Lexicon and Grammar: The English Syntacticon

Studies in Generative Grammar 50

Editors

Henk van Riemsdijk
Harry van der Hulst
Jan Koster

Mouton de Gruyter
Berlin · New York

Lexicon and Grammar: The English Syntacticon

by

Joseph E. Emonds

Mouton de Gruyter
Berlin · New York 2000

Mouton de Gruyter (formerly Mouton, The Hague)
is a Division of Walter de Gruyter GmbH & Co. KG, Berlin.

The series Studies in Generative Grammar was formerly published by
Foris Publications Holland.

∞ Printed on acid-free paper which falls within the guidelines
of the ANSI to ensure permanence and durability.

Die Deutsche Bibliothek – Cataloging-in-Publication Data

Emonds, Joseph E.:
Lexicon and grammar : the English syntacticon / by Joseph E. Emonds. –
Berlin ; New York : Mouton de Gruyter, 2000
(Studies in generative grammar ; 50)
ISBN 3-11-016981-9 brosch.
ISBN 3-11-016689-5 Gb.

Printing: Werner Hildebrand, Berlin.
Binding: Lüderitz & Bauer GmbH, Berlin.

For Henk van Riemsdijk
Scientist, scholar, bon vivant,
research organizer and fellow-traveller

Preface

This book focuses on the urgent need for a *formal, constrained* and *empirically revealing* theory of a syntactic lexicon. To satisfy this need, it proposes new principles regulating subcategorization which determine how syntactic structures project from a language's lexicon. In the theory developed here, the trees projected by lexical subcategorization frames are not always copies of the frames themselves, as they are in the first classical proposals of Chomsky's (1965) *Aspects of the Theory of Syntax*. This study's highly restrictive sub-theories of grammatical categories and features, abstract case, derivational levels and economy principles have the consequence that sisterhood between a selecting and selected element is only subcategorization's "simplest case."

The data and paradigms here are for the most part drawn from English, but properties of Romance languages also form an essential part of the argumentation, and some points are discussed in terms of constructions from yet other languages, especially Japanese.

This study defends a strictly syntactic approach to the lexicon, i.e. it elaborates a theory of c-selection (subcategorization) and argues against the use of any thematic grids or lexical conceptual structures in grammatical computation. Constructions which have been widely invoked as necessarily involving semantic selection, such as the *spray/ load* alternation, propositional complements, light verbs and understood arguments are shown to be better analyzed without it (cf. especially Chapters 2, 6 and 9). A central organizing factor for this approach to the lexicon is a crucial distinction between an item's "cognitive syntactic" features F used in syntactic derivations and its "purely semantic" features *f* which are not (Chapter 1). The first use of the former (Chapter 2) sharpens the theory of c-selection by using these features in lexical frames (e.g., +___ANIMATE rather than +___DP and +___PATH rather than +___PP).

The principal innovation based on the F/*f* distinction is the proposal in Chapter 3 that the lexicon consists of two quite different components, a grammatical lexicon bereft of purely semantic features (the "Syntacticon") and a mental lexicon which consists of the open classes of the more specified contentful lexical items (the "Dictionary," which is the faculty of human linguistic memory and culture). There are only four categories in the Dictionary (N, V, A and P), what I term "nature's bottleneck" in a final

discussion in Chapter 10. Dictionary items are always inserted at the outset of transformational computations on a domain, as in Chomsky (1965). The perennial question "why do transformations exist?" is answered as follows: to assemble sets of disparate open class items in structures which can be communicated at the Phonological Form interface (PF) and interpreted at the Logical Form interface (LF).

In contrast to the Dictionary, the Syntacticon is regulated by a theory of *multi-level lexical insertion*: the feature composition of a lexical item in the Syntacticon determines at which stage in a derivation it is inserted, i.e., satisfies its c-selection properties. I take it that sweeping statements about levels of lexical insertion are empty unless tied to predictions about particular items satisfying specific insertion contexts. Chapters 4 and 5 explore the predictive consequences of this study's theory of multi-level insertion for the Syntacticon's bound morphemes and Chapters 6 and 7 for its free morphemes. In the limiting case, entirely uninterpreted Syntacticon items are inserted at PF; in particular, the entire class of items traditionally called inflections are characterized as "late-inserted" in this way. Inflections which apparently contribute to meaning (past tense, noun plurals, etc.) in fact serve only to license empty nodes in other interpretable free morpheme positions (I, D).

While developing the theory of the Syntacticon and multi-level insertion, I argue not only against semantic selection but also against any autonomous "morphological component" with combinatorial properties, in the lexicon or elsewhere. I claim the only statements needed for morphology are those with phonological effects, possibly conditioned by syntactic factors. The entire content of morphology is thus akin to interface statements such as a rule of Classical Greek, "verbs and adjectives receive penultimate stress." Except for the effects of such statements, the combinatory principles of bound morphology are exactly the same as those governing the syntax of free morphemes, especially those of compounding (Chapter 3). Supposed differences between compounding and morphology in e.g. Romance languages are shown to be differences between free and bound morphemes.

The notion that some categories in trees are associated with lexical items in the course of derivations raises the question of how these categories act prior to such insertion. Many analyses in this book demonstrate that a category in a head position *does not act like a head until it is lexically filled.* This principle regulating underlying empty heads is introduced in Chapter 4 and used throughout the rest of the study. Since grammatical elements can

be inserted at later derivational levels, the empirical properties of particular grammatical elements are typically due to (i) their delayed status as head of a construction and/ or (ii) whether they are interpreted in Logical Form. Simple and uniform lexical specifications leading to insertion of single morphemes at two or three derivational levels provide novel explanations for many previously poorly understood complex grammatical patterns, for example in both the passive and perfect periphrastic constructions of Germanic and Romance (Chapter 5).

In Chapter 6, Syntacticon entries of the form X, +___Y are shown to induce syntactic "flat structures" if both X and Y are the same category and X lacks purely semantic features *f* (i.e., is subject to late lexical insertion). This peculiar conjunction of properties provides analyses which solve many recalcitrant syntactic puzzles of especially the Romance languages. Causative, restructuring, linking and light verbs, which have all been treated differently in the literature, actually realize very similar structures. The explanations all crucially exploit the notion of "empty underlying head." Another area which demonstrates the explanatory power of syntactically empty heads is a range of PPs headed by grammatical P, including adjuncts as well as complements (Chapter 7).

Chapter 8 builds to what is in some way the intellectual climax of the book. It begins by extending case theory and refining certain formal properties of subcategorization in a way that fully defines, in conjunction with the category and feature theory of Chapter 1, the notion of "possible syntactic part of a lexical entry." *This chapter demonstrates that quite general and familiar classes of complements in the English lexicon instantiate all and only the structures predicted by the category and subcategorization theories of this book.* This reformalization of the notion "lexical entry," after twenty years of promissory notes in the ever novel coinages of theta theory, semantic selection, conceptual structures and the like, finally proposes limits imposed by Universal Grammar on the construct "possible lexical entry." This step removes the vagueness that has been associated with the lexicon for decades and enables us to embed the study of syntactic derivations in a fully generative system, in which all modules are formalized.

The last two chapters pursue another hypothesis that privileges syntax over any active role for the lexicon in language use. These chapters argue that all understood arguments are represented syntactically; a case is made against various notions of "unprojected arguments" which have appeared in the literature. In particular, Chapter 9 justifies unexceptional, purely syntactic representations for "null complement anaphora" and "null generic objects," and derives obligatory PRO as a by-product of subcategorization for

a V-headed rather than I- or C-headed complement. Finally, Chapter 10 argues for a discourse-governed reference of optional PRO subjects, which turn out to have a broader distribution than usually envisioned and include subjects of imperatives and agent phrases in verbal passives. All understood arguments are thus syntactically present arguments.

A central theme in this book's approach to lexical theory is that contextual features in individual entries must be *formally simple, uniform in format,* and stated in terms of *psychologically accessible categories* – i.e., categories which classify the concrete words of PF, not the abstract phrases of LF; as real as these latter are, they are not present in the data used by the child. That is, the many detailed properties of complex constructions (e.g., participles, nominalizations, passives) must derive from syntactically combining extremely transparent lexical entries. These entries can contain no diacritics or purely formal features. The motivation for this is simply a classic Chomskyan argument from poverty of the stimulus. What individual children do in learning a specific language is acquire a hoard of lexical entries for "small" grammatical items, both bound and free. Now we observe that not only some but essentially *every one of these learned grammatical items differs syntactically* from its closest counterparts in closely related languages. For example, if there is a handful of grammatical free or bound English morphemes lexically specified as exact translations of French ones, they are exceptions which prove the rule (I have yet to encounter a single one). Given that pre-school or unschooled children learn hundreds of these language-particular grammatical items fast and well, their lexical representations must have a simple and uniform format, expressed in categories readily accessible in the linguistic data, i.e. subcategories of words.

On the other hand, the facts of rapid acquisition do not suggest that lexical insertion theory, presumably uniform across the species, has to itself be transparent, any more than transformational theory or phonological theory are transparent. The structure of the uniform lexical theory developed in this study is subject only to general considerations of parsimony, elegance and empirical coverage which we require of any scientific enterprise.

