjects. In addition, Spec-TP is a potential position for EPP-feature checkers in non-finite clauses that exhibit V-to-T.

I have claimed that the correlation between V-to-T and availability of Spec-TP as a checking position is predicted by *MPLT* and I have tried to show that this prediction is supported empirically. Thus Icelandic and Faroese both have

V-to-T and exhibit the range of subject positions that I have argued involve Spec-TP as a subject Case-checking position (TECs, passive expletive constructions, expletives with "extraposed" subjects) or intermediate (non-Case) position (raising expletive constructions, ECM constructions). MSc on the other hand does not exhibit V-to-T and lacks Spec-TP, overtly as a Case-checking or intermediate (EPP-checking) position, covertly as a Case-checking position.

Acknowledgements

This paper contains material further developed in Jonas (1995). Therefore there are some differences in details of the analyses presented here and in Jonas (1995). I thank members of my dissertation committee (Noam Chomsky, Samuel D. Epstein, Höskuldur Thráinsson, Kenneth Wexler) for useful discussions of many of the issues raised here.

Parts of this work have been presented at various locations, including Harvard University, Cornell University and the University of Maryland (spring of 1994), University of Bergen (fall of 1994), University of Connecticut and Yale University (spring of 1995), and I thank audiences there for their comments.

I also wish to thank Jonathan Bobaljik, Chris Collins, Samuel D. Epstein, Höskuldur Thráinsson and anonymous reviewers of this volume for comments on an earlier version of this paper. Special thanks to Zakaris Hansen, Jógvan í Lón Jacobsen and Eivind Weyhe and their colleagues at Fróðskaparsetur Føroya for help with collection of the Faroese data, and to Tor Åfarli, Anders Holmberg, Tarald Taraldsen, Höskuldur Thráinsson and Øystein Vangsnes for help with other Scandinavian data. The data collection in the Faroe Islands was supported by a research grant from the American-Scandinavian Foundation. All errors are my own.

Raising Quantifiers without Quantifier Raising

Hisatsugu Kitahara
The University of British Columbia

... the language faculty is nonredundant, in that particular phenomena are not "over determined" by principles of language.

— Chomsky (1993: 2)

Abstract

This paper, couched within the theory of feature-checking which incorporates the chain-formation algorithm (Chomsky 1993, Chomsky & Lasnik 1993), develops a chain-based theory of scope interpretation (Aoun & Li 1989, 1991), which explains the core cases of scope phenomena concerning argument-quantifiers without Aoun & Li's (1991) central stipulation concerning NP-trace. The proposed analysis of such scope phenomena also reveals significant redundancy between feature-checking operations (independently motivated within the minimalist program) and the (widely assumed) LF-rule of QUANTIFIER RAISING (May 1977, 1985), suggesting that QUANTIFIER RAISING is (at least partially) eliminable.

1. Introduction

The goal of this paper is to show that, within the theory of feature-checking which incorporates the chain-formation algorithm (Chomsky 1993; Chomsky & Lasnik 1993), a chain-based theory of scope interpretation renders the LF-rule of QUANTIFIER RAISING (QR) (May 1977, 1985) unnecessary to explain the core cases of scope phenomena concerning argument-quantifiers. The paper is organized as follows: Section 2 reviews Aoun & Li's (1989, 1991) chain-based theory of scope interpretation including a stipulation concerning NP-trace.

Section 3 reviews the theory of feature-checking which incorporates the chain-formation algorithm (Chomsky 1993; Chomsky & Lasnik 1993). Section 4 proposes an analysis of the core cases of scope phenomena, within which Aoun & Li's stipulation is eliminated; and more importantly, the LF-rule of QR is rendered unnecessary. Finally, section 5 briefly discusses the overall role of QR within the minimalist program.

2. A Chain-Based Theory of Scope Interpretation

As is well known, the following sentence exhibits two distinct scope interpretations.[1]

(1) *Someone loves everyone.* some >< every

The sentence (1) can be interpreted as "there is a particular person who loves every person" (collective reading i.e. *someone* taking wide-scope) or as "for every person, there is someone or other who loves that person" (distributive reading i.e. *everyone* taking wide-scope). Adopting the VP-Internal Subject Hypothesis (see among others, Fukui & Speas 1986; Kitagawa 1986; Kuroda 1988; Koopman & Sportiche 1991), in the overt syntax, *someone* undergoes NP-movement from the Spec-VP to the Spec-IP, yielding the S-structure (2a). Given the LF-rule of QR (May 1977, 1985), in the LF-component, *someone* and *everyone* each undergo QR. The S-structure (2a) is thus mapped into the LF-structure (2b), in which: (i) $someone_1$ c-commands both $everyone_2$ and its trace t_2, and (ii) $everyone_2$ c-commands the VP-internal trace of $someone_1$, namely t_1 occupying the Spec-VP.

(2) a. *S-structure* (*resulting from NP-movement*):

$[_{CP}\ [_{IP}\ someone_1\ [_{VP}\ t_1\ loves\ everyone\ ]]]$

NP-mvt

b. *LF-structure* (*resulting from QR*):

$[_{CP}\ [_{IP}\ someone_1\ [_{IP}\ t'_1\ [_{VP}\ everyone_2\ [_{VP}\ t_1\ loves\ t_2\ ]]]]]$

QR QR

---

1. I use the following notations for distinct scope interpretations:
   (i) a. X > Y iff X obligatorily has scope over Y.
   b. X >< Y iff either X or Y may have scope over the other.

Given that the LF structure (2b) is assigned to the ambiguous sentence (1); Aoun and Li (hereon, A&L) (1989, 1991) propose the following two principles.[2,3]

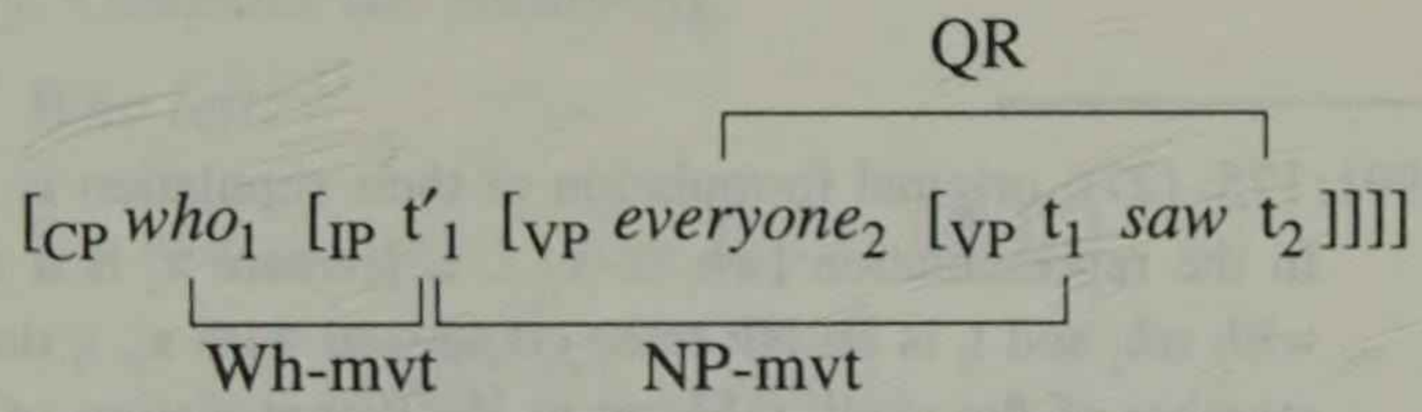
(3) Minimal Binding Requirement (A&L 1991: 164: (I))
A variable must be bound by the most local potential A′-binder.

(4) Scope Principle (A&L 1991: 164: (II))
A quantifier A has scope over a quantifier B in case A c-commands a member of the chain containing B.

Given that traces created by QR are variables, the LF-structure (2b) satisfies the Minimal Binding Requirement (3). Finally, given the Scope Principle (4), A&L correctly predict the scope ambiguity in (2b), i.e. either *someone*$_1$ (c-commanding both *everyone*$_2$ and its trace t_2) or *everyone*$_2$ (c-commanding the VP-internal trace of *someone*$_1$, namely t_1) may have the scope over the other. However, this analysis also incorrectly predicts the existence of a wide-scope reading of *everyone* in (5a), in which *who*, in fact, obligatorily takes wide-scope.[4] The unambiguous sentence (5a) is assigned the LF-structure (5b) (containing a chain-structure quite similar to the one in (2b)), in which *everyone*$_2$ c-commands a trace of *who*$_1$, namely t_1 occupying the Spec-VP:

(5) a. *Who saw everyone?* *who > every*

b.



---

2. A&L assume the definition of C-COMMAND in (i) below, but they note that the definition in (ii) could have been incorporated instead.
   (i) Node A c-commands node B if neither A nor B dominates the other and the first branching node dominating A dominates B (Reinhart 1976).
   (ii) A c-commands B if A and B are XP, A≠B, and every maximal projection dominating A dominates B (Aoun & Sportiche 1983).

3. A&L (1991: 171) define POTENTIAL $\overline{\text{A}}$-BINDER as follows:
   (i) A qualifies as a potential $\overline{\text{A}}$-binder for B iff A c-commands B, A is in an $\overline{\text{A}}$-position, and the assignment of the index of A to B would not violate any grammatical principle.

4. (5a) can be interpreted only as a single question (collective reading i.e. *who* taking wide-scope) whose appropriate answer would be e.g. "John saw everyone."

Consequently, the Scope Principle (4) incorrectly predicts the existence of a wide-scope reading of *everyone*. To avoid this problem, A&L (1991: 175) stipulate that the NP-trace t_1 is excluded from scope interpretation in this particular structure (5b) (containing a *wh* and a QP) but the NP-trace t_1 is not excluded from scope interpretation in the structure (2b) (containing two QPs).[5]

In what follows, I will show that A&L's insightful chain-based theory of scope interpretation can be maintained without stipulation. I first review the theory of feature-checking which incorporates the chain-formation algorithm (Chomsky 1993; Chomsky & Lasnik 1993). Then, I propose a natural revision of A&L's Scope Principle (4). Within the proposed analysis, the contrast between (1) and (5a) is shown to be explicable without stipulation. Further, our analysis of such scope phenomena reveals significant redundancy between feature-checking operations (independently motivated within the minimalist program) and the LF-rule of QR, suggesting that QR is (at least partially) eliminable.[6]

3. Feature-Checking Operations and Chain-Formation

Assuming the VP-Internal Subject Hypothesis, Chomsky (1991, 1993) adapts the articulated IP-structure of Pollock (1989) and proposes a theory of

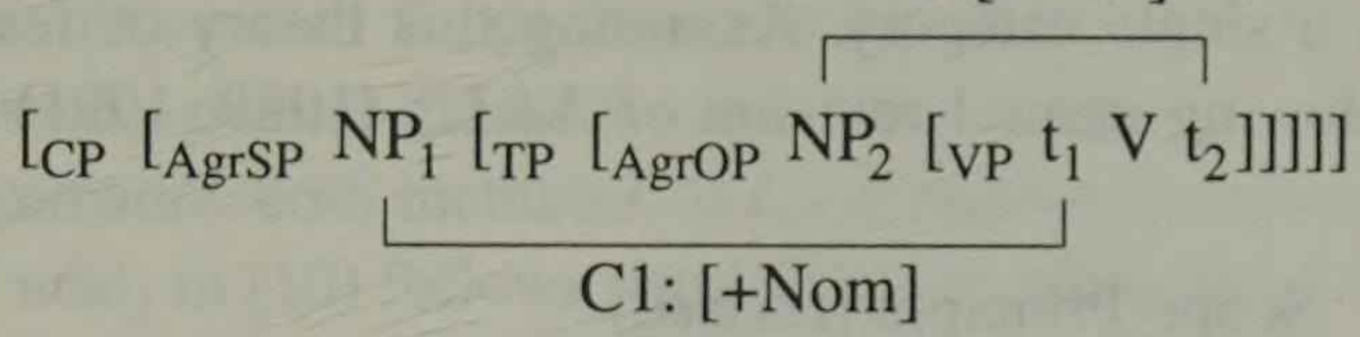
---

5. A&L's (1991: 175: (27)) original formulation of their stipulation is as follows:
   (i) In the representation [wh_i ... x_i ... t_i] where x_i is a variable coindexed with wh_i and t_i is an NP-trace coindexed with x_i, t_i does not constitute a member of the chain relevant to the determination of scope; whereas in the representation [QP_i ... x_i ... t_i] where x_i is a variable coindexed with QP_i and t_i is an NP-trace coindexed with x_i, t_i constitutes a member of the chain relevant to the determination of scope.
   It is important to note that (i) is not a stipulation about NP-trace in general, but is rather a stipulation about a certain type of NP-trace, an entity having no independent theoretical motivation.

6. A&L (1993) eliminate their stipulation (A&L 1991: 175: (27)) by excluding all NP-traces from scope interpretation and propose a new analysis of (1), in which two classes of QR (i.e. QP-raising, Q-raising) interact with their new (disjunctive) Scope Principle: "An operator A may have scope over an operator B iff A c-commands B or an A′-element in the chain headed by the operator" (A&L 1993: 88). Since the scope of this paper is limited to suggesting an alternative way of eliminating their stipulation, and most crucially, one which leads us to an analysis without any appeal to the LF-rule of QR, I will not review their new QR-based analysis here. For detailed discussion of their new proposal, see A&L (1993).

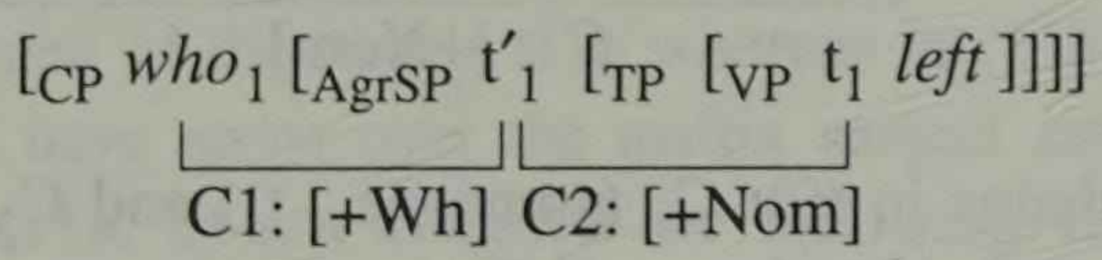
feature-checking, within which categories lexically specified for morphological features (e.g. Case-features) must move to a position where these features can be checked off. Making the assumption that Agr mediates the checking of Case-features, subject and object undergo the following movements triggered by the need to check Case-features.[7]

(6)



The subject NP_1 moves to the Spec-AgrSP to have its Nominative Case (henceforth, [+Nom]) checked off whereas the object NP_2 moves to the Spec-AgrOP to have its Accusative Case (henceforth, [+Acc]) checked off.[8] Further, following the insight of Chomsky & Lasnik (1993), I assume that each feature-checking operation (i.e. each feature-checking movement) forms a distinct chain.[9] Given this chain-formation algorithm, there are two distinct chains in (6): C_1 (NP_1, t_1) formed by the checking of [+Nom] and C_2 (NP_2, t_2) formed by the checking of [+Acc]. These two chains are associated with NP_1 and NP_2, respectively. Also, under this analysis, two distinct chains can be associated with a single category. Consider the following:

(7) a. *Who left?*

b. 

There are two chains in (7b): C_1 (who_1, t'_1) formed by the checking of [+Wh] and C_2 (t'_1, t_1) formed by the checking of [+Nom]. These two distinct chains are formed by the two distinct feature-checking movements of the single category, namely who_1.

---

7. This paper ignores the movement of heads involved in the checking of Case features, since it is irrelevant to our discussion. For the precise formulation of the theory of feature-checking, see Chomsky (1993).

8. In this paper, I adopt a simplified nominal structure, namely, NP. This simplification does not affect any of the arguments presented below.

9. For relevant discussion of this point, see also Epstein (1993), Kitahara (1992, 1994).

4. Articulated Chain-Structure and the Scope Principle

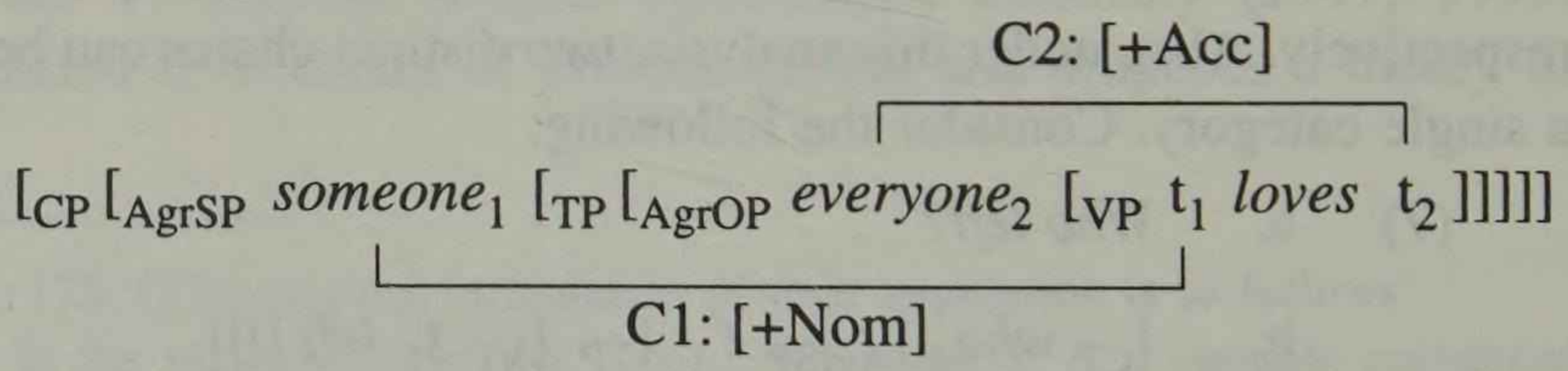
As reviewed in Section 3, within the theory of feature-checking which incorporates the chain-formation algorithm, categories lexically specified for morphological features must undergo feature-checking operations, each of which triggers a movement forming a distinct chain. Thus, one or more chains can be associated with a single category. Assuming this theory of feature-checking, I propose the following natural revision of A&L's (1989, 1991) Scope Principle (4):

(8) Scope Principle (revised)
A quantifier X may take scope over a quantifier Y iff X c-commands a member of *each* chain associated with Y at LF.

I will now demonstrate that A&L's stipulation concerning NP-trace follows naturally from the revised Scope Principle (8). Moreover, within the analysis proposed, we shall see that the LF-rule of QR plays absolutely no role.

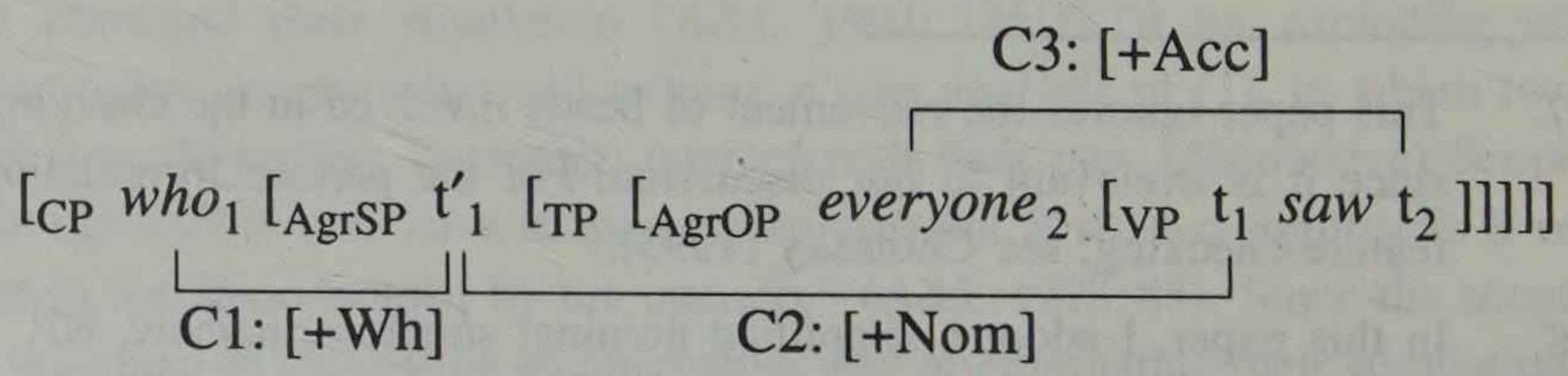
Recall the (scope-ambiguous) sentence (1). The movements triggered by the need to check Case-features (e.g. [+Nom], [+Acc]) produce the following LF-structure of (1):

(9)



There are two distinct chains in (9): C_1 (someone$_1$, t_1) and C_2 (everyone$_2$, t_2). *Someone$_1$* c-commands both members of C_2, and *everyone$_2$* c-commands a member of C_1 (i.e. t_1). Thus, the revised Scope Principle (8) predicts the scope ambiguity in (9). Now recall the sentence (5a), exhibiting the obligatory wide-scope reading of *who* (which, recall, was a problem for A&L's analysis). The movements triggered by the need to check [+Nom], [+Acc], and [+Wh] produce the following LF-structure of (5b):

(10)



There are three chains in (10): C_1 (who_1, t'_1), C_2 (t'_1, t_1), and C_3 ($everyone_2$, t_2). Notice that C_1 and C_2 are each associated with who_1. Therefore, according to the revised Scope Principle (8), in order for $everyone_2$ to have scope over who_1, $everyone_2$ must c-command a member of C_1 and a member of C_2. Although $everyone_2$ c-commands a member of C_2 (i.e. t_1), $everyone_2$ fails to c-command any member of C_1. Thus, the absence of a wide-scope reading of $everyone_2$ in (10) follows from the revised Scope Principle (8). Now, does the revised Scope Principle (8) correctly predict the wide-scope reading of who_1 in this example? Since who_1 c-commands both members of C_3 (containing $everyone_2$), the wide-scope reading of who_1 in (10) follows naturally from the revised Scope Principle (8). Crucially, notice that A&L's stipulation concerning NP-trace is entirely eliminated; and, more importantly, the LF-rule of QR plays absolutely no role in the analysis of scope interpretation presented here. That is, within the theory of feature-checking which incorporates the chain-formation algorithm, the chain-based theory of scope interpretation (incorporating the revised Scope Principle (8)) renders the LF-rule of QR unnecessary to explain the core cases of scope phenomena discussed here.

The proposed analysis of scope interpretation further explains the following seemingly paradoxical contrast (see among others, Lasnik & Saito 1992):

(11) a. *Who do you think everyone saw?* *who > < every*
b. *Someone thinks everyone saw John.* *some > every*

In (11a), the embedded subject *everyone* can (optionally) have scope over *who* occupying the Spec of the matrix CP; whereas in (11b), the embedded subject *everyone* cannot have scope over the matrix subject *someone*.[10] Under the proposed analysis, (11a, b) are assigned the following LF-structures, respectively.[11]

---

10. More precisely, (11a) can be interpreted as a single question (collective reading, i.e. *who* taking wide-scope), whose appropriate answer would be e.g. "I think everyone saw John", or as a distributive question (distributive reading i.e. *everyone* taking wide-scope), whose appropriate answer would be e.g. "I think Mary saw Tom, Sally saw Bill, Nancy saw John ..." On the other hand, (11b) can be interpreted only as "there is a particular person who thinks everyone saw John" (collective reading i.e. *someone* taking wide-scope) (see among others, May 1977, 1985).

11. The NP-movement of *John* (triggered by the checking of [+Acc]) is omitted from (12b) since it is irrelevant to our discussion.

(12) a. $[_{CP}$ *who*$_2$ *do you think*

C3: [+Nom]

$[_{CP}$ $[_{AgrSP}$ *everyone*$_1$ $[_{TP}$ $[_{AgrOP}$ t′$_2$ $[_{VP}$ t$_1$ *saw* t$_2$]]]]]]

C1: [+Wh] C2: [+Acc]

b. $[_{CP}$ $[_{AgrSP}$ *someone*$_1$ $[_{TP}$ $[_{VP}$ t$_1$ *thinks*

C1: [+Nom]

$[_{CP}$ $[_{AgrSP}$ *everyone*$_2$ $[_{TP}$ $[_{AgrOP}$ $[_{VP}$ t$_2$ *saw John*]]]]]]]]

C2: [+Nom]

There are three chains in (12a): C_1 (*who*$_2$, t′$_2$), C_2 (t′$_2$, t$_2$), and C_3 (everyone$_1$, t$_1$).[12] *Who*$_2$ c-commands both members of C_3, and *everyone*$_1$ c-commands a member of each chain associated with *who*$_2$; that is, *everyone*$_1$ c-commands a member of C_1 (i.e. *t′*$_2$) and two members of C_2 (i.e. *t′*$_2$, *t*$_2$). Thus, the revised Scope Principle (8) correctly predicts the scope ambiguity in (12a). In (12b),

---

12. Notice that in (12a), the tail of the chain created by the checking of [+Wh] occupies the Spec-AgrOP at LF (i.e. *t′*$_2$ of C_1). The following derivation illustrates how this chain-structure of a *wh*-phrase bearing both [+Wh] and [+Acc] is formed:

(i) a. I wonder $[_{CP}$ $[_{AgrSP}$ John $[_{TP}$ $[_{AgrOP}$ $[_{VP}$ bought what]]]]]

b. *Overt Wh-Movement:*

I wonder $[_{CP}$ **what**$_1$ $[_{AgrSP}$ *John* $[_{TP}$ $[_{AgrOP}$ $[_{VP}$ *bought* **X**$_1$]]]]]

C1: [+Wh]

c. *LF-Object Shift of* X_1*:*

I wonder $[_{CP}$ *what*$_1$ $[_{AgrSP}$ *John* $[_{TP}$ $[_{AgrOP}$ **X**$_1$ $[_{VP}$ *bought* **Y**$_1$]]]]]

C1: [+Wh] C2: [+Acc]

I assume that in the overt syntax, *what*$_1$ undergoes the movement triggered by the checking of [+Wh] (i.e. syntactic WH-movement), leaving the trace X_1 (bearing [+Acc]) in its departure site (i.e. the complement of *bought*), as in (ib). Then, in the LF-component, the direct object trace X_1 undergoes movement triggered by the checking of [+Acc] (i.e. LF-object-shift of X_1), leaving the trace Y_1 in its departure site. That is, the tail of the chain created by the checking of [+Wh] itself moves to the Spec-AgrOP (see also Epstein 1993 for relevant discussion of such articulated chain-structures).

there are two chains: C_1 (someone$_1$, t_1) and C_2 (everyone$_2$, t_2). While *someone*$_1$ c-commands both members of C_2, *everyone*$_2$ fails to c-command any member of C_1. Thus, the revised Scope Principle (8) correctly predicts the obligatory wide-scope reading of *someone*$_1$ in (12b). As shown above, the contrast in (11) follows naturally from the revised Scope Principle (8). Notice that, here again, the LF-rule of QR plays absolutely no role.

5. Concluding Remarks: Is QR Eliminable?

As we have seen in the preceding sections, the core cases of scope phenomena discussed here were shown to follow from the revised Scope Principle (8) without stipulation, and crucially without any appeal to the LF-rule of QR. One remaining question is: Is QR eliminable? Suppose that the apparent obligatoriness of QR is to be deduced from LF-principles, which require quantifiers to occupy non-θ-marked positions (the Predicate Condition, May 1977) and/or to have propositional scope i.e. to have scope over (at least) VP, given the VP-Internal Subject Hypothesis (Tancredi 1990). Then, for at least argument-quantifiers bearing the structural Case-feature: [+Nom] or [+Acc], which (respectively) occupy the Spec-AgrSP and the Spec-AgrOP at LF, QR is unnecessary since these LF-principles are already satisfied by virtue of Case Theory. Given Economy of Derivation (see among others, Chomsky 1991, 1993; Epstein 1992; Kitahara 1994), which disallows superfluous steps in derivations, such quantifiers are then prohibited from undergoing (unnecessary) QR.[13]

Although further research is necessary to provide a more comprehensive determination of the precise empirical content of feature-checking theory, and hence of the overall role of QR within this theory, the analysis of scope interpretation presented here, reveals that independently motivated feature-checking operations (i.e. the checking of [+Nom], [+Acc], and [+Wh]) proposed within the minimalist program, generate structures quite similar to those QR was formulated to produce. The precise extent (beyond that presented here), to which feature-checking operations subsume QR, is a question I leave for future research.

---

13. For relevant discussion concerning problematic aspects of QR, see Hornstein (1994), Lasnik (1993), and Vanden Wyngaerd & Zwart (1991).

Acknowledgements

This is a revised version of Kitahara (1992). I would like to thank Noam Chomsky, Howard Lasnik, Mamoru Saito, Höskuldur Thráinsson, and especially Samuel D. Epstein for valuable comments and helpful discussion. In addition, I am grateful to Jun Abe, Erich Groat, Ruriko Kawashima, Masatoshi Koizumi, Susumu Kuno, Shigeru Miyagawa, Yoichi Miyamoto, Masayuki Oishi, Toshifusa Oka, Geoff Poole, Shigeo Tonoike, Akira Watanabe, Jan-Wouter Zwart, and participants in the Harvard Syntax Practicum, Spring 1992 for clarifying remarks and helpful suggestions. All remaining errors are, of course, my own.

Optional Movement in the Minimalist Program

Geoffrey Poole
Harvard University

Notice that this approach tends to eliminate the possibility of optionality in derivation. Choice points will be allowable only if the resulting derivations are all minimal in cost.... This may well be too strong a conclusion, raising a problem for the entire approach.

— Chomsky (1991: 433)

Abstract

This paper proposes a solution to the problem of derivational optionality in the Minimalist Program. Beginning with Chomsky's (1993) distinction between the singularly Generalized Transformation (Move α) and the basic transformational operation (Form Chain), I show that only applications of Form Chain are counted by Economy of Derivation. Thus any operation that consisted of Move α alone, without Form Chain, is cost-free. Prior to the examination of any data, I explore the properties that Move α without Form Chain is predicted to possess. I then examine two optional movements, Icelandic Stylistic Fronting and "semantically vacuous" cases of Japanese scrambling, and show that they possess precisely the predicted properties.

1. Introduction

One potentially serious problem facing current work in the Minimalist Program (Chomsky 1991, 1993, 1995a) is the question of derivational optionality. As Chomsky (1991) notes, given the principle of Economy of Derivation, optionality should only be allowed in those cases in which there are two or more derivations which are equally minimal in cost. However, there exist constructions in natural language in which the optionality appears to reside in moving an element as opposed to leaving it *in situ*. The existence of this sort of optionality

in derivation leads Chomsky (1991:433) to observe that "[allowing optionality in derivation only when both derivations are equally economical] may well be too strong a conclusion, raising a problem for the entire approach."

In this paper, I argue that in fact Economy of Derivation allows a derivation D in which an element E undergoes movement to be just as economical as the corresponding derivation D′ in which E is not moved. New phenomena concerning Icelandic Stylistic Fronting (see Maling 1990; Jónsson 1991; Poole 1992; Anderson 1993) are presented and unified with Saito's (1989, 1992) analysis of semantically vacuous $\bar{A}$-movement in Japanese. Unification is obtained under the assumption that such optional movements leave a trace, but fail to form a chain with their trace. This contrasts with non-optional movements, which leave a trace and *do* form a chain. I argue that this distinction between optional (non-chain forming) and obligatory (chain-forming) movement can be captured by appealing to a distinction independently motivated within the Minimalist Program between Move α (the singularly Generalized Transformation) and Form Chain (the basic transformational operation).

I begin in Section 2 with a discussion of Economy of Derivation, and demonstrate that it counts applications of Form Chain, rather than Move α. Thus any operation involving *only* Move α, and *not* Form Chain, is cost-free. The remainder of the section examines what it would mean for there to be an application of Move α only (and not Form Chain) and what properties the theory predicts such movement must have (crucially, prior to the examination of any data whatsoever). Section 3 contains an analysis of two optional movements in natural language, Icelandic Stylistic Fronting and Semantically Vacuous $\bar{A}$-movement in Japanese, and argues that they possess precisely the properties entailed by the analysis in Section 2. Section 4 concludes.

2. Move α, Form Chain, and Optionality

Given the principle of Economy of Derivation, a derivation in which an element has moved can be just as economical as the corresponding derivation in which the element was left in-situ only if the movement is cost-free. In this section, I propose a new, explanatory definition of "cost-free" movement, and

discuss some of the properties that the theory forces "cost-free" movement to have.[1]

2.1. *Economy of Derivation: Move α versus Form Chain*

Chomsky (1993: 15) argues that the basic transformational operation must be Form Chain, rather than Move α. As he notes, there is a seeming paradox between two natural interpretations of Economy of Derivation: fewest moves vs. shortest moves. A derivation which has fewer moves will naturally have chains with links that are farther apart and a derivation in which the chain links are as close together as possible will have a greater number of moves. Chomsky notes that this paradox can be resolved by taking the basic transformational operation to be Form Chain, an operation in which chains of arbitrary length with intermediate traces included are created in "a single step". However, it is not the case that Move α has been eliminated from the theory. Chomsky (1993: 22) writes, "Alongside the binary substitution operation [Generalized Transformation], which maps (K, K′) to K*, we also have the singularly substitution operation Move α, which maps K to K*".

One way this can be interpreted is to say that Form Chain merely "hooks together" moved categories with traces that have been created by iterated applications of Move α. That is to say, Move α moves an element, leaving behind a trace of that element. This process is repeated as many times as desired, creating a representation in which there are lexical items and some number of traces. Then Form Chain targets a lexical item and any of its copies and creates from this set of categories a new syntactic object, namely a chain.[2,3]

---

1. For a different account of optional movement, see Fukui (1993).

2. I assume that it is in fact not necessary to stipulate that Form Chain specifically targets a lexical item and the traces of that very lexical item. Presumably, various conspiracies will ensure that the derivation will crash if Form Chain targets, for example, two lexical items or a lexical item and a trace that is not a copy of it, but rather a copy of something else. However, this is an empirical issue, and as such I leave it open.

3. An anonymous reviewer notes "[a]nother way to think of it would be to say that intermediate traces are generated by binary operations, and that the movement operation in fact takes only one step." A thorough investigation of the empirical difference between the two proposals, which would center on the presence of phonologically unrealized intermediate traces in intermediate levels of representation, would be well

If both Move α and Form Chain are operations, a question arises as to which operation is counted by Economy of Derivation.[4] I argue that it must be the case that Form Chain is the operation counted by Economy of Derivation, rather than Move α. Otherwise, all derivations involving successive cyclic movement would be prohibited by virtue of being less economical than their (ECP/Subjacency-violating) one-fell-swoop counterparts. Take, for example, the sentence in (1), with two possible representations in (2):

(1) *Why do you think Bill hit Barney?*

(2) a. $[_{CP}$ *Why*$_i$ *do* $[_{IP}$ *you think* $[_{CP}$ $[_{IP}$ *Bill hit Barney* t_i]]]

b. $[_{CP}$ *Why*$_i$ *do* $[_{IP}$ *you think* $[_{CP}$ t'_i $[_{IP}$ *Bill hit Barney* t_i]]]

Both the $\overline{A}$-chain in (2a) and the $\overline{A}$-chain in (2b) are the result of one application of Form Chain. Furthermore, both derivations in (2) converge. No unchecked morphological features remain and all chains are uniform. However, the one-fell-swoop case in (2a) requires only one application of Move α, whereas the successive cyclic movement in (2b) requires two applications of Move α. If Economy of Derivation counted applications of Move α instead of Form Chain, the one-fell-swoop derivation in (2a) would block the successive-cyclic derivation in (2b).[5] However, (2a) violates the Minimal Link Condition (Chomsky 1993, 1995a) and the resulting prediction is that (1) is strongly ungrammatical — clearly the wrong result.

One might be tempted to say here that, although (2a) is the more economical convergent derivation (if Economy of Derivation counts applications of Move α), (2b) will actually be preferred because it violates fewer principles of grammar (in this case, none). However, this way out is not possible. The most economical convergent derivation, whatever it is, must *always* be picked. Consider the following LF representation:

(3) * $[_{CP}$ *Why*$_i$ *were* $[_{IP}$ *you wondering* $[_{CP}$ *whether* $[_{IP}$ *Bill said* $[_{CP}$ *t'_i $[_{IP}$ *Mary left* t_i]]]]

---

beyond the scope of this paper, and for this reason I will leave the issue open.

4. I leave aside the possibility that Economy of Derivation counts neither Form Chain nor Move α.

5. I assume that the zero-move derivation (in which *why* remains *in situ*) is not a member of Economy of Derivation's competitor or "reference" set (Chomsky 1995a) since it fails to converge. (The strong *+wh* morphological feature on *why* is never checked.) Hence the shortest derivation that Economy of Derivation considers is the one-move one.

If it were the case that Economy of Derivation allowed one to pick less economical derivations if the more economical ones violated principles of grammar, then the prediction would be that Economy of Derivation would allow deletion of t'_i, with the result that (3) would have the status of a mere Subjacency violation, and not an ECP violation.[6,7] However, (3) is strongly ungrammatical and appears to have the status of an ECP violation. This suggests that the most economical convergent derivation must *always* be picked.

Therefore, the fact that the sentence in (1) above is perfect, and not an ECP violation, indicates that the derivation in (2b) must be at least as economical as the derivation in (2a). Hence, Economy of Derivation cannot be counting applications of Move α. If, on the other hand, Economy of Derivation counts applications of Form Chain, then (2a) and (2b) are equally economical, and the fully grammatical status of (1) is explained.

Therefore, any operation that consisted only of Move α, without Form Chain, would not be counted by Economy of Derivation, and hence would be cost-free. This would then allow a derivation in which some element has been moved to be as economical as the corresponding derivation in which the element was left *in situ*.

2.2. *The Question of Greed*

In additional to Economy of Derivation, Chomsky (1993) introduces

---

6. Recall from Chomsky and Lasnik (1993: 547) that Subjacency violations occur if a starred trace is created during the course of a derivation and is subsequently deleted. Additionally, an ECP violation occurs if a starred trace is created and remains at LF.

7. An anonymous reviewer notes:

> In the MPLT approach (basing this on class lectures) [the status of examples like (3)] has nothing to do with the ECP, but with difficulties in interpreting the moved element (i.e., linking it with its trace). The ECP–Subjacency distinction then reduces to an argument adjunct distinction (barring additional locality factors).

I take the reviewer to be referring here to Chomsky's suggestion in class lectures that argument/adjunct asymmetries in *wh*-island extraction configurations are not due to the ECP, but rather arise from some kind of parsing effect. Since no discussion of this point has appeared in any published work of Chomsky's of which I am aware, and since I have no particular suggestions to make concerning this alternate analysis, I will leave it open.

another principle which constrains movement operations: the principle of GREED, given in (4):[8]

(4) *Greed* (Chomsky 1993: 33)
Move α applies to an element α only if morphological properties of α itself are not otherwise satisfied.

Although Chomsky (1993: 33) suggests that Greed might be reducible to Economy of Derivation, I believe that both principles should be addressed in the context of optional movement. I argued in section 2.1 that an operation which consisted of Form Chain without Move α is cost-free; although the principle of Greed, if correct, would still require that there be some morphological property of the optionally moved element which is satisfied.

However, there is some question as to whether or not Greed is obeyed in optional movement. In Section 2.3 below, I argue that applications of Move α without Form Chain undergo obligatory LF "reconstruction". More precisely, I assume the element which undergoes optional movement is obligatorily deleted in the LF component, leaving the copy in the base-position to be interpreted. This has the result of making optional movement semantically vacuous, since the derivation in which optional movement took place is LF-identical (though not PF-identical) with the derivation in which no movement took place. It seems methodologically suspect in the absence of independent motivation to postulate a functional head bearing a strong feature that must be checked off, but which has no effect on interpretation. It would seem preferable to say that optional movement of the sort under consideration here is not driven by Greed.

Providing independent empirical motivation for this conclusion is not entirely straightforward. However, given Chomsky's (1993) division between Move α and Form Chain, together with the copy theory of movement, there is a plausible conceptual argument: Under the copy theory of movement, there is a close relationship between feature-checking and chain-formation. A trace of a moved element is precisely identical with that moved element, save for the fact that the trace is unpronounced. Thus, whatever features are being targeted for

---

8. The relationship between Economy of Derivation and Greed is not entirely clear. Chomsky (1993: 33) suggests that Greed might ultimately reduce to Economy of Derivation: "the notion of Last Resort operation [of which Greed is a subcase — GP] is in part formulable in terms of economy: a shorter derivation is preferred to a longer one, and if the derivation D converges without the application of some operation, then that operation is disallowed."

movement are present both on the moved element and on its copy. Applying Form Chain to the moved element and its copy welds both sets of features into a single syntactic object, the chain; allowing both sets of features to be checked off. To take a concrete example, if Tense were to adjoin to AgrS without forming a chain with the trace of Tense, there would be *two* elements in the structure bearing strong Noun features, and unless *both* were somehow checked off, the derivation would crash. But if Tense did form a chain with its trace, then anything in a checking relationship with the head of Tense would be sufficient to check off the features of the tail.

The Minimalist Program distinction between Move α and Form Chain provides a natural expression for this close relationship between feature-checking and chain-formation: Greed constrains applications of Form Chain, but not applications of Move α. If this were correct, we would then expect there to be movement operations which were not feature-checking, namely those that consisted of Move α without Form Chain.

2.3. *The Predicted Properties of "Move α without Form Chain"*

I have argued above that optional movement of an element E can be allowed, provided that is an instance of Move α without Form Chain. In this subsection, I investigate the properties that such an operation would have. First, the theory predicts that optional movement must obligatorily undergo LF "reconstruction", at least for optional movement of any θ-role assignor or assignee. Second, the theory predicts that optional movement of a verb does not extend the minimal domain of that verb (in the sense of Chomsky 1993).

2.3.1. *LF "Reconstruction"*

The first property that the theory predicts optional movement should possess is LF "reconstruction". However, following Chomsky (1993), I assume that literal reconstruction does not take place. Rather, under the "copy" theory of movement, LF "reconstruction" consists of deletion of the optionally moved element in the LF component, leaving only the copy in the original position. This has the effect of "undoing" or "reconstructing" the original movement.

If an element E which assigns or is assigned a θ-role is moved by an application of Move α, but never "hooked up" with its trace by Form Chain, the moved element must be deleted, in accordance with Full Interpretation, as the moved element will not constitute a legitimate LF-object.

Deletion is required because, if Form Chain never applies to the moved element E and its trace, marking them as being members of a single chain, both elements will be treated as separate, unrelated one-membered chains. For the trace, residing in the element's base position, this is unproblematic. The element can assign or receive its θ-role in the normal manner, just as if the element had never moved. However, since the moved element is not a chain-mate with its trace, but rather is a completely separate entity, the moved element *also* will either require a θ-role or have one to discharge. Since it is no longer in a position to either receive or assign a θ-role, a θ-Criterion violation will result.

Following Chomsky (1995a: 429), I assume that a θ-Criterion violation causes the derivation to crash (rather than to "converge as gibberish"). If this is the case, the Last Resort Principle will license and require deletion of the offending element prior to the LF interface:

(5) *Last Resort Principle* (Chomsky 1993: 32)
A step in a derivation is legitimate only if it is necessary for convergence—had the step not been taken, the derivation would not have converged.

Of course, in the vast majority of cases in which a θ-Criterion violation occurs (e.g., *John is likely that Bill left*), deletion of the offending element is blocked by the principle of Recoverability of Deletion, and the derivation has no choice but to crash. However, in the case of an element that has undergone optional movement, the deletion is fully recoverable, since there is a copy of the deleted element that is already in exactly the correct position for interpretation, namely the base-position. Therefore, in the case of optional movement of an element that either requires a θ-role or has θ-roles to discharge, the deletion of the moved element, necessary to save an otherwise crashing derivation, is not blocked by any principle.[9,10]

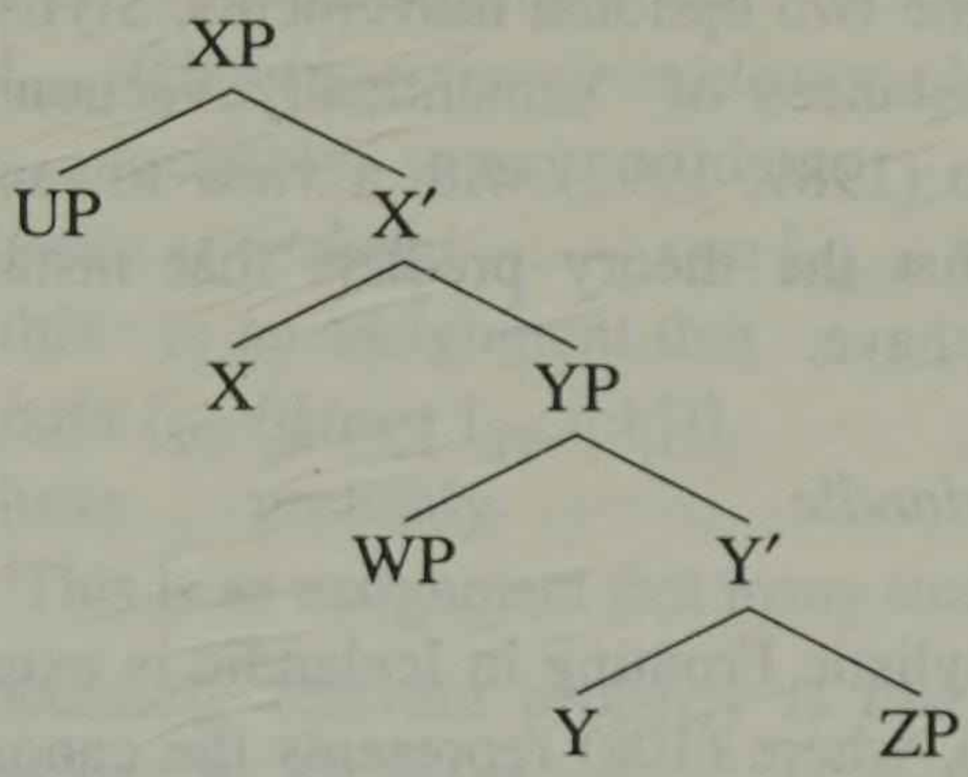
---

9. All cases of optional movement to be discussed in this paper do in fact involve elements that either require a θ-role (optional movement of an NP) or have θ-roles to discharge (optional movement of a verb). A case of optional movement of an element that neither requires a θ-role nor has one to discharge (say, optional movement of an adverb) would be slightly more complicated. It might be the case that the derivation also crashes if precisely the same element is interpreted as being in two different places, no matter what sort of element it is. One would have to see if the arguments advanced in Chomsky (1995a: 426–429) concerning θ-Criterion violations carry over to "interpretive" or "predication" violations generally. The required argumentation would be subtle however, and I must, with regret, leave it for further research.

2.3.2. *Domain Extension*

A second property that Move α without Form Chain should possess is the inability of a verb to extend its domain. This is because extending a verb's domain requires forming a chain with a trace, which is exactly what *fails* to happen in the case of Move α without Form Chain. To illustrate, consider the general schema in (6) and the relevant definitions in (7)–(9):

(6)



(7) "Let us define the domain of [a chain $(\alpha_1,...,\alpha_n)$] to be the set of nodes contained in $Max(\alpha_1)$ and not containing any α_i." Chomsky (1993: 13)

(8) "For any set S of categories, let us take Min(S) (minimal S) to be the smallest subset K of S such that for any $\gamma \in$ S, some $\beta \in$ K reflexively dominates γ." Chomsky (1993: 12)

(9) X reflexively dominates Y iff X dominates Y or X = Y.

If Y is not a member of a chain with X, the members of the domain of Y are WP, ZP and anything they dominate. YP, Y′, and Y are not members of the domain of Y since they all contain Y. Therefore, the minimal domain of Y is exactly {WP, ZP}. However, if Y is in a chain with X, for example, if X is a verb and Y is the trace of that verb, the minimal domain of the *chain* (X, Y) would be "extended", since it now would include UP (as well as WP and ZP). But, if X moved from the position of Y, left a trace in the position of Y, but failed to form a chain with that trace, there would be no minimal domain

---

10. If the deletion operation is able to target features within a lexical item, it could be the case that only those features that are causing the derivation to crash are deleted; in this case, perhaps the item's semantic features. However, this is an empirical issue, and I leave it open.

"extension"; that is, no minimal domain that included both ZP and UP. This has consequences for the relationship between verb raising and the movement of ZP, a matter to which I return in the following section.

3. Two Optional Movements

In this section, I examine two optional movements, Stylistic Fronting (SF) in Icelandic and certain instances of "semantically vacuous" scrambling in Japanese discussed by Saito (1989, 1992) with a view toward whether or not they have the properties that the theory predicts that instances of Move α without Form Chain should have.