Curiously, most generative treatments of the lexicon, including those of Chomskyan inspiration, search for extreme simplicity just where it seems least likely to be found: areas of species-wide genetic predisposition which have taken eons to perfect are felt to be governed by simple statements such as "affect alpha." Yet in the area where poverty of the stimulus and rapid acquisition have some force, the largely non-genetic lexical lists, essentially any quasi-formalized linguistic or conceptual properties are attributed to

individual items and no principles regulate either lexical form or the lexicon-syntax interface. I do not see how this combination of abstract principles operating on unconstrained lexicons contributes to explaining acquisition of actual languages. Worse, the unformalized nature of the lexicon renders largely untestable any empirical claims made for the derivational component of UG (which depends on "projection from the lexicon").

The literature under the rubric of Head Driven Phrase Structure Grammar, which human time limitations have prevented me from seriously investigating, does nonetheless seem to have a failing in this area as well, though one that is not so maddening as the hand-waving approach to the lexicon widely practiced in transformational syntax. While HPSG pays careful attention to formalizing the lexicon, it seems to care less about constraining it. Practitioners seem to be complacent and satisfied that if it works (i.e. can be computationally modeled), it's (relatively at least) good enough. In my tentative excursions into this literature, I sense little concern with trying to formulate a theory of lexical entries all and only of whose possible instantiations are realized in the lexicon of a language or some collection of languages.

A potentially controversial psychological implication of this study is that language can express only those combinatory meanings which simply "arise" from valid syntactic combinations. In the view developed here, syntax determines entirely any *propositions* we can formulate about our mental world and yet is largely independent of and provides very little insight into how we otherwise conceptualize it. Moreover, a little reflection shows (see in particular the concluding sections of each of the last three chapters) that our mental world and the form of natural language are almost entirely incommensurable, at least in terms of our ability to consciously reflect on them. So for example, while models of propositions are always discrete and almost invariably two-dimensional, our conceptualizations are obviously continuous and three or four dimensional: *The room slowly filled with smoke.*

Thus, even if thematic relations between a verb and its arguments are just "convenient mnemonics for particularly prominent configurations" in conceptual structures (Jackendoff 1987: 385), a given verb's lexical conceptual structure, while psychologically applicable to the world, is still linguistically opaque and unlikely to correspond to any structure in a grammatical phrase marker outside the verb. Consequently, any pieces of conceptual structure attached to lexical items remain unanalyzable and are hence next to useless as a guide to more general knowledge about the interface of conceptual structure with syntax.

Acknowledgments

This book was entirely researched and written in the atmosphere of Generative Linguistics in the Old Worlds of Europe and Japan. I am very grateful to the intellectuals, institutions, cooks, house mates, and surgeon (Mr. Gaynor of North Tyneside General Hospital) who have welcomed me as an emigrant or perhaps vagrant, and kept me respectively stimulated, comfortably paid, well fed, entertained and reasonably free of pain during this work.

Especially the following friends have been involved with this work for sustained periods of time, supported the intellectual ideas throughout and integrated them into their own work: Henk van Riemsdijk, Miori Kubo, Lida Veselovská, Andrew Caink and Mi-Jeung Jo. These intellectual companions have given me the sense that the ideas presented here are a basis for fruitful research in areas where I previously had very little conception of how they could be. This enthusiasm for my approach has been an indispensable aid in helping me to persevere. They also seem to share an approach to linguistic analysis which values working out consequences over a period of time and maintaining some fidelity to the insights and analyses of the past; theirs is therefore not only a valued companionship of ideas, but also an even more treasured companionship of method. I am especially grateful to Lida for the careful reading and many comments she has provided for the final version.

I have been particularly fortunate in this last decade to have benefited from generous fellowships and visiting professorships arranged by some of the most effective and tireless organizational supporters of the generative enterprise. As always, these times have proved the most fruitful in terms of developing the material in this study. For providing proper working conditions and research and writing time and for many unexpected professional extras, I am therefore deeply indebted especially to the following friends and colleagues:

To Prof. Henk van Riemsdijk who invited me both as visiting professor to Tilburg in 1992 and as Fellow to the Netherlands Institute for Advanced Study in 1997. The project of extending Alternative Realization to Romance clitics began during my lectures at Tilburg, which set me on the road to using it to encompass all of inflection; the linguistic community at Tilburg was a real source of initial inspiration for this study. The excellent research atmosphere Henk and Co-director Martin Everaert created in the SYNCOM

project at NIAS enabled the writing of the book to get to the crucial halfway point. I very much benefited from the constant interchanges on syntax with my project colleagues there, Norbert Corver, Denis Delfitto, Sten Vikner and Martin. I am of course also indebted to the Institute and its Officers for the excellent conditions for collaborative research they provide.

I also wish to thank the doctoral students and colleagues in linguistics at the University of Durham, who over several years have developed a challenging and active research community. I especially appreciate the explicit and implicit encouragement of my research direction provided by my colleague Dr. Bonnie Schwartz and former supervisees Andrew Caink, Cécile de Cat, Dalina Kallulli and Roger Maylor. This linguistic community has been made possible by the efforts of Vice-Chancellor Evelyn Ebsworth, who gave the strongest possible support to our field during his tenure and encouraged me to take every opportunity to pursue this research both at Durham and when invited elsewhere. This support culminated with a Christopherson Research Fellowship for the academic year 1998-1999, which was indispensable in bringing the scholarship for this book to a timely conclusion.

The greatest part of this work was written and revised during my several invited stays in Japan since 1994. Consequently a major share of my gratitude goes to the figures of Japanese generative grammar who have provided a pattern of invitations and support for research and advanced teaching that I think are practically unparalleled in present-day linguistics. The first versions of some material were undertaken as revisions of the summer 1995 TEC lectures in Tokyo, given at the invitation of Professors K. Hasegawa, K. Inoue and M. Kajita, and the last revisions were made during my first semester at the new Kobe-Shoin Graduate School, organized by Dr. Taisuke Nishigauchi. Thus, this book began, came to maturity and has been completed in Japan, in settings very conducive to research.

The support of my long-time friend and wide-ranging intellectual model Prof. S.-Y. Kuroda has provided crucial links in several of these initiatives, and I earnestly thank him for his sustained efforts on my behalf.

Three chapters of this book were largely written while I was visiting professor at Kanda University of International Studies in Japan. The research atmosphere and facilities there were superb, and I am most grateful to its President, Prof. Kazuko Inoue, for integrating me into the scholarly community she has created. And not the least benefit of my three years at Kanda was the introduction it facilitated into the wider world of Japanese linguistics.

Three further chapters of this book were written while a Fellow of the Japanese Society for the Promotion of Science at Nanzan University. I am most grateful for this excellent arrangement, and especially to my host scientist Prof. Mamoru Saito at Nanzan. His successful efforts in obtaining this grant and his scientific consultation and encouragement during my stay were indispensable contributions.

During all these stays in Japan, Miori Kubo has provided me with feedback and commentary on several chapters and sections, as well as often being extremely helpful in technical matters.

I gratefully acknowledge permissions from Taishukan Press, Tokyo, to revise and reprint material from Emonds (1997) forming the bulk of Chapter 2 and from Kaitakusha Press (Tokyo) to revise and reprint material for Chapter 9 from my "Subcategorization as the Unique Source of Null Complements," in their 1999 volume *In Search of the Human Mind – a Festschrift for Prof. Kazuko Inoue*, M. Muraki and E. Iwamoto (eds.).

Finally, I am also grateful to Sean Burke, José Deulofeu, Shalom Lappin and Pat Waugh for their personal enthusiasm and collegial support, which have also sustained me at particular stages of this work.

In producing this book, stellar performances in absolutely central roles have been delivered by the series editor Henk van Riemsdijk, an anonymous referee for Mouton de Gruyter, the editor Ursula Kleinhenz, the technical editor Monika Wendland, and the tireless creator of the subject index, Lida Veselovská. Thanks to them, bringing this book to a physical incarnation has been a *relatively* painless project.

Author's academic biography

Born in 1940 in North Dakota; his father Joseph was a highway engineer for the Bureau of Indian Affairs in the American West, and his mother Margaret (Embley) was a primary school teacher. He attended high school in Milwaukee, Wisconsin, received a BA in mathematics from Loras College, Dubuque, Iowa (1962), and an MA in mathematics at the University of Kansas, Lawrence, Kansas (1964).

After a year of teaching at the U.S. Naval Academy, he studied for the PhD in linguistics at the Massachusetts Institute of Technology (1965–1970), including a final fellowship year at the newly inaugurated University of Illinois Center for Advanced Study.

An intermittent European career began at the Université de Paris VIII (Vincennes) in 1969–1970, after which he began his first permanent position at the University of California at Los Angeles (1970–1979). His son Peter was born in 1973. While still a novice at the joys of fatherhood, he lectured at Princeton University (1973–1974) and the Université de Paris VII (1976–1977), where a Guggenheim Fellowship permitted extending his stay. He also gave a summer school course at the International Christian University in Japan. The decade closed with a self-granted half-year respite from academia in London.