3.1. *Stylistic Fronting in Icelandic*

The phenomenon of Stylistic Fronting in Icelandic is exemplified by word order contrasts such as (10), where (10a) represents the canonical word order, and (10b) the order after Stylistic Fronting has applied.[11]

(10) a. *Þetta er versta bók sem hefur verið skrifuð.*
this is the worst book that has been written

b. *Þetta er versta bók sem* ***skrifuð*** *hefur verið.*
this is the worst book that written has been
'This is the worst book that has (ever) been written'

What about with Aux?

There is no meaning difference whatsoever associated with Stylistic Fronting. The sentences in (10) are simply considered stylistic variants. Crucially, there is not even any associated emphasis in (10b). Thus SF appears to be a different phenomenon from Topicalization (which is also found in Icelandic; see Maling 1990 for extensive discussion concerning the differences between SF and Topicalization). I take this to be convincing evidence that SF is a truly optional, i.e. non feature-driven, movement.

There are two central properties of SF that are relevant for my analysis. First, there is a restriction on the distribution of elements that can occur in the subject position of the finite clause within which Stylistic Fronting has applied.[12]

---

11. Here and below, the Stylistically Fronted element is in boldface.

12. Following Jonas (1992) and subsequent work, I assume that subjects in Icelandic can move to either Spec-AgrSP or Spec-TP. I refer only to Spec-AgrSP here and below for typographical convenience only.

The descriptive generalization, due to Maling (1990), is that Stylistic Fronting requires a "subject gap". SF is licit when the subject remains in Spec-VP (11a) (possible only if the subject is a member of a certain class of indefinites). However, raising of the subject to Spec-AgrSP results in ungrammaticality (11b).[13]

(11) a. *Þetta er verkefni sem* [$_{AgrSP}$ ***lokið***$_j$ *hafa*
this is an assignment that finished have
[$_{VP}$ *líklega* [$_{VP}$ *margir stúdentar* t_j]]].
probably many students

b. **Þetta er verkefni sem* [$_{AgrSP}$ *margir stúdentar*$_i$ *lokið*$_j$
this is an assignment that many students finished
hafa [$_{VP}$ *líklega* [$_{VP}$ t_i t_j]]].
have probably
'This is an assignment that many students have probably finished.'

The second immediately relevant property is that SF is an instance of head movement, as opposed to XP-movement.[14] As evidence, notice that in (12) SF appears to have fronted the adjective head of an AP small clause.[15] In (13), SF has moved a past participle, stranding the past participle's direct object.[16]

---

13. The examples are due to Höskuldur Thráinsson (personal communication).

14. I follow Jónsson (1991) here, contra Maling (1990). See Poole (1992) for a more comprehensive discussion concerning the bar-level of the category which undergoes SF.

15. Assuming that *be* takes a small-clause complement.

16. A third property of SF, noted first by Maling (1990), is the presence of an accessibility hierarchy. Although both negation and past participles undergo SF, (iia) is the only grammatical application of SF from the source sentence (i) ((iia–c) are taken from Maling 1990: 31):

(i) *Þetta er glæpamaðurinn sem hefur ekki verið dæmdur.*
this is the criminal that has not been convicted

(ii) a. *Þetta er glæpamaðurinn sem* ***ekki*** *hefur* t *verið dæmdur.*

b. **Þetta er glæpamaðurinn sem* ***verið*** *hefur ekki* t *dæmdur.*

c. **Þetta er glæpamaðurinn sem* ***dæmdur*** *hefur ekki verið* t.

I have no particular proposal to offer regarding this property of SF (but see Jónsson 1991 for an attempt to deduce facts like those in (iia–c) from the Head Movement Constraint). Note however, that the existence of Maling's accessibility hierarchy might argue against an analysis like the one suggested by Noam Chomsky (personal communication) in which SF is movement which takes place in the PF-component.

(12) *Þeir sem* ***ánægðir*** *eru* [$_{AP}$ t *með kaupið*] *kvarta ekki.*
those who content are with the pay complain not
'Those who are content with the pay do not complain.'

(13) *Þetta er maður sem* ***leikið*** *hefur* [$_{VP}$ t *níutíu leiki*].
this is a man that played has ninety games
'This is a man that has played ninety games.'

The contrast in (11) seems to indicate that something prevents NPs from moving from Spec-VP to Spec-AgrSP when SF has applied. One possible explanation is that Spec-AgrSP is already filled by the Stylistically Fronted element. However, under the assumption that SF is head-movement, rather than XP-movement, I assume that this possibility is excluded.[17]

If it is not the case that Spec-AgrSP is filled, then it must be the case that movement to that position is blocked. The central principle in the Minimalist Program by which movement is licensed is Greed, repeated below:[18]

---

17. Notice further that those adverbial elements which can intervene between the subject and the finite verb in Icelandic cannot intervene between the SF'ed element and the finite verb. In (i), we see that the adverb *kannski* 'possibly' can either precede or follow the finite verb. However, in (ii), having left the subject *margir Íslendingar* 'many Icelanders' in Spec-VP, we see that when SF of the past participle has applied, *kannski* can appear only after the finite verb. It cannot appear between the SF-ed element and the finite verb. (The examples are due to Höskuldur Thráinsson; personal communication.)

(i) a. *Þetta er fugl sem* [$_{AgrSP}$ *margir Íslendingar kannski hafa*
this is a bird that many Icelanders possibly have
[$_{VP}$ *oft* [$_{VP}$ *séð*]]]
often seen

b. *Þetta er fugl sem margir Íslendingar hafa kannski oft séð.*
'This is a bird that many Icelanders have possibly often seen'

(ii) a. **Þetta er fugl sem séð kannski hafa* [$_{VP}$ *oft* [$_{VP}$ *margir Íslendingar]]*

b. *Þetta er fugl sem séð hafa kannski* [$_{VP}$ *oft* [$_{VP}$ *margir Íslendingar]]*

I take this as further evidence that the SF'ed element is not moving to the subject position. I assume, following Jónsson (1991), that the SF'ed element is adjoined to the verb.

18. I put aside here the reformulation of Greed that would be necessary given the discussion in Section 2.2 above, as it is not relevant for the point at hand.

(4) *Greed* (Chomsky 1993: 33)
Move α applies to an element α only if morphological properties of α itself are not otherwise satisfied.

If SF somehow were to render Spec-AgrSP a non-Case-checking position, then movement from Spec-VP to Spec-AgrSP would correctly be prohibited. One way in which this possibility might be implemented would be to say that the Stylistically Fronted element, a head adjoined to Tense, creates a new morphological entity—crucially one which is not a Nominative Case-checker. Thus, movement from Spec-VP to Spec-AgrSP would not be movement to check a morphological feature and would be prohibited.[19]

However, recall again (11a) above. If the subject can surface in Spec-VP, SF is possible. This suggests that while *overt* movement of the subject to Spec–AgrSP may be blocked by Greed (the movement is not to a feature-checking position), *covert* movement at LF must not be blocked. Otherwise the NP *margir stúdentar* in Spec-VP would never get its Case feature checked, and the grammatical (11a) would incorrectly be predicted to crash at the LF interface. Since, by hypothesis, it is the presence of the Stylistically Fronted element that blocks nominative Case-checking, deletion of the Stylistically Fronted element in the LF component would "restore" a T^0 capable of checking Nominative Case, and correctly account for the grammaticality of (11a).[20,21]

Recall now from Section 2 that general considerations concerning what movement operations exist lead us to the conclusion that optional movement, if correctly analyzed as Move α without Form Chain, would result in the moved category being deleted in the LF component, provided that it was either a θ-role assigner or assignee. In (11), the Stylistically Fronted element is a verb, a θ-role

---

19. SF must also, of course, "weaken" the strong N features on Tense, so that no strong features remain at PF, causing the derivation to crash.

20. In fact, I believe that the exact nature of the problem induced by SF (i.e. blocking nominative Case-checking by T versus e.g. blocking agreement by AgrS) is not relevant for my analysis. The crucial fact is that whatever the Case or Agreement-checking problem turns out to be, checking seems to be able to take place at LF — a fact to which I return below.

21. These properties of SF are purely stipulative at this point, but seem to be well-motivated both conceptually and by the data. As always with stipulations of this nature, I assume that they are ultimately deducible from independently motivated principles, but I must leave this question for further research.

assigner. The grammaticality of (11a) indicates that the Stylistically Fronted element does in fact undergo obligatory LF deletion and our prediction is thus confirmed.

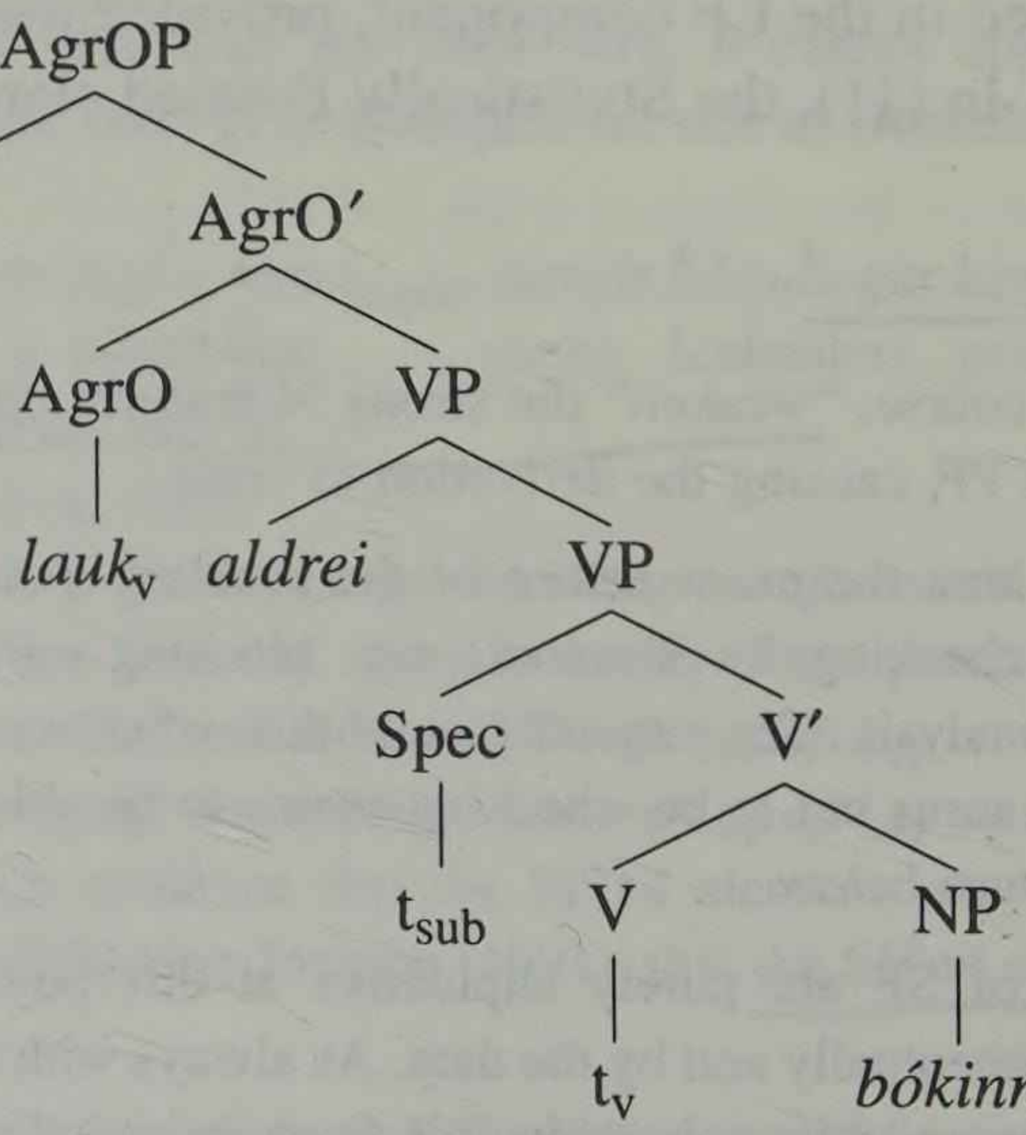
In addition, because this case of SF happens to be movement of a verb, we can test the second prediction from Section 2 above. If, as I propose, SF is correctly analyzed as an instance of Move α without Form Chain, then SF of a verb should not license object shift. This is because object shift is permitted only if the verb's domain is extended (in the sense of Chomsky [1993]) and chain-formation is a prerequisite for domain extension.

As first noted by Holmberg (1986), Icelandic allows full NP direct objects to move to a position on the left periphery of the VP. Following Vanden Wyngaerd (1989a) and others, I assume that this movement is to Spec-AgrOP (within the expanded system of inflection proposed by Chomsky 1991 — based originally on proposals by Pollock 1989), as indicated in (14):

(14) *Þetta er maðurinn* [$_{CP}$ *sem lauk*$_i$ [$_{AgrOP}$ *bókinni*$_j$
this is the man that finished the book
[$_{VP}$ *aldrei* [$_{VP}$ t_i t_j]]]
never
'This is the man that never finished the book.'

Holmberg further observed that OBJECT SHIFT of this sort was possible only when verb raising had taken place. An explanation for this descriptive generalization was offered by Chomsky's (1993) "Equidistance" analysis (see the "Introduction" to this volume). Looking at (14), before the object shifts, but after verb raising to AgrO has taken place, we have the partial structure in (15):

(15)



The problem is as follows: the object must move to Spec-AgrOP, but the closest po-

tential landing site is Spec-VP. Moving directly to Spec-AgrOP would thus result in an economy violation, as the object *bókinni* would have failed to make the shortest move.

However, crucially involved in calculating whether an element has made the "shortest move" is the notion "equidistance".[22]

(16) If α, β are in the same minimal domain, they are equidistant from γ.

Since *lauk* has raised from V to AgrO, the minimal domain of the chain (*lauk*, t) is Spec-AgrOP, Spec-VP and the NP *bókinni*. Therefore, since Spec-AgrOP and Spec-VP are both "equidistant" from *bókinni*, it can move directly to Spec-AgrOP while still making "the shortest move".

All of this is directly relevant to SF and the question of optional movements. If SF is correctly analyzed as an instance of Move α without Form Chain, then, unlike the case of overt verb raising in (14), SF of a verb does not form a chain with the trace left behind. Therefore, there can be no "domain extension". Spec-VP and Spec-AgrOP will not be equidistant from the object position. Therefore, the prediction is that movement of the object to Spec-AgrOP should not be possible when the verb has undergone SF. Although this prediction is a little difficult to test, since Icelandic has obligatory V2 both in main and embedded clauses, consider the following contrast:[23]

(17) *Þetta er maðurinn* [$_{CP}$ *sem* ***ljúka***$_{v}$ *vildi* [$_{CP}$ [$_{VP}$ *aldrei*
this is the man that finish wanted never
[$_{VP}$ t_v *bókinni*]].
the book
'This is the man who wanted to never finish the book.'

(18) **Þetta er maðurinn* [$_{CP}$ *sem* ***ljúka***$_{v}$ *vildi* [$_{CP}$ [$_{AgrOP}$ *bókinni*$_{i}$
this is the man that finish wanted the book
[$_{VP}$ *aldrei* [$_{VP}$ t_v t_i]].
never
'This is the man who wanted to never finish the book.'

SF out of a control clause is possible, as in (17), but in (18), Object Shift within the control clause is not possible when the verb has undergone SF. If SF involved movement that formed a chain, *ljúka*'s minimal domain should have been extended, just as in the case of object shift discussed above. The fact that it is not further suggests that SF is an instance of Move α without Form Chain.

---

22. For a different "equidistance" metric, see Ferguson (this volume).

23. See Thráinsson (1993a) for a discussion of SF in control clauses.

3.2. *Semantically Vacuous Ā-Movement in Japanese*

Saito (1989, 1992) argues that some instances of Japanese scrambling similarly represent optional, semantically vacuous movement. Japanese, unlike English, is a language which has *wh*-in-situ, rather then overt *wh*-movement. Thus the scope of a *wh*-element is not determined by the position of that element, but rather by the position of the Q-morpheme *ka*. At LF, the *wh*–element must move to the specifier of the CP headed by *ka*.

Saito argues that the trace left by LF wh-movement is subject to the Proper Binding Condition (PBC), which requires that a trace be c-commanded by its antecedent. He notes the following contrast (his examples are from Harada 1972):

(19) a. [$_{IP}$ *Hanako.ga* *Masao.ni* [$_{CP}$ [$_{IP}$ *dare.ga* *kuru*] *ka*] *osieta*]
Hanako.NOM Masao.to who.NOM come Q taught
koto
fact
'the fact that Hanako told Masao [Q [who is coming]]'

b. *[$_{IP}$ *Hanako.ga* *dare.ni* [$_{CP}$ [$_{IP}$ *Masao.ga* *kuru*] *ka*] *osieta*]
Hanako.NOM who.to Masao.NOM come Q taught
koto
fact
'the fact that Hanako told who [Q [Masao is coming]]'

This pattern is explained under the assumption that, at LF, the wh-element must move to the Specifier of the CP headed by the Q-morpheme, and that the movement is constrained by the Proper Binding Condition. In (19a) the wh-element *dare* moves to a position that c-commands its surface position. Therefore the trace that is created by LF movement will be c-commanded by its antecedent. However, in (19b), *dare* would be moving downward, to a position which does not c-command its surface position. Therefore, the trace created by LF movement will not be c-commanded by its antecedent, in violation of the PBC.

But now, observe the following contrast:

(20) a. *Masao.ga* [$_{CP}$ [$_{IP}$ *Hanako.ga* *dono* *hon.o*$_i$ *tosyokan.kara*
Masao.NOM Hanako.NOM which book.ACC library.from
karidasita] *ka*] *siritagatteiru*] *koto*
checked.out Q want.to.know fact
'the fact that Masao wants to know [Q [Hanako checked out which book from the library]]'

b. ?*dono hon.o*$_i$ [*Masao.ga* [$_{CP}$ [$_{IP}$ *Hanako.ga* t$_i$ *tosyokan.kara*
which book.ACC Masao.NOM Hanako.NOM library.from
karidasita] *ka*] *siritagatteiru*] *koto*
checked.out Q want.to.know fact
'the fact that which book$_i$, Masao wants to know [Q [Hanako checked out t$_i$ from the library]]'

Saito notes that (20b), which is only marginally ungrammatical, appears to exhibit exactly the same configuration as (19b), which is strongly ungrammatical. In both instances, a *wh*-element is, on the surface, in a position which is not c-commanded by the Specifier of the CP to which it must move at LF. The prediction would then be that (20b) should have the same status as the strongly ungrammatical (19b). But Saito claims that (20b) is far better than (19b) and does not appear to have the status of a PBC violation.

Saito's conclusion from this is that the scrambled *wh*-phrase *dono hono* must be able to move downward to a position inside the c-command domain of the Q-morpheme at LF. As he points out (1992: 86) "...if this is the case, nothing seems to prevent [scrambled phrases] from moving back to their D-structure positions in LF."[24]

However, under the COPY theory of movement (which I assume, but which Saito does not), another analysis consistent with the data becomes available. It could be the case that it is the *copy* of the *wh*-phrase which is moving to the Specifier of the CP headed by *ka* (recall that the copy of the *wh*-phrase is identical to the original in every respect, except that it is unpronounced). But (20b) is not interpreted as a multiple *wh*-question, only as a single one. Therefore, the sentence-initial *wh*-phrase must have been deleted somewhere prior to LF.

Once again, we have a convergence of evidence. The leading idea throughout this subsection has been that Japanese scrambling, like SF in Icelandic, is an instance of Move α without Form Chain. General considerations concerning the nature of such an operation led us to the conclusion that the moved element would have to undergo LF-deletion. What we find, upon examination of one particular case of (optional) Japanese scrambling, is that the scrambled element in fact appears to undergo LF deletion.

---

24. This assumes (contra Chomsky 1993) that trace-leaving is optional, not obligatory. Otherwise all cases of downward movement would result in PBC violations.

4. Conclusions and Questions for Further Research

In this paper I have attempted to resolve the problem observed by Chomsky (1991: 433) concerning optional movements where the optionality resides in moving an element as opposed to leaving it *in situ.* I have argued that movements of this sort are correctly analyzed as instances of Move α without Form Chain. Prior to the examination of *any* data, I was able to determine what properties that the theory would require Move α without Form Chain to have. I then turned to Stylistic Fronting in Icelandic and Japanese Scrambling, two optional movements, and argued that they have precisely the properties predicted.

Since the primary goal of this paper has been to develop a theory of optional movement within the Minimalist Program, the central question for further research concerns the application of this theory to other cases of optional movement (see Kim; in progress, and the references cited there). However, there would seem to be two sources for optionality that should in principle be distinguished. One is true optional movement of the sort described in this paper. Such movement would, by hypothesis, be characterized by the failure to form a chain, semantic vacuity, LF "reconstruction", and so forth. There is, however, another potential source of optionality, namely optional introduction of a functional head to which movement is obligatory. Those cases of optionality which are not semantically vacuous, and which do seem to have an effect on binding relations, etc. might well be instances of the second sort. However, a fuller consideration of these issues awaits further research.

Acknowledgements

I would like to express my appreciation to Erich M. Groat, Dianne Jonas, Hisatsugu Kitahara, John O'Neil, Eríkur Rögnvaldsson, an anonymous reviewer, and especially Samuel D. Epstein and Höskuldur Thráinsson for their many helpful comments. An earlier version of this paper was presented at the 1994 Western Conference on Linguistics, held at the University of California, Los Angeles, and I thank the audience there for much helpful discussion. Any remaining errors, however, are my own.

Morphology and Word Order in Germanic Languages

Jaume Solà
Autonomous University of Barcelona

UG provides a unique computational system, with derivations driven by morphological properties to which syntactic variation of languages is restricted.

— *Chomsky (1993: 44)*

Abstract

This paper makes a general proposal on the relation between inflection and syntax, which constitutes an alternative instantiation of Chomsky's (1993) idea that word order variation across languages can be (at least partly) explained by relating it to morphological variation. Unlike in Chomsky's minimalist proposal, though, I will argue that the crucial parameter is not weak/strong morphological features, but rather presence/absence of inflectional morphology for a given functional category on a given word. This account crucially relies on close morphological analysis of inflected words. I will also adopt Kayne's (1994) antisymmetry hypothesis, both because it sharply narrows down morphologically-driven word-order predictions, and because it allows a more elegant account of West Germanic word order alternations, in the spirit of Zwart (1993b).

1. Introduction

One salient idea in the Minimalist program advanced in Chomsky (1993) is that syntactic representation (conceived as the interface with interpretation — LF[1]) essentially does not vary across languages. Word order, however, varies

---

1. It is not entirely clear to me that Chomsky's proposal is not bound to do "some" syntax at (or before) PF (the phonetic interface). In any event, the present proposal allows just one syntactic representation.

because, other things being equal, there is variation as to whether movement is visible (takes place before Spell-out) or not. Movement, in turn, is conceived as necessary in order for the morphological features of inflected lexical items to be checked against the corresponding features in the functional heads. Given the sentence structure in (1):[2]

(1) $[_{AgrSP}\ AgrS^0\ [_{TP}\ T^0\ [_{AgrOP}\ AgrO^0\ VP]]]$

the inflected verb (the head of VP) has to move, overtly or covertly, to the functional heads $AgrS^0$, T^0 and $AgrO^0$ to check its agreement and tense features against the corresponding features in these heads. Similarly, the NP-arguments in the VP have to move to the specifiers of the functional projections in order to check their Case and phi-features against the corresponding functional features. Thus differences in word order between languages are predicted as a result of the different options of moving, overtly or covertly, the inflected verb and the NPs to the different heads and specifiers.

According to Chomsky, whether movement is overt or covert is ultimately derived from the strength of morphological features: strong features trigger overt movement, while weak features do not (and economy ensures that overt movement is not even an option for weak features). This is a welcome alternative to the earlier Government-Binding idea that both raising and lowering are available options for putting together (functional) affixes and (lexical) stems. In Chomsky's words "we can eliminate S-structure conditions on raising and lowering in favor of morphological properties of lexical items" (Chomsky 1993: 31).

The predictive power of such a proposal depends on whether we can reduce other possible factors for word order variation to a minimum: if we allow for many other factors (the head-initial/final parameter, language specific rules of adverb placement, etc.), there will be a good deal of indeterminacy for the derivation of a given pattern of word order, which will pose a problem for learnability. In this sense, a proposal like Kayne (1994), which provides a highly constrained X-bar structure and does not allow for parameterization of the order of constituents, constitutes a good background against which the minimalist program can be implemented. A theory combining the minimalist program with

---

2. In Chomsky's proposal, movement (at least overt movement) and structure-building transformations are manifestations of a single syntactic device, the Generalized Transformation. Here I abstract away from this fact, as has become usual practice.

Kayne's proposal has already been fruitfully developed in Zwart (1993b), among others. Here I will adopt Kayne's proposal, and some of Zwart's crucial ideas on the nature of verb second in West Germanic.

I will concentrate mainly on head movement, and will illustrate my proposals mainly in West Germanic languages and English. In section 1, I will review Chomsky's theoretical assumptions concerning head-movement parametrization and highlight some problems in it. In Section 2, I will introduce a general definition of what inflection is and which kinds of variation we can expect depending on whether functional categories correspond to bound morphemes (inflection), free morphemes (usually "short" particles), or are (morphematically) absent (as, for example, tense in Chinese). I will also introduce a definition of how this variation predicts word order. The essential idea is that an inflected word overtly moves if and only if it contains overt morphology of the target of movement; and that there is no covert movement. In Section 3 I will address the issue of English lexical verbs, and provide an alternative account of their special behavior with respect to verb movement. My view is that the minimalist account of the contrast between French and English (based on Pollock 1989) overstates the crosslinguistic representativeness of the English case, which I believe is a very special case. In Section 4, I will apply and slightly adapt my proposal in Section 2 to account for the main/embedded clause word order alternations in Dutch and German. This account is compatible with Kayne's (1994) antisymmetry hypothesis. Section 5 is a tentative proposal on how the antisymmetry hypothesis, combined with a closer morphological analysis, can derive the apparent (and traditionally assumed) head-finalness of West Germanic IPs (or VPs). Finally, I end with a conclusion and some speculations.

2. Strong/Weak Features and Functional Projections

In Chomsky's (1993) proposal, whether a lexical head (or an NP) moves overtly or covertly depends on whether the corresponding functional features it has to check in the functional domain are strong or weak, respectively. As features are checked, they are eliminated. A strong feature must be checked (eliminated) overtly because otherwise it would be present in the phonological component (PF, the output of Spell-out), where it would be an illegitimate object. Weak features, instead, are harmless (by definition) at PF, and can be eliminated covertly, only before LF. Economy principles (procrastination) ensure

that, if they need not be eliminated overtly, they cannot be. In this way, strong/weak features uniquely determine overt/covert movement, respectively.

The idea that strength of features is the factor that determines movement can be traced back to Pollock's (1989) proposal (which was further developed and modified in Chomsky 1991) that inflectional heads are either transparent or opaque, and only the former allow movement of lexical verbs. Although Pollock is not very explicit about the issue, his suggestion seems to be that transparency/opacity is related to the robustness of the inflectional paradigm. In this way, French Agr (= AgrS in Chomsky 1993) is transparent because the verbal agreement in this language displays more morphological distinctions than the English one, which is opaque. Although this suggestion is vague, it constitutes an attempt at providing independent motivation for the strong/weak parameter, on the basis of observable morphology.[3]

In the minimalist proposal, no attempt is made at providing independent motivation (morphological or other) for the weak/strong parameter. It is clear, however, that if no independent motivation is provided, deriving overt/covert movement from the strong/weak character of features does not constitute an explanation, but simply a metaphoric restatement of the facts. In other words the claim that movement is driven by the weak/strong parameter runs the risk of becoming circular, if any observable case of overt/covert movement can be accommodated by an empirically innocuous decision on the strength of the corresponding features.

There is a closely related aspect of Chomsky's proposal that deserves comment. The idea that inflected lexical heads have to move to the functional heads to check their functional features looks like a natural solution to the mismatch between the proposed sentence structure, where functional categories project separately from the lexical ones, and inflection, which, under the Strong Lexicalist Hypothesis, attaches the functional morphemes to the lexical head. It

---

3. Also see Roberts (1993: Chapt. 5) and Vikner (1995b: Chapt. 5) for some interesting attempts at establishing correlations between the strength of agreement features and the robustness of agreement paradigms (number and quality of the distinctions). I cannot do justice to these proposals here. Let me just say that my proposal does not have room for robustness: a feature (such as person) is present or not (however impoverishedly if present). My speculation is that these authors do not bear in mind other Infl features which crucially distinguish Germanic languages: Mood (present in Icelandic, but not in Scandinavian), Aspect (whose morphological expression varies across these languages), etc.

is clear, however, that Chomsky's proposal pays little attention to inflectional morphology: a finite verb is assumed to always be inflected for AgrS, T and AgrO, whether this is observable or not in the verb morphology.

The proposal in this paper is that features are present in an inflected word *only if* they are observable in the morphological alternations of the paradigm the word belongs to. As a simple illustration, the French verbal form *mangeons* 'we eat' contains agreement features, as it alternates with *mangez* 'you eat' and *mange(s/nt)*, the default form ('non first/second-plural'). The English verbal form *ate*, instead, does not contain any agreement features, as it does not contrast with any other (past) form for these features. In the remainder of this paper, we will provide evidence that overt movement (the only one allowed in this proposal in the sense we will make precise) correlates with the presence of morphologically detectable features.[4]

Finally, there is another aspect of the minimalist program that sooner or later has to be revised: it rather crucially relies (for minimality effects, for instance) on the existence of three functional categories (AgrS, T, and AgrO). On the one hand, most languages seem to display a richer ontology of inflectional paradigms (which includes Mood and Aspect, at least). In some languages, as we will see, even Comp is an inflectional category.[5] On the other hand, it is not entirely clear to me that agreement should constitute an independent category: Subject agreement is suspiciously closely related to Tense (morphemes for AgrS and Tense are very often merged into a single morph). Subject agreement could be conceived as a morphological manifestation of T (perhaps the expression of

---

4. An anonymous reviewer points out as a problem for the present proposal that non-finite verbs undergo movement in Icelandic control complements. And this is also true of infinitives in most Romance languages. I do not see this as a problem: infinitives are plausibly inflected for Mood (they are, roughly speaking, non-finite subjunctives). What would constitute a problem, then, is just the opposite: an infinitive-marked verbal form that does not move. Notice that in English, the infinitival marker is not morphological, it is a free particle (*to*), which appears higher than the verb (as split infinitives show). How to deal with infinitives in other Germanic languages is beyond the scope of this paper.

5. In view of this, I think Chomsky's concept of L-RELATEDNESS (which tries to derive the traditional A/$\overline{\text{A}}$-distinction) is not morphologically motivated: Comp can be an inflectional category (an affix), and tense can be a free particle (as is the Case in the Haitian creole). I do not mean that L-relatedness cannot be defined somehow, but the situation where the Infl morphemes are bound to the verb while the Comp ones are not seems to be just an option (attested in most Indo-european languages).

Nominative assignment or checking). An alternative idea is that both tense and person agreement are deictic categories, the most basic ones (their anchoring relies on who speaks and when). This would somehow make them an interpretatively significant unit.[6] In any event, details of definition aside, I will assume, for reasons that will become crucial in Section 4, that Agr is part of the T morphology.

3. Inflection and Its Relationship to Syntactic Structure

The present proposal, like the minimalist program, is crucially based on the Strong Lexicalist Hypothesis: both derivational and inflectional morphology are dealt with in the lexicon. What is then the difference between inflection and derivation? Derivational morphology is a (more or less regular and productive) compositional device for obtaining new lexical items: as we combine morphemes or stems, we obtain a new item the meaning of which is obtained compositionally (*pace* regularity/opacity). Inflection is different. Let me illustrate in what sense. Consider the sentence in (2), with the constituent bracketing roughly defining the relevant structure:

(2) *I wonder...* [$_{CP}$ [$_{C^0}$ *if*] [$_{IP}$ *he* [$_{I^0}$ *will*] [$_{VP}$ *be happy*]]]

In (2), it is reasonable to assume that both syntactic and semantic compositionality are expressed by this structure: *will* is a head (a functor) which takes the proposition [x be happy] as its complement (argument) and gives as output a "future-marked" proposition. *If*, in turn, takes the proposition *he will be happy* as its complement (argument), and gives an embedded *yes/no* question (or a conditional sentence) as output.[7] Now consider the Italian and Irish verbal forms in (3):[8,9]

---

6. See Laka (1990: Chapt 1) for arguments that subject agreement is parasitic on Infl.
7. Of course alternatives are available for expressing compositionality much more independently of syntactic structure: function composition, as in categorial grammar, or feature percolation as in HPSG: but they would just waste the opportunity the GB-like structure has of being compositional in a trivial and straightforward way.
8. The Irish data are freely adapted from Ó Siadhail (1989).
9. Even though *an mbeidh* is spelled as two words, it is a single word, as we will see directly.

(3) a. Italian: *sarà* 'he.will.be'
b. Irish: *an mbeidh* 'if.s/he.will.be'

We can see that the counterparts of morphemes that are free in English (*if*, *will*) are bound morphemes in these languages. Now, if we assume that Italian and Irish have essentially the same sentence structure as English, we have a mismatch between syntax/semantics (which require different positions for the morphemes 'if', 'will' and 'be') and inflectional morphology, which glues (some of) these morphemes together in a single word (we will call words that contain morphemes of different categories MULTICATEGORIAL words). The idea is that this mismatch is the trigger for head movement: since multicategorial words contain morphemes of several categories, it should be possible to somehow tear these morphemes apart in order for them to appear in their respective syntactic positions. Suppose this is feasible in the following way:

(4) In order to insert a multicategorial word in a syntactic structure:
a. Insert a copy of this word in each of the positions it contains features of.
b. In each copy, read only the relevant features, and ignore the other features.
c. Pronounce only the highest copy.[10]

Let me illustrate these definitions with an example. Take Italian *sarà* 'will be' and Irish *an mbeith* 'that (he) will be'. Since they are multicategorial (they contain the morphemes future and 'be'; and 'yes/no', future and 'be, respectively'),[11] in accordance with (4a) they will be multiply inserted, with a copy in each relevant position, as illustrated in (5). In accordance with (4b), each feature will be read in its relevant position, while, in accordance with (4c) only the highest position (Comp in Irish; Future in Italian) will be pronounced (the non-pronounced copies are represented as scratched out). The conditions in (4)

---

10. In Kayne's (1994) syntactic structure, the highest copy will also be the first copy in the relevant head-movement cases. So an alternative option could be "Pronounce the first copy." I will not be dealing with any cases where this distinction makes a difference.

11. Here I abstract away from matters of detail: future seems to belong to Mood rather than Tense; it is not clear whether *be* heads a VP or stands for some Functional Category. The important point is that they should be inserted in the same positions as the corresponding English free morphemes ('that', 'will', and 'be')

should be understood as both obligatory and exhaustive (as if-and-only-if conditions):[12]

(5)	a.	*if*	*John*	*will*	*ever*	*be*	*happy*
	b.	*se*	*Gianni*	*sarà*	*mai*	~~*sarà*~~	*contento*
		if	Gianni	will.be	ever	happy	
	c.	*an mbeidh*	*Seán*	~~*an mbeidh*~~	*ariamh*	~~*an mbeidh*~~	*sásta*
		if.will.be	Seán		ever		happy

In a way, this is just a different implementation of the traditional idea of movement: the different copies will constitute a Chain, which will be subject to locality conditions (which I will not discuss here).[13] This notion of movement has at least two nice properties for comparative linguistics: (a) it allows for a cross-linguistically uniform syntactic (and semantic) analysis and (b) it predicts word order variation, as can be seen in (5).[14]

---

12. Irish adverbs such as *ariamh* 'ever' tend to appear in sentence final position more often than their English or Italian counterparts. This is an independent factor I ignore here.

13. The present proposal is *not* derivational. I will nevertheless use the term movement as a descriptive term for convenience, but there is no question here as to which copy is inserted first or how the structure is built. In other words, this proposal is compatible with a constraint-based approach to syntax, where it does not matter how structure is built, provided it abides by all principles and constraints. If, however, the reader finds it easier to think derivationally, nothing in this paper will very much interfere with this view. My point would rather be that derivationalism (and much of economy) is unnecessary and, besides, for reasons I cannot develop here, highly risky of becoming untractably intricate.

    Furthermore, this is *not* a checking theory of head-movement: a word "moves" in order for all morphemes it contains to be able to appear where they belong, not to check any independently inserted features. In fact, Chomsky's checking theory is lexicalist in a very special way: all of the functional terminal nodes are initially filled with non-lexical material. The checking theory is computationally very baroque and mostly empirically vacuous (if a morphological feature does not match a silent functional one, the derivation crashes for no observable reason). In this proposal, head adjunction is entirely dispensed with (and also XP adjunction, on the lines of Kayne 1994): head chains are a set of simple terminal nodes.

14. Chomsky (1993) suggests that VSO Celtic languages are to be analyzed as having the verb in Infl and the subject in its basic position. This is hardly sustainable if we use Pollock's (1989) adverb and negation placement tests: both adverbs (and negation in colloquial Modern Welsh) follow the subject, which is almost strictly adjacent to the verb (as is usually the case in V-to-Comp structures: **Where does generally John stay?*). The present proposal more accurately predicts these facts.

This notion of head movement, however, is intended to be much more restrictive than the traditional notion based on Move α: overt head movement is triggered if and only if a word is multicategorial (contains morphemes that belong in different syntactic positions). The point I want to emphasize is the following: in order for this proposal not to become circular, movement, on the one hand, and the multicategorial status of a word, on the other hand, have to be independently motivated on empirical grounds. So we cannot postulate that a word is multicategorial if the only piece of evidence for this claim is that it moves (overtly).[15] In other words, we need some independent criteria to identify movement, on the one hand, and the morphemes which are conveyed by a word, on the other hand. We therefore also need independent criteria to decide what a word is. The prediction should then be that, once they are independently identified, overt movement and word morphology correlate.

To identify what is a word is no easy task. For the purposes of this paper, and restricting myself to inflection, I will keep on the safe side and assume that the following are sufficient (even if not necessary) clues for the word status of a cluster of morphemes:

(6) Two or more morphemes are in the same word if they form a cluster which cannot be split by any phrasal constituent and some of the following phenomena are detectable:

a. The boundaries between them show word-internal (as opposed to word-boundary) phonological phenomena.

b. Some of them trigger allomorphy on some of the others.

c. They are not segmentable (there is no one-to-one correspondence between morphemes and morphs).

Consider, in the light of (6), the Irish expression *an mbeidh*: even if the spelling suggests that we are dealing with two words, (6) tells us otherwise: complementizer particles are strictly adjacent to the verb and trigger allomorphy (suppletion) on it, in addition to initial consonantal mutations:[16]

---

15. Or, conversely, we could not assume (even if we allowed covert movement, which we do not) that it moves covertly if the only evidence is that it is (supposedly) inflected for a feature, with no morphological evidence for the presence of this inflection.

16. In fact, in both Modern Irish and Modern Welsh, some of these particles tend to disappear, so that what is left is just the initial mutation on the verb and the verb suppletion. This further reinforces the single word status of the expression.

(7) a. *a bhi*
'that.s/he.was' (relative clause)
b. *an raibh:*
'if.s/he.was' (*yes/no* question)

In addition, the complementizer particle sometimes redundantly expresses the past/present morpheme (*go* = 'that + present'; *gur* = 'that + past'). If allomorphy and morpheme overlapping are inflectional phenomena to be dealt with in the lexicon, then these forms can only be inserted as a unit in syntactic structure. I will propose that head movement is the only option to interpret them.

Once we have identified words, we have to identify which morphemes they contain. There is no guarantee that this can always be done by morphological parsing (segmentation), for the reasons just mentioned. The criterion I will use is the following:

(8) A morpheme (feature value) can be identified in a word if this word minimally contrasts with another word for this morpheme (feature value).[17]

The set of all forms that contrast for a feature value we call a *paradigm*. In the following sections I will present evidence and arguments to support the view that, once we identify the morphemes and paradigms in words, we can make predictions about head movement.[18]

---

17. A word A minimally contrasts with a word B with respect to a feature α if A and B contain exactly the same feature values except for the value of α. Example: Italian *sei* 'you are' and *eri* 'you were' minimally contrast for tense: both are 2SG, indicative and imperfective of 'be', but one is present and the other is past.

18. Notice that in the present proposal it is immaterial which morpheme is the head of an inflected (multicategorial) word (which may of course be a morphophonologically meaningful question), since the several morphemes it conveys are not hierarchically organized (they are just stuck together, and will be read independently each in a different copy). Traditional wisdom would have it that the head of an inflected verb is the "verb" itself, but this is circular: in my view crosslinguistically a verb should be defined as any set of words which can be grouped into paradigms for sentence functional category morphemes (this is why *to cross* is a verb and *across* is not). A lexical verb will then be a verb of which one of the morphemes is a lexical predicate (as opposed to an auxiliary verb, which would only convey functional morphemes). For convenience, I will keep talking of Vs and VPs as lexical categories, but I think notions such as "category V" are derivative, crosslinguistically non-uniform, and ultimately epiphenomenal (it is impossible to universally define necessary and sufficient conditions for verbhood).

In order to make predictions about movement, we have to establish which Functional Categories are present in the structure (and in which hierarchy) and which morphemes correspond to them, not a trivial task. For the sake of argumentation and illustration, and given limitations of space, let me just sketchily introduce a rough proposal on what the functional structure of a sentence must look like in order to provide room for the functional paradigms we can identify in English. The English sentence seems to involve at least the following functional paradigms:[19]

a. Complementizer: *if*, *that*, possibly a null complementizer for main clauses.
b. Mood: *can*, *may*, *will*, etc. The future marker (*will*) is in this paradigm, rather than in the Tense paradigm, which is for the [±Past] feature.
c. Tense (± Past): its morphemes are always bound to modals (*can/could*, *will/would*, etc.) or to other verbs (*is/was*, *has/had*). In the next section I will argue that in English lexical verbs are not really inflected for Tense or subject agreement.
d. Relative tense: its only morpheme is auxiliary *have* (i.e., the one that is reducible in all dialects, *not* 'main verb' *have*), which, in Reichenbachian terms, conveys the meaning e > r (event time precedes reference time). Its null counterpart (absence of *have*) would convey the default value: e = r.
e. Aspect: the participial morphemes *-ed* and *-ing* (As in *John has worked/John is working*) are likely candidates for this category. In the next section we will make a crucial proposal on aspectual participles.
f. Telicity (Aktionsart): particles such as *up*, *down* in *He ate it up*, *He wrote it down* seem to have a telicity import (they convey a telic interpretation). I will later suggest that, since objects play a role for telicity, the Telicity functional category is what would correspond to Chomsky's AgrO.

To exemplify how these morphemes can combine in a sentence structure, let me illustrate it with the following example, where I stick to the assumption that inflected words are inserted in several positions:

(9)

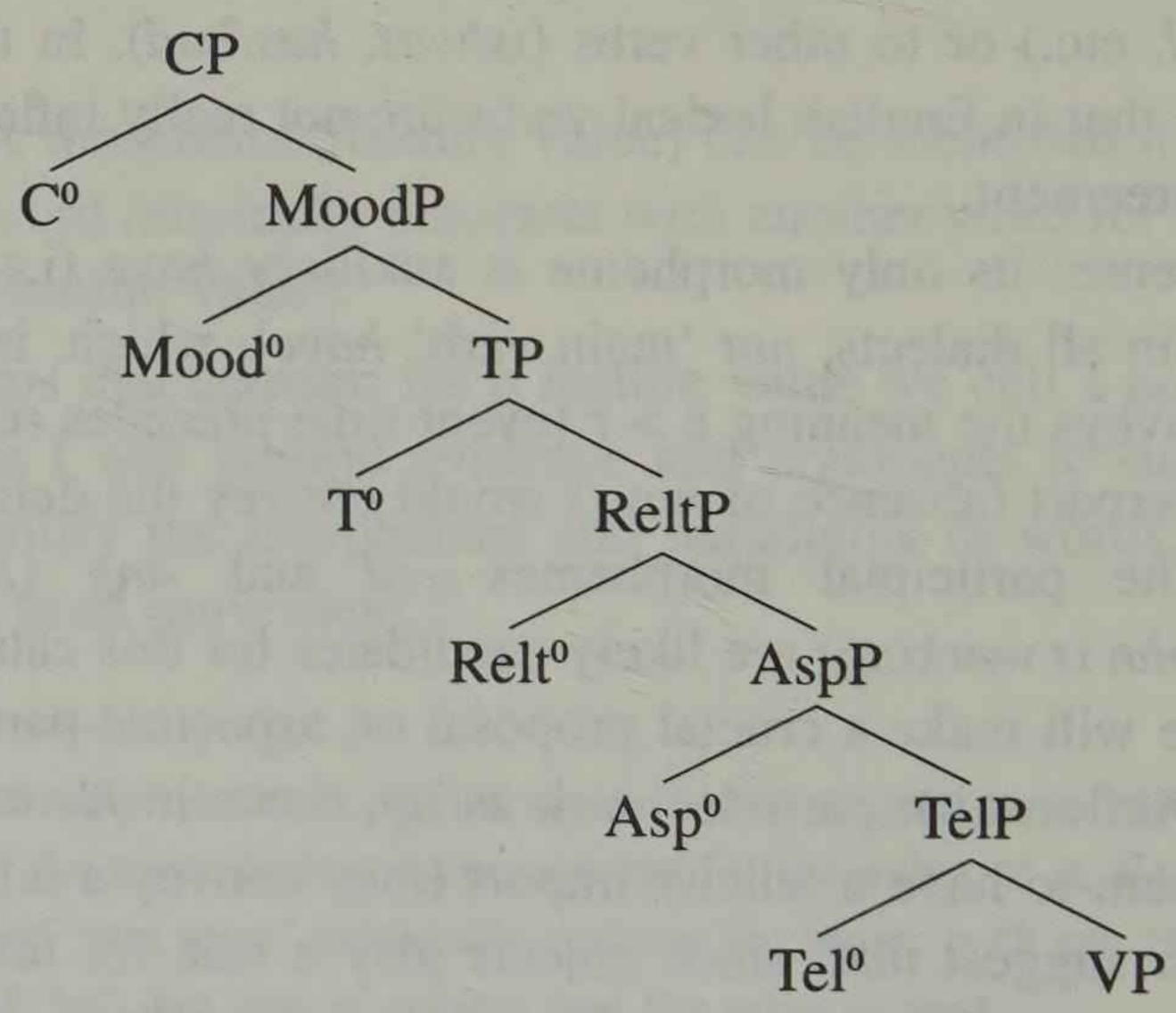
$Comp^0$	$Mood^0$	$Tense^0$	$Relt^0$	Asp^0		Tel^0	V^0
if	he	would	~~would~~	have	eaten	it up	~~eaten~~

---

19. A very enriched ontology of sentence functional categories has been advanced by Cinque in a course at the Girona International Summer School in Linguistics (1994), where he motivated it mainly on the basis of one to one correspondence between types of adverbs and the functional categories they can be specifiers of.

In (9) there are only two inflected (multicategorial) words, *would* ('will' + Past) and *eaten* (lexical predicate + perfective aspect). It is beyond the scope of this paper to extensively argue for this view on the functional domain of the English sentence. However, it is not controversial that the structure of the English sentence has to allow for a good deal of this structural complexity. Whether we call Relative Tense, Aspect or Telicity functional categories or not might turn out to be, to some extent, a terminological issue. For instance, one can chose to label as VPs the phrases headed by auxiliary verbs like *have*, but they are still functional in the reasonable sense that they are structurally higher than lexical verbs and behave like propositional functors, not as predicates. If we adopt Kayne's (1994) proposal about X′-structure, the above proposal on functional morphemes can only be given a structure as in (10) (specifiers omitted):

(10)



where the order of constituents is fixed. We will see that a good deal of Germanic word order can be derived from this sentence structure plus the present proposal concerning head movement.

If we compare the present proposal with Chomsky's minimalist proposal, one could translate the latter in terms of the former by saying that strong features, which trigger (overt) movement would correspond to features that are present in an inflected word, while weak features would be features that are simply absent in the relevant inflected word.[20] But if weak features are "absent" in the inflected word, there is no reason why they should trigger covert

---

20. A similar conclusion is reached by Groat & O'Neil (this volume).

movement, whatever "covert movement" might correspond to in the present approach.[21] So, in the present proposal, as we will see, there is no need for covert movement, and "weak"/"strong" have no theoretical status. In addition, the present proposal does not define functional morphemes as essentially inflectional (as Chomsky's checking theory suggests): there is variation as to whether a functional morpheme is free or bound, and in the latter case, as to which other morpheme(s) it is bound to. And, as we will see, there is also variation as to whether a functional morpheme is overtly present at all in a given language.[22] Thus, considering, for a simple case, the relation between V and T, there can be three situations: (a) V is inflected for tense, and there is overt movement (a copy of the verb is inserted in both T and V); (b) the verb is not inflected for tense because the tense morpheme is a free particle: there will be no movement of the verb to T; the free particle itself will be inserted under T; and (c) There is no tense morpheme in the language, and there will not be no movement either (possibly a null tense morpheme will be inserted in T).