A second permanent position, including five years as chair, was at the University of Washington (1980–1991). Research and teaching leaves were spent at the Massachusetts Institute of Technology (1984), the Stanford Center for Advanced Study in the Behavioral Sciences (1985–1986), the Université de Provence (1987) and the Université de Paris VIII (1989). He was also co-director of the Dubrovnik summer school (1987) and made frequent academic visits to the Netherlands. The last four Seattle years coincided with experiencing high school a second and more exciting time.

Expatriation began in earnest with a teaching fellowship at Tilburg University (1992), followed by taking up the Chair of Linguistics and English Language at the University of Durham in 1992. During this period, he was active in H. van Riemsdijk's European Science Foundation Eurotyp Group 8 and visiting professor at Kanda University of International Studies from 1994 through 1996. After fellowships at the Netherlands Institute for Advanced Science in 1997 and at Nagoya's Nanzan University in 1998 (sponsored by the Japanese Society for the Promotion of Science), he accepted a professorship at Kobe Shoin University's new graduate school in 2000.

Table of Contents

Chapter 1
Categories and feature inventories of Universal Grammar

1.1 A theory and practice of well-formed lexical entries

1.1.1 Specifying the well-formed sentences

Generative grammar originally took and can still take as its goal, to formally specify all and only the grammatically well-formed strings of a language. "The fundamental aim in the linguistic analysis of a language L is to separate the *grammatical* sequences which are the sentences of L from the *ungrammatical* sequences which are not sentences of L and to study the structure of the grammatical sequences." Chomsky (1957, 13)[1] The earliest period focused on the form of phrase structure and transformational rules, which were independent of individual lexical items. Although the actual statements of many early generative rules in fact mentioned specific morphemes, their dependence on these items was accidental; the development of an appropriate formal theory was not concerned with lexical statements as such.

Subsequently, the "standard theory" period of generative grammar (from Chomsky's *Aspects of the Theory of Syntax* through his "Conditions on Transformations," 1965–1973) recognized the necessity of succinctly formalizing co-occurrence relations between lexical items and the word classes they appear with; e.g., some verbs are obligatorily transitive, others are optionally so, and still others are incompatible with object nouns. The burden of expressing these well-formedness conditions was placed on base rules and syntactic subcategorization, the latter much later renamed "c-

---

[1] "The grammar, then, is a device that (in particular) specifies the infinite set of well-formed sentences and assigns to each of these one or more structural descriptions. Perhaps we should call such a device a *generative grammar* to distinguish it from descriptive statements that merely present the inventory of elements that appear in structural descriptions, and their contextual variants." (Chomsky, 1964, 9).

There are well-formed strings not technically sentences, which should still be generated by the grammar: *Off the porch with you both! (*With you both off the porch!); Well, what a nice present! (*What a nice present, well!)* As observed in Banfield (1982, Ch. 1), these constructions are not generally embeddable. Perhaps for this reason, both traditional and generative grammar have concentrated on sentences. Nonetheless, the syntax of non-embeddable constructions is intriguing.

selection". Base rules constituted the "outer limit" on the complexity of lexical subcategorization.

In a later third period (say from "Conditions on Rules" through Mark Baker's *Incorporation*, 1974–1988), base rules were rightly replaced by universal conditions on category inventories and projections via a parsimonious bar notation. The residual content of such rules were to be expressed by a few parameters such as the head-initial/head-final parameter. This reduced role of base rules could have suggested expanding and refining the role of subcategorization. In fact, in studies of morphology and compounding, Lieber (1980, 1983) moved in this direction. With the demise of base rules, there also arose a (largely unacknowledged) need to limit the form of possible lexical entries.

Instead, in syntax proper, a disturbingly anti-theoretical option was suddenly almost universally embraced around 1980, one which tended to entirely discount any role for subcategorization. Most syntacticians abandoned work on the lexicon, but whether one worked on the lexicon or not, the field as a whole seemed to pin its hopes on a future theory of semantic selection or "s-selection". Additionally, any concern with the limits on "possible lexical entry" was deemed premature. It thus became widely accepted that deep structures (i.e., inputs to a syntactic derivation) were to be determined via s-selection and the projection principle. For example, this step is endorsed in Chomsky's *Lectures on Government and Binding* (1981) and *Barriers* (1986). That is, heads combine semantically with some complement(s) by "assigning theta roles" to them, somehow thereby projecting or "licensing" deep structures. A consensus that deep structures project from (unformalized, intuitive) lexical-semantic structures effectively replaced the previous formal syntactic theory; syntacticians focused their efforts on constraining possible steps in derivations and on finding well-formedness conditions on co-indexing in the outputs.

Oddly, mainstream syntax showed almost no interest in the very lexical semantic structure that was newly considered the basis of syntax; it found little use for articulated proposals, notably by Jackendoff (1987, 1990), for specifying well-formed semantic combinations of elements such as lexical heads and complements. Nor did later lexico-semantic formalizations of head-complement structure find much favor in analyses using the framework of *Lectures on Government and Binding*.[2] Nonetheless, a moment's

[2] Theories developed to answer the need within government and binding theory for a formal mechanism to generate deep structures include Zubizarreta (1987) and Grimshaw (1990). Revealingly, one finds few references to precise aspects of such lexical formalisms in theoretical syntactic research.

reflection about a generative model in which all syntactic structures project from the lexicon reveals the key role of characterizing "possible lexical entry" – of exactly determining and formalizing lexical properties.

Instead of pursuing this goal, influential syntactic studies after 1980 have focused almost exclusively on how to relate steps and levels in a derivation. Though such studies often allude to combinatorial devices (s-selection, feature checking, well-formed objects at Logical Form), their lexical specifications for the most part remain undefined and underexemplified, and so fail to advance understanding of the lexical base of syntactic structures. In fact, research in government and binding and minimalism (the framework of Chomsky, 1995) seems to have abandoned a goal of working out some type of formalized lexicon to actually generate the syntactic structures which then undergo the carefully studied derivations. Perhaps this testifies to an unspoken realization that lexicalist s-selection for describing co-occurrence has simply failed to improve on earlier and more restrictive syntactic formulations in capturing co-occurrence generalizations.[3]

Inescapably, if formal syntax has little to say about how syntactic categories and specific lexical items combine, but concentrates only on transforming structures *assumed* to be lexically well-formed, it is not generative grammar in the original sense. One might object that the original goal was misformulated or too ambitious. However, if a particular framework actually revealingly approaches the original goal, such objections become a priori and irrelevant. This study's purpose is to develop a theory of subcategorization which accurately captures the patterns and structural generalizations projected from the English lexicon (i.e., a theory of c-selection). Additionally, it succeeds (in Chapter 8) in imposing a very strict limit on the notion "possible lexical entry" and argues that this limit nicely corresponds to the observed range of lexical complementation.

By incorporating many results of generative grammar's third period as well as more recent concepts such as Principles of Economy, I hope that this study's deliberately syntactic approach to the lexicon and its syntactic projections will help re-instate the generative ideal, which studies not only derivations but also characterizes the formal objects to which they apply.

---

[3] This is my argument in Emonds (1991b). What is at issue is not the priorities of particular researchers, but why syntactic researchers as a (large) group feel no need to address the flagrant weakness of an unformalized lexicon, which undermines any theory's claim to be a formal model of language.

1.1.2 Judging the well-formed sentences

It seems appropriate to explain at the outset how my grammatical argumentation treats data for which native speakers make conflicting acceptability judgments.

Sometimes, conflicting judgments are clear-cut enough to justify different rule systems, parameter settings or, in terms to be developed in this study, different syntactic entries in the lexicon. For example, while many Standard American speakers clearly accept patterns with post-verbal particles separating double objects as in (1.1), others simply do not.

(1.1) Mary brought her kids out some clothes.
Some students may send companies back their free samples.
Write the customers up an order.

In constructions where comparative judgments are murkier, I follow a heuristic proposed by N. Chomsky in 1967 class lectures: if native speakers assess differently the acceptability of an example which has a clear meaning and presents no processing difficulty (such as center embedding, undue length, homonyms, ambiguity, etc.), *the example should be taken as ungrammatical*. The reason is that, by hypothesis, there is no reason for its partial rejection other than it being ungrammatical. In varying degrees, speakers obviously accept and process all sorts of ill-formed utterances (e.g., by children, foreigners, the speech-impaired, or speakers making performance errors) without the slightest consciousness of deviance. Thus, if no processing factors intervene, a reaction of acceptability cannot be trusted in the face of conflicting judgments, while one of unacceptability may be.[4] For this reason, the questionable sentences as in (1.2) should be analyzed as ungrammatical.

(1.2) ?John believes that the earth the moon circles weekly.
?Mike probably considers some professors fools and Ann many knaves.
What did other students { think that/ ?wonder whether } Mary bought?

---

[4] Perhaps facetiously, Chomsky also suggested that being more sensitive to unacceptability might be an evolutionary disadvantage, since not noticing it might enhance communication. Then, if awareness of ungrammaticality has negative survival value, any vestigial consciousness of it suggests an underlying mental reality.

Bill or { his brothers/ ?they } Mary doesn't believe can win.
Which paints did Sue buy before { showing/ ?her brother showed } to Sam?

Consequently, in this study, among sentences easily understood and processed, *only those universally judged as acceptable* are taken as grammatical. This simple heuristic conflicts with widespread justifications of analyses which treat data accepted only by some speakers as grammatical, a practice which more often than not undermines the most interesting and contentful theoretical formulations.

1.2 Types of syntactic categories and features

1.2.1 Canonical matching of features and categories

Any framework purporting to characterize syntactic well-formedness requires an approximate inventory of the categories whose co-occurrence is to be determined.[5] Since language clearly distinguishes four "open" or "lexical" categories whose members number in the hundreds or thousands (nouns, verbs, adjectives, and pre/post-positions: N, V, A, P), these categories must play a central role. Moreover, the great membership of these lexical categories never fails to astonish; informal sampling today still confirms Jespersen's (1905, 227) observation: "People who had never been to college, but, ... were regular readers of books and periodicals, ... reported generally from 25,000 to 30,000 words ..." That is, the combined membership of the four lexical categories X must typically be well over 20,000. Outside of compounds in a well-formed syntactic structure, each of these lexical categories X has a "maximal projection" XP which obligatorily contains ("dominates") its structural head X as well as any modifiers and complements which may modify X.

In addition to and in contrast to the open categories, a number of syntactic categories of limited membership modify and help extend the projections of the lexical categories. Each of these non-lexical "closed" or "functional"

---

[5] An exact inventory of the categories in a theory cannot be required a priori, since otherwise the theory would be fully specified before investigation begins. No scientist can demand that an exact inventory of all distinctive features precede research in formal phonology, that a full inventory of chemical elements precede experimentation based on the periodic table, or that a complete inventory of elementary particles precede work in atomic physics. In fact, one rationale for pursuing a science is to arrive at such inventories.

categories seems to contain at most twenty or so morphemes. The most prominent functional categories are the elements I making a verb phrase into a finite Fregean judgment and a class of quantificational and/or definite items D determining the referential properties of noun phrases. I and VP together form an "extended projection" IP of V, while D and NP together an extended projection DP of N.

For a number of other closed class elements which are characteristic modifiers of at least the four lexical categories X, I continue to use the early cover term specifier SPEC(XP) as in Chomsky (1970); in this study, specifiers can include both closed classes of grammatical elements and/or phrases. For example, there seems to be a single "slot" in which modifiers of English adjectives appear: certain degree words such as *very* and *so*, measure phrases (*two hours, ten feet*) and short adjective phrases (*less believably, somewhat tolerably*) are in complementary distribution:

(1.3) *That lecture was { very two hours/ two hours very } long.
*That { unbelievably too/ too unbelievably } optimistic estimate was foolish.
*Be sure to install a { tolerably ten feet/ ten feet tolerably } deep pool.

I will consider all these elements to be in SPEC(AP) and do not commit myself to analyzing morphemes such as *very* as either a head or a phrase; but little if anything in my arguments hinges on this decision. Nonetheless, this use of SPEC does not conform to the practice of authors who reserve this symbol for phrasal positions.[6]

Syntactic categories thus fall into two separate classes, whose diverging nature can be derived from the following five properties of Universal Grammar:

a. UG provides a restricted set of morpheme categories {B}: lexical heads X, specifiers SPEC(XP), I, D and perhaps a few others.[7]

---

[6] It is not obvious that expletives such as *there* in the supposedly phrasal position SPEC(IP) or unmodifiable question and relative markers such as *whether, why*, French *dont* 'of which' or Dutch *wer* 'where' in SPEC(CP) are phrasal either, other than by assumption.

On the other hand, certain finer analyses of extended projections of A (Corver, 1997) and P (van Riemsdijk, 1998b) argue that head and phrasal positions can be kept distinct, in line with suggestions in Chomsky (1986).

[7] These include (i) coordinate conjunctions, (ii) the discourse particles or "delimiters" of several languages *even, only, also,* etc. which attach to any XP (Kuroda, 1965), and (iii)

b. UG matches a small range of *cognitive syntactic features* F (upper case) with each B whose combinations [B, ±F] characterize up to a maximum of twenty or so members of B.[8]

c. The cognitive syntactic features F on B contribute *centrally* to meaning (that is, to "Logical Form") in all syntactic classes.

d. Finer distinctions of meaning in terms of purely semantic features *f* (lower case script) which play no role in syntax (Chomsky, 1965, Ch. 4) appear *only in the four open lexical classes N, A, V and P.*

e. All non-lexical categories are "closed" because they crucially lack these purely semantic features *f*, which apparently proliferate and recombine fairly freely. Hence, the closed categories have few members and disallow coining (Emonds, 1985, Ch. 4).

Some plausible examples of purely semantic features *f*: (i) I would imagine that color terms share a feature f_i, permitting for example unexceptional compound adjectives such as *dark pink* and *light magenta*, in contrast to (compounds) **dark smooth* or **light dirty*. (ii) Among verbs of say "harm" (which itself might be a semantic feature), some exclude further normal use of an object (*!He destroyed/ ruined/ totaled/ wrecked the bicycle and then rode away on it*) while others don't (*He damaged/ harmed/ messed up/ misused the bicycle and then rode away on it*). Plausibly, some feature f_j meaning roughly "usable for intended purpose" seems involved. It seems incontrovertible that elaboration of syntactic theory does not in general depend on such features *f*.

I now discuss the properties (a–e) in more detail. Properties (a–b) are reflected in the partial and tentative table (1.4), which may well be modified as research proceeds. Most of these matches are inherited from traditional grammar, with generative grammar providing a number of non-trivial modifications, as in note 8. Many features F have unique canonical positions in

---

the emphatic particles *too, so, either* which appear only with finite verbs. (ii) and (iii) seem to escape classification as specifiers precisely because they can occur with them. Aoyagi (1998) nonetheless considers the discourse particles to be specifiers which can further adjoin to otherwise maximal phrasal projections. The category of emphatic particles is uncertain; perhaps, as argued in Laka (1990), they constitute a separate class of functional heads.

[8] I do not identify all the features which traditional grammar associates with a lexical category as its features in UG; e.g. generative analyses of English, Kru languages (Koopman, 1984) and Chinese (Huang, 1990) have shown that TENSE is not a feature on V but rather a separate category. But neither do I assume a priori that every syntactic feature projects as a separate head.

which they are realized, such as PAST on I, but for some F, UG may provide more than one possible host. Thus, POTENTIAL is matched with I and A in English (*can, able*) but with V in French (*pouv-*), and the COMPARATIVE features can occur in several specifier positions (Bresnan, 1973: *more into Zen, more of a man*).

(1.4) Examples of *probable* UG matches:

syntactic features F	*categories B*
tense and modal features	I
quantifier features	D or NUM
space-time co-ordinates	P
ACTIVITY	V
PERFECTIVE (aspect)	V
ANIMATE, COUNT	N
comparative features	SPEC(XP)

Certain features F may cross-classify the syntactic categories B. Thus, a [+N] feature subsumes N, D and A, while [-N] subsumes V, I and P.[9] The much discussed feature ±WH most plausibly occurs on the "specifier" (SPEC) of both the marked [+N] categories D and A in e.g. *which, what, how*, and so also does ±PROXIMATE; e.g. *{this / that} + {bread / tall}*.

A central tenet in this study's approach to the lexicon is thus a general condition on syntactic features (1.5), which expresses (b) and also (c) above.

(1.5) *Canonical Realization. UG associates a few cognitive syntactic features F with each syntactic category B. These features F contribute to semantic interpretation (Logical Form) only in these "canonical positions" on B, and appear elsewhere only via language-particular lexical stipulation.*

Chapter 4 returns to examples of syntactic features which are *not* canonically realized and to mechanisms which severely limit the distribution of these uninterpreted features.

---

[9] Whether plus values of features can always correspond to marked values is discussed below; the feature +N unifying nouns and adjectives was historically adopted without regard to this issue. A little reflection shows, however, that -N (V and P) is a marked value; N is certainly the unmarked word class, while P is the least numerous and hence plausibly the most marked. We might think of case-marking of N and A (+N, accepts +*f*) by V, P, I and D as categories in a "higher energy" or marked state discharging this "energy" (markedness) onto the less active categories N and A.