With these ideas as background, let us proceed to more specific and controversial cases.

4. English Lexical Verbs

This section deals with a very specific (and problematic) case: the special behavior of English finite lexical verbs as opposed to that of auxiliaries with respect to head movement. Chomsky's minimalist proposal presents the contrast between English and French as a paradigmatic instance of a UG parametric option (weak vs. strong V-features). I think that English is a very special case in this connection, and it should be dealt with accordingly. If it were true that English verbs do not move to Infl despite their finite status, this would be a fatal counterexample for the present proposal. I will therefore challenge the idea that

---

21. In Section 5 I will propose that movement is not always visible, but this is not because it is delayed to some later stage of the derivation: it is rather a case where the highest copy is not the one to be pronounced.

22. This opens the question whether, in the absence of (overt) morphemes for a given functional category, the category itself is present or not (for instance: does Chinese project a TP, if it has no tense morphology or particle?). I would assume it does, since specifiers seem to be available (rather crucially so in Kayne's 1994 proposal) in the absence of an overt head morpheme. But nothing hinges on the issue in this paper.

apparent finite lexical verbs in English are indeed finite.

Pollock (1989), based on Emonds (1978), convincingly argued, on the basis of word order facts, that English lexical verbs differ from French lexical verbs in that only the latter move to inflectional categories, skipping over some adverbs and the negative morpheme *pas*:

(11) a. *Jean voit souvent Marie.*
b. *John often sees Mary.*
c. **John sees often Mary.*

(12) a. *Jean (ne) voit pas Marie.*
b. **John sees not Mary.*

Consider first French. It can easily be shown that the French finite forms can convey morphemes of at least the following categories (I exemplify it with the verb *aller* 'go'. I use the 3rd person singular form throughout):[23]

	Tense	
	Present	Past
Mood		
Indicative	*va*	*allait*
Subjunctive	*aille*	*(allât)*
Future	*ira*	*irait*

All these forms, in addition, have agreement distinctions, which we take (not crucially yet) to be a further manifestation of Tense morphology. Then we can reasonably expect that these verbal forms move to the Functional heads that host these features, Mood and Tense, and that this movement skips over the adverbs and negation.

Before addressing the issue why the English lexical verbs do not behave similarly, let us consider some historical facts. In Middle English, lexical verbs behaved very much like in French:

I. They moved over negation (and adverbs):[24]

23. Modern French has lost the subjunctive past Tense: the present form covers past and present.

24. See Jespersen (1927-V: 427 ff.).

(13) *I (ne) seye not*
I NEG say not

II. They had a range of interpretation closer to French finite verbs: the example in (13) could be interpreted as 'I'm not speaking'.[25]

III. Both the present and the past forms had agreement distinctions. According to what I proposed above, subject agreement is a Tense morpheme:[26]

	Present	Past
1st sg	*speak*	*spoke*
2nd sg	*speakst*	*spokest*
3rd sg	*speaks/eth*	*spoke*
plural	*speak (en/eth)*	*spoke(n)*

IV. In addition, Mood was still in the verb paradigm.

Modern English, after a transitional period,[27] lost three of these characteristics: inflected lexical verbs do not move, they do not allow a progressive interpretation, and they do not show agreement in the past form (while they show a very limited form of agreement in the present form). Subjunctive Mood, we can assume, also disappeared and is nowadays a normatively preserved option outside the core grammar. My suggestion is that these three properties were lost together because they are manifestations of the same diachronic change. Let us see how.

The past form lost all of the agreement suffixes (the plural suffix was lost just as in the present, and the 2nd person singular affix was lost due to the disappearance of the pronoun *thou*). In addition, for regular verbs, the past form happened to become homophonous with the participial form. Suppose, then, that

---

25. See for instance Visser (1973.II: 711, 746) for examples.

26. See Görlach (1978: 88) for the agreement paradigms, which I roughly summarize here.

27. A "transitional period" can of course mean a stage where the English grammar had changed with respect to Middle English but had not attained the present day status. But it can also mean that the grammar had already changed (in the relevant respects) but writers still tried to imitate their older classics, as they usually do, only that they often failed in their attempt, thus showing a "mixed" grammar. Here I cannot address the details of the transitional period in the XVIth and XVIIth centuries, but probably both factors (intermediate stage and writers' failed attempts at using an older grammar) play a role in the grammatical facts we can observe in this period.

the past form was reanalyzed as a past participle. The essence of my proposal is then: if *worked* is a participle in both *John has worked* and *John worked*, then in the latter case it should not move any more than the participle does in the former case.

For this proposal to be made plausible, several issues have to be addressed: (a) Why do some irregular verbs still have two differentiated forms (*ate/eaten*)?; (b) What morpheme do these participial forms convey?; (c) Why can a participial form appear without an auxiliary (*John worked*)?; (d) If the past form is a participle, what is the present form (*works*)?; and e) Why is *do* required in negative, interrogative and emphatic-assertive sentences? We will address these issues in turn.

Concerning the fact that some irregular verbs have two distinct forms (*drove/driven*), the claim that they are both the "same" participle can be maintained at the cost of postulating that one form (*eaten*) is selected by the overt auxiliary, whereas the other is selected by a null auxiliary (or is the unmarked form). However clumsy this solution may seem, there is suggestive evidence that Modern English speakers tend to blur the contrast between the "past participle" and the "past" form. Several facts point in this direction.

First, the existence of a common form is not restricted to regular verbs (*loved*): according to my counts, out of some 135 irregular verbs in Standard English, only a 25% has a differentiated form (e.g. *stole/stolen*), while 75% have a common form (e.g. *thought*). There are only some 10–15 high frequency verbs that have two distinct forms.[28]

Second, there has been, in the Modern English period, a strong tendency for the distinct-form paradigm to be replaced by the same-form paradigm. In some cases this is so because the verb has become a regular verb (*strove/striven* → *strived*). In many cases, though, neutralization of the two forms has taken place without the verb becoming regular (the changes in the examples are standard English or dialectal):

(14)

cleave:	*clove*	*cloven*	→	*cleft*	*cleft*
bid:	*bade*	*bidden*	→	*bid*	*bid*
stink:	*stank*	*stunk*	→	*stunk*	*stunk*
spin:	*span*	*spun*	→	*spun*	*spun*
sting:	*stang*	*stung*	→	*stung*	*stung*

28. Many of these are of the Ablaut pattern *sing/sang/sung*. Still, some of the two-form paradigms are unsteady and tend to regularize (→ *stung* for both).

This tendency shows up in early Modern English.[29] It is also very strong in dialectal American English (with sporadic interference into the standard colloquial speech), where almost all verbs have undergone this neutralization, and speakers seem to have trouble in learning the standard forms:[30]

(15) *go: went gone → went went*
see: saw seen → seen seen

It is also remarkable that these changes tend to obey a general pattern: the new neutralized form is usually the one that most looks like a regular past participle. The changes tend to go from left to right in the following scale of participial morphs (which is a scale of productiveness and regularity):

(16) front vowel > back vowel > *-n* suffix > *-t* suffix > (underlying) *-d* suffix

So in the dialectal American examples, *went* (*–t* suffix) tends to be chosen over *gone* (*-n* suffix) and *seen* (*-n* suffix) is preferred over *saw* (back vowel). The reader can check that the same criterion applies to the changes in (14) (assuming that *bid* is a case of the irregular *-t* suffixation: *bid+t*). All these tendencies clearly suggest that the participial status of the "past" form has some psychological reality that triggers language change.

We have argued that the "past" form is actually a participle on the basis of diachronic change. Now let us address the issue of which kind of inflection is involved in English past participles. My proposal will be that these participles are inflected for Aspect, and convey the feature value [–progressive]. Then the difference in interpretation between *John has/had left* and *John left* (both [–progressive]) should be a function of the presence versus absence of the auxiliary, not of the meaning of *left* itself.

Before proceeding, let us consider the simple present form (*works*). If my proposal is correct that the "past" form was reanalyzed as a [–progressive] participle, then the "present" form had to be reanalyzed too: a Present morpheme cannot exist if it does not stand in opposition to a Past morpheme (just

---

29. See Jespersen (1927-VI: 44): writers such a Shakespeare often used the etymological past form instead of the participle (*chose*, *broke*, *spoke)*.

30. See Visser (1973-II: 1297–1298). This is a comment I have heard from several American speakers. It is nevertheless understandable that the more the level of education raises, the less the phenomenon will be detectable and robust. In addition, using the paradigms in (15) is subject to strong normative contempt.

like singular cannot exist without plural). It should then also have been re-analyzed as a participle.

Here, we seem to face an obvious problem: the contrast between *arrives* and *arrived* seems to be a contrast in Tense. In fact, both seem to be marked for the feature [–progressive], as both allow an habitual and a perfective interpretation, depending on the context, but not the progressive interpretation:[31]

(17) a. *John **arrived** at ten* (habitual/perfective)
b. *John **arrives** at ten* (habitual/*perfective)
c. *I hope that John **arrives** at ten* (habitual/perfective)
d. *When/before John **arrives/arrived*** (*habitual/perfective)
e. *If he **comes/came** at ten ...* (habitual/perfective)
f. *John **comes** to me and **says**...*
(habitual/perfective-historical present)

The restrictions on which readings we can get can probably be derived from independent interpretative constraints. For instance, we can assume that referential present tense (i.e., when present is anchored in speech time) and perfective aspect exclude each other, which would account for the impossibility of a perfective reading in (17b).[32] If *arrives* and *arrived* do not contrast in Aspect (both are [–progressive], which allows for only the perfective and the habitual reading) and do contrast in Tense, then the proposal that they are both aspectual participles seems to be on the wrong track. For in order to maintain the claim that they are aspectual participles, we have to assume that they are selected by a null Tense morpheme: when [–Past], it would select *arrives*, when [+Past] it would select *arrived*. Then one should ask why we could not more simply assume that these Tense morphemes are encoded in the lexical verb itself. From a different perspective: why should the learner ever hypothesize that the contrast *arrives/arrived* is not one of tense, but rather one of selection by two distinct null tense morphemes?.

There is one reason why, though: if our proposal is correct that the simple

---

31. Perfective and habitual are the [–progressive] options for eventive predicates, which I use in my examples. For simplicity, I will ignore stative predicates like *know* or *love*, assuming that their aspectual interpretation in *He loves/loved* is the stative counterpart of habituality.

32. As pointed out by A. Giannakidou, the Greek present aorist, which is perfective, cannot be used in present declarative sentences: it can only be used when combined with the subjunctive or future particles.

past form and the past participle are the same form (putting aside some irregular verbs), then, since this form cannot be inflected for Tense features when it co-occurs with the auxiliary (it is the auxiliary that is inflected for Tense), then it cannot be inflected for Tense features when it occurs without any overt auxiliary either. The only option is then for the learner to assume that a null [+Past] Tense morpheme selects this form when the auxiliary is not present. Then since the simple present form will not stand in opposition to a past form, it has to be analyzed as a participle which is selected by a null [–Past] Tense morpheme. At the bottom of all these (acquisition) entailments is the quite robust fact that 'past' forms mostly look like participles and, in Modern English, steadily appear in the same position as the past participle (after several testing adverbs and right before the object: *Papa completely lost his patience*).[33]

My proposal is then that English has the following set of aspectual participles:

		Selected by:
[+progressive]:	*eating*	*be*
[–progressive]:	*eaten*	*have*
	ate	[+Past]
	eat(s)	[–Past]
	eat	modals, *to*

It can be shown that, except for the progressive form, all the other forms are interpreted as [–progressive], which allows for the perfective or habitual meaning only. This is shown in (17) and in the following examples:[34]

(18) a. (*Until now*) *John has* ***come*** *early* (perfective/habitual)
b. (*From now on*) *He will* ***come*** *early* (perfective/habitual)
c. *When you come, I'll* ***smoke*** *a cigarette* (perfective/habitual)
d. *I may* ***smoke*** *a cigarette when you come* (perfective/habitual)

Since none of the lexical verbal forms in English allows for a progressive interpretation except for the progressive participle itself, we can conclude that all of them are marked for [±progressive]. The [–progressive] forms, however, are

---

33. For the transitional period, the reanalysis may have been accidental, since at a given moment only the morphological similarity could be a trigger.

34. These data are hard to elicit, and some linguists and an anonymous reviewer have cast some doubt on it. I still think that the exclusion of the progressive reading with nonfinite forms is a strong tendency, if not a sharp contrast.

subject to selection by an Auxiliary or a tense null morpheme.

It remains to be explained why the form *eat(s)* agrees with the subject in 3rd singular in finite sentences whenever no auxiliary or modal verb appears (*John eats*). I will make two suggestions: that this agreement is not person agreement, but only number agreement; and that number (and gender) agreement between the subject and the participle is independent of standard subject agreement, and is attested in other languages.

That the 3rd person singular ending *-s* is just number (singular) agreement has been proposed by Kayne (1989b). In a nutshell, he claims that agreement with *you* is always (formally) plural, even in the singular usage of this pronoun; and that 1st person agreement is radically unmarked. My suggestion is that the confinement of number agreement to 3rd person is related to the participial status of the agreeing form. That participle number (and gender) agreement can be restricted to the 3rd person is attested in Italian clitic–participle agreement, where 3rd person object clitics always agree with the participle, while for 1st and 2nd clitics there is variation as to whether there is agreement or not:

(19) a. *Le ha viste / *visto.*
them.has seen.FEM.PL seen.MASC.SG
'S/he has seen them.'

b. *Ci ha viste / visto.*
us-has seen.FEM.PL seen.MASC.SG
'S/he has seen us.'

The existence of number agreement restricted to 3rd person is probably a marked option. Many dialects of British English have either lost this form of agreement or have generalized the *-s* morpheme to all persons.[35] In the latter case, *-s* has simply become a participial suffix.

As for the possibility of agreement between subject and participle, it is attested in languages like Bulgarian:

(20) a. *Toj e rabotil.*
he is worked.MASC.SG
'He has worked.'

b. *Te sa rabotili.*
they are worked.PL
'They have worked.'

---

35. See Trudgill (1990: 94).

How agreement between a subject and a (non passive) participle takes place is no easy question. The answer should probably be based on the VP Internal Subject Hypothesis. The version of this hypothesis developed by Sportiche (1988) and Koopman & Sportiche (1991) claims that the basic position of the subject can be identified by the distribution of floating quantifiers. Whether floating quantifiers are in the subject basic position or not, they can certainly appear right before the participial form in English, which suggests that subject–participle agreement can be established locally given some implementation of cyclic NP-movement of subjects:[36]

(21) *The boys (have) both come to the meetings.*

Let us now address the issue of the appearance of *do* in negative, interrogative and emphatic sentences. The present proposal that a null Tense morpheme selects a participial form is very similar to the classic idea that these constructions involve a null counterpart of *do*:[37] both overt and null *do* would carry the Tense feature. This idea was developed in Pollock (1989). According to his proposal, the different distribution of null and overt *do* would be accounted for on the basis of structural differences in the sentence structure (presence of negation, movement to Comp) that determine the (im)possibility of null or overt *do*. In Chomsky's (1991) account, insertion of *do* is forced whenever the option of lowering the inflectional affixes to V (with subsequent raising of V + Infl at LF) is not available.

In both Pollock's and Chomsky's account, the presence of *do* is determined by purely formal properties of the construction: the presence of negation or interrogation creates configurations in which null *do* or affix lowering are impossible. There is another view of the facts that has been advanced by Laka (1990): the contexts where *do* appears (negation, interrogation and *also* emphatic

---

36. An anonymous reviewer points out that this kind of agreement suggests that the past participle construction involves a Small Clause, with the external argument having gone through its subject position. If we stick to Kayne's (1994) antisymmetry hypothesis, a Small Clause will be structurally indistinguishable from any Spec-X′ structure (Spec-XP in Kayne's theoretical terms). Thus in order to give theoretical content to the term Small Clause we need something else, which, in the Government-Binding tradition, has often been described as its predication status. Whether this intuition (or any other) can be interestingly applied to this case, I leave as an open question. It also remains to be explained why, crosslinguistically, passives and unaccusatives trigger past participle agreement with the subject more often than unergatives and transitives.

37. The essentials of this idea are already in Chomsky (1957).

assertion) share some interpretative feature having to do with the assertive force of the sentence. Laka locates these features in a functional category she calls Σ, which constitutes a functional projection of its own and replaces and generalizes Pollock's Negation Phrase. Adapting this proposal, we can assume that *do* is an inflected form which contains both Tense and Σ features (while its null counterpart is unmarked for Σ features). That *do* is a Σ morpheme explains that the negative morpheme can be affixed to it and even trigger allomorphy on it ([du:]/[dount]). Σ morphology is not restricted to *do*: all auxiliaries and modal verbs should be inflected for Σ in interrogative, negative and emphatic assertive contexts. Like *do*, all of them can be attached to the negative particle, which can trigger allomorphy on them ([wil]/[wount]; [kan]/[ka:nt]).

In sum, *do* is not just a pro-verb required as a last resort, as suggested in Chomsky (1991), but an inflected word that contains both Tense and Σ features. The reason why sentences containing *do* have the same [–progressive] interpretation as corresponding sentences without *do* is that, as we have seen in (18b,c,d), the lexical verb bare form, which *do* selects, is also marked as [–progressive].

In sum, however tentative the above proposal may be, I think there are good reasons for deeply reconsidering the English facts, which have tended to be analyzed in the Government-Binding framework in a rather traditional-grammar fashion with respect to verb inflection.

5. West Germanic Morphology and Word Order

Dutch and German seem to pose a problem for the hypothesis that head movement is driven by inflection. As is well known, verb movement shows a clear contrast between main and embedded clauses: only the former involve verb movement to a fronted verb second position. If we assume the sentence structure in (10) above, and given our conclusions thus far, the prediction should be that finite verbs, since they are (robustly) inflected for Tense and Agreement, should move to T, even in embedded clauses, and thus precede the object in Spec-TelP. In addition, at first glance, morphology seems not to predict at all whether a verb moves to Comp or not. I will argue, though, that in spite of appearances, morphology is crucially involved in determining whether overt verb movement takes place or not.

Zwart (1993b), who develops a minimalist proposal on Dutch syntax compatible with Kayne (1994), argues that Infl categories (AgrS, T and AgrO) are head initial in Dutch (and German), and therefore precede the VP. One of

his main points can be summarized as follows. Root clauses feature movement of the verb to Infl (specifically, to AgrS):

(22) [$_{AgrSP}$ *Jan* [$_{AgrS^0}$ *leest*] *het boek*]
Jan reads the book

Additional movement to Comp only takes places in XP–V–S root clauses, where XP-movement to Spec-CP (*wh*-movement or topicalization) triggers Agr-to-Comp movement (for licensing reasons of the kind of Rizzi's *wh*-criterion). It does not take place in subject-initial main clauses. He provides both empirical and conceptual arguments for this view. The main conceptual argument, which was already advanced by Travis (1984), is that generalized verb movement to Comp in all main clauses is unmotivated, and only *wh*- or topic movement should trigger it.

Adapting his proposal to the previous assumptions in this paper, similar conclusions should be reached: since verbs in Dutch (and German) are inflected for Tense (and agreement) movement to T^0 should take place, which would show up as fronting. Additional movement to Comp is not in principle required.

The problem for both Zwart's analysis and its adaptation into the present proposal is then: why does verb movement to Infl (AgrS or T) not take place in embedded sentences:

(23) **dat* [$_{AgrSP}$ *Jan* [$_{AgrS^0}$ *leest*] *het boek*]
that Jan reads the book

Zwart's account for this is based on the observation that many West Germanic dialects show complementizer agreement. Let us exemplify it here with West Flemish:[38]

(24) a. *da-n-k ik komen*
that.1SG I come
b. *da-se zie komt*
that.3P.SG.FEM she comes

Zwart's proposal can be summarized as follows. On the one hand, complementizer agreement makes verb movement unnecessary, because it provides an

38. See Zwart (1993b: 160) for further examples of various dialects. He makes a distinction between complementizer agreement suffixes and complementizer pronominal clitics. I will crucially assume that clitics are also a form of Comp-agreement.

alternative way of checking AgrS features (moving AgrS to Comp).[39] Since it is unnecessary, economy principles make it impossible. This account extends to standard Dutch and German, where complementizer agreement is not apparent, by assuming that these languages have an abstract form of complementizer agreement.

I cannot go into a discussion of this analysis for reasons of space. I will assume, though, that it is on the right track, and will try to make sense of it within the present proposal. I have suggested above that subject agreement is a Tense morpheme (now this idea becomes crucial). This means that, in dialects with complementizer agreement, both the complementizer and the finite verb are inflected for Tense morphology. Given the above proposal on insertion of inflected words, this implies that the inflected complementizer should be inserted both in Comp and in T, while the finite verb should also be inserted in T (and in whatever other position it is inflected for). Then the resulting structure for (25a) (West Flemish) would be (25b), where copies of both *da-se* and *komt* end up dumped together in the same insertion site (under T^0).

(25) a. *da-se zie komt*
that.she SHE comes

b. $Comp^0$ T^0 [... V^0 ...]
da-se zie *da-se* *komt*
komt

Recall that the present proposal excludes head adjunction: if two copies are required to be inserted in the same head position because of their shared categorial content (both are Ts in this case), then dumping is the only option. Assume this is not a problem for interpretation: dumped words are licensed if their features can be unified. But it is reasonable to assume that it is a problem for pronunciation: the two copies are not even linearized. Assume then that when two copies of different words are inserted in the same position, neither can be pronounced. The copy of the verb that is pronounced is then a lower copy.[40] Thus we change (4c) above into:

(4c′) Pronounce the highest pronounceable copy (where two copies in the same head position are not pronounceable).

---

39. I am not very clear about what Zwart precisely means by moving AgrS to Comp, if affixes do not move in the minimalist program he amply assumes. This issue will not arise in my alternative proposal below.

40. See Section 5 for an account of the sentence-final position of the verb.

For standard Dutch or German (and any dialects with no apparent complementizer agreement), it is not enough to assume that there is an abstract form of complementizer agreement: it is essential to the present proposal that only "overt" morphology is relevant for movement. My suggestion is then that all West Germanic dialects do show overt complementizer agreement: it consists in SUBJECT CLITICIZATION.

It is a well-known fact that subject clitics in embedded clauses have to be strictly adjacent to the complementizer. So the Dutch subject clitic pronoun *ze* 'she', unlike full NPs and stressed pronouns (*zij* 'she') has to appear adjacent to the complementizer in embedded clauses.

(26) a. *Ik denk dat ze /zij/ Marie zal komen.*
I think that she SHE Marie will come

b. *Ik denk dat uiteindelijk zij / Marie zal komen.*
I think that eventually SHE Marie will come

c. **Ik denk dat uiteindelijk ze zal komen.*
I think that eventually she will come

It is not only the case that subject clitics have to be adjacent to the complementizer. They also show a strong degree of phonological attachment to it (which is not usually shown by the standard spelling):[41]

(27)

	Spelling	Colloquial pronunciation
a.	*of ik* if I	[ɔvək]
b.	*dat je* that you	[dɑtjə] / [dɑjə]
c.	*dat ze* that she	[dɑtsə] / [dɑsə]

However convincing morphophonological evidence may be, these facts have not been traditionally analyzed as instances of complementizer agreement. The reason is that, unlike in the Flemish examples, subject clitics are in complementary distribution with overt subjects in standard Dutch (and German):

---

41. In the remainder of this section, I will leave German aside and use only Dutch examples. For extensive evidence for the occurrence this kind of cliticization in Germanic verb second languages, see Abraham & Wiegel (1993), Vikner (1995a: Ch. 3), and Christensen (1984) (on a Mainland Scandinavian dialect). I do not have any data about Icelandic.

(28) **Ik denk dat-ze zij/ Marie komt.*
I think that.she SHE Marie comes

This, however, is not conclusive evidence that complementizer subject clitics are not an instance of agreement. My suggestion is that the difference between standard Dutch and Flemish complementizer agreement is of the same nature as the difference between Irish and Welsh subject agreement: only the latter can co-occur with an overt (pronominal) subject:

(29) Irish: *Taim (*me) sásta.*
am I happy
'I am happy.'
Welsh: *Yr wyf i yn darllen.*
PRT am I in reading
'I am reading.'

In Irish, when the subject is overt the default form of agreement (3rd person singular) has to be used:

(30) *Ta me sásta.*
is I happy
'I am happy.'

We can say that it is then inflected for the default form.[42]

Now consider what the default form should be for Dutch (and similarly for the other Germanic verb second languages). The clitic-inflected forms for 3rd person singular are either masculine, feminine or neuter:

(31) Dutch: *dat-ie*
'that.he'
dat-ze
'that.she'
dat-'t
'that.it'

Let me propose that neither of these forms is sufficiently unmarked, namely that the default form must not be more than just 3rd person singular. We can then

---

42. I am not providing any account for these agreement facts: my sole point is that postulating an agreement form which does not co-occur with an overt subject is not an unheard of theoretical whim: see McCloskey and Hale (1984) for the view that Irish indeed has subject agreement and not syntactic subject pronoun incorporation.

assume that the unmarked form is simply *dat*.[43] The idea, then, is that complementizers are always inflected for agreement (either with a clitic or in the default form), and this predicts that they always will be inserted in T, which prevents the copy of the finite verb in T from being pronounced.

In sum, accounting for the complementary distribution of complementizers and overt movement of the verb to T can be done with no stipulation specific to this case: the morphophonological attachment between the complementizer and the subject clitic requires, under the Strong Lexicalist Hypothesis, that complementizer-clitic must be generated as a word; and Irish, among other languages,[44] provides an independent case where agreement cannot co-occur with an overt subject.

Now let us address another issue: independently of the question of verb movement to T, main clauses allow verb movement to Comp in verb second languages. If movement is always triggered by morphology, there should be some morphological Comp feature on the verb that triggers this movement. Since verb-to-Comp always licences either a Topic or a Focus phrase in clause initial position, we can assume that the content of this morpheme has something to do with topic/focus licensing. Is there, then, overt morphological evidence for this morpheme? I think the answer is yes if we again look at subject clitics. As in the case of the complementizer, subject clitics show strong attachment to the verb when the verb is in Comp. In colloquial Dutch, for instance, subject clitics can trigger reduction (thus allomorphy) on some verbal forms:[45]

---

43. This does not contradict my earlier claim that morphology must be overt. In the relevant sense here, overt morphology means morphology that can be phonologically detected, not that it must have an overt affix: it may be detected by opposition. Thus, for instance, English singular nouns are overtly marked as such, not by any affix, but precisely by the absence of an affix and by their standing in opposition to plural forms. Similarly, *dat* would be overtly marked as default by the absence of any pronominal clitic. Quite commonly (but not always) functional paradigms consist of one or several overt morphemes (affixes or particles) and one affixless default form.

44. Standard Arabic (with post verbal subjects); Trentino and Fiorentino (concerning subject agreement clitics: see Brandi & Cordin 1989).

45. See Booij (1985) for examples and discussion. He specifically proposes that these forms should be generated in the lexicon and dealt with as agreement forms that do not allow doubling, as in the Irish case.

    An anonymous reviewer raises a warning flag: reduction is such a common phenomenon between pronouns and auxiliaries that it would be unreasonable to take all cases as lexical (s/he mentions cases like stop palatalization and affrication in sequences like

(32)

heb ik	'have **I**'	→	[hɛk]
wil ik	'want **I**'	→	[vɪ(l)k]
moet ik	'must **I**'	→	[muːk]

As in the preceding case of complementizer agreement, the Strong Lexicalist Hypothesis forces us to generate these irregular (unpredictable) verb-clitic forms in the lexicon, so that the morphology of the clitic-inflected verb is different from the clitic-less form and forces movement to Comp. Like in the preceding case, we would analyze the clitic-less forms in Comp as cases where an overt subject forces the use of the default form for Comp-agreement.[46]

Now, can it be argued that the presence of enclisis on the verb constitutes a morpheme that belongs in Comp? In fact, neither the finite verb nor the clitics seem to be morphemes which are related to Comp. My proposal is that the Comp morpheme is neither in the finite verb nor in the clitic, but in verb-pronoun enclisis itself (i.e., enclisis constitutes its morph).[47] That enclisis can

---

did you). Here I cannot address these English cases (or cases like *you're* → [joːr)] in British English), but I think that similar approaches are called for. Notice, though, that English, unlike West Germanic, does not have any special reduction or adjacency phenomena between complementizer particles and subject pronouns: the phenomenon is restricted to V-in-Comp + pronoun and pronoun + auxiliary sequences. Notice also, for the latter, that only heads can be involved (**Mary and you're here*).

46. The idea that the verb moves to Comp in (residual) verb second in order to licence some Comp-feature (*wh*-, topic, etc.) has been in the literature for years (see Rizzi 1991a, and Vikner 1995a: Chapt. 3 for a good review). My innovation is that this feature is morphologically marked on the verb by pronoun enclisis. Vikner's conclusion about the issue is that the trigger for V-to-Comp is the *wh*-criterion, and that pronoun cliticization is a side effect due to case reasons (in the spirit of Rizzi and Roberts 1989): V-to-Comp prevents Spec–head Case assignment. In Vikner's (and Rizzi & Roberts') proposal, however, whether or not a language requires pronoun incorporation for case reasons in V-to-Comp structures does not follow from anything; and the *wh*–feature is not morphologically visible on the verb (so just stipulated). My proposal overcomes both problems by stipulating what must be independently stipulated in the lexicon: morphophonologically irregular attachment of two morphemes (verb + pronoun), which should not be dealt with by postlexical (morpho)phonology if we stick to the strict lexicalist view that all idiosyncrasies are stored in the lexicon.

47. Or that, as an anonymous reviewer suggests, enclisis is indicative of the presence of a null morpheme. Recall, though, that our requirement that morphology must be overt in these cases is not affected by this possibility: morphology is overt if there is some pronounceable clue to detect it.

convey a morpheme in itself should not be surprising: in the domain of inflection, the presence of a morpheme can be indicated through a varied set of possibilities: affixes, ablaut, reduplication, suprasegmental morphemes, etc. That enclisis can convey a morpheme independently of the content of the host and guest is, I think, a well-attested fact. Northern Italian dialects provide examples of verb subject–clitic order that constitutes an interrogative morpheme (as opposed to clitic-subject forms). Consider the following examples from Trentino:[48]

(33) a. *el ven*
he comes
'He comes.'

b. *vegne-lo*
comes.he
'Does he come.'

In (33) we can see that the clitic shows a different form in the V–Cl order. In addition, the palatal [ɲ] (spelled as *gn*) appears in this case just like when the stem is followed by an inflectional suffix (*vegno* 'I come'): there is non-word-boundary evidence. The V–Cl forms should then be generated in the lexicon as forms inflected for an interrogative morpheme.

Summarizing, we have argued that in West Germanic:

— there is a main/embedded clause asymmetry because complementizer particles are inflected for subject agreement: this forces insertion of two words (complementizer particle and verbal form) in the same position (T^0), which prevents the verbal form from being pronounced in this position.

— subject pronoun cliticization on the finite verb is also a form of inflection which licences (and forces) verb-movement to Comp, and conveys a focus/topic morpheme.

It is to be noticed that the present proposal is contrary to usual practice: cliticization is stipulated as a lexical fact, and not as something to be derived from syntactic processes (movement). But to the extent that morphophonological criteria single out these cases as candidates to be generated in the lexicon because of allomorphy and phonological attachment, the Strong Lexicalist Hypothesis should not allow for the syntactic cliticization option. And if this morphologically (independently) motivated decision makes nice predictions for syntax, we have a real explanation.

---

48. The examples are from Brandi & Cordin (1989).

6. West Germanic Head-finalness

As a last point, let me make a tentative proposal on West Germanic head-finalness. Although this section deals with object movement, the crucial assumption will concern V-movement, which is the topic of this paper. I will assume that object movement is uniformly overt in all languages: it is then V–movement (or lack thereof, depending on its morphology) which will crucially make the verb precede or follow the object in a given language.[49]

Zwart (1993b) proposes that the fact that the participle, and, in embedded clauses, the finite verb, appears after its complements is not to be derived from the head-final status of the VP (or IP) in these languages, but rather from overt movement of complements to functional specifiers:

(34) a. *Ik heb* [$_{AgrOP}$ *het boek*$_i$ [$_{VP}$ *gelezen* t_i]]
I have the book read.PASTPART
b. *dat ik* [$_{AgrOP}$ *het boek*$_i$ [$_{VP}$ *lees* t_i]]
that I the book read.PRES

Let us concentrate on object movement. Zwart's proposal is that object movement to Spec-AgrOP is overt in Dutch and German, but not in English, which accounts for the contrast in word order (O–V/V–O). Here I want to explore an alternative possibility. I suggested above that in both English and West Germanic objects sit in the specifier of the Telicity Phrase (TelP), whose head can be occupied by terminative particles such as English *up*:

(35) a. *I ate* [$_{TelP}$ *the cake up* [$_{VP}$...]]
b. *Ik at* [$_{TelP}$ *de koek op* [$_{VP}$...]] (Dutch)

TelP would be equivalent to Chomsky's AgrOP. The reason why objects move there would be that they convey crucial information for the telic/atelic status of the sentence:[50]

---

49. See Johnson (1991) for the argument that object shift is overt in English.

50. See Abraham (1993a) for evidence, mainly based on Middle High German, that Telicity (Aktionsart) morphology correlates ("agrees") with the form of the object.
In fact, it is internal arguments (not just objects) which determine telicity, as can be seen with unaccusative verbs:
(i) *Stones were falling.* (atelic)
(ii) *The stones were falling.* (telic)

(36) a. *to read (books)* (atelic)
b. *to read a/the book* (telic)

Thus movement of objects to Spec-TelP would be motivated by the need to create a local configuration for telicity interpretation. Accusative marking could then be seen as the overt manifestation of Spec–head agreement in TelP (just like *wh*-marking can be seen as a manifestation of agreement with the *wh*–feature in C^0). But XP-movement is beyond the scope of this paper.

Assuming that object movement to Spec-TelP is overt in both English and West Germanic, then the contrast in word order between the verb and the object should be accounted for in terms of movement of the verb itself.

In English, the strict adjacency requirement between lexical verb and object would be due to the fact that, as we have argued above, English lexical verbs are always inflected for Aspect, and are therefore inserted in Asp^0, which is the head immediately preceding TelP.[51] If this is the right account for English, it should be the case that in West Germanic the verb (finite or not) is not inflected for Aspect, and therefore does not move to (is not inserted in) Asp^0. Then it should be only inserted below, inside the VP, which would account for the object–verb word order (whenever it does not move, or is not pronounceable, in a higher functional projection). Let us see if the claim that West Germanic verbs are not inflected for Aspect is tenable.

This would predict that unaccusative subjects move through Spec-TelP, which fits in with the general phenomenon that unaccusative subjects trigger past-participle agreement. This is not compatible with Chomsky's view that AgrS vs. AgrO checking are complementary alternatives.

51. I cannot address the issue why full NP objects can also follow these telicity particles in English. My tentative suggestion is that the option V–particle–object has been, for a long time, promoted by normative pressure and has triggered the (perhaps non-core) grammatical option of having the particle as a verb clitic. Since there is no evidence for morphological attachment in the V–clitic sequence, the V–object–particle is felt as more natural (*pace* heavy object extraposition), at least according to some native speakers. This would lead to the conclusion that object pronouns are also verb clitics (here the evidence is fairly robust: *I hate'm*).

I will not address several other issues, such as the position occupied by preverbal adverbs if the verb is in Asp, or why the strict adjacency between participle and object is not attested in similar languages such as French (see Pollock 1989).

For finite verbs, this seems to be correct. Finite verbs convey no aspectual information: both a perfective and a progressive reading seem possible with finite verbs.

(37) a. *dat Jan het boek leest* (Dutch)
daß Hans das Buch liest (German)
that John the book reads
'that Jan reads/is reading the book'

b. *dat Jan het boek las* (Dutch)
daß Hans das Buch las (German)
that John the book read.PAST
'that Jan read/was reading the book'

Let us now consider infinitives, which also appear after the object in both main and embedded clauses. They also appear to be unmarked for aspect:

(38) a. *Wanneer je komt, zal ik werken.* (Dutch)
Wenn du kommst, werde ich arbeiten. (German)
when you come will I work
'When you come, I will work/be working.'

b. *Jan scheen te werken.* (Dutch)
Hans schien zu arbeiten. (German)
John seemed to work
'Jan seemed to work/be working.'

Let us consider past participles. At least for Southern German, it seems unproblematic to assume that they are not inflected for Aspect: a sentence containing a past participle is completely neutral between a progressive and a perfective reading:[52]

(39) *Ich habe das Buch gelesen.*
I have the book read
'I (have) read the book.' or 'I was reading the book.'

It is less obvious how to extend this account to Northern German and Dutch, where sentences containing a past participle never allow the progressive reading:

(40) *Ik heb het boek gelezen.* (Dutch)
I have the book read
#'I was reading the book.'

52. Thanks to Karen Lattewitz for the data.

My suggestion is that the participle in the above example is not marked for Aspect ([+perfective]), for if it was, it would not have any other verbal form to contrast with, since both finite and infinitival forms are not marked for Aspect. In the absence of any contrastive form, the participial form cannot be assigned an aspectual feature.[53] Then the fact that (40) is only interpreted as perfective must follow from independent factors: probably the construction the participle is embedded in.

An account of word order in West Germanic and English cannot be solely based on object placement, because many other constituents precede the verb in West Germanic that follow the verb in English. Abstracting away from scrambling (i.e., word order alternations concerning phrases preceding the verb in final position), an issue I cannot address here, it seems that at least predicative PPs and datives obligatorily precede the sentence final verb (examples from Dutch):

(41) Dutch:

a. *Ik heb Marie het boek gegeven.*
I have Mary the book given

b. *Ik heb het boek op de tafel gelegd.*
I have the book on the table layed

Both Zwart (1993a) and Koster (1995) propose that, like objects, predicative phrases move to some designated functional projection in Dutch, skipping over the verb. My suggestion is that both datives and predicative phrases are relevant, like objects, for telicity, in that they define terminative points of the event.[54]

---

53. Past participles in West Germanic take the *ge-* prefix (but see just below). From a historical point of view, *ge-* is a telicity (Aktionsart) morpheme in earlier stages of these languages, but it has lost this status (see Abraham 1993a: 4). Nowadays it is in complementary distribution with several prefixes (of applicative, inchoative, etc. nature). It can be then characterized as the unmarked form of this paradigm of affixes. Unlike them, though, it only appears in participles.

In any event, if past participles in modern West Germanic languages were still marked for Telicity (but *not* for Aspect), my proposal would all the same correctly predict that they should follow the object, which occupies Spec-TelP.

54. To see this, consider:

(i) *to put (the) books on (the) tables*

(ii) *to assign (the) jobs to (the) professors*

The reader can check that only if the definite article (or any other determiner) is present *both* in the object and inside the PP is the construction telic.

If so, both objects and datives or predicative PPs should move to Spec-TelP. Since a specifier can only host one phrase, this means that they should form a Constituent: a SMALL CLAUSE. The essence of (object) Small Clauses would then be that they contain all the information that is relevant for Telicity: the internal argument and the termination point of the event. Then the fact that in English both Objects and Datives or predicative PPs follow the verb, whereas they precede the verb in Continental West Germanic, can be seen as a manifestation of the same phenomenon. I will leave the issue here.

7. Conclusions and Speculations

The present proposal explores the highly restrictive view that there is a strict correlation between (overt) morphology and (overt) head movement, here defined in a non-derivational way. This alternative is only broadly related to the Minimalist Program, but I think it gives much more content to Chomsky's claim that morphology is the key to language variation. By incorporating Kayne's antisymmetry hypothesis the proposal becomes easily falsifiable and, if empirically accurate, highly explanatory. As the reader may have perceived throughout the paper, though, identifying overt morphology becomes extremely complex and controversial. I firmly believe, though, that the usual practice of basically relying on the spelling (perhaps because of the feeling that it is accurate enough for syntax) should be abandoned when we deal with such slippery entities as functional morphemes.

In addition, the success of the present proposal requires a fine-grained analysis of the sentence functional architecture and the nature of its functional categories. It is my belief that the split Infl hypothesis should develop into its ultimate consequences through detailed analysis of all the semantic functions that are involved in sentence interpretation. On the one hand because semantics can provide independent motivation and criteria for otherwise highly speculative and often untestable speculations about functional categories. On the other hand because it seems that the ontology of propositional functors (the most natural candidates for functional categories) is very large and needs to be incorporated into syntax if we do not want to give birth to a monster transformational component for semantics. In my view, research on morphological paradigms (and sentence particles) should go hand in hand with research on the sentence functional structure, and the present hypothesis can serve as a guiding test.

This paper probably bites off more than it can chew and digest. The

purpose of going into several quite unrelated issues has been both to give some plausibility to the proposal and to suggest how much can be rethought in the light of a closer morphological examination.

That morphology is what makes languages different is a very attractive programmatic idea for the universalist syntactic enterprise. In this view, it is my belief that syntax should give up taking care of rather irregular attachment phenomena, in order to gain, as a bargain, a richer inflectional morphology, since what is a hopelessly irregular process of morpheme attachment or reduction can be reanalyzed as an idiosyncratically stored clue for functional features in the lexicon.

Acknowledgements

This research is being supported by the joint research grant PB93-0893-C04-04, awarded by the Spanish Ministry of Eduction, and the Quality Research Group GRQ 93-2.035 financed by the Catalan autonomous government Generalitat de Catalunya, to both of which I belong as a reincorporation researcher (RED94/-206, Generalitat de Catalunya). I would like to thank people at the University of Groningen (Jan-Wouter Zwart, Jan Koster, Laurie Stowe, Charlotte Koster, Paulien Rijkhoek, Anko Wiegel, Jack Hoeksema, Elly van Gelderen, Dicky Gilbers, Karen Lattewitz, Lily Grozeva, and Anastasia Giannakidou) for data and discussion, and people at the Autonomous University of Barcelona for discussion and support.

Gs vary with particular
Rates determined
by experimentation

On the (Non-)Universality of Functional Categories

Höskuldur Thráinsson
Harvard University/University of Iceland

The structure of CP [see (1) below] is largely forced by ... properties of UG, assuming the minimalist program with Agr abstracted as a common property of inflectional systems, a reasonable assumption, given that agreement appears without Case (as in NP-AP agreement) and Case appears without agreement (as in transitive expletives, with the expletive presumably in [Spec,Agr$_s$] position and the subject in [Spec,Tense], receiving Case. [...] Suppose that VP contains only one NP. Then one of the two Agr elements will be "active" (the other being inert or perhaps missing)

— *Chomsky (1993: 9)*

Abstract

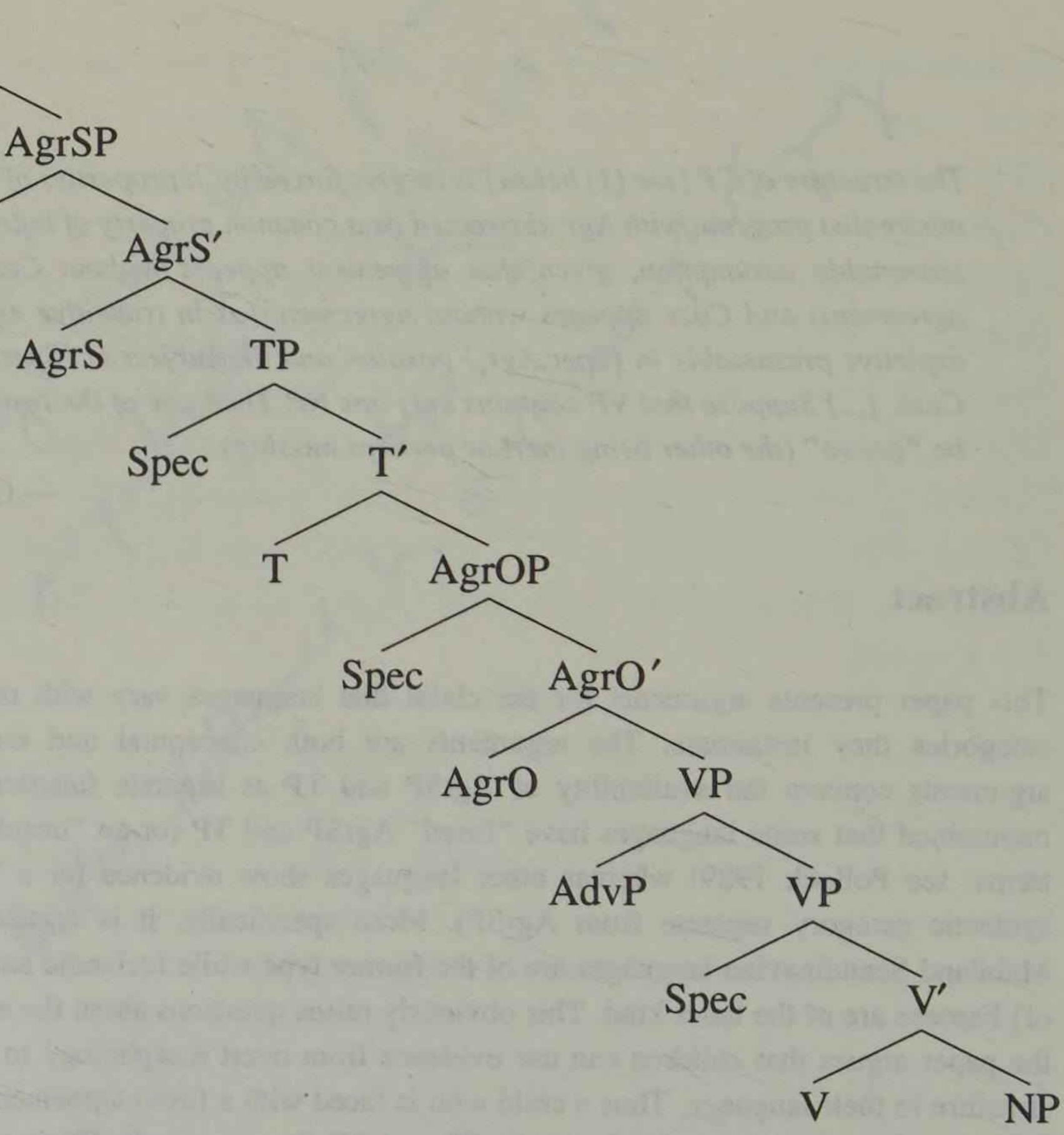
This paper presents arguments for the claim that languages vary with respect to the functional categories they instantiate. The arguments are both conceptual and empirical. The empirical arguments concern the availability of AgrSP and TP as separate functional categories and it is maintained that some languages have "fused" AgrSP and TP (or an "unsplit IP" in pre-Pollockian terms; see Pollock 1989) whereas other languages show evidence for a "split IP" (i.e. TP as a syntactic category separate from AgrSP). More specifically, it is argued that English and the Mainland Scandinavian languages are of the former type while Icelandic and (at least some dialects of) Faroese are of the latter kind. This obviously raises questions about the acquisition of syntax and the paper argues that children can use evidence from overt morphology to determine the syntactic structure in their language. Thus a child who is faced with a fused agreement and tense morphology will take that as evidence for fused AgrSP and TP (i.e. an unsplit IP) in the syntax while a child presented with morphological evidence for separate agreement and tense will assume that the language in question has a split IP. In this sense the paper argues for a more direct relationship between overt morphology and syntactic structure than is usually assumed in recent minimalist work.

1. Introduction

1.1. *On the (Non-)Universality of Basic Clause Structure*

The purpose of this paper is to discuss a fundamental question raised in the above quote from Chomsky's (1993) paper "A Minimalist Program for Linguistic Theory" (henceforth MPLT): Which aspects of clause structure in general and which functional categories in particular are "forced by ... properties of UG" (Universal Grammar)? The CP-structure referred to in the above quote is given in (1) (with minor notational modifications and Spec-TP added for the sake of consistency):

(1)



The structure in (1) differs from early X′-structure analyses in that the IP has been "split" into separate projections for agreement (Agrs) and tense (T),[1] an

---

1. The AgrO projection is also an innovation in this kind of structure, as discussed below (see in particular Section 4.3).

analysis standardly attributed to Pollock (1989) and adopted and revised by Chomsky (1991; originally published 1989, 1993) and Chomsky & Lasnik (1993), for instance. Pollock suggested that differences in the nature of functional categories (in particular what he referred to as the "opacity" vs. "transparency" of Agr) could explain certain cross-linguistic syntactic differences, including well known differences between French and English. This was further developed by Chomsky (1991, 1993) and others, including Ouhalla (1988, 1991), and Ouhalla's work was pioneering in extending this line of research to other language families. But despite the rapidly growing research on the nature of functional categories, relatively little attention seems to have been devoted to the following basic question:

(2) Do all languages have the same inventory of functional categories (and if so, what is it)?