Notice that properties (a–c) do *not* distinguish open from closed classes; the distinction follows rather from properties (d–e). These two properties which set off closed class categories can be formally linked as follows:

(1.6) *Outside the lexical categories N, V, A and P, the only features allowed are the cognitive syntactic features F (and the small sets of morphemes they generate).*

Observe now that the lexical categories are defined as those whose features *may* extend beyond the inventory F; this possibility of additional semantic features *f* makes these classes "open." The question arises as to whether every N, V, A and P *must* have *f* distinct from F. In fact, there is no reason to assume this, i.e., Canonical Realization (1.5) can apply as its stands to the four lexical categories. Indeed, certain subclasses of N, V, A and P have properties characteristic of the non-lexical classes, such as post s-structure insertion and unique syntactic behavior (cf. Emonds, 1985, Ch. 4, and 1987). *The lexical categories are in this way like the others: each has a subset of say up to twenty or so elements fully characterized by cognitive syntactic features F and entirely lacking purely semantic features f.*

These closed subsets of open categories can be called "grammatical" N, V, A or P; in contrast, open class subsets which have purely semantic features will be called "lexical" N, V, A or P. English grammatical verbs include *be, have, do, get, go, come, let, make, say*, etc. Its grammatical nouns include *one, self, thing, stuff, people, other(s), place, time, way, reason.* A widely acknowledged distinction between grammatical and lexical or "contentful" prepositions also falls into place as a predicted subcase of a more general property of a theory of lexical categories: lexical P are specified with purely semantic *f* and grammatical P are not.[10]

(1.7) *Definition. A closed grammatical class X (including N, V, A, P) is one whose members have no purely semantic features f, but only cognitive syntactic features F.*

In Chapter 3 of this work, I define the subpart of the lexicon which contains all and only the closed grammatical classes of elements as the

---

[10] Studies of types of aphasia invariably show that forms of the English copula group with closed class elements. But since such research is not examining any proposal to group morphemes except as members or non-members of the four lexical classes, it does not systematically investigate whether other grammatical members of these classes differ from open class items.

"Syntacticon" of a language, and show that it has properties quite distinct from an open class "Dictionary". In the first two chapters, however, this distinction will not yet play a central role. Nonetheless, it will be necessary to retain the definitions and proposals in this and the next subsections, for the explanatory accounts throughout this study will repeatedly utilize the distinctions between grammatical and lexical X and between features in canonical and non-canonical positions.

1.2.2 Marked feature values, including Absence of Content

This study takes no position as to whether lexical entries are specified for all their features, for marked feature values only, or in some other way. However, this study will carefully develop a conception whereby *marked syntactic features* on lexical items can appear in syntax outside their canonically matched base positions under very restricted syntactic conditions. Thus, to facilitate properly defining these conditions, it is important to have criteria for when a feature value is marked.

The canonical positions of syntactic features such as PLURAL, PERSON or DEFINITE are somewhere in the extended projections of nouns, most likely on D. Thus, when such features appear under V in e.g. an agreement morpheme or a Romance pronominal clitic, they are not in their canonically realized positions inside DP. How complex and stipulative such "alternative realizations" are should be calculated only in terms of marked features. Thus, PLURAL and DEFINITE are marked (and hence costly) features to realize under V, or in other words, the Romance pronominal clitic system is a lexically marked subsystem. But once a PLURAL clitic appears in Romance, it is probably not additionally more costly or stipulative if the clitic is further Masculine and 3rd Person, since these are unmarked feature values with PLURAL.

Use here of a plus value on a feature will usually reflect some previous or obvious argumentation that a certain feature is a marked option: e.g., +PLURAL number, 2nd Person,[11] Romance +FEMININE gender, +PAST tense, IRREALIS = +MODAL, the +SOURCE value of PATH, +NEGATIVE polarities of antonymous adjectives, etc. In other cases, it turns out that certain feature *combinations* are unmarked, such as possibly [ACCUSATIVE, -DEFINITE] and [NOMINATIVE, +DEFINITE]. Ideally, one

[11] Benveniste (1966) argues extensively that 3rd Person is the least marked of the Person features, and Banfield (1982) that 2nd Person is the most marked.

would like to equate plus values, marked values, and simple presence, and similarly identify minus values, unmarked values, and absence. But this is not possible if unmarked values of a feature such as ±DEFINITE vary according to syntactic context. When this situation arises, I will simply remark in the text that a minus value is the marked value.

An important but rarely remarked correlation within the open classes is that their more numerous syntactic subclasses have what one might call "more content." For example, the vast majority of verbs are activity verbs, that is, not +STATIVE; the vast majority of English nouns are countable, and a great majority of prepositions indicate spatial or temporal relations. It seems plausible that these more numerous and more contentful subclasses of X^0 are the cognitively primitive or unmarked ones, and that *the less numerous subclasses of less content are lexically marked* (stative verbs, mass nouns, logical or grammatical prepositions), that is, less expected by the language learner.

If so, the appropriate way to reflect these facts is to let certain cognitive syntactic features on N, V, A and P represent a *marked Absence of Content.* Thus, +STATIVE is the marked lexical value for V; there are fewer stative verbs and their defining property is *failing to* contribute certain typical content of V to interpretation. For instance, Chomsky's rule for determining agenthood of subjects (cf. Chapter 2) does not apply to (precisely) STATIVE verbs. From this point of view, then, the -STATIVE verb *do* is less marked than the +STATIVE verb *have.*

Similarly, the unmarked class of P have the feature +LOCATION. Those P which fail to specify spatial or temporal location (*of, despite,* most uses of *for* and *with*) are lexically marked for this feature and hence are not as finely subdivided for semantic nuances as are the locational P. In Chapter 2, we will see that certain classes of V also receive +LOCATION, but as a *marked* value, so as with ±DEFINITE above, there is no consistent ± notation meaning "marked for location": +LOCATION is lexically marked for V, while -LOCATION is marked for P. -LOCATION is an "Absence of Content," and so for P it is a "marked Absence of Content".

As we proceed, it will become evident that among the syntactic features F (the only features within closed classes), three subclasses are not interpreted in LF: (i) purely contextual features, (ii) copies of a feature F outside of F's canonical position(s), referred to subsequently as "alternatively realized F," and (iii) features which specify a marked absence of content. These three classes of syntactic F which are not used at LF will be called "purely syntactic" or uninterpreted. Other cognitive syntactic features which are used at LF, for instance those discussed above, will be called "cognitive

syntactic" or simply cognitive or interpreted features notated in upper case. Lexical elements will turn out to have very different syntactic properties depending on which of these types of features they realize.

There are thus three main types of features on categories which interact differently with respect to lexical insertion, syntactic derivation and logical form ("LF"):

(i) Purely semantic features *f*, which are present *only* on the head categories X = N, V, A and P. They are not used in syntax and are not present on closed subclasses of grammatical X.

(ii) Cognitive syntactic features F in canonical positions, which can occur with all syntactic categories. They play a central role in both syntax and at Logical Form.

(iii) Purely syntactic features F, also occurring with all syntactic categories. They indicate contexts, realize the features in (ii) in non-canonical positions, or stipulate a marked lack of content (e.g, +STATIVE on V, -LOCATION on P). They are centrally used in syntax but play no role in Logical Form.

Summarizing, as stated earlier, both the types of features used in syntax, i.e. in (ii) and (iii), are uniformly notated with upper case F.

1.3 A theory of phrase structure as Extended Projections

1.3.1 Lexical Projections

The co-occurrence properties of the categories and features for which the preceding section gives a rough inventory must be accounted for by a universal theory of phrase structure and language-particular lexicons. So before turning in especially Chapters 2, 3 and 8 to the task of formalizing the notion "possible lexical entry," we need a general theory of how, independent of lexical specification, the categories of section 1.2 can combine.

Around 1986, generative grammar reached a seeming consensus that a restrictive and universal set of syntactic categories, called the "bar notation", was empirically adequate and cross-linguistically appropriate. Under this conception, any language contains at most the four lexical heads X, a very few functional heads such as D (associated with NP) and I (associated with VP), non-maximal or maximal projections of these six or so heads (respectively X' and XP), and specifier (SPEC) daughter nodes of XP. As argued in Speas (1990, section 2.2), non-maximal X' differ from maximal projections XP only in that the latter do not head some larger XP; i.e.,

adjoining @ to a phrase renders this phrase non-maximal. My own research has argued that an additional functional head C for introducing clauses ("complementizer") reduces to a class of grammatical P with sentence complements (Emonds, 1985, Ch. 7); this revision of the bar notation is adopted throughout this work.

Such a conception of syntax suggests strong restrictions on the category inventories available in Universal Grammar. Consequently, in that work's Introduction, I formulated two principles which are still tenable if, as just outlined, we are justified in postulating only a very small set of syntactic categories.

(1.8) *Categorial Uniformity. The categories defined in terms of the bar notation, X^j and SPECIFIER(X), do not differ from language to language, but their subcategories which are realized in each language's syntax may vary.*

(1.9) *Hierarchical Universality. The range of hierarchical combinations of syntactic categories does not vary from language to language at the level of deep structure.*

Although many analysts in intervening years have reanalyzed the syntactic features of the six or so lexical and functional heads as a greatly expanded set of functional heads and projections, I remain unconvinced of the fruitfulness of this line of research.[12] Rather, the earlier, more parsimonious theory retains its appeal and promise. I thus utilize a conservative category system, elaborated along the following lines:

a. Lexical category heads X together with their complements and adjunct phrases constitute units of syntax, called *maximal projections* XP of these X.
b. Finite verbs V and *all* their arguments are syntactic units, called clauses IPs.
c. Analogously, nouns N including derived nominals and all their linguistic (as opposed to pragmatic) arguments are syntactic units, called DPs.