This question is discussed by Iatridou (1990: 553) who suggested that languages might "vary with respect to the functional categories they instantiate and that therefore evidence for an AgrP [...] will have to be found in each language separately." This issue is obviously very important for the discussion of the acquisition of syntax in general and the debate about the acquisition of functional categories (FCs) in particular (see, for instance, Radford 1990; Rizzi 1993; Wexler 1994; Weinberg 1994; Poeppel and Wexler 1993; and references cited there). Ultimately, it will depend on our ideas about the way UG constrains possible grammars. Consider, for instance, the following hypotheses (where the weak version is weak in the sense that it is necessarily true if the strong version is true but the converse does not hold):

(3) *The Structural Uniformity Hypothesis (SUH)*

a. *Strong Version*
Clausal architecture is completely determined by UG in the sense that all clauses in all languages have the same set of functional categories and their sequence (c-command relations) is uniform.

b. *Weak Version*
Clausal architecture is determined by UG in the sense that UG defines a set of functional categories, $\{F_1, F_2, \ldots, F_n\}$, that languages "select" from. For any functional categories F_i and F_j, the sequence will be uniform whenever they occur, i.e. if L_1 and L_2 each instantiate both F_i and F_j and F_i c-commands F_j in L_1, then F_i c-commands F_j in L_2.

According to the strong version of the SUH, the answer to the question in (2) would be *yes*. But it is clear that many linguists reject the SUH, either implicitly or explicitly. One can distinguish at least three different reasons for rejecting (at least the strong version of) the SUH.

First, we may not want to claim that all languages "select" the whole set of FCs made available by UG, as claimed by the strong version of the SUH. This was Iatridou's (1990: 552) original concern: "... are data from one language in favor of a functional projection sufficient for us to postulate that the same functional category exists in all languages?" Van Gelderen (1993) argues that the answer is *no* and so do Reuland & Kosmeijer (see e.g. Reuland & Kosmeijer 1988; Reuland 1990; and Kosmeijer 1993; especially Chapter 5). I will return to this question below and will agree with the basic conclusion of these linguists, although my arguments will be rather different from theirs.

Second, the question is whether the sequence of FCs is the same in all the languages they occur in. It seems that a part of the sequence is given by the definition of the FCs (or stipulated), e.g. that CP must dominate the TP of "its" clause and not the other way around since CP is taken to define the boundaries of a clause (= dominate all elements within a clause). But it is not clear that the order of the FCs that are directly related to morphological distinctions, such as AgrS and T, could not be reversed. For example, Pollock (1989) actually suggested originally that TP dominated AgrSP (his AgrP — he did not assume AgrOP) in French and English and not the other way around, while Chomsky (1991) and Belletti (1990) argued for the other order. As a matter of fact, Ouhalla has argued (most recently 1994) that "... in VSO languages T is higher than AgrS, whereas in SVO languages the reverse relation is found" (1994: 46). Rizzi (1993: 6), by contrast, assumes that "... the usual sequence of functional heads and projections is determined by UG", which suggests that he accepts at least the weak version of the SUH (unless "the usual sequence..." is taken to mean "the unmarked" or "default" order). One possibility is that the sequence of those FCs that are directly related to morphological distinctions mirrors the order of the morphological markers in the language in question (in the spirit of Baker's Mirror Principle; 1988: 13). That means, however, that the sequence of FCs (their c-command relations) must vary from language to language if the order of the relevant morphemes (say, agreement morphemes and tense morphemes) varies, as observed by Ouhalla (1994: 45).

Third, discussions of the categorical status of different types of complement clauses, such as analyses of various kinds of infinitival complements (see e.g. Thráinsson 1993 and references cited there), obviously presuppose that clauses

can vary within a given language with respect to their architecture, in that some clause types contain fewer FCs than others. Note also that Chomsky suggests that in clauses that contain an intransitive verb, one of the Agr-projections will be "inert or perhaps missing" (see above). The question is then when a given FC can be "inert" or "missing" and whether there is more than a notational difference between the two concepts.[2]

There are thus three different reasons why one might want to adopt something like (4) instead of either version of the SUH:

(4) *The Limited Diversity Hypothesis (LDH)*
Clausal architecture is determined by UG in the sense that UG defines the set of functional categories, $\{F_1, F_2, \ldots, F_n\}$, that languages "select" from. Cross-linguistic and intra-linguistic variations are limited to the following:

a. It is not the case that all FCs are instantiated in all languages.

b. The FCs selected by a given language may not be present in all clause types of that language.[3]

c. The sequence (c-command relations) of those functional categories (dominance relations between the functional projections) that are directly related to morphological distinctions may vary from language to language, consistent with the Mirror Principle.

Note that it follows from (4c), i.e. the Mirror Principle, that the sequence of any two functional categories that are directly related to morphological distinctions is uniform for all clause types within each language, as long as the order of morphological markers within words does not vary in the language in question.[4]

---

2. See also the discussion in Chomsky (1995a: 399) of unergative "intransitive" verbs, where he assumes that unergative verbs are in fact transitive in the sense that they project a transitive structure, as suggested by Hale & Keyser (1993).

3. Many of the ideas presented in this paper are similar in spirit to ideas proposed in Grimshaw (1993), such as her idea that a "clause is only as big as it needs to be. It is an IP unless it has to be a CP (a VP unless it has to be an IP, cf. Safir [1993])" (Grimshaw 1993: 5).

4. I do not know of any languages where the order of morphological markers varies in this sense, such that, say, a subject agreement marker in a verb form sometimes precedes a tense marker and sometimes follows it. It would be interesting to know whether such languages exist, and if so, whether there us any evidence that this is reflected in syntactic structure. For a claim that the hierarchical relationship between the syntactic categories Agr and T can vary within a given language (e.g. between

Although it is by no means obvious that the Mirror Principle should hold if one assumes a strong lexicalist approach to morphology as is done in MPLT (i.e., the approach which assumes that words "emerge" fully inflected from the lexicon and their morphological features are checked off in the course of the syntactic derivation), I will assume as a working hypothesis in this paper that the Mirror Principle does indeed hold. As pointed out by Chomsky (1993: 28), it is possible to explain why it does (if it does!) while still assuming a strong lexicalist approach to morphology. The explanation could go along the following lines: Let us make the natural assumption that morphological features are not associated with each lexical element as a set (i.e. in an unordered fashion) but rather with the relevant overt morphemes (when such morphemes are available). Thus the morphological feature of tense is associated with the tense morpheme of a verb, and morphological features of person and number (e.g. subject agreement features) of a finite verb are associated with the agreement morpheme of the verb form in question. These morphological features are then checked off in a cyclic fashion as the verb form is adjoined to the relevant functional head, beginning with the features associated with the morphemes closest to the stem of the verb since features associated with "outer" morphemes are not visible until features associated with morphemes closer to the stem have been checked off. In a language where a tense morpheme is closer to the stem of a finite verb than a subject agreement morpheme for instance, this would mean that the functional head T (i.e. the head against which the morphological feature of tense is checked) would have to be lower in the syntactic structure than AgrS (the head against which the subject agreement feature is checked). I will refer to this version of the Mirror Principle as the LEXICALIST MIRROR PRINCIPLE. We will return to it in Section 3.3 below with illustrative examples from Icelandic, Mainland Scandinavian and English, although a detailed investigation of how this would work in different languages is outside the scope of this paper.[5]

Note that the Limited Diversity Hypothesis (LDH) about FCs is rather similar to commonly accepted ideas about the universal set of phonological features made available by UG to the languages of the world (see, e.g., Chomsky 1965: 28). First, some phonological features (or contrasts) are not

---

sentences in Pres and Past tenses) see Campbell (1991).

5. The Lexicalist Mirror Principle as outlined here is very similar to the idea in Chomsky (1993: 28). Bobaljik (1995a, b) discusses a different way of relating syntactic movement and morphology, assuming a non-lexicalist version of morphology in the spirit of Halle & Marantz 1993 (see in particular Chapter 1 of Bobaljik 1995b).

considered relevant in certain languages and thus do not play any role in a description of their phonetic inventories or in phonological rules in these languages. Thus there is just one liquid in Japanese whereas English distinguishes between /l/ and /r/. Second, the dominance relations in the geometry of phonological features are largely predetermined, but there may be room for some minor cross-linguistic variation. Third, some phonological features may be irrelevant or non-existent for the description of certain speech sounds. The child acquiring the phonology of a given language will learn which distinctions are relevant in the language in question and for each type of sound. The LDH maintains that the same is true of the acquisition of syntax, *mutatis mutandis.*

There is also a certain parallelism between FCs and lexical categories under this approach. Thus it is well known that not all languages have adjectives as a lexical category distinct from verbs or nouns, nor do all languages have prepositions (or postpositions) (for a brief discussion of crosslinguistic differences of this kind see, e.g., O'Grady, Dobrovolsky & Aronoff 1993, Section 5.4). A child acquiring Korean, for instance, will not find any evidence for adjectives as an independent lexical category. Similarly, I will argue below that a child acquiring English will not find any reason to distinguish AgrS and T as syntactic categories but a child acquiring Icelandic will.

While the LDH does not imply that clausal architecture can vary "without limits and in unpredictable ways", it does imply that at least something about clausal architecture has to be learned. This paper will not discuss any language acquisition data but it is necessary to consider the acquisition issue in general terms before proceeding.

1.2. *Acquiring Functional Categories*

According to the LDH in (4) above, the task of acquiring the FCs of any given language is at least threefold.

First, a child has to determine what FCs play a role in the language being acquired (L_1) and what their nature is, to the extent that this is not completely determined by UG. If one were to assume that all languages have Case and that Case is always checked against the functional head F_x, then F_xC (the functional category whose head is F_x in syntactic structure) would have to be present in all languages and there would be no learnability problem involved in determining whether L has F_xC or not. But this is apparently not how Case works. It is generally assumed in the minimalist framework that the Case of subjects is not checked against the same functional head as that of objects, for instance.

Furthermore, I will argue below that the functional categories checking Case vary cross-linguistically. Hence there is something for the child to learn about FCs checking subject Case. It may very well be, on the other hand, that some aspects of functional structure are universal and need not be learned. Thus it is possible that all finite embedded clauses are CPs by definition in all languages, i.e. are dominated by a functional projection which is headed by a functional head (C) which typically hosts a finite complementizer. If this is so, and if all languages have finite embedded clauses, then a child acquiring L will not have to learn this, although it is a matter of debate whether young children "have" CPs in their grammar at the early stages (see again Radford 1990; Rizzi 1993; Wexler 1994; Weinberg 1994; Poeppel & Wexler 1993, and references cited there).

The second task for a child acquiring the FCs of L is to determine their sequence (their c-command relations). As mentioned above, it is possible that this is largely determined by UG and I am assuming that variability in the order of FCs that are directly related to morphological distinctions follows from the Mirror Principle.

Third, a child acquiring L has to determine whether a given functional category F is "present" in clauses of type A or not? Consider the following:

(5) *The Functional Projection Alibi (FPA)*
Let F be a functional head and Φ the set of ϕ-features (and/or other features) which F contains. Now assume that the different values for the ϕ-features in Φ are systematically represented in overt morphology in language L in clauses of type A but never in clauses of type B. Assume further that there is evidence for overt syntactic movement to F and/or to Spec-FP in language L in clauses of type A but no such evidence can be found in clauses of type B. Then we may hypothesize that F is absent in clauses of type B in language L although it is present in clauses of type A.

It seems that many linguists are reluctant to accept reasoning of the type exemplified in (5). There are two obvious reasons. First, since linguists assume in many cases that certain morphological features are "present" although they are not represented in overt morphology (such that English nouns check [Nom and Acc] Case features even if there is no Nom/Acc distinctions in the overt morphology of English nouns), they are reluctant to take lack of overt morphological distinctions as valid evidence for hypothesizing the absence of a func-

tional projection. Second, since MPLT assumes covert syntactic movement (LF-movement) in addition to overt movement (movement before spellout), lack of overt movement to F^0 or Spec-FP in a given clause type (where F^o is some functional head) does not necessarily mean that F is absent in this particular type of clause. I want to argue, however, that the reasoning behind the FPA is in the spirit of Chomsky's general principle of Economy of Representation (see, e.g., Chomsky 1991: 437; cf. also Grimshaw 1993) and I want to claim that a child acquiring L is guided by the principle stated in (6):

(6) *The Real Minimalist Principle* (RMP):
Assume only those functional categories that you have evidence for.

This means, then, that a child acquiring L will not assume that F_x is present in L unless (s)he finds evidence suggesting that it is. Furthermore, a child will not automatically assume that F_x is present in all clause types in L although it is clearly present in some. (S)he will look for evidence in each clause type. What counts as "evidence" in this connection is by no means obvious, but the RMP is clearly very close to Iatridou's original suggestion referred to above, namely that "evidence for an AgrP [...] will have to be found in each language separately" (Iatridou 1990: 553).

1.3. *Proposals and Organization*

The above discussion has concerned possible cross-linguistic and intra-linguistic variation with respect to FCs. In the rest of the paper I will be mainly concerned with one particular instance of cross-linguistic variation of this type.[6] The main proposal I want to argue for in this paper is the following:

---

6. Certain aspects of the intra-linguistic variation of FCs are discussed in Thráinsson 1993a, in preparation. One could also mention here the Germanic controversy about main clauses as CPs or IPs in the so-called verb-second languages. Contrary to common analyses, Rögnvaldsson and Thráinsson (1990) argue that subject-initial main clauses in Icelandic are typically IPs rather than CPs and this is also the position argued for in Ottósson (1989) for Icelandic, Abraham (1993b) for German, and Zwart (1993a, b) for Dutch (*mutatis mutandis,* since Zwart is assuming an expanded IP and thus his AgrSP corresponds to Abraham's IP, for instance). But if main clauses are IPs (or AgrSPs) (unless a non-subject is topicalized) but embedded clauses are typically CPs, then that is a case of "intra-linguistic variation of FCs".

(7) *The Split-IP Parameter (SIP)*
Languages that have a positive value for SIP have AgrSP and TP as separate functional projections. Languages with a negative value of SIP are characterized by an unsplit (pre-Pollockian) IP.

The SIP is to be understood in the spirit of the discussion in 1.2 above, in particular the Real Minimalist Principle (RMP). I am assuming that unless a child finds morphological or syntactic evidence for a split IP in L, (s)he will not assume separate FPs for agreement and tense. The relevant features of verbs will just be checked in a "fused" IP projection.[7]

The organization of the remainder of the paper is as follows. In Section 2, I briefly review common assumptions about the nature of AgrP and TP. In Section 3, I present well known facts about verb movement and subject-verb agreement in Scandinavian and argue that the Mainland Scandinavian (MSc) languages do not have a split IP. I then show that Chomsky (1993) in fact "comes very close" to claiming that English does not have a split IP either and I will assume that it does not.[8] In Section 4, I argue that certain syntactic

---

7. The discussion centers around Agr and T here since these are the FCs usually included in a split IP. But as Iatridou (1990) points out, Pollock's approach immediately raises the question whether we could have a MoodP and an AspectP, not just in some languages but universally. Interestingly, most syntacticians working on Germanic languages, for instance, seem to assume that Mood and Aspect do not play any role as FCs in these languages. I have seen no syntactic evidence suggesting that they do and the approach advocated in the present paper would predict that they do not since Mood and Aspect are not marked separately in the overt morphology of these languages. See the discussion in Section 3.3 below and the discussion in Bobaljik 1995b, especially Chapter 1.

8. Interestingly, Chomsky (1995b) suggests that Agr may be eliminated entirely, not just in languages like English and MSc as suggested here. The basic idea is that in cases where there seems to be syntactic evidence for a specifier position associated with an Agr-projection, we are in fact dealing with an additional specifier position of a lower syntactic category. So instead of Spec-AgrSP and Spec-TP we would be dealing with two specifier positions of TP (or of an unsplit IP). As we will see in Section 4 below, an analysis along these lines seems quite problematic for Icelandic, for instance. See also Jonas (1995, Chapter 2) for arguments against eliminating Agr entirely, and Bobaljik (1995b) for a very interesting and thorough discussion of why one might (Part I) or might not (Part III) want to relate morphology and syntax as closely as suggested in the present paper.

differences between languages like Icelandic on the one hand and MSc and English on the other can be accounted for if one assumes that Icelandic has a positive value for SIP and MSc and English have a negative one. Section 5 is the conclusion.

2. On the Nature of AgrP and TP

Chomsky (1993) adopts the widely accepted VP-internal subject hypothesis (see e.g. Kitagawa 1986; Koopman & Sportiche 1991; Speas 1990; Fukui & Speas 1986; Zagona 1982; compare also Burton & Grimshaw 1992 and McNally 1992), with subjects "originating" in Spec-VP but moving (overtly or covertly) to the specifier position of a higher (functional) projection. He discusses the functional heads Agr and T and their role in the checking of morphological features. If we want to argue that some languages do not separate Agr (or AgrS) and T, it is necessary to make clear what the morphological and syntactic role of these categories is taken to be. (9) provides a representative sample of the proposed ideas:

(9) a. "Agr is a collection of ϕ-features (gender, number, person); these are common to the systems of subject and object agreement, though Agr_S and Agr_O may of course be different selections, just as two verbs or NPs ... may differ" (Chomsky 1993: 7).[9]

b. "We now regard both agreement and structural Case as manifestations of the Spec–head relation (NP, Agr). But Case properties depend on characteristics of T and the V head of VP. We therefore assume that T raises to Agr_S ... and V raises to Agr_O...; the complex includes the ϕ-features of Agr and the Case feature provided by T, V" (Chomsky 1993: 7). "I overlook here the possibility of NP raising to [Spec, T] for Case assignment, then to [Spec, Agr_S] for agreement. This may well be a real option" (Chomsky 1993: 45, fn. 11).

---

9. Chomsky 1995b maintains, on the other hand, that Agr "lacks ϕ-features" whereas V and T do have such features. The claim is that when it seems that Agr does have such features, they are in fact provided by a T or a V that has raised. This is then used as an argument for eliminating Agr from UG (see note 8 above).

These quotes clarify the basic ideas. Note in particular that the case marking (checking) and agreement properties that were associated with INFL in pre-Pollockian X′-structures are now divided between AgrS and T.

3. Functional Projections in Scandinavian and English

3.1. *Verb Movement and Subject–Verb Agreement in Scandinavian*

First, consider the following well known contrast between Danish (a representative of Mainland Scandinavian) and Icelandic:

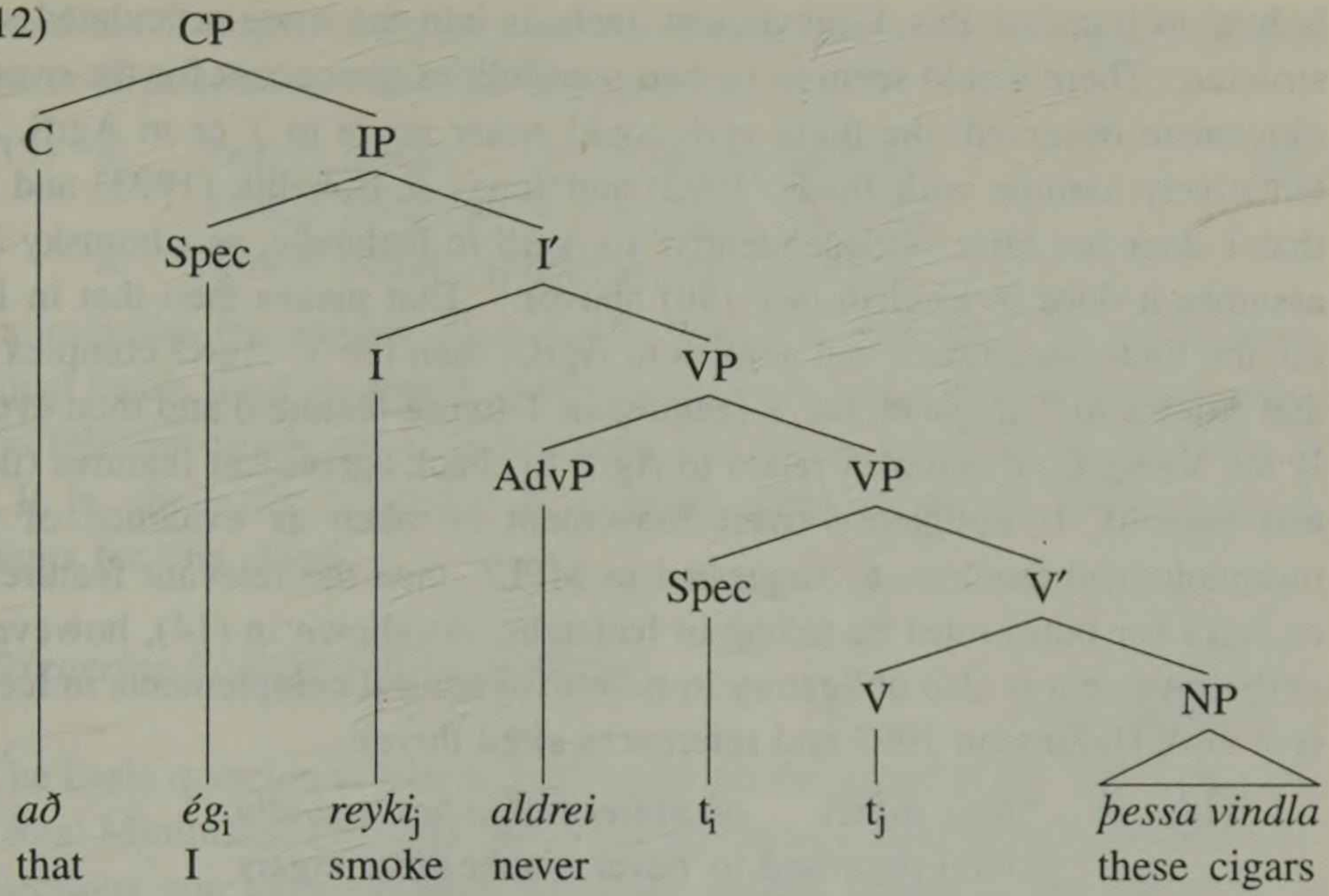
(10) a. *Peter tvivler på at jeg aldrig ryger disse cigarer.* (Da)
Peter doubts on that I never smoke these cigars
b. **Peter tvivler på at jeg ryger*$_i$ *aldrig* t_i *disse cigarer.*
(cf. Vikner 1994: 123)

(11) a. **Pétur efast um að ég aldrei reyki þessa vindla.* (Ic)
Peter doubts about that I never smoke these cigars
b. *Pétur efast um að ég reyki*$_i$ *aldrei* t_i *þessa vindla.*

As shown here, movement of the finite verb across a sentential (or "medial") adverb is impossible in "non-bridge verb" complements in Danish but obligatory in Icelandic. Vikner (1991, 1994) takes this as evidence that Danish does not have "independent V-to-I" as he calls it, namely movement of the finite verb to I independent of the finite verb movement to C in main clauses (and main-clause-like clauses) that Vikner takes as the defining characteristic of verb-second (V2) languages. This "independent V-to-I" movement is illustrated in (12), given the (non-expanded) IP structure Vikner assumes and the adjunction of the sentential (or "medial") adverb to VP he argues for:[10]

10. Rögnvaldsson & Thráinsson (1990) assumed that medial adverbs in Icelandic were adjoined to V′. This was criticized by Vikner (1994: 140), who pointed out that adjunction to V′ would be "impossible" in the X′-theory outlined by Chomsky (1986b). It is therefore of some historical interest at least to note that Chomsky (1995a) suggests that adjunction to VP (and other semantically active maximal projections) should be impossible but adjunction to functional projections and all single-bar level categories should be legitimate. The empirical arguments that Vikner (1994) provides against V′-adjunction of medial adverbs in Icelandic seem valid, however (see also Thráinsson 1994). Hence, I here follow Vikner and others in assuming that these adverbs are adjoined to VP.

(12)



Like many other linguists before (and after) him, Vikner wants to relate differences in verb movement to differences in verbal morphology (Vikner 1994: 118 ff. and Vikner 1995b; for similar ideas see, e.g., Kosmeijer 1986; compare also Kosmeijer 1993; Holmberg & Platzack 1991, 1995; Platzack 1988; Roberts 1993; Rohrbacher 1994). Languages with "sufficiently rich" verbal morphology (such as Icelandic) have independent V-to-I, whereas languages with "poor" verbal morphology do not. As shown by Vikner (1994: 119), there is a clear difference between Icelandic and Danish in this respect:

(13)

		Icelandic	Danish
	inf.	*reykja*	*ryge*
sg.	1.	*reyki*	*ryger*
	2.	*reykir*	*ryger*
	3.	*reykir*	*ryger*
pl.	1.	*reykjum*	*ryger*
	2.	*reykið*	*ryger*
	3.	*reykja*	*ryger*
		'smoke'	'smoke'

Leaving aside the "cause" of verb movement for now, it is clear that if we accept the basic argumentation and assumptions described above, the question

is how to translate this V-movement analysis into the more articulated split IP structure. There would seem to be two possibilities to account for the overt verb movement observed: the finite verb could either move to T or to AgrS. Let us tentatively assume with Jonas (1992) and Jonas & Bobaljik (1993) and others that T does not raise "independently" to AgrS in Icelandic, as Chomsky (1993) assumes it does in English (see (9b) above).[11] That means then that in Icelandic the finite verb raises and adjoins to AgrO, then the V–AgrO complex raises and adjoins to T to check the V-features in T (tense features) and then eventually the V–AgrO–T complex raises to AgrS to check agreement features (number and person). If obligatory overt movement is taken as evidence of strong morphological features, as suggested in MPLT, then the relevant features of T or AgrS (or both) must be strong in Icelandic. As shown in (14), however, this verb movement is also obligatory in infinitival control complements in Icelandic (see also Thráinsson 1993 and references cited there):

(14) a. **Jón lofaði að aldrei reykja þessa vindla.*
John promised to never smoke these cigars
b. *Jón lofaði að reykja*$_i$ *aldrei* t$_i$ *þessa vindla.*

The facts illustrated in (14) are theoretically important. Since non-agreeing verbs must move in complements of this sort, the "cause" of movement cannot be that the verb needs to "pick up" (overt) agreement morphemes "situated" in some functional head position, as pre-checking theories of movement would have it. The facts are, however, consistent with a checking theory account that maintains that obligatory verb movement is triggered by strong features of functional heads. It is clear from examples like (11) and (14) that the relevant features are strong in Icelandic, regardless of the actual morphology of the verb form involved. Note also that if Thráinsson (1993) is right in claiming that there is no verb movement to AgrS in infinitival complements in Icelandic (since the infinitival marker is in AgrS and the non-agreeing verb does not move there to check agreement features since it does not have any), then the observed verb movement in (14) must be to T (more accurately: the verb moves and adjoins to

---

11. "Independently" means here "independent of any V-movement to T", not "independent of the V2–characteristic V-to-C movement", as in Vikner's terminology. I apologize for this somewhat confusing terminology here.

AgrO and the V–AgrO complex then moves and adjoins to T). Again, no such movement is possible in MSc, as indicated by (15b):[12]

(15) a. *Anders lovade att inte dricka mjölken.* (Sw)
Anders promised to not drink the milk
b. **Anders lovade att dricka*$_i$ *inte* t$_i$ *mjölken.*

In a minimalist framework, this means that the V-features of the relevant functional heads must be strong in Icelandic and weak in MSc. I want to claim that the relevant heads are T in Icelandic and I (in an unsplit IP structure) in MSc. It is obviously necessary to provide some empirical and theoretical arguments for this claim.

3.2. *Triggering Positive Setting of the SIP*

The basic question is how a child could set the value of the SIP. According to the Real Minimalist Principle (RMP) in (6) above ("assume only the functional projections you have evidence for"), the default setting of the SIP must be negative. Evidence for positive setting could include the following:

(16) Evidence for positive setting of the SIP:
a. "Rich" agreement and tense morphology.
b. Syntactic evidence for the additional positions provided by a split IP structure.[13]

It is well known that it has proven to be notoriously difficult to define "rich verbal morphology" in such a way that it makes the right predictions regarding overt verb movement (see e.g. Rohrbacher 1994; Vikner 1994, 1995b and references cited there). Note, however, that if the approach advocated here is correct, the correlation is between rich verb morphology and split IP rather than rich verb morphology and overt verb movement *per se*. I will argue below that it is possible to explain this correlation whereas it has always seemed puzzling

---

12. Since the negation always precedes the infinitival marker in Danish, it is easier to use Swedish examples to illustrate lack of non-finite verb movement in MSc. For discussions of differences between MSc infinitives see Platzack (1986), Johnson & Vikner (1994) and references cited there.

13. By "additional positions" I am mainly referring to the fact that instead of just Spec-IP and I, a split IP structure has Spec-AgrSP, AgrS, Spec-TP and T as "positions" for maximal projections and heads. We will return to the relationship between AgrS and AgrO below.

why there should be any kind of correlation between, say, a distinction between two or more persons in verbal morphology and whether or not there is overt verb movement.[14]

3.3. *What Counts as Morphological Evidence for a Positive Setting of the SIP and Why?*

I have already discussed the difference between Icelandic and MSc with respect to verbal morphology. English, where the only distinction in overt agreement is between the third person singular and everything else (*smokes*, *smoke*) is obviously more similar to MSc than to Icelandic. But how rich does the verbal morphology have to be, and why should it matter?

An overview of various definitions of "rich verbal morphology" is given in Vikner (1994, 1995b). He rejects, for instance, a suggestion to the effect that a distinction between different persons (which MSc does not have but English has a minimal residue of) is what matters for the relevant definition of "rich" verbal morphology. He attributes this suggestion to Platzack & Holmberg (1989). Similar ideas have been developed by Rohrbacher (e.g. 1993, 1994). Vikner adopts a proposal he attributes to Roberts (1993: 267): verbal morphology is rich in the relevant sense if there are person distinctions both in the singular and plural (Vikner 1994: 121; see Vikner 1995b for modifications). Neither MSc nor English would have rich verbal morphology under this definition. Hence there would be no morphological evidence for a positive setting of the SIP in these languages if this is the relevant distinction.

Even if this definition of rich verbal morphology were to make the correct empirical predictions, it is rather mysterious why a distinction between two or more persons in the agreement morphology should trigger a positive setting of a parameter concerning the inclusion of both Agreement and Tense (or the distinction between them) in the inventory of syntactic categories. Bobaljik &

---

14. It remains to be seen whether there are any languages that do not have a split IP but do have overt verb movement to I. More specifically, could a language "lose" rich verbal morphology and not lose verb movement at the same time. See the discussion of the history of English verb movement in van Gelderen (1993) and Jonas (1995) and references cited there. See also the discussion in Bobaljik (1995b), Sections 1.3 and 1.4, on relevant morphological differences between the Germanic languages. I am assuming that V-to-C is irrelevant in this respect since it is presumably triggered by factors independent of inflection.

Jonas (1993) make a somewhat different proposal. They argue that Spec-TP is "licensed at S-structure" only in a language having "independent Tense and Agreement morphology". Tense and Agreement morphology is "independent" in this sense, according to B&J, if a verb form can bear both an Agreement marker (person, number) and a Tense marker simultaneously. The basic idea can be illustrated by the following examples from Icelandic and English:[15]

(17)

	Icelandic		English	
	Present	Past	Present	Past
sg. 1	*reyk.i*	*reyk.t.i*	*smoke*	*smoke.d*
2	*reyk.ir*	*reyk.t.ir*	*smoke*	*smoke.d*
3	*reyk.ir*	*reyk.t.i*	*smoke.s*	*smoke.d*
pl. 1	*reykj.um*	*reykj.t.um*	*smoke*	*smoke.d*
2	*reyk.ið*	*reyk.t.uð*	*smoke*	*smoke.d*
3	*reykj.a*	*reykj.t.u*	*smoke*	*smoke.d*
	smoke.AGR	smoke.T.AGR	smoke.AGR	smoke.T

Here we see clearly that Icelandic has person and number agreement in present and past tense whereas English has minimal agreement in the present tense but none in the past tense. There is thus no verb form in English that carries an overt agreement marker and an overt tense marker simultaneously. In this sense morphological agreement is not "independent" of morphological tense in English. The same obviously holds for MSc since verbs in the Mainland Scandinavian languages do not have any overt agreement markers at all.

Now consider how this could relate to the acquisition of functional

15. The exact number of overt person distinctions in the verbal paradigm in Icelandic depends on the inflectional class of the verb. Sometimes 2nd and 3rd person are not overtly distinguished (as in Pres Sg in the example in (17)), sometimes 1st and 3rd person are not (as in Past Sg in (17)). The main point to note here is that the agreement morphemes are easily separated from the (Past) tense morpheme. The orthography is a quite reliable indicator of the distinctions involved in Icelandic, whereas it can be dangerous in many languages to rely on conventional spelling when trying to determine what the morphological distinctions are, as discussed in the paper by Solà in this volume. Jan-Wouter Zwart also points out to me (p.c.) that Dutch verbal morphology is not as rich as the spelling suggests. The same is true of French verbal morphology. I will return to this question below.

categories. According to the Limited Diversity Hypothesis (LDH) advocated here, the sequence of morphologically related FCs may vary from language to language, as stated in (4c) above, repeated here for convenience:

(18) The sequence (c-command relations) of those functional categories (dominance relations between the functional projections) that are directly related to morphological distinctions may vary from language to language, consistent with the Mirror Principle.

Recall the version of the Mirror Principle outlined in Section 1.1 above referred to as the Lexicalist Mirror Principle. According to that principle, a morphological feature associated with an overt morpheme close to the stem of a verb has to be checked off before a morphological feature associated with an overt morpheme farther from the stem can be checked. Assuming the Strict Cycle, this means that there is morphological evidence for the c-command relation between AgrS and T in Icelandic but no such evidence can be found in English (nor in MSc). The morphological evidence in Icelandic suggests AgrSP over TP (because the tense morpheme is "inside" the agreement morpheme), just like in French. Morphological evidence in Arabic suggests the reverse relation since there the tense morpheme is "outside" the agreement morpheme according to Ouhalla (1994: 45):

(19) a. *sa.ya.zuuru l.?awlaad.u xaal.a.hum* (Ar)
T.AGR.visit the.boys.NOM uncle.ACC.their
'The boys will visit their uncle.'

b *Ils arriv.er.ont demain* (Fr)
they come.T.AGR tomorrow
'They will arrive tomorrow.'

It is important to make clear what is being claimed here. The first claim is that a child acquiring Icelandic, French and Arabic will find morphological evidence for the independence of agreement and tense morphemes because they can be isolated in the morphology of these languages. The second claim is that because of the Lexicalist Mirror Principle this evidence tells the child that these languages must have separate functional projections where agreement and tense features are checked and that they must be hierarchically ordered in a particular way (with AgrS above T in Icelandic and French but the other way around in Arabic). A child acquiring English and MSc will not encounter any such evidence. Hence (s)he will make the default assumption that there is no split IP in these languages and the SIP will retain its initial negative value (unless additional evidence is encountered to suggest otherwise).

As mentioned in Section 1.3, Chomsky (1993) actually "comes very close" to saying that AgrS and T are not split in English:

(20) "We therefore assume that T raises to Agr_s ... and V raises to Agr_o ..." (Chomsky 1993: 7)

If T *always* raises to AgrS, independently of any verb movement, then T and AgrS do not really act as separate functional heads the way they do in languages where this independent T-to-AgrS does not take place and the verb may raise to T and the complex head thus formed will check the Case of a subject in Spec-TP. The return to an unsplit IP analysis for English, as suggested here, can thus be seen as a reformulation of Chomsky's analysis. Rather than saying that AgrS and T do not act as separate functional heads because T always raises to AgrS independent of any verb movement, I suggest instead that the IP is not split in languages like English and MSc. There is no AgrS head which is separate *at any level* from the T head but only an unsplit Infl-head. This is an analysis empirically distinct from Chomsky's (1993) "invariant T-to-AgrS movement" analysis. In particular, if it can be shown that a T-trace or the T-to-AgrS chain plays any role in languages like MSc and English, then the present analysis would be problematic while such facts would corroborate Chomsky's movement analysis. Let us therefore turn to this question and to possible syntactic evidence for the setting of the SIP.

4. Finding Syntactic Evidence for a Positive Setting of the SIP

4.1. *Theoretical Plausibility of an Unsplit IP*

Before proceeding, we need to consider whether it is theoretically necessary that T and Agr be distinct functional heads. Checking theory could demand, for instance, that Case and agreement be checked against different functional heads. The checking theory outlined in MPLT does not appear to require this, however. Thus, it is not clear that the trace of T plays any crucial role in Case checking in languages like English and MSc, i.e. a role that could not be fulfilled by the adjunction structure T–AgrS and hence also the unsplit Infl head under the analysis advocated here.[16] It is also suggestive that a theory incorporating an unsplit IP seemed to work quite well for English (and MSc) for many years and

16. See Jonas (1995, this volume), however, for a T-raising analysis of English and MSc.

the most suggestive arguments for a split IP come from languages like French (Pollock 1989 — see also Vanden Wyngaerd 1989a) or Icelandic (Jonas 1992; Jonas & Bobaljik 1993; Bobaljik & Jonas 1993; Bobaljik 1995b; Jónsson 1995) and not from languages like English or MSc. This section reviews some of this evidence.

4.2. *Subject Positions in Transitive Expletives*

Consider again the quote at the beginning of this paper, the relevant part of which is repeated here for ease of reference:

(21) "The structure of CP [see (1) above] is largely forced by ... properties of UG, assuming the minimalist program with Agr abstracted as a common property of inflectional systems, a reasonable assumption, given that agreement appears without Case (as in NP–AP agreement) and Case appears without agreement (as in transitive expletives, with the expletive presumably in [Spec,Agr_s] position and the subject in [Spec,Tense], receiving Case" [Chomsky 1993: 9]).

Note that Chomsky cites "transitive expletives" here as a type of construction where "Case appears without agreement" and should therefore be checked against a separate functional head, namely in Spec-TP against the functional head T. While it is not entirely clear which constructions Chomsky has in mind, it is usually assumed that only a subset of Germanic languages, for instance, allow the transitive expletive construction (TEC). This is pointed out in Holmberg & Platzack (1991) and Vikner (1991) and discussed extensively in Jonas & Bobaljik (1993) and Jonas (1992, 1994, 1995, this volume), for instance (see also Platzack 1983, Maling 1988, Vikner 1991, and Zwart 1992a). Typical examples are given in (22):

(22) a. *Það luku einhverjir stúdentar ekki þessu verkefni* (Ic)
there finished some students not this assignment

b. *There ate a student the pudding

c. **Det åt en student puddingen* (Sw)
there ate a student the pudding

If Chomsky (1993) is right in maintaining that in TECs the overt expletive is in Spec-AgrSP and the lexical subject is in Spec-TP in TECs, then the grammaticality of (22a) in Icelandic vs. the ungrammaticality of its English and

MSc counterparts in (22b,c) suggests that Icelandic has "two subject positions" available (namely Spec-AgrSP and Spec-TP) where English and MSc have only one. This is in fact what Bobaljik & Jonas (1993, henceforth B&J), Jonas & Bobaljik (1993, J&B hereafter), and Jonas (1992, 1994, 1995, this volume) argue. They claim that Spec-TP is not "available" as a Case checking position for subjects in MSc and English (see the discussion around (18) above). This would follow, of course, if AgrS and T are not separate functional heads in English and MSc, as proposed here. Under such an analysis only Spec-IP would be available as an "external" (i.e. VP-dominating) subject position in these languages.[17]

Note in passing that the difference between Icelandic on the one hand and English and MSc on the other does not follow straightforwardly from a "multiple specifier" analysis (along the lines discussed in Chomsky 1995b) which maintains that TP can have two specifiers in Icelandic but only one in English and MSc. This is so because the word order we get in Icelandic is expletive–finite verb–NP and not expletive–NP–finite verb as one might expect if the head which the verb moves to in Icelandic (i.e. T) could project two specifiers.

In addition, it should be pointed out here that Jonas (1992, 1994, 1995, this volume) argues that there is evidence for the "additional subject position" provided by Spec-TP in various other constructions in Icelandic but not in MSc or English. These constructions include Passive Expletive Constructions and Raising Expletive Constructions. If Jonas is right, these constructions provide the child acquiring Icelandic with additional pieces of evidence for the positive setting of SIP whereas no such evidence is available in MSc or English.[18]

---

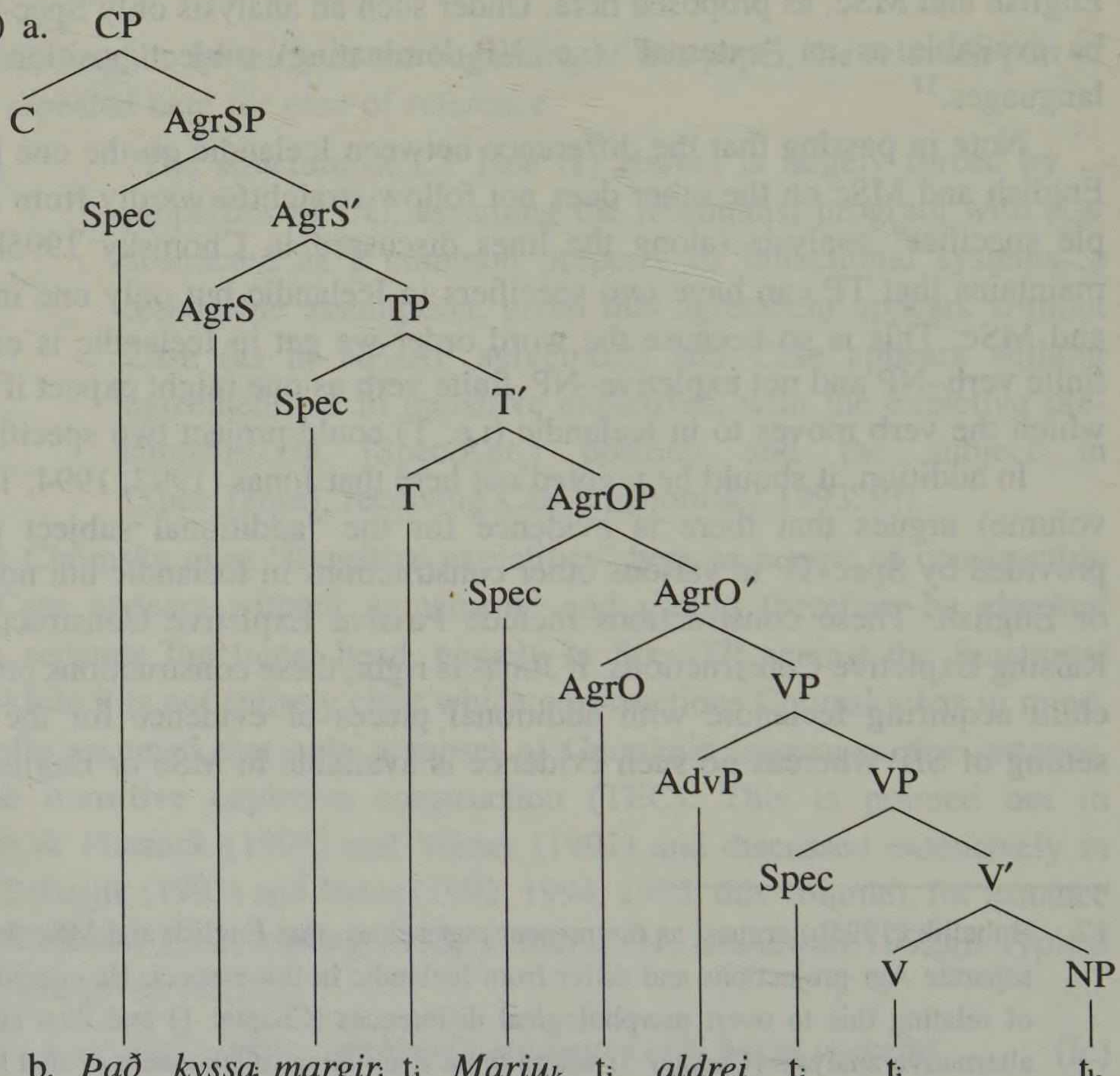
17. Bobaljik (1995b) argues, as the present paper does, that English and MSc do not have separate Agr-projections and differ from Icelandic in this respect. He considers a way of relating this to overt morphological differences (Chapter 1) and then suggests an alternative analysis (Chapter 5) where there is no direct relationship of that kind. Both analyses are framed in a non-lexical (or late lexical insertion) model of morphology-syntax interaction in the spirit of Halle and Marantz 1993. Under the former analysis, the "syntax is filtered by the morphology", under the latter "the syntactic computation is uniformly blind to the ramifications of its output in the morphological component" and the cross-linguistic differences are supposed to reflect different values of the "Free Agr Parameter" (1995b: 21). In connection with the latter analysis, Bobaljik suggests a "single output syntax", much in the spirit of Groat and O'Neil (this volume).

18. Jonas (1995, this volume) argues for a different analysis for MSc and English, however, namely an independent T-to-AgrS analysis along the lines of Chomsky (1993,

4.3. *Evidence from Object Shift in Icelandic, MSc and English*

J&B and B&J and Bures (1993) argued that full NP-Object Shift (NP-OS) is available only in languages that have Spec-TP available as a checking position for subject Case. Icelandic has NP-OS but neither English nor MSc do. They argue that Icelandic sentences like the (23b) provide excellent evidence for the expanded IP structure illustrated in (23a):

(23) a.



b. *Það kyssa*$_j$ *margir*$_i$ t_j *Mariu*$_k$ t_j *aldrei* t_i t_j t_k
there kiss many Mary never
'Many never kiss Mary.'

1995a). But her analysis differs from Chomsky's (1995a) analysis in that she argues that a bare T (T which no V adjoins to) does not project a specifier position. Hence Spec-TP is not an available Case checking position (nor EPP-feature checking position) for the subject in these languages.

In (23) it is assumed that the object has shifted out of the VP and across the adverb *aldrei* 'never'. Evidence for this analysis comes from the fact that *aldrei* is an adverb that cannot occur VP-finally when no NP-OS takes place, as illustrated in (24):[19]

(24) **Stúdentarnir hafa lokið verkefninu aldrei.*
the students have finished the assignment never

J&B and B&J argue that sentences of the type illustrated in (23) crucially involve Spec-TP as a checking position for the subject since otherwise the minimalist principle of Shortest Move would be violated (see note 19 and the discussion below). If this is correct, such sentences provide the language learner with evidence for a positive setting of the SIP in Icelandic. No comparable evidence is available to the child acquiring MSc or English.

It is necessary, however, to reconsider the arguments presented by J&B and B&J under the unsplit IP analysis advocated here. First, consider a structure such as the following (with a MSc (Swedish) example):

---

19. As Holmberg (1986) first observed, NP-OS in Icelandic cannot occur if there is an auxiliary verb in the clause. This is known as Holmberg's Generalization. It is illustrated by the contrasts in (i-a)–(i-d):

(i) a. *Hann las aldrei bókina.*
he read never the book

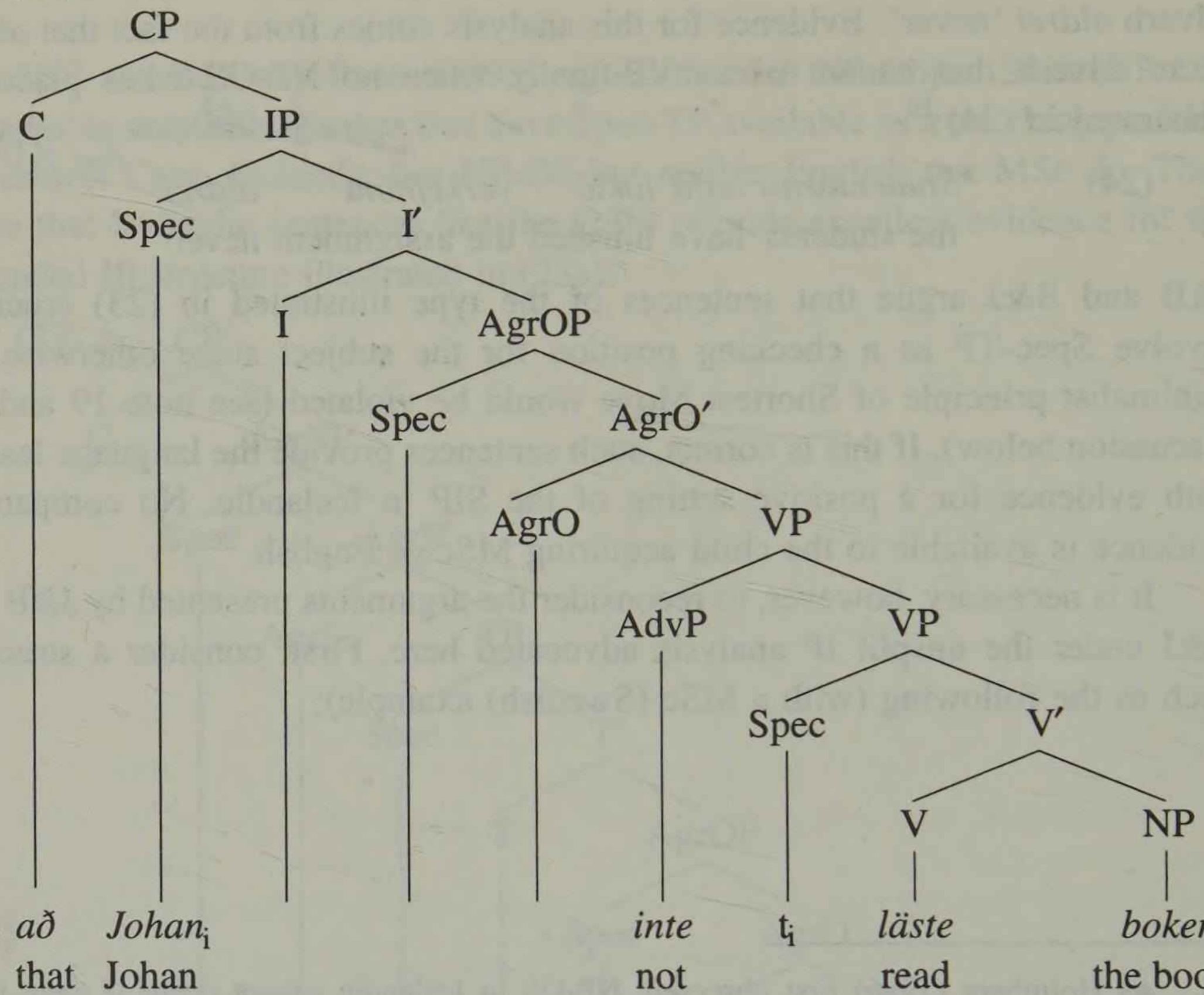
b. *Hann las bókina aldrei.*
he read the book never

c. *Hann hefur aldrei lesið bókina.*
he has never read the book

d. **Hann hefur bókina aldrei lesið.*
he has the book never read

J&B provide the following account in minimalist terms: If there is no auxiliary verb in the clause, the main verb leaves the VP and moves (at least) to T. If there is an auxiliary verb in the clause, it will occupy T (or move through T on its way to AgrS) so the main verb cannot move there. NP-OS is dependent on (main) verb movement because otherwise the domain of the main verb would not be extended. Hence Spec-AgrOP and Spec-TP would not be equidistant from Spec-VP so (NP-OS and subsequent) subject raising to Spec-TP would violate the principle of Shortest Move. See also the discussion in Thráinsson (1993a: 202–203). For a different view see Bobaljik (1995b), especially Parts II and III.

(25)



Since there is no V-movement in (25) (cf. the discussion in 3.1 above), the checking domain of the verb is not extended. According to the analysis advocated in J&B and B&J, Spec-VP and Spec-AgrOP will thus not be equidistant from the object so the object cannot shift across Spec-VP to Spec-AgrOP without violating the principle of Shortest Move.[20] Hence (26) is correctly predicted to be ungrammatical in Swedish:

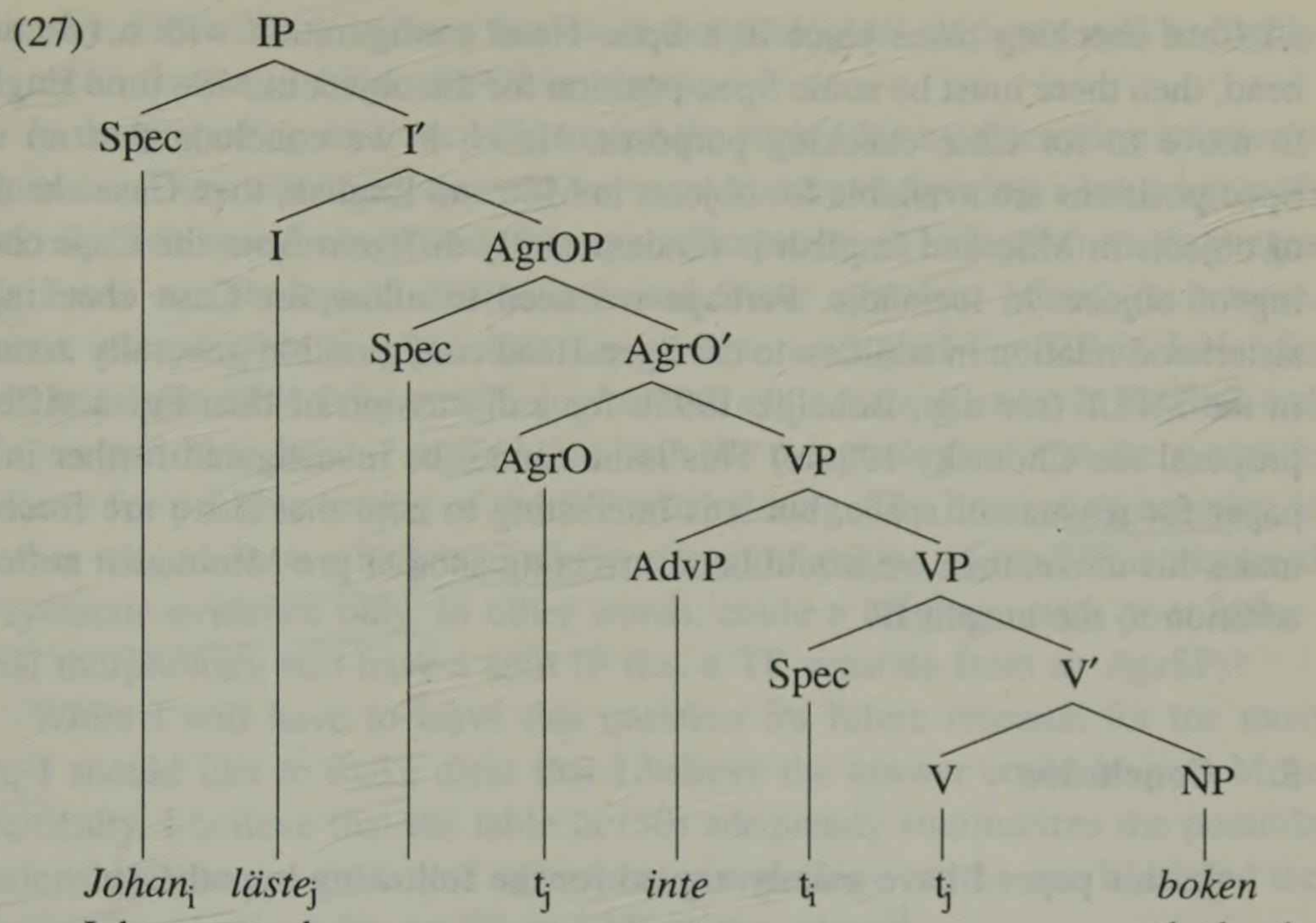
(26) *...att *Johan*$_i$ *boken*$_j$ *inte* t$_i$ *läste* t$_j$
that Johan the book not read

But now recall that there is verb movement in main clauses in MSc. This is illustrated in (27), assuming (with Travis 1991, for instance) that this verb movement is only to I and not to C (contra Vikner 1991, 1994 and others), and that it adjoins to AgrO on the way:

---

20. For a slightly different analysis involving the notion of "closest checking position" rather than Equidistance see Ferguson (1993, this volume) and Ferguson & Groat (1994).

(27)



Despite this verb movement, NP-OS is just as impossible in Swedish main clauses as it is in embedded clauses. This is shown in (28) (compare (26)):

(28) **Johan*$_i$ *läste*$_j$ *boken*$_k$ t_j *inte* t_i t_j t_k (Sw)
Johan read the book not

The ungrammaticality of (28) cannot be predicted by reference to the principle of Shortest Move, given the clausal architecture in (27). So we must look for another explanation.

One possible explanation is that the principle of Shortest Move does not apply the way I have assumed and needs to be revised. Another possibility is that the relevant N-features of AgrO are weak in MSc although they are (optionally) strong in Icelandic. A third possibility is that Chomsky (1991, 1993) is right in assuming that AgrS and AgrO are in fact two sides of the same coin. If that is the case, then it seems natural to assume that if a language has no AgrS then it has no AgrO either. That would explain why there is no NP-OS in MSc (and English) — there simply is no Spec-AgrOP for the object to shift to.[21] But this suggestion obviously has serious implications for Case theory. If

---

21. I am assuming with Josefsson (1992), Déprez (1989), J&B and others that the pronominal object shift found in Swedish, for instance, is different from NP-OS and does

all Case checking takes place in a Spec–Head configuration with a functional head, then there must be some Spec-position for the object in MSc (and English) to move to for Case checking purposes. Hence if we conclude that no such Spec-positions are available for objects in MSc and English, then Case checking of objects in MSc and English is fundamentally different from the Case checking of objects in Icelandic. Perhaps we need to allow for Case checking in sisterhood relation in addition to the Spec-Head configuration generally assumed in the MPLT (see e.g., Bobaljik 1995b for a discussion of this. For a different proposal see Chomsky 1995b.) This issue cannot be investigated further in this paper for reasons of space, but it is interesting to note that if we are forced to make this move, then we would be resurrecting another pre-Minimalist notion in addition to the unsplit IP.

5. Conclusion

In this paper I have mainly argued for the following hypothesis:

(29) There is parametric variation between languages with respect to the nature of functional categories (FCs) implemented. In particular, some languages have a split IP whereas others do not.

If the relationship between morphology and syntactic structure is as close as the spirit of MPLT would lead one to believe, this claim seems very natural. This is so because languages vary greatly with respect to overt morphology and some of the FCs are supposed to be directly related to morphological distinctions that are not made overtly in all languages. If the sequence (c-command relations) of such FCs obeys the Lexicalist Mirror Principle described above (based on Baker's original idea; 1988: 13, and related to an idea outlined by

---

not involve movement to Spec-AgrOP. It may be some sort of cliticization instead. It was in fact already pointed out in Holmberg (1986: 229–230) that pronominal object shift has "something in common with cliticization of object pronouns as found in many languages of the world" although Holmberg also argues that they are not French-type clitics. Pronominal object shift is illustrated in (i) with a Swedish example:

(i) *Johan läste den*$_i$ *inte* t_i
Johan read it not

I am also making the standard assumption, contra Johnson (1991), that English does not have NP-OS.

Chomsky; 1993: 28), it is clear that overt morphological evidence will be highly relevant to the acquisition of FCs in syntax.

In the preceding sections I illustrated the basic idea by discussing relatively easy cases. They were easy because we were mainly comparing a language with relatively rich morphology and rich syntactic patterns (Icelandic) to languages with fewer morphological distinctions and fewer syntactic "slots" as it were. Thus it was possible to argue that the positive setting of the SIP in Icelandic could be triggered both by morphological and syntactic evidence. In English and MSc, on the other hand, I argued that both the morphological and the syntactic evidence for positive setting of the SIP was missing. The interesting question is whether we can have a language where a positive setting of the SIP is triggered by syntactic evidence only. In other words, could a language with poor inflectional morphology still have a split IP (i.e. a TP separate from an AgrSP)?

While I will have to leave this question for future research for the most part, I should like to make clear that I believe the answer could be *yes*. More specifically, I believe that the table in (30) adequately summarizes the possible relationship between tense and agreement morphology on the one hand and the functional projections IP, AgrSP and TP on the other:[22]

(30)

Tense and agreement morphology	Fused IP	Split IP
fused morphology	necessary	impossible
split morphology	impossible	necessary
ambiguous morphology	possible	possible

The table in (30) is to be interpreted as follows: If the overt morphology of a given language clearly indicates that the tense and agreement morphemes are fused, i.e., that there is one overt inflectional morpheme which simultaneously marks tense and agreement, then the language cannot have a split IP.[23] This is

---

22. This is very similar in spirit to the discussion in Bobaljik (1995b: 49–52).

23. Obviously, something more specific has to be said about what counts as "fused" Agr and Tense morphology in the relevant sense. Bobaljik (1995b: 43) proposes the definition in (i), based on a survey of the Germanic languages:

    (i) If the appearance of Tense morphology blocks the appearance of Agreement morphology, then Tense and Agreement Vocabulary Items are in complementary distribution and T and Agr must be fused.

    Although Bobaljik is assuming a Distributed Morphology framework (cf. Halle and

so because according to the Lexicalist Mirror Principle the tense and agreement features will have to be checked simultaneously in cases of this sort as they will be associated with the same inflectional morpheme. Conversely, if a language clearly has separate morphemes for tense and agreement (like Icelandic), then it cannot have a fused IP since the tense and agreement features will not be visible for checking at the same time but rather one after the other (the "innermost" first, as explained in Sections 1.1 and 3.3 above). The third possibility is that the morphology gives no clear indication as to whether the tense and agreement features are associated with a single overt morpheme or not. In such a situation it would seem that the child acquiring the language would have to rely on syntactic evidence only for setting the SIP parameter.[24]

Needless to say, this paper is to be seen as a research hypothesis rather than a detailed analysis. It makes certain predictions about the relationship between overt morphology and syntax, not only between tense and agreement morphology and TP, AgrSP and IP but also more generally, and these predictions should be testable principle. An important question not addressed here is the relationship between the syntactic FCs and the semantic correlates of the ϕ-features or interpretable features associated with them (such as tense — and perhaps also under different circumstances aspect, mood, etc.).

Finally, note that under the view advocated here, it is not the case that MSc and English simply select fewer functional categories than Icelandic from the set made available by UG (see the discussion of the LDH around (4) above). The FCs in these two groups of languages are also different in their feature composition since Infl in MSc and English contains features that are distributed between AgrS and T in Icelandic.

Acknowledgements

I would like to thank Werner Abraham, Jonathan Bobaljik, Noam Chomsky, Chris Collins, Elisabet Engdahl, Samuel D. Epstein, Erich Groat, Dianne Jonas, John O'Neil, Halldór Ármann Sigurðsson, Sten Vikner and Jan-Wouter Zwart

---

Marantz 1993), his ideas should be translatable into a lexicalist account like the one assumed in this paper.

24. Interesting Germanic languages to study from this point of view include Afrikaans, Dutch and Faroese, cf., e.g., Bobaljik (1995a, b), Jonas (1995, this volume), Solà (this volume).

for valuable comments on earlier versions of this paper, or on some of the ideas presented here. The main idea in this paper was originally presented at the Spring Colloquium at the University of Maryland, College Park in May 1994, and later at the Sixth European Summer School in Logic, Language and Information in Copenhagen in August 1994; in a course on Scandinavian Linguistics at Harvard in the fall of 1994; and at the XIVth Conference of Nordic and General Linguistics in Oslo in January 1995. I would like to thank the audiences for interesting discussions and comments. All errors are, unfortunately, my own.

Participles and Bare Argument Structure

Guido Vanden Wyngaerd
NFWO/KU Brussel

The operation Merge, then, is asymmetric, projecting one of the objects to which it applies, its head becoming the label of the complex formed. There can be no non-branching projection.

— Chomsky (1995a: 398)

Abstract

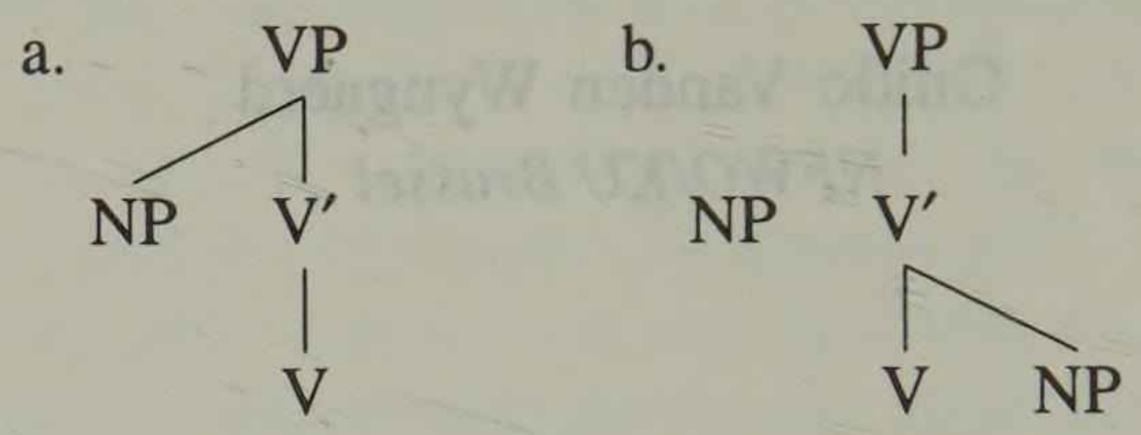
It is potentially undesirable consequence of Chomsky's Bare Phrase Structure framework that ergative and unergative verbs would share the same argument structure. The present paper proposes to address this problem by arguing that all verbs are potentially transitive. An analysis of the structure of complex verb forms (Dutch and German past participles containing an aspectual prefix *ge-* and a tense suffix *-t*) in terms of Kayne's Linear Correspondence Axiom (LCA) suggests that the prefix occurs in the complement of the verbal root, thus in effect making unergatives transitive. The structure proposed for the participle furthermore allows one to explain two familiar restrictions on verb raising, the first being its in compatibility with particle verbs, the second the so-called INFINITIVUS PRO PARTICIPIO (IPP) effect. It is argued that IPP involves the deletion of an inflectional prefix, required to bring the structure of a verbal cluster in agreement with LCA. A particular case of this phenomenon in certain Flemish dialects is discussed, where the prefix gets deleted but the suffix remains unaltered. This variety of IPP supports the hypothesis that the prefix is responsible for the IPP effect.

1. Introduction

Chomsky (1995a: 398) argues that within in the framework of Bare Phrase Structure, "there can be no non-branching projection." This is because projection

is a subcase of Merge, and Merge always applies to two elements, projecting one. A potentially undesirable consequence of this assumption is that there would be no difference in terms of argument structure between unergative and ergative VPs. Earlier theory permitted the distinction to be made in virtue of letting (unergative) V project to a nonbranching V′, as in (1a), so that the unergative argument could be a sister to V′, whereas an ergative argument was a sister of V, as depicted in (1b).

(1)



If, however, non-branching projection is excluded, the single argument is a sister of V both with ergative and unergative verbs. In order to solve this problem, Chomsky suggests that unergative verbs are really transitive, i.e., that in (1a) V′ is really branching, despite the absence of an overt internal argument. This idea he attributes to Hale and Keyser (1993), who argue that unergative verbs are derived from underlying transitive verbs through a process of noun-incorporation.[1] One thus sees that an adoption of a minimalist view on projection or on possible phrase structures has immediate consequences in terms of what are possible and impossible argument structures.

The present paper likewise proposes to establish a link between argument structure and (certain aspects of) minimalism. The minimalist tenet is the LINEAR CORRESPONDENCE AXIOM (LCA; Kayne 1994), and the link with argument structure is established in a somewhat roundabout manner. The first step involves the structure of complex verb forms, notably Dutch and German participles containing an aspectual prefix *ge-* and a tense suffix *-t*. Adopting the LCA, I shall argue in favor of a particular structure for the participle. This structure will then be shown to have important consequences for the argument

---

1. Actually, Hale and Keyser's theory leaves open another option for making a distinction between ergatives and unergatives, in that it allows subjects to occupy either a VP-internal position or a VP-external position at D-structure. If one accepts this, the conclusion that unergatives are transitive is not compelling from the point of view of bare phrase structure. As it is unclear how Hale and Keyser's alternative would translate into a Bare Phrase Structure Framework, I shall ignore it.

structure of unergative verbs, which go into the direction of confirming the conclusion established in the previous paragraph, i.e., verbs invariably have an internal argument.

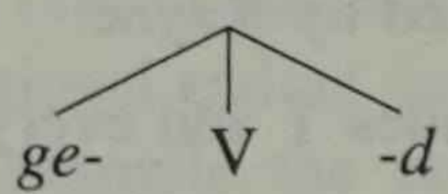
The paper is organized as follows: in the second section, I shall develop the structure of the Dutch past participle as it follows from Kayne's LCA. The third section establishes the link between participial structure and argument structure. The analysis proposed will receive support from certain restrictions on verb raising in Dutch, i.e., the clustering of a series of infinitives in sentence-final position. One of these restrictions involves particle verbs, which are incompatible with verb raising (section 4); a second restriction concerns the idiosyncratic INFINITIVUS PRO PARTICIPIO (IPP) phenomenon, which involves the replacement of a past participle by an infinitive in a verb cluster derived by verb raising. In section 5, I shall propose an analysis of the IPP phenomenon which relates it to the prefixal nature of the participle. The sixth section discusses evidence from certain Flemish dialects, in which IPP takes a surprising form.

2. The Structure of the Participle

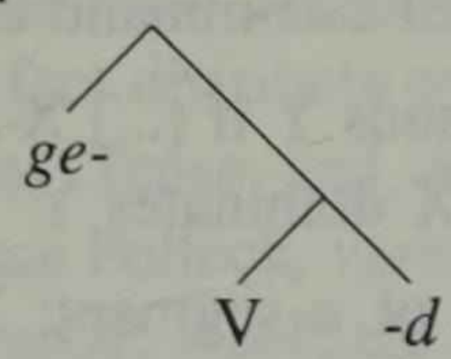
The past participle in Dutch and German is a tripartite structure, consisting of an aspectual prefix *ge*, a verbal root, and a suffix (*-d* in Dutch and *-t* in German, e.g. Du. *gewandel-d* 'walked', Ge. *gewander-t* 'roamed').[2] In principle, three possible structures could be assigned to such a participle.

(2)

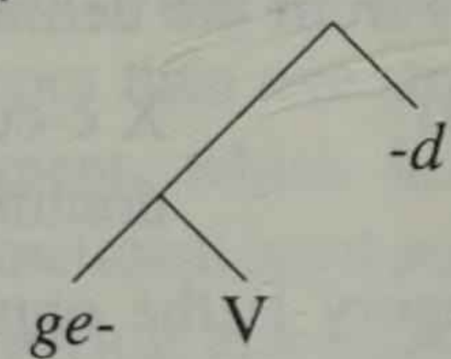
a.



b.



c.



The Linear Correspondence Axiom as proposed by Kayne (1994) has the effect of ruling out both (2a) and (2b), leaving (2c) as the only possible alternative. Let

---

2. With certain verbs, both in German and Dutch, the suffix is *-en*, e.g., Du. *bakken-bakte-gebakken*, Ge. *backen-backte-gebacken* 'bake-baked-baked'. This exhausts the possibilities, i.e., although there are irregular verbs, these always conform to the pattern *ge-V-d/en* (e.g., Du. *zoeken-zocht-gezocht* 'look for'; *moeten-moest-gemoeten* 'must'; *mogen-mocht-gemogen* 'may').

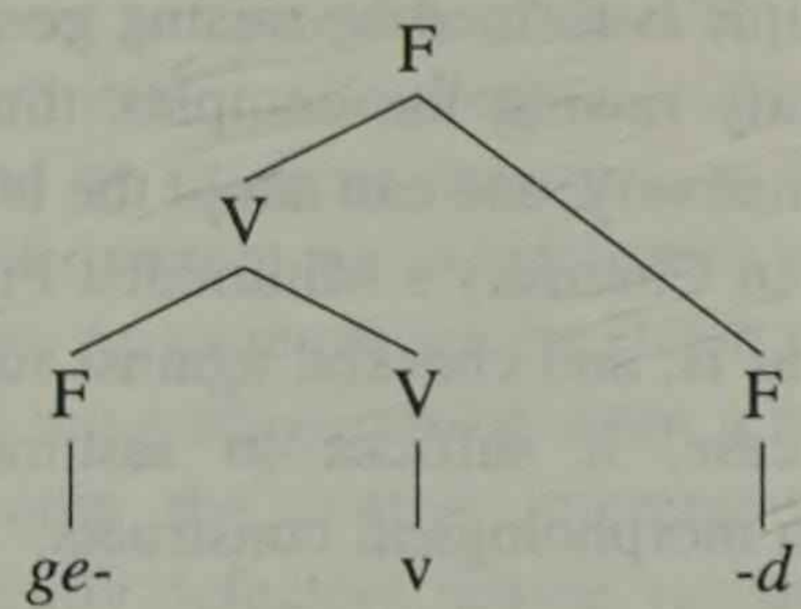
us discuss why this is the case. The idea behind Kayne's theory is, briefly put, the following: linear precedence relationship must mimic asymmetric c-command relationships, i.e., if α a-symmetrically c-commands β (that is, α c-commands β but not *vice versa*), then α must precede β. It follows that a given hierarchical representation can only be associated with one linear order. For example, any head X asymmetrically c-commands the Y head of its complement YP, so that in all languages heads precede their complements. By the same reasoning, Spec-XP universally precedes X, since Spec-XP asymmetrically c-commands X. A further condition imposed by Kayne's theory is that any linear ordering must be *total*, i.e., all terminal elements need to be ordered with respect to one another, ordering being determined by asymmetric c-command. If the ordering is not total, the representation is ruled out. As a consequence of this view, the three-way branching structure (2a) can be ruled out: the three terminals all c-command one another, i.e., no asymmetrical c-command relationships exist between them, so that no linear order is defined for these three terminals with respect to one another. The structure (2b) is also ruled out. Explaining why takes some more elaborate argumentation, however. For one thing, Kayne's theory rules out the possibility that the node immediately dominating two heads is a projection of either one of them. This is because the two heads would c-command one another, and hence be unordered. The only possibility for two heads to be sisters is for one to be adjoined to the other. This still leaves two options, viz. left adjunction and right adjunction; only the former is allowed, however, because in an adjoined structure the adjoined element asymmetrically c-commands the element adjoined to, and hence must occur to its left. This follows from the definition of c-command assumed by Kayne:

(3) X c-commands Y if [...] X excludes Y and every category that dominates X dominates Y.

A category is the opposite of a segment; X excludes Y if no segment of X dominates Y. In an adjunction structure $[_Y$ X Y], X c-commands Y, but Y does not c-command X because Y does not exclude X. As a result, X must precede Y, and right adjunction is barred. Now the label of the node immediately dominating verbal root and inflectional ending could either be identical to the label of the left-hand node (the verb) or the right-hand one (the suffix). However, in the former case, we would have a case of right adjunction, which is illicit for the reasons just discussed. It follows that V+*d* is dominated by the same label as the suffix -*d*; for convenience, I shall use the label F for the prefixal and suffixal functional elements *ge-* and -*d*; the representation of the

lower part of (2b) is therefore necessarily as follows: [$_F$ V [$_F$ *-d*]]. With this much established, we can go on to answer another question, viz. what is the label for the topmost node in (2b)? Two possibilities exist, but neither of them is possible on account of the LCA. One possibility is to assign to the top node the same label as the suffix, yielding the following: [$_F$ *ge-* [$_F$ V [$_F$ *-d*]]]. In this case, there would be two adjunctions to the suffix, which is illicit, the reason being that the two adjuncts, the prefix and the verbal root, c-command one another, and as a result they would not be ordered with respect to one another (cf. Kayne 1994: 19). The second possibility would be to assume that the top node is a projection of the prefix: [$_{ge}$ *ge-* [$_F$ V [$_F$ *-d*]]]. This would again make the resulting structure one of right adjunction, however, which is also ruled out (cf. above). Hence only the structure in (2c) remains, with only one possible labeling:

(4)

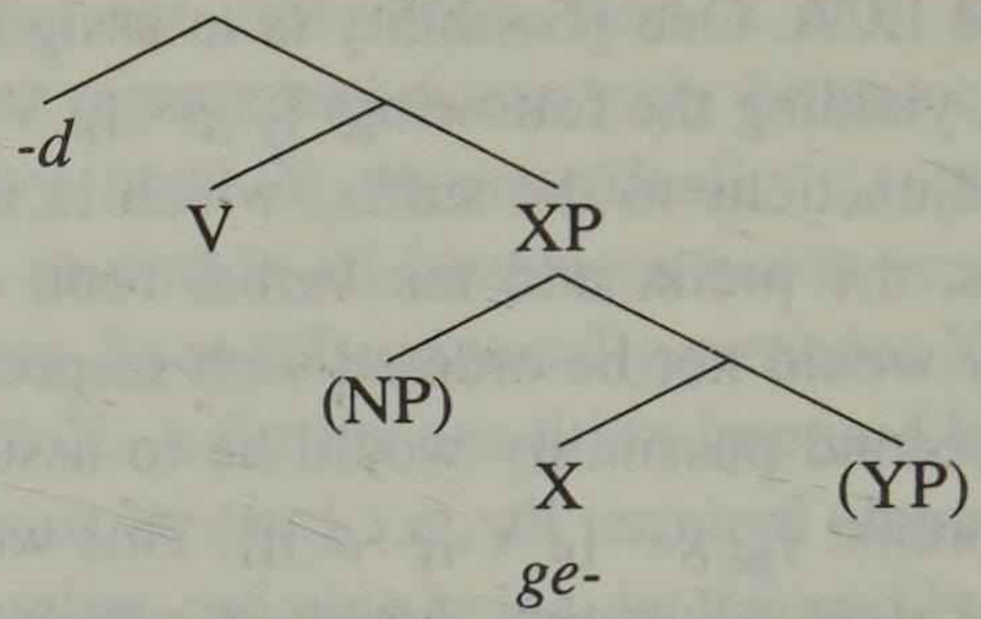


Now it has been assumed at least since Chomsky (1965) that verbal roots and the inflectional endings attached to them, despite appearances, occupy an altogether different structural position in the tree, i.e., head their own projections. Pollock (1989) has refined this view, suggesting that what appears to be a single inflectional ending may in fact distribute over more than one structural position, for example the positions Tense and Agreement, which head the respective phrases TP and AgrP. Since Pollock, various functional head positions have been proposed (e.g. AgrSP, AgrOP, AspP, etc.). Applying this logic to the Dutch and German participle with *ge*, I shall assume that both the prefix and the suffix occupy a functional head position.[3] The question now is, which are these, and where are they situated? Given the structure in (4), there is only one

3. The implicit assumption here is that *ge-* is inflectional, not derivational. Independent evidence in support of this claim is hard to come by. Still, some circumstantial evidence can be given, in that participle formation does not seem to involve a category change; see Hoekstra (1984: 163, 201) for arguments showing that the Dutch participle is indeed verbal rather than adjectival.

compatible tree, viz. one in which the *ge-* prefix occurs in the complement of the verb. In the diagram below, *ge-* occupies the head position of a complement small clause, XP, with NP and YP the optional complements of V.[4]

(5)



If the verb is transitive or ergative, the internal argument can be assumed to occupy the Spec position of the complement XP, much as in Larson's VP-shell analysis (Larson 1988). The participle is formed by raising *ge-*, left-adjoining it to the verbal root, and subsequently raising the complex thus formed to the suffix, left-adjoining it there. Alternatively, one can adopt the lexicalist checking approach to morphology assumed in Chomsky's Minimalist Program, in which case the inflected verb is inserted as is, and checked against functional features at a later stage. In the latter case, it suffices to assume, as in Kayne (1994: 38–41), that LCA applies to morphological constructs.

3. Consequences for Argument Structure

The position of the prefix *ge-* in the complement of the verb puts it in its subcategorization domain, instead of the functional domain dominating the verb, where inflectional information usually goes (e.g. T, AgrS, and AgrO all c-command the verb they belong to). As a result, all verbs, including unergative ones, are potentially transitive, in that a verb in the perfect always has a complement containing *ge-*. This conclusion is particularly interesting in the context of a minimalist theory of phrase structure, as outlined in the introduc-

---

4. Following Chomsky (1995a: 417), *ge-* would need to be a sister of V if *ge-* were the only complement of the verb (as with unergative verbs), because of the ban on non-branching projection. Kayne's LCA then does not define an order for the terminals V and *ge-*, because they c-command each other. This problem can be overcome, Chomsky argues, by moving either terminal, leaving a trace, which is erased at PF and hence need not be linearly ordered, i.e., is not subject to LCA. I shall ignore these details of implementation here.

tion: since there can be no nonbranching projection, the distinction between ergatives and unergatives resides in the presence or absence of the *external* argument, rather than the internal one, i.e., all verbs have an internal argument. Now there are still two forms that this claim can take: one could argue that all verbs have a complement regardless of their aspect, or, alternatively, that argument structure changes with perfectivity. The former position is taken by, for example, Hale & Keyser (1993). The latter position, which would relate perfective aspect to argument structure, also receives support from certain indications in the literature, though it is hard to establish hard and fast conclusions. What follows is therefore intended merely as an indication of potentially promising avenues of research, rather than as a conclusive argument one way or the other. For example, the claim that perfective aspect and argument structure correlate agrees well with the argument made in Verkuyl (1987) to the effect that perfective aspect is a property of the VP, notably the verb and its complement, rather than the verb itself.[5] Also, Postma (1994) observes that argument structure is sometimes affected by changing the aspect to the perfect. For example, verbs with defective argument structure are also defective in the perfect, i.e., they cannot appear in the perfect tense. An example are the Dutch verbs *dreigen* 'threaten' and *beloven* 'promise', which are ambiguous between a raising and a control interpretation. In the perfect, only the control interpretation remains, suggesting that the argument structurally defective raising variant cannot be put in the perfect.[6]

(6) a. *Jan dreigt/ dreigde te vallen.*
Jan threatens threatened to fall
'Jan is almost falling/almost fell.' (raising reading)
'Jan threatens/ed that he would fall.' (control reading)

---

5. One piece of evidence given by Verkuyl in support of this view is the observation made in Carlson (1977) to the effect that the determiner on the object affects the perfectivity of the VP.
(i) a. *John discovered a rabbit in his garden* (**for a whole day/in five minutes*)
b. *John discovered rabbits in his garden* (*for a whole day/*in five minutes*)
If the object is a bare plural, the aspect is nonperfective (or atelic), otherwise it is perfective.

6. A reviewer points out, correctly, that the argument structurally defective raising verbs in (6) are ergative, so that thay have an internal argument anyway. For these verbs, then, the point that every verb has a complement needs no further demonstration. Still, the correlation with aspect is suggestive, which is why the facts are brought up here.

b. *Jan belooft/ beloofde een beroemd geleerde te worden.*
Jan promises promised a famous scientist to become
'Jan will/would probably become a famous scientist.'
(raising reading)
'Jan promised to become a famous scientist.' (control reading)

(7) a. *Jan heeft gedreigd te vallen.*
Jan has threatened to fall
'Jan threatened that he would fall.' (control reading only)

b. *Jan heeft beloofd een beroemd geleerde te worden.*
Jan has promised a famous scientist to become
'Jan promised to become a famous scientist.'
(control reading only)

Similarly, Hopper & Thompson (1980: 271) observe that "if the Aspect is perfective, the interpretation — other things being equal — has properties allowing the clause to be classified as more transitive; but if the Aspect is imperfective, the clause can be shown on independent grounds to be less transitive." These remarks are admittedly somewhat vague, and not every factor that Hopper & Thompson classify as contributing to a higher level of transitivity could probably be seen as affecting argument structure, but at least some of them might. For example, Postma (1994) quotes the case of Georgian from Hopper & Thompson. In the present tense, the subject appears in the nominative, and the object in the dative. In the perfect, however, the Case pattern is different: the subject gets ergative Case, and the object nominative.

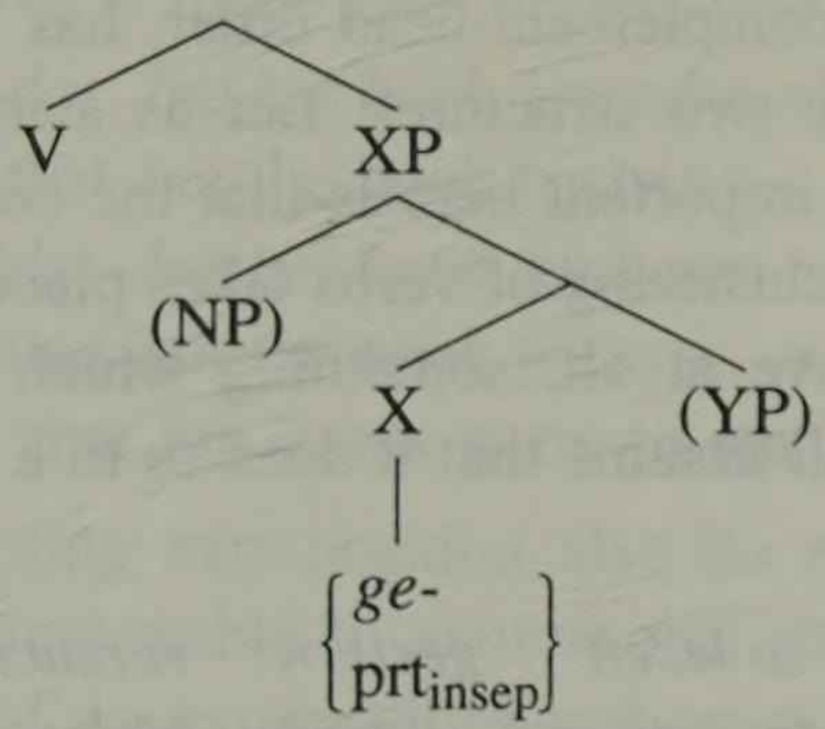
(8) a. $NP_{Nom}\ V_{Pres}\ NP_{Dat}$
b. $NP_{Erg}\ V_{Perf}\ NP_{Nom}$

The former Case pattern is antipassive, the second passive-like, i.e., an object with a structural Case like the one also assigned to (in)transitive subjects, and a subject with oblique Case. Now why would the present and perfect tenses exhibit different Case patterns? One way to think about this is that the argument structure of both tenses differs. Now it is often the case that passivisability correlates with higher transitivity (e.g., English intransitive verbs are not passivisable). The fact that the Georgian perfect, unlike the present, would display a passive-like Case-pattern could then also be attributed to the higher level of transitivity of the perfect tense in comparison with the present.

4. Particle Verbs

Dutch particle verbs provide additional support for the structure given in (5) above. Particle verbs come in two kinds: those with a separable particle and those with an inseparable particle. The latter are interesting because of the way their participle is formed: inseparable particles are incompatible with *ge-*, i.e., there is no prefix other than the particle itself in the participial form of these verbs (e.g., *over.tuigd* 'convinced', rather than **ge.over.tuigd* or **over.ge.tuigd*). This suggests that the particle and *ge-* occupy the same structural position, as indicated in (9).

(9)



Mulder (1991, 1992) independently proposes a structure that is closely analogous to (9) in a discussion of object control verbs. The evidence he provides concerns verb raising, i.e., the clustering of a series of verbs in sentence-final position as a result of successive adjunctions, as described in Evers (1975).[7] Before going on to discuss Mulder's argument, some preliminary remarks about verb raising are in order. Consider the example in (10), with the optional verb raiser *proberen* 'try'.

(10) a. *dat Jan het boek heeft proberen te lezen.*
that Jan the book has try to read

b. *dat Jan heeft geprobeerd [(om) het boek te lezen]*
that Jan has tried COMP the book to read
'that Jan tried to read the book.'

The traditional picture, as sketched by Evers, is that (10) represents the following two possibilities: starting out from an underlying OV order, i.e., [...[...[... *te*

---

7. Movement to C does not fall under the term "verb raising" as I shall be using it here, since it has quite different characteristics: it only involves the finite verb, and never more than one verb.

lezen] *proberen*] *heeft*], the embedded infinitive *te lezen* 'to read' in (10a) right-adjoins to the dominating verb *proberen* 'try', and the complex thus formed right-adjoins to the matrix verb *heeft* 'has', yielding the cluster [*heeft* [*proberen* [*te lezen*]]]; in (10b), on the other hand, the entire complement clause extraposes from a preverbal position to the right periphery of the clause, through right-adjunction to VP or IP. Verb raising and extraposition constructions differ empirically on a number of counts, having to do with transparency of the complement clause for processes like clitic climbing, negation, etc. (see Evers 1975 for details). Although I shall therefore by and large want to adopt the picture just sketched, it will be obvious that the LCA, with its ban against right adjunction and underlying complement-head order, has important consequences for the analysis of these two structures. Let us start with the extraposition structure (10b). What is important here is that the complement clause remains independent and that no clustering of verbs takes place. In so far as the complement clause needs to move at all, something which is suggested by German examples like (11), I shall assume that it does so to a Spec position on the left.

(11) *daß er das Buch zu lesen gestern versucht hat.*
that he the book to read.INF yesterday try.PART has
'that he has tried to read the book.'

One may consequently replace the term "extraposition" with the term CP-raising. Dutch does not allow the analogue of (11); in line with minimalist assumptions, I shall assume that the difference between Dutch and German resides in the locus of application of CP-raising: overtly in German, covertly in Dutch. Let us next turn to the verb raising structure (10a). By the LCA, all languages, including Dutch (Zwart 1993a), have head–complement order, i.e., are VO underlyingly. As a result, languages which have V1–V2–V3 order in clusters (VP1 highest, VP3 most deeply embedded), like Dutch, directly reflect this underlying order. A language like German, by contrast, which has V3–V2–V1 order in clusters, shows the results of verb raising overtly. Again, this difference can be seen as one between overt verb raising (German) and covert verb raising (Dutch). The application of verb raising not being signalled by surface word order in Dutch, one may look for other clues betraying its presence. I shall side with authors like Bennis & Hoekstra (1989) and Den Dikken (1988, 1989) in assuming that the Infinitivus Pro Participio or IPP effect, i.e., the appearance of an infinitive (*proberen* 'try' in (10a)) where one would expect a participle (*geprobeerd* 'tried' as in (10b)), signals the presence of

cluster formation by verb raising (see section 5 below for details). This effect remains absent in the third construction (Den Besten *et.al.* 1988), so called because it represents a mixed case, which shares with verb raising constructions the appearance of the object to the left of the verbal cluster, but, like extraposition constructions, lacks IPP.

(12) *dat Jan het boek heeft geprobeerd te lezen.*
that Jan the book has try.PART to read.INF

As noted in Bennis & Hoekstra (1989: 39n), such examples in all likelihood do not involve the creation of a cluster, however, as the participle can occur to the left of *heeft*, which is impossible in true verb raising constructions with IPP.

(13) a. *dat Jan het boek geprobeerd heeft te lezen.*
that Jan the book try.PART has to read.INF
b. **dat Jan het boek proberen wil te lezen.*
that Jan the book try.INF wants to read.INF
'that Jan wants to try to read the book.'

Verbs not permitting extraposition also do not permit the third construction. Hence it seems reasonable to assume that the third construction illustrated in (12) is a variant of the extraposition construction illustrated by (10b), rather than a variant of the verb-raising construction (10a).

Having thus outlined the main facts of verb raising, we can return to the issue at hand, which is the structure of particle verbs. Evers (1975: 40) already observed that particle verbs never trigger verb raising; this restriction is illustrated in (14a); instead, the complement clause must remain independent, as in (14b).

(14) a. **dat hij Jan het boek heeft over.tuigen te lezen.*
that he Jan the book has convince.INF to read
b. *dat hij Jan heeft over.tuigd het boek te lezen.*
that he Jan has convince.PART the book to read
'that he convinced Jan to read the book.'

Mulder's argument referred to above concerns object control verbs, but since, as he observes, these are almost all prefixal, his analysis is relevant for particle verbs in general.[8] Mulder proposes the following structure:

(15) $V\ [_{XP}\ NP\ [\phi_{X^0}\ [_{PP}\ P\ IP\ ]]]$

---

8. An exception to this generalization is *dwingen* 'force', which Mulder analyzes as covertly prefixal.

This structure must be understood as follows: object-control verbs take a small-clause complement hosting the "object" NP and the sentential complement (PP in (15)). The head of the small clause, indicated by ϕ, contains the particle; P is an infinitival complementizer, which gives the sentential complement the status of a PP. It is easy to see that this structure is virtually identical to the one proposed in (9) above.

Now since object-control verbs are almost all prefixal, they allow no verb raising, like particle verbs in general. Mulder further (1991: 307, 1992: 112) argues that the impossibility of verb raising with object-control verbs follows from his analysis, the particle blocking verb raising. This explanation presupposes that the particle cannot become a member of the verbal cluster, so that its presence will effectively block the application of verb raising through some form of minimality. If however, a particle can become a part of a verbal cluster, there is no reason why it should induce a minimality violation: verbs raising up could adjoin to the particle (thus treating it essentially as a verb), and the complex thus formed could subsequently adjoin to a higher verb. The assumption that particles cannot be a part of verbal clusters, essential to the explanation of the ban on verb raising with particle verbs, is problematic however, in a framework that assumes underlying OV order for Dutch and German, like Mulder's. This is because of the existence of examples like the following, where the lowest verb is the particle verb *op-bellen* 'phone up':

		1	2	3
(16)	*dat*	*ik hem (op)*	*had (op)*	*willen (op) bellen.*
	that	I him up	had	want phone
	'that I had wanted to phone him up.'			

In an OV framework, underlying order is V3–V2–V1 (VP1 topmost, VP3 lowest): [...[...[...*bellen*] *willen*] *had*]. The observed order in (16) is the reverse as a result of the application of verb raising, i.e., successive head-adjunction yielding the cluster *[had [willen bellen]*. But if indeed the verbs in (16) form a cluster, it is unclear how the particle could occur in the positions 2 and 3 other than by becoming a part of that cluster. And if particles can become parts of adjunction clusters, it is unexplained why they should block verb raising just in case the particle belongs to the higher verb, as in (14a) and with object-control verbs, but not if the particle happens to belong to the lowest verb, as in (16).

In the universal VO framework adopted here, however, an explanation for the inability of particle verbs to trigger verb raising is forthcoming. Assume, as in Hoekstra (1988), that particles are predicates of (resultative) small clauses.

Further assume, as in Zwart (1993b) and Koster (1995), that small-clause predicates must move to the left, landing in the specifier position of the functional projection PredP. I propose that this leftward movement to PredP is what is responsible for the positions of the particle in (16). However, as remarked at the outset of this section, not all particles can be separated from their verbal root: a distinction can be made between separable and inseparable particle verbs, which differ on a number of counts aside from separability (see Vanden Wyngaerd 1994a, 1994b for detailed discussion of the two classes of particle verbs). I assume that an inseparable particle cannot move to PredP (though, like *ge-*, it must raise to the verbal root). Now an example like (14a) contains a verb with an inseparable particle, *overtuigen* 'convince'. Hence it is necessarily the case that the particle, like *ge-*, becomes part of the verbal cluster if such a cluster is formed through verb raising. Crucially, the sentence could not have arisen through leftward movement to PredP, as that would separate the particle from its verb. The fact that the result is ungrammatical shows that verbal clusters indeed cannot contain particles. The reason for this presumably lies in their nonverbal nature, i.e., verb raising can only affect verbs and the inflectional elements that they support, not preposition-like elements. This claim is supported by German data. Recall that surface word order in Dutch clusters reflects that before the application of verb raising. As a result, particles moved to PredP can intervene between the verbs that will eventually come to form a cluster. In German, by contrast, verb raising takes place in overt syntax, so that one expects a particle to be unable to intervene between the clustered verbs. As (17) reveals, this prediction is borne out.

		1	2	3
(17)	*daß ich ihn (an) rufen (*an) können (*an) wollte.*			
	that I	him up	phone.INF can.INF	wanted
	'that I wanted to be able to phone him.'			

The particle can be stranded in PredP as in Dutch (position 1), but not in positions 2 and 3, which are necessarily cluster-internal positions.

In sum, the analysis of particle verbs as in (9) coupled with an underlying VO order for Dutch and leftward particle scrambling allows one to account for the restriction observed by Evers (1975) to the effect that particle verbs never trigger verb raising. In addition, the analysis presented here of the structure of past participles allows one to generalize over the case of particle verbs and the IPP effect, which will be the topic of the following section.

5. IPP

The previous section has shown that the particle in the tree in (9) could block raising of a verb embedded in YP. The question that naturally arises at this point is whether *ge-*, which occupies the same position, has a similar blocking effect. At a superficial level, the answer is affirmative, i.e., no infinitive can be embedded under a participle if the participle is an obligatory inducer of verb raising, such as Dutch *willen* 'want'. The example (18) testifies to this.

(18) **dat Jan het boek heeft gewild kopen.*
that Jan the book has want.PART buy.INF
'that Jan has wanted to buy the book.'

I shall come back to the structure of the cluster in (18) below; the idea that I will defend is that *ge-* is the offending factor that blocks the raising of the infinitive embedded under the participle. Accordingly, the sentence can be made grammatical by dropping the *ge-* prefix and replacing the *-d* suffix by an *-en* suffix, thus turning the participle into an infinitive, whence the name of the phenomenon, the Infinitivus Pro Participio (IPP) effect, or Double Infinitive Construction (DIC).

(19) *dat Jan het boek heeft willen kopen.*
that Jan the book has want.INF buy.INF
'that Jan has wanted to buy the book.'

The effect is found whenever the participle itself has an infinitival complement (with or without *to*) that undergoes verb raising, but not when the participle has a nonverbal complement:

(20) *dat Jan het zo heeft gewild / gewild heeft.*
that Jan it that way has want.PART want.PART has
'that Jan wanted it that way.'

As already shown above, the participle also remains with extraposed infinitives: thus in (10b) the sentential complement remains independent, and there is consequently no IPP; this is in contrast with (10a) where cluster formation occurs, with concurrent IPP.