---

[12] Chomsky (1995, section 4.10) seems less enthusiastic about such analyses, e.g. about separate projections for agreement features. My skepticism about a greatly expanded set of functional categories does not preclude the possibility that the set widely recognized in the mid-eighties is too small.

Formally specifying how these heads combine with complements is the subject matter of succeeding chapters, which each modify or expand or limit the role of subcategorization. Chapter 7 also includes discussion of adjuncts, i.e. phrases which are both a sister and a daughter of the same type of phrase. Provided that CPs are special cases of PPs, one may impose a general syntactic restriction: *Adjuncts are always PPs or APs.*[13] APs are sometimes "adjectival" (they agree with a modified NP) and sometimes "adverbial" (they lack such agreement).

The divisions in (a-c) above suggest that a certain subclass of "external arguments" of V and N is generated outside VP and NP. The remainder of this chapter focuses not on the internal structure of XP, but rather on how XP containing all of X's complements and adjuncts combine into larger phrases with functional heads, namely IP and DP.

1.3.2 The Subject as a special phrase: I and IP

Within finite clauses, one argument NP/DP of the verb, a "subject phrase," is external to the VP. Indirect evidence for its special status is of many sorts: (i) the role of the subject – predicate pairing in Aristotle's logic and its wide and long-standing acceptance; (ii) its central role in Cartesian *grammaire générale* and the resulting "traditional grammar"; and (iii) the still valid descriptive value of Chomsky's (1957) original phrase structure rule, S → NP - VP. Many of the thirty odd properties examined in Keenan (1976) that "cluster" around subjecthood are prototypically exhibited by subjects of finite clauses.

Several direct, empirical syntactic arguments for the VP-external nature of the subject are provided by processes such as e.g. VP-fronting and VP-ellipsis in English and Japanese, English tag questions, widespread subject-verb agreement in languages lacking other agreement patterns, etc. Notably, these arguments all concern or work best with the subject phrase of *finite* clauses.

By the same arguments and others, the finite elements I are just as external to VP as the subject phrase is. This can be seen easily with English

---

[13] Finite adverbial clauses are invariably introduced by complementizers, and in turn, as mentioned above, CPs can be shown to be PPs. "Bare NP adverbials" as well as adjunct purpose clause infinitivals are argued to be PPs with empty heads in Emonds (1987). Present participles, which might be thought of as adjunct VPs or reduced IPs, are argued to have the structural form of APs is Emonds (1991a); Chapter 5 extends this idea to passive participles.

modals, which are prototypical instances of the finiteness category. Like subject phrases, modals don't front or ellipt with VPs and they appear in tag questions. They furthermore contract with the subject and with the negative (in the form *n't*) in ways that Vs do not.

(1.10) *Deep structure finite clauses are of the form IP = DP - I - VP, where I is a grammatical "finiteness" head paired with V.*

For uniformity with the rest of the bar notation, we can assume that this DP is in SPEC(IP) and that the combination I + VP constitutes an I' (Chomsky, 1986).[14]

The structure in (1.10) differentiates a special argument called the subject from other "internal arguments" or complements, which is then defined by (1.11). In most cases, Keenan's special properties of a subject can be related to its VP-external status. A counterpart to such a definition is needed in any theory of grammatical relations which differentiates subjects and objects.

(1.11) *Definition of Subject (tentative): The subject (or external argument) of a head V is a DP/NP which c-commands VP within the minimal IP containing V.*

1.3.3 The DP Hypothesis and a generalized definition of Subject

Chomsky (1970), Jackendoff (1977), George and Kornfilt (1981) and Abney (1987) all argue that the internal structure of noun phrases parallels that of sentences (IP), in that many classes of head nouns have a "subject", analogous to the subject of verbs, realized as a possessive noun phrase within a traditionally conceived larger noun phrase. Abney further proposes, following Brame (1984), that the noun phrases of the earlier generative literature contain two heads, where one is a grammatical "reference or quantification" head, say D (= Determiner), paired with N much as I is paired with V. This is Abney's DP-hypothesis. Under this conception, the earlier noun phrases

---

[14] CPs and PPs can appear to be subjects of English finite clauses. However, investigation of a range of embedded contexts reveals these non-DP subjects to be a root phenomenon akin to topicalization (See Emonds, 1976, Ch. 4 for CPs and Emonds, 1985, Ch. 7 for PPs). Moreover, Higgins (1973) shows that these topicalized CPs have their source only in DP positions: *That Mary would be late [*it didn't seem / few believed (*it) / we thought (*it) obvious].*

have the form (1.12) analogous to finite clauses as in (1.10). This structure then defines an external argument for N, which again we call SPEC(DP).[15]

(1.12) *Deep structure nominal phrases are of the form DP = (DP) - D - NP, where D is a grammatical "reference or quantification" head paired with N.*

The parentheses around the possessive DP in (1.12), in contrast to an IP subject's obligatory status in (1.10), reflects the fact that nouns, including derived nominalizations of verbs, often lack a linguistic subject within the larger DP (Wasow and Roeper, 1972). That is, their understood subject can only be pragmatically determined.

One argument for this is that a syntactically represented understood subject, a so-called "arbitrary PRO", must have animate reference at least in the positions SPEC(IP) and SPEC(DP), as shown in (1.13a-b). In contrast, the pragmatically understood subjects of derived nominals are often inanimate (1.13c).

(1.13) a. Inferred PRO must be animate; the following cannot refer to a volcanic eruption:
To explode like that does a lot of harm.
Exploding like that does a lot of harm.
(inferred subject = +ANIMATE)

b. A V which cannot have an animate subject is incompatible with a PRO subject:
*Few people like to unexpectedly occur.
*Few people like unexpectedly occurring.

c. Noun-headed nominalizations may have a pragmatic subject or no subject:
An explosion like that does a lot of harm.
(inferred subject = ±ANIMATE)
Few people like unexpected occurrences.

Serious questions remain as to exactly which morpheme classes realize the functional head D paired with N (= Determiner). Often without argument, the best candidates are assumed to be the definite demonstratives and

---

[15] The advent of the DP hypothesis has given rise to especially interesting work on Semitic languages focusing on the "construct state" of DPs which modify Ns: Borer (1984, 1989), Ritter (1988), Fassi-Fehri (1993) and Siloni (1997), among others.

articles. Nonetheless, Jackendoff's (1977, Ch. 4) detailed empirical work on English noun phrase specifiers isolates two distinct closed class or functional category positions which can occur in sequence, a definiteness / quantifier position and a numeral / quantifier position. Both are exemplified in the sequences *those many books, any three vertices, the little nitrogen remaining.*

On the face of it, both classes of morphemes might qualify for D, or there may be two rather than one functional projections matched with N. For example, Szabolcsi (1987) argues on the basis of extractions in Hungarian that a definiteness category D shares properties of C, while an additional functional head between D and N, analogous to I, assigns case to subjects. In other work in principle compatible with this analysis, Cardinaletti and Giusti (1991) and Giusti (1991) argue that both the definiteness and numeral positions should be functional heads.

On the other hand, these approaches seem to exclude a priori a possibility I retain, that closed classes of morphemes might be listed as SPEC(DP) or SPEC(NP). Lobeck (1995) explores parallelisms between English VP and NP ellipsis, both presumably licensed only if a functional head is present outside the ellipted phrase. Clearly, Jackendoff's second numeral / quantifier position may play this role and thus qualify as D, while his first position (for definite elements and certain quantifier morphemes) might conceivably be SPEC(DP).[16] In this work I will not try to definitively establish criteria for membership in D, nor determine whether two independent functional heads can appear above nouns in a DP.

Leaving aside the exact nature of D, let us proceed to formalize further parallels between NP and VP, besides the fact that each is the complement of a functional head. For these we need definitions.

(1.14) a. *N and the projections of N and D are "N-projections"; V and the projections of V and I are "V-projections."*
b. *DPs and IPs are "extended projections" of N and V respectively.*
c. *IP and DP are "cyclic domains" of V and N respectively.*

---

[16] Under this view (to which I am not committed), the ellipsis in *Your wine but not Sam's was tasty* would require a null functional head D. Such null heads might then be licensed by the possessive morpheme *'s* as an "alternative realization" of D on the SPEC(DP) sister of D', as in section 4.3.2.

A reviewer observes that if non-phrasal morphemes can be SPEC(XP), then they must still be specified as ± ___X', a head-like property. Thus, if *every* and *all* are SPEC(DP), *every* is +___D' and *all* is ± ___D'.

It is a natural step to extend the notion of "subject of V" in (1.11) so as to encompass "subject of N", including of nominalizations:

(1.15) *Generalized Subjects: The subject or external argument of a head X is the lowest N-projection which c-commands a phrasal projection of X within its minimal cyclic domain.*[17]

(1.15) correctly requires that a *linguistic* subject of N (including nominalizations) must be inside DP. It further predicts that if a head noun has no such DP-internal subject, it can only have an unprojected, pragmatic subject (Wasow and Roeper, 1972). This generalized definition of subject will play an important and recurring role throughout this study.