The idea that the IPP effect relates to the formation of a cluster is one that the present account shares with the ones proposed by Bennis & Hoekstra (1989) and Den Dikken (1988, 1989). A question that has to date not received a satisfactory answer, however, is why *ge-*, or a participle in general, should block cluster formation. The Kaynean framework of antisymmetry allows one to

explain this curious fact. In order to show this, however, some comments on the structure of verb clusters are in order. Verb clusters cannot involve the simple adjunction of one verb to another:

(21) $[_V\ V\ V\ ]$

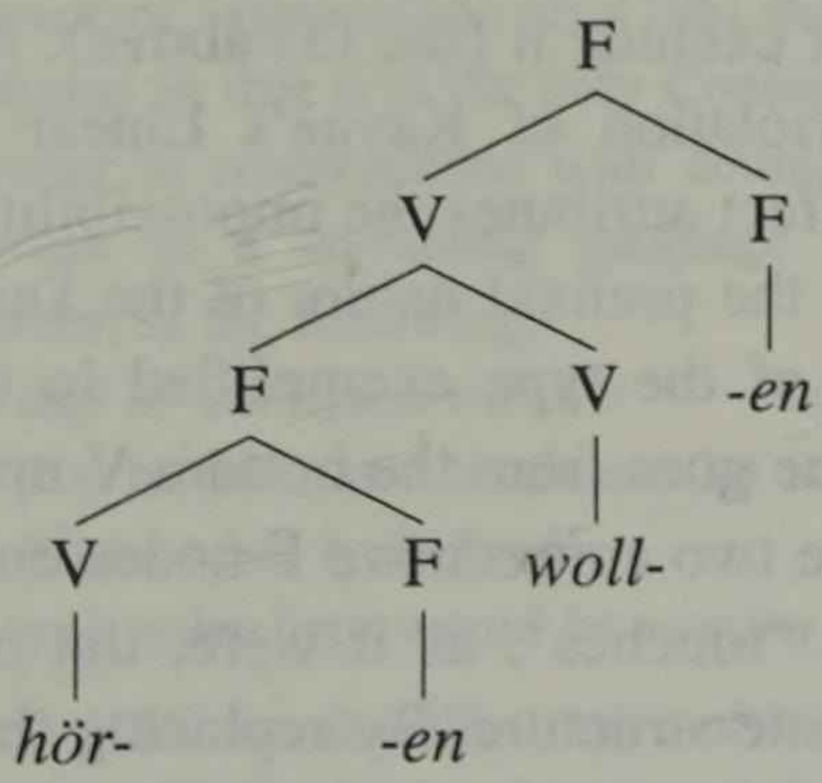
The reason for this is that the two terminal nodes would be too symmetrical, i.e., each would c-command the other so that no linear order would be defined for them. This problem can be solved by assuming that verbs, both finite and infinitival, have internal structure, i.e., that they consist of a lexical and an inflectional head, as follows:

(22) $[_\alpha\ V\ F\ ]$

As we saw above, F is the only possible choice for α in (22), as otherwise the structure would be one of right adjunction. Put in movement terms, an infinitive must be analyzed as resulting from left-adjoining a verbal root to an inflectional head F.[9] Now how can two infinitives be combined in an adjunction structure? Again, there is only one possibility, which is left adjunction of the lower verb+F complex to the verbal root of a higher verb, yielding the structure in (23); an example with the cluster *hören wollen* 'hear want' is given in (24):

(23) $[_F\ [_V\ [_F\ V\ F\ _F]\ V\ _V]\ F\ _F]$

(24)



More complex clusters are built on the same pattern. One thus sees that, while LCA on the whole generates uniformly right-branching trees, it generates *left*-branching subtrees in the domain of multiple (head)-adjunction.

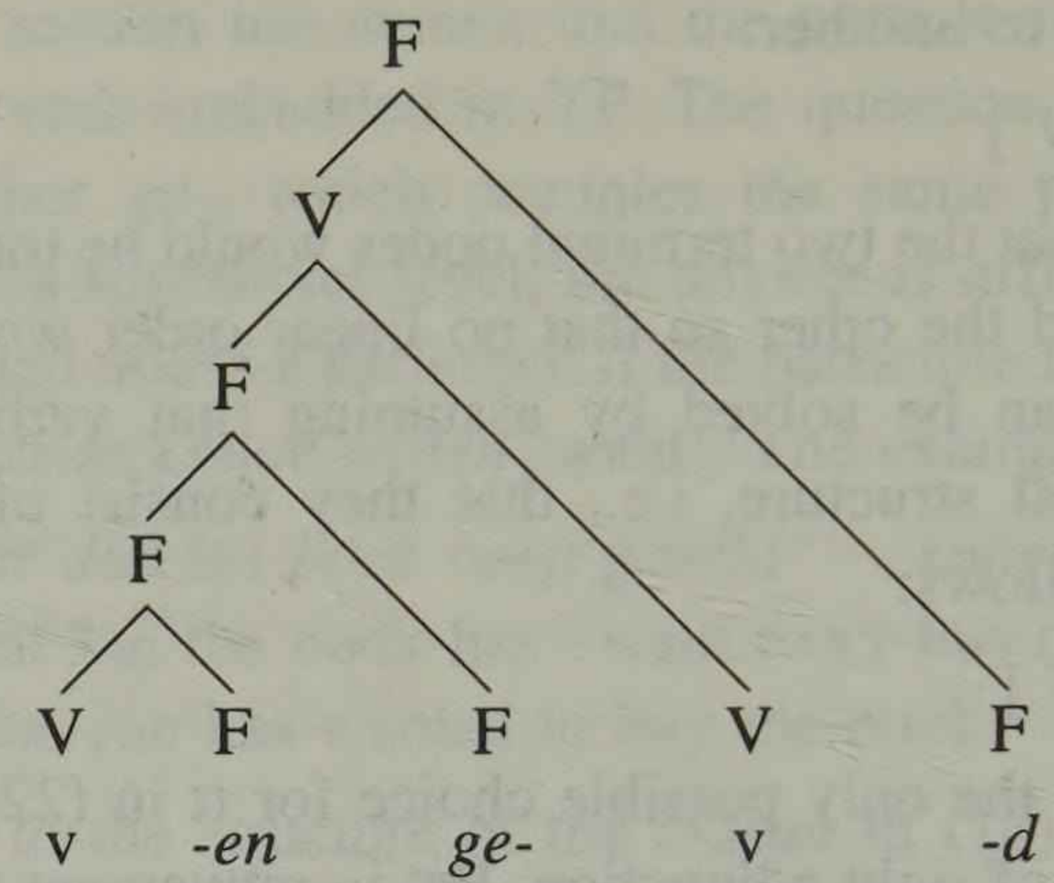
Now consider what happens if an infinitive adjoins to a participle. Given

---

9. In a checking approach to inflectional morphology, the entire inflected form would be moved. This perspective is compatible with the text argument, which concerns the internal structure of inflected verbs and verbal clusters.

the ban on right adjunction, the infinitive must adjoin to the left periphery of the participle, concretely to *ge-*. The resulting structure would be as follows:

(25)



Now what is wrong with this? I should like to suggest that the infinitival suffix and the participial prefix, being inflectional elements, are both dominated by a node with the same categorical label (F in (25)).[10] It then follows that neither c-commands the other, since neither excludes the other; in particular, a segment of the infinitival suffix dominates the participial prefix and hence does not exclude it; conversely, a segment of the participial prefix dominates the infinitival suffix and therefore does not exclude it (see (3) above). As a result, there is no total linear ordering, in violation of Kayne's Linear Correspondence Axiom. This account of the IPP effect attributes the impossibility for a participle to be a link in a verbal cluster to the prefixal nature of the Dutch and German participle. In legitimate clusters, of the type exemplified in (24), there is an alternation of V and F nodes as one goes from the bottom V up to the top node. In the configuration (25), there are two consecutive F nodes on this path, at the point where the infinitival suffix "touches", as it were, the participial prefix. This is what makes it an illegitimate structure. By replacing the participle with

---

10. This assumption is potentially controversial. It cannot be the case, for example, that all inflection bears the same categorical label F, as this would rule out attested cases of verbal roots with both tense and agreement suffixes attached to them on the same side. In order to allow these, one must assume that tense and agreement suffixes are dominated by distinct categorial labels. Another case that would require distinct labels is that of multiple clitics, in so far as they would be adjoined to one another, a possibility entertained by Kayne (1994: 20): in order to allow this, the two clitics must be dominated by distinct categorial labels. I shall leave this matter aside here.

an infinitive, the offending functional prefix is removed and the V–F alternation is restored.

Taking a cross-linguistic perspective on the matter, one is struck by the rareness of the IPP phenomenon and the restricted distribution of prefixes in the participle. Moreover, the two phenomena are found to occur, i.e., IPP only occurs in languages with a prefixal participle. Thus Lange (1981: 64) notes that the Double Infinitive Construction (i.e., IPP) is found in all Germanic languages in which the participle is built with the prefix *ge-*, and is absent where the participle is merely suffixal.[11] By way of illustration, consider the following data from Frisian and low German, respectively (from Den Besten & Edmondson 1983: 157), which lack *ge-* and do not have IPP either.[12]

---

11. An exception to this generalization is Yiddish, which has the *ge-* prefix in the participle but no IPP (example from Jacobs, Prince & Van der Auwera 1994: 414).

(i) *hot zi nit gekent farshteyn ver mit vemen es schlogt zikh.*
has she not can.PART understand who with whom it hits self
'so she couldn't understand who was fighting with whom'

Here the participle *gekent* 'could' embeds an infinitive *farshteyn* 'understand', without triggering the IPP effect. Possibly, Yiddish does not permit any verb raising at all, and this is an instance of extraposition (cf. (10b) above). A reviewer points out that Yiddish is also exceptional in that it is the only Continental West Germanic language that has the V2–V3 order in constructions with double participles, which would support the view of Yiddish as a no-raising language. A German example, with the usual V3–V2–V1 order, is the following:

(ii) *daß er entlassen worden ist*
that he fire.PART become.PART is
'that he has been fired'

The Yiddish type order here would be *worden entlassen* 'been fired'.

12. The examples in (26) are also illustrative of a correlation, observed by Weijnen (1966: 320), between word order in the cluster and the existence of IPP. The correlation goes as follows: V1–V2–V3 order correlates with IPP, and V3–V2–V1 order with absence of IPP. The former situation holds in standard Dutch, whereas Frisian, low German, and the Groningen dialect instantiate the latter situation. Both possibilities are attested in Standard German:

(i) a. *daß er das Buch lesen gekonnt hat.*
that he the book read can.PART has
b. *daß er das Buch hat lesen können.*
that he the book has read can.INF
'that he has been able to read the book.'

In (i-a) the participle occurs to the left of HAVE (V3–V2–V1), and there is no IPP; in

(26) a. *dat er it boek lêze ken.t hat.*
that he the book read can.PART has
b. *dat he dat book lesen kunn.t het.*
that he the book read can.PART has
'that he has been able to read the book.'

The prefix *ge-* does not occur in all Dutch dialects either: Van Loey (1964: 137) notes that it is absent in the north of North-Holland, Groningen, Drente, the east of Twente, and the east of the Achterhoek (cf. also Weijnen 1966: 286). In the Groningen dialect at least, IPP likewise remains absent (Schuringa 1923: 105):

(27) *Ik heb 't zeg'n heurd.*
I have it say.INF heard.PRT
'I heard someone say it.'

Still, one would like to explain why *ge-* has such a limited distribution cross-linguistically. A possibility that one could entertain is that, despite appearances, its distribution is not so limited after all; that is, languages not overtly displaying *ge-* could have empty or covert *ge-* in participles. An argument to this effect may be drawn from Bennis & Hoekstra (1989), who have argued that a participle is an impossible link in a verb cluster. While this agrees well with the claim defended here to the effect that the participial prefix *ge-* blocks cluster formation, not all the evidence discussed by Bennis & Hoekstra involves prefixal participles. For the account given above to extend to these cases, one would

---

(i-b) the participle appears to the right of HAVE (V1–V3–V2), with IPP. In an OV framework, one could propose to account for this correlation by assuming that cluster formation does not apply in the case of V3–V2–V1 order, whence the absence of IPP. Such an account is presented by Den Dikken (1988, 1989). It will be clear that this account crucially relies on the assumption that V3–V2–V1 order is the underlying order, in other words, that VP is head final. If VP is always head initial underlyingly, however, this account no longer stands. Assuming the validity of the correlation observed by Lange between IPP and *ge-*, the question can be phrased as follows: why does the absence of *ge-* tend to correlate with V3–V2–V1 order, and why does presence of *ge-* tend to correlate with V1–V2–V3 order? I am grateful to Erik Hoekstra for helpful discussion on this point.

In a VO framework, the further question arises how (i-a) must be analyzed: the order of the verbs indicates rearrangement, yet there is no IPP. I shall assume that no verb raising has taken place here; instead, the entire complement clause has been raised to the left (see (11) above).

need to assume that in such cases the participle is covertly prefixal as well. Relevant examples from English and Swedish, respectively, are given in (28).

(28) a. **Kaatje was heard sing a song.*[13]

b. **Peter blev hörd sjunga en sang.*

Peter was hear.PART sing.INF a song

These would be cases analogous to Dutch and German IPP, in so far as they illustrate the impossibility of a participle to occur in a verbal cluster. This prefix would block the raising of the embedded infinitive to the matrix verb. As far as Romance is concerned, Kayne (1993: 12) argues that unergatives are covert transitives; his argument concerns certain patterns of past participle agreement and auxiliary selection in Romance. Notably, unergative verbs need an internal argument in Spec-AgrOP to prevent the subject from landing there and triggering (illicit) participial agreement. This analysis could be assimilated to the present proposal, by assuming that the covert internal argument with unergatives is the Romance analogue of the Germanic participial prefix *ge-*. In sum, such cases as (28) can be accounted for by assuming a covert prefix in the participle, as can the Romance facts discussed by Kayne.

6. A Prefixless Participle

Some particularly interesting evidence in support of the account of IPP presented here can be found in certain dialects of Flemish. These dialects have prefixal participles and behave like Dutch in terms of displaying IPP, but deviate from the standard Dutch pattern in one particular construction. The construction is exemplified by the (Standard Dutch) sentence in (29):

---

13. With respect to English three verb clusters like the one in (i), there is an independent reason why these are ruled out.

(i) **John has could read the book.*

It seems reasonable to assume that the reason for the deviance of (i) reduces to the fact that the English modals are defective in the nonfinite paradigm (i.e., they lack infinitive, present and past participle). This is confirmed by the fact that, even if the participle does not embed another verb, (i) remains bad (**John has could*), whereas this does yield an acceptable result in Dutch and German (*Jan heeft dat gekund* 'Jan has could that').

The grammatical variant of (28a) with *to* could be analyzed as a case where the infinitival complement has been extraposed, analogous to (10b).

(29) *Jan is voetball.en.*
Jan is football.INF
'Jan has gone out to play football.'

If this construction is put into the perfect tense, one expects the participle to be replaced by the infinitive as a case of IPP, but the result is equally ungrammatical.

(30) **Jan is geweest/ zijn voetballen.*
Jan is be.PART be.INF football.INF

Postma (1993) argues that this is because the auxiliary *is*, which needs to appear in the perfect of *zijn* in Dutch, is incompatible with the *zijn*-infinitive. As Postma observes, this case instantiates a more general restriction against BE+BE, which can be observed in a wide variety of languages:

(31) **est été* (French)
**is been* (English)
**is gezijnd* (Dutch)
**es sido* (Spanish)

Postma attributes this restriction to Condition B of the Binding Theory. Faced with this restriction, languages have two options: they can turn to HAVE as the auxiliary of the perfect (e.g., English, French, Spanish), or use a suppletive form for the participle (e.g., German, Italian). The latter strategy is also employed in standard Dutch, where the suppletive infinitive *wezen* can be used instead of *zijn*:

(32) *Jan is wez.en voetball.en.*
Jan is be.INF football.INF
'Jan has been out to play football.'

However, the Flemish dialects do not have this alternative form of *zijn* (Van Haeringen 1954). What happens instead is that the participle, which is normally prefixal, is stripped of its prefix:

(33) *Jan is weest voetball.en.*[14]
Jan is be.PART football.INF

---

14. The geographical distribution of the *weest+infinitive* construction is studied in De Schutter (1974): it occurs in West Flanders, the western half of East Flanders, the southern half of the Antwerp province, and sporadically in Flemish Brabant and East Flanders. Other dialects of Flemish display a *weesten* alternant to the *wezen* infinitive in this construction.

This is despite the fact that prefixless participles otherwise do not occur in the Flemish dialects. The situation with *zijn* can be summarized as follows:

(34) *geweest* → **zijn* → *wezen*
→ *weest*

That is, the participle cannot be replaced bij *zijn* because of the restriction against BE+BE, in which case it is replaced by *wezen*, but if *wezen* is unavailable as well, *weest* occurs.[15] In this way, the structure of the verbal cluster is brought in line with the principles following from the Linear Correspondence Axiom. This case then constitutes rather striking confirmation of our claim that the prefixal nature of the participle, besides the fact of cluster formation, is responsible for the IPP effect.

7. Conclusion

It is a potentially undesirable consequence of Chomsky's Bare Phrase Structure framework, and of the ban on nonbranching projection in particular, that ergatives and unergatives would come to share the same argument structure. Hale and Keyser (1993) develop a theory of argument structure that eliminates this undesirable consequence, by assuming that unergatives are underlyingly transitive. Approaching the matter form a different angle, I have presented some evidence that goes in the same direction: an analysis of the Dutch and German prefixal participles along Kaynean lines suggests that the prefix occurs in the complement of the verbal root, thus in effect making unergatives transitive. The structure proposed for the participle furthermore allowed the development of an explanation for certain restrictions on verb raising, such as its incompatibility

---

15. One would expect that those dialects that form the perfect of *zijn* with HAVE would allow the *zijn* infinitive in this construction. One such dialect is the Gent dialect, but as Wim de Geest (personal communication) informs me, the prediction is not borne out.

(i) *Jan heeft *zijn/weeste vissen.*
Jan has be.INF/be.SUPPL angle.INF
'Jan has been out to angle.'

While the absence of *heeft zijn vissen* does not undermine the validity of Postma's generalization that the infinitive *zijn* is incompatible with the auxiliary *zijn*, (i) suggests that this cannot be the only motivation for shifting from the *zijn* to the *wezen* root. Another problem is the existence of dialects, reported by De Schutter (1974), which allow *geweest vissen* or *vissen geweest* in this construction; possibly, the latter do not form a cluster, so that IPP remains absent.

with particle verbs and *ge-* participles, the latter phenomenon being known as the IPP effect. A particular case of IPP was discussed, which occurs in certain Flemish dialects, and which involves the mere deletion of the prefix, leaving the affix intact. This phenomenon was argued to support the claim that the prefix of the participle is responsible for the IPP effect.

Acknowledgements

The research in this paper was made possible thanks to a postdoctoral research grant to the author from the National fund for Scientific Research (Belgium), whose support is gratefully acknowledged. This paper has benefitted from the detailed comments by Teun Hoekstra and two anonymous reviewers. Needless to say, I alone assume responsibility for its shortcomings.

"Shortest Move" versus "Fewest Steps"

C. Jan-Wouter Zwart
Groningen University

I will have to put aside a careful development here, but it is intuitively clear how certain basic aspects will enter. Take the phenomena of Superiority (..) and of Relativized Minimality (..). Looking at these phenomena in terms of economy considerations, it is clear that in all the "bad" cases, some element has failed to make "the shortest move".

— Chomsky (1993: 14)

Abstract

This article argues that economy of derivation does not contain a requirement that movement steps be as short as possible. Restrictions on head movement, superraising, and *wh*-movement follow from well-established principles and mechanisms made explicit in recent minimalist work. It is furthermore argued that Chomsky's FORM CHAIN mechanism must be decomposed into a "trace" insertion mechanism and a (long distance) movement mechanism, where the intermediate "traces" are inserted before the movement takes place, in accordance with strict cyclicity. This resolves the conflict between the "fewest steps" requirement and the "shortest move" requirement in Chomsky (1993).

1. Introduction

In much recent work in generative grammar, derivations are considered to be subject to principles of economy. In Chomsky (1993), economy of derivation is implemented in at least two ways: derivations should involve the shortest possible movements and the fewest possible steps. As Chomsky notes, these two

requirements appear to be contradictory. He proposes a FORM CHAIN mechanism to resolve the contradiction.

In this paper, I will argue that the Form Chain mechanism should be described as follows. Form Chain consists of two processes, one of which generates empty *wh*-elements in the intermediate specifier positions required for chain formation over a longer distance. The other process consists in movement *across* these intermediate *wh*-elements, in compliance with the fewest steps requirement, but in violation of the shortest move requirement. This implies that the shortest move requirement does not apply in long distance movement processes.

In the remainder of this paper, I will argue that the shortest move requirement is equally redundant in the analysis of the other phenomena that seem to call for it, head movement and superraising. My conclusion will be that economy of derivation consists in the single requirement that the number of movement steps should be as small as possible.

2. Economy

Let us start by assuming that subjects are generated inside VP, the VP constituting a kind of kernel sentence consisting of a subject, a verb, and a complement of the verb (an object or a subject–predicate combination, referred to as SMALL CLAUSE). Let us also assume that the elements of a kernel sentence have to be formally licensed in well-defined syntactic configurations, called functional projections. Syntax then consists in moving the elements of a kernel sentence to designated positions in functional projections. The output of the syntactic component is an interface representation serving as the input to other cognitive systems, such as those involved in speech processing and interpretation.

As a minimalist principle, Chomsky (1993) assumes that the interface representations should be pure and simple, stripped of all features that are not relevant to the cognitive systems they provide input for. This he calls ECONOMY OF REPRESENTATION, summarized in (1):[1]

(1) *Economy of representation*:
Use as few symbols as possible in the output of a derivation.

---

1. Economy of representation ultimately reduces to the principle of FULL INTERPRETATION, formulated in (36) below.

In addition to (1), Chomsky proposes a second minimalist principle, stating that interface representations should be arrived at in the most economical way. This paper discusses the proper formulation of this second principle, called ECONOMY OF DERIVATION. I will argue for the following formulation:

(2) *Economy of derivation*:
Use as few steps as possible in deriving an output representation.

(2) is a standard feature of the minimalist program of Chomsky (1993). Chomsky argues that derivations are governed by principles summarized here under the label *inertness*:

(3) *Procrastinate*:
Move as late as possible.

(4) *Greed*:
Move α only if movement contributes to licensing of α.

(3) and (4) can be grouped together as in (5):[2]

(5) *Inertness*
Move as little as possible.

(5) and (2) are equivalent.

The formulation of economy of derivation in (2) is more interesting for what it leaves out than for what it contains. In particular, (2) makes no reference to the *length* of the steps involved in a derivation. According to conventional wisdom, short steps are more economical than long steps. Thus, it has been proposed that economy of derivation contains (6) in addition to (2):

(6) *Economy of derivation part 2*:
In deriving a representation, make the shortest possible movements.

(6) underlies the concept of *minimality* (Chomsky 1986b; Rizzi 1990), paraphrased in (7):

(7) *Minimality*:
Don't move α across a place where α could have landed.

(6) also plays a major part in Chomsky (1993).

---

2. The principle PROCRASTINATE reduces to INERTNESS, if the latter is considered to apply at each point in the derivation.

3. Equidistance

To illustrate the workings of (6) in Chomsky (1993), consider the derivation of a simple sentence consisting of a subject, a verb, and an object:

(8) *John loves Mary.*

Let us assume that all elements in (8) have been moved out of their kernel sentence positions to their licensing positions.[3] Let us also assume that the architecture of the functional domain is as in (9), where subjects (SUBJ) are licensed in the Spec position in AgrSP, objects (OBJ) in the Spec position in AgrOP, and the finite verb (V) in the head position in TP, and where VP equals the kernel sentence:

(9) $[_{AgrSP}$ Spec AgrS $[_{TP}$ Spec T $[_{AgrOP}$ Spec AgrO $[_{VP}$ SUBJ [V OBJ]]]]]

In the syntactic derivation of (8), the object has to move from the position indicated by OBJ in (9) to the Spec position in AgrOP. In doing so, the object crosses the position indicated by SUBJ in (9). Assuming this to be a position where the object could have landed, movement of OBJ to the Spec position in AgrOP violates the shortest steps requirement (6).

A strict application of (6), then, makes all object movement impossible in the presence of a subject. To remedy this problem, Chomsky proposes that overt movement of the verb to the head position of AgrOP makes the position indicated by SUBJ in (9) and the Spec position of AgrOP "equidistant" from the position indicated by OBJ. Spec-AgrOP and SUBJ being equidistant from OBJ, movement of the object to Spec-AgrOP is in keeping with (6). Verb movement, then, is the way to make the VP transparent.

As Chomsky notes, this analysis makes the prediction that overt object movement never occurs without overt verb movement to the head of AgrOP. But this is clearly wrong for languages like German and Dutch, as can be concluded from Vikner 1991, Section 4.2.5. In German and Dutch, objects always leave the VP, which must be described as movement to Spec-AgrOP in a minimalist approach (Vanden Wyngaerd 1989a; Zwart 1993b).[4] Since the object and the

---

3. It is irrelevant whether this is the correct analysis of English sentences.

4. The argument in the texts abstracts away from possible additional $\overline{A}$-movement of the object, as argued by Vanden Wyngaerd (1989a) and Mahajan (1990) to account for the observation that object movement in Germanic displays both A-movement properties and $\overline{A}$-movement properties.

verb are not adjacent, the verb cannot have moved to the head of AgrOP in these constructions:

(10) *..daß Peter das Buch gestern nicht gekauft hat.*
that Peter the book yesterday not bought has

The only way to align the Germanic object movement facts with Chomsky's equidistance principle is to assume that the head of AgrOP is located to the right of the VP, an option excluded in minimalist work such as Zwart (1993a, 1993b, 1994a), and Kayne (1994).

The generalization that object movement is dependent on verb movement is based on the position of objects in the Scandinavian languages (Holmberg 1986). For example, in Icelandic the object moves to the right of the negative adverb *ekki*, but not if the verb selecting the object is a participle (PART):

(11) a. *..að Jón keypti bókina ekki.*
that John bought the book not
b. *..að Jón hefur bókina ekki keypt.*
that John has the book not bought-PART

It is assumed that the finite verb *keypti* in (11a) moves to the left, opening up the possibility for the object *bókina* to move as well, whereas the participle *keypt* in (11b) remains in its basic position inside VP, forcing the object to remain inside the VP as well. The equidistance principle of Chomsky (1993) captures this state of affairs in a straightforward way. Without verb movement, the position indicated by SUBJ in (9) and the Spec-AgrOP position would not be equidistant to the object, blocking object movement by virtue of the shortest move condition.

However, this empirical argument in support of the shortest move condition faces a conceptual problem. According to the Procrastinate principle (3), covert movement (i.e., after the Spell-out point in the derivation) is preferred over overt movement. Chomsky (1993) describes violations of Procrastination in terms of the strength of the features represented in the functional heads. If these are *strong*, they are visible (but not interpretable) at the representation interfacing with the acoustic-perceptual component of the cognitive system (PF), and must be eliminated before the derivation reaches that stage. Therefore, strong features trigger overt movement, in violation of Procrastinate. The movement is nevertheless inevitable, as without it, economy of representation (1) would be violated.

(11a) suggests that the feature represented in AgrO triggering object movement must be characterized as strong. If so, the absence of object

movement in (11b) should lead to a violation of economy of representation, as the relevant features in AgrO are not eliminated in overt syntax.[5]

We must therefore conclude that the equidistance principle is problematic. In what follows, I will argue that the shortest move requirement is not a part of economy of derivation. Since the equidistance principle was prompted by the shortest move requirement, abolishing the latter removes the need for the former, thereby avoiding the problems it poses.

4. Shortest Move Phenomena

There are several phenomena for which the shortest move requirement appears to be relevant. These include head movement, superraising, and wh-movement. I will discuss these phenomena one by one, arguing that the shortest move requirement is irrelevant for the analysis of their properties.[6]

4.1. *Head Movement*

The history of the *Head Movement Constraint* (Travis 1984: 131) is instructive for our purpose:

(12) *Head Movement Constraint*:
An X^0 may only move into the Y^0 which properly governs it.

This constraint bars the derivation of (13a) from (13b), where *kiss* and *will* are both X^0s (heads), and one of these heads must move to C:

---

5. See Zwart (1994b) for more extensive discussion of this problem. A way out might be to stipulate that features can be optionally strong. The fact that the object in (11a) may appear both to the right and to the left of the negative adverb *ekki* could be taken to support this move. However, adverbial elements in Germanic in general do not appear to occupy a single fixed position. It is not clear that the variation in the order of object and (negative) adverb in (11a) cannot be described in terms of freedom of placement of the adverb, rather than in terms of optional movement. More generally, it is not clear that allowing features to be optionally strong is necessarily a part of the minimalist approach.

6. Chomsky (1993: 14) also mentions superiority phenomena as illustrating the need for a shortest move requirement. I will refrain from discussing these here, partly because phenomena from other languages (e.g. Dutch) suggest that the shortest move condition is not an inviolable condition in the domain of superiority phenomena.

(13) a. **Who have John will kissed?*
b. *Who* C [*John will* [*have* [*kissed*]]]
c. *Who will John have kissed?*

As (13c) shows, only the higher auxiliary *will* may move to C.

From the outset it has been clear that the Head Movement Constraint is part of a principle with a larger scope, the *Empty Category Principle* (see Travis 1984: 133, Chomsky 1986b):

(14) *Empty Category Principle:*
Empty categories must be properly governed.

A trace is *properly governed* if it is governed by its antecedent.[7] In (13b), movement of *have* to C would leave a trace which is not properly governed by its antecedent (*have*), since the trace would be separated from its antecedent by two maximal projections functioning as *barriers* for government (indicated by the brackets in 13b).[8]

Chomsky (1991) argues that the Head Movement Constraint, if it is reducible to the Empty Category Principle, can be dismissed "as a descriptive artifact, valid only insofar as it does in fact reduce to the ECP." Importantly, Chomsky (1991:429), continuing where Chomsky (1986b:88) left off, considers the ECP a *condition on chains*.

This suggests that in (13) the process of *deriving* (13a) from (13b) is legitimate, whereas the resulting *representation* is illegitimate. The phenomena associated with the Head Movement Constraint thus traditionally fall in the domain of conditions on representations.

In Chomsky (1993), government does not play a role, and hence conditions on head movement cannot be reduced to (12) or (14). It appears that here, the shortest move requirement of economy of derivation becomes crucial. It takes a longer step to derive (13a) from (13b) than it does to derive (13c) from (13b).

However, closer scrutiny indicates that (13a) can be excluded in a more minimalist way. The question that has to be asked first is: What is the trigger for verb movement to C in *wh*-constructions? If this trigger does not apply to infinitives, (13a) will never be derived because of the inertness principle.

There is ample evidence that verb movement to C in Germanic is closely linked to tense. Consider the following facts from Dutch:

---

7. I ignore θ-government for expository purposes, cf. Chomsky (1986b:69).

8. (This is the analysis in Chomsky (1986b, Chapt. 11).

(15) *Koopt Jan een huis?*
buys John a house
'Is John buying a house?'

(16) a. *Jan een huis kopen?*
John a house buy-INF
'John buy a house?'

b. **Kopen Jan een huis?*
buy-INF John a house

Assuming that the structure of *yes/no* questions matches that of *wh*-questions, (15) and (16) are comparable to (13). We may consider the counterpart to the *wh*-word in (13) to be empty in (15) and (16). This suggests that the verb movement in (15), as in (13c), targets C. As can be seen in (16), such verb movement takes place only when the verb is finite.

In terms of Chomsky (1993), we may suppose that C hosts a tense feature, comparable to the V-features of AgrS etc., which must be checked by moving T(ense) to C (cf. Wilder & Ćavar 1993). This triggers movement of the finite auxiliary in (13).

The proper test case for the Head Movement Constraint, then, contains a choice of two finite verbs, only the higher of which may be moved to C:

(17) a. *John did not think I could help someone*
b. *Who did John t not think I could help?*
c. **Who could John did not think I* t *help?*

(17c) is correctly excluded by the Head Movement Constraint and the shortest move requirement. However, (17c) can also be excluded on standard minimalist assumptions, once we assume that C hosts a tense feature, to be eliminated by T.

T itself hosts features which must be eliminated by the verb, and conversely, the tense features of the verb must be checked by the corresponding features in T. So, V and T are mutually dependent, as are T and C.

The sentences in (17) contain two finite clauses, each containing a C, a T, and a finite V. In (17a) and (17b), the finite verb *did* links up with the matrix T, and, ultimately, with the matrix C. The finite verb *could* links up with the embedded T and C. In (17c), however, the embedded verb *could* links up with the matrix C. There are various derivations possible to yield this result, but all of these derivations have in common that the proper linking up of V, T, and C is not established. For instance, *could* could move to the embedded T first and then on to the matrix C. This would have the result that the embedded C is not linked up with T, and that the matrix verb *did* is not linked up with the matrix

T and C. Or *could* could skip the embedded T and move to the matrix C via the matrix T, but this would again rob the matrix verb *did* of its licensing position T, and it would leave the tense feature in the embedded C unchecked. Finally, movement of *could* via the embedded C would be blocked, as *could* and the embedded T would have all their features checked in C, and there would be no trigger for further movement (by the principle of Greed [4]).

Hence, the derivation yielding (13c) would always leave certain features unchecked, ultimately leading to a violation of economy of representation (1).

More generally, the observation that heads move stepwise rather than long distance is explained by the feature checking requirements of the minimalist approach. If a functional head Y contains a feature to be checked by the verb X, and X moves to a functional head Z, crossing Y, the feature in Y will remain unchecked, and the derivation will not converge. This covers the core cases of the Head Movement Constraint. Hence, for deriving stepwise head movement, no special constraint needs to be formulated.[9]

Taking (13) and (17) to be representative of Head Movement Constraint phenomena, the minimalist analysis of them appears to make the shortest move requirement redundant for the proper understanding of this class of facts.

4.2. *Superraising*

The shortest move requirement likewise appears to play a role in barring *superraising* phenomena:

(18) a. **John seems is likely to win.*
b. **John seems it is likely to win.*

The sentences in (18) are derived from more basic representations in which *John* is the subject of *win*, generated inside the VP as previously assumed. As (19) shows, the subject position (Spec-AgrS) of the embedded clause is a legitimate target for subject movement:

(19) *It seems John is likely to win.*

It seems, then, that the sentences in (18) are derived by moving *John* across a legitimate target for subject movement, in violation of the shortest move

---

9. The reduction of the Head Movement Constraint to general requirements of feature checking predicts that if no features in a head Y need to be checked by the verb X, X may move to Z crossing Y. In Zwart (1993b) I discuss several of these cases (see also Ouhalla 1989).

requirement of economy of derivation (Chomsky 1993: 14).

However, it is immediately obvious that (18a), at least, is excluded on standard minimalist assumptions of movement and feature checking. As discussed in section 3, all movement operations are assumed to be triggered by feature checking requirements. Let us assume that in (19), the subject, *John*, moves to the specifier position of AgrS. This creates the proper configuration for checking the features of AgrS against the corresponding features of the subject. As a result, the features of AgrS and the corresponding features on the subject are eliminated.

Consider the consequences for (18a). Suppose (18a) is derived from an intermediate representation corresponding to (19), (18a′):

(18a′) *seems John is likely to win*

At this point, the subject's features will be eliminated as a result of the checking operation taking place in Spec-AgrSP in the embedded clause, just like in (19). This will make it impossible for (18a) to be derived from (18a′): the subject has lost its features, and cannot serve to check the features of the AgrS in the matrix clause. This derivation of (18a) is excluded.

Another possible derivation of (18a) moves the subject in one swoop from the VP-internal position to the specifier position of the matrix AgrSP. This creates the proper configuration for checking the features of the matrix AgrSP. However, it leaves the features of the embedded AgrSP unchecked. This derivation, therefore, is also excluded.

A final possibility would be to move the subject in one swoop from the VP-internal position to the specifier position of the matrix AgrSP, and to insert an expletive *it* in the specifier position of the embedded AgrSP, in order to check the latter's features. As this would yield (18b), this derivation must be excluded as well. This derivation assumes that *it* is a dummy subject that can be inserted when the derivation requires it. However, it is questionable whether *it* has this dummy status.

Bennis (1986) argues convincingly that expletives of the type of *it* are not dummy subjects, but part of the complement domain of the verb *seem* (see also Moro 1993). If so, *it* can never be simply *inserted* in the specifier position of the embedded AgrSP in (18b). If it ends up in that position, it must have *raised* just like *John* in (19). The ungrammaticality of (18b), then, shows that *it* can only raise to the matrix AgrSP, and *John* can only raise to the embedded AgrSP. How can this be explained?

I would like to propose here that raising verbs like *seem* select a

propositional complement (SMALL CLAUSE, including AgrP), not a sentential complement (CP). Consequently, if a CP appears in the complement domain of a raising verb, it must be either the subject or the predicate of a Small Clause (see Moro 1993). Considering that Small Clauses generally do not have sentential subjects, the starting hypothesis would be that a CP in the complement domain of a raising verb is a Small Clause predicate, needing a subject. My proposal is that *it* is the unique subject that can be associated with the CP predicate.

According to this proposal, (19) must have the structure in (20), with *it* generated as the subject of the complement of the higher raising verb:

(20) $[_{AgrSP1}$ AgrS1 $[_{VP1}$ V1 [*it* $[_{CP}$ $[_{AgrSP2}$ AgrS2 $[_{VP2}$ V2 [*John (to) win*]]]]]]]

It is easy to see that starting from the structure in (20), we will never be able to derive (18b). *John* cannot move to the specifier position of AgrSP1, because this would leave the features of AgrS2 unchecked. *It* is not available for checking the features of AgrS2, because *it* is generated in the matrix clause (and lowering is not allowed). The only correct outcome, then, is (19).

The proposal made here implies that (21) is not a legitimate structure:

(21) $[_{AgrSP1}$ AgrS1 $[_{VP1}$ V1 $[_{CP}$ $[_{AgrSP2}$ AgrS2 $[_{VP2}$ V2 [*it* [*John (to) win*]]]]]]]

In (21), V1 does not have a Small Clause/AgrP complement, but a direct sentential complement CP. This, we have assumed, is not allowed.[10]

What evidence do we have that raising verbs take propositional complements only? Consider first an empirical argument.

Take the Dutch raising verb *schijnen* 'seem, appear'. *Schijnen* appears in standard raising constructions, illustrated in (22):

(22) a. *Het schijnt dat Jan ziek is.*
it seems that John sick is

---

10. In Zwart (1993b), I explained the illegitimacy of (21) by assuming that *it* cannot be combined with a nonfinite clause. This appears to be a correct generalization, but it would not exclude all possible cases of superraising. For example, we could leave out *it* in (21), and turn the verb *win* into a passive verb. A possible outcome of that structure would be the ungrammatical *The race seems John appears to be won.* This cannot be explained by conditions on the distribution of *it*, but it can on the assumption that CP cannot be a direct complement of a raising verb.

b. *Jan schijnt ziek te zijn.*
John seems sick to be
c. *Jan schijnt ziek.*
John seems sick

In (22c), *schijnen* takes a propositional complement [*Jan ziek*], in (22b) an infinitival complement [*e te zijn* [*Jan ziek*]]. (22a) is the expletive construction. Crucially, *schijnen* cannot take a single noun phrase argument. Thus, the Dutch counterpart of *John appeared* is (23a), not (23b):

(23) a. *Jan ver-scheen.*
John appeared
b. **Jan scheen.*

The prefix *ver-* has been analyzed as an incorporated predicate by Mulder (1992). If this is correct, *schijnen* in (23a) again takes a propositional complement [*Jan ver*].[11]

(22b ,c) and (23), then, indicate that *schijnen* takes a propositional complement. The only exception would be (22a), unless we adopt the proposal advanced here, according to which *schijnen* in (22a) would take a complement [*het* [*dat Jan ziek is*]].[12]

A second observation supporting the analysis proposed here, is that CP complements are *objects*, whereas raising verbs (of the type of *be* and *seem*) select *states* (i.e. subject predicate combinations). Assuming categorical selection to be a function of semantic selection (Grimshaw 1981, Pesetsky 1982 and 1994, Chomsky 1986a), this makes it impossible for raising verbs to directly select a CP complement.

I assume Small Clauses (i.e. subject–predicate combinations) to be the canonical structural realization of states. This would make the Small Clause the prototypical raising complement. In the absence of evidence to the contrary, we may conclude that in typical raising constructions like (22a) the prototypical

---

11. In (23b), the only permissible interpretation is that John is an element that emits light (e.g. a celestial body). In that case, *schijnen* is not a raising verb, as indicated by standard tests. For example, *schijnen* in the sense of 'emitting light' takes a *have*-auxiliary instead of a *be*-auxiliary.

12. The prefix *ver-* does not appear in the sentences in (22). This might be taken to indicate that in (22), the predicate position is taken by other predicates. This is straightforwardly true in (22b) and (22c), but in (22a) again only if *schijnen* takes a propositional complement in (22a) as well.

raising complement is realized. This implies that in (22a), too, the complement of *schijnen* is a Small Clause. The analysis of "expletive" *het* as a Small Clause subject captures this.[13]

In other words, superraising does not result from a violation of the shortest move requirement, but from a violation of the selection requirements of raising verbs (in addition to the violation of Greed discussed in connection with (18a)).

Other superraising constructions can be explained along similar lines. Consider (24):[14]

(24) *$John_i$ *seems* [*it is likely* [*that we assured them* [t_i *to win the race*]]]

This sentence is excluded because the raising predicates *seems* and *likely* cannot take a CP complement directly. The CP would have to be part of a Small Clause with *it* as its subject. In (24), this condition is met with *likely*, but not with *seems*. The only possible output, therefore, would involve two instances of *it*, but then there would be no position left in which *John* could be licensed:

(25) **It*$_i$ *seems* t_i [*it*$_j$ *is likely* t_j [*that we assured them* [*John to win the race*]]]

Consider also (26):

(26) **John*$_i$ *seems* [*that it was told* t_i [*that he would win*]]

Again, *seems* lacks a Small Clause subject in its complement domain. Inserting *it* there would again rob *John* of its licensing position:

(27) **It*$_i$ *seems* t_i [*that it*$_j$ *was told John* t_j [*that he would win*]]

This explanation for the nonexistence of superraising, introduced to account for the ungrammaticality of (18b), in fact also extends to the cases of superraising illustrated by (18a). If *seems* cannot have a CP in its complement without also having *it* in its complement, (18a) will never be generated.

In sum, the superraising constructions in (18) can both be excluded without taking recourse to the shortest move requirement of economy of derivation.

---

13. I have been assuming that infinitival complements are structurally comparable to Small Clause complements (i.e. Small Clause complements of the category AgrP). It is not clear to me at this point that infinitival complements should be characterized as states also. I would like to postpone discussion of this issue until after further study.

14. Brought to my attention by Phil Branigan (p.c.).

4.3. *Wh-Islands*

Wh-island facts present a third class of phenomena which have been presumed to motivate a shortest move requirement.

(28) **What did he wonder where John put* **t**?

In (28), *what* moves out of an embedded interrogative clause, crossing the position occupied by *where*. This position (the specifier position of the embedded CP) being a potential landing site for *what*, minimality (7) is violated.

The analysis of *wh*-movement as being *successive cyclic*, first put forward in Chomsky (1973), had a major impact on grammatical theory. It is probably correct to state that the desirability of having short steps in a derivation goes back to this pioneering work.

However, the shortest steps requirement goes much further than the successive cyclic movement requirement. Successive cyclic movement involves employing intermediate landing sites on the edge of a local domain (in the core cases, CP). But movement to the edge of the local domain (i.e., to Spec-CP) is not the shortest movement imaginable. The possibility of moving via an adjunction position to VP has been explored in Chomsky (1986b), for instance.

The discussion of how many intermediate landing sites are involved in *wh*-movement does not center around the length of the steps, but around the opacity of a certain local domain. There has to be an intermediate landing site between the boundaries of two opaque local domains.

As Koster (1987) argues at length, locality conditions on movement can be regarded as a subcase of the more general locality conditions on dependency relations. Movement creates a dependency relation between an antecedent and a trace, which therefore have to be in the same local domain. Intermediate elements may serve to link an antecedent and a trace which are not in the same local domain.

From this point of view, intermediate landing sites are not needed to comply with conditions on derivation, but forced by conditions on interpretation.

There are three additional reasons to be skeptical about the relevance of the shortest move requirement in the domain of *wh*-movement.

First, as Chomsky (1993: 15) notes, the shortest move requirement appears to be incompatible with the fewest steps requirement. The number of steps increases with the shortness of the movements. The fewest steps requirement appears to be an essential ingredient of the minimalist program. Procrastinate (3) and Greed (4) are both instantiations of the fewest steps requirement (2).

Therefore, if the incompatibility cannot be resolved, it seems that the shortest move requirement must be abandoned.

Chomsky (1993: 15), noting this problem, states that successive cyclic movement does not consist in a succession of identical adjunction operations but in a single operation *Form Chain*. This operation turns a representation like (29a) into, for instance, (29b), without the intermediate step that yields (29c):

(29) a. e [*you think* [e [*you love who*]]]
b. ***who*** [*you think* [***t*** [*you love* **t**]]]
c. e [*you think* [***who*** [*you love* ***t***]]]

It would take two steps to get from (29a) to (29b) via (29c), whereas Form Chain derives (29b) from (29a) in one step.

This proposal raises the question whether Form Chain is a new type of structure building process, or a combination of well-known structure building processes. Chomsky (1993) distinguishes two structure building processes, each operating in a bottom-up fashion (both referred to as GENERALIZED TRANSFORMATION). The first structure building process combines two independent phrase markers. The second structure building process adjoins to a phrase marker K a proper subpart of K. Insertion involves the first process, movement involves the second process.

Form Chain now appears to be a combination of movement and insertion. Therefore, we might want to decompose it into an insertion part, combining the 'intermediate' empty category and the phrase marker *you love who* (30a) and a movement part, adjoining *who* to the phrase marker *you think e you love* (30b):

(30) a. [**e** [*you love* ***who***]]
b. [***who*** [*you think* [***e*** [*you love* ***t***]]]]

This has the advantage that Form Chain is not a mysterious, novel structure building process. It has the disadvantage that the movement operation yielding (30b) violates the shortest move requirement. But this disadvantage disappears if the shortest move requirement is not part of economy of derivation, as argued here.

In (30b), *who* and its trace *t* are arguably not in the same local domain. We may hypothesize, however, that the intermediate empty element *e* can serve as a link between the antecedent and its trace, thus replicating the effect of successive cyclic movement.

The second reason to be skeptical about the relevance of the shortest move requirement in the domain of *wh*-movement is that it is a *vacuous* requirement, given the assumptions on structure building entertained in Chomsky (1993).

Chomsky (1993: 22) assumes that the structure building process is subject to a condition of Strict Cyclicity (also known as the EXTENSION CONDITION). In particular, insertion and movement are allowed only if these processes *extend* the phrase marker affected by them. Thus, the only way to move a subpart α of a phrase marker K is by adjoining α on the outside of K. Adjunction of α at any hierarchically lower point in K is excluded by the Strict Cycle condition.

In terms of conditions on movement, this means that at any point in a derivation, the only movement that is allowed is the *longest* possible movement. Given that a shorter movement is blocked by the Strict Cycle condition, there really appears to be no option at all in terms of the length of a movement.[15]

A third reason to have doubts about the shortest move condition is that in the domain of *wh*-movement violations of "shortest move" yield various gradations of ungrammaticality. The status of nonlocal *wh*-movement appears to be dependent on the nature of the local domain's opacity (strong versus weak islands), and on the status of the moved category (D-linked arguments versus non-D-linked arguments and adjuncts; see Cinque 1990 for a survey).

It is generally assumed since Chomsky (1991) that violations of economy conditions yield the worst kind of ungrammaticality. In Chomsky (1993), violations of economy conditions lead to a crashing derivation, which cannot be repaired by other components of the cognitive system. To the extent that nonlocal head movement and superraising can be reduced to violations of economy of representation (i.e., certain features remain unchecked), the relevant constructions have the expected quality of ungrammaticality. Nonlocal *wh*-movement constructions do not in general, suggesting that no economy violation is involved in the derivation of these constructions.

4.3.1. *Form Chain*

In the previous section, we have argued that Form Chain can be decomposed into two standard structure building processes (cf. (30)). First, the embedded clause is extended by adjunction of an independent empty *wh*-element (30a). Second, the matrix clause is extended by movement and adjunction of the *wh*-word generated in the embedded clause. We have assumed that the

---

15. Chomsky (1993: 24) tentatively restricts the Strict Cycle condition to overt movement to specifier positions. This implies that for covert movement, the shortest move requirement is not vacuous.

intermediate empty *wh*-element then serves to link the antecedent to its trace through a succession of local dependency links.