The construction known as "subject small clauses" confirms the appropriateness of (1.15). Kubo (1993b, 103–107) provides evidence that such predications have both the internal structure and external distribution not of clauses but of DPs. In her structures (1.16), the brackets indicate the smallest cyclic DP domains containing italicized heads X of predicate phrases; the bold N or N' are the nominal projections which are the subjects of these X.

(1.16) [**Young workers** *angry* over their pay] looks revolutionary.
[**Children** *in* dangerous parks] is a scene used to convince women to quit jobs.
[The **flags of many lands** *flying* over the plaza] is a good scene for a postcard.
[**Paris** *and* its perfumes] fascinates American women.
[**Sake** *and* tofu (together)] makes me sick.

By interpreting the predicate category as covering A, P, non-finite V and even co-ordinate conjunctions, Kubo shows that a definition similar to (1.15) correctly characterizes the (bold) nominal projections in (1.16) as the subjects of the italicized predicates. Her analysis makes no appeal to any clause-like structure not independently motivated as DP-internal.

---

[17] Generalized Subjects permits subjects of X to be within maximal projections of X but does not require them to be. It is thus neutral with respect to the issue of the "VP-internal subject hypothesis" (Zagona, 1982) and in fact allows for the possibility that in some constructions a subject is outside VP and in others within it.

1.3.4 The EPP: explaining the "strong D feature on Tense"

Let us now return to the discrepancy between (1.10) and (1.12): subjects of IPs are obligatory while those of DPs are optional. To maintain that, uniformly, only heads of projections and extended projections are obligatory and that argument positions are optional, we can revise (1.10) to IP = (DP) - I - VP, parallel to the formula for DP (1.12). The obligatory linguistic presence of the subject of a V under IP is then factored out as Chomsky's (1981) Extended Projection Principle, which can be expressed as (1.17).

(1.17) *Extended Projection Principle. Every head verb present in Logical Form must have a structural subject phrase to which a semantic role may be assigned.*[18]

We will see in later chapters that the EPP plays an explanatory role well beyond allowing the IP projection (i) to parallel that of DP and (ii) to conform to the generalization that only heads are structurally obligatory.

The EPP and the formulation of Generalized Subjects (1.15) guarantee that the SPEC(IP) will always contain a structural DP subject for the highest V in IP. While this seems correct for English and many languages whose finite I obligatorily agrees with its subject, there is evidence that SPEC(IP) in a non-agreeing language such as Japanese can be absent, i.e. Japanese seems to directly reflect the optionality in the formula IP = (DP) - I - VP.

Sells (1996) contrasts two types of raising and obligatory control complements of Japanese verbs, both of which are uniformly IPs with past – non-past tense alternations. He shows how one type permits an overt (reflexive) subject DP presumably in SPEC(IP), while the other entirely lacks such a DP position. This latter type of IP, for which he constructs I believe incontrovertible arguments, Sells calls "sub-clausal" and observes that their understood subject appears in either subject or object position in the main clause. His arguments thus establish that Japanese robustly instantiates a subjectless (head-final) structure [VP - I]. In a somewhat more general and

---

[18] Impersonal verbs with so-called "theta-bar" subjects (like *seem* and French *falloir* 'be necessary') do not actually assign such roles in LF, but they still require a subject position. The EPP has come to be known in the Minimalist Program as the "strong D feature on Tense," i.e. the Tense projection (here IP) in many languages must have a DP in its Specifier position at Spell Out. As far as I can tell, this is simply stipulated and is moreover significantly less general than (1.17). For example, (1.17) will crucially explain why the lower lexical verbs in both verbal passives (Ch. 5) and Romance causatives (Ch. 6) must have (possibly understood) subjects, which are outside SPEC(IP).

theory-based treatment, Kuroda (1992, 321) reaches the same conclusion: SPEC(IP) "can be left vacant" in Japanese but not English.

These arguments, taken at face value, suggest that a Generalized Subject @ of a V need not always be present within the smallest finite IP; i.e., @ can be the lowest DP which c-commands VP within some domain *larger* (in Japanese) than its minimal IP.[19] We get this consequence if we include a condition – not stated in (1.14c) – that a cyclic domain of V must additionally contain *a potential subject,* i.e. a DP c-commanding a projection of V. Under this condition, SPEC(IP) can, like SPEC(DP), be entirely absent in certain instances. Nonetheless, cyclic domains for a VP do not under this conception extend upward in the tree without limit; the lowest IP containing a potential DP controller of VP will then be the latter's cyclic domain, and the EPP (1.17) guarantees that there must be such a DP.

This now raises the question: if the EPP doesn't force SPEC(IP) to contain DP, why must it do so in English and other agreeing languages? To account for this, I extract a part of a general hypothesis of Kuroda (1992, 325) that "English is a forced Agreement language." While he derives several contrasting properties of English and Japanese from the respective presence and absence of his generalized notion of agreement, I focus here on perhaps the morphologically most salient difference between the two languages:

(1.18) *Number Filter. The functional heads (I and D) must be specified for number (±PLURAL) at PF in certain languages (e.g., English but not Japanese).*

Now, the cognitive syntactic feature PLURAL is canonically associated with D – at least for this discussion I mean D to be precisely the canonical locus of PLURAL as in (1.5). (1.18) then specifies the fairly uncontroversial property that the English but not the Japanese DP must express the ±PLUR distinction; for related discussion of (optional) Japanese noun classifiers, see Kubo (1996) and section 6.4.3 here.

But equally, PLURAL is *not* a canonical feature of I. Number can only surface on I and hence satisfy (1.18) when the latter agrees with some phrasal projection in its SPEC which *is* specified for ±PLUR, i.e. SPEC(IP) must be a DP in a "Number Filter" language. We thus derive not only the

[19] Neither Kuroda's nor Sells' framework can accept this conclusion because of commitments to other hypotheses (VP-internal subjects and argument structures respectively).

obligatory presence of a phrase in SPEC(IP) from a more general property of English but also the fact that its SPEC(IP) must be DP rather than just any XP.[20] In light of (1.18), the specifiers of both IP and DP can remain optional at the level of UG, as can all non-heads in the bar notation.

In conclusion, the combinations of the basic syntactic categories discussed in section 1.2 are the classic bar notation XP projections of the four lexical heads N, V, A and P and the extended projections given by the formula (1.19). Strictly speaking, these latter expansions are not rules, but simply definitions of the extended projections of N and V.

(1.19) *Functional Projections. FP = (DP) - F - XP; when F is I, then X is V and when F is D, then X is N.*[21]

The subject positions of FP can be referred to as SPEC(IP) and SPEC(DP). The left to right order of the elements in these projections is related to general word order principles in section 3.1. Beyond projections of these six heads, I do not envision greatly expanding the inventory of possible phrases.

---

[20] If the XP in SPEC(IP) need not be DP in certain "V second" languages such as Icelandic and Yiddish, then number ageement on their I must be ensured differently than in English. This obviously holds independently of (1.18).

It is tempting to say that a null and caseless DP in SPEC(IP), though sufficiently specified for number for ensuring compliance with the Number Filter (1.18), is not a visible source at PF for any concrete morphemes which realize agreement on I. From this would follow the correct result that a null caseless subject in English implies a null I at s-structure. (The converse of course doesn't hold.)

[21] In my view, A and P do not have extended projections, at least in the sense that these extended projections *contain DP subjects of A and P.* For example in English, no movement of an AP or a PP which may happen to be adjacent to its subject, as in *We consider Bill stupid* or *We found John in the library,* ever includes that subject:

(i) *What we considered was Bill stupid.
*Bill stupid though she considered, she married him anyway.

(ii) *It was John in the library that we found.
*Whose father in the library did they find?
Cf. How many feet behind the house was it placed?

I don't exclude the possibility that full APs and PPs include functional projections, but such projections, unlike those of N and V, don't contain subjects of A and P. My use of the term "extended projection" covers only functional projections which include subjects.

1.4 The interplay among derivations, the Lexicon, and Economy Principles

1.4.1 Transformational derivations

The descriptive and explanatory successes of transformational generative grammar have been based on a model which assembles choices from a lexicon, puts them together in trees (or phrase markers) in terms of a theory of phrase structure, and then subjects these underlying trees to a sequence of Chomskyan "transformational operations" which maps them into some sort of "well-formed surface structures" in which, possibly among other elements, overt morphemes appear in the order pronounced. Abstracting away from actual phonological processes (e.g. assimilations, epentheses, deletions, aspirations, etc.), we can, along with many syntacticians, still refer to the results of these transformational derivations as "Phonological Form" or "PF".

Chomsky (1976) and Chomsky and Lasnik (1977), building on a decade of "interpretivist semantics" on the relation of syntax to a range of generalizations about meaning, codified the notion that the basic locus of the interface of syntactic structure with the use and understanding of language was "late" in a transformational derivation, i.e. subsequent to most or all of the transformations that had been scrutinized up to that time. The transformationally derived syntactic structures which speakers use to interpret and understand language are widely called "Logical Form" or "LF", following Chomsky (1976).

A number of often rather abstract debates have surrounded the relation between LF and PF. The most usual positions are that a transformational derivation of PF "branches" to LF (or vice-versa, depending on one's point of view), and that subsequent to this branching, a number of invisible or "covert" transformational operations further rearrange syntactic structure before it reaches the LF interface with understanding and use. Many "minimalist" models postulate even more covert syntax than is observed in the overt part of derivations. At the other extreme, the "linear model" of van Riemsdijk and Williams (1981) and van Riemsdijk (1982) simply defines LF as a certain intermediate point in the derivation of PF; this model does not countenance covert movements. However these debates are eventually settled, cyclic domains (1.14) at the branch point (also called Spell Out) can be called s-structures, without claiming or denying that they exhibit interesting sets of properties. In all current competing derivational models (i.e. those which transformationally map combinations of lexical choices to PF),

the derivational structures at the branch point, whether close or far from either LF or PF in terms of operations, are agreed to contain various types of empty categories (traces, "silent copies," empty pronouns, empty discourse anaphors, certain null morphs), as well as indices which link these empty categories (and other co-referring categories) to antecedents of various sorts. That is, any current derivational conceptions of s-structures are much more than structuralist-style trees, whose only function is to impose structure and classification on sequences of overt morphemes.

Throughout this work, I adhere to such a derivational model, in which transformations map sets of lexical choices combined in phrase structures into s-structures with empty categories and co-indexing and subsequently into LF and PF; hence I frequently employ the acronyms PF and LF. A transformational operation which precedes s-structure is said to occur "in syntax," while one which applies subsequent to s-structure in deriving PF is said to take place "on the way to PF" or "late" or even, somewhat misleadingly, "at PF." Since I am unconvinced that morphemes are actually reordered between s-structure and LF, my operations "on the way to LF" or "in LF" refer, unless explicit mention is made of movement, to either deletions as in Lasnik and Saito (1984) or to copyings. I tend to think of LF copying as applying to indices, which express identity of reference for projections of D and identity of sense for other constituents.

1.4.2 The Lexicon

Chomsky (1995) systematically refers to PF and LF as "interfaces" of syntax with respectively a perception/articulation system and an interpretation/use system. Although perception, articulation and use of language can common-sensically be thought of as part of the "real world," the systems or components to which Chomsky refers are clearly mental faculties. PF is then the interface with a mental faculty that allows us to produce and understand phonological material in terms of lexical items packaged in syntactic structures, while LF is the interface with a mental faculty which allows us to use lexical items packaged in syntactic structures to exchange linguistic information in the here and now (and/or plan, reflect on the past, reason things out, express emotion or art, etc.).

While I agree with this "minimalist" line of thought (and hope I have not grossly misrepresented it), there seems to be a missing element in many discussions of how syntax mediates between the PF and LF interfaces. The missing element is the nature of the lexicon, or at least that major part of the lexicon which is not purely syntactic. Along these lines, in fact, Chomsky

(1991, 46) assumes:

> "... that there are three 'fundamental' levels of representation: D-structure, PF, and LF. Each constitutes an 'interface' of the syntax (broadly construed) with other systems: D-structure is a projection of the lexicon, via the mechanisms of X-bar theory; PF is associated with articulation and perception, and LF with semantic interpretation."

At this point, I am not yet concerned with how a syntactic representation is built up from the lexicon,[22] but rather with maintaining that whatever the mechanisms for assembling lexical items, the mental location of these items (the lexicon) is just as much a system or mental faculty as are, broadly speaking, the phonology-articulation-perception and semantics-interpretation-use faculties. More precisely, the open class lexicon, perhaps together with principles for assembling its items into trees, *is a syntactic interface*, and what it interfaces with is the mental faculty of culture and human memory.

In fact, the structure of the mental lexicon, as we will see in this book, is eminently more knowable than various aspects of the other two faculties. This is because the elements of the lexicon leave so many syntactic signatures, and syntactic behavior is the area of language to which we have the most direct and extensive access. In contrast, though at relatively abstract levels we know quite a bit about PF and LF interface representations, *our knowledge of what they are interfacing with* is largely shrouded in mystery.

I am not sure if taking the lexicon as a mental interface in its own right really requires defense, but if so the following considerations may be offered. Basic aspects of linguistic perception (PF interface) and use (LF interface) are plausibly shared with primates.[23] Syntax itself is defined in terms of the cognitive categories of the lexicon and must additionally be largely designed to "fit" the LF and PF interfaces. Consequently, the lexicon presents itself as the best candidate for a *purely human* mental faculty.

---

[22] On this point I have always agreed with arguments given in class lectures by E. Klima (Emonds, 1985, Ch. 2) that lexical items should be inserted in trees from the bottom up, subsequent to transformational operations on their complements, rather than simultaneously at a single level of "deep structure."

[23] It may well be that an oral production faculty is not shared with primates, but it is harder to exclude the possibility that primates share in a non-oral or "signed" mode the perception/ production faculties with which human language interfaces.

(1.20) a. *Species-specific nature of the Lexicon. The open class lexicon is a fully independent faculty of "human knowledge."*

There is thus no close counterpart to an open class lexicon among even higher primates.[24]

When we think that the English mental lexicon probably fully characterizes various complex concepts such as *advantage, flaw, furniture, game, history, religion, vacation, vegetable, collect, intricate* and *outdoors*, it is not so clear that any other knowledge that is both systematic (rather than intuitive) and *effortlessly shared throughout a culture* (neither specialized nor consciously taught) still remains outside the lexicon. It thus seems reasonable to further conjecture:

(1.20) b. *The Lexicon as Knowledge. The open class lexicon is the systematic shared part of human memory.*

This formulation is fairly close to de Saussure's (1916) conception of "Langue", as articulated in his *Course on General Linguistics.*[25]

---

[24] The specifically human "design features" of natural language in Hockett (1960), provided certain of his definitions are clarified, are displacement, duality of patterning, and traditional (rather than genetic) transmission of a large lexicon.

Somewhat speculatively, it is plausible to trace everything specifically human here to the property of duality of patterning, whereby both (i) the smallest elements concatenated in the system (roughly, phonemes) are meaningless *and* (ii) the smallest meaningful elements (roughly, morphemes) again concatenate according to a system. Natural animal communication systems fail to *simultaneously* satisfy both conditions.

This duality is plausibly a necessary precondition for a large, human-like lexicon, which then is passed on through culture. There is no reason to assume that primates couldn't amass a large lexicon if they could divorce phonological units from meaning.

Most likely, really massive lexical transmission, in order to be successful, must additionally be divorced from the here and now, in that we can't learn our large lexicons (containing abstract nouns, illnesses, internal organs, virtues, verbs of mental activities, etc.) merely by pointing. That is, full lexical learning must involve displacement. Thus, displacement, as well as duality of patterning, may be a necessary condition for a large lexicon; in this case, the two design features together would be sufficient for the large lexicon and human language.

A final reduction would consist in showing that either a large lexicon or a division between phonology and syntax (duality) is sufficient to provide a user with a means of displacing thought away from the conceptual template of the here and now.

[25] I share the generative critique of de Saussure, that he simply omitted Universal Grammar from his object of study, identifying "Langue" with a store of signs, i.e., the lexicon. An earlier version of this criticism was that de Saussure omitted all of syntax. But a recent

Finally, as de Saussure also emphasized in the first chapters of the *Course,* while mature individuals may trivially add to or modify their "inventories of signs," they do not lose or even seriously modify the vast majority of their lexical entries over a lifetime. Upon reflection, it seems plausible that the very permanence of lexical entries (in contrast to the notoriously fleeting nature of individual sentences, almost immediately inaccessible to exact recall) may be the basis of Hockett's (1960) natural language "displacement," which is either unavailable to animals or present on a much reduced scale. In any case, its permanence and resistance to modification further support the lexicon's claim to being a mental faculty separate from syntactic computation.

(1.20) c. *Permanence of the Lexicon. The entries of the lexicon, once acquired in youth, remain basically unchanged and unaffected by use throughout life.*

The three conditions (1.20) justify treating the lexicon as a separate mental faculty. The inventories of syntactic categories and features (section 1.2) and the theory of phrase structures as extended projections (section 1.3) then become interface conditions with this faculty. In particular, the fact that syntactic features and categories F are drawn entirely from the cognitive categories which organize the lexicon suggests that syntax in this respect at least is a "perfect system," in the sense discussed in Chomsky (1995). Less obvious is why precisely four of these categories should give rise to (extended) phrasal projections and be at the center of the system; I will not investigate this rather abstract question in this work, though it is a subtheme throughout Emonds (1985).

1.4.3 Economy Conditions

Under the conception that the lexicon is a mental faculty or system, a transformational derivation becomes a device which mediates between *three* syntactic interfaces, the Lexicon, PF, and LF.[26] This suggests that there may

---

generative tendency, which this work adheres to, tends to ascribe all language-particular syntax to the lexicon (cf. Chapters 3 and 4). Perhaps de Saussure also made this assumption, thereby including any syntax linked to a particular language's grammatical morphemes under "Langue." If so, he stands accused not of omitting all of syntax, but only its universal component from his object of study.

[26] Conditions on lexical insertion are, under this conception, properly called interface conditions.