The more traditional successive cyclic movement faces a problem that the Form Chain process, made explicit in this way, avoids. Suppose the intermediate empty *wh*-element were not an independently inserted element, but a trace of the *wh*-moved category itself. This implies that the first step in the derivation involves movement and adjunction of the *wh*-word generated in the embedded clause to the embedded clause (which, at that point in the derivation, is not yet "embedded") (31a). In the second step, this *wh*-word will move on and adjoin to the matrix clause (31b):

(31) a. [***who*** [*you love* ***t***]]
b. [***who*** [*you think* [***t*** [*you love* ***t***]]]]

The principle of Greed (4) requires that there is a trigger for each of the movements in (31a) and (31b). Arguably, there can be a ⟨+*wh*⟩-feature in the embedded C, which needs to be checked against the *wh*-features of an element in Spec-CP.[16] This ⟨+*wh*⟩-feature provides a trigger for the movement in (31a), leading to elimination of both the ⟨+wh⟩-feature in C and the *wh*-feature of the moved *wh*-word. But at this point, no trigger is left for the movement of the *wh*-word in (31b). Its *wh*-features have been checked in the intermediate Spec-CP, and no further movement is allowed, by Greed.

In the Form Chain approach pursued here, the ⟨+*wh*⟩-features of the embedded clause are checked by an independent empty *wh*-element, leaving the *wh*-word generated in the embedded clause free to move to the specifier position of the matrix CP.

---

16. Long distance *wh*-movement shows effects of the presence of a *wh*-element in the specifier position of the embedded CP in a number of languages. For example, in Dutch, a *wh*-complementizer *ofdat* can appear in long distance *wh*-movement constructions and embedded interrogatives only (cf. Hoekstra & Zwart 1994):

(i) a. *Piet denkt (*of) dat Jan het gedaan heeft.*
Pete thinks if that John it done has
'Pete thinks that John did it.'
b. *Wat denkt Piet (of) dat Jan gedaan heeft?*
what thinks Pete if that John done has
'What does Pete think John did?'
c. *Piet vraagt zich af wat (of [dat]) Jan gedaan heeft.*
Pete asks himself off whatif that John done has
'Pete wonders what John did.'

Chomsky (1995b) avoids the problem for successive cyclic movement posed by the principle of Greed by assuming that checking of features does not automatically result in the elimination of features. In particular, the *wh*-word in (31a) would check and eliminate the ⟨+wh⟩-feature of the embedded C, but would retain its own *wh*-features for further checking operations to take place in the matrix CP.

There is reason to believe that a *wh*-element in the specifier position of an embedded CP still has certain features left which need to be checked in the specifier position of the matrix CP. This becomes clear from partial *wh*-movement constructions in German (McDaniel 1989; Huybregts 1992; Gamon 1995). In these constructions, the step in (31a) is overt, but the step in (31b) does not take place in overt syntax. Instead, a dummy *wh*-element *was* 'what' is inserted in the specifier position of the matrix CP:

(32) [***Was*** *glaubst du* [***mit*** ***wem*** *ich* ***t*** *gesprochen habe*]]
what believe you with whom I spoken have
'Who do you think I talked to?'

Gamon (1995) observes that the construction in (32) is ungrammatical if the *wh*-element in the specifier position of the embedded CP is located inside a weak island (e.g., a *wh*-island):

(33) *[***Was*** *fragst du* *dich* [***weshalb*** *er glaubt*
what ask you yourself why he believes
[***mit*** ***wem*** *ich* ***t*** *gesprochen habe*]]]
with whom I spoken have
'Who do you wonder why he thinks who I talked to?'

This suggests that the *wh*-phrase in the embedded clause *mit wem* 'with whom' moves covertly (at LF) to the matrix Spec-CP occupied by the dummy *wh*-word *was* (cf. also Huybregts 1992).[17]

It is not clear that the trigger for the covert movement would involve *wh*-features. The dummy element in the matrix Spec-CP must have the morphology of a *wh*-word, suggesting that checking of the *wh*-features has been taken care of in overt syntax. Nevertheless, (33) appears to present clear evidence for covert movement from the embedded Spec-CP to the matrix Spec-CP.

Crucially, however, the same evidence is lacking in nonpartial

---

17. In Chomsky (1995b), the covert movement would involve movement of the relevant *features* to the matrix C, not movement of the entire *wh*-phrase to the matrix Spec-CP.

wh-movement constructions in English, as weak islands only yield marginally ungrammatical sentences:[18]

(34) ?***Who*** *do you wonder whether Bill believes* **e** *we talked to* **t** *?*

Apparently, the relation between *who* and the intermediate empty element *e* in (34) is not subject to the same constraints as the relation between the dummy *wh*-element *was* and *mit wem* in (33). This suggests that the empty element *e* in (34) is not a trace of *who* but an independently inserted intermediate *wh*-element, as proposed here.

This leads to the following conclusion. Even if it is possible for *wh*-elements to move successive cyclically, the evidence from partial vs. full *wh*-movement suggests that a derivation in which the *wh*-phrase moves long distance in one step must always be available. This would be impossible if the shortest move condition is part of economy of derivation.

4.4. *Conclusion*

It appears that head movement, superraising, and *wh*-movement do not provide decisive evidence in support of the validity of the shortest move condition. Nonlocal head movement is excluded by economy of representation: skipping relevant functional heads would leave certain features unchecked. Part of the superraising cases can be explained as economy of representation violations as well. For the remaining cases, I have argued that they could only be derived by starting from illegimitate base structures (i.e., structures that would never be generated). Nonlocal *wh*-movement yields gradations of ungrammaticality and therefore appears to belong to a different category. It can be argued that successive cyclic movement does not consist in a stepwise movement procedure, but in a combination of inserting intermediate empty *wh*-elements and moving the *wh*-phrase long distance to the matrix CP.

---

18. The same contrast between partial and full *wh*-movement can be observed in Frisian, in which both types of *wh*-construction are possible (Pytsje van der Veen, p.c.).

5. Consequences

5.1. *Equidistance*

Let us now return to the Equidistance Principle of Chomsky (1993). This principle allows elements to cross a position where they could have landed, provided the target position is in the same minimal domain as the position which is crossed.

(35) *Equidistance* (Chomsky 1993: 17):
If α,β are in the same minimal domain, they are equidistant from γ.

As we have seen, positions end up being in the same minimal domain as a function of head movement. This leads to the prediction that objects can leave the VP, crossing the subject position inside VP, only if the verb has moved to AgrO first.[19]

As argued in Section 3, this prediction is not borne out. This suggests that the equidistance principle as stated in (35), under the definitions understood, is not relevant for conditions on movement. But this is not surprising if the shortest move condition does not exist. Since the equidistance principle is merely an explicitation of the shortest move requirement, we are not surprised to find that it has no empirical substance.

The shortest move requirement, including the equidistance principle, does make one important contribution that is lost in the present approach. Chomsky (1993: 19) derives from this requirement the fact that movement of subject and object to their respective licensing positions (L-RELATED MOVEMENT) is always crossing instead of nesting. It is well-known that movement to operator positions (NON-L-RELATED MOVEMENT) is nesting rather than crossing (Pesetsky 1982).

This follows in Chomsky's analysis, because (for reasons not discussed here) a minimal domain contains at most two equidistant specifier positions.[20] As a result, if the subject moves to the specifier position of AgrO, so that the

---

19. See Chomsky (1993: 12) for the definition of MINIMAL DOMAIN. See also Zwart (1993b: 231 f.).

20. As pointed out by a reviewer, if the Head Movement Constraint does not hold, it becomes possible for more than two specifier positions to end up in the same minimal domain. The specifier of a head that is skipped in the head movement process would also be included in the minimal domain of the chain resulting from the head movement.

subject and its trace occupy the two positions that are equidistant from the object, the object cannot move at all. Consequently, the object has to hop over the subject first, and only then is the subject allowed to cross the object, again as a function of head movement creating the proper equidistant landing site.

This important result is lost.[21] We could try to achieve the same result by stipulating the order of agreement projections in the functional domain, but that would clearly be inferior to Chomsky's approach.[22] This, then, is the problem to solve if we accept the point argued for here, namely that the shortest move requirement does not exist.

5.2. *Reductions*

Accepting the arguments against the existence of the shortest move requirement, economy of derivation can be stated as in (2):

(2) *Economy of derivation*:
Use as few steps as possible in deriving an output representation.

The question arises whether economy of derivation can be reduced to economy of representation:

---

21. However, as pointed out to me by Liliane Haegeman and Gereon Müller, the Equidistance Principle as stated is not able to accommodate derivations involving two internal arguments and one external argument. Assuming both internal arguments to have designated licensing positions in the functional domain (i.e. there is a sequence of two AgrOPs), the higher internal argument will at some point in the derivation have to cross the basic position of the external argument (i.e. the Spec-VP) and the licensing position of the lower internal argument (i.e. the specifier of the lower AgrOP) on its way to its licensing position (i.e. the specifier of the higher AgrOP). Spec-VP and the two Spec-AgrOPs, however, cannot all three be equidistant from the basic position of the higher internal argument. The Equidistance Principle appears to be unable to derive these structures. If so, the Equidistance Principle captures the crossing requirement on A-movement only partly. However, see Collins and Thráinsson (1993) for an analysis of double object constructions that does not incur this problem.

22. As pointed out by a reviewer, even stipulating the order of agreement projections does not help much. If we stipulate that AgrSP must be higher than AgrOP, we still cannot exclude that a subject is generated with objective Case morphology and licensed in AgrOP (yielding *She loves him* meaning *He loves her*). It seems, then, that the organization of the functional domain expresses a (partial) link between argument status, morphology, and word order.

(1) *Economy of representation*:
Use as few symbols as possible in an output representation.

Inasmuch as steps reduce the number of symbols, (2) can be reduced to (1).

Movement is generally regarded as a chain creating process. If a chain counts as one symbol, movement will not decrease the number of symbols. However, features must count as symbols, or otherwise they would not be relevant for the principle of Full Interpretation:

(36) *Full Interpretation*:
In an interface representation, do not use useless symbols.

Therefore, only movements that result in checking and elimination of features is allowed by (1).

We can also state that strong features count as symbols at the PF interface, but weak features do not. Movement to check and eliminate strong features then reduces the number of symbols, but movement to check and eliminate weak features does not. *Procrastinate* (3), then, can be reduced to economy of representation holding for the PF interface.

Greed (4) follows from (1) if we take (1) to apply in a nonglobal fashion, that is: every element must carry as few symbols as possible in the output representation. Thus, (1) forces a feature bearing element α to move, but prohibits movements that do not result in the elimination of a feature carried by α.

If this conception of *Greed* is correct, economy of derivation (i.e. the "fewest steps" requirement) and the principles that fall under economy of derivation (the *inertness* principles) can be seen to reduce to economy of representation.

Finally, note that economy of representation can be reformulated as in (37):

(37) *Economy of representation (reformulated)*
In an output representation, do not use superfluous symbols.

It is clear from this reformulation that economy of representation, now incorporating economy of derivation, is equivalent to the principle of *Full Interpretation* (36).

6. Conclusion

I have argued that economy of derivation does not contain a requirement that steps be as short as possible. Restrictions on head movement, superraising,

and *wh*-movement follow from well established principles and mechanisms made explicit in recent minimalist work. This resolves the conflict between the "fewest steps" requirement and the "shortest move" requirement in Chomsky (1993). I argued that Chomsky's *Form Chain* mechanism must be decomposed into a "trace" insertion mechanism and a movement mechanism, where the intermediate "traces" are inserted before the movement takes place, in accordance with strict cyclicity. Finally, I offered some speculation on the possibilities of reducing economy of derivation to economy of representation, and economy of representation to the principle of Full Interpretation.

It is interesting to note that, if we were correct in the above, generative grammar, even in the purely derivational approach that incorporates the structure building process of generalized transformations (including movement), does not need any constraints beyond conditions on representation.[23]

Acknowledgements

This material was presented at the TABU-dag at the University of Groningen, June 25, 1993, at the Stafseminar of the Department of Linguistics of the University of Tilburg, September 8, 1993, and at the Research Seminar of the Department of Linguistics of the University of Geneva, March 8, 1994. I would like to thank the audiences at these occasions, in particular Jaume Solà, Henk van Riemsdijk, Riny Huybregts, Luigi Rizzi, Ur Shlonsky, and Liliane Haegeman. Also thanks to Eric Hoekstra, Phil Branigan, Gereon Müller, Anthony Kroch, Erich Groat, and K. Scott Ferguson and to two reviewers for this volume.

---

23. This final remark presupposes that locality conditions and conditions that ensure the correct order of projections in the functional domain are to be considered as representational rather than derivational conditions. As for the conditions on interpretation requiring a moved element to be related to its trace, it was assumed in the text that these are either representational conditions, or extragrammatical conditions of interpretation.

References

Abney, Steven. 1987. "The English Noun Phrase in Its Sentential Aspect." Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge.

Abraham, Werner. 1993a. "The Focus Interface Hypothesis and the Historical Decay of the Verbal Genitive in German." de Boer, de Jong & Landeweerd 1993.1–10.

———. 1993b. "Null Subjects in the History of German: From IP to CP." *Lingua* 89.117–142.

———. 1994. "Kayne's Asymmetriehypothese und die Abfolge im V-Komplex." Zwart, ed. 1994.1–17.

———, ed. 1983. *On the Formal Syntax of the Westgermania.* Amsterdam: John Benjamins.

Abraham, Werner & Josef Bayer, eds. 1993. *Dialektsyntax* [*Linguistische Berichte*, Thematic issue.]. Opladen: Westdeutscher Verlag.

Abraham, Werner, Wim Kosmeijer & Eric Reuland, eds. 1991. *Issues in Germanic Syntax.* Berlin: Mouton de Gruyter.

Abraham, Werner & Anko Wiegel. 1993. "Reduktionsformen und Kasussynkretismus bei deutschen und niederländischen Pronomina." Abraham & Bayer 1993.12–50.

Agouraki, Yoria. 1993. "Spec-Head Licensing: The scope of the theory." Ph.D. Dissertation, University College London.

Allen, Barbara J., Donna B. Gardiner & G. Donald. 1984. "Noun Incorporation in Southern Tiwa." *International Journal of American Linguistics* 50.292–311.

Anderson, Stephen R. 1993. "Wackernagel's Revenge: Clitics, Morphology, and the Syntax of Second Position." *Language* 69.68–98.

Anderson, Stephen R. & Paul Kiparsky, eds. 1973. *A Festschrift for Morris Halle.* New York: Holt, Rinehart & Winston.

Aoun, Joseph & Audrey Yen-Hui Li. 1989. "Constituency and Scope." *Linguistic Inquiry* 20.141–172.

———. 1991. "The Interaction of Operators." Freidin 1991.163–181.

———. 1993. *The Syntax of Scope*. Cambridge, Mass.: MIT Press.

Aoun, Joseph & Dominique Sportiche. 1983. "On the Formal Theory of Government." *The Linguistic Review* 2.211–236.

Bach, Emmon. 1962. "The Order of Elements in a Transformational Grammar of German." *Language* 38: 263–269.

Baker, Mark. 1985. "The Mirror Principle and Morphosyntactic Explanation." *Linguistic Inquiry* 16:373–416.

———. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: The University of Chicago Press.

———. 1991. "On Some Subject/Object Non-Asymmetries in Mohawk." *Natural Language and Linguistic Theory* 9.537–576.

Baker, Mark C. & Kenneth Hale. 1990. "Relativized Minimality and Pronoun Incorporation." *Linguistic Inquiry* 21.289–297.

Baker, Mark C., Kyle Johnson & Ian G. Roberts. 1989. "Passive Arguments Raised." *Linguistic Inquiry* 20.219–251.

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

Belletti, Adriana. 1990. *Generalized Verb Movement*. Torino: Rosenberg & Sellier.

———. 1993. "Case Checking and Clitic Placement." *GenGenP* 1(2).101–118.

Belletti, Adriana & Luigi Rizzi, eds. 1995. *Parameters and Functional Heads: Essays in comparative syntax*. New York: Oxford University Press.

Benincà, Paola, ed. *Dialect Variation and the Theory of Grammar*. Dordrecht: Foris.

Bennis, Hans. 1986. *Gaps and Dummies*. Dordrecht: Foris.

Bennis, Hans & Frits Beukema, eds. 1985. *Linguistics in the Netherlands 1985*. Dordrecht: Foris.

Bennis, Hans & Teun Hoekstra. 1989. "Why Kaatje Was Not Heard Sing a Song." Jaspers *et al.* 1989.21–40.

Besten, Hans den & Jerold Edmondson. 1983. "The Verbal Complex in Continental West Germanic." Abraham 1983.155–216.

Besten, Hans den *et al.* 1988. "Verb Raising, Extrapositie en de Derde Constructie." Ms., University of Amsterdam.

Besten, Hans den & Gert Webelhuth. 1990. "Stranding." Grewendorf & Sternefeld 1990.77–92.

Bloomfield, Leonard. 1933. *Language*. New York: Holt.

Bobaljik, Jonathan D. 1995a. "The Morphological Determination of Germanic Syntax." Paper presented at the 10th Comparative Germanic Syntax Workshop, Brussels.

———. 1995b. "Morphosyntax: The syntax of verbal inflection." Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge. [Distributed by *MITWPL*.]

Bobaljik, Jonathan & Andrew Carnie. 1992. "A Minimalist Approach to Some Problems of Irish Word Order." Paper presented at the XIIth Harvard Celtic Colloquium 1992. [To appear in the proceedings of the colloquium, edited by B. Hillers & J. Hunter. A revised version to appear in Robert Borsley & Ian Roberts, eds.: *Celtic and Beyond*, Cambridge: Cambridge University Press.]

Bobaljik, Jonathan D. & Dianne Jonas. 1993. "Subject Positions and the Role of TP." Paper presented at GLOW 16, Lund. [Revised version appeared in *Linguistic Inquiry* 27.195–236.]

Bobaljik, Jonathan D. & Colin Phillips, eds. 1993. *Papers on Case and Agreement.* Vol. 1. *MIT Working Papers in Linguistics* 19.131–173. Massachusetts Institute of Technology, Cambridge.

Boer, Ale de, Jelly de Jong & Rita Landeweerd, eds. 1993. *Language and Cognition 3. Yearbook 1993*. Groningen, University of Groningen.

Boer, Ale de, Helen de Hoop & Henriëtte de Swart, eds. 1994. *Language and Cognition 4. Yearbook 1994.* Groningen, University of Groningen.

Booij, Geert E. 1985. "Lexical Phonology, Final Devoicing and Subject Pronouns in Dutch." Bennis & Beukema 1985.21–26.

Brandi, Luciana & Patrizia Cordin. 1989. "Two Italian dialects and the Null Subject Parameter." Jaeggli & Safir 1989.111–142.

Branigan, Philip. 1992. "Subjects and Complementizers." Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge.

———. 1995. "Improper Movement and French Participle Agreement." Ms., Memorial University of Newfoundland.

Bresnan, Joan. 1982. "Control and Complementation." Bresnan 1982.282–390.

———, ed. 1982. *The Mental Representation of Grammatical Relations.* Cambridge, Mass.: MIT Press.

Brody, Michael. 1993. "Θ-Theory and Arguments. *Linguistic Inquiry* 24.1–24.

Broekhuis, Hans. 1993. "Verb Projection Raising." *Spektator* 22.28–47.

Browning, Margaret, ed. To appear. *Working Minimalism*. Cambridge, Mass.: MIT Press.

Bures, Anton. 1992. "Re-Cycling Expletives (and Other Sentences)." Unpublished general examination paper, MIT.

———. 1993. "There Is an Argument for a Cycle at LF Here." *Chicago Linguistic Society* 28/2.14–35.

Burton, Strang & Jane Grimshaw. 1992. "Coordination and VP-internal Subjects." *Linguistic Inquiry* 23.305–313.

Campbell, Richard. 1991. "Tense and Agreement in Different Tenses." *The Linguistic Review* 8.159–183.

Cardinaletti, Anna. 1992a. "On the Typology of Pronouns and the Notion of Syntactic Support." Paper presented at the ESF Plenary Conference, San Sebastian, September 2–5.

———. 1992b. "On Cliticization in Germanic Languages." Ms., University of Geneva.

Cardinaletti, Anna & Michal Starke. 1993. "On Dependent Pronouns and Pronoun Movement." Paper presented at GLOW 16, Lund.

Carlson, Greg. 1977. "A Unified Analysis of the English Bare Plural." *Linguistics and Philosophy* 1.413–457.

Cecchetto, Carlo. 1994. "Clitics, Specificity and Some Other Stories." Paper presented at the Seminaire de Recherche, Geneva.

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.

———. 1965. *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.

———. 1970. "Remarks on Nominalization." In: Roderick A. Jacobs & Peter S. Rosenbaum, eds. *Readings in English Transformational Grammar*, pp. 184–221. Waltham: Ginn and Co. [Also published in Noam Chomsky 1972: *Studies on Semantics in Generative Grammar*, pp. 11–61. The Hague: Mouton.]

———. 1973. "Conditions on Transformations." Anderson & Kiparsky 1973.232–286.

———. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

———. 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, Mass.: MIT Press.

———. 1986a. *Knowledge of Language. Its Nature, Origin, and Use*. New York: Praeger.

———. 1986b. *Barriers*. Cambridge, Mass.: MIT Press.

———. 1987. "On the Nature, use, and Acquisition of Language." In *Generative Grammar: Its Basis, Development, and Prospects*. Kyoto University of Foreign Studies [*Studies in Linguistics and Literature*, special issue].

———. 1991. "Some Notes on Economy of Derivation and Representation." Freidin 1991.417–454.

———. 1993. "A Minimalist Program for Linguistic Theory." Hale & Keyser 1993.1–52.

———. 1995a. "Bare Phrase Structure." Webelhuth 1995.383–439.

———. 1995b. "Categories and Transformations." Chomsky 1995c.219–394.

———. 1995c. *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Chomsky, Noam & Howard Lasnik. 1977. "Filters and Control." *Linguistic Inquiry* 8: 425–504.

———. 1993. "Principles and Parameters Theory." Jacobs, von Stechow, Sternefeld & Vennemann 1993.506–569.

Christensen, Kirsti Koch. 1984. "Subject clitics and $\overline{A}$-Bound traces." *Working Papers in Scandinavian Syntax* 15.

Cinque, Guglielmo. 1990. *Types of $\bar{A}$-dependencies*. Cambridge, Mass.: MIT Press.

Cinque, Guglielmo *et al.*, eds. 1995. *Paths towards Universal Grammar. Studies in Honor of Richard S. Kayne*. Washington, D.C.: Georgetown University Press.

Collins, Chris. 1994. "Merge and Greed." Ms. Cornell University, Ithaca.

Collins, Chris & Höskuldur Thráinsson. 1993. "Object Shift in Double Object Constructions and the Theory of Case." Phillips 1993.131–173.

Corver, Norbert & Henk van Riemsdijk, eds. 1994. *Studies on Scrambling: Movement and non-movement approaches to free word-order phenomena*. Berlin: Mouton de Gruyter.

Dahl, Östen & Anders Holmberg, eds. 1986. *Scandinavian Syntax*. University of Stockholm, Stockholm.

Déprez, Viviane. 1989. "On the Typology of Syntactic Positions and the Nature of Chains. Move α to the Specifier of Functional Projections." Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge. [Distributed by MIT Working Papers in Linguistics.]

Dikken, Marcel den. 1988. "The Syntax of Movement Rules Affecting Verbs and Verbal Constituents." Master's Thesis, University of Leiden.

———. 1989. "Verb Projection Raising en de Analyse van het IPP-effect." *Tabu* 19.59–75.

———. 1994a. "Auxiliaries and Participles." *NELS* 24.65–79. Amherst, GLSA.

———. 1994b. "Minimalist Verb Projection Raising." Zwart ed. 1994.71–88.

———. 1995a. "Verb Projection Raising, Scope and Uniform Phrase Structure." *NELS* 25, Amherst: GLSA.

———. 1995b. "Binding, Expletives and Levels." *Linguistic Inquiry* 26.347–354.

———. 1995c. "Copulas." Paper presented at GLOW 18, Tromsø. Ms., Vrije Universiteit Amsterdam/HIL.

Dikken, Marcel den & Hans Broekhuis. 1993. "Locality in Minimalist Syntax." *HIL Manuscripts* 1(3).1–15.

Dikken, Marcel den & Eric Hoekstra. 1995. "Parasitic Participles." Paper presented at the 10th Comparative Germanic Syntax Workshop, Brussels. Ms., Vrije Universiteit Amsterdam/HIL & P.J. Meertens-Instituut.

Drijkoningen, Frank & Ans van Kemenade, eds. 1991. *Linguistics in the Netherlands*. Dordrecht: Foris.

Einstein, Albert. 1954. *Ideas and Opinions*. New York: Crown.

Emonds, Joseph E. 1978. "The verbal complex V′–V in French." *Linguistic Inquiry* 9.151–175.

Epstein, Samuel David. 1987. *Empty Categories and Their Antecedents*. Doctoral dissertation, University of Connecticut.

———. 1989. "On Adjunction and Pronominal Variable Binding." *Linguistic Inquiry* 20.307–319.

———. 1991. *Traces and Their Antecedents*. New York: Oxford University Press.

———. 1992. "Derivational Constraints on Ā-chain Formation." *Linguistic Inquiry* 23.135–159.
———. 1993. "Superiority." Thráinsson, Epstein & Kuno 1993.14–64.
———. 1994. " On the Derivation of Syntactic Relations." Paper presented at the Harvard University Forum in Synchronic Linguistic Theory, December 1994.
———. 1995. "Un-Principled Syntax and the Derivation of Syntactic Relations." Paper presented at the University of Maryland at College Park. To appear in Browning, ed., to appear.
———. To appear. "Overt Scope-Marking and Covert V2." To appear in *Linguistic Inquiry*.
Epstein, Samuel David, Erich Groat, Ruriko Kawashima, & Hisatsugu Kitahara. Forthcoming. *Non-Representational Syntax*. New York: Oxford University Press.
Evers, Arnold. 1975. "The Transformational Cycle in Dutch and German." Ph.D. Dissertation, University of Utrecht [Distributed by the Indiana University Linguistics Club, Bloomington].
Ferguson, K. Scott. 1993. "Notes on the Shortest Move Metric and Object Checking." Thráinsson, Epstein & Kuno 1993.65–80.
———. 1995. "A Feature-Relativized Shortest Move Requirement." Ph.D. dissertation, Harvard University, Cambridge, Mass.
Ferguson, K. Scott & Erich M. Groat. 1994. "Defining 'Shortest Move'." Paper presented at GLOW 17, Vienna.
———. 1995. "Defining 'Shortest Move'." Ms., Harvard University, Cambridge, Mass. Paper presented at GLOW 17. Vienna.
Frampton, John. 1991. "Relativized Minimality: A review." *Linguistic Review* 8.1–46.
Frank, Robert & Anthony S. Kroch. 1992. "Generalized Transformations in Successive Cyclic and Long Dependencies or Substitution + Extending = Adjoining." Paper presented at GLOW 16, Lund.
Frank, Robert, Young-Suk Lee & Owen Rambow. 1992. "Scrambling as Non-operator Movement and the Special Status of Subjects." Ms., University of Pennsylvania.
Freidin, Robert. 1978. "Cyclicity and the Theory of Grammar." *Linguistic Inquiry* 9.519–549.
———. 1991. "Introduction." In: Robert Freidin, ed. *Principles and Parameters in Comparative Grammar*, pp. 1–6. Cambridge, Mass.: MIT Press.
———, ed. 1991. *Principles and Parameters in Comparative Grammar*. Cambridge, Mass.: MIT Press.
Fukui, Naoki. 1993. "Parameters and Optionality." *Linguistic Inquiry* 24.399–420.
Fukui, Naoki & Margaret Speas. 1986. "Specifiers and Projections." *MIT Working Papers in Linguistics* 8.128–72.
Gamon, Michael. 1995. "Partial *Wh*-movement in German: A non-homogeneous *wh*-dependency." Samiian & Schaeffer 1995.106–119.

Gelderen, Elly van. 1993. *The Rise of Functional Categories*. Amsterdam: John Benjamins.

Gilbers, Dick & Sietze Looyenga, eds. 1992. *Language and Cognition*. University of Groningen.

Görlach, Manfred. 1978. *Introduction to Early Modern English*. Cambridge: Cambridge University Press.

Greenberg, Joseph. 1963. "Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements." Greenberg, ed. 1963.73–113.

———, ed. 1963. *Universals of Language*. Cambridge, Mass.: MIT Press.

Grewendorf, Günther & Sternefeld, Wolfgang, eds. 1990. *Scrambling and Barriers*. Amsterdam: John Benjamins.

Grimshaw, Jane. 1989. "Complement Selection and the Lexicon." *Linguistic Inquiry* 10.279–326.

———. 1993. "Minimal Projections, Heads, and Optimality." Ms., Rutgers University, New Brunswick.

Groat, Erich. 1995. "On the Redundancy of Syntactic Representations." Paper presented at GLOW, Tromsø, Norway, June 1995.

———. In progress. *A Derivational Program for Syntactic Theory*. Doctoral dissertation, Harvard.

Groat, Erich M. & John O'Neil. This volume. "Spell-Out at the LF Interface."

Groenendijk, Dick, Jeroen de Jongh & Martin Stokhof, eds. 1987. *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*. Dordrecht: Foris.

Haeberli, Eric & Liliane Haegeman. 1992. "Clause Structure in Old English: Evidence from negative concord." Ms., University of Pennsylvania/Geneva.

Haegeman, Liliane. 1988. "Verb Projection Raising and the Multidimensional Analysis: Some empirical problems." *Linguistic Inquiry* 19.

———. 1991. *Introduction to Government and Binding Theory*. Cambridge: Basil Blackwell.

———. 1992. *Generative Syntax: Theory and description. A case study of West Flemish*. Cambridge: Cambridge University Press.

———. 1993a. "Some Speculations on Argument Shift, Clitics and Crossing in West Flemish." Abraham & Bayer 1993.131–160.

———. 1993b. "The Morphology and Distribution of Object Clitics in West Flemish." *Studia Linguistica* 47.57–94.

———. 1993c. "Object Clitics in West Flemish and the Identification of A/$\overline{\text{A}}$-positions." *GenGenP* 1(1).1–30.

———. 1994. "Root Infinitives, Tense and Truncated Structures." *GenGenP* 2(1).12–41.

———. 1995a. *The Syntax of Negation*. Cambridge: Cambridge University Press.

———. 1995b. "Root Infinitives and Root Null Subjects in Early Dutch." Paper presented at GALA, 1995. Groningen.

———. 1995c. "IPP Constructions and V-movement in West Flemish." Paper presented a the Manchester Workshop on Functional Categories. To appear in *GenGenP*.

———. In press. "Object Clitics in West Flemish." Halpern & Zwicky in press.

———, ed. To appear. *The New Comparative Syntax*. London: Longman.

Haegeman, Liliane & Henk van Riemsdijk. 1986. "Verb Projection Raising, Scope and the Typology of Rules Affecting Verbs." *Linguistic Inquiry* 17.417–466.

Haeringen, C. B. van. 1954. "Zijn en Wezen." *Taal en Tongval* 6.167–171. Reprinted in van Haeringen, 1962.

———. 1962. *Gramarie*. Assen: Van Gorcum.

Hale, Kenneth & Samuel J. Keyser, eds. 1993. *The View from Building 20: Essays in linguistics in honor of Sylvain Bromberger.* Cambridge, Mass.: MIT Press.

———. 1993. "On Argument Structure and the Lexical Expression of Syntactic Relations." Hale & Keyser 1993.53–109.

Halle, Morris & Alec Marantz. 1993. "Distributed Morphology and the Pieces of Inflection." Hale & Keyser, ed. 1993.111–176.

Halpern, Aaron & Arnold Zwicky, eds. In press. *Approaching Second: Second Position Clitics and Related Phenomena*. Stanford: CSLI.

Harada, Kazuko I. 1972. "Constraints on *wh*-Q Binding." *Studies in Descriptive and Applied Linguistics* 5.180–206.

Hedlund, Cecilia & Anders Holmberg, eds. 1994. *Papers from the Workshop on Scandinavian Syntax.* University of Göteborg, Göteborg.

Hellan, Lars. 1991. "A Two-Level X-bar System." *Lingua* 84.239–257.

Hellan, Lars & Kirsti Koch Christensen, eds. 1986. *Topics in Scandinavian Syntax.* Dordrecht: Reidel.

Hewitt, J. N. B. 1903. "Iroquoian Cosmology." *21st Annual Report of the Bureau of American Ethnology*. Washington, D. C.: Smithsonian Institution.

Hoekstra, Eric. 1991. *Licensing Conditions on Phrase Structure.* Doctoral dissertation, University of Groningen. [Distributed as vol. 2 of Groningen Dissertations in Linguistics.]

Hoekstra, Eric & C. Jan-Wouter Zwart. 1994. "De structuur van de CP: Functionele projecties voor topics en vraagwoorden in het Nederlands." *Spektator* 23.191–212.

Hoekstra, Teun. 1984. *Transitivity*. Dordrecht: Foris.

———. 1988. "Small Clause Results." *Lingua* 74.101–139.

Holmberg, Anders. 1986. "Word Order and Syntactic Features in the Scandinavian Languages and English." Ph.D. dissertation, University of Stockholm.

———. 1991. "The Distribution of Scandinavian Weak Pronouns." van Riemsdijk & Rizzi 1991.155–173.

Holmberg, Anders & Christer Platzack. 1991. "On the Role of Inflection in Scandinavian Syntax." Abraham, Kosmeijer & Reuland 1991.93–118.

———. 1995. *The Role of Inflection in Scandinavian Syntax.* New York: Oxford University Press.

Hoop, Helen de. 1992. "Case Configuration and Noun Phrase Interpretation." Ph.D. dissertation, University of Groningen.

Hopper, Paul & Sandra Thompson. 1980. "Transitivity in Grammar and Discourse." *Language* 56.251–299.

Hornstein, Norbert. 1990. "Verb Raising in Icelandic Infinitives." *NELS* 20.215–229.

———. 1994. "An Argument for Minimalism: The case of antecedent-contained deletion." *Linguistic Inquiry* 25.455–480.

———. 1995. *Logical Form: From GB to minimalism.* Oxford: Basil Blackwell.

Huybregts, Riny. 1992. "Scattering: A derivational approach without S-structure." Paper presented at the 8th Workshop on Comparative Germanic Syntax, Tromsø, November 21.

Iatridou, Sabine. 1990. "About AgrP." *Linguistic Inquiry* 21.551–576.

Jackendoff, Ray. 1977. *X′-Syntax: A study of phrase structure.* Cambridge, Mass.: MIT Press.

Jacobs, Joachim, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann, eds. 1993. *Syntax: An international handbook of contemporary research.* Berlin: Walter de Gruyter.

Jacobs, Neil, Ellen Prince & Johan van der Auwera. 1994. "Yiddish." König & van der Auwera 1994.388–419.

Jaeggli, Osvaldo & Kenneth Safir, eds. 1989. *The Null Subject Parameter.* Dordrecht: Kluwer.

Jaspers, Danny *et al.*, eds. 1989. *Sentential Complementation and the Lexicon.* Dordrecht: Foris.

Jespersen, Otto. 1927. *A Modern English Grammar on Historical Principles.* Copenhagen: Ejnar Munksgaard.

Johnson, Kyle. 1991. "Object Positions." *Natural Language and Linguistic Theory* 9.577–632.

Johnson, Kyle & Sten Vikner. 1994. "The Position of the Verb in Scandinavian Infinitives: V^0 or C^0 but not in I^0." *Working Papers in Scandinavian Syntax* 53.61–84.

Jonas, Dianne. 1992. "Checking Theory and Nominative Case in Icelandic." *Harvard Working Papers in Linguistics* 1.175–196.

———. 1994. "The TP Parameter in Scandinavian Syntax." Hedlund & Holmberg 1994.33–60.

———. 1995. "Clause Structure and Verbal Syntax on Scandinavian and English." Ph.D. dissertation, Harvard University, Cambridge, Mass.

———. This volume. "Clause Structure, Expletives and Verb Movement."

Jonas, Dianne & Jonathan D. Bobaljik. 1993. "Specs for Subjects: The role of TP in Icelandic." Bobaljik & Phillips 1993.59–98.

Jónsson, Jóhannes Gísli. 1991. "Stylistic Fronting in Icelandic." *Working Papers in Scandinavian Syntax* 48.1–44.

———. 1995. "Constituent Structure and Case in Icelandic." Ph.D. dissertation, University of Massachusetts, Amherst.

Josefsson, Gunlög. 1992. "Object Shift and Weak Pronominals in Swedish." *Working Papers in Scandinavian Syntax* 49.59–94.

Kaan, Edith. 1992. "A Minimalist Approach to Extraposition of CP and Verb Projection Raising." Gilbers & Looyenga 1992.169–79.

Karlsson, Fred, ed. 1983. *Papers from the Seventh Scandinavian Conference of Linguistics.* University of Helsinki, Helsinki.

Kayne, Richard. 1984. *Connectedness and Binary Branching.* Dordrecht: Foris.

———. 1989a. "Facets of Romance Past Participle Agreement." Benincà 1989.85–103.

———. 1989b. "Notes on English Agreement." *CIEFL Bulletin* 1.40–67.

———. 1989c. "Null Subjects and Clitic Climbing." Jaeggli & Safir 1989.239–261. Dordrecht: Kluwer.

———. 1991. "Romance Clitics, Verb Movement and PRO." *Linguistic Inquiry* 22.647–686.

———. 1993. "Toward a Modular Theory of Auxiliary Selection." *Studia Linguistica* 47.3–31.

———. 1994. *The Antisymmetry of Syntax.* Cambridge, Mass.: MIT Press.

Kim, Soo-Yeon. In progress. *The Interaction of Structural and Non-Structural Constraints on Binding and Scrambling.* Ph.D. dissertation, Harvard University, Cambridge, Mass.

Kitagawa, Yoshihisa. 1986. "Subjects in Japanese and English." Ph.D. dissertation, University of Massachusetts, Amherst.

Kitahara, Hisatsugu. 1992. "Checking Theory and Scope Interpretation without Quantifier Raising." Thráinsson & Kuno 1992.51–71.

———. 1993. "Deducing 'Superiority' Effects from the Shortest Chain Requirement." Thráinsson, Epstein & Kuno 1993.

———. 1994. "Target-α: A Unified Theory of Movement and Structure Building." Ph.D. dissertation, Harvard University, Cambridge, Mass.

———. 1995. "Target-α: Deducing Strict Cyclicity from Principles of Derivational Economy." *Linguistic Inquiry* 26.47–77.

———. 1996. *Elementary Operations and Optimal Derivations.* Cambridge, Mass.: MIT Press.

Kitahara, Hisatsugu & Ruriko Kawashima. 1994. "Strict Cyclicity, Linear Ordering, and Derivational C-Command." *Proceedings of the West Coast Conference on Formal Linguistics* 14. Stanford: CSLI.

König, Ekkehard & Johan van der Auwera, eds. 1994. *The Germanic Languages.* London: Routledge.

Koopman, Hilda. 1995. "On Verbs that Fail to Undergo Verb Second." *Linquistic Inquiry* 26.137–163.

Koopman, Hilda & Dominique Sportiche. 1991. "The Position of Subjects." *Lingua* 85.211–258.

Kosmeijer, Wim. 1986. "The Status of the Finite Inflection in Icelandic and Swedish." *Working Papers in Scandinavian Syntax* 26. University of Lund.

———. 1993. "Barriers and Licensing." Ph.D. dissertation, University of Groningen. [Distributed as vol. 9 of *Groningen Dissertations in Linguistics*.].

Koster, Jan. 1975. "Dutch as an SOV Language." *Linguistic Analysis* 1.111–136.

———. 1978. *Locality Principles in Syntax.* Dordrecht: Foris.

———. 1986. "The Relation between Pro-drop, Scrambling, and Verb Movements." Ms., University of Groningen.

———. 1987. *Domains and Dynasties: The radical autonomy of syntax.* Dordrecht: Foris.

———. 1995. "Predicate Incorporation and the Word Order of Dutch." Cinque *et al.*, 1995.255–277.

Kroch, Anthony. 1989. "Asymmetries in Long Distance Extraction in a Tree Adjoining Grammar." In: Mark Baltin & Anthony Kroch, eds. *Alternative Conceptions of Phrase Structure,* pp. 66–98. Chicago: The University of Chicago Press.

Kroch, Anthony & Aravind Joshi. 1985. "The Linguistic Relevance of Tree Adjoining Grammar." *Technical report MS-CIS-85–16*, Department of Computer and Informational Sciences, University of Pennsylvania.

Kuroda, Shige-Yuki. 1988. "Whether We Agree or Not." *Lingvisticae Investigationes* 12.1–47. [Also Poser 1988.103–143.]

Laka, Itziar. 1990. "Negation in Syntax: On the nature of functional categories and projections." Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge.

Lange, Klaus-Peter. 1981. "Warum Ersatzinfinitiv?" *Groninger Arbeiten zur germanistischen Linguistik* 19.62–81.

Larson, Richard. 1988. "On the Double Object Construction." *Linguistic Inquiry* 19.335–391.

Lasnik, Howard. 1993. "Lectures on Minimalist Syntax." Occasional papers 1. *Working Papers in Linguistics*, University of Connecticut. [Distributed by *MIT Working Papers in Linguistics.*]

Lasnik, Howard & Joseph J. Kupin. 1977. "A Restrictive Theory of Transformational Grammar." *Theoretical Linguistics* 4.173–196.

Lasnik, Howard & Mamoru Saito. 1992. *Move α: Conditions on its application and output.* Cambridge, Mass.: MIT Press.

Lasnik, Howard & Juan Uriagereka. 1988. *A Course in GB Syntax.* Cambridge, Mass.: MIT Press.

Lebeaux, David. 1988. *Language Acquisition and the Form of the Grammar.* Ph.D. dissertation, University of Massachusetts, Amherst.

Lehmann, Sabine. In preparation. *Clitic climbing in the Jura dialect.* Mémoire DES. University of Geneva.

Li, Yafei. 1990. "X^0–Binding and Verb Incorporation." *Linguistic Inquiry* 21.399–426.

Lightfoot, David & Norbert Hornstein, eds. 1994. *Verb Movement.* Cambridge: Cambridge University Press.

Loey, A. van. 1964. *Schönfelds Historische Grammatica van het Nederlands.* Zutphen: Thieme.

Mahajan, Anoop. 1990. "The A/$\overline{\text{A}}$-distinction and Movement Theory." Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge.

Maling, Joan. 1988. "Variations on a Theme: Existential sentences in Swedish and Icelandic." *McGill Working Papers in Linguistics: Special issue on comparative Germanic syntax*, 168–191. McGill University, Montreal.

———. 1990. "Inversion in Embedded Clauses." Maling & Zaenen. 1990. [Originally published as Maling 1980 "Inversion in Embedded Clauses in Modern Icelandic." *Islenskt Mál* 2.175–193].

Maling, Joan & Annie Zaenen, eds. 1990. *The Syntax of Modern Icelandic* [*Syntax and Semantics*, 24]. San Diego: Academic Press.

Manzini, Maria Rita. 1992. *Locality: A Theory and Some of Its Consequences.* Cambridge: MIT Press.

Marantz, Alec & Morris Halle. 1993. "Distributed Morphology and the Pieces of Inflection." Hale & Keyser 1993.111–176.

May, Robert. 1977. "The grammar of Quantification." Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge.

———. 1985. *Logical Form: Its structure and derivation.* Cambridge, Mass.: MIT Press.

McCloskey, James & Kenneth Hale. 1984. "On the Syntax of Person–Number Inflection in Modern Irish." *Natural Language and Linguistic Theory* 1.487–533.

McDaniel, Dana. 1989. "Partial and Multiple *Wh*-movement." *Natural Language and Linguistic Theory* 7.565–604.

McNally, Louise. 1992. "VP Coordination and the VP-internal Subject Hypothesis." *Linguistic Inquiry* 23.336–341.

Moro, Andrea. 1993. *I predicati nominali e la struttura della frase.* Padova: Unipress.

Mulder, René. 1991. "An Empty Head for Object Control." *Proceedings of NELS* 21.293–307. GLSA, University of Massachusetts, Amherst.

———. 1992. "The Aspectual Nature of Syntactic Complementation." Ph.D. dissertation, University of Leiden.

Müller, Gereon. 1993. "On Deriving Movement Type Asymmetries." Ph.D. dissertation, University of Tübingen.

Muysken, Pieter. 1982. "Parametrizing the Notion 'Head'." *Journal of Linguistic Research* 2.57–75.

Nash, Leonard. 1963. *The Nature of the Natural Sciences*. Boston: Little, Brown and Company.

Næss, Alma. In prep. [No title.] Ph.D. dissertation, University of Oslo.

Ó Siadhail, Mícheál. 1989. *Modern Irish: Grammatical structure and dialectal variation*. Cambridge: Cambridge University Press.

O'Grady, William, Michael Dobrovolsky & Mark Aronoff. 1993. *Contemporary Linguistics: An introduction.* New York: St. Martin's Press.

O'Neil, John. 1992. "Relativized Minimality and the A/$\overline{\text{A}}$-distinction." Kuno & Thráinsson 1992.143–153.

Ottósson, Kjartan. 1989. "VP-specifier Subjects and the CP/IP Distinction in Icelandic and Mainland Scandinavian." *Working Papers in Scandinavian Syntax* 44.89–100.

Ouhalla, Jamal. 1988. "The Syntax of Head Movement: A study of Berber." Ph.D. dissertation, University College London, London.

———. 1989. "Against the Head Movement Constraint." *University College London Working Papers* 1.193–205.

———. 1991. *Functional Categories and Parametric Variation*. London: Routledge.

———. 1994. "Verb Movement and Word Order in Arabic." Lightfoot & Hornstein 1994.41–72.

Pesetsky, David M. 1982. "Paths and Categories." Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge.

———. 1987. "*Wh*-in-situ: Movement and unselective binding." Reuland & ter Meulen 1987.98–129.

———. 1994. *Zero Syntax.* Cambridge, Mass.: MIT Press.

Phillips, Colin, ed., 1993. *Papers on Case and Agreement.* Vol. 2. *MIT Working Papers in Linguistics* 19.131–173. Massachusetts Institute of Technology, Cambridge.

Pica, Pierre & William Snyder. 1994. "Weak Crossover, Scope and Agreement in a Minimalist Framework." *Proceedings of the West Coast Conference on Formal Linguistics XIII*. Stanford: CSLI.

Platzack, Christer. 1983. "Existential Sentences in English, Swedish, German and Icelandic." Karlsson 1983.80–100.

———. 1986. "The Structure of Infinitive Clauses in Danish and Swedish." Dahl & Holmberg 1986.123–137.

———. 1988. "The Emergence of a Word Order Difference in Scandinavian Subordinate Clauses." *McGill Working Papers in Linguistics: Special issue on comparative germanic syntax* 215–238. McGill University, Montreal.

Platzack, Christer & Anders Holmberg. 1989. "The Role of AGR and Finiteness." *Working Papers in Scandinavian Syntax* 43.51–76.

Poeppel, David & Ken Wexler. 1993. "The Full Competence Hypothesis of Clause Structure in Early German." *Language* 69.1–33.

Pollock, Jean-Yves. 1989. "Verb movement, Universal Grammar and the Structure of IP." *Linguistic Inquiry* 20.365–424.

Poole, Geoffrey 1992. "The Case Filter and Stylistic Fronting in Icelandic." Kuno & Thráinsson 1992.19–31.

———. 1994. "On the Nature of Syntactic Adjunction." Ms., Harvard University, Cambridge, Mass.

———. Forthcoming. "Deducing the X'-Structure of Adjunction." In: Anna-Maria di Sciullo, ed. *Essays on X'-Structure and Adjunction*. Ithaca: Cascadilla Press.

Poser, William, ed. 1988. *Japanese syntax. Papers from the second international workshop on Japanese syntax*. Stanford: CSLI.

Postal, Paul. 1962. "Some Syntactic Rules of Mohawk." Ph.D. dissertation, Yale University. [Published New York 1979: Garland Press].

Postma, Gertjan. 1993. "The Syntax of the Morphological Defectivity of BE." *HIL Manuscripts* 1.31–67.

———. 1994. "The Interaction between Argument Structure and Perfect Tense." Ms., Leiden.

Radford, Andrew. 1990. *Syntactic Theory and the Acquisition of English Syntax*. Oxford: Basil Blackwell.

Reinhart, Tanya. 1976. "The Syntactic Domain of Anaphora." Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge.

———. 1991. "Elliptic Conjunctions—Nonquantificational LF." In Asa Kasher, ed. *The Chomskyan Turn,* pp. 360–384. Oxford: Basil Blackwell.

Reuland, Eric J. 1990. "Head Movement and the Relation between Morphology and Syntax." *Yearbook of Morphology* 3.129–161.

Reuland, Eric J. & Wim Kosmeijer. 1988. "Projecting Inflected Verbs." *Groninger Arbeiten zur germanistischen Linguistik* 29.

Reuland Eric J. & Alice G.B. ter Meulen, eds. 1987. *The Representation of Indefiniteness*. Cambridge, Mass.: MIT Press.

Riemsdijk, Henk van & Luigi Rizzi, eds. 1991. *Clitics and Their Hosts*. [ESF Working Papers] University of Tilburg.

Riemsdijk, Henk van & Edwin Williams. 1981. "NP Structure." *The Linguistic Review* 1.171–217.

Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.

———. 1990. *Relativized Minimality*. Cambridge, Mass.: MIT Press.
———. 1991a. "Residual Verb Second and the *wh*-criterion." *Technical Reports in Formal and Computational Linguistics* 3. University of Geneva.
———. 1991b. "Proper Head Government and the Definition of A-positions." Paper presented at GLOW 14, Leiden.
———. 1993. "Some Notes on Linguistic Theory and Language Development." Ms., University of Geneva, Geneva.
———. 1994. "Some Notes on Romance Cliticization." Ms, University of Geneva, Geneva.
Rizzi, Luigi & Ian G. Roberts. 1989. "Complex Inversion in French." *Probus* 1.1–33.
Roberts, Ian. 1985. "Agreement and the Development of English Modal Auxiliaries." *Natural Language and Linguistic Theory* 3.21–58.
———. 1993. *Verbs and Diachronic Syntax: A comparative history of English and French*. Dordrecht: Kluwer.
Rögnvaldsson, Eiríkur & Höskuldur Thráinsson. 1990. "On Icelandic Word Order Once More." Maling & Zaenen 1990.3–40.
Rohrbacher, Bernhard. 1993. "V-to-Agr Raising in Faroese." Ms., University of Massachusetts, Amherst.
———. 1994. "The Germanic VO Languages and the Full Paradigm: A theory of V-to-I raising." Ph.D. dissertation, University of Massachusetts, Amherst.
Rouveret, Alain, & Jean-Roger Vergnaud. 1980. "Specifying Reference to the Subject: French Causatives and Conditions on Representation." *Linguistic Inquiry* 11. 97–202.
Rutten, Jean. 1991. "Infinitival Complements and Auxiliaries." Amherst, Mass.: Graduate Linguistic Student Association, Department of Linguistics, Ph.D. dissertation, University of Amsterdam.
Safir, Kenneth. 1993. "Perception, Selection and Structural Economy." Ms., Rutgers University, New Brunswick.
Saito, Mamoru. 1989. "Scrambling as Semantically Vacuous $\overline{A}$-Movement." Baltin & Kroch 1989.
———. 1992. "Long Distance Scrambling in Japanese." *Journal of East Asian Linguistics* 1.69–118.
Samiian, Vida & Jeannette Schaeffer, eds. 1995. *Proceedings of the twenty-fourth Western Conference on Linguistics (WECOL 94)*. Department of Linguistics, California State University, Fresno.
Schuringa, Frans. 1923. *Het Dialect van de Veenkoloniën*. Groningen: Wolters.
Schutter, Georges de. 1974. "Wezen Vissen: Dialectgeografie van een constructie." *Taal en Tongval* 26.70–85 [Reprinted in Stroop 1983].
Shlonsky, Ur. In press. *Clause Structure and Word Order in Hebrew and Arabic: An essay in comparative semitic syntax*. New York: Oxford University Press.

Sigurðsson, Halldór Ármann. 1991. "Icelandic Case-marked PRO and the Licensing of Lexical Arguments." *Natural Language and Linguistic Theory* 9.327–364.
Speas, Margaret. 1990. *Phrase Structure in Natural Language*. Dordrecht: Kluwer.
Sportiche, Dominique. 1983. *Structural Invariance and Symmetry in Syntax*. Ph.D. dissertation, MIT.
Sportiche, Dominique. 1988. "A Theory of Floating Quantifiers and its Corollaries for Constituent Structure." *Linguistic Inquiry* 19.425–450.
———. 1992. "Clitics, Voice and Spec/Head Licensing." Paper presented at Glow Lisbon, May 13.
Sternefeld, Wolfgang. 1994. "Subjects, Adjuncts, and SOV-Order in Antisymmetric Syntax." Zwart, ed. 1994.227–246.
Stowell, Tim. 1981. *Origins of Phrase Structure*. Ph.D. dissertation, MIT.
———. 1983. "Subjects Across Categories." *The Linguistic Review* 2.285–312.
Stroop, Jan, ed. 1983. *Nederlands Dialectonderzoek*. Amsterdam: Huis aan de drie grachten.
Stuurman, Frits. 1985. *Phrase Structure Theory in Generative Grammar*. Dordrecht: Foris.
Tancredi, Christopher D. 1990. "Not only even, but even only." Ms., Massachusetts Institute of Technology, Cambridge.
Thráinsson, Höskuldur. 1986. "On Auxiliaries, AUX and VPs in Icelandic." Hellan & Koch Christensen 1986.235–265.
———. 1993a. "On the Structure of Infinitival Complements." *Harvard Working Papers in Linguistics* 3.181–213.
———. 1993b. "Functional Projections and Infinitives." Talk given at the University of Massachusetts. Amherst, October 1993.
———. 1994. "Comments on the Paper by Vikner." Lightfoot & Hornstein 1994.149–162.
———. This volume. "On the Non-Universality of Functional Categories."
———. In prep. "Functional Projections and Infinitives." Ms., Harvard University, Cambridge, Mass.
Thráinsson, Höskuldur, Samuel D. Epstein & Susumu Kuno, eds. 1993. *Harvard Working Papers in Linguistics* 3. Harvard University, Cambridge, Mass.
Thráinsson, Höskuldur & Susumu Kuno, eds. 1992. *Harvard Working Papers in Linguistics*. Department of Linguistics, Harvard University, Cambridge, Mass.
Travis, Lisa deMena. 1984. "Parameters and Effects of Word Order Variation." Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge.
———. 1991. "Parameters of Phrase Structure and Verb-Second Phenomena." Freidin 1991.339–364.
Trudgill, Peter. 1990. *The Dialects of England*. Oxford: Basil Blackwell.

Ura, Hiroyuki. 1993. "L-relatedness and its Parametric Variation." *MIT Working Papers in Linguistics* 19.377–399.

Vangsnes, Øystein. 1995. "Referentiality and Argument Positions in Icelandic." *Working Papers in Scandinavian Syntax* 55.89–109.

Vergnaud, Jean-Roger. 1982. *Dépendances et Niveaux de Representation en Syntaxe.* Ph.D. dissertation, University of Paris VII. [Published by John Benjamins, Amsterdam, 1985.]

Verkuyl, Henk. 1987. "Nondurative Closure of Events." Groenendijk, de Jongh & Stokhof 1987.87–113.

Vikner, Sten. 1990. "Verb Movement and the Licensing of NP-positions in the Germanic Languages." Ph.D. dissertation, University of Geneva [Abridged version published as Vikner 1995a].

———. 1991. "Verb Movement and the Licensing of NP-Positions in the Germanic Languages" Ph.D. dissertation, University of Geneva. [Revised Version, Institut für Linguistik, Universität Stuttgart].

———. 1994. "Finite Verb Movement in Scandinavian Embedded Clauses." Lightfoot & Hornstein 1994.117–147.

———. 1995a. *Verb Movement and Expletive Subjects in the Germanic Languages.* New York: Oxford University Press.

———. 1995b. "V^0–to-I^0 and Inflection for Person in All Tenses." *Working Papers in Scandinavian Syntax* 55.1–27.

Vikner, Sten & Bonnie D. Schwartz. 1995. "The Verb Always Leaves IP in V2 Clauses." Belletti & Rizzi 1995.

Visser, Fredericus Theodorus. 1973. *An Historical Syntax of the English Language* (Vol. II). Leiden: Brill.

Watanabe, Akira. 1992. "*Wh*-in-situ, Subjacency and Chain Formation." *MIT Occasional Papers in Linguistics* 2, Department of Linguistics and Philosophy, Massachusetts Institute of Technology, Cambridge.

———. 1994. "Conceptual Basis of Cyclicity." Ms., University of Tokyo.

Webelhuth, Gert. 1992. *Principles and Parameters of Syntactic Saturation.* New York: Oxford University Press.

———, ed. 1995. *Government and Binding Theory and the Minimalist Program.* Oxford: Basil Blackwell.

Weijnen, Antonius. 1966. *Nederlandse Dialectkunde.* Assen: Van Gorcum.

Weinberg, Amy. 1994. "Comments on the Paper by Wexler." Lightfoot & Hornstein 1994.351–362.

Wexler, Ken. 1994. "Optional Infinitives, Head Movement and the Economy of Derivations." Lightfoot & Hornstein 1994.305–350.

Wilder, Christopher & Damir Ćavar. 1993. "Word Order Variation, Verb Movement, and Economy Principles." Ms., University of Frankfurt am Main.

Williams, Edwin. 1987. "Introduction." In: Thomas Roeper & Edwin Williams, eds. *Parameter Setting*, pp. vii-xix. Dordrecht: Reidel.

———. 1988. "Is LF Distinct from S-structure? A Reply to May." *Linguistic Inquiry* 19: 135–146.

Woodbury, Hanni. 1975. "Noun Incorporation in Onondaga" Ph.D. dissertation, Yale University.

Wyngaerd, Guido Vanden. 1989a. "Object Shift as an A-movement Rule." *MIT Working Papers in Linguistics* 11.256–271.

———. 1989b. "Verb Projection Raising and the Status of Infinitival Complements." Jaspers *et al.* 1989.423–438.

———. 1994a. "IPP and the Structure of Participles." *Groninger Arbeiten zur Germanistischen Linguistik* 37.265–276. University of Groningen, Groningen.

———. 1994b. "Partikelwerkwoorden in het Nederlands." *Tabu* 24.141–148.

———. 1995. "Gapping Constituents." Ms., KU Brussel.

Wyngaerd, Guido Vanden & Jan-Wouter Zwart. 1991. "Reconstruction and Vehicle Change." Drijkoningen & van Kemenade 1991.151–160.

Zagona, Karen. 1982. "Government and Proper Government of Verbal Projections." Ph.D. dissertation, University of Washington, Seattle.

Zanuttini, Raffaella. 1991. "Syntactic Properties of Sentential Negation: A comparative study of Romance languages." Ph.D. Dissertation, University of Pennsylvania.

Zwart, C. Jan-Wouter. 1992a. "Dutch Expletives and Small Clause Predicate Raising." *NELS* 22.477–491.

———. 1992b. "Matching". Gilbers & Looyenga 1992.349–361.

———. 1993a. "Verb Movement and Complementizer Agreement." Bobaljik & Phillips 1993.297–340.

———. 1993b. "Dutch Syntax. A Minimalist Approach." Ph.D. dissertation, University of Groningen. [Distributed as vol. 10 of *Groningen Dissertations in Linguistics*].

———. 1994a. "Dutch is Head Initial." *The Linguistic Review* 11.377–406.

———. 1994b. "On 'Holmberg's Generalization'." de Boer, de Hoop & de Swart 1994.

———. This volume. "'Shortest Move' versus 'Fewest Steps'."

———. To appear. "The Germanic SOV Languages and the Universal Base Hypothesis". Haegeman to appear.

———, ed. 1994. *Minimalism and Kayne's Asymmetry Hypothesis* [*Groninger Arbeiten zur Germanistischen Linguistik*, 37]. Department of German, University of Groningen.

Åfarli, Tor A. 1992. *The Syntax of Norwegian Passive Constructions.* Amsterdam: John Benjamins.

Index

Aa

A- vs. Ā-
  dichotomy, 18, 20, 53f.
  chain, 102
  movement, 46f., 149ff.
  position, 41f., 53f., 102, 142ff.
Ā-
  chain, 202
  movement, semantically vacuous, 200, 214ff.
acquisition
  of functional categories, 255, 279
  of phonology, 259
adequacy
  , descriptive (empirical coverage), 3
  , explanatory 3
adjunction, 24f., 100*fn.*, 101, 117, 122f., 286, 321
  as a restricted option in Minimal Theory, 32
  as a specifier, 32*fn.*
  of sentential (or "medial") adverb, 264
  to arguments, 129
  to functional projections and all single-bar level categories, 264
  to specifiers, 25
  to T, 176
  to VP (and other semantically active maximal projections), 171, 264
  under LCA, 25
  vs. substitution, 31f.
  , multiple, 26
adverb, 161
  intervening between subject and finite verb, 210*fn. 17*
  , VP-adjoined manner, 172, 178, 230
  , sentential (or "medial"), 264
affix
  hopping, 12
  lowering, 237
Agr (= Agreement (category))
  as a collection of φ-features (person, number, and gender), 18
  , eliminated, 262
  ; free Agr parameter, 273*fn. 17*

Agr-based theory, 95
of feature checking, 93
Agr-less ("bare") theory, 94, 253 ff.
agreement, 18, 221, 263
and Case features, 176
distinctions, 231
markers, overt, 269
morpheme, 258, 266
paradigms, robustness of, 220*fn.*
relations, 9 f.
, rich verb
and split IP, 267
and overt verb movement, 267
and tense, 268
, sufficiently, 267 f., 279
, separate
and tense, 253
, Spec–head, 247
, subject–passive participle, 237
, subject–verb, 264 f.
AgrO, 12, 36, 93, 218, 220
, movement to T, 173
AgrP, 263 ff.
AgrS, 12, 36, 218, 220, 315
and T as non-separate functional heads, 273
AgrSP
and TP, 279
over TP, 270
allomorphy, 225
American English, dialectal, 233
antipassive, 290
antisymmetric, 70, 92
VO-perspective, 72
Antisymmetry Hypothesis (Kayne 1994), 217
A–P (= Articulatory–Perceptual) system, 7, 15
Arabic, 243*fn. 44*, 270
argument
structure, 284, 289
, bare, 283 ff.
, DP-, 149
, ergative, 284
, three zones for, 148 ff.
, unergative, 284
aspect, 221, 227, 233, 247 f., 262, 290
inflection, 248 f.
aspectual participle, 235
asymmetry, 35, 126, 149
attract, 93 f.
attraction of elements, 37 f.
auxiliary, 233
selection, 300 ff.
verb, 275

Bb

bare
indefinite, 184
-output conditions, 8 f.
phrase structure (Chomsky 1995), 283 f., 303
T(ense), 177
theory (= Agr-less theory, multiple Spec; Chomsky 1995), 94 f.
base structure, 323
base-generation, 131, 135
, optional, 135, 136
BE + BE, 302
binary operation, 9 f.
binder, potential $\overline{A}$-, 191*fn. 3*
binding
condition, proper (= PBC), 214
requirement, minimal, 191
theory, 302
bridge verb
complement, 264
, non-, 173
building representation, 7 ff.
Bulgarian, 236

Cc

C/Comp (complementizer)
categorial feature, 9
c(ategorial)
-command, 19f., 25f., 33, 103, 191*fn. 2*, 256, 286
, asymmetric, 24, 25f., 33f.
, morphological evidence for, relation, 270
Case, 14, 39
and agreement, dissociation of, 14, 173
assignment, 22*fn. 16*, 72, 110
, structural, 18
checking
at LF (weak checking features), 115
in Spec–head configuration, 22*fn. 16*, 278
of objects, 278
feature, 18
of subject, 167
filter, 39, 110
and the visibility condition, 97f., 110
in ergative languages, 290f.
in sisterhood relation, 278
position, 273
marking in Spec-Agr configuration, 18
of objects, 260
of subjects, 259
reduced to feature checking, 22*fn. 16*
theory, 110, 197
, abstract, 39
, feature
in Spec–head configuration, 22*fn. 16*
, null, 91
, object
checking, 97ff.
, structural, 263
category
, type, 52ff., 123
, null, 91f.
, phonological, 125
Celtic, 224*fn. 14*
chain, 115, 180, 224
algorithm, 189, 193, 194, 195
-based theory of scope interpretation, 189, 192, 195
-formation, 82f., 115, 124, 306
structure, 194f.
, conditions on, 311
checker
, intervening, 98f.
checking
agreement features, 222, 266
Case feature in Spec-TP, 173
D-feature of Spec-AgrSP, 181
domain, 15f., 19, 99ff., 276
EPP-feature, 183
features, 11ff., 38ff., 152f., 315
in a cyclic fashion, 258
N-feature, 177
plural agreement features, 173
, closest
position, 276*fn. 20*
V-features in T, 266
, object Case, 97ff., 104, 111
, feature, 11ff.
and form chain, 204
CL. *See* clitic
clause
architecture, 255f., 259, 277
asymmetry, 245
structure, 18, 167ff., 170
, basic, 17f., 99, 254
-initial position, 181
clitic (= CL), 153ff.
as DP, 158
as Comp-agreement, 239*fn.*
cluster, 163f.
movement, 156ff.
theory, 13
, free-standing, 155

clitic (*Continued*)
, order relative to
subject, 245
verb, 245
, subject, 243 f.
cliticization, 278
as a form of inflection, 245
, subject, 241
clitics
in Dutch, 241
in Romance, 154
in West Flemish, 156
cluster formation (clustering), 163, 292, 296
cognitive system (embracing the syntactic component), 20
Comp (= complement(izer))
domain, 9 f., 19 f., 135
feature, 244 f.46
hosting T, 312
complement
, propositional, 316
, raising, 316
complementizer (= Comp), 12, 227, 241
agreement, 242
particle, 226 (Irish), 245
computational system (= syntax) 3, 113 ff., 217
concatenate, 7 f.
C–I (Conceptual–Intentional) system, 7 f.
consonantal mutation, 225
contain relation, 101
contained in, 9 f.
contentful element, 9 f.
converge, 15 f., 20
convergent derivation, 114
Copy Theory of movement, 215
covert *ge-* (prefix of past participle, in German and Dutch), 300
movement, 261, 320
syntax, 16 f.
verb raising, 176
CP, complement(izer) phrase
, structure of, 253 ff.
CP-projection, 18
-raising, 292
crash (at PF, at LF), 20, 22, 99
cycle/cyclicity
, strong, 121 f.
, strict, 41 f., 50 ff.

Dd

Danish, 265 f.
D(efinite)
argument, 320
-feature, 171, 179
-linked arguments vs. non D-linked
D(eep)-structure, 4 f., 6
default assumption, 270
form, 221, 242
setting of the Strict Inflection Parameter (SIP), 267
definiteness feature, 171, 179
deletion
(operation), 207
under identity, 16*fn. 11*
dependency
relation, 318
, shortest, 11 ff.
derivational, 13 f.
morphology, 222
optionality, 199 f.
syntax, 17
directionality, 17 ff., 62 ff.
and word order, 47 f.
parameter, 72
distributed morphology, 279
do, 200 ff., 237 f.
domain 19 f., 101, 135, 207
-extending head movement, 69 f., 93 f.
extension, 133, 207 f.

of chain, 69*fn.*
of the main verb, 275
, checking, 101
, complement, 101
, internal, 101, 135
, minimal, 69*fn.*
, residual/residue, 15f., 101
dominance, dominate, 34, 101
, reflexive, 207
, immediate, 123
double object, 82
constructions, 77, 84, 325
double infinitive construction (= IPP), 296ff.
Dutch, 104, 238, 241ff., 246, 248, 295, 315f.
and Aux-selection, 302
, head-initial, 238f.

Ee

ECM. *See* Exceptional Case Marking
economy, 118f., 305
and restrictions, 11ff.
condition, 15ff., 34f., 43ff.
, local, 116ff.
constraint, 67f.
of derivation, 34f., 200ff., 307
of representation, 261, 306f., 309, 313, 320, 323, 326f.
ECP. *See* Extended Projection Principle
ELCA. *See* Extended Linear Correspondence Axiom
elimination, 12f.
of features, 15, 310, 326
of N-feature of AgrS, 176
emphatic
assertion, 237f.
-assertive sentence, 232
Empty Category Principle, (=ECM), 311
empty element, 9f.
en- (passive morpheme) 148
enclisis, 244*fn. 47*
English
and Aux-selection, 302
and IPP, 302
tense paradigms, 269
without NP-OS (-object shift), 278
EPP. *See* Extended Projection Principle
equidistance, 67f., 94, 97f., 102f., 105, 109ff., 133, 168, 172, 180, 212f., 275f., 308ff., 324f.
to Spec, 213
, movement and, 179, 212f.
Equidistance Principle, 310, 325*fn. 21*
ergative
language and case distribution, 290
VPs, 284
Ersatzinfinitiv (German). *See* I(nfinitivus) P(ro) P(articipio)
Exceptional Case Marking (ECM), 18
complement, 184
construction, 188
, non-finite, 174
explanation through minimization, 4
expletive, 87, 167ff., 187, 272, 314, 317
as a small clause, 317
construction, 186
in Mainland Scandinavian, 179
with "extraposed" subjects, 188
in Spec-AgrSP, 171
passive, 181ff.
, overt. *See* overt, expletive
, passive, 181, 188, 273
, raising, 188, 273
, transitive construction (=TEC). *See* transitive ergative construction
Extended Projection Principle (= EPP), 88, 91, 176, 176*fn. 13*, 276
feature, 167
-checking, 188

Extended Linear Correspondence Axiom (= ELCA), 19f., 24ff.
extension
condition (*see also* Strict Cycle Condition), 5, 10, 17, 27f., 276, 320
requirement, 115ff.
extraposition, 292
of subjects, 186

Ff

Faroese, 168, 174*fn.*, 187
FC. *See* functional category
feature, 10, 17
checking, 11ff., 30f., 189, 192ff., 197, 314
at LF (weak features) vs. at PF (strong features), 100
elimination, 16, 22
-relativized approach to minimality, 80
strength, 22, 175
, strong ("offensive"), 22, 124, 219ff.
, theory of, 189, 192–197
, weak ("inoffensive"), 22, 124, 219ff.
fewer morphological distinctions, 279
fewest steps, 34ff., 48ff., 305ff.
FI. *See* full interpretation
filter
, descriptive universal, 8
, disjunctive, 110
finite verb, 258
, movement of
to T or to AgrS, 265
finite complementizer, 260
Fiorentino, 243*fn. 44*
Flemish, 241, 301f.
focus, 243
form chain, 29f., 42f., 49f., 91, 119, 125, 199ff., 306ff., 319ff., 327
rather than Move-α, 201
FPA. *See* Functional Projection Alibi
French, 104, 125, 230, 247*fn. 51*
Frisian, 299
fronting
, stylistic (= SF), 200, 208f.
full interpretation (= FI), 7, 12f., 15, 306*fn. 1,* 326f.
functional
category (= FC), 3, 9ff., 28f., 36f., 44ff., 57ff., 159ff., 263ff., 274–278
checking Case, 260
head, 9f., 71, 136, 218
projection, 219ff.
Σ (sigma), 258f., 278
structure, 260
, incorporation of, 255
, acquisition of, 259ff.
, inventory of, 255
, incorporation
into auxiliary, 76*fn.*
of, 75f.
, inventory of, 255
, non V-related, 162
, sequence (c-command relations) of 257, 270, 278
, sequence of
and projections, 256
, sequence (c-command relations), 255
, set of, 255
Functional Projection Alibi (= FPA) 261f.
fused
agreement (AgrSP) and tense (TP)
AgrSP and TP, 253
IP morphology, 253

Gg

GB, Government and Binding (= Chomsky 1983), 141ff., 222
ge- (past participle prefix, Dutch and German), 249*fn. 53*, 285, 287, 295, 298
, non-occurrence of, 300

generalized transformation (= GT), 4f., 8f., 121, 319, 327
Gent dialect, 303
Georgian, 290
German, 109*fn.*, 238, 241, 248, 295, 299*fn. 12*, 322
, head-initial, 238f.
government
proper, 17f., 311
directionality of, 22
Government and Binding model, 3f.
Greed, 48f., 99, 116, 119f., 203ff., 211, 307, 313, 317f., 321, 326
vs. Economy of Derivation, 204*fn.*
Groningen dialect, 299
GT. *See* generalized transformation

Hh

H^{max} (maximal head), 19f.
habitual, 234, 235
head, 9f., 21f.
initiality (Kayne's Antisymmetry Hypothesis), 23f., 63, 238f.
movement, 18, 136, 225, 238, 306, 310, 323, 326
chain, 69
, overt, 59
constraint (= HMC), 97, 102, 111, 155, 310, 324
(reducible to ECP, 311)
effects, 163
of T to AgrS, independent functional, 181
-final language, 124, 246ff.
-skipping head movement, 79*fn.*
-to-head movement, 162
, domain of, 19
headedness parameter (directionality), 48, 62f., 246ff.
Hindi, 104
HMC. *See* head, movement, constraint
Holmberg's Generalization, 35f., 46ff., 64, 70, 70, 94, 97f., 109*fn.*, 132, 212, 275*fn.*

Ii

Icelandic, 104, 108, 131, 138, 178ff., 200, 208ff., 265, 269
V-to-T in non-finite clauses, 138, 183. *See also* stylistic fronting; transitive expletive construction
illegitimate object, 17f., 103
include/inclusion 117
incorporation, 97, 105ff., 110
of object, 98, 105f., 108, 111
of predicate, 316
indefinite, 170
inert or missing, 257
AgrO projection, 186
infinitival
complements, 316f.
construction, 181ff.
marker in AgrS, 266
infinitive, 248, 297, 311
Infinitivus pro Participio (= IPP; = German *Ersatzinfinitiv*), 27f., 285, 292f., 296ff., 301
Infl(ection(al)), 217, 222
categories (AgrS, T and AgrO), 238
structure, articulated, 170
, split, 18
, unsplit, 253, 271, 275, 278
inflected
verb, 218
word, morphological analysis of, 217
inflection (phrase = IP), 217, 222
inflectional
affix, 237
category (IP)
, fused, 279f.
, split, 279f.
ending, 287

inflectional (*Continued*)
morphology, 16, 221
, poor, 279
, presence vs. absence of, 217ff.
paradigm, 221
Insular Scandinavian (vs. Mainland Scandinavian), 173
intermediate
position, 167, 188
projection, 21f.
internal
domain, 15f., 135
Subject Hypothesis (subject base-generated in Spec-VP), 98f.
interpretable features, 16f., 280
interpretation
, conditions on, 318
interrogative
, embedded, 321
intervening position, 98
intransitive (= unergative)
verb, 257*fn. 2*
, true, 186
invisibility, 17f. *See also* visibility condition
IP. *See also* Inflection Phrase
, Spec-, 68
, split, 254ff.
IPP. *See* Infinitivus pro Participio (= German *Ersatzinfinitiv*)
Irish, 222f., 225*fn. 16*, 242f.
irregular verb, 232f.
island
effects, 50f.
, strong vs. weak, 320
Italian, 222f., 245

Jj

Japanese, 127, 200, 214ff.

Kk

knowability
of natural language 2
, contribution of organism 2
, contribution of environment 2

Ll

L. *See* lexical
L-feature, 19
movement, 324
positions, 145
-related, 14f., 20, 53f., 145ff., 221*fn. 5*
, broadly vs. narrowly, 20, 42f.
, non-, 20
movement, 324
position, 145
landing site
, potential, 34, 45f.
Last Resort Principle, 206
LCA *See* Linear Corresponding Axiom
Last Resort Principle, 12f., 15, 206
LDH. *See* Limited Diversity Hypothesis
learnability problem (*see also* acquisition), 259
left-branching, 297
legitimate object, 12f., 15f.
level of representation, 5f.
lexical
category, independent, 259
subject in Spec-TP, 272
verb, 229
lexicalist checking approach to morphology, 288
Lexicalist
Approach, 39f.
to morphology, strong, 258
Mirror Principle (LMP), 258,*fn. 5*, 269f., 278ff.
lexicon, 9, 115, 118
LF. *See* Logical Form

licensing, 10f., 30f., 38ff.
of a lexical subject, 177
light verb (projection) (= *v*, *v*P), 93
Limited Diversity Hypothesis (= LDH), 257ff., 270, 280
linear
ordering, 18f.
precedence (Kayne's Antisymmetry Hypothesis), 24f.
Linear Corresponding Axiom (= LCA), 18f., 24, 63*fn. 34*, 284, 297f.
linguistic theory, central goals of, 1f.
link
, minimal, 67
condition, 12f., 91
constraint, 67
local
domain, 44f.
locality condition, 12, 44f., 224, 318
as final stage of a derivation, 21
interface, 21, 121
Logical Form (= LF)
interface, 113ff.
-movement, 261
object, legitimate, 205
"reconstruction", 204ff.
LMP. *See* Lexicalist Mirror Principle
Low German, 299
lowering, 218

Mm

Mainland Scandinavian (= MSc), 107, 169, 267
manner adverbs, 178
match, 10f.3
maximal
projection, 9f., 21f.
medial
adverb, 171
Merge, 5, 7ff., 22f., 27f., 30, 36, 284
bottom-up, 28
Middle English, 230f.
Middle High German, 246*fn. 50*
minimal
binding requirement, 191
contrast, 226
domain, 15f., 324
link condition/constraint (= shortest derivation), 12f., 15, 67f., 91
constraint, 67
S (= Min(S)), 207
Minimalist
analysis, 73
program, 1ff., 98f., 189, 192, 217
minimality, 34f., 136, 307
effects, 221
, feature-relativized, 79
, relativized, 34f., 97f., 101f., 107, 305
minimization
(of axioms), method of, 4
of levels of representation, 4f.
Mirror Principle (= MP), 28f., 56ff., 217ff., 256ff., 270
Modern English, 231, 232
mood
and MoodP, 221, 227, 230f., 262
indicator (e.g. +*wh*), 18
morphological
agreement dependent on morphological tense, 269
features, 218, 258, 263
of person and number, 258
, strength of, 218
evidence for the independence of agreement and tense morphemes, 270
strength, 43f., 46ff., 55ff., 60ff.
variation, 217
morphologically-driven word-order predictions, 217
morphology
, independent tense and agreement, 269
, overt, 57ff.

morphology (*Continued*)
, poor vs. rich
, inflectional, 279
, verbal, 265, 268 ff.
Move-α, 200 ff., 225
, filter-based, 17
movement, 4 f., 7 f., 9, 22 f., 28, 218, 224 f., 319
across a potential landing site, 68 ff.
into the functional domain, 23
motivated by the need to check features, 99 f.
of finite and non-finite verbs, 175, 264
of functional heads, 84
targets, equidistant, 94 f.
to license and check, 9 ff.
, cause of, 266
, chain forming (vs. non-chain forming), 200 ff.
, checking approach to, 83
, copy theory of, 215
, covert
preferred over overt (= Procrastination), 309
, cyclic, 322
, direction of, 22, 23 ff.
, domain-extending, 69
, feature-driven, 82 ff.
, head, 225
parameterization, 219
, non-L-related, 324
, object
dependent on verb movement, 309
to AgrOP, 246
, optional, 199 ff.
, relativization, types of, 97, 102 f., 107, 109, 111
, restrictions on, 15 ff., 43 ff.
, semantically vacuous $\overline{A}$-, 200
, shortest, 97, 101 f., 111, 136
requirement, 47 f., 97 f., 102 f., 105 ff., 111
, successive cyclic, 37, 50 f., 318
, verb
to Infl (to AgrS), 239
to T, 239
MPLT "Minimalist Program for Linguistic Theory" (= Chomsky 1993) 116 ff.
MSc. *See* Mainland Scandinavian, 184 f.
multicategorial
status of word, 223, 225
multiple
specifier, 32 *fn. 20*, 88
analysis, 273
positions, 92
-Spec theory, 93
subject/specifier constructions, 92
successor, 95

Nn

N-feature, 11 f., 133, 147, 277
Neg-operator, 144 ff.
NegP. *See* Negation Phrase
negation, 132, 230, 237, 263, 267 *fn. 12*
Negation Phrase (NegP), 78 f., 148, 230, 238
negative concord, 78
NI (= noun incorporation), 106 f.
nominative Case
assignment, 221
checking, 193
non-branching condition, 288
non-L-related. *See* L-related
NP (= nominal phrase)
movement
, successive cyclic, 75 *fn.*
, case driven, 71
-object shift (in MSc), 274, 277 f.
null
Case, 91

category, phonological, 125
operator, 125
numeration, 35

Oo

object
case checking, 104 ff.
clitic, 153
control, 291
incorporation, 105 ff.
movement, 36 f., 246, 308
raising
, covert, 186 f.
, overt, 46
shift (= OS)
70, 97 f., 104 f., 131 f., 135, 108, 111, 274 f., 277 f.
, pronominal, 277 *fn.*
and verb raising (*see also* Holmberg's Generalization), 212
at LF, 196 *fn.*
split, 84
IP (vs. fused), 279
, double construction, 83 f.
, full NP- (NP-OS), 170, 274
, overt, 179
, pronominal, 277, 278
to Spec-AgrOP, 168
and Spec-TP as an available A-position, 187
Onondaga, 105 *fn.*
operator, 144
features, 144
, null, 125, 130
Optional Movement, 199 ff.
optionality, 48 ff., 199 ff.
, observed, 162
order
of agreement projections, 325
, hierarchical, 270
ordering constraint, rigid, 150
organization of the grammar, 64 ff.
OS. *See* object shift
OV (= object–verb) order (vs. VO), 71 ff.,, 246, 294
overt
agreement
markers, 268 f.
expletive in Spec-AgrSP, 27
morphological evidence, 279
morphology, 44 ff., 278 f.
object shift (= overt OS) 179
order, 44 f.
syntax, 16 f., 22, 40 f.
verb movement
to AgrS, 176
to T, 168, 176
verb raising, 175

Pp

paradigm, verbal, 226
parametric variation, 44 ff., 55 f., 57 ff.
among languages, 278
, universality and, 15 ff.
parasitic gap, 152
participle, 233, 246, 283 ff.
, prefixal, 299
, prefixless, 303
, structure of, 285 ff.
particle
verb, 291
, inseparable, 295
passive
expletive construction, 181, 188, 273
past
morpheme *-d*, 287
participle, 232 f., 248, 285, 301
agreement, 156
PBC (= proper binding condition) 214
perfective, 234 f., 289 f.
performance systems (articulatory–perceptual, conceptual–intentional), 16 f., 20

person distinctions, 268
PF (= Phonetic Form, the output of Spell-out), 219
-object, legitimate, 20
φ (=phi)-features, 9, 144, 263
, lack of
in Agr, 263*fn.*
, semantic correlates of, 280
Phonetic Form (PF), 15f., 21
phonetic inventories, 259
phonetically realized, 5f.
phonological
attachment, 241
feature 50f., 258
filter, descriptive, 6f.
rules, 259
phrase
level, 29
structure, 24f., 40f., 52ff.
, building up, 10ff., 27ff.
, relational definition of, 30
in Kayne's sense, 32ff.
, head initial, 24f.
position
, types of, 146f.
potential landing site, 34f.
PP arguments, 149
pre-Spell-Out, 50f.
precedence relation, 19f., 26
predicate
condition, 197
phrase (= PredP), 249, 295
, incorporated (Dutch/German *ver-*), 316
PredP. *See* predicate, phrase
prefixal participle, 299
prefixless participle, 303
present tense, 231ff.
Principles and Parameters Approach, 2f., 48, 56f.
PRO, 125
infinitival, 182
Procrastinate (Procrastination Principle), 12f., 16f., 21, 49f., 115, 118f., 307, 309, 318, 326
productivity, 233
progressive, 235
projection, 8f., 23f., 29ff., 286
line, 21f.
, non-branching, 105*fn. 9*
pronominal object shift, 277f.
proper
government, 17f., 311
part, 117

Qq

Q. *See also* question; quantifier
(= question) morpheme (in Japanese), 214
quantifier
raising (= QR), 31f., 40f., 189ff., 191, 192, 195, 197
scope interactions, 82
QR. *See* quantifier, raising

Rr

raising, 71ff., 218, 315
complements
as VP, 184
as TP, 178
expletive constructions, 174, 188, 273
of tense to AgrS, 167, 184
to Spec-AgrSP at LF, 173
, non-verb
language (= English), 176
, string vacuous (VP-final V to IP-final I), 178
Real Minimalist Principle (= RMP), 261, 267
reduction to independent conditions, 6
reference (= R-) projection, 165

regular verb, 232
relative tense, 227
relativization of movement types, 98
relativized
minimality, 34f., 44f., 51, 97ff., 101f., 107, 305
relativity, 107
RelTP (= relative tense phrase), 228
remnant topicalization, 81
representation
abstract of sound (PF), 5
of meaning (LF), 5
, building of, 8ff.
, conditions on, 327
, levels of, 3f.
rewrite rule (Chomsky 1970), 28
right
-adjunction, 71
-branching, 297
rigid ordering constraint, 150
RMP (= Real Minimalist Principle), 261
Romance, 154, 301
clitics, 154, 162ff.
rule application, 16

Ss

S-structure, 218
, elimination of, 6
Scandinavian, 182ff.
schijnen (Dutch 'seem')
selecting a propositional complement, 315f.
scope, 40, 68, 81f., 85, 189ff.
interactions, 82
in double-object, 83
interpretation, 190*fn. 1*, 191
, narrow vs. wide, 195*fn.*
Scope Principle, 82, 191f.
, 192*fn. 5, 6*, 194
, revised, 194ff.
scrambling, 36f., 71, 78f., 80f., 85, 87, 89, 93, 147
as optional movement, 214
in Continental West Germanic, 109*fn.*
second position
, clausal, 169, 174
seem, 314
segment, 25f., 26*fn. 18*, 123, 286
segmental theory, 26*fn.*
selection, 314ff.
separable, 295
sequence
(c-command relations) of functional categories, 255, 257, 270, 278
of functional heads and projections is determined by UG, 256
SF (= stylistic fronting), 209,*fn. 16*
shortest
derivation (= Minimal Link Condition), 11ff., 15, 306ff.
move, 33f., 97ff., 101f., 168, 172, 275ff., 305ff.
movement condition/requirement (= SMR), 67f., 101ff., 107f., 277
, naturalness of, 103ff.
steps (requirement), 34f., 44ff.
movement and equidistance, 179
Σ (=sigma)-features, 238
single output syntax, 273
singulary, 9f.
SIP. *See* Strict Infl(ection) Parameter
sister relation, 121, 284
small clause, 237*fn. 36*, 250, 294f., 306, 315f.
as a raising complement 316
SMR. *See* shortest, movement requirement
Spanish, 302
Spec. *See also* specifier
of (= Spec-)
AgrOP, 68, 135, 146f., 180
AgrSP, 146f.

Spec (*Continued*)
CP, 128f., 146f.
IP, 68, 143
NegP, 146f.
TP, 146f., 180
as an intermediate position, 177
available as Case, 167
available as a subject Case-checking position at LF, 185
Spec–head relation
and telicity, 247
before and after after Spell-out, 12*fn. 4*
for Case-checking, 39ff.
; agreement is with immediate projection of head
specifier. *See also* Spec
as an adjunct (sister of XP), 32*fn. 20*
skipping NP-movement and domain-extending head movement, 69
, multiple, 32*fn. 20*
specificity interpretation, 153
Spell-out, 5–7, 16f., 21, 99f., 113ff., 118, 121, 218f.
of a D-feature, 179
, pre-, 50f., 65f., 118
, post-, 65f., 118
split
clitic, 163
IP, 253, 262, 272, 278f.
Parameter (=SIP), 57f., 262
, evidence for positive setting of, 275
, morphological evidence for positive setting of, 268
, positive value for, 263, 267, 279
strong
cycle, 121, 123
feature, 17f., 266
vs. weak, 124, 219
island, 320
N-feature, 176
Strong Lexicalist Approach Hypothesis, 222ff., 244
of functional heads, 266
to morphology, 258
structural Case, 263
Structural Uniformity Hypothesis (= SUH), 255f.
structure building process (see: generalized transformation), 8ff., 319
stylistic fronting, 200, 209f., 213f.
subcategorization, 288
subjacency
effects, 66
in LF, 127
violation vs. ECP violation, 203*fn. 6*
subject
agreement features, 221, 258
and object agreement, 263
Case-checking, 167, 176
checking
at LF, 185f.
of the N-features in T, 179
clitic, 242f.
following finite auxiliary, 181
gap, 209
hypothesis, VP-internal, 263
in base position (= Spec-VP), 37, 87
in Icelandic TECs in Spec-TP, 178
in Spec-TP, 42*fn. 25, 26*, 42f.
in TEC, 178
movement to Spec-TP, 91
position
state
selected by raising verb, 316
strict cycle
condition (= extension condition), 5, 10, 17, 27f., 65, 115, 117, 123, 172, 270, 320, 327
strict cyclicity (= Extension Condition), 320
Strict Inflection Principle (= SIP), 279f.

in expletive constructions, 178 ff.
, range of, 187
, two, 273
pronoun cliticization, 245
raising, covert, 187
-verb agreement, 262, 264
, dummy, 314
, expletives with "extraposed", 188
, extraposed, 186
, heavy
in final position, 185 f.
, VP-Internal
Hypothesis, 74, 190 ff., 197, 237, 263
subjunctive, 231
substitution, 24 f., 31, 39 f., 117
successive cyclic movement, 252, 255, 318, 321
suffixation, 233
superiority, 97 ff., 305
superraising, 97 ff., 102, 306, 313 ff., 323, 326
suppletion, 225
Swedish, 183, 272, 301
syntactic
categories, 259
, inventory of, 268
, syntactic evidence for, 267

Tt

T (= tense), 12, 134
as possessor of strong null case, 91
higher than AgrS, 218, 256
in VSO higher than in SVO, 256
projecting specifier, 183
raising to AgrS, 175
-to-AgrS raising, independent, 187, 271
, independent raising of, 175 f.
, relative (= RelTP), 228
TEC (= transitive expletive construction) (Dutch, Icelandic, West Flemish), 89, 168 f., 177 *fn.*, 187
, passive, 169 f.
transitive expletive construction (= TEC), 188
telicity, 227, 246
and definiteness (specificity), 246 f.
TelP (= telicity phrase), 247
temporal modification, 79, 90 ff.
tense (= T), 36, 230, 238,
and agreement
morphemes, fused, 279
morphology, independent, 269
feature, 235, 237
morpheme, 231
"inside" the agreement morpheme 270
, relative, 227
, separate morphemes for and agreement, 280
theta (Θ) criterion, 6
relation, 19
role, 135
assignment at LF, 15
, internal, 135
violation, 206
third construction, 293
Tiwa, 107
topic, 243
topicalization remnant, 81
tough-construction, 6
TP (= tense phrase)
complementation, 70 f., 78 ff., 86
, Spec-, 85, 91, 263 ff.
trace, 124
; copy, 315
transformation
, binary and singulary, 5, 114, 121

transformation (*Continued*)
, generalized (= binary substitution operation), 5, 10, 13, 115, 121, 201, 319
transitive
expletive construction (TEC), 33f., 40-ff., 71, 86ff., 88f., 91, 93, 167, 188, 272ff.
VP-raising construction, 92
verb, 68, 179
vs. unergative, 187
transitivity
, high and passivizability, 290
transparency, 292
Trentino, 244*fn. 44*, 245
trigger
verb movement to Comp, 37f., 311

Uu

unaccusative (= ergative), 105*fn.*, 186
unergative (= intransitive), 284
UG, *see* Universal Grammar
Universal Grammar (= UG), 1ff.
universality
and parametric variation, 15ff., 53ff.
of basic clause structure, 254ff.

Vv

v, *v*P. *See* light verb (projection)
V. *See* verb
V1 (= modal verb in V-cluster), 73ff., 77, 95
V2
main lexical verb in V-cluster, 73ff., 77, 95
verb second, as in Dutch and German in main clauses, 146f.
verb (= V)
–clitic order, 244f.
-feature, 11f., 19, 131, 147
of T, 181f.
lacking an external argument, 179
movement, 148, 167, 174, 246, 262ff., 266, 276, 308
and subject Case-checking, interaction of, 178
chain, 182
in complement of (non-)bridge verb, 173f., 264
in embedded clauses in Faroese, 174
in non-finite complement clauses, 181
in Scandinavian, 173
in V2 clauses, 173
to AgrO, 173
to Comp, 244*fn. 46*
triggering of, 311
to I, independent, 173, 264
to T, 168
chain, 136
in finite clauses in MSc, absence of, 182
overt, 168
, non-finite
in Icelandic, 174
, lack of in Mainland Scandinavian, 266
, obligatory in infinitival control complements, 266
, overt, 57*fn.*
to T, 168
(projection) raising, 63f., 67ff., 71, 73ff., 108, 120, 125, 285, 291
, lack of in English
, object control (and verb raising), 294
, overt, 175
-second language, 261
, irregular, 232
, overt, 57ff.
to T, 168

, particle, 291f., 294
, pronoun enclisis on, 244
, raising, 289*fn. 6,* 291*fn. 7*
, regular, 231f.
verbal
agreement and tense, 268
complex, 292
morphology
and split IP, 267
and overt verb movement, 267
and adjunction to AgrS, 171
differences in, 265
, loss of rich, 268
, poor, 265
, rich, 267f.
, sufficiently rich, 267f., 279
paradigm, 226
visibility(/visible), 17f.
condition (for theta marking), 110
VO (verb–object) order. *See* head, initiality
VP (= verb phrase)
-adjoined manner adverb, 178
-Internal Subject Hypothesis, 74, 190ff., 197, 237, 263
VP1 (= modal verb cluster), 73ff.
VP2 (= main verb cluster), 73ff.
VPR (= verb projection raising), 72f., 91
VR (= verb raising) 72ff.

Ww

weak
feature, 17f.
island, 322
V-features, 176
vs. strong morphological features, 217, 229
weest + infinitive construction (West Flemish), 302
Welsh, 225*fn. 16*, 242
West Flemish, 71ff., 78, 86f., 142ff., 146ff., 239f.
West Germanic, 238, 245, 247, 250
head-finalness, 246
wh (also: Q-element)
-criterion, 144, 239
-(question) elements, 306
-element in CP, 37, 214
-feature, 247, 322
in situ (in Japanese), 214
-island, 101, 318ff., 322
-movement, 21f., 127, 318, 323, 326
, long distance, 321
, non-local, 323
operator, 127f., 144
, overt 196*fn.*
, partial, 321
, dummy, 321
, weak, 130
word order
and directionality, 23f., 62ff.
variation (see: optional movement), 217ff.
, overt (see: feature checking), 60ff., 218
, covert (see: feature checking), 218

Xx

X-bar (= X/X′)
schema, 11
Theory, 28
, concept of, 29f.
X′-level
, intermediate, 31

Yy

yes/no question, 312
Yiddish, 299*fn. 11* (German)

In the series LINGUISTIK AKTUELL/LINGUISTICS TODAY (LA) the following titles have been published thus far, or are scheduled for publication:

1. KLAPPENBACH, Ruth (1911-1977): *Studien zur Modernen Deutschen Lexikographie. Auswahl aus den Lexikographischen Arbeiten von Ruth Klappenbach, erweitert um drei Beiträge von Helene Malige-Klappenbach.* 1980.
2. EHLICH, Konrad & Jochen REHBEIN: *Augenkommunikation. Methodenreflexion und Beispielanalyse.* 1982.
3. ABRAHAM, Werner (ed.): *On the Formal Syntax of the Westgermania. Papers from the 3rd Groningen Grammar Talks (3e Groninger Grammatikgespräche), Groningen, January 1981.* 1983.
4. ABRAHAM, Werner & Sjaak De MEIJ (eds): *Topic, Focus and Configurationality.Papers from the 6th Groningen Grammar Talks, Groningen, 1984.* 1986.
5. GREWENDORF, Günther and Wolfgang STERNEFELD (eds): *Scrambling and Barriers.* 1990.
6. BHATT, Christa, Elisabeth LÖBEL and Claudia SCHMIDT (eds): *Syntactic Phrase Structure Phenomena in Noun Phrases and Sentences.* 1989.
7. ÅFARLI, Tor A.: *The Syntax of Norwegian Passive Constructions.* 1992.
8. FANSELOW, Gisbert (ed.): *The Parametrization of Universal Grammar.* 1993.
9. GELDEREN, Elly van: *The Rise of Functional Categories.* 1993.
10. CINQUE, Guglielmo and Guiliana GIUSTI (eds): *Advances in Roumanian Linguistics.* 1995.
11. LUTZ, Uli and Jürgen PAFEL (eds): *On Extraction and Extraposition in German.* 1995.
12. ABRAHAM, W., S. EPSTEIN, H. THRÁINSSON and C.J.W. ZWART (eds): *Minimal Ideas. Linguistic studies in the minimalist framework.* 1996.

Tina Fry
43 Haddington Road
Whitley Bay
NE25 9XE

The articles in this volume are inspired by the Minimalist Program first outlined in Chomsky's MIT Fall term class lectures of 1991, and his seminal paper "A Minimalist Program for Linguistic Theory" (1993). They seek to develop further some key idea in the Minimalist Program, sometimes in ways deviating from the course taken by Chomsky. Reference is taken also to Chomsky's "Bare Phrase Structure" (1995a, first distributed 1994), and to the basic ideas in Kayne's book, *The Antisymmetry of Syntax* (1994, distributed in part in manuscript form in 1993). The volume starts off with a thorough introduction into the minimalist framework with special attention to the question how the minimalist framework developed out of the Principles and Parameters (Government and Binding) framework.

ISBN 90 272 2733 0 (Eur.) / 1-55619-231-2 (US